Vietnam Journal of Science and Technology ***56*** *(1) (2018) 47-63*

DOI: 10.15625/2525-2518/56/1/9732

NUCLEOTIDE DIVERSITY OF 15 CONIFER SPECIES IN VIETNAM' S CENTRAL HIGHLAND BASED ON THE ANALYSIS OF *ITS*, *trn*H-*psb*A, *mat*K, *trn*L AND *rpo*C1 GENE REGIONS

Dinh Thi Phong[1, 2, *], Vu Thi Thu Hien[1], Tran Thi Lieu[1]

[1]*Vietnam National Museum of Nature, VAST, 18 Hoang Quoc Viet, Cau Giay, Hanoi*

[2]*Graduate University of Science and Technology, VAST, 18 Hoang Quoc Viet, Cau Giay, Ha Noi*

[*]Email: *dinhthiphong@hotmail.com*

Received: 26 April 2017; Accepted for publication: 30 November 2017

Abstract. In this study, five DNA sequences from *ITS*, *trn*H-*psb*A, *mat*K, *trn*L and *rpo*C1 gene regions were used to explore relationships of 15 conifer species in Highland of Vietnam. All target gene segments have been cloned at size as predicted by the theory for all 15 species of conifers. Nucleotide-level change of 15 coniferous species in five gene regions showed from the highest to the lowest as follows: the *ITS* gene region (0.428), the *trn*H-*psb*A region (0.378), the *trn*L (0.354), the *mat*K gene (0.192) and the *rpo*C1 gene (0.105). The *mat*K gene region showed the highest level of conservation (671 nucleotides) and the *trn*H-*psb*A gene region showed the lowest (78 nucleotides). Phylogenetic tree showed that the species in the same family are formatted in a separate evolutionary branch with bootstrap values obtained from the branching nodes of each species ranging from 52 to 97 % for the *ITS* gene, from 50 to 100 % for *trn*H-*psb*A gene region, from 66 to 100 % for *mat*K gene region, from 50 to 100 % for *trn*L gene region and from 57 to 100 % for *rpo*C1 gene region. Of the three gene regions of *mat*K, *trn*L and *rpo*C1, the grouping of species in the same family showed the most obvious. This result suggests the three gene regions of *mat*K, *trn*L and *rpo*C1 could be used as barcode for the 15 conifer species in Central Highland of Vietnam.

Keywords: conifer, gene regions, Highland, taxonomic classification.

Classification numbers: 1.3.2; 3.1.2.

1. INTRODUCTION

Viet Nam is considered one of ten 'hot spot' for pine conservation in the world, with more than half of its 34 species includes in the Red Book of Global endangered species. Central Highland is regarded as the cradle of coniferous species of Viet Nam. There are 15 species of conifers found in this area [1], of which six conifer species have been recently evaluated as globally threatened. They are: *Pinus krempfii* (VU B1+2c), *Pinus dalatensis* (VU B1+2c), *Pinus latteri* (NT), *Fokienia hodginsii* (NT), *Calocedrus macrolepis* (VU B1 + 2b), *and Cephalotaxus mannii* (VU A1d) [2].

DNA sequences are considered suitable for taxonomic classification by different levels of nucleotide to maintain the conservation in taxa. Today, DNA barcoding is considered an effective technique for distinguishing species. Herbert et al. [3] developed this technique into a tool of classification using short DNA sequence fragments, the nucleotide sequences of the mitochondrial genome in animals and sequences of the chloroplast genome in plants. DNA barcoding has been applied very successfully in the animal taxa. However, despite the fact that it has become an effective tool for plant classification in many studies [4, 5], DNA barcoding application in plants is still controversial [6, 7]. One of the biggest challenges for plant barcoding is distinguishing the sister species in the same geography. The main concern is that genetic barcoding-based system may not be able to distinguish them if the nucleotide variety is too small. Therefore, researchers have focused on specific DNA regions to classify plant species. Of the gene regions, *rbc*L and *mat*K are being widely used as the "DNA barcode" for plants [8, 9]. In this study, we present results on nucleotide diversity of 15 conifer species in Vietnam's Central Highland based on the analysis of *ITS*, *trn*H-*psb*A, *mat*K, *trn*L and *rpo*C1 gene regions, aiming to further extend the scientific basis for molecular analysis-based research. The results of this study provide data of the gene regions which can be used to identify not only some conifer species of high values in the Central Highland of Vietnam, but also some species that have been challenging for field botanists to morphologically classify.

2. MATERIALS AND METHODS

2.1. Materials

In this study, the inner barks from 29 to 38 mature trees (> 20 cm dbh) were randomly sampled from 45 sites, representing the natural range of 15 species (three sites for each species, depending on their distribution in Kon Tum, Lam Dong, Dak Lak and Gia Lai provinces, Figure 1 and Table 1). The nucleotide sequences and the theoretical size of the five primer pairs used in the study is shown in Table 2. Total genomic DNA was isolated from leaves using the method described by Doyle and Doyle [10].

2.2. Methods

MEGA v. 4.0.2 [11] was also used to calculate the proportion of sites differing between pairs of sequences. These are quoted in the text as percentage differences. In the phylogenetic inference method, gaps were not considered as character states. The phylogenetic trees were reconstructed using Neighbor joining method and bootstrap test of phylogeny with 1000 replications [12]. The sequence distances were calculated using nucleotide substitution model: Maximum Composite Likelihood [11]. The gaps and missing data were deleted. Substitutions include the transition and transversions. Pattern among lineages was homogeneous and rates among sites were uniform.

Table 1. Samples' codes, genotypes' locations and conservation status of the fifteen coniferous species.

