

## Identification of Vietnamese Native *Rhododendron* and *Enkianthus* Species of *Ericaceae* Family Based on Ribosomal DNA Internal Transcribed Spacer (ITS) Sequence

T.V. Tam<sup>1</sup>, D.V. Dong<sup>1</sup>, D.V. Lam<sup>1</sup>, T.H. Anh<sup>1</sup>, N.D. Van<sup>2</sup>, D.M. Trung<sup>3</sup>,  
T.D. Duong<sup>4</sup>, K.H. Trung<sup>4</sup>, N.T. Duong<sup>4</sup>, T.D. Khanh<sup>4\*</sup>

<sup>1</sup>Fruit and Vegetable research Institute, Trau Quy, Hanoi, Vietnam

<sup>2</sup>Department of Science and Technology, Quang Ninh province, Vietnam

<sup>3</sup>Institute of Biomedicine and Pharmacy Vietnam Military Medical University (VMMU)

<sup>4</sup>Agricultural Genetics Institute, Hanoi, Vietnam

\*Corresponding author: [khanhkonkuk@gmail.com](mailto:khanhkonkuk@gmail.com)

**Keywords:** *Rhododendron*, *Ericaceae* family, Internal Transcribed Spacer (ITS) sequences, DNA barcode

**Abstract.** The genus *Rhododendron* and *Enkianthus* are belonging the *Ericaceae* family. In Vietnam there have diversified with 12 genera and about 90 species, of which more *rhododendrons* 30 species with beautiful flowers and two of *Enkianthus* species, widely distributed across the regions of the country. In this study, twelve Vietnamese native *Rhododendron* and *Enkianthus* species collected from Quang Ninh province were identified based on sequences of the internal transcribed spacer (ITS) regions of nuclear ribosomal DNA. Comparison on sequencing among 12 Vietnamese native *Rhododendron* and *Enkianthus* with world's *Rhododendron* and *Enkianthus* species, we have determined the accurate same pattern the 5/10 species by recognition molecule forms such as DQ2, DQ5, DQ6, DQ7, DQ8 and five species as DQ1, DQ9, DQ10, DQ11 and DQ12 were determined in genus *Rhododendron* species. Two varieties of *Enkianthus* as DQ3 and DQ4 were identified as *Enkianthus quinqueflorus*. Although the species is still limited but the ITS sequences application revealed the strength of promotion of universal DNA barcode in the delimitation of species and subspecies level for plant group flowering.

### Introduction

The *Ericaceae* belongs a family of flowering plants, widely known as the heath or heather family, adapted in areas with acid and infertile conditions. The family is large, with more than 4.250 known species spreading across 124 genera, making it the 14th most species-rich family of flowering plants [1]. The *Ericaceae* contain a morphologically diverse range of taxa, including herbs, dwarf shrubs, shrubs, and trees. Their leaves are usually alternate or whorled, simple and without stipules. Up to date, the genus *Rhododendron* and *Enkianthus* belonging to the *Ericaceae* family included 12 genera and about 90 species were reported [2-5].

The genus *Rhododendron* consists of 1.025 species, across occurs in Asia, Europe and North America [6] which is a popular genus for its stunning variety of colors with beauty, and long and rich horticultural history [7-9]. *Rhododendrons* are spread in ecological zones ranging from lowland to high altitude forest and grasslands at various altitude bogs and swamps, as well as montane and alpinzones [9, 10].

Living forms of *Rhododendrons* include trees, shrubs, groundcovers, and even epiphytes [10]. The ethno pharmacological relevance of *Rhododendrons* corresponds to their geographical distribution, most species being used in traditional Chinese and Tibetan medicine, Nepalese herbalism and Ayurvedic medical system and only a few in North American and European folk medicine were reported [12].

The genus *Enkianthus* Lour. (*Ericaceae*) is a genus of shrubs or tiny trees and often used as ornamental plants in the gardens. The genus contains 12-17 species [13, 5]. Most the species are

distributed from the Himalayas to Japan [14]. Molecular phylogenetic studies showed that *Enkianthus* is sister to all other representatives of *Ericaceae* s.l. including *Empetraceae*, *Epacridaceae*, *Monotropaceae* and *Pyrolaceae* [15-17]. A cladistic analysis using morphological, anatomical, embryological and cytological characters suggested monophyly of the genus [13]. The species of *Enkianthus* are distinct from the other *Ericaceae* in several typical traits such as embryological features and morphologies of anther, pollen, seed and hair [5].

