Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2020, 7(1):57-64**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Genetic Diversity of *Morinda officinalis* Based on *ITS* and *matK, rbcL* and Interspace *trnH-psbA* Genes on the Chloroplast

Pham Van Kien[1], Doan Cao Son[1], Bui Hong Cuong[2], Tran Minh Ngoc[3], Tran Ngoc Thanh[3], Tran Dang Khanh[4], Pham Thi Minh Tam[5], Tran Viet Hung[5*]

[1]National Institute of Drug Quality Control, Vietnam
[2]Hanoi University of Pharmacy, Vietnam
[3]National Institute of Medicinal Materials, Vietnam
[4]Agricultural Genetics Institute, Vietnam
[5]Institute of Drug Quality Control, Ho Chi Minh City, Vietnam
*Corresponding author: tran.viethung168@gmail.com

Abstract The species of Morinda is native to China and Vietnam. Among them, *Morinda officinalis* How., is traditionally used as ailments for many diseases and developed into various healthy food for human health. The DNA barcoding is a useful and reliable method for the identification of species, especially the processed material, including *Radix Morindae officinalis*. In this study, the internal transcribed spacer (*ITS*) region of nuclear ribosomal RNA genes and three other chloroplast genes including maturase K (*matK*), ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*), *psbA-trnH* intergenic spacer region (*psbA–trnH*) were sequenced. The results showed that the *matK* and *ITS* genes regions are very similar in the four *M. officinalis* samples, suggesting that they may stem from ordinary roots in a common origin. The *rbcL* and *trnH-psbA* sequences begin to accumulate mutations and on the *trnH-psbA*, as a sample has separated from the group. In this case, the *rbcL* and *trnH-psbA* sequences are useful to analyze the variation under species and geographic impacts. The combination of these genes is suitable for the identification *Baiji tian* at species and sub-species level.

Keywords *Morinda officinalis*, *ITS*, chloroplast, DNA barcoding

Introduction

The *Morinda* genus is estimated over 102 species and is widely distributed tropical, subtropical and temperate areas [1]. Among this genus, *Baji tian*, radix *Morindae officinalis* How., is one of the four traditional Chinese medicinal herbs and availably growing in southern China and in Vietnam. *M. officinalis* is a lianoid shub with stiff, slender bristles and protected with fine soft hair. The branches covered with a tiny leaf-like appendage. The root succulent hypertrophy, irregular, cylindrical and intermittent swelling with a rosary. The flower is merged for the half receptacle, the corolla is white, bell-shaped or urn inside, and hair finely hirsute outside. The aggregate fruit is orbicular to flatten and fully fused with red and subglobose as shown in Fig 1 [1,2]. This plant is traditionally used as an ailment for many diseases such as diabetes, hypertension, impotence, spermatozoa, irregular menstruation, male infertility, cold uterine infertility, cold-induced lower abdominal pain and weak tendon disease. Moreover, some reports have shown that this plant has had great biological activities potential including antioxidant, antitubercular, analgesic, antibacterial, anticancer, anti-inflammatory, cancer-chemopreventive and cardiovascular actions [3-5]. Numerous secondary metabolite compounds which involve in pharmacological activities have been isolated and identified from the roots which belong to the iridoid glycosides, anthraquinones, saccharides, organic acids, volatile oils and homogeneous polysaccharides [2, 3].


Figure 1: Dried roots and panorama of M. officinalis [6]

In Vietnam, *Baji tian*, *M. officinalis* mainly grows wild in the hills and low mountains of the midland and mountainous areas of northern Vietnam provinces such as Quang Ninh, Lang Son, Bac Giang, Cao Bang, Ha Giang and Tuyen Quang and are also distributed in central Vietnam like Quang Nam province. The application of the three-dimensional medicinal identifier and genetic diversity analysis using molecular biology methods in Vietnam has been carried out in a number of previous studies, but only based on the *ITS* gene sequence [7]. At present, over 120 genes of *Morinda officinalis* which are available in the database of NCBI. There are about 90 sequences related to the *ITS* region and 34 sequences on chloroplast including 15 sequences in the *matK* gene, 7 sequences in the genome *rcbL*, 5 trnH, 6 trnL and 1 trnT. Thus, the sequences on chloroplast have not been studied as much as *ITS*. The objective of this paper was to study on species diversity and identify the scientific names of the *Baji tian* samples collected by a geometric method combining gene sequencing method in order to find an appropriate gene region for the establishment of DNA barcoding. The gene regions selected for research are *ITS* on chromosomes and *matK, trnH-psbA, rbcL* on the chloroplast.

