https://doi.org/10.17221/88/2020-CJGPB

# Phylogenetics of native conifer species in Vietnam based on two chloroplast gene regions *rbcL* and *matK*

*Mai Phuong Pham[1], Viet Ha Tran[2], Dinh Duy Vu[1,3]*, Quoc Khanh Nguyen[1], Syed Noor Muhammad Shah[4]**

[1]*Vietnam-Russia Tropical Centre, Hanoi, Vietnam*
[2]*College of Forestry, Vietnam National University of Forestry, Hanoi, Vietnam*
[3]*Graduate University of Science and Technology (GUST), Vietnam Academy of Science and Technology (VAST), Hanoi, Vietnam*
[4]*Department of Horticulture, Faculty of Agriculture, Gomal University, Dera Ismail Khan, Pakistan*

**Corresponding authors: duydinhvu87@gmail.com; noor.aup@gmail.com*

*Mai Phuong Pham and Viet Ha Tran contributed equally to this work and should be regarded as co-first authors.*

**Citation:** Pham M.P., Tran V.H., Vu D.D., Nguyen Q.K., Shah S.N.M. (2021): Phylogenetics of native conifer species in Vietnam based on two chloroplast gene regions *rbcL* and *matK*. Czech J. Genet. Plant Breed., 57: 58–66.

**Abstract:** We used two chloroplast gene regions (*matK* and *rbcL*) as a tool for the identification of 33 local conifer species. All 136 sequences, 101 newly generated (14 species for gene *matK*; 16 species for gene *rbcL*) and 35 retrieved from the GenBank, were used in the analysis. The highest genetic distance (*matK* region) was recorded between the species in Cupressaceae with an average of 5% (0.1–8.5), Podocarpaceae with an average of 6% (0–8.5), Taxaceae with an average of 5% (0.2–0.5) and Pinaceae with an average of 20.4% (0.8–54.1). The *rbcL* region showed a low genetic distance between the species in Cupressaceae 2% (0–3.3), Podocarpaceae 3% (0.6–3.4), Taxaceae 1% (0–2.1) and Pinaceae 1.2% (0–5.82). The phylogenetic analyses using the Maximum likelihood (ML) and Bayesian inference (BI) bootstrap values obtained at the branching nodes of each species ranged from 62 to 100% (Maximum likelihood bootstrap – MLBS and Bayesian posterior probabilities – BPP) for the *matK* gene; from 66 to 100% (MLBS) and 60 to 100% (BPP) for the *rbcL* region. The *rbcL* region was not identified between the species of Taxaceae and Cephalotaxaceae. The *matK* gene region was very clear in the different species among the families (Cupressaceae, Podocarpaceae, and Cephalotaxaceae) and unsuitable for identifying closely related species in *Amentotaxus* (Taxaceae) and *Pinus* (Pinaceae). The gene (*matK*) is a useful tool as a barcode in the identification of conifer species of Cupressaceae, Podocarpaceae, and Cephalotaxaceae in Vietnam.

**Keywords:** DNA barcoding; endangered; phylogeny; taxonomy

Gymnosperms are the dominant species of temperate regions. Plant fossil records show its existence since the late Palaeozoic era. The current gymnosperm distribution is affected by the past continental drift and climatic change effect (Contreras-Medina & Vega 2002). According to Forest et al. (2018) 1 090 species, 18 genera, and 14 families have been discovered. Conifers are ancient gymnosperm species and were reported about 300 million years ago (Asaf et al. 2018). Conifers have naked seeds, while the cones are dioecious or monoecious and usually cross-pollinate. Conifers produce woody cones. Co-

Supported by Vietnam-Russia Tropical Center basis Project, 2020–2022: Application of geographic information method (GIS) and molecular biology in investigation, monitoring and development *Cunninghamia konishii* Hayata species of Vietnam.

https://doi.org/10.17221/88/2020-CJGPB

nifer species are found above 800 m altitudes in the Northwest, Northeast, in the Truong Son, and Tay Nguyen of Vietnam.

Based on the morphological characteristics using a traditional method, 33 species, 2 subspecies and 5 varieties of conifer belonging to 19 genera and 5 families have been identified in Vietnam up till now. The largest family is Pinaceae (5 genera, 12 species) followed by Cupressaceae (7 genera, 8 species), Podocarpaceae (4 genera, 6 species), Taxaceae (2 genera, 6 species), and Cephalotaxaceae (1 genus and 1 species) (Hiep et al. 2005, 2015; Loc et al. 2017). Among them, 30 species are listed in the endangered species list (MOST & VAST 2007; Loc et al. 2017). The identification of conifer species is not an easy task in Vietnam. The traditional method for plant identification based on the morphological parameters, such as the plant height, leaf shape, and size, the characteristics of the flowers and fruits, is not an authenticated method, because most of the conventional method parameters vary with the plant's age and habitats. Advance methods are in a dire need of time to identify a species.

