ACADEMIA JOURNAL OF BIOLOGY 2021, 43(3): 1–8
DOI: 10.15625/2615-9023/15888

*rbc*L DNA REGION REVEALED AS THE BEST DNA BARCODE FOR IDENTIFICATION OF *Mahonia* AND *Berberis* SPECIES (Berberidaceae)

Nguyen Thi Phuong Trang[1,2]

[1]Institute of Ecology and Biological Resources, VAST, Vietnam
[2]Graduate University of Science and Technology, VAST, Vietnam

Received 22 February 2021, accepted 12 July 2021

ABSTRACT

Berberidaceae contains 17 genera and nearly 650 species, in which the genus *Berberis* and the genus *Mahonia* are sisters with a very close relationship and share many similar characteristics. In the past, the two genera were merged into the genus *Berberis.* At present, the Berberidaceae is at risk and is prioritized for conservation, especially the species of *Mahonia* and *Berberis*. In the fact, identification for the Berberidaceae species based on the morphology of parts for sale (such as roots, stems, leaves) without reproductive organs (flowers and fruits) is impossible.

In recent years, molecular biology techniques are being applied widely and effectively in research on the evolution, classification and genetic diversity of populations. Study based on DNA has highly accurate and particularly useful for closely related species which morphological observations are not sufficient to distinguish. Results from DNA analysis allow authentic species, populations or individuals from un-intact specimens accurately and especially, it is not affected by objective factors such as the environment or human.

In this study, we assessed the taxonomy ability of three commonly DNA barcoding regions used for classification including *rbc*L, *trn*H-*psb*A and *ITS*2 on 12 samples of the Berberidaceae family, in which 7 samples are genus *Mahonia*, 4 samples are genus *Berberis* and one sample of *Epimedium* was used as the control sample. The results of the study will contribute to the selection of suitable DNA barcode for the identification of *Mahonia* and *Berberis* samples.

The results demonstrated that the *rbc*L region showed the most obvious ability to distinguish the 12 Berberidaceae species. The *ITS*2 sequences of *Berberis julianeae* and *Mahonia bealei* of Vietnam were submitted to GenBank with accession numbers MT073031 and MT008067, respectively. The sequence of *rbc*L of *Mahonia bealei* was also submitted to GenBank with accession number MT457415.1.

Keywords*:* Berberidaceae, *ITS*2, *rbc*L, *trn*H-*psb*A, selection of candidate DNA markers.

s

Citation: Nguyen Thi Phuong Trang, 2021. *rbc*L DNA region revealed as the best DNA barcode for identification of *Mahonia* and *Berberis* species (Berberidaceae). *Academia Journal of Biology*, 43(3): 1–8. https://doi.org/10.15625/2615-9023/15888

**Corresponding author email*: nptrang@gmail.com

INTRODUCTION

The first person to mention the taxons of the Berberidaceae family was Linnaeus in 1753. Berberidaceae is a well-known traditional herbal medicine family including 17 genera and about 650 species that are mainly distributed in the tropical, subtropical and temperate regions (Ban Nguyen Tien, 2003). According to Nguyen Tien Ban (2003), in Vietnam, the Berberidaceae family has 4 genera and 9 species, mainly distributed in the highlands of the northern mountainous provinces and Lam Dong province. All the species of the Berberidaceae family are valuable medicinal plants with high economic value, but due to over-exploitation and trading, they have been sharply declined in both quantity and quality. At present, all the species of the Berberidaceae family are listed in the Vietnam Red Book at Endangered Level (EN) and are classified in the IA group of Cites in Vietnam (the endangered species, prohibited from export, import, re-export, import from the sea and transit of specimens exploited from the wild for commercial purposes). Identifying species from parts of an organ (such as stems, roots, or leaves) without reproductive organs (flowers and fruits) based on morphology is impossible.

