INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
22–0319/2023/29–3–227–234
DOI: 10.17957/IJAB/15.2024
http://www.fspublishers.org

Full Length Article

DNA Barcoding and Phylogenetic Analysis of the Species in the Genus *Alpinia*

Nguyen Nhat Linh[1†], Le Quang Trung[2†], Nguyen Quoc Binh[3], Le Thi Thu Ha[1], Pham Le Bich Hang[1], Nguyen Phuong Hanh[4], Luu Han Ly[1], Nguyen Hai Ha[1,5] and Le Thi Thu Hien[1,5*]

[1]*Institute of Genome Research, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam*

[2]*VNTEST Institute for Quality Testing and Inspection, Nguyen Xien, Thanh Xuan, Hanoi, Vietnam*

[3]*Vietnam National Museum of Nature, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam*

[4]*Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam*

[5]*Graduate University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam*

[†]Co-first authors and contributed equally to this work

[*]For correspondence: hienlethu@igr.ac.vn

Received 11 August 2022; Accepted 17 November 2022; Published 27 February 2023

Abstract

Alpinia Roxb. is the largest and most taxonomically complex genus of the flowering plant family Zingiberaceae. Internal transcribed spacer (ITS) region of ribosomal DNA has been used to resolve the phylogenetic relationships in Zingiberaceae family and its genera including *Alpinia*, *Globba*, and *Amomum*. In this study, the ITS region was used as additional molecular data for species classification in the genus *Alpinia*. Genomic DNAs from leaf samples of 23 *Alpinia* species and 11 taxa collected in Vietnam were isolated and used as templates for PCR amplifications of the ITS region. Phylogenetic trees were constructed *via* both Neighbor Joining and Maximum Likelihood methods using 36 reference *Alpinia* sequences along with 30 sequences obtained in this study. Results demonstrated that DNA barcoding using the ITS region is a reliable tool for supporting the identification of *Alpinia* species. Morphological characteristics and ITS sequences help to better understand the phylogenetic relationship of the species in the genus *Alpinia* distributed throughout Vietnam. © 2023 Friends Science Publishers

Keywords: *Alpinia*; DNA barcode; ITS; Phylogenetic relationship; Species Identification

Introduction

Alpinia Roxb. is the largest and most taxonomically complex genus of the flowering plant family Zingiberaceae. This genus has approximately 250 species distributed among areas of tropical and subtropical climates, including Asia, Australia, and the Pacific Islands (Smith 1990; Larsen 1998; Vu *et al*. 2019). Most *Alpinia* species are commonly grown for their flowers (*e.g.*, *A. purpurata),* while others have economic potential and are used as spices (*e.g.*, *A. galanga*) and medicines (*e.g.*, *A. bracteata*) (Wu 2000; Kress *et al*. 2005; Uma and Muthukumar 2014). According to Smith's classification, the genus *Alpinia* consists of two subgenera, *Alpinia,* and *Dieramalpinia*. In Vietnam, most of the *Alpinia* species belongs to the subgenus *Alpinia* which was divided into four groups, *Dydimanthus*, *Alpinia*, *Guillania*, and *Allughas* (Smith 1990). Recently, several species have been recorded for Vietnam's flora such as *A. rugosa* in Thua Thien-Hue province, *A. graminifolia* in Quang Ninh and Bac Giang provinces, and *A. coriandriodora* in Bac Kan province (Le *et al*. 2017; Nghiem *et al.* 2018; Vu *et al.* 2019). By early 2019, 36 species of *Alpinia* have been recorded to be found throughout the country (Vu *et al*. 2019). Sixteen species are traditionally used as medicines by the Vietnamese people to treat common illnesses, such as stomachache, indigestion, cholera, dysentery, diarrhea, vomiting, excessive urination at night, and natural ejaculation. Various parts of the plant (rhizome, tuber, leaves, flowers, seed, and fruit) are used to remedy these symptoms, with the rhizome being the most commonly used part (Nguyen *et al*. 2014; Nghiem *et al*. 2018).

To cite this paper: Linh NN, LQ Trung, NQ Binh, LTT Ha, PLB Hang, NP Hanh, LH Ly, NH Ha, LTT Hien (2023). DNA barcoding and phylogenetic analysis of the species in the genus *Alpinia*. *Intl J Agric Biol* 29:227–234

Over the last two decades, DNA barcoding has been rapidly developed as an useful tool for species classification, biodiversity investigation and conservation, molecular phylogeny and evolutionary studies (Kang *et al*. 2017). The method is based on the principle of comparing short and universal DNA sequences from standard regions of the genome that have efficiently high evolution rates, allowing it to be appropriate for classifying members of a specific genus (Hebert *et al*. 2003). The advantage of this molecular approach is that the starting material can be as small as a sample of a plant tissue, and the identification process is fast and reproducible (Hartvig *et al*. 2015). DNA barcodes utilized for plant taxonomic classification belong to the internal transcribed spacer (ITS) region in the nuclear genome and *psb*A-*trn*H, *mat*K, *rbc*L, *trn*L-*trn*F in the chloroplast genome (Kress *et al*. 2005; Kress and Erickson 2007; CBOL Plant Working Group *et al*. 2009; Panaligan *et al*. 2021).