In Vietnam, it is found the diversity of this plant with 12 genera and about 90 species [2], of which more *rhododendrons* 30 species beautiful flowers and two of *Enkianthus* species. *Rhododendron* and *Enkianthus* grows naturally in mountainous areas across in this country such as Sapa in Lao Cai, Bach Ma in Da Nang, Tam Dao in Vinh Phuc, Yen Tu in Quang Ninh, Da Lat in Lam Dong provinces and 30 species of *Rhododendron* in the Hoang Lien National Park in Lao Cai province [18].

The assessment of genetic relationships is important for breeding and species classification. There are a large number of *Rhododendron* resources, including wild resources and cultivars which should be evaluated. Previously, the genetic relationships and diversity of wild *Rhododendron* plants were estimated [19], but few studies have investigated genetic relationships among *Rhododendron* cultivars. Sequence data from the internal transcribed spacer region (ITS) is used to investigate the phylogenetic relationships within the tribe. Previous studies have shown ITS to be useful for phylogenetic analysis in *Ericaceae* [20, 21]. This investigation is intended to lay the ground work for future, more detailed studies within the group. Therefore, in this study, we constructed a sequence database of the internal transcribed spacer ITS1- 5.8S-ITS2 region for Vietnamese native *Rhododendrons* and *Enkianthus* species, all of which are to accurately identify based on the morphological characters of vegetative and reproductive organs. The database covers most of the *Rhododendrons* and *Enkianthus* species collected from Yen Tu, Quang Ninh province. The sequence readout of ITS1 and ITS2 regions of nuclear ribosomal DNA appeared to contain sufficient variation to discriminate *Rhododendrons* and *Enkianthus* at the interspecific level. Also, in this study, ITS DNA sequencing was to apply to identify Vietnamese native *Rhododendrons* and *Enkianthus* species obtained in Yen Tu, Quang Ninh province of the country.

## Materials and Methods

Ten samples of Vietnamese native *Rhododendron* and two *Enkianthus* (aged 2-4 years) species were collected in the Yen Tu, Quang Ninh province, grown at Fruit and Vegetable Research Institute, Hanoi, Vietnam. Genomic DNA was isolated from the fresh leaves or flowers by a DNeasy Plant Mini Kit (QIAGEN, Hamburg, Germany) following the manufacturer's instructions and was used as a template for polymerase chain reaction (PCR) amplification

Genomic DNA extraction and PCR amplification: Total genomic DNA was isolated from fresh of leaf by DNA Plant Mini Kit (Qiagen). In brief, a 15 $\mu$ l PCR reaction mixture contained 1.5 $\mu$ l 10xPCR reaction buffer (including MgCl<sub>2</sub>), 0.2 $\mu$ l dNTPs mix (10mM/ $\mu$ l), 0.1  $\mu$ l *Taq* ADN polymerase (5u/ $\mu$ l), and 1.5  $\mu$ l forward and reverse primers (10pmol/ $\mu$ l). The primers used for amplification of ITS region were ITS1 forward primer (5'-GGT GAA CCT GCG GAA GGA TC-3') and ITS4 reverse (5'-TCC TCC GCT TAT TGA TAT GC-3') following the protocol of Trung et al [22]. Approximately 50-100 ng of genomic DNA was used as template for reaction. The mixture was denatured at 94°C for 3 min, and subjected to 30 cycles at 94°C for 15s, 58°C for 30s, and 72°C for 30s, and final extension step of 72°C for 7 min. The PCR production were subjected to 1% agarose gel electrophoresis and detected by ethidium promide staining under UV. DNA was then purified using DNA Purification System (Qiagen Kit).