Materials and Methods

Materials

A total of 6 samples of *Baji tian* root were collected from the different provinces including Bac Giang, Quang Ninh, Quang Nam, Ha Giang, and numbered from BK1 to BK 6 (BK1: *Baji tian* of Luc Ngan, Bac Giang, BK2: *Baji tian* of Ha Giang, BK3: *Baji tian* Van Don, Quang Ninh, BK 4: *Baji tian* Hoanh Bo, Quang Ninh, BK 5: *Baji tian* Quang Nam, BK6: *Baji tian* Quan Ba, Ha Giang, respectively). In this study, some chemicals, reagents and equipment were used such as Dneasy Plant Miniki Qiagent extraction kit (Germany); RunSafe DNA Loading Dye; TBE buffer (Tris base 10.4 g, boric acid 5.5 g, EDTA 2 ml 1M in 100ml); Plant direct Master Mix 2x, Transilluminator UV gel scanner; Veriti PCR instrument (AB Aplied BioSystems); AIB 3130XL sequencing instrument.

Methods

Total DNA extraction: The sample was cleaned, crushed with a mortar and extracted DNA whole by Dneasy Plant Miniki Qiagent-Germany extraction kit. DNA extraction test: DNA extracted by dyeing of RunSafe DNA Loading Dye, electrophoresis on Agarose gel in TBE buffer (Tris base 10.4 g, boric acid 5.5 g, EDTA 2 ml 1M in 100ml), with DNA ladder 1 kbp / 100 bp and read the results on the transilluminator UV gel scanner. Amplification and DNA sequencing: Amplification of *ITS*1-5,8S-*ITS*2, *matK*, *rbcL* and gene bullets inserted between trnH and psbA with the primers in Table 1.

Table 1*:* The information of the sequence of primers was used in this study

Name	Sequences	Reference	Place
ITS U4- *ITS* P5	RGTTTCTTTTCCTCCGCTTA CCTTATCAYTTAGAGGAAGGAG	[8]	*ITS*1-5,8S-*ITS*2
1326R- 390F	TCT AGC ACA CGA AAG TCG AAG T CGA TCT ATT CAT TCA ATA TTT C	[9]	*matK*
*rbcL*a-F	ATGTCACCACAAACAGAGACTAAAGC	[10]	*rbcL*

*rbcL*a-R	GTAAAATCAAGTCCACCRCG	[10]	
psbA- trnH	CGAAGCTCCATCTACAAATGG ACTGCCTTGATCCACTTGGC	[11]	*trnH-psbA*

The PCR reaction was performed with the following reaction component: Plant direct Master Mix 2x (0.05 U / µL Phire Hot Start II DNA polymerase, reaction buffer: 4 mM $MgCl_2$, 0.4 mM per dNTP (dATP, dCTP, dGTP and dTTP): 10 µl; primer (10 µM): forward primer: 1 µl; reverse primer: 1 µl; DNA template (template) 8-20 ng / µl: 3 µl; sufficient water 23 µl, respectively.

PCR was performed on the Veriti machine of AB Aplied BioSystems with the program set as follows: Starting 5 min at 98^{o}C; follow 30 repetition cycles including 98^{o}C for 30 sec (for *ITS*, *matK*), 15 sec (for *trnH-psbA*, *rbcL*), 55^{o}C (for *ITS*, *rcbL*) for 40 sec (for *ITS*, *matK*) 30 seconds (for *ITS*, *matK*) *rbcL* and *trnH-psbA*), 72^{o}C for 40 sec - 60 sec; finishing with a period of 72^{o}C for 10 min.

Sequencing and searching, evaluating results: Sequencing at Nam Khoa Service and Trading Co., Ltd. (Ho Chi Minh City) by Sanger method on AIB 3130XL.The amplified DNA sequence was assembled by DNA STAR, searched on NCBI (Blast Search on National Center for Biotechnology Information Genbank database), aligned by MEGA 6. Other gene sequences used for comparison, alignment and drawing tree subspecies in the article taken from NCBI.