DNA barcode using a standardised gene region is considered a productive tool of taxon identification (Kress et al. 2015; Wu et al. 2019). DNA barcoding is a diagnosing tool using short DNA sequences of the chloroplast genome in plants, developed by Hebert et al. (2003). DNA barcoding use in plants is controversial for some researchers (Chase et al. 2005; Kress et al. 2005; Cowan et al. 2006; Pennisi 2007). However, many studies have proved that DNA barcoding is an effective tool for the identification of plants (Newmaster et al. 2006; Kress & Erickson 2007; Lahaye et al. 2008; Fazekas et al. 2012). The *rbcL* and *matK* genes are widely used as "DNA barcode" regions for plants (Kress et al. 2005; Chase et al. 2007; Bafeel et al. 2011; de Vere et al. 2012; Little et al. 2013). First, Kress et al. (2005) proposed the *trnH-psbA* plastid as a barcode and then later, in Kress and Erickson (2007), proposed the *TrnH-psbA* and *rbcL* genes as a suitable, universal barcode for plant species. Lahaye et al. (2008) compared eight gene (*trnH-psbA*, *matK*, *ycf5*, *rbcL*, *rpoB*, *ndhJ*, *accD*, and *rpoC1*) regions and identified the plastid *matK* gene as a universal DNA barcode for flowering plants. The *matK* gene can explore cryptic species and prove useful in identifying species. However, the low variation between species is a hurdle in DNA barcode applications in plants. Hollingsworth et al. (2009) recommended the 2-locus combination of ribulose-1,5-bisphosphate carboxylase (*rbcL*) and maturase (*matK*) as a plant barcode based on the assessments of the recoverability, sequence quality, and levels of species discrimination for the seven plastid DNA regions (*atpF-atpH*, *matK*, *rbcL*, *rpoB*, *rpoC1*, *psbK-psbL* and *trnH-psbA*). This 2-locus barcode can identify species and will contribute to the discovery of new species (Burgess et al. 2011; de Vere et al. 2012; Kress et al. 2015). Different researchers have studied the identity, genetic variation and verified the taxonomic status of conifer species at the molecular level (de Vere et al. 2012; Zheng et al. 2016; Asaf et al. 2018; Phong et al. 2018; Kang et al. 2019;). However, fewer researchers have studied conifers in Vietnam. Therefore, the current study was designed to understand the phylogenetic relationships among the native conifer species in Vietnam, to find out the most suitable DNA barcode region for the identification of the conifer species in Vietnam and to develop a background for the conservation, evolution, and systematics of the species.

## MATERIAL AND METHODS

**Taxon sampling.** The conifer species location and GenBank accession numbers of previously published DNA sequences are shown in Table 1. The 101 samples (young leaves) of twenty-one species (Table 1) were collected in plastic bags having silica gel and immediately transferred to the laboratory, stored at –30 °C until the DNA extraction took place. The samples were also collected from the Local Herbarium of the targeted locations and confirmed by Dr. Nguyen Tien Hiep, a Botanist at the Institute of Ecological Resources, Vietnam Academy of Science and Technology.

**DNA isolation.** The genomic DNA was extracted by the modified cetyl trimethylammonium bromide (CTAB) method according to Doyle and Doyle (1990). The DNA was quantified through fluorometry and then diluted to a 10ng/μL concentration.

**PCR amplification of the two chloroplast genes region.** Two chloroplast genes (*rbcL* and *matK*) were amplified through the following polymerase chain reaction (PCR) cycling profile: an initial denaturing at 94 °C for 3 min; 40 cycles for 1 min at 94 °C; 30 s at a 55–56 °C annealing temperature for the primer pairs, a 1 min extension at 72 °C and 10 min at 72 °C for the final cycle before holding the samples at 4 °C. All the PCR reactions were performed in 25 μL volumes using Gene Amp PCR Systems 9700