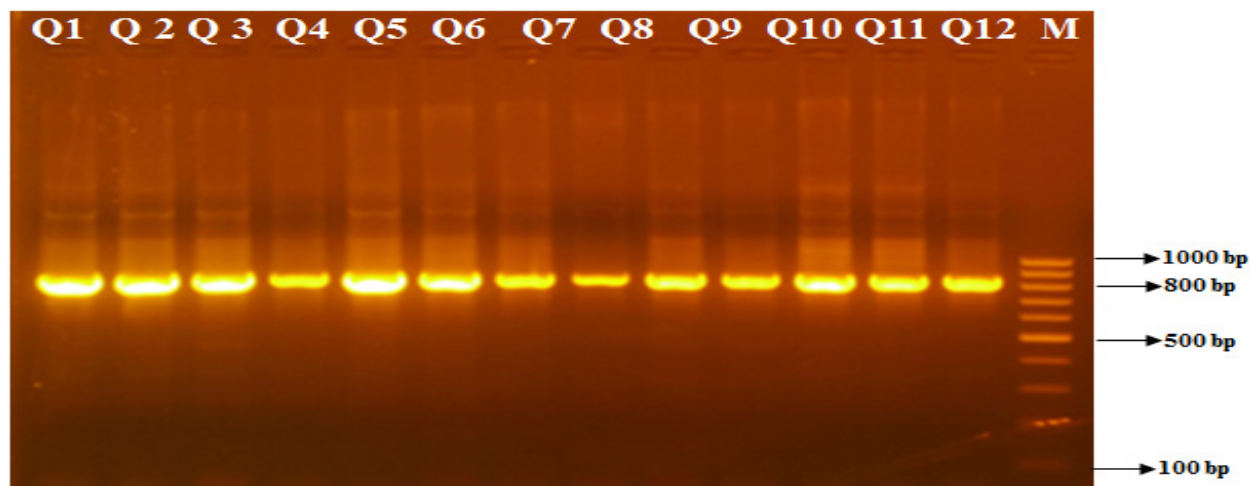
Sequencing: The purified PCR products were sequenced directly by an ABI PRISM™ 310 Genetic Analyzer (Applied Biosystems). The primers ITS1 and ITS4 were used for the sequence reaction. The ITS region of each individual was sequenced in the both the 5'direction and 3'direction at least twice to avoid mutation introduced by *Taq* polymerase. The boundaries of the ITS1 and ITS4 were recorded by comparing them with the sequences of the published species family in the GenBank.

Statistical analyses: The sequences were aligned and compared using the ClustalW software and analyzed using the MEGA 5.2.1 program.

## Results and Discussion

### Molecular makers to identify *Rhododendron* and *Enkianthus* species based on ITS region sequences

By using ITS1 and ITS4 primer pairs, we successfully amplified ITS region by PCR products. These results were high quality with only one band in size ranged from 750-800 bp (Fig. 1). The bands size were clear, correct size which should be able to use for sequencing.



**Figure 1.** Electrophoresis of amplification ITS segment on 12 *Rhododendrons* and *Enkianthus* species by PCR with ITS1 and ITS4 primer pairs.

Lanes: 1, 4, 5, 6, 7, 8, 9, 10, 11, 12 *Rhododendrons* samples; Lanes: 2, 3- *Enkianthus* samples; M: 100bp ladder.

### Analysis of 12 *Rhododendrons* and *Enkianthus* species by sequencing

The results of sequencing showed that the 12 of *Rhododendrons* and *Enkianthus* species were amplified and sequenced the ITS region include partial 18S region, entire regions ITS1, 5.8S, ITS2 and partial 28S area, total length of nucleotides obtained from 750 to 778 nucleotide with average 769.3 nucleotides.

**Table 1.** Length of sequencing of Vietnamese native *Rhododendron* and *Enkianthus* species.

Samples	T(U)	C	A	G	Length of sequencing (bp)
DQ1	24.0	25.8	23.0	27.1	770
DQ2	23.3	26.6	23.2	26.9	772
DQ3	16.5	32.0	22.0	29.6	751
DQ4	16.5	32.1	21.9	29.5	750
DQ5	23.1	26.3	23.5	27.2	776
DQ6	22.8	26.4	23.6	27.2	776
DQ7	22.8	26.4	23.4	27.3	772
DQ8	23.0	26.3	23.1	27.5	778
DQ9	23.3	26.4	23.3	27.0	777
DQ10	23.0	26.5	23.4	27.2	758
DQ11	23.7	26.0	23.4	26.9	773
DQ12	23.8	25.8	23.5	26.9	778
Average	22.2	27.2	23.1	27.5	769

---

The percent of nucleotide as T (U) = 22.2%; C = 27.2%; A = 23.1%; and G = 27.65%. This result has been quite consistent with some previous reports of Xu et al. [23]; Liu et al. [17]; Duong et al. [24] who reported that the ITS in some plants such as *Dendrobium*, *Cymbidium*, *Rhododendron* flowers had a success rate of about 90%.