Results and Discussion

Results of sequencing and searching on NCBI. The sequencing and identification results on the *ITS* gene segment showed that only 4 of the 6 initial samples (BK1, BK3, BK4 and BK5) were *Morinda officinalis*. The other two samples (BK2 and BK6) are one species of the genus Polygala. The sequencing of three genes *matK*, *rbcL* and *trnH-psbA* on chloroplast also confirmed that BK2 and BK 6 were not *Morinda officinalis*.

Results of sequencing and identification of 4 samples BK1, BK3, BK4, BK5 on *ITS* and 3 genes *matK*, *trnH-psbA* and *rbcL* on chloroplast, comparing coverage and similarity with *ITS* segment of *Morinda officinalis* AY 551330 and genes on chloroplast of *Morinda officinalis* KR869730, showed in Table 2.

Table 2: Results of similarity and coverage of the three halves compared with *M. officinalis* KR869730 and *M. officinalis* AY 551330

Sample	**Genes**							
	ITS		**matK**		***rbcL***		**trnH-psbA**	
	Query cover (%)	**Ident (%)**	**Query cover (%)**	**Ident (%)**	**Query cover (%)**	**Ident (%)**	**Query cover (%)**	**Ident (%)**
BK1	98	99.07	100	100.00	89	99.62	93	93.80
BK3	99	99.38	100	100.00	90	99.81	98	99.50
BK4	98	99.42	100	99.66	85	97.7	99	99.75
BK5	98	97.27	99	99.33	100	98.19	98	99.19

The results in Table 2 show that the *ITS* gene region of all four triggers is quite similar and coincides with the *ITS* gene of *M. officinalis* AY 551330 and the sequence on the *matK* gene of all four triggers is similar coincide with *matK* gene of *M. officinalis* KR869730. This proves that the three sizes of BK1, BK3, BK4, and BK5 can be from one origin. The gene regions *rbcL* and *trnH-psbA* are highly variable and can be distinguished by species or geographic influences.

Genetic Diversity Analysis

Gene *ITS*

The alignment results showed that, on *ITS*, all 4 samples have had 1 wrong position (133/520) compared to *M. officinalis* AY 551330. BK 1 has 2 more wrong positions (519/520 and 520 / 520), BK3 has had 1 more position (520/520), BK4 possessed 4 more wrong positions (1/520, 3/520 10/520 and 520/520), respectively, and BK 5 was more than 11 wrong positions compared to *M. officinalis* AY 551330. Thus, if considering the genetic diversity of the *ITS* gene, there were 4 different varieties of the same species.