No.	Name of family	Name of species and conservation status (IUCN 2013)	Collection locality	Sample code	Altitude (◦N)	Longitude (◦E)	Elevation (m)
1	Cephalotaxaceae	*Cephalotaxus mannii* Hook.f. - VU A4acd, B1, 2ab, C	Ta Nung, Da Lat, Lam Đong	VNMN000325	11°56'01.3"	108°23'12.5"	1364
			Hiep An, Duc Trong, Lam Dong	VNMN000330	11°50'13.3"	108°25'37.5"	1392
			Hiep An, Duc Trong, Lam Dong	VNMN000336	11°50'13.3"	108°25'37.5"	1392
2	Cupressaceae	*Fokienia hodginsii* A. Henry & H.H. Thomas - EN A4acd B2ab(iii,v)	Da Chais, Lac Duong, Lam Dong	VNMN000359	12°11'02.7"	108°41'24.3"	1400-1500
			Da Chais, Lac Duong, Lam Dong	VNMN000362	12°03'43.8"	108°37'93"	1400-1500
			Da Chais, Lac Duong, Lam Dong	VNMN000363	12°08'11.2"	108°38'44.9"	1400-1500
3		*Calocedrus macrolepis* Kurz - EN A2acd, A3acd, B2ac, C1	Datanla, Da Lat, Lam Dong	VNMN000364	11°54'02.5"	108°26'56.8"	1315
			Hoa Son, Krong Bong, Dak Lak	VNMN000396	12°25'05.0''	108°22'17.0"	1200
			Son Lang, K' Bang, Gia Lai	VNMN000430	14°30'52.0"	108°33'21.0"	1040-1057
4		*Glyptostrobus pensilis* (Staunt.) K. Koch - CR A1ac, B1 + 2 bc, D1	Eaho, Krong Nang, Dak Lak	VNMN000434	12°59'08.0"	108°17'01.0"	712
			Eaho, Krong Nang, Dak Lak	VNMN000436	12°59'07.0"	108°17'03.0"	712
			Eaho, Krong Nang, Dak Lak	VNMN000438	12°60'01.0"	108°18'06.0"	712
5	Pinaceae	*Keteleria evelyniana* Mast. VU A4acd, B1+2b(ii,iii,v), C	Suoi Vang, Da Lat, Lam Dong	VNMN000439	11°59'58.8"	108°21'59.3"	1464
			Hoa Son, Krong Bong, Dak Lak	VNMN000465	12°25'05.2"	108°22'17.1"	1116
			Dak Glei, Dak Glei, Kon Tum	VNMN000470	15°01'17.0''	107°48'04.0"	1553
6		*Pinus dalatensis* Ferré - VU A2acd, A3acd, B2ac, C1	Da Chais, Lac Duong, Lam Dong	VNMN000500	12°11'02.7"	108°41'24.3"	1482
			Hoa Son, Krong Bong, Dak Lak	VNMN000529	12°29'33.2"	108°18'17.1"	1116
			Xa Hieu, Kon Plong, Kon Tum	VNMN000564	14°40'40.0"	108°23'36.0"	1159
7		*Pinus kesiya* Royle ex Gordon (VU A4 acd, B2ac, C1)	Suoi Vang, Da Lat, Lam Dong	VNMN000579	11°59'58.8"	108°21'59.3"	1464
			Da Chais, Lac Duong, Lam Dong	VNMN000584	12°11'02.7"	108°41'24.3"	1482
			Da Chais, Lac Duong, Lam Dong	VNMN000585	12°08'11.2"	108°38'44.9"	1400-1500
8		*Pinus latteri Mason* (VU)	Hiệp An, Duc Trong, Lam Dong	VNMN000586	11°49'52.5"	108°25'26.0"	1390
			Hiệp An, Duc Trong, Lam Dong	VNMN000589	11°50'36.1"	108°28'21.5"	1390

			Dak Glei, Dak Glei, Kon Tum	VNMN000591	15°01'17.0''	107°48'04.0"	1553
9		*Pinus krempfii* Lecomte (VU A2acd, A3acd, B2ac, C1)	Da Chais, Lac Dương, Lam Dong	VNMN000592	12°11'01.3"	108°41'20.3"	1482 – 1485
			Lat, Lac Dương, Lam Dong	VNMN000606	12°05'17.9"	108°22'13.0"	1659-1757
			Hoa Son, Krong Bong, Dak Lak	VNMN000642	12°25'05.2"	108°22'17.1"	1110- 1120
10	Podocarpaceae	*Dacrycarpus imbricatus* (Blume) de Laub. (LC)	Da Chais, Lac Duong, Lam Dong	VNMN000662	12°11'02.5"	108°41'24.0"	1482
			Da Chais, Lac Duong, Lam Dong	VNMN000665	12°11'02.3"	108°41'24.1"	1482
			Ngoc Linh, Dak Glei, Kon Tum	VNMN000669	15°04'23.0''	107°57'31.0"	1935
11		*Dacrydium elatum* Roxb.) Wall.) (VU A4acd, B2b(ii,iii,v), C1)	Da Chais, Lac Duong, Lam Dong	VNMN000670	12°11'02.7"	108°41'24.3"	1482
			Hoa Son, Krong Bong, Dak Lak	VNMN000693	12°25'05.2"	108°22'17.1"	1116
			Xa Hieu, Kon Plong, Kon Tum	VNMN000718	14°40'06.8"	108°24'30.3"	1194
12		*Nageia wallichiana* (C. Presl) Kuntze (VU B2ab(iii,v)	Ta Nung, Da Lat, Lam Dong	VNMN000740	11°56'01,3"	108°53'12,5"	1364
			Hoa Son, Krong Bong, Dak Lak	VNMN000775	12°25'05,2"	108°22'17,1"	1016
			Xa Hieu, Kon Plong, Kon Tum	VNMN000803	14°35'05.0"	108°24'55.0"	1267
13		*Podocarpus neriifolius* D. Don (LC)	Da Chais, Lac Duong, Lam Dong	VNMN000810	12°11'13.1"	108°42'55.3"	1593
			Hoa Son, Krong Bong, Dak Lak	VNMN000812	12°29'28.3"	108°18'38.4"	1120
			Đak Glei, Dak Glei, Kon Tum	VNMN000814	15°01'17.0''	107°48'04.0"	1553
14	Taxaceae	*Amentotaxus poilanei* D.K. Ferguson (VU D2)	Ngoc Linh, Dak Glei, Kon Tum	VNMN000815	15°03'20.0''	107°58'31.0"	1935
			Ngoc Linh, Dak Glei, Kon Tum	VNMN000817	15°04'23.0''	107°57'31.0"	1935
			Ngoc Linh, Dak Glei, Kon Tum	VNMN000818	15°04'25.1''	107°52'30.0"	1935
15		*Taxus wallichiana* Zucc (EN A4acd, B1b,2, C1)	Da Chais, Lac Duong, Lam Dong	VNMN000819	12°03'43.8"	108°37'93.0"	1533
			Da Chais, Lac Duong, Lam Dong	VNMN000823	12°08'11.2"	108°38'44.9"	1400-1500
			Da Chais, Lac Duong, Lam Dong	VNMN000827	12°11'02.7"	108°41'24.3"	1400-1500