Among these DNA barcodes, ITS is the most widely used marker in plant phylogenetic studies because of its high resolution of inter- and intraspecific discrimination (Cheng *et al*. 2016; Keskin *et al*. 2017). Previous studies indicated that ITS possessed greater discriminatory ability when compared to other markers from chloroplast genomes (Hollingsworth *et al*. 2011; Huang *et al*. 2015). However, the main drawbacks of using this region as a core universal DNA barcode for plant classification are results of the incomplete concerted evolution of multiple copies, different alleles from paternal and maternal parents, DNA contamination of different species, amplification and sequencing success rate, and other technical problems (China Plant BOL Group *et al*. 2011; Hollingsworth *et al*. 2011). This region belongs to ribosomal DNA in the nuclear genome (Kang *et al*. 2017) and is comprised of the ITS1 intergenic spacer, *5.8S* rDNA, and the ITS2 intergenic spacer, whose size ranges from 400 to 1000 bp in total (Álvarez and Wendel 2003). Among these three partial sequences, 5.8S is the most conserved region while other two spacers possess high discriminatory ability with an abundance of variable sites (Hollingsworth *et al*. 2011). ITS helped resolve the phylogenetic relationships in Zingiberaceae family and its genera including *Alpinia*, *Globba*, and *Amomum* (Vinitha *et al*. 2014). Using ITS and *trn*L-F sequences, the phylogeny of tribe Zingibereae was studied (Ngamriabsakul *et al*. 2003). ITS along with *trn*K-*mat*K were used for investigating the phylogeny, evolution, and classification of the *Globba* genus (Williams *et al*. 2004). The molecular phylogenetic analysis based on multiple accessions of ITS and *mat*K regions of *Alpinia*, *Amomum*, *Elettaria*, *Elettariopsis*, *Geocharis*, *Geostachys*, and *Hornstedtia* genera revealed that *Alpinia* genus consists six clades (Boer *et al*. 2018). Within the genus *Alpinia*, Kress *et al*. (2005) reported the most extensive phylogenetic analysis based on molecular characteristics using ITS and *mat*K regions. This study combined the morphology (Smith 1990) and molecular based analyses to build a six-clade classification system for the genus *Alpinia* (Kress *et al*. 2005). ITS1 was also used as molecular evidence in Qiao's analysis to differentiate an *Amomum* species from *Alpinia* (Qiao *et al*. 2009). Tan *et al*. (2020) demonstrated the high species identification of *Alpinia* species collected in Peninsular Malaysia using the ITS2 region. The efficacy of 4 barcoding loci including *ycf*1b, *rbc*L, ITS and ITS2 were evaluated on 13 species belonging to 4 genera of Zingiberaceae (Saha *et al*. 2020).

In Vietnam, there are still difficulties in species identification among the genus *Alpinia* due to similarities in morphological characteristics and the lack of DNA barcode studies. Thus, in the present study, the ITS region was used as additional molecular data for species classification in the genus *Alpinia*. Our aim is to obtain a better understanding of phylogenetic relationship of the *Alpinia* species distributed throughout Vietnam. These molecular data provide supportive information for identification of sampled species and the phylogeny data are useful for further investigation on the divergence and branching of species and selected clades within the genus *Alpinia* and family Zingiberaceae.

Materials and Methods

Materials

Forty-four leaf samples from 23 species *Alpinia* and 11 taxa were collected from different regions throughout Vietnam from 2010 to 2018 and stored on silica-gel within 24 hours of collection till further use (Table 1 and Fig. 1). All specimens were morphologically identified by Nguyen Quoc Binh and Nguyen Phuong Hanh using comparative morphological method (Nguyen *et al*. 2017) and deposited at the Vietnam National Museum of Nature (VNMN). All laboratory work and bioinformatics analysis were performed at the Institute of Genome Research, Vietnam Academy of Science and Technology.

Methods

Total DNA extraction, and amplification of ITS region: Twenty milligrams of each of the lyophilized leaf specimens were used for total genomic DNA extraction using GeneJET Plant Genomic DNA Purification Kit (Thermo Fisher Scientific, USA), according to the protocol supplied by the manufacturer. The ITS region was amplified from the genomic DNA using DreamTaq DNA polymerase (Thermo Fisher Scientific, USA). The forward and reverse primers used to amplify the ITS sequence in this study were ITS-F (5'-ACG AAT TCA TGG TCC GGT GAA GTG TTC G-3') and ITS-R (5'-TAG AAT TCC CCG GTT CGC TCG CCG TTA C-3') (Sun *et al*. 1994). PCR was performed on a Mastercycler Pro (Eppendorf, Germany) under the following conditions: an initial denaturation step at 94°C for 5 min, followed by 40 cycles of denaturation at 94°C for 2 min, primer annealing at 54°C for 30 s, extension at 72°C for 50 s