DNA barcode is a modern technique that uses short DNA fragments to standardize differentiation between species [4, 5]. They have become a new tool for forensics, classification, evaluation of genetic relationships, quality management, the origin of the biological products [6]. In flowering plants, some chloroplast genomic regions (such as *mat*K, *rbc*L, *psb*A-*trn*H, *atp*F-*atp*H ...) and nuclear genome region (such as *ITS*-rDNA, *18S*,...) are widely used in studies of phylogeny, taxonomy, and species identity [7, 8]. However, for each different family, the classification ability of these gene regions is different. The study of Song et al. (2019) [9] published in the journal Horticulture showed that 3 genomic regions *trn*C-*psb*D, *ndh*D and *atp*A-*atp*H are the 3 most effective gene regions to distinguish species in the genus Plum-apricot (*Prunus*). Muellner et al. (2011) [10] demonstrated that the ITS gene region was most effective for the identification of the species in Meliaceae family. Trang et al. (2015) [11] demonstrated that the combination of three gene regions *rbc*L, *mat*K and *trn*H-*psb*A was effective in distinguishing species of genus *Hopea* (Dipterocapaceae). Stoeckle et al. (2011) [12] used a combination of *rbc*L and *mat*K gene regions to identify rare medicinal plants that are heavily traded in North Africa, but their one obstacle is the lack of DNA comparable data. Up to now, *rbc*L *ITS* and *trbH-psb*A are the most popular gene regions often used for identifying forensic of the plant species.

In this study, we sequenced the *rbc*L, *trn*H-*psb*A and *ITS*2 genes of 12 Berberidaceae species to test the classification ability of three DNA barcoding markers on the species of *Mahonia* and *Berberis* (Berberidaceae). This study also supplemented the DNA database for the International gene bank, contributing to a more complete database for the genebank of Berberidaceae species distributed in Vietnam.

MATERIAL AND METHODS

Materials

A total of 12 leaf samples from seven *Mahonia* samples, four *Berberis* samples and one *Epimedium* sample were collected by the research team of the Department of Plant Ethnology, Institute of Ecology and Biological Resources during the trips by 2020 in several provinces of the north and highland Vietnam. Symbol and collected places of the studied samples were shown in Table 1.

Methods

DNA extraction

Total genomic DNA was extracted according to the method of Doyle and Doyle (1990) [13] under local laboratory conditions.

PCR amplification

*ITS*2 region with 300 bp, *trn*H-*psb*A with 400 bp and *rbc*L region with 600 bp in length were amplified using universal ITS1/ITS2 primer (Cheng et al., 2016) [14], *rbc*LF/R and

*trn*H-*psb*A (Kress et al., 2005) [5]. PCR was performed in 25 µl of reaction system containing 7 µl deionized H_2O, 12.5 µl of PCR Master mix kit (2X); 1.25 µl of each primer (10 pmol/µl); 3 µl of DNA template (10–20 ng). The PCR reaction was performed using PCR Model 9700 (GeneAmp PCR System 9700, USA) for 3 min at 94 ºC for denaturation, 35 amplification cycles (45 s at 94 ºC for denaturation, 30 s at 55 ºC annealing and 30 s at 72 ºC for extension), then 10 min at 72 ºC for extension and hold at 4 ºC.

Table 1. Symbol and collected places of the studied samples

Order	Symbol	Scientific name	Collected places
1	BHG07	*Mahonia* sp.4	Quan Ba, Ha Giang province
2	BHG04	*Mahonia* sp.5	Dong Van, Ha Giang province
3	DL 01	*Mahonia klossii*	Da Lat, Lam Dong province
4	DL 02	*Mahonia klossii*	Da Lat, Lam Dong province
5	BHG09	*Mahonia* sp.3	Vi Xuyen, Ha Giang province
6	BCB05	*Mahonia* sp.1	Thong Nong, Cao Bang province
7	BBK03	*Mahonia* sp.2	Ngan Son, Bac Kan province
8	SB 01	*Berberis julianea*	Sa Pa, Lao Cai province
9	SB 02	*Berberis julianea*	Sa Pa, Lao Cai province
10	SB 03	*Berberis julianea*	Sa Pa, Lao Cai province
11	SB 04	*Berberis julianea*	Sa Pa, Lao Cai province
12	DDH	*Epimedium* sp.	Sample bought from China