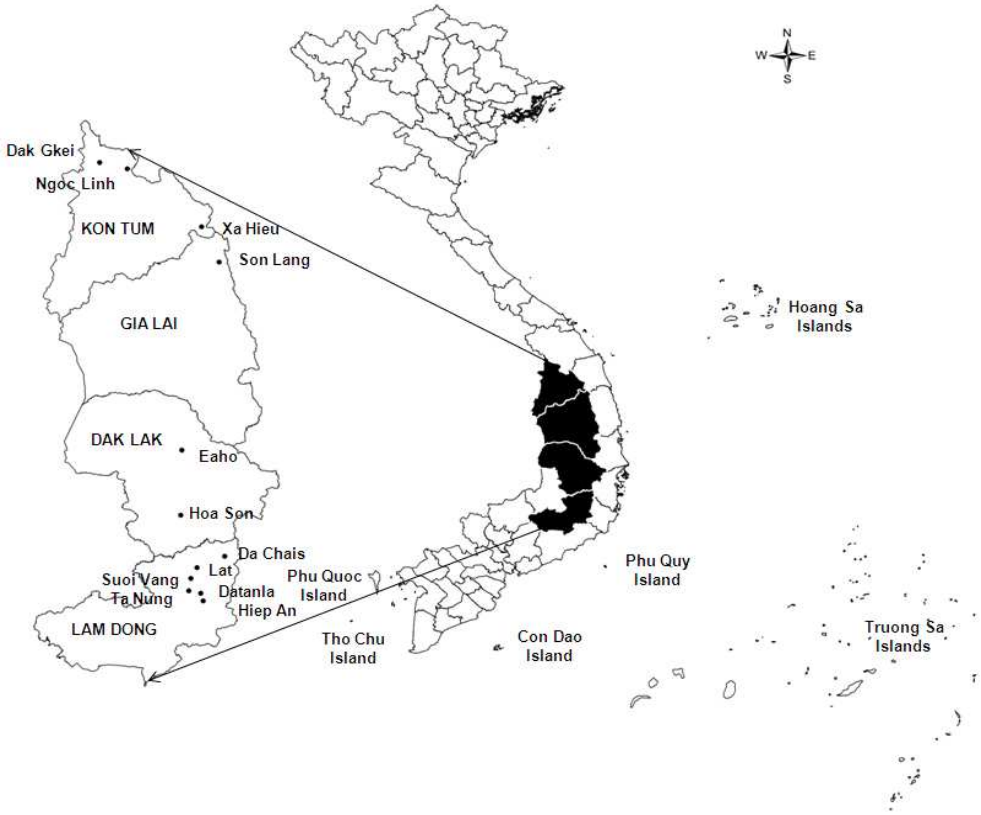

Table 2. List of primer pairs used in this study.

Primer names	Primer sequences (5'– 3')	Expected size (bp)	Annealing temperature (°C)	Origin
*trn*H/ *psb*A	GTTATGCATGAACGTAATGCTC CGCGCATGGTGGATTCACAATCC	680	50-53	Sang et al., 1997 [22] Tate et al., 2003 [23]
*ITS*F/ *ITS*R	CCTGCGGAAGGATCATTGTC TTAAACTCAGCGGGTAGTC	1100	50-53	Designed based on *Elleanthus conifer* coded EU490666 Genbank (2014)
*rpo*C1F/ *rpo*C1R	GTGGATACACTTCTTGATAATGG TGAGAAAACATAAGTAAACGGGC	600	50-52	http://www.kew.org/barcoding/protocols.html [24]
*trn*LF/ *trn*FR	CGAAATCGGTAGACGCTACG ATTTGAACTGGTGACACGAG	1000	50-53	Taberlet et al., 2006 [13]
*mat*KF/ *mat*KR	TGGCAFTGCAATCAAAAC ATCGCTAATCAATAAATCATCT	950	50-55	Designed based on *Taxus wallichiana* coded HM590991 Genbank (2014)

3. RESULTS AND DISCUSSIONS

3.1. Character analysis of gene regions

We successfully cloned target gene fragments at sizes as theoretically predicted for all 45 samples of the 15 coniferous species. Gene fragments of three samples of the same species were amplified with the same size. Different sizes of gene segments were amplified for each species, ranging from 975 to 1200 bp for the nuclear *ITS* region; from 400 to 780 bp for the *trn*H-*psb*A region; from 650 to 1000 bp for *mat*K region; from 400 to 970 bp for the *trn*L region and 600 bp for *rpo*C1 region (Figure 2). In this study, a total of 165 nucleotide sequences of the five gene regions have been obtained for fifteen species, which have been deposited in Genbank/EMBL databases, including 165 accession numbers KR780651- KR780665; KR907882- KR907890; KR920092- KR920100; KT001114- KT001125; KT150253- KT150258; KT222871- KT222882; KT236090- KT236092; KT247644- KT247646; KT265683- KT265685; KT272170- KT272172; KT008100- KT008105; KT037124- KT037129; KR855700- KR855711; KR674115- KR674126; KT072777- KT072785; KR605490- KR605498 and KU940072- KU940107) available online.

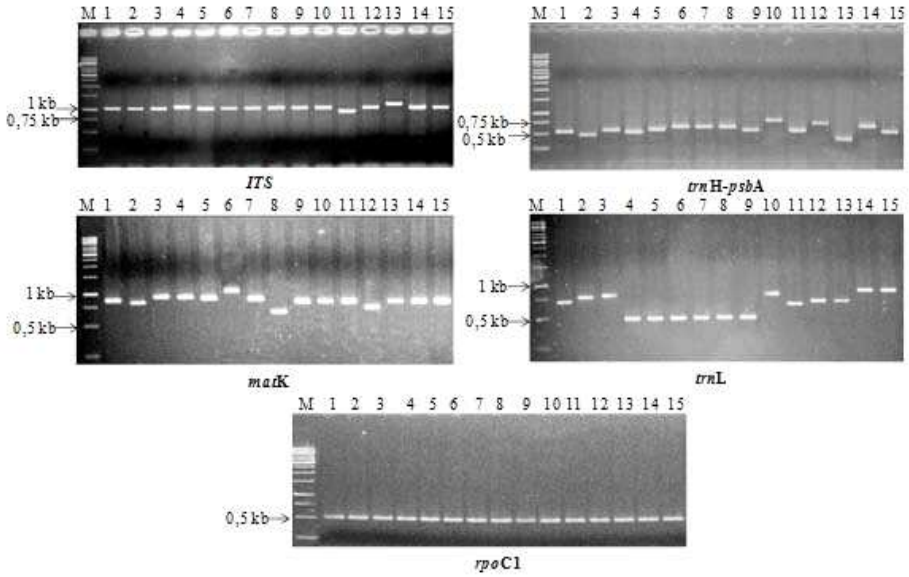

Figure 2. PCR representative products of coniferous species analyzed five gene regions on the 1% agarose (M: Molecular ladder1 kb; 1: *Cephalotaxus mannii,* 2*: Fokienia hodginsii,* 3: *Calocedrus macrolepis,* 4: *Glytostrotrobus pensilis,* 5: *Keteleria evelyniana,*6: *Pinus dalatensis,*7: *Pinus kesiya,* 8: *Pinus latteri,* 9: *Pinus krempfii,* 10: *Dacrycarpus imbricatus,* 11: *Dacrydium elatum,* 12: *Nageia wallichiana,* 13: *Podocarpus neriifolius,* 14: *Amentotaxus poilanei* and 15: *Taxus wallichiana*).