Table 1: *Alpinia* samples used in this study

No.	Sample ID	Morphology identification	Collected location	Collection time
1	PD1	*A.* aff. *Calcarata*	Phong Dien, Thua Thien-Hue	12 October 2016
2	SH83	*A.* aff. *Coriandriodora*	Na Ri, Bac Kan	07 April 2016
3	SH84	*A.* aff. *Coriandriodora*	Na Ri, Bac Kan	11 April 2016
4	2179	*A.* aff. *Coriandriodora*	Trang Dinh, Lang Son	07 April 2015
5	2197	*A.* aff. *Coriandriodora*	Kim Hy, Bac Kan	13 March 2013
6	SH87*	*A. blepharocalyx*	Tam Dao National Park, Tam Dao, Vinh Phuc	14 April 2016
7	SH650	*A. blepharocalyx*	Forest Inventory and Planning Institute, Thanh Tri, Ha Noi	20 April 2018
8	1093*	*A. bleviligulata*	Bach Ma National Park, Phu Loc, Thua Thien-Hue	29 August 2010
9	SH89	*A. calcicola*	Tam Dao National Park, Tam Dao, Vinh Phuc	15 April 2016
10	SH91	*A. conchigera*	Bach Ma National Park, Phu Loc, Thua Thien-Hue	15 May 2016
11	2190*	*A. galanga*	Binh Gia, Lang Son	03 May 2016
12	SH669	*A. galanga*	Mai Chau, Hoa Binh	27 June 2018
13	SH06*	*A. globosa*	Tam Dao National Park, Tam Dao, Vinh Phuc	08 October 2014
14	SH94*	*A. gramineum*	Son Dong, Bac Giang	29 May 2016
15	2189	*A. kwangsiensis*	Loc Binh, Lang Son	03 May 2016
16	SH649	*A. latilabris*	Forest Inventory and Planning Institute, Thanh Tri, Hanoi	20 April 2018
17	SH90	*A. maclurei*	Xuan Son National Park, Tan Son, Phu Tho	07 May 2016
18	SH93	*A. maclurei*	Xuan Son National Park, Tan Son, Phu Tho	22 May 2016
19	SH163	*A. menghaiensis*	Tam Dao, Vinh Phuc	12 April 2017
20	2186*	*A. oblongifolia*	Tam Dao National Park, Tam Dao, Vinh Phuc	09 October 2014
21	2182*	*A. oxymitra*	Phu Quoc, Kien Giang	19 May 2015
22	SH185	*A. oxymitra*	Phu Quoc, Kien Giang	19 May 2016
23	SH661	*A. oxymitra*	Phu Quoc, Kien Giang	05 June 2018
24	SH156	*A. pinnanensis*	Tam Dao National Park, Tam Dao, Vinh Phuc	05 November 2016
25	SH85*	*A. polyantha*	Son Dong, Bac Giang	28 April 2016
26	SH88	*A. pumila*	Tam Dao National Park, Tam Dao, Vinh Phuc	15 April 2016
27	SH125	*A. purpurata*	Krong Bong, Dak Lak	08 July 2016
28	2188	*A. strobiliformis*	Loc Binh, Lang Son	03 May 2016
29	2194*	*A. zerumbet*	Tan Son, Phu Tho	18 May 2016
30	SH101	*A. "kontumensis"*	Dak Glei, Kon Tum	01 July 2016
31	SH176	*A. "kontumensis"*	Dak Glei, Kon Tum	19 July 2017
32	SH86	*A. "tamdaoensis"*	Tam Dao National Park, Tam Dao, Vinh Phuc	14 April 2016
33	2183*	*A. "tamdaoensis"*	Tam Dao, Vinh Phuc	10 January 2015
34	SH167*	*A.* spp. 1	Tam Dao, Vinh Phuc	08 July 2017
35	SH97*	*A.* spp. 2	Dak Glei, Kon Tum	01 July 2016
36	SH155*	*A.* spp. 3	Tam Dao National Park, Tam Dao, Vinh Phuc	04 November 2016
37	2180*	*A.* spp. 5	Trang Dinh, Lang Son	22 April 2015
38	SH651	*A.* spp. 6	Forest Inventory and Planning Institute, Thanh Tri, Hanoi	20 April 2018
39	SH652	*A.* spp. 7	Forest Inventory and Planning Institute, Thanh Tri, Hanoi	20 April 2018
40	SH653	*A.* spp. 8	Bidoup Nui Ba National Park, Lac Duong, Lam Dong	23 April 2018
41	SH479	*A.* spp. 9	Cu Jut, Dak Nong	15 October 2017
42	SH486	*A.* spp. 10	Cu Jut, Dak Nong	15 October 2017
43	SH532	*A.* spp. 11	Dak Song, Dak Nong	18 October 2017
44	SH538	*A.* spp. 12	Dak Song, Dak Nong	18 October 2017

*Samples failed in amplification and sequencing were marked in dark and light grey, respectively

and final extension at 72°C for 10 min. For each reaction, the PCR mixture consisted of 2.0 μL 10X DreamTaq buffer, 1.0 μL each 10 μM primer, 0.5 μL 2.5 mM dNTPs, 0.15 μL of 5 U/μL DreamTaq DNA polymerase, 18.85 μL milliQ, and 1.0 μL template DNA for a total volume of 20 μL. PCR products were detected by 0.8% agarose gel electrophoresis and purified using GeneJET PCR Purification kit (Thermo Fisher Scientific, USA).

Sequencing and alignment of ITS region: Sanger sequencing of ITS region was performed on an ABI 3500 Genetic Analyzer system using BigDye Terminator v. 3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, USA). Raw sequencing results were compared and aligned using the program BioEdit 7.0.5. Sequences obtained in this study were submitted to GenBank with accession number from MN545627-MN545656. BLAST (Basic Local Alignment Search Tools) searches for evaluating the species identification ability were performed using reference sequences on GenBank.