https://doi.org/10.17221/88/2020-CJGPB

Table 1. List of the sampled species, locations and GenBank accessions

| Species | Collected location | No. of individuals | Chloroplast gene (GenBank accession No.) | |
|---|---|---|---|---|
| | | | *rbcL* and reference | *matK* and reference |
| **Cephalotaxaceae** | | | | |
| *Cephalotaxus mannii* Hook. f. | Bao Lac, Ha Giang, 1 895 m, 22°52'N–105°50'E | 05 | JF941236 | JX099351 |
| **Cupressaceae** | | | | |
| *Calocedrus macrolepis* Kurz | Da Lat, Lam Dong, 1 500 m, 11°56'N–108°25'E | 03 | JN039267 | JN657251 |
| *Calocedrus rupestris* Aver., H.T. Nguyen & L.K. Phan | Hang Kia, Mai Chau, Hoa Binh, 1 047 m, 20°44'N–104°55'E | 03 | JN039266 | JN657252 |
| *Cunninghamia konishii* Hayata | Bat Mot, Thanh Hoa, 970 m, 19°52'N–104°58'E | 05 | JN039274 | JN657259 |
| *Cupressus tonkinensis* Silba | Huu Lien, Lang Son, 440 m, 21°49'N–106°22'E | 03 | JN039269 | AY988355 |
| *Fokienia hodginsii* Dunn | Hang Kia, Mai Chau, Hoa Binh, 1 047 m, 20°44'N–104°55'E | 05 | JN039273 | JN657258 |
| *Glyptostrobus pensilis* (Staunt.) K. Koch | EaRal, Dak Lak, 570 m, 13°09'N–108°18'E | 05 | JN039281 | JN657263 |
| *Taiwania cryptomerioides* Hayata | GenBank | | L25756 | AB030127 |
| *Xanthocyparis vietnamensis* Farjon & Hiep | GenBank | | AY380895 | AY380850 |
| **Pinaceae** | | | | |
| *Abies delavayi* Franch. | GenBank | | JF940555 | MH230852 |
| *Keteleeria davidiana* (Bertrand) Beissn. | GenBank | | JN935652 | AB161020 |
| *Keteleeria evelyniana* Mast. | GenBank | | MH069638 | KT150255 |
| *Pinus dalatensis* Ferré | Da Lat, Lam Dong, 1 500 m, 11°56'N–108°25'E | 05 | JQ062979 | KT236090 |
| *Pinus kesiya* Royle | Ho Tien, Lam Dong, 1 390 m, 11°48'N–108°29'E | 05 | JN039276 | KT247644 |
| *Pinus krempfii* Le comte | Xuan Tho, Lam Dong, 1 450 m, 11°57'N–108°36'E | 03 | JN039275 | KT272170 |
| *Pinus kwangtungensis* Chun ex Tsiang | Hang Kia, Mai Chau, Hoa Binh, 1 047 m, 20°44'N–104°55'E | 05 | JN039280 | EF546713 |
| *Pinus latteri* Manson | GenBank | | JF943429 | JF955524 |
| *Pinus wangii* Hu & W.C. Cheng | GenBank | | NC039613 | KP128409 |
| *Pseudotsuga sinensis* Dode | GenBank | | MH069634 | AB601120 |
| *Tsuga chinensis* (Franch.) Pritz. ex Diels | GenBank | | AF145462 | LC095866 |
| *Tsuga dumosa* (D. Don) Eichler | GenBank | | AF145460 | MH116937 |
| **Podocarpaceae** | | | | |
| *Dacrycarpus imbricatus* (Blume) de Laub. | Hang Kia, Mai Chau, Hoa Binh, 1 047 m, 20°44'N–104°55'E | 05 | JN039279 | JN657262 |
| *Dacrydium elatum* (Roxb.) Wall. | Da Lat, Lam Dong, 1 500 m, 11°56'N–108°25'E | 05 | JQ062981 | JN657255 |
| *Nageia fleuryi* (Hickel) de Laub. | Bat Dai Son, Quan Ba, Ha Giang, 1 250 m, 23°08'N–104°56'E | 03 | JN039271 | JN657256 |

https://doi.org/10.17221/88/2020-CJGPB

Table 1 to be continued

| Species | Collected location | No. of individuals | Chloroplast gene (GenBank accession No.) | |
|---|---|---|---|---|
| | | | *rbcL* and reference | *matK* and reference |
| **Podocarpaceae** | | | | |
| *Nageia wallichiana* (C. Presl) O. Kuntze | Da Lat, Lam Dong, 1 500 m, 11°56'N–108°25'E | 05 | HM593616 | KR855701 |
| *Podocarpus neriifolius* D. Don | Da Lat, Lam Dong, 1 500 m, 11°56'N–108°25'E, | 05 | JN039282 | JN657264 |
| *Podocarpus pilgeri* Foxw. | Lac Duong, Lam Dong, 1 482 m, 12°11'02.7"N–108°41'24.3"E | 05 | JN039283 | KY021422 |
| **Taxaceae** | | | | |
| *Amentotaxus argotaenia* (Hance) Pilg. | Dak Glei, Kon Tum, 1 935 m, 15°04'23"N–107°57'31"E | 03 | JN039268 | JN657253 |
| *Amentotaxus hatuyenensis* N.T. Hiep | GenBank | | NS | KX059357 |
| *Amentotaxus poilanei* (Ferré & Rouane) D.K. Ferguson | GenBank | | JF940824 | KT072780 |
| *Amentotaxus yunnanensis* H.L. Li | Bao Lac, Cao Bang, 1 895 m, 22°52'N–105°50'E | 03 | JF940826 | JN657254 |
| *Taxus chinensis* Pilg. | Hoang Lien, Sa Pa, Lao Cai, 1 950 m, 22°12'N–103°05'E | 10 | JN039277 | JN657260 |
| *Taxus wallichiana* Zucc. | Nui Voi, Lam Dong, 1 474 m, 11°50'N–108°25'E | 10 | JN039278 | JN657261 |