Sequence analysis and alignment

The PCR product was screened by electrophoresis on 1% agarose gel, then sequenced at FirstBase Co. Ltd. (Malaysia). Raw sequences obtained were assembled and edited by Chromas-Pro 2.1.6 [15]. All sequences were then aligned on BLAST, Genbank (http: www.ncbi.nlm.nih.gov/BLAST). The pairwise distance was determined using Mega 7.0 [16]. The phylogenetic trees were constructed using Maximum Likelihood (ML) and Bayesian Inference (BI) with a bootstrap value of 1000.

RESULTS AND DISCUSSION

Total DNA of 12 study samples were successfully DNA extracted with ratio $OD_{260/nm}/OD_{280nm}$ of all goten total DNA range from 1.81–1.83. Total DNA of all samples were used as templates for amplifying *ITS*2-rDNA (300 bp), *trn*H-*psb*A (400bp) and *rbc*L (600 bp) regions (Fig.1).

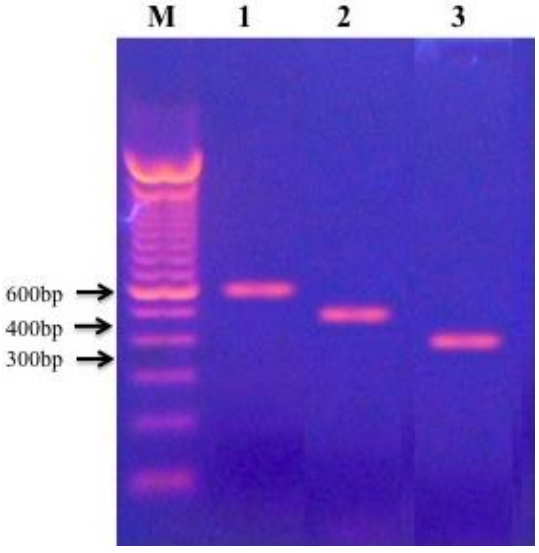


Figure 1. PCR products of 3 samples representing 3 amplified gene regions tested by electrophoresis on 1% agarose gel (Lane M: DNA 100 bp ladder, lane 1: PCR products of *rbc*L, lane 2: PCR products of the trnH-psbA and lane 3: PCR products of the ITS2

Sequencing analysis

All PCR products after purification are sequenced at Firstbase Co. Ltd. The results of sequencing were then checked for similarity with the available sequences on the GeneBank (GB) using the BLAST tool. Usually, BLAST results do not give exact conclusions about the species, but in the cases where BLAST has high coverage and similarity (above 98%), it may suggest the species which have the most closely relationship to the study sample. In this study, the sequences of 3 DNA regions *ITS*2, *rbc*L, *trn*H-*psb*A of all 12 study samples showed high similarity index (over 95%) corresponding to *ITS*2, *rbc*L, *trn*H-*psb*A DNA regions of the *Mahonia* and *Berberis* species on GenBank. This demonstrated that the initial morphological identification of the samples was correct, as well as demonstrated that our cloned DNA sequences were successful.

The sequence of three DNA regions including *rbc*L, *ITS*2 and *trn*H-*psb*A were analyzed by Mega 7.0 software, aligned and assessed the classification ability of these three DNA regions base on the genetics distance of 12 samples. The genetic relationship diagram of 12 study samples based on sequences of *rbc*L (A), *trn*H-*psb*A (B), *ITS* (C) was shown in Figure. 2.