3.2. Nucleotide diversity

Results from compared nucleotide sequences using MEGA v.4.0.2 software between samples of the same species at the *ITS*, *trn*H-*psb*A, *mat*K, *trn*L and *rpo*C1 gene regions showed that their length and nucleotide sequence similarity were 100 %. Therefore, the following study took only one representative sample of each species. The degree of nucleotide diversity among 15 species ranged from 0.000 (such as *P. latteri* with *P. dalatensis*, *P. krempfii* with *P. dalatensis*, *P. kesiya* with *P. krempfii,* etc...) to 2.642 % (*P. neriifolius* with *C. macrolepis*) for the *ITS* gene (Table 3); From 0.000 (*P. latteri* with *P. kesiya*) to 3.552 % (*P. wallichiana* with *P. kesiya* and *P. latteri)* for *trn*H-*psb*A region (Table 4); from 0.009 (*P. latteri* with *P. kesiya*) to 0.367 % (between *P. neriifolius* with *P. kesiya* and *P. latteri*) for the *mat*K region (Table 5);

From 0.000 (between *F. hodginsii* and *C. macrolepis*) to 1.237 % (between *K. evelyniana* and *P. dalatensis*) for *trn*L gene region (Table 6); from 0.000 (between *P. dalatensis* and *P. krempfii)* to 0.177 % (between *C. macrolepis* with *K. evelyniana, P. dalatensis, P.latteri* and *P. krempfii*) for the *rpo*C1 region (Table 7). Among the five regions, the *trn*H-*psb*A region showed the highest level of nucleotide diversity (from 0.000 to 3.552 %) and the *rpo*C1 region showed the lowest (from 0.000 to 0.177 %).

In this study, characters of conservation, variation and parsimony information were 124, 1146 and 801 for the nuclear *ITS* gene; 78, 726 and 546 for the *trn*H-*psb*A region; 671, 487 and 331 for the *mat*K gene; 218, 778 and 493 for the *trn*L gene; and 372, 190 and 127 for the *rpo*C1 gene. The number of variable sites (V) for individual loci among fifteen species ranged from 190 (*rpo*C1) to 1146 (*ITS*). The value of the highest conservative characters (C) across fifteen species was 671 for the *mat*K region and the lowest for *trn*H-*psb*A region (78) (Table 8). Analysis results of the five gene regions showed that among 15 species, the nuclear *ITS* gene had the highest level of nucleotide diversity (0.428), followed by the *trn*H-*psb*A region (0.378), the *trn*L gene (0.354), is the *mat*K gene (0.192). The *rpo*C1 gene had the lowest level (0.105) (Table 8). They also showed that, the chloroplast gene regions had the more conservative characters than the nuclear *ITS* gene region.

We also checked the divergence to distinguish species among the 15 ones, the *mat*K is the most powerful, with 100 % discriminated species pairs. Another species pairs similarity was founded in each of three gene regions: *trn*H-*psb*A (VNMN000586_ *P. latteri* and VNMN000579_ *P. kesiya),* *trn*L (VNMN000359_*F. hodginsii* and VNMN000364_*C. macrolepis)* and *rpo*C1 (VNMN000509_ *P. dalatensis* and VNMN000592_ *P. krempfii)* (Table 4, 6 and 7).

3.3. Phylogeny based on gene regions

Because samples (individuals) of the same species have the identical sequence, only one individual of each species was selected to reconstruct the phylogenetic tree from five DNA regions of 15 coniferous species. The species were separated supported by bootstrap values > 50 %. The phylogenetic tree of 15 coniferous species based on the analysis of method NJ (Neighbor - Joining) in Figure 3 showed that all 15 species of conifers formed a separate subsidiary and joint evolution closely together with bootstrap values obtained at the branching nodes of each species ranged from 52 to 97 % for the *ITS* gene (Figure 3A); from 50 to 100 % for the *trn*H-*psb*A region (Figure 3B); from 66 to 100 % for the *mat*K gene region (Figure 3C); from 50 to 100 % for the *trn*L gene region (Figure 3D) and from 57 to 100 % for the *rpo*C1 gene region (Figure 3E).

Table 3. Nucleotide diversity of the 15 coniferous species analyzing ITS region.

No.	Name of samples	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	VNMN000325_ *C. mannii*															
2	VNMN000359_ *F. hodginsii*	0.725														
3	VNMN000364_ *C. macrolepis*	2.614	2.268													
4	VNMN000434_ *G. pensilis*	0.640	0.427	2.617												
5	VNMN000439_ *K. evelyniana*	1.638	1.902	2.351	1.673											
6	VNMN000509_ *P. dalatensis*	1.647	1.890	2.336	1.682	0.001										
7	VNMN000579_ *P. kesiya*	1.647	1.890	2.336	1.682	0.001	0.000									
8	VNMN000586_ *P. latteri*	1.647	1.890	2.336	1.682	0.001	0.000	0.000								
9	VNMN000592_ *P. krempfii*	1.647	1.890	2.336	1.682	0.001	0.000	0.000	0.000							
10	VNMN000662_ *D. imbricatus*	1.638	1.902	2.322	1.673	0.003	0.004	0.004	0.004	0.004						
11	VNMN000670_ *D. elatum*	1.699	1.918	2.447	1.754	0.011	0.011	0.011	0.011	0.011	0.013					
12	VNMN000740_ *N. wallichiana*	1.728	1.998	2.218	1.768	0.020	0.021	0.021	0.021	0.021	0.023	0.026				
13	VNMN000810_ *P. neriifolius*	1.030	1.152	2.642	1.077	2.258	2.241	2.241	2.241	2.241	2.228	2.378	2.368			
14	VNMN000815_ *A. poilanei*	0.378	0.632	2.292	0.639	1.865	1.876	1.876	1.876	1.876	1.865	1.945	1.985	0.953		
15	VNMN000819_ *T. wallichiana*	0.510	0.735	2.164	0.745	1.809	1.820	1.820	1.820	1.820	1.812	1.864	1.904	1.093	0.472	

Table 4. Nucleotide diversity of the 15 coniferous species analyzing *trn*H-*psb*A region.