that included: 2 µL of genomic DNA (total 20 ng), 12.5 µL of a master mix 2X, 1 µL of each 10 µM primer, and 8.5 µL of deionised $H_2O$. Two pairs of primers: *rbcL* F: 5'-CAC TGT TTG GAC CGA TGG ACT TAC-3' and *rbcL* R: 5'-CTT CGC GGA TCA CTT CAT TAC CTT C-3' and *matK* F: 5'-TTT GAC AGT TAT CTT GGA AGT TTC-3' and *matK* R: 5'-TAC TAA TTG GCT GCC CTG AGA TGT-3' were used to amplify the chloroplast genes with about 802–915 and 619–828 nucleotides for *rbcL* and *matK*, respectively. We designed the primers for *rbcL* and *matK* on the basis of the *rbcL* sequence of *Taxus brevifolia*, GenBank accession number AF249666 and *matK* of the *Taxus wallichiana* var. *chinensis*, GenBank accession HM590991.

**Sequencing of the two chloroplast genes region.** The sequencing was performed on an Avant 3100 Automated DNA sequencer using a Dye Terminator Cycle sequencing kit (PE Applied Biosystems, USA). The sequencing of the studied conifer species used the primers *rbcLF*/*rbcLR* and *matKF*/*matKR*.

**Phylogenetic analyses.** All of the 136 sequences, including 101 newly generated ones in this study and 35 retrieved from the GenBank, were used in the phylogenetic analyses. ChromasPro software (Ver. 2.1.6, Technelysium Pty Ltd., Tewantin, Australia) was used to edit and combine the sequences. The sequence alignments were made with BioEdit Ver. 7.0.9 (Hall 1999). The phylogenetic trees were produced using the Maximum likelihood (ML) and Bayesian inference (BI) approaches implemented in Treefinder (Jobb 2011) and MrBayes Ver. 3.2.1 (Ronquist & Huelsenbeck 2003), respectively. Optimum substitution models for each partition were selected by Kakusan Ver. 4.0 (Tanabe 2011), based on the Akaike information criterion (AIC). To obtain an estimate of the strength of the support for each node, a bootstrap method with 1 000 replicates with a heuristic search was performed as well. We used Mega Ver. 7.0 (Kumar et al. 2016) to analyse the p-distance between the conifer species.

## RESULTS

All of the studied 101 samples from 21 species were amplified and sequenced for the *matK* and *rcbL* regions. The sequences obtained for *matK* ranged from 619 to 828 bp in fourteen conifer species and

https://doi.org/10.17221/88/2020-CJGPB

Table 2. Summary ability of distinguishing the species for each gene region

| Gene regions | *matK* | *rcbL* |
|---|---|---|
| Number of species | 14 species in this study<br>19 species in GenBank | 16 species in this study<br>15 species in GenBank |
| Number of aligned nucleotide sites (bp) | 619–828 | 802–915 |
| Mean interspecific distance among species in Cupressaceae (%) | 5 (0.1–8.5) | 2 (0–3.3) |
| Mean interspecific distance among species in Podocarpaceae (%) | 6 (0–8.5) | 3 (0.6–3.4) |
| Mean interspecific distance among species in Taxaceae (%) | 5 (0.2–0.5) | 1 (0–2.1) |
| Mean interspecific distance among species in Pinaceae (%) | 20.4 (0.8–54.1) | 1.2 (0–5.82) |

for *rcbL*, ranging from 802 to 915 bp in sixteen conifer species in Vietnam (Table 1). All the sequences have been deposited in the GenBank under accession numbers JN039266-83, JQ062979-81 (*rcbL*) and JN657251-64, JX099351-58 (*matK*).

The number of base divergences per site and the recommended standard distance model showed high interspecific and low intraspecific divergences. The genetic distance of the *matK* region was higher than the genetic distance of the *rbcL* region (Table 2).