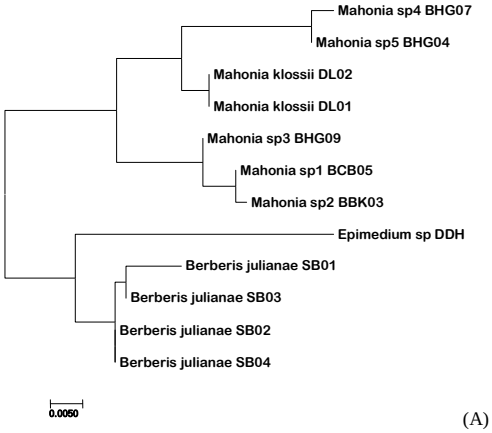
Results of phylogeny based on the method of Maximum Likelihood (ML) of 12 study samples (Fig. 2) showed that:

When analyzing based on the sequences of *rbc*L with 600 bp in length, 12 samples of Berberidacae family were separated into 3 different branches, 7 *Mahonia* samples separated into the first branch, 4 Berberis samples in the second branch, and *Epimedium* sp. in the third clade and closer to the species of the genus *Berberis* than *Mahonia* (Fig. 2A).

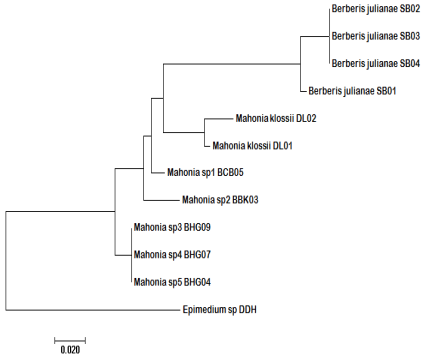
When analyzing based on the sequence of *trn*H-*psb*A with 400 bp in length, 4 *Berberis* samples were separated into one branch, 7 *Mahonia* samples in another branch and the control species *Epimedium* sp. always in an independent branch from the two groups of *Berberis* and *Mahonia*. However, 7 *Mahonia* samples continued to be divided into 2 groups, of which 3 *Mahonia* samples that were collected in Ha Giang (BHG 04, BHG 07, BHG 09) concentrated on a small branch, while the remaining 4 *Mahonia* samples separated into other small branches. BBK03 and BCB05 samples collected in Cao Bang and Bac Kan province were separated into 2 different branches, 2 samples DL01 and DL02 that collected in Da Lat (Lam Dong province) were in the same branch. This separation can be considered as an attractive result because in terms of taxonomy at the species level, *trn*H-*psb*A gave correct results when accurately dividing 7 *Mahonia* samples and 4 *Berberis* samples into several smaller branches following exactly their collected places. So, it seems that the *trn*H-*psb*A region can be divided by geographical samples. However, to prove this, it is necessary to have a larger number of samples as well as need more samples of the same species collected at different locations for control. In the present study, the goal was to evaluate the ability of identification at the species level, so if the *trn*H-*psb*A DNA region is really able to distinguish both species and subspecies (or geographic species) is considered inappropriate because it is easy to confuse, leading to false conclusions about the specimen to be identified. Therefore, from the perspective of using molecular markers to classify at the species level, the *trn*H-*psb*A gene region would be inconsistent. However, in the study, *trn*H-*psb*A will be the DNA region containing many attractions, because it is seemly that this DNA region will help assess the dissociation capacity of geographic species or subspecies of Berberidaceae.

For the *ITS*2-rDNA region: Unlike 2 genomic regions *rbc*L-600 and *trn*H-*psb*A, 12 samples belonging to 3 genera when analyzed based on ITS2-sequence were divided only into 2 main branches, the sample of *Epimedium* sp. in one large branch and the remaining 11 samples in the second largest branch. The *Mahonia* and *Berberis* samples are located alternately in small branches, not clearly divided into 2 branches as the obtained