Phylogenetic analysis and species classification: The matrix for phylogenetic analysis consisted of ITS sequences obtained in this study and reference sequences, and the global alignment was performed using MAFFT version 7.407 (Katoh *et al*. 2002) with local re-alignment using MUSCLE version 3.8.1551 (Edgar 2004). Phylogenetic tree of the aligned ITS sequence sets was separately reconstructed by Neighbor Joining (NJ) and Maximum Likelihood (ML) methods with Kimura 2-parameter model of 1000 replicates using MEGA.X (Kumar *et al*. 2018) with 1000 replicates. Phylogenetic variation was estimated with bootstrap values (%), which indicated confidence interval between phylogenetic lineages of the studied samples on the

Fig. 1: Different types of *Alpinia* genus collected in Vietnam
(**a**) *A.* aff. *calcarata*; (**b**) *A.* aff. *coriandriodora*; (**c**) *A. blepharocalyx*; (**d**) *A. bleviligulata*; (**e**) *A. calcicola*; (**f**) *A. conchigera*; (**g**) *A. galanga*; (**h**) *A. gramineum*; (**i**) *A. kwangsiensis*; (**j**) *A. latilabris*; (**k**) *A. maclurei*; (**l**) *A. menghaiensis*; (**m**) *A. oblongifolia*; (**n**) *A. oxymitra*; (**o**) *A. pinnanensis*; (**p**) *A. polyantha*; (**q**) *A. pumila*; (**r**) *A. purpurata*; (**s**) *A. strobiliformis*; (**t**) *A. zerumbet*; (**u**) *A. "kontumensis"*; (**v**) *A. "tamdaoensis"*; (**w**) *Alpinia* spp. 2; (**x**) *Alpinia* spp. 7

Table 2: Success rate of PCR amplification and sequencing of ITS region in sample set

	Number of success samples	Success rate (%)
Genomic DNA extraction	44	100
PCR amplification	37	84
Sequencing	30	81

tree. Information of the ITS fragments of studied samples, including accession numbers of referred taxa as showed in Fig. 3a, b. Outgroup selection for phylogenetic analysis was ITS sequence from *A. longipetiolatum*. BLAST searches results were used as an initial classification to localize each sample into sister species groups. Comparison between NJ and ML phylogenetic trees was performed based on nodes with bootstrap value greater than 50. Ambiguous branches and nodes were excluded from the analysis. Results from both phylogenetic tree construction methods and BLAST searches were then compared to morphology based classification. Results from the comparison were used to evaluate the discriminating ability of ITS regions in certain groups of *Alpinia* species.

Results

Total DNA extraction and amplification of ITS region

Genomic DNAs were isolated from 44 leaf samples of species *Alpinia* and had sufficient quality for further uses. After the extraction step, genomic DNA was used as template for PCR amplification of the ITS region. The length of the amplicons obtained with universal primers for ITS amplification was approximately 850 bp, as expected (Fig. 2). The success rate of PCR amplification was 84% due to failures in the amplification of 7 DNA samples (Table 2).

Sequencing and alignment of ITS region

Total 37 PCR products were purified and sequenced using Sanger-based sequencing system. Among those samples, 30 sequences were obtained, which contributed to 81% of the success rate for sequencing the ITS region. Most of the samples that were failed to amplify and sequence were collected during the period from 2010 to 2016. The above proportions showed difficulties in amplification and sequencing of ITS region for prolonged storage samples despite optimizing effort.

Raw sequences obtained from the sequencing step were proceeded to a rough editing process. Ambiguous

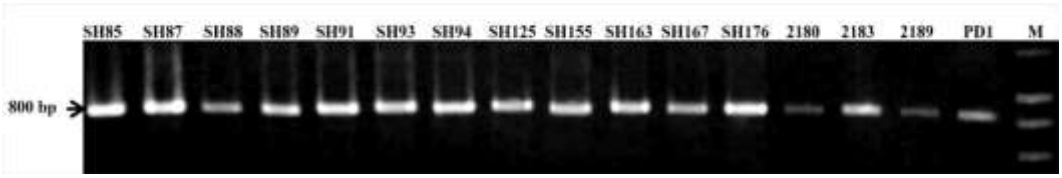

Fig. 2: Electrophoresis of PCR products of 800 bp amplified between ITS-F and ITS-R primers and gDNA of representative samples from 44 samples of *Alpinia* species. SH85-PD1: ID of the samples with detailed in Table 1*;* M: Hyper ladder 1kb (Bioline, UK)