#### **Alignment on Vietnamese native *Rhododendron* and *Enkianthus* species**

To compare the difference of Vietnamese native *Rhododendron* and *Enkianthus* species and *Rhododendron* and *Enkianthus* species of the world, we have conducted the analysis of the sample of Vietnamese and the world species based on the analysis coming aligned columns. The results were compared the ITS sequences of some Vietnamese native *Rhododendron* and *Enkianthus* species and the world based on GeneBank.

#### **Alignment of DQ1 (*Rhododendron*) species**

Through the analysis aligns upcoming column (alignments), ITS sequences of DQ1 Vietnamese species with three species of *Rhododendron* with accession number |AJ626913.1|, |AJ626910.1| and |DQ295782.1| in the world, these results showed that in some statistical tables nucleotides 4 ITS sequences of species with fluctuations *Rhododendron* number of nucleotide sequences from 712- 809 nucleotides when it was compared with the sample sequence revealed the DQ1 had 7 different nucleotides (Fig. 2). This result demonstrated that DQ1 species was not much different from the world's *Rhododendron* species.

#### **Alignment of DQ3 (*Enkianthus*) species**

Based on the results of the Vietnamese native *Enkianthus* DQ3 sequences with two species of *Enkianthus* with accession number KP092597.1, KU350168.1 in the world, we found that these results disclosed the number of nucleotides in the 3 ITS sequences of the *Enkianthus* genome had the nucleotide number variation of the sequences from 656-711 nucleotides and the QD3 had 10 difference in sequence when compared with two reference species *Enkianthus*, KP092597.1 and KU350168.1. Considering the two reference species, it can be concluded that the DQ3 species may be *Enkianthus perulatus* because it did not differ much from the world's with KU350168.1 *Enkianthus perulatus* species as shown in Fig. 3.