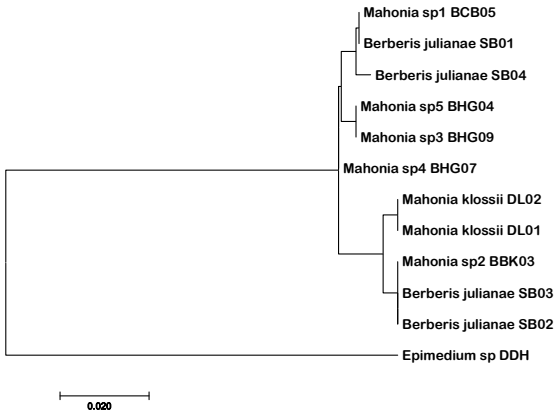
results when analyzing based on the sequences of *rbc*L-600 bp and *trn*H-*psb*A. *ITS*2 in the nuclear genome is more stable than the DNA in the chloroplast genome. Normally, samples that have genetic distance are not large enough, it is difficult to separate if based on the sequence of the stable nuclear DNA.



(A)



(B)



(C)

Figure 2. The genetic relationship diagram of 12 study samples based on the sequence of *rbc*L (A), *trn*H-*psb*A (B), *ITS* (C)

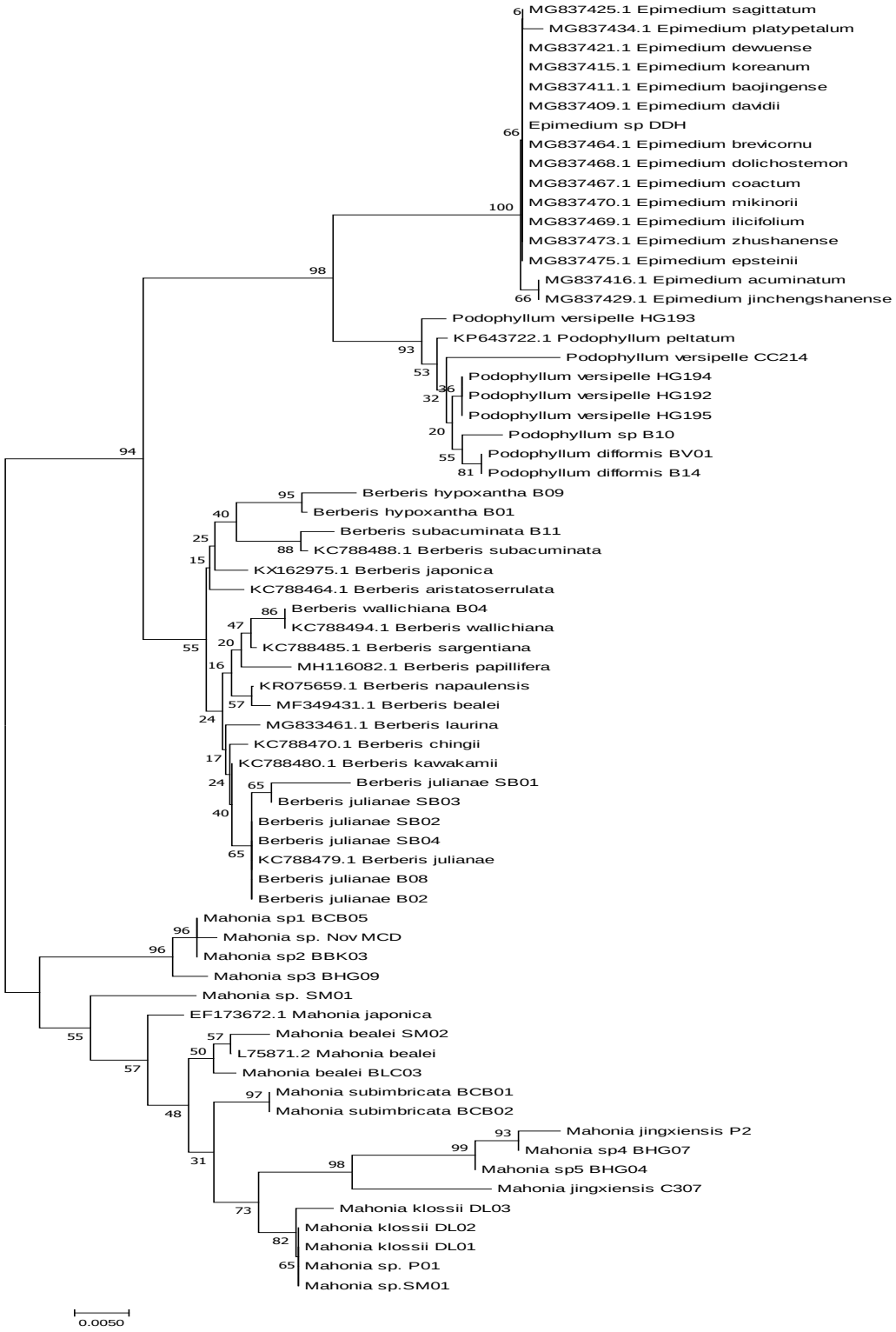


Figure 3. Diagram of genetic relationship of the 12 studied species, compared with other species of Berberidaceae on the GenBank base on *rbc*L sequence

Mahonia and *Berberis* are sisters that share many similar morphological characteristics and often confuse together. So far, based on morphology, 2 genera were classified in 1 genus namely *Berberis*. So, in this study, it is not surprising that *ITS*2-300 bp could not clearly divide the species of these two genera into two separate groups. This is also demonstrated that *Mahonia* and *Berberis* are closely genera with very small distance genetic.

From gotten result as above, we can see *rbc*L is the best DNA region that can be used to identify Berberidaceae species compared to *ITS*2 and *trn*H-*psb*A.

*rbc*L DNA with 600 bp in length was then used to construct phylogenetic trees for the 12 study samples, compared with other Berberidaceae species (data from GenBank) (Fig. 3). The results showed that 4 samples including SB01, SB02, SB03, SB04 were initially identified as *Berberis julianae*, and gave 100% match with *Berberis julianae* (accession number KC788479.1). This proves that the initial morphological identification for these 4 species was completely accurate. The two samples BHG07 and BHG04 have not been accurately identified by morphology. After comparison, it is in the same clade with *Mahonia jingxiensis*, however the rate of similarity was only 98%. Following 3 specimens of BHG09, BCB05 and BBK03 are also 3 *Mahonia* species. However, due to the lack of data on GenBank, they can not conclude exactly the scientific names for these 3 species. Based on genetic distance analysis, these 3 species have a close genetic relationship, they were in the same group in the *Mahonia* branch (Fig. 3).