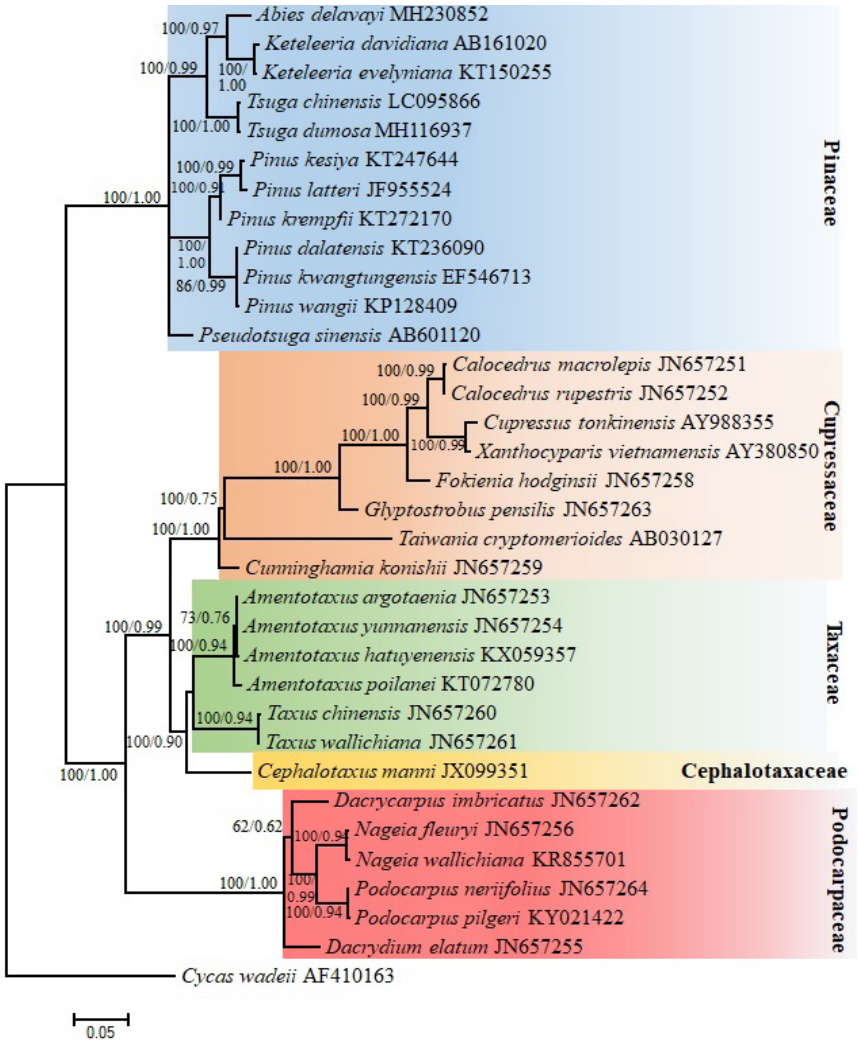


Figure 1. Phylogenetic relationships among the conifer species based on the partial sequence of the chloroplast genes (*matK*) using the Maximum likelihood (ML) tree; the numbers above and below the branches represent the bootstrap 1 000

https://doi.org/10.17221/88/2020-CJGPB

In the Cupressaceae family (8 studied species), the mean interspecific divergence was 5% ranging from 0.1% (between *Calocedrus macrolepis* and *C. rupestris*) to 8.5% (between *Cupressus tonkinensis* and *Taiwania cryptomerioides*) for *matK* (Table 2; Table S1a in the Electronic Supplementary Material (ESM)). The mean genetic distance (2%) ranged from 0.1% (between *C. macrolepis* and *C. rupestris*) to 3.3% (between *Cupressus tonkinensis* and *Taiwania cryptomerioides*) for *rbcL* (Table 2; Table S2a in the ESM)

In the Podocarpaceae family (6 species), the mean interspecific divergence was 6% ranging from 0.0% (between *Podocarpus pilgeri* and *P. neriifolius*) to 8.5% (between *Dacrydium elatum/Nageia fleuryi* and *Dacrydium elatum/Nageia wallichiana*) for *matK* (Table 2; Table S1b in the ESM). The mean genetic distance (3%) ranged from 0.6% (between *Podocarpus neriifolius* and *P. brevifolius*) to 3.4% (between *Dacrycarpus imbricatus/Nageia fleuryi* and *Dacrycarpus imbricatus/Podocarpus brevifolius*) for *rbcL* (Table 2; Table S2b in the ESM).

In the Taxaceae family (6 species), the mean interspecific divergence was 5% ranging from 0% between *Amentotaxus argotaenia* and *A. yunnanensis* to 9% between *Amentotaxus hatuyenensis* and *Taxus chinensis*. The interspecific divergence was 0.2% between *T. chinensis* and *T. wallichiana* for *matK* (Table 2; Table S1c in the ESM). The mean genetic distance (1%) ranged from 0% (between *Amentotaxus argotaenia*/*A. poilanei*; *A. argotaenia*/*A. yunnanensis*; *A. poilanei*/*A. yunnanensis* and *Taxus chinensis*/*T. wallichiana*) to 2.1% (between *A. argotaenia/T. chinensis*; *A. argotaenia/T. wallichiana*; *A. poilanei/T. chinensis*; *A. poilanei*/*T. wallichiana*; *A. yunnanensis*/*T. chi-*