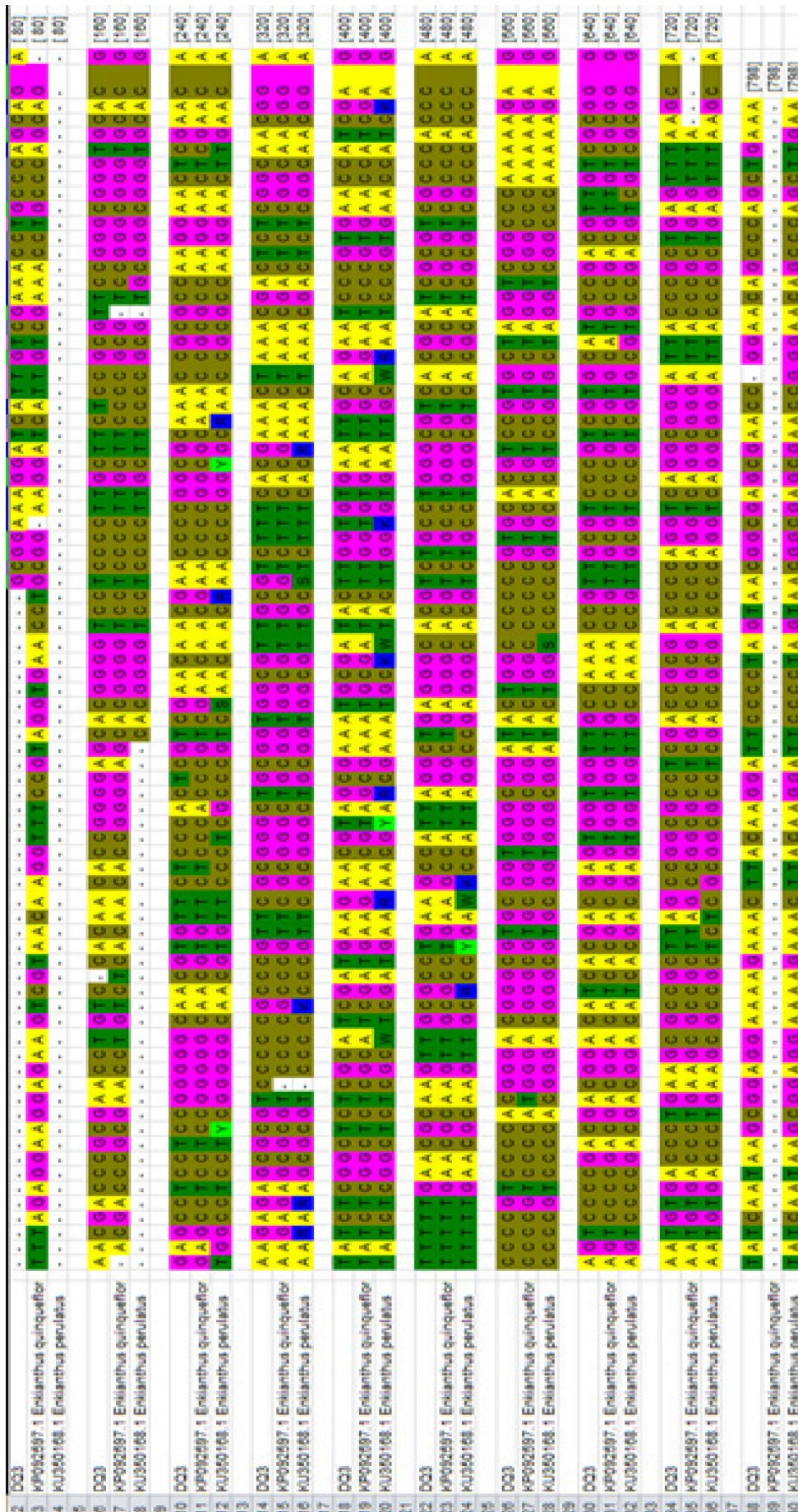


Figure 3. Alignment of Vietnamese native *Enkianthus*- DQ3 with two sequences of accession number (*Enkianthus*) KP092597.1, and KU350168.1.

Similarly, most of Vietnamese native *Rhododendron* and *Enkianthus* species were also analyzed by alignment with *Rhododendron* and *Enkianthus* species of the world to determine the difference in the order identify species/subspecies based on the ITS regions. However, the comparison with the GenBank database on aim for a comparable result with similar taxa with the reference sequence. BLAST results could not conclude exactly the species. For instance BLAST similarities coverage and high identity sequences (99%), could not reverse impairment species name. BLAST results were only available showing the most homologous sequences in the GenBank. Therefore, in order to determine the correct species name which should conduct more of other analyses as compared with morphological data and determined the relationship through phylogenetic tree based on ITS region sequences.