Table 3: Species identification of *Alpinia* species using ITS region

No.	Sample ID	GenBank accession number	Morphological classification	Molecular based classification
1	PD1	MN545627	*A.* aff. *calcarata*	*A. calcarata*/ *A. galanga*
2	SH83	MN545628	*A.* aff. *coriandriodora*	*A. coriandriodora*
3	SH84	MN545629	*A.* aff. *coriandriodora*	*A. coriandriodora*
4	2179	MN545630	*A.* aff. *coriandriodora*	*A. tonkinensis*
5	2197	MN545631	*A.* aff. *coriandriodora*	*A. coriandriodora*
6	SH650	MN545632	*A. blepharocalyx*	Generated a separated branch
7	SH89	MN545633	*A. calcicola*	*A. tonkinensis*
8	SH91	MN545634	*A. conchigera*	*A. calcarata*/ *A. galanga*
9	SH669	MN545635	*A. galanga*	*A. galanga*
10	2189	MN545636	*A. kwangsiensis*	*A. kwangsiensis*
11	SH649	MN545637	*A. latilabris*	Generated a separated branch
12	SH90	MN545638	*A. maclurei*	*A. maclurei*
13	SH93	MN545639	*A. maclurei*	*A. maclurei*
14	SH163	MN545640	*A. menghaiensis*	*A. kwangsiensis*
15	SH185	MN545641	*A. oxymitra*	*A. oxymitra*
16	SH661	MN545642	*A. oxymitra*	*A. oxymitra*
17	SH156	MN545643	*A. pinnanensis*	*A. pinnanesis*
18	SH88	MN545644	*A. pumila*	*A. pumila*
19	SH125	MN545645	*A. purpurata*	*A. purpurata*
20	2188	MN545646	*A. strobiliformis*	*A. strobiliformis* var. *glabra*
21	SH101	MN545647	*A. "kontumensis"*	Generated a separated branch
22	SH176	MN545648	*A. "kontumensis"*	*A. nutans*
23	SH86	MN545649	*A. "tamdaoensis"*	*A. chinensis*/ *A. japonica*/ *A. pumila*
24	SH651	MN545650	*A.* spp. 6	Generated a separated branch
25	SH652	MN545651	*A.* spp. 7	Generated a separated branch
26	SH653	MN545652	*A.* spp. 8	*A. nutans*
27	SH479	MN545653	*A.* spp. 9	Generated a separated branch
28	SH486	MN545654	*A.* spp. 10	Generated a separated branch
29	SH532	MN545655	*A.* spp. 11	*A. conchigera*
30	SH538	MN545656	*A.* spp. 12	Generated a separated branch

nucleotides and background noises in obtained sequences were removed to enhance the accuracy of the analysis. Afterwards, sequences from 30 samples were searched and compared to reference sequences in GenBank using web-based BLAST server. Results in identity reference were used to evaluate the species identification ability and to find the relationship of species within the genus *Alpinia*. Sequence alignment was performed using both global and local approaches to reduce overall error rate caused by a wide range of sequence variations. A total of 36 reference sequences of species in the genus *Alpinia* from GenBank, along with 30 sequences in this study, were included in the alignment (Suppl. material 1). The alignment matrix had a total length of 593 bp, covering partial sequence of ITS1, 5.8S, and ITS2 regions.

Phylogenetic analysis

Based on the nucleotide matrix, phylogenetic trees were constructed using both Neighbor Joining (Fig. 3a) and Maximum Likelihood methods (Fig. 3b) with 1000 replications. *Amomum longipetiolatum*, a species of closely related genus of *Alpinia* was used as an outgroup sequence. Bootstrap values were estimated in both methods. Only bootstrap values greater than 50 were displayed in the phylogenetic tree for better observation and comprehension (Fig. 3). Therefore, only branches with reliable support were useful for species discrimination process. Table 3 summarized the species classification of 30 samples from the genus *Alpinia* in Vietnam. In general, there were 14 samples including PD1, SH83, SH84, 2197, SH669, 2189, SH90, SH93, SH185, SH661, SH156, SH88, SH125, and 2188 belonged to 10 species had identities between morphological and phylogenetic specification. Four out of 30 samples were classified as different species from morphological discrimination including samples 2179, SH89, SH91, and SH163. PD1 was the only sample that showed incongruence between the two phylogenetic trees. Remaining 12 samples were either generated separated branches or considered belonged to distinct taxa that have sequences currently not