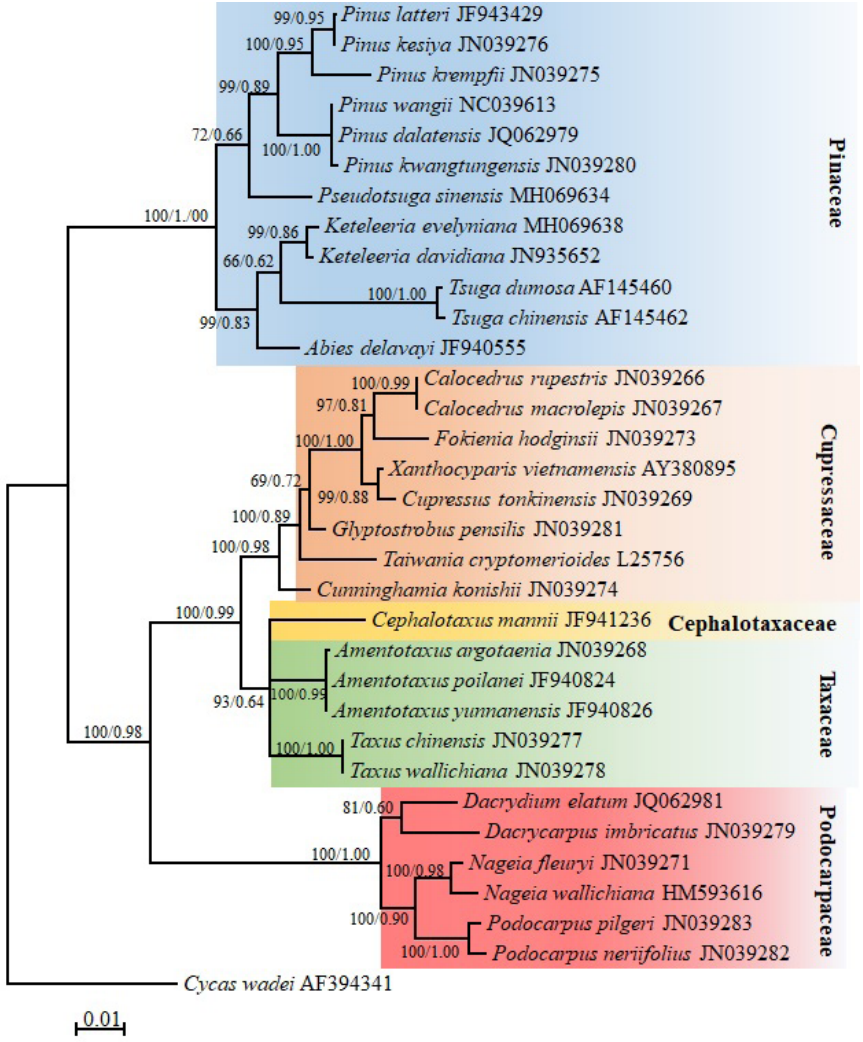


Figure 2. Phylogenetic relationships among the conifer species based on the partial sequence of the chloroplast genes (*rbcL*) using the Maximum likelihood (ML) tree; the numbers above and below the branches represent bootstrap 1 000

https://doi.org/10.17221/88/2020-CJGPB

*nensis* and *A. yunnanensis*/*T. wallichiana*) for *rbcL* (Table 2; Table S2c in the ESM).

In the Pinaceae family (12 species), the mean interspecific divergence was 20.4% ranging from 0.8% between *Pinus kwangtungensis* and *Pinus wangii* to 54.1% between *Tsuga dumosa* and *Pseudotsuga sinensis* for *matK* (Table 2; Table S1d in the ESM). The mean genetic distance (1.2%) ranged from 0% (between *Pinus kesiya* and *P. latteri*) to 5.82% (between *Keteleeria evelyniana*/*P. wangii*) for *rbcL* (Table 2; Table S2d in the ESM).

To identify the conifer species by two DNA barcodes of *matK* and *rcbL*, the maximum likelihood (ML) and Bayesian inference (BI) methods were performed (Figures 1 and 2). The ML and BI produced topologies of (-lnL) 10 009.783 and 10 051.266, respectively for the *matK* region and (-lnL) 5 046.6173 and 5 092.02, respectively for the *rbcL* region.

The phylogenetic relationships of thirty-three local conifer species based on the analysis of the ML and BI showed that all the conifer species form a monophyletic clade with a bootstrap value (Figures 1 and 2). The bootstrap values obtained at the branching nodes of each species ranged from 62 to 100% (Maximum Likelihood bootstrap – MLBS and Bayesian posterior probabilities – BPP) for the *matK* gene (Figure 1); and from 66 to 100% (MLBS) and 60 to 100% (BPP) for the *rbcL* region (Figure 2).

Figure 1 explains that the five family groups were distinctly separated with a high bootstrap value. The analysis indicated that two clades were distinctly separated. The first group was composed of twelve species in Pinaceae with a bootstrap value (MLBS = 100% and BPP = 100%) and the second group was composed of twenty-one species in four families (Podocarpaceae, Cupressaceae, Taxaceae and Cephalotaxaceae) with a high bootstrap value (MLBS = 100% and BPP = 90–100%). Two families (Taxaceae and Cephalotaxaceae) had a high genetic distance with a bootstrap value (MLBS = 100% and BPP = 90%).