### Phylogenetic trees base on ITS region sequences

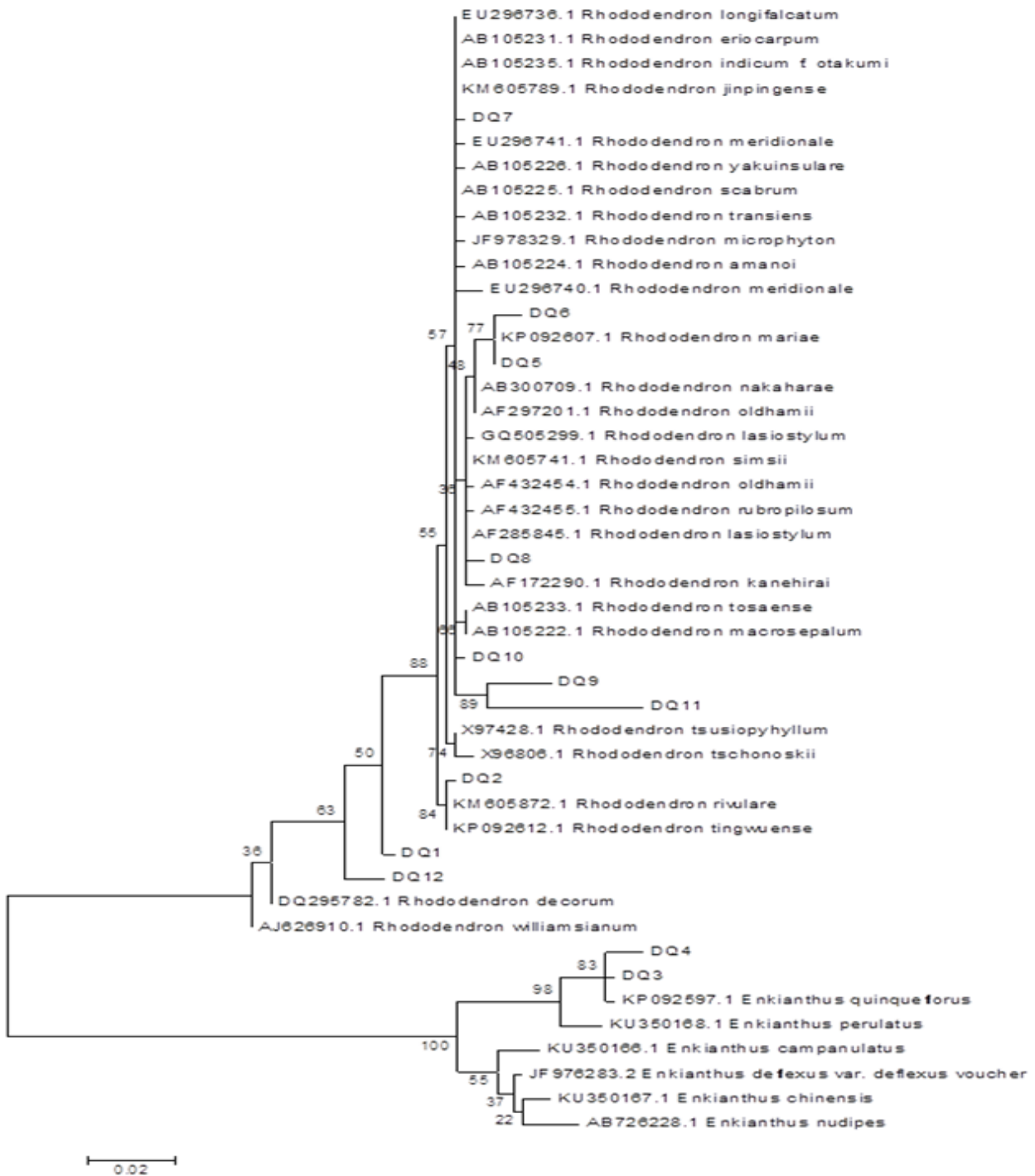
In order to launch the ITS database, it needs to construct the 12 species of *Rhododendron* and *Enkianthus* species in this study, we collected 34 references ITS sequences of (*Rhododendron*) and (*Enkianthus*) species which are closely related to 12 species (*Rhododendron*) and (*Enkianthus*). The alignment process was done using the SeaView software with the MUSCLE option [25]. The final dataset consists of 46 sequences consisting of 34 control groups (*Rhododendron*) and (*Enkianthus*) with an analytical sample and 12 ITS sequences from the analytical samples.

According the phylogenetic trees, 12 Vietnamese native *Rhododendron* and *Enkianthus* species trees were divided into 2 different groups (Fig. 4).

Group I included 10 Vietnamese native *Rhododendron* and 29 *Rhododendron* species in the world and they were divided into 7 different sub-groups.

The subgroup I.1 included DQ7 species and 11 *Rhododendron* species in the world. At the early divergence, the DQ7 species was in the same large group as the 11 *Rhododendron* sequences of the world but it was closed to the *Rhododendron jinpingense* as accession number KM605789.1. When it was comparing the DQ7 species on the Blast Sequence, the DQ7 species had 99% identity with of 99% *Rhododendron jinpingense* KM605789.1 species. Therefore, the DQ7 species were accurately identified as *Rhododendron jinpingense*.

The subgroup I.2 consisted of three Vietnamese *Rhododendron*: DQ5, DQ6 and DQ8 and they were divided into two subgroups. Two species of DQ5 and DQ6 were in the same group with three of the world's *Rhododendron* species. However, the DQ5 and DQ6 species were in the indistinguishable group as the *Rhododendron mariae* KP092607.1 in the world with a boost percentage of 77%, while the DQ5 species was identical and divided with accession number KP092607.1 at 100% bootstrap index. Therefore, DQ5 species was accurately identified as *Rhododendron mariae* species and DQ6 species would be *Rhododendron mariae* species but it was slight difference of genetic distance (Fig. 4).



**Figure 4.** Phylogenetic trees based on ITS sequences

The DQ8 species was same group with 6 *Rhododendron* accession number as GQ505299.1, KM605741.1, AF432454.1, AF432455.1, AF285845.1, AF285845.1 and AF172290., respectively. The bootstrap is quite far with 36%. However, DQ8 species was grouped with *Rhododendron kanehirai* AF172290.1 at 80% bootstrap index, so it could be confirmed that DQ8 was *Rhododendron kanehirai* species.