CONCLUSION

All species of Berberidaceae have medicinal and high economic value, they are exhausted in the wild. Currently, they are listed in the Red Book and Group IA of CITES (Conversion on International Trade in Endangered Species of Wild Fauna and Flora). In Berberidaceae, specifically, *Mahonia* genus and *Berberis* genus are two genera has closely related, the species in two genera shared many similarities of morphological characteristics (when plants are still in their immature stage), it is really difficult to identify species based on morphology. Therefore, sofar, *Mahonia* genus and *Berberis* genus were classified into the genus *Berberis.*

Young Dong Kim was the first person who used DNA sequences to classify species of the Berberidaceae family. All research results of Kim et al. (1996, 2004) (were sequenced of *ITS*, *rbc*L (1996) and *ndh*F (2004)) were supported for the division species of *Mahonia* genus and species in B*erberis* genus into two separate genera even they are closely related.

So, apply molecular biology techniques for identification the species of Beberidaceae family is necessary. The selection of the most suitable DNA barcode for the species identification of Berberidaceae family was performed in the present study and the results showed that the *rbc*L region is suitable for Berberidaceae identification at species level. There is also still lack of data on the DNA sequence of the species of Berberidaceae family, so updating and adding molecular data about these species to have enough database for comparison, identification is also necessary and important.

The sequences of *ITS*2 of *Berberis julianeae* and *Mahonia bealei* of Vietnam have been submitted to GenBank with accession numbers MT073031 and MT008067, respectively. The sequence of *rbc*L of *Mahonia bealei* is also submitted to GenBank with accession number MT457415.1.