Among the two gene regions, the *rbcL* region (Figure 2), was a weak level of identification among the species (even between species in the same family). Two family species (Taxaceae and Cephalotaxaceae) had closed relationships and were not separated. The *rbcL* region was not identified between the species in the Taxaceae family due to the difficulty of the gene region in the species identification. Therefore, gene regions (*matK*) can be used for the identification of thirty-three local conifer species.

## DISCUSSION

The correct identification of a species is essential for the management and conservation of the species (Trias-Blasi & Vorontsova 2015). Identification of a species on the basis of morphology is difficult and inaccurate most of the time, while the DNA barcoding method can provide rapid and accurate species identification (Kress et al. 2015; Bui et al. 2019; Wu et al. 2019). Hollingsworth et al. (2009) recommended *rbcL* and *matK* for the core DNA barcoding in the identification of plants. Burgess et al. (2011) stated that the core barcode has successfully identified 93% of the species of the temperate flora of Canada. The study by de Vere et al. (2012) revealed that *rbcL* + *matK* identified 69.4%–74.1% of flowering plants in Wales, UK. Kress et al. (2015) studied 296 woody species in Panama and found that the species identification rate through *matK* + *rbcL* was high, i.e., 98%. Phong et al. (2018) suggested the use of three gene regions, i.e., *matK*, *trnL* and *rpoC1* as a barcode for the identification of fifteen conifer species in the Central Highland of Vietnam. In our study, the core barcode *matK* region covered the identification of all thirty-three local conifer species (Figure 1), while the *rbcL* regions were weak in some species (Figure 2). This indicates that our results are in line with previous studies and confirmed the effectiveness of the core barcode.

The efficiency of the two gene regions in the species identification was not similar, but can support the effectiveness of the method against the morphological method. Our results suggested the use of the *matK* gene regions as a DNA barcode sequence in three families: Cupressaceae, Podocarpaceae, and Cephalotaxaceae in Vietnam.

## CONCLUSION

In the current study, we had phylogenetic relationships among the native conifer species in Vietnam and used the most suitable genomic marker (*matK*) to identify the local conifer species of Vietnam. The species identification will help in the conservation, and to study the evolution, as well as the systematics of the species.

## REFERENCES

Asaf S., Khan A.L., Khan M.A., Shahzad R., Kang S.M., Al-Harrasi A., Al-Rawahi A., Lee I.-J. (2018): Complete

https://doi.org/10.17221/88/2020-CJGPB

chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species. PLoS One, 13: e0192966.

Bafeel S.O., Arif I.A., Bakir M.A., Khan H.A., Al Farhan A.H., Al Homaidan A.A., Ahamed A., Thomas J. (2011): Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant Omics, 4: 195.

Bui X.P., Do H.T., Shah S.N.M., Nguyen P.D.N., Pham Q.M., Nguyen H.M., Dao M.D., Bui L.M., Vu D.D. (2019): Molecular identification of *Boletus speciosus* (Boletaceae) using nuclear ribosomal internal transcribed spacer region in Vietnam. Research Journal of Biotechnology, 14: 53–67.

Burgess K.S., Fazekas A.J., Kesanakurti P.R., Graham S.W., Husband B.C., Newmaster S.G., Percy D.M., Hajibabaei M., Barrett S.C. (2011): Discriminating plant species in a local temperate flora using the rbcL+ matK DNA barcode. Methods in Ecology and Evolution, 2: 333–340.

Chase M.W., Salamin N., Wilkinson M., Dunwell J.M., Kesanakurthi R.P., Haidar N., Savolainen V. (2005): Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transactions of the Royal Society B: Biological Sciences, 360: 1889–1895.

Chase M.W., Cowan R.S., Hollingsworth P.M., van den Berg C., Madriñán S., Petersen G., Seberg O., Jørgsensen T., Cameron K.M., Carine M. (2007): A proposal for a standardised protocol to barcode all land plants. Taxon, 56: 295–299.

Contreras-Medina R., Vega I.L. (2002): On the distribution of gymnosperm genera, their areas of endemism and cladistic biogeography. Australian Systematic Botony, 15: 193–204.

Cowan R.S., Chase M.W., Kress W.J., Savolainen V. (2006): 300 000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. Taxon, 55: 611–616.

de Vere N., Rich T.C., Ford C.R., Trinder S.A., Long C., Moore C.W., Satterthwaite D., Davies H., Allainguillaume J., Ronca S. (2012): DNA barcoding the native flowering plants and conifers of Wales. PLoS One, **7**: e37945.

Doyle J.J., Doyle J.L. (1990): Isolation ofplant DNA from fresh tissue. Focus, 12: 39–40.

Fazekas A.J., Kuzmina M.L., Newmaster S.G., Hollingsworth P.M. (2012): DNA barcoding methods for land plants. In: DNA Barcodes. Methods in Molecular Biology (Methods and Protocols). Totowa, Humana Press: 223–252.

Forest F., Moat J., Baloch E., Brummitt N.A., Bachman S.P., Ickert-Bond S., Hollingsworth P.M., Liston A., Little D.P., Mathews S. (2018): Gymnosperms on the EDGE. Scientific Reports, 8: 1–11.