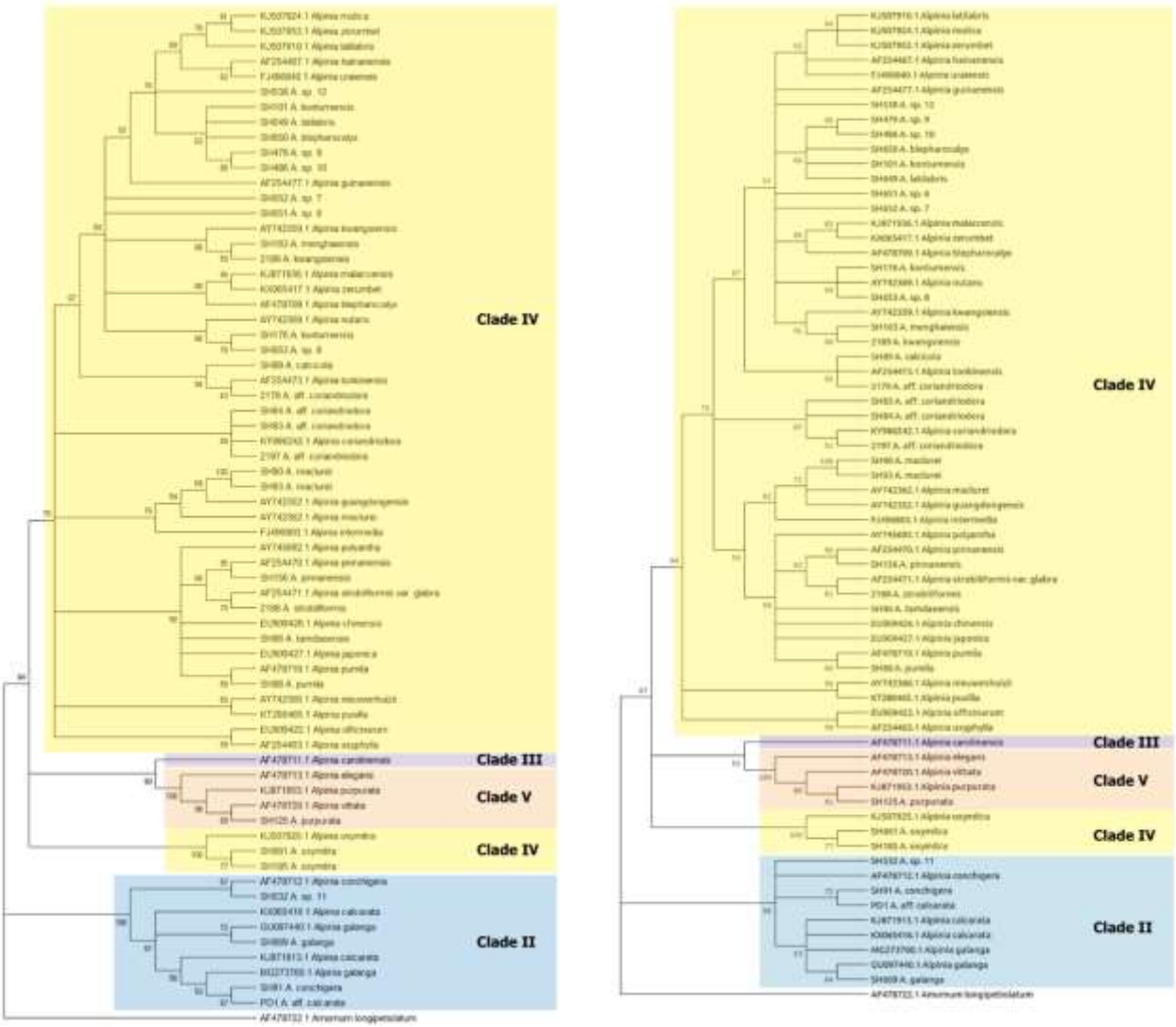

Fig. 3: Phylogenetic trees of *Alpinia* species constructed using Neighbor Joining (**a**) and Maximum Likelihood (**b**) methods

available. Besides, there were incongruences in molecular based species identification between samples in the same species such as *A.* aff. *coriandriodora* and *A.* "*kontumensis*". These conflicts were results of distinct geographical characteristics of collected locations, differences in collection time, and lack of reference sequences for *Alpinia* species in Vietnam on GenBank.

Discussion

The low amplification and sequencing success rates of the ITS region were observed in several previous studies due to divergent paralogous copies within individuals and fungal contamination in a certain group of plants (Hollingsworth *et al*. 2011; Vinitha *et al*. 2014). In this study, the success rates of more than 80% for PCR amplification and sequencing of interested samples are similar to the previous study of China Plant BOL Group *et al.* (2011).

The alignment matrix of 66 ITS sequences consisted of a multitude of differences in nucleotide sequences among both samples used in this study and reference sequences. The conserved regions observed in the matrix were from the 5.8S, while most of the variations were distributed in ITS1 and ITS2. These regions had a potential in species classification due to their high resolutions of inter- and intraspecific relationship (Cheng *et al*. 2016). However, in the genus *Alpinia*, ambiguous nucleotides in ITS1 and ITS2 of GenBank reference sequences generated difficulties in alignment and species identification. Therefore, a combination of sequence alignment and BLAST searches were necessary to enhance the accuracy of species identification.

The phylogenetic tree of the ITS region showed the relationship between Vietnamese *Alpinia* species used in

this study and *Alpinia* species available in GenBank. According to the phylogenetic analysis, SH91 (*A. conchigera*), PD1 (*A.* aff. *calcarata*), and SH532 (*A.* spp. 11) were sisters to the group of *A. calcarata* and *A. galanga*. This relationship between *A. conchigera*, *A. calcarata* and *A. galanga* was supported by fruit wall anatomy study of Liao & Wu (Liao and Wu 1996) and molecular based classification of Kress *et al.* (Kress *et al*. 2005). This species group belongs to the subsection *Alpinia*, *Catimbium* (section *Alpinia*), and *Strobidia* (section *Allughas*) according to Smith (Smith 1990) and Clade II (*Galanga* clade) in Kress's classification system (Kress *et al*. 2005). The only sample belongs to Clade V (*Eubractea* clade) in this study was SH125 (*A. purpurata*, section *Guillainia*) along with *A. elegans* (section Kolowratia) and *A. vittata* (section *Dieramalpinia*). Other species used in the present study belong to Clade IV (*Zerumbet* clade). Among these samples, SH101 (*A. "kontumensis"*), SH649 (*A. latilabris*, subsection *Catimbium*, section *Alpinia*), SH650 (*A. blepharocalyx*, subsection *Catimbium*, section *Alpinia*), SH651 (*A.* spp. 6), SH652 (*A.* spp. 7), SH479 (*A.* spp. 9), SH486 (*A.* spp. 10), and SH 538 (*A.* spp. 12) generated a separated branch, indicating that these samples were distinct from all the *Alpinia* species sequences in GenBank. These species also formed a distinct group in Kress's study (Kress *et al*. 2005). SH176 (*A. "kontumensis"*) and SH653 (*A.* spp. 8), which were not identified by morphological characteristics and were closely related to *A. nutans* from section *Dieramalpinia* (bootstrap value equal 82). The sample SH163 (*A. menghaiensis*, subsection *Catimbium*, section *Alpinia*) and SH479 were closely related to *A. kwangsiensis* (subsection *Catimbium*, section *Alpinia*). The sample 2179 (*A.* aff. *coriandriodora*, subsection *Alpinia*, section *Alpinia*) and SH89 (*A. calcicola*, subsection *Catimbium*, section *Alpinia*) were placed in the same branch with *A. tonkinensis* (subsection *Alpinia*, section *Alpinia*) with strong support (bootstrap values equal 95 and 94 in NJ and ML trees, respectively). Another sample SH86 (*A. "tamdaoensis"*), was closely related to *A. chinensis* (subsection *Alpinia*, section *Alpinia*), *A. japonica* (subsection *Alpinia*, section *Alpinia*), and *A. pumila* (section *Didymanthus*).