The subgroup I.3 was included three Vietnamese native *Rhododendron* species as DQ10, DQ9 and DQ11 and they were divided same group with two *Rhododendron* accession number AB105233.1 (*Rhododendron tosaense*) and AB105222.1 (*Rhododendron macrosepalum*). DQ9 and DQ11 species were in the similar group and have a boost percentage of 89%, while the DQ10 species was far away with two accession number: AB105233.1 and AB105222.1. Therefore, it could not identify three of DQ9, DQ10 and DQ11 species which they are kinds of *Rhododendron*. To



accurately determine the above species, we need to sequence some other genes such as *matK*, *rbcL*, *PsbA-trnH* *trnL-F* to classify *Rhododendron* species. The results also coincided with the reports of Kress et al. [26]; Sundberg et al. [27] and Milne et al. [28] by using only 1-2 ITS gene regions and *matK* only for recognizable in plant.

The subgroup I.5 included the sequence of the Vietnamese native *Rhododendron* DQ2 and two accession numbers as *Rhododendron rivulare* KM605872.1 and *Rhododendron tingwuense* KP092612.1. The DQ2 species has 84% bootstrap index when it compared to *Rhododendron rivulare* KM605872.1, so it could be confirmed that DQ2 was *Rhododendron rivulare* species.

The subgroup I.6 and I.7 were consisted of two Vietnamese native *Rhododendron* DQ1 and DQ12 species, which were isolated in two subgroups comparing to the world's *Rhododendron* species and they were relatively large genetic distance with each other. Based on the ITS sequencing, these two species were identified as belonging to the genus *Rhododendron*, but it was unable to identify which species belongs to the genus *Rhododendron*. Therefore, it is necessary to sequence some regions to accurately determine which samples of DQ1 and DQ12 as be a member of the species of *Rhododendron*.

Group II consisted two Vietnamese native *Enkianthus* DQ3 and DQ4 species and six accession numbers *Enkianthus* species that were divided in two subgroups as two Vietnamese native *Enkianthus* DQ3, DQ4 species and accession numbers KP092597.1 and KU350168.1 in group. Base on the phylogenetic trees, the DQ3, DQ4 species were same group with *Enkianthus quinqueflorus* KP092597.1 at 83% bootstrap index. It can be confirmed that two varieties of Vietnamese native *Enkianthus* were *Enkianthus quinqueflorus*.

The information through the ITS region sequencing have shown the majority of the collected samples in this study was ascertained compatible with morphological identification. Five per ten species were indentified by recognition molecule forms such as DQ2, DQ5, DQ6, DQ7, DQ8 and five species as DQ1, DQ9, DQ10, DQ11 and DQ12 samples were determined as in genus *Rhododendron* species. Two varieties of *Enkianthus* as DQ3 and DQ4 were identified as *Enkianthus quinqueflorus*. The success of this study would generate important prerequisite information for further expanding the sample size like (the same sample in the species, or more species in *Rhododendron* and *Enkianthus*), or selection some molecular markers (16 rRNA genes, gene *matK* ...) to construct DNA barcodes for valuable plants in Vietnam [24].

In summary, the species is still limited but the ITS sequences application revealed the strength of promotion of universal DNA barcode in the delimitation of species and subspecies level for plant group flowering. The results demonstrated that the location of varieties used in the phylogenetic tree based on sequence analysis of the ITS which helps to confirm correct scientific names of those varieties.

### Conflict of Interest

The authors declare no conflicts of interest.

### References

- [1] M.J.M. Christenhusz, J.W. Byng, The number of know plants species in the world and its annual increase, *Phytotaxa*. 261 (2016) 201–217.
- [2] D.J. Mabberly, *The plant book: A portable dictionary of the higher plants*, Cambridge University Press, 1987.
- [3] N.T.T. Huong, Key to the genera of the family *Ericaceae* (Juss.) in Vietnam, 4<sup>th</sup> National Conference on Ecology and Natural Resources. 25 (2011) 148–150.
- [4] W.S. Judd, K.A. Kron, Circumscription of *Ericaceae* (Ericales) as determined by preliminary cladistic analyses based on morphological, anatomical and embryological features, *Brittonia*. 45 (1993) 99–114.