No.	**Name of samples**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	VNMN000325_ *C. mannii*														
2	VNMN000359_ *F. hodginsii*	0.206													
3	VNMN000364_ *C. macrolepis*	0.716	0.636												
4	VNMN000434_ *G. pensilis*	0.441	0.403	0.134											
5	VNMN000439_ *K. evelyniana*	1.400	1.617	1.219	0.991										
6	VNMN000509_ *P. dalatensis*	1.801	1.805	1.532	1.245	0.109									
7	VNMN000579_ *P. kesiya*	1.801	1.805	1.516	1.227	0.117	0.006								
8	VNMN000586_ *P. latteri*	1.801	1.805	1.516	1.227	0.117	0.006	0.000							
9	VNMN000592_ *P. krempfii*	1.736	1.739	1.545	1.287	0.123	0.026	0.020	0.020						
10	VNMN000662_ *D. imbricatus*	2.018	2.412	3.139	2.531	3.118	3.282	3.305	3.305	3.267					
11	VNMN000670_ *D. elatum*	1.663	2.056	2.867	2.361	3.127	3.247	3.247	3.247	3.205	1.540				
12	VNMN000740_ *N. wallichiana*	2.151	2.424	3.200	2.680	3.284	3.529	3.552	3.552	3.514	0.089	1.821			
13	VNMN000810_ *P. neriifolius*	1.913	2.381	3.045	2.538	3.101	3.263	3.285	3.285	3.247	0.041	1.554	0.074		
14	VNMN000815_ *A. poilanei*	2.243	2.356	2.388	1.674	2.450	2.724	2.743	2.743	2.633	2.490	2.740	2.532	2.656	
15	VNMN000819_ *T. wallichiana*	0.529	0.517	0.547	0.270	1.130	1.268	1.249	1.249	1.187	2.739	2.530	2.797	2.595	2.026

Table 5. Nucleotide diversity of the 15 coniferous species analyzing *mat*K region.

No.	**Name of samples**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	VNMN000325_ *C. mannii*														
2	VNMN000359_ *F. hodginsii*	0.174													
3	VNMN000364_ *C. macrolepis*	0.167	0.026												
4	VNMN000434_ *G. pensilis*	0.168	0.054	0.045											
5	VNMN000439_ *K. evelyniana*	0.232	0.310	0.288	0.273										
6	VNMN000509_ *P. dalatensis*	0.282	0.344	0.317	0.314	0.093									
7	VNMN000579_ *P. kesiya*	0.287	0.353	0.331	0.328	0.093	0.051								
8	VNMN000586_ *P. latteri*	0.282	0.348	0.325	0.322	0.096	0.061	0.009							
9	VNMN000592_ *P. krempfii*	0.282	0.348	0.326	0.319	0.086	0.039	0.018	0.020						
10	VNMN000662_ *D. imbricatus*	0.254	0.267	0.256	0.255	0.290	0.340	0.348	0.348	0.339					
11	VNMN000670_ *D. elatum*	0.248	0.270	0.258	0.258	0.294	0.341	0.345	0.345	0.335	0.041				
12	VNMN000740_ *N. wallichiana*	0.262	0.303	0.291	0.290	0.309	0.332	0.360	0.360	0.341	0.072	0.079			
13	VNMN000810_ *P. neriifolius*	0.288	0.316	0.304	0.297	0.334	0.346	0.367	0.367	0.366	0.079	0.086	0.060		
14	VNMN000815_ *A. poilanei*	0.090	0.169	0.151	0.137	0.221	0.254	0.266	0.266	0.262	0.235	0.229	0.256	0.255	
15	VNMN000819_ *T. wallichiana*	0.110	0.159	0.141	0.131	0.230	0.264	0.280	0.275	0.263	0.238	0.232	0.247	0.262	0.054

Table 6. Nucleotide diversity of the 15 coniferous species analyzing *trn*L region.

No.	Name of samples	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	VNMN000325_ *C. mannii*														
2	VNMN000359_ *F. hodginsii*	0.280													
3	VNMN000364_ *C. macrolepis*	0.280	0.000												
4	VNMN000434_ *G. pensilis*	0.297	0.049	0.049											
5	VNMN000439_ *K. evelyniana*	0.367	0.392	0.392	0.393										
6	VNMN000509_ *P. dalatensis*	1.043	1.122	1.122	1.198	1.237									
7	VNMN000579_ *P. kesiya*	1.008	1.054	1.054	1.122	1.090	0.039								
8	VNMN000586_ *P. latteri*	1.043	1.119	1.119	1.195	1.180	0.044	0.024							
9	VNMN000592_ *P. krempfii*	1.024	1.111	1.111	1.187	1.222	0.034	0.044	0.029						
10	VNMN000662_ *D. imbricatus*	0.296	0.391	0.391	0.402	0.412	0.911	0.893	0.895	0.878					
11	VNMN000670_ *D. elatum*	0.350	0.488	0.488	0.502	0.438	1.089	1.065	1.086	1.065	0.101				
12	VNMN000740_ *N. wallichiana*	0.287	0.379	0.379	0.390	0.377	0.913	0.895	0.911	0.895	0.039	0.101			
13	VNMN000810_*P. neriifolius*	0.302	0.411	0.411	0.408	0.409	0.932	0.913	0.930	0.913	0.039	0.085	0.034		
14	VNMN000815_ *A. poilanei*	0.172	0.294	0.294	0.303	0.359	1.176	1.116	1.190	1.182	0.318	0.339	0.295	0.294	
15	VNMN000819_ *T. wallichiana*	0.277	0.308	0.308	0.320	0.423	1.155	1.133	1.208	1.199	0.330	0.395	0.338	0.338	0.216

Table 7. Nucleotide diversity of the 15 coniferous species analyzing *rpo*C1 region.