Species identification results of 14 samples were similar to those concluded by morphological classification. However, in several complex groups of *Alpinia* genus, there were conflicts between species classification based on morphology and molecular marker. SH91 (*A. conchigera*) was not in the same grouped with *A. conchigera* species (AF478712.1). References sequence for other ambiguous sample, SH163 (*A. menghaiensis*) was currently unavailable. Therefore, except for the sample 2179 which showed clear difference between morphological and molecular based identification, other conflict samples had insufficient amount of reference ITS sequences, which might lead to unreliable discriminating results. Previous studies have indicated the effectiveness of ITS region in resolving phylogenetic relationships at different taxonomic levels (Vinitha *et al*. 2014; Boer *et al*. 2018). The conflicts may occur due to the lack of reference sequences, and high variation of ITS sequence.

The main results in this present study were supported by the phylogenetic research and molecular based classification of Kress *et al.* (2005). According to Kress's classification system, *Alpinia* species in Vietnam belong mainly to Clade IV. The BLAST searches and phylogenetic analysis showed the high species identification ability of the region ITS as molecular marker.

Conclusion

This study clearly indicated that DNA barcoding using ITS region is a reliable method for supporting species classification in the genus *Alpinia*. ITS can be applied to rapid identification of these medicinal and ornamental plants, along with their products.

Acknowledgements

This work was supported by Vietnam Academy of Science and Technology under the program: "Support scientific research activities for senior researchers" (grant no. NCVCC40.02/22-23).

Author Contributions

Le Thi Thu Hien and Nguyen Quoc Binh initiated this study. Nguyen Quoc Binh and Nguyen Phuong Hanh collected and identified plant materials. Nguyen Nhat Linh, Le Thi Thu Ha, Pham Le Bich Hang, and Luu Han Ly performed the experiments. Nguyen Nhat Linh, Le Quang Trung, Nguyen Hai Ha, and Le Thi Thu Hien performed data analysis and drafted the manuscript. All authors have read, commented and approved the final manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability

DNA barcoding sequences used in this study were available on GenBank (https://www.ncbi.nlm.nih.gov/nucleotide) and their accession numbers were provided in Suppl. material 1.

Ethics Approval

Ethical approval is not applicable in this study.

References

Álvarez I, JF Wendel (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol* 29:417–434

Boer HD, M Newman, AD Poulsen, AJ Droop, T Fer, LTT Hen, K Hlavata, V Lamxay, JE Richardson, K Steffen, J Leong-Skornickova (2018). Convergent morphology in Alpinieae (Zingiberaceae): Recircumscribing *Amomum* as a monophyletic genus. *Taxon* 67:6–36

CBOL Plant Working Group, PM Hollingsworth, LL Forrest, JL Spouge, M Hajibabaei, S Ratnasingham, MVD Bank, MW Chase, RS Cowan, DL Erickson, AJ Fazekas (2009). A DNA barcode for land plants. *Proc Natl Acad Sci USA* 106:12794–12797

Cheng T, C Xu, L Lei, C Li, Y Zhang, S Zhou (2016). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Mol Ecol Res* 16:138–149

China Plant BOL Group, DZ Li, LM Gao, HT Li, H Wang, XJ Ge, JQ Liu, ZD Chen, SL Zhou, SL Chen, JB Yang, CX Fu, CX Zeng, HF Yan, YJ Zhu, YS Sun, SY Chen, K Wang, T Yang, GW Duan (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc Natl Acad Sci USA* 108:19641–19646

Edgar RC (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res* 32:1792–1797

Hartvig I, M Czako, ED Kjær, LR Nielsen, I Theilade (2015). The use of DNA barcoding in identification and conservation of rosewood (*Dalbergia* spp.). *PLoS One* 10:e0138231

Hebert PD, S Ratnasingham, JRD Waard (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Roy Soc Lond Ser B Biol Sci* 270:96–99

Hollingsworth PM, SW Graham, DP Little (2011). Choosing and using a plant DNA barcode. *PLoS One* 6:e19254