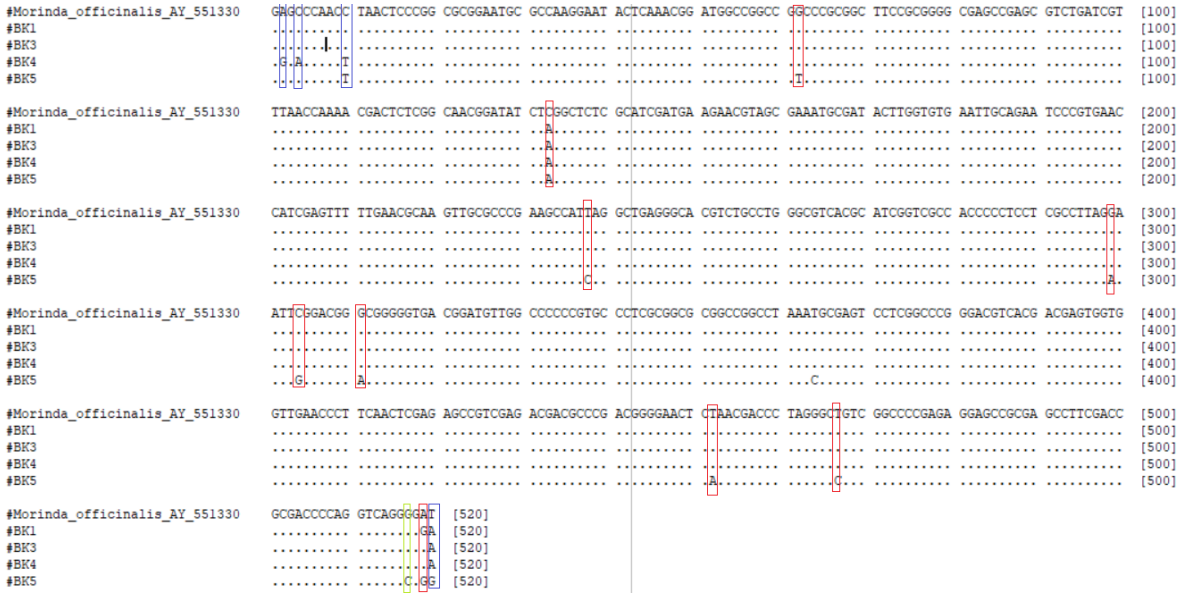

Figure 2: ITS alignment analyses among the M. officinalis samples

Table 3*:* Genetic distances of the samples in *ITS* gene regions

No.	Species	1	2	3	4	5
1	*M. officinalis*_AY_551330					
2	BK1	0.0125				
3	BK3	0.0078	0.0078			
4	BK4	0.0116	0.0115	0.0057		
5	BK5	0.0658	0.0317	0.0286	0.0254	

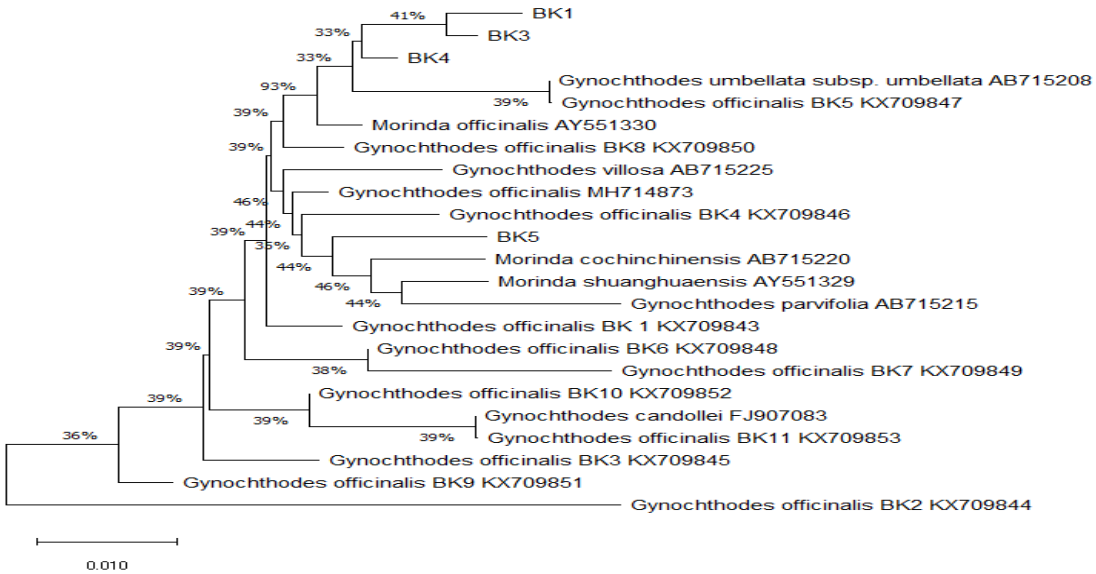

Figure 3: The neighbour-joining tree subspecies on ITS

Tree subspecies were drawn to compare the *ITS* sequence of 11 species of Ba size of Vietnamese origin [7] and some other species stored on NCBI. From subspecies, it can be observed that BK1, BK3 and BK4 have had similar sequences to *M. officinalis* AY 551330 with 93% bootstrap. BK5 and *G. officinalis* BK4 KX 709846 [7] which are close to each other and close to *G. officinalis* MH 714873 but the degree of seed is not large (bootstrap <50%).

Gene *matK*

Figure 4: Results alignment on matK

The alignment results revealed that, on *matK*, BK3, and BK4 genes, the deviation from *M. officinalis* KR869730 was 2/866 nucleotides (BK3: 863/866 and 866/866, BK4: 399/866 and 422/866 positions), BK 5 has had 1 wrong position (81/866), BK 1 closely matched KR869730. Thus, when evaluated on *matK* gene, it can be confirmed that 4 samples BK1, BK3, BK4 and BK5 are different varieties of the same species.

Table 4: Spacing between samples on the *matK*

No.	Species	1	2	3	4	5
1	*Gynochthodes officinalis* KR869730					
2	BK5	0.0023				
3	BK4	0.0035	0.0058			
4	BK3	0.0023	0.0046	0.0058		
5	BK1	0.0000	0.0023	0.0035	0.0023	

Species distance also shows that BK1 completely coincides with *G. officinalis* KR869730 on *matK*.