No.	**Name of samples**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	VNMN000325_ *C. mannii*														
2	VNMN000359_ *F. hodginsii*	0.082													
3	VNMN000364_ *C. macrolepis*	0.087	0.021												
4	VNMN000434_ *G. pensilis*	0.078	0.047	0.050											
5	VNMN000439_ *K. evelyniana*	0.171	0.171	0.177	0.174										
6	VNMN000509_ *P. dalatensis*	0.164	0.165	0.177	0.162	0.028									
7	VNMN000579_ *P. kesiya*	0.167	0.164	0.173	0.164	0.042	0.018								
8	VNMN000586_ *P. latteri*	0.171	0.168	0.177	0.168	0.039	0.015	0.003							
9	VNMN000592_ *P. krempfii*	0.164	0.165	0.177	0.162	0.028	0.000	0.018	0.015						
10	VNMN000662_ *D. imbricatus*	0.138	0.151	0.156	0.141	0.135	0.126	0.132	0.135	0.126					
11	VNMN000670_ *D. elatum*	0.151	0.161	0.166	0.151	0.142	0.133	0.139	0.142	0.133	0.023				
12	VNMN000740_ *N. wallichiana*	0.144	0.154	0.160	0.145	0.139	0.127	0.133	0.136	0.127	0.021	0.034			
13	VNMN000810_ *P. neriifolius*	0.148	0.151	0.156	0.142	0.142	0.130	0.136	0.139	0.130	0.023	0.036	0.003		
14	VNMN000815_ *A. poilanei*	0.061	0.081	0.099	0.078	0.165	0.152	0.155	0.158	0.152	0.123	0.136	0.130	0.133	
15	VNMN000819_ *T. wallichiana*	0.053	0.067	0.078	0.069	0.142	0.135	0.138	0.141	0.135	0.123	0.136	0.133	0.136	0.045

Table 8. Summary of characteristics of 5 DNA barcodes evolution and ability of distinguishing species for each gene region.

Gene regions	m	C	V	Pi	π	P%
ITS	15	124	1146	801	0.428	58.2
*trn*H-*psb*A	15	78	726	546	0.378	97.1
*mat*K	15	671	487	331	0.192	100
*trn*L	15	218	778	493	0.354	97.1
*rpo*C1	15	372	190	127	0.105	97.1

m: *Number of species; C: Consevative characters; V: Variable characters; Pi: Parsimony informative characters; π: nucleotide diversity. P: The power to distinguish species (%).*

The data also indicated that species in the same genus of the family clustered in the same branch of evolution. For example, five species in Pinaceae family, including *Keteleeria evelyniana* (VNMN000439_*K. evelyniana*), *Pinus dalatensis* (VNMN000500_*P. dalatensis*), *Pinus kesiya* (VNMN000579_*P. kesiya*), *Pinus latteri* (VNMN000586_*P. latteri*) and *Pinus krempfii* (VNMN000592_*P. krempfii*) formed a branch of evolution with bootstrap valuesranging from 70 to 100 % for the *trn*H-*psb*A region; from 93 to 100 % for the *mat*K region, from 84 to 100 % for the *trn*L region and from 68 to 100 % for the *rpo*C1 gene region. Branching level of nuclear *ITS* region (Figure 3A) is the weakest in the five gene regions. The branching level was weak even between species in the same family. For example, some species in the family Pinaceae (such as VNMN000592_*P. krempfii*, VNMN000589_*P. latteri*, VNMN000500_*P. dalatensis* and VNMN000579_*P. kesiya*) formatted with the species of family Podocarpaceae (such as VNMN000662_*D. imbricatus*, VNMN000670_*D. elatum*, VNMN000740_*N. wallichiana*). Position classification between families on the phylogenetic tree were the most obvious in three regions of *mat*K, *trn*L and *rpo*C1. Principally, in the same species, they made up a branch with bootstrap values ranging from 50 % (Taxaceae) for the gene region *trn*L (Figure 3D) to 100 % (Pinaceae) with the three gene regions of *mat*K (Figure 3C), *trn*L (Figure 3D) and *rpo*C1 (Figure 3E). Therefore, the three gene regions of *mat*K, *trn*L and *rpo*C1 can be used to identify 15 coniferous species in Central Highland of Viet Nam.

The five gene regions of *ITS*, *trn*L, *mat*K, *rpo*C1 and *trn*H-*psb*A were used as barcode objects in many cultivars, but they have still limitations. For example, the nuclear *ITS* region has not been successfully cloned for some species groups, or in the *trn*H-*psb*A gene region there are still more "indels" leading to difficulties in comparing the nucleotide sequences, or the two regions of *trn*L and *mat*K have very low nucleotide variations [13, 14]. However, in this study, we successfully cloned the gene fragments for all of five gene regions of all 15 species. Among them, the three gene regions of *mat*K, *trn*L and *rpo*C1 were capable of grouping most of species in the same family together (Figure 3C, 3D and 3E). Among them, the bootstrap value of the *mat*K region was the highest at the nodes among species of the same family (from 66 to 100 %) and among families (from 67 to 100 %), followed by the *rpo*C1 region (from 57 to 100 % between species in the same family and from 63 to 100 % between different families), and lastly is the region *trn*L (from 50 to 100 % between species in the same family and from 51 to 79 % between family together). The two nuclear gene regions of *ITS* and *trn*H-*psb*A did not group the species in the same family (Figure 3A and 3B). Contrast to this result, the nuclear *ITS* region to

be very effective to identify the 8 *Dalbergia* species in the genus *Dalbergia* of Viet Nam [15]. However, the region *trn*H-*psb*A were proposed as DNA barcoding for species of Taxaceae [16, 17], but previous research revealed that they were not suitable for a number of *Dalbergia* species in the family Fabaceae in Viet Nam [13]. The *mat*K gene region is suggested as DNA barcoding in plants [18]. In this study, this region also demonstrated its capability, which can be seen on Figure 3D. However, this gene region was not able to perform as barcode for some wood species of the genus *Dalbergia* of Viet Nam in the study of Phong et al. [15], while the *mat*K gene has been unable to separate the two species *D. entadoides* and *D. dialoides* or *D. hencei* and *D. oliveri* with bootstrap value of 90% and 96%, respectively. As announced by the group CBOL [19], the three gene regions of *trn*H-*psb*A, *mat*K and *trn*L may be appropriate for the study of DNA barcoding in plants because they are exact clones of the target gene fragments by specific primers, and decoding their sequences may make it possible to distinguish plant taxa. Meanwhile, the *trn*H-*psb*A region [17, 20] and the *matK*-barcode [21] have been proposed as DNA barcoding in plants. Our study reconfirmed the appropriate use of the three gene regions of *mat*K, *trn*L and *rpo*C1 to discriminate the 15 coniferous species in Central Highland of Viet Nam (Figure 3C, 3D and 3E). Although the efficiency of each of the gene region for these species were not similar, our results suggested that the technique of decoding and comparing different nucleotide sequences could effectively support traditional identification by morphology. Our study also proposed that the three gene regions of *mat*K, *trn*L and *rpo*C1 are the best option for DNA barcoding for interspecific variation of 15 coniferous species in the Central Highland of Vietnam.