REFERENCES

Ban Nguyen Tien, 2003. List of plant species in Vietnam. Volume 2. Agricultural Publishing House, Hanoi.

CBOL Plant Working Group, 2009. A DNA barcoding for land plants. *PNAS*, 106(31): 12794–12797.

Chen S., Hui Y., Jianping H., Chang L., Jingyuan S., Linchun S., Yingjie Z., Xinye M., Ting C., Xiaohui P., Kun L., Ying L., Xiaocheng J., Yulin L., Christine L., 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species, *PLoS One*, 5(1): e8613.

Cheng T., Chao X., Li L., Changchao L., Yu Z. and Shilian Z., 2016. Barcoding the kingdom plantae: new PCR primer for ITS region of plants with improved universality and specificity. *Molecular Ecology Resources*, 16: 138–149.

Chromas Pro2.1.6. https://chromas.software.informer.com/download/.

Doyle J. J., Doyle J. L., 1990. Isolation of plant DNA from fresh tissue. *Focus*, 12: 13–15.

He J. M., Mu Q., 2015. The medicinal uses of the genus Mahonia in traditional Chinese medicine: An ethnopharmacological, Phytochemical and pharmarcological review. *J. Ethnopharmacol*, 175: 668–683.

Kress W. John, Kenneth J. Wurdack, Elizabeth A. Zimmer, Lee A. Weigt, and Daniel H. Janzen, 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences*, 102(23): 8369–74.

Kumar S., Stecher G., Tamura K., 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets, *Molecular Biology and Evolution*, 33(7): 1870–1874.

Lahaye R., van der Bank M., Bogarin D., Warner J., Pupulin F., Gigot G., Maurin O., Duthoit S. Barraclough T.G., Savolainen V., 2008. DNA barcoding the floras of biodiversity hotspots, *Proceedings of the National Academy of Sciences of the United States of America*, 105: 2923–2928.

Liu J., Provan J., Gao L. M., Li D. Z., 2012. Sampling strategy and potential utility of indels for DNA barcoding of closely related plant species: A case study in taxus, *International Journal of Molecular Sciences*, 13: 8740–8751.

Muellner A. N., Schaefer H., Lahaye R., 2011, Evaluation of candidate DNA barcoding loci for economically important timber species of the mahogany family (Meliaceae). *Molecular Ecology Resources*, 11(3): 450–460.

Nguyen Thi Phuong Trang, N. M. Duc, N. V. Sinh, Triest L., 2015. Applictaion of DNA barcoding markers for identification of *Hopea* species. *Genetic Molecular Research (GMR)* 14(3): 9181–9190.

Stoeckle M. Y., Catherine C., Gamble R., Kirpekar G. Y., Selena A., Damon P. L., 2011. Commercial teas highlight plant DNA barcode identification successes and obstacles. *Nature Research Journal*, 1(42): 1–7

Song X., Ting S., Wenjie L., Xiaopeng N. Shahid I., Zhaojun N., Xiao H., Dan Y., Zhijun S. And Zhihong G., 2019. Comparative analysis of the complete chloroplast genome among *Prunus mume*, *P. armeniaca*, and *P. salicina*. *Horticulture Research*, 6(89): 2–13.

Young Dong Kim, Robert K. Jansen, 1996. Phylogenetic implications of *rbc*L and ITS sequence variation in the Berberidaceae. *Systematic Botany* vol 21(1): 381–396.

Young Dong Kim, Sung Hee Kim, Chul Hwan Kim, Robert K. Jansen, 2004. Phylogeny of Berberdaceae based on sequences of the chloroplast gene ndhF. *Biochemistry Systematics and Ecology*, 32: 291–301.

Vietnam Ministry of Sciences and Technology, 2007. Vietnam Red Book Part II - Plants. Natural Science and Technology Pub., Hanoi.