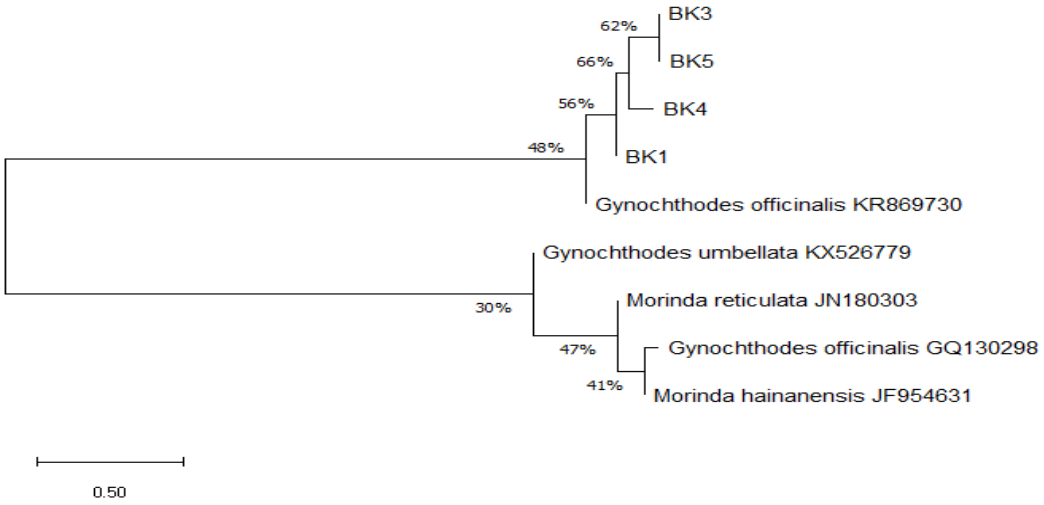

Figure 5: The neighbour-joining tree on the matK

The subspecies show all four specimens collected in Vietnam in the same branch as *G. officinalis* KR869730 on *matK*, but the bootstrap between samples is quite small, the largest being about 66%, indicating that these varieties have differences in genetic.

Gene *rbcL*

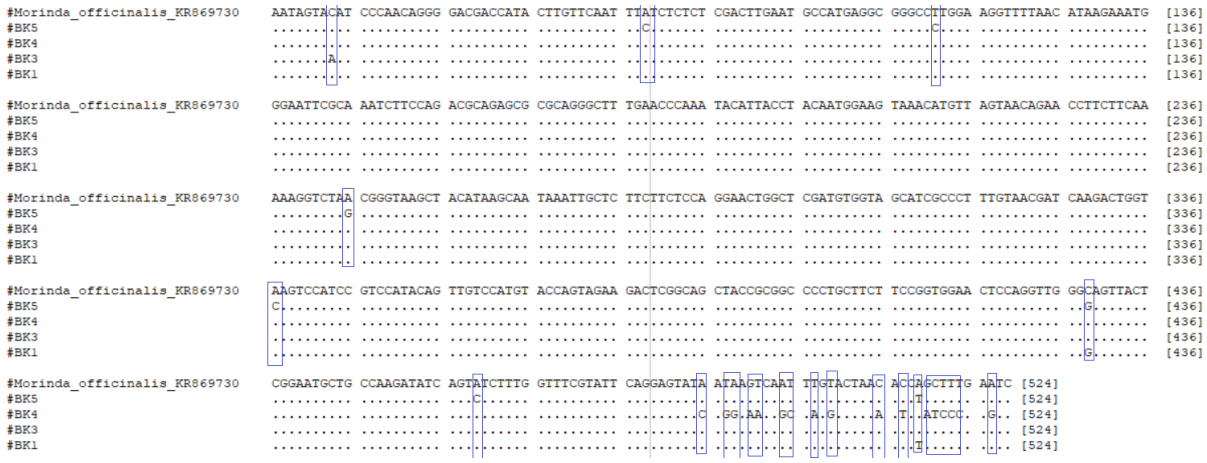

Figure 6: Results of alignment on rbcL

The alignment results on the *rbcL* gene showed a large difference between BK5 (7/524 nucleotides) and BK4 (17/524 nucleotides) compared with *G. officinalis* KR869730.