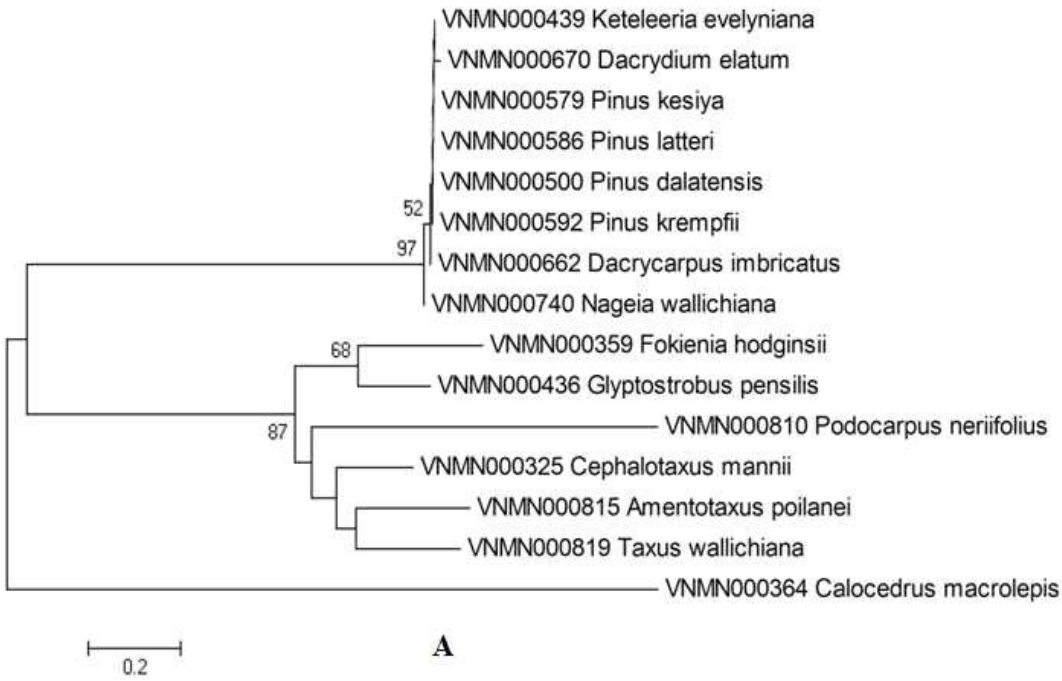

Figure 3. Phylogenetic tree reconstruction using NJ method for five studied DNA regions, *ITS* (A), *trn*H-*psb*A (B); *mat*K (C); *trn*L (D) and *rpo*C1(E). Numbers above branches indicate bootstrap values.

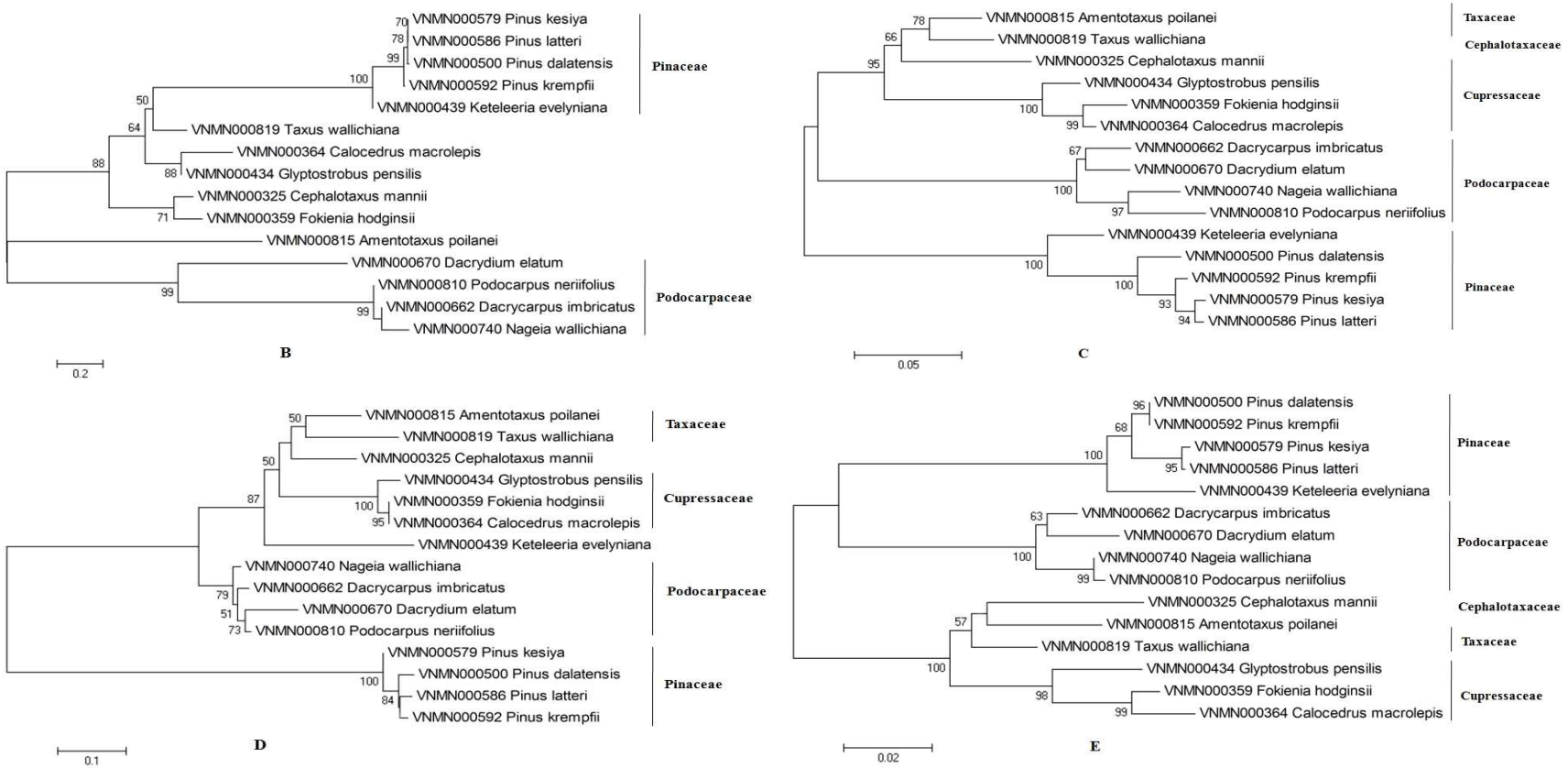

Figure 3. (continued).