- 
- [5] K.A. Kron et al., Phylogenetic classification of Ericaceae: molecular and morphological evidence, *Bot. Rev.* 68(3) (2002) 335–423.
- [6] D.F. Chamberlain et al., *The genus Rhododendron: its classification and synonymy*, first edition, Royal Botanic Garden Edinburgh, 1996.
- [7] J. Brown, *Tales of the rose tree: ravishing Rhododendrons and their travels around the world*, First edition, David R. Godine Publisher, 2004.
- [8] A. Gunduz et al., Wild flowers and mad honey, *Wild Environ. Med.* 18 (2007) 69–71.
- [9] P. Kumar, Assessment of impact of climate change on *Rhododendrons* in Sikkim Himalayas using Maxent modelling: limitations and challenges, *Bio Conserv.* 21 (2012) 1251–1266.
- [10] P.A. Cox, K.N.E. Cox, *The encyclopedia of Rhododendron species*, Glendoick Publishing, Scotland, 1997.
- [11] M. MacKay, S.E. Gardiner, A model for determining *ex situ* conservation priorities in big genera is provided by analysis of the subgenera of *Rhododendron* (Ericaceae), *Bio Conser.* 26 (2016) 189–208.
- [12] R. Popescu, B. Kopp, The genus *Rhododendron*: An ethnopharmacological and toxicological review, *J. Ethnopharmacol.* 147(1) (2013) 42–62.
- [13] A.A. Anderberg, Cladistic analysis of *Enkianthus* with notes on the early diversification of the *Ericaceae*, *Nordic J. Bot.* 14 (1994) 385–401.
- [14] A.K.M. Golam Sarwar, H. Takahashi, Pollen morphology of *Enkianthus* (Ericaceae) and its taxo-nomic significance, *Grana.* 45(3) (2006) 161–174.
- [15] K.A. Kron, Phylogenetic relationships of *Empetraceae*, *Epacridaceae*, *Ericaceae*, *Monotropaceae*, and *Pyrolaceae*: evidence from nuclear ribosomal 18S sequence data, *Ann. Bot.* 77 (1996) 293–303.
- [16] A.A. Anderberg et al., Phylogenetic classification of *Ericaceae*: molecular and morphological evidence, *Bot. Rev.* 68 (2002) 335–423.
- [17] Z.N. Liu et al., Phylogeny of *Pyroleae* (Ericaceae): Implications for character evolution, *Plant Res.* 124 (2011) 325–337.
- [18] M.V. Pho, The genus *Rhododendron* in Vietnam, *J. Res. Dev.* 73 (2009) 51–54.
- [19] B. Zhao et al., Genetic relationship among nine *Rhododendron* species in Qinling Mountains, China using amplified fragment length polymorphism markers, *Pak. J. Bot.* 47(3) (2015) 1069–1074.
- [20] K.A. Kron, W.S. Judd, Systematics of the *Lyonia* group (*Andromedeae*, *Ericaceae*) and the use of species as terminals in higher-level cladistic analyses, *Syst. Bot.* 22 (1997) 479–492.
- [21] C.C. Tsai et al., Genetic relationships of *Rhododendron* (Ericaceae) in Taiwan based on the sequence-of the internal transcribed spacer of ribosomal DNA, *J. Hortic. Sci. Biotechnol.* 78 (2003) 234–240.
- [22] K.H. Trung et al., Molecular phylogeny of the endangered Vietnamese *Paphiopedilum* species based on the internal transcribed spacer of the nuclear ribosomal DNA, *Adv. Biol. Studies.* 5(7) (2013) 337–346.
- [23] W. Xu et al., Differentiation of *Dendrobium* species used as "Huangcao Shihu" by rDNA ITS sequence analysis, *Planta Med.* 72(1) (2005) 89–92.
- [24] T.D. Duong et al., Identification of Vietnamese native *Dendrobium* Species based on ribosomal DNA internal transcribed spacer sequence, *Adv. Stud. Biol.* 10 (2018) 1–12.
- [25] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research.* 32(5) (2004) 1792–1797.
- [26] W.J. Kress et al., Use of DNA barcodes to identify flowering plants, *Proc. Nat. Acad. Sci.* 102 (2005) 8369–8374.
- [27] P. Sundberg et al., DNA barcoding should accompany taxonomy - the case of *Cerebratulus* spp. (Nemertea), *Mol. Ecol Res.* 10(2) (2010) 274–281.
- [28] R.I. Milne, Phylogeny and biogeography of *Rhododendron* subsection *Pontica*, a group with a Tertiary relict distribution, *Mol. Phylog. Evol.* 33(2) (2004) 389–401.