Huang XC, XQ Ci, JG Conran, J Li (2015). Application of DNA barcodes in Asian tropical trees–a case study from Xishuangbanna Nature Reserve, Southwest China. *PLoS One* 10:e0129295

Kang Y, Z Deng, R Zang, W Long (2017). DNA barcoding analysis and phylogenetic relationships of tree species in tropical cloud forests. *Sci Rep* 7:12564

Katoh K, K Misawa, K Kuma, T Miyata (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucl Acids Res* 30:3059–3066

Keskin F, İ Kaya, M Usta, İ Demir, HM Sipahioglu, Y Nemli (2017). Molecular cloning and sequence analysis of the its region of nuclear ribosomal DNA for species identification in dodders (*Cuscuta*; Convolvulaceae). *Intl J Agric Biol* 19:1447–1451

Kress WJ, DL Erickson (2007). A two-locus global DNA barcode for land plants: The coding *rbc*L gene complements the non-coding *trn*H-*psb*A spacer region. *PLoS One* 2:e508

Kress WJ, AZ Liu, M Newman, QJ Li (2005). The molecular phylogeny of *Alpinia* (Zingiberaceae): A complex and polyphyletic genus of gingers. *Amer J Bot* 92:167–178

Kumar S, G Stecher, M Li, C Knyaz, K Tamura (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549

Larsen K (1998). Zingiberaceae. *In*: *The Families and Genera of Vascular Plants*, Vol. 4, pp:474–495. Kubitzki K (Ed.). Springer, Dordrecht, The Netherlands

Le HT, ND Do, QB Nguyen, TT Nguyen (2017). *Alpinia rugosa* S.J.Chen & Z.Y.Chen – a new record for flora of Vietnam. *VNU J Sci Nat Sci Technol* 33:101–104

Liao J, Q Wu (1996). Fruit anatomy of Chinese *Alpinia* and its taxonomic significance. *In: Proceedings of the 2nd Symposium on the Family Zingiberaceae*, pp:82–90. Wu T-L (Ed.). Zhongshan University Press, Guangzhou, China

Ngamriabsakul C, M Newman, Q Cronk (2003). The phylogeny of tribe Zingibereae (Zingiberaceae) based on ITS (nrDNA) and *trn*l–f (cpDNA) sequences. *Edinb J Bot* 60:483–507

Nguyen PH, QB Nguyen, BS Adhikari (2014). Distribution of *Alpinia* (Zingiberaceae) and their use pattern in Vietnam. *J Biodivers Endanger Spec* 2:121–125

Nguyen QB, PH Nguyen, LTA Hoang (2017). The genus *Alpinia* Roxb. (Zingiberaceae) in Vietnam. *In: Proceedings of the 7th National Science Conference on Ecology and Biological Resources*, pp:50–53. Vietnam

Nghiem DT, QB Nguyen, VO Tran, PH Nguyen, TT Nguyen (2018). A new species of the genus *Alpinia* (*Alpinia graminifolia* D. Fang & G.Y. Lo) recorded for Vietnam's flora. *VNU J Sci Nat Sci Technol* 34:54–58

Panaligan AC, MD Baltazar, GJD Alejandro (2021). Molecular authentication of commercially cultivated coffee (*Coffea* spp.) in the Philippines using DNA barcodes. *Intl J Agric Biol* 25:227–230

Qiao CF, QB Han, ZL Zhao, ZT Wang, LS Xu, HX Xu (2009). Sequence analysis based on ITS1 region of nuclear ribosomal DNA of *Amomum villosum* and ten species of *Alpinia*. *J Food Drug Anal* 17:142–145

Saha K, B Dholakia, R Sinha, S Sinha (2020). DNA barcoding of selected Zingiberaceae species from North-East India. *J Plant Biochem Biotechnol* 29:494–502

Smith R (1990). *Alpinia* (Zingiberaceae): A proposed new infrageneric classification. *Edinb J Bot* 47:1–75

Sun Y, DZ Skinner, GH Liang, SH Hulbert (1994). Phylogenetic analysis of Sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theor Appl Genet* 89:26–32

Tan WH, LC Chai, CF Chin (2020). Efficacy of DNA barcode internal transcribed spacer 2 (ITS 2) in phylogenetic study of *Alpinia* species from Peninsular Malaysia. *Physiol Mol Biol Plants* 26:1889–1896

Uma E, T Muthukumar (2014). Comparative root morphological anatomy of Zingiberaceae. *Syst Biodivers* 12:195–209

Vinitha MR, US Kumar, K Aishwarya, M Sabu, G Thomas (2014). Prospects for discriminating Zingiberaceae species in India using DNA barcodes. *J Integr Plant Biol* 56:760–773

Vu XD, QB Nguyen, TL Dang, DT Nghiem, PB Cao, TC Vu, X Ye, N Xia (2019). *Alpinia coriandriodora* D. Fang, a new record for flora of Vietnam. *J Trop Subtrop Bot* 27:99–101

Williams KJ, WJ Kress, PS Manos (2004). The phylogeny, evolution, and classification of the genus *Globba* and tribe Globbeae (Zingiberaceae): Appendages do matter. *Amer J Bot* 91:100–114

Wu TL (2000). Family Zingiberaceae. *In*: *Flora of China*, Vol. 24, pp:322–377. Wu ZG, PH Raven (Eds.). Science Press, Beijing, China