4. CONCLUSIONS

Of the five gene regions, the highest level of nucleotide diversity was shown in the *trn*H-*psb*A region (from 0.000 to 3.552 %), while the lowest was in the *rpo*C1 region (from 0.000 to 0.177 %). The *mat*K gene is the most conservative (671 nucleotides) and the *trn*H-*psb*A gene region is the least (78 nucleotides). The capability to distinguish species among15 species of the *mat*K region was the highest, with 100 % discriminated species pairs. The *ITS* region did not have sufficient capability to distinguish 6 species pairs (58.3 %). Those results suggested that the three gene regions of *mat*K, *trn*L and *rpo*C1 could be used as barcode for 15 conifer species in Central Highland of Vietnam.

***Acknowledgments*:** This research was funded by Tay Nguyen 3 Program (Project code TN3/T15). The authors gratefully acknowledge the assistance and support in sample collection of Ngoc Linh Nature Reserve (Kon Tum province), Kon Ka Kinh National Park (Gia Lai province), Bidoup – Nui Ba National Park (Lam Dong province), Chu Yang Sin National Park (Dak Lak), Kon Tum Science and Technology Department, Dak Lak Science and Technology Department, and Lam Dong Science and Technology Department in Vietnam. We are grateful for Dr. Nguyen Tien Hiep for his help in the field survey and collection.

REFFERENCES

1. Nguyen Tien Hiep, Phan Ke Loc, Nguyen Duc To Luu, Philip Ian Thomas, Aljos Farjon, Leonid Averyanov and Jacinto Regalado Jr. - Viet Nam Conifers: Conservation Status Review 2004. Flora and Fauna International, Viet Nam Programme, Hanoi, 2005.
2. IUCN. 2013. IUCN Red List of Threatened Species (ver. 2013.1). Available at: http://www.iucnredlist.org. (Accessed: 12 June 2013).
3. Hebert P. D. N., Cywinska A., Ball S. L. and Waard J. R. - Biological identifications through DNA barcodes, Proc. R. Soc. Lond. B Biol. Sci. **270** (2003) 313–321.
4. Newmaster S. G., Fazekas A. J. and Ragupathy S. - DNA barcoding in land plants: evaluation of *rbc*L in a multigene tiered approach, Can. J. Bot. **84** (2006) 335-341.
5. Lahaye R. M., Bank M. V. D. and Bogarin D. - DNA barcoding the floras of biodiversity hotspots, Proc. Nat. Acad. Sci. USA. **105** (2008) 2923-2928.
6. Cowan R. S., Chase M. W., Kress W. J. and Savolainen V. - 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants, Taxon **55 (**2006) 611-616.
7. Pennisi E. - Taxonomy. Wanted: a barcode for plants, Sci. **318 (**2007) 190-191.
8. Little D. P., Knopf P. and Schulz C. - DNA Barcode Identification of Podocarpaceae – The Second Largest Conifer Family, PLoS ONE **8** (11) (2013) e81008.
9. Bafeel S. O., Arif I. A., Bakir M. A., Khan H. A., Ahmad H., Farhan., Ali A., Homaidan A., Ahamed A. and Thomas J. - Comparative evaluation of PCR success with universal primers of maturase K (*mat*K) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (*rbc*L) for barcoding of some arid plants, POJ **4**(4) (2011) 195-198.
10. Doyle J. J. and Doyle J. - A rapid DNA isolation procedure for small quantities of fresh leaf tissue, Phytochem. Bull. **19 (**1987) 11-15.

11. Tamura K., Dudley J., Nei M. and Kumar S. - MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, Mol. Biol. Evol. **24** (2007) 1596-1599.
12. Saitou N. and Nei M. - The neighbor-joining method: A new method for reconstructing phylogenetic trees, Mol. Biol. Evol. **4 (**1987) 406-425.
13. Taberlet P., Coissac E., Pompanon F., Gielly L., Miguel C., Valentini A., Vermat T., Corthier G., Brochmann C. and Willerslev E. - Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding, Nucleic Acid Research **35 (**2006) 14.
14. Liu Y., Yan H. F. and Ge X. J. - Evaluation of 10 plant barcordes in Bryophyta (Mosses), J. Syst. Evol. **48** (2010) 36-46.
15. Phong D. T., Tang D. V., Hien V. T. T., Ton N. D. and Hai N. V. - Nucleotide diversity of a nuclear and four chloroplast DNA regions in rare tropical wood species of *Dalbergia* in Vietnam: A DNA barcode identifying utility, Asian Journal of Applied Sciences **2** (2) 2014 116-125.
16. Song J., Yao H., Li Y., Li X., Lin Y., Liu C., Han J., Xie C. and Chen S. - Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique, J. Ethnop. **124 (**2009) 434-439.
17. Hao D. C., Chen S. L. and Xiao P. G. - Sequence characteristics and divergent evolution of the chloroplast *psb*A-*trn*H noncoding region in gymnosperms, J. App. Genet. **51** (2010) 259-273.
18. Peter M. H., Laura L. F. and John L. S. - A DNA barcode for land plants, Proc. Natl. Acad. Sci. USA. **106 (**2009) 12794-12797.
19. CBOL Plant Working Group - A DNA barcode for land plants, Proc. Natl. Acad. Sci. **106 (**2009) 12794-12797.
20. Julian R. S., Robert F. C. N. and Btianna N. C. - Plant DNA barcodes and species resolution in sedges (Carex, Cyperaceae), Mol. Ecol. Res. **9** (2009) 151-163.
21. Fazekas A. J., Kuzmina L., Newmaster S. G. and Hollingsworth P. M. - DNA barcoding methords for land plants. In: Kress W. J., Erickson D. L. (ed) DNA barcodes: Methods and Protocols, Meth. Mol. Biol. **858 (**2012) 223-252.
22. Sang T., Crawford D. J. and Stuessy T. F. - Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae), Am. J. Bot. **84** (8) (1997) 1120-1136.
23. Tate J. A. and Simpson B. B. - Paraphyly of Tarasa (Malvaceae) and diverse origins of the polyploidy species, Syst. Bot. **28 (**2003) 723-737.
24. http://www.kew.org/barcoding/protocols.html. Accessed July 16, 2009.