Hall T.A. (1999): BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic Acids Symposium Series. London, Information Retrieval Ltd.: 95–98.

Hebert P.D., Cywinska A., Ball S.L., Dewaard J.R. (2003): Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270: 313–321.

Hiep N.T., Loc P.K., Luu N.D.T., Thomas P., Farjon A., Averyanov L.V., Regalado Jr. J. (2005): Viet Nam Conifers: Conservation Status Review 2004. Hanoi, Fauna & Flora International, Vietnam Programme.

Hiep N.T., Khang N.S., Loc P.K. (2015): Gymnosperms of Indochina. St. Petersburg, Komarov Botanical Institute of the Russian Academy of Sciences.

Hollingsworth P.M., Forrest L.L., Spouge J.L., Hajibabaei M., Ratnasingham S., van der Bank M., Chase M.W., Cowan R.S., Erickson D.L., Fazekas A.J., Graham S.W., James K.E., Kim K.-J., Kress W.J., Schneider H., van AlphenStahl J., Barrett S.C.H., van den Berg C., Bogarin D., Burgess K.S., Cameron K.M., Carine M., Chacón J., Clark A., Clarkson J.J., Conrad F., Devey D.S., Ford C.S., Hedderson T.A.J., Hollingsworth M.L., Husband B.C., Kelly L.J., Kesanakurti P.R., Kim J.S., Kim Y.-D., Lahaye R., Lee H.-L., Long D.G., Madriñán S., Maurin O., Meusnier I., Newmaster S.G., Park C.-W., Percy D.M., Petersen G., Richardson J.E., Salazar G.A., Savolainen V., Seberg O., Wilkinson M.J., Yi D.-K., Little D.P. (2009): A DNA barcode for land plants. Proceding of National Academy of Sciences, 106: 12794–12797.

Jobb G. (2011): TREEFINDER Manual. Version of March 2011. Munich. Available at https://www.treefinder.de/tf-march2011-manual.pdf

Kang H.-I., Lee H.O., Lee I.H., Kim I.S., Lee S.-W., Yang T.J., Shim D. (2019): Complete chloroplast genome of *Pinus densiflora* Siebold & Zucc. and comparative analysis with five pine trees. Forests, 10: 600.

Kress W.J., Erickson D.L. (2007): A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS One, 2: e508.

Kress W.J., Wurdack K.J., Zimmer E.A., Weigt L.A., Janzen D.H. (2005): Use of DNA barcodes to identify flowering plants. Proceding of National Academy of Sciences, 102: 8369–8374.

Kress W.J., García-Robledo C., Uriarte M., Erickson D.L. (2015): DNA barcodes for ecology, evolution, and conservation. Trends in Ecology and Evolotion, 30: 25–35.

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

https://doi.org/10.17221/88/2020-CJGPB

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. Proceding of National Academy of Sciences. 105: 2923–2928.

Little D.P., Knopf P., Schulz C. (2013): DNA barcode identification of Podocarpaceae – the second largest conifer family. PLoS One, 8: e81008.

Loc P.K., Long P., Regalado J., Averyanov L., Maslin B. (2017): Native conifers of Vietnam – A review. Pakistan Journal of Botany, 49: 2037–2068.

MOST and VAST (2007): Vietnam Red Data Book. Part II. Plants. Hanoi, Natural Science and Technology Publishing House.

Newmaster S., Fazekas A., Ragupathy S. (2006): DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Canadian Journal of Botany, 84: 335–341.

Pennisi E. (2007): Taxonomy. Wanted: a barcode for plants. Science, 318: 190–191.

Phong D.T., Hien V.T.T., Lieu T.T. (2018): Nucleotide diversity of 15 conifer species in Vietnam's central highlands based on the analysis of ITS, trnH-psbA, matK, trnL and rpoC1 gene regions. Vietnam Journal of Science and Technology, 56: 47.

Ronquist F., Huelsenbeck J.P. (2003): MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19: 1572–1574.

Tanabe A.S. (2011): Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional and separate models for combined molecular phylogenetic analyses of multilocus sequence data. Molecular Ecology Resources, 11: 914–921.

Trias-Blasi A., Vorontsova M. (2015): Plant identification is key to conservation. Nature, 521: 161–161.

Wu F., Li M., Liao B., Shi X., Xu Y. (2019): DNA barcoding analysis and phylogenetic relation of mangroves in Guangdong Province, China. Forests, 10: 56.

Zheng W., Chen J., Hao Z., Shi J. (2016): Comparative analysis of the chloroplast genomic information of *Cunninghamia lanceolata* (Lamb.) Hook with sibling species from the Genera *Cryptomeria* D. Don, *Taiwania* Hayata, and *Calocedrus* Kurz. Intenational Journal of Molecular Science, 17: 1084.

Received: September 23, 2020
Accepted: January 29, 2021
Published online: February 15, 2021