Table 5: Spacing between samples on *rbcL*

No.	Species	1	2	3	4	5
1	*Morinda officinalis* KR869730					
2	BK5	0.0145				
3	BK4	0.0376	0.0749			
4	BK3	0.0019	0.0607	0.0703		
5	BK1	0.0038	0.0562	0.0742	0.0071	

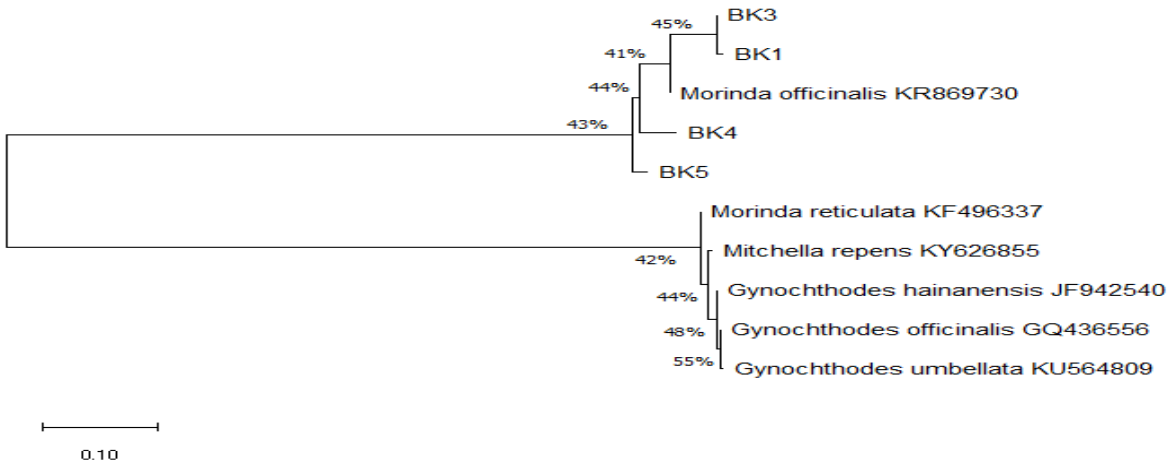

Figure 7: Neighbour-joining tree in rbcL

On the tree, all four samples of the jackets still belong to the branch with *G. officinalis* KR869730, although BK4 and BK5 have changed positions. The *rbcL* gene may have a mutation from some original root. The bootstrap in this branch is about 45% largest, indicating that these individuals have a big difference in *rbcL*. This is probably due to the large cumulative mutations in this region.

Gene *trnH-psbA*

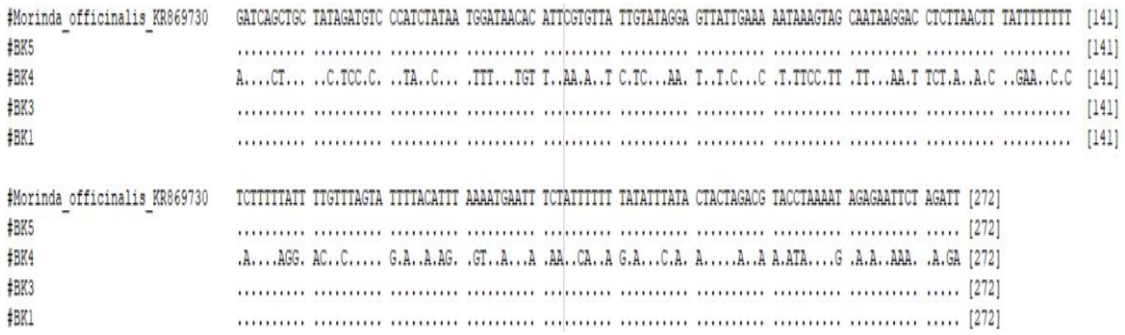

Figure 8: Results of alignment on trnH-psbA

This gene segment only has about 300 nucleotides but there is a big difference between BK4 *G. officinalis* KR869730 and samples BK1, BK3, BK5 (with a genetic distance of 1.3173; 1.4444; 1.3740; and 0.9738, respectively.

Table 6*:* Spaces between strains on *trnH-psbA*

No.	Species	1	2	3	4	5
1	*Morinda officinalis* KR869730					
2	BK5	0.0041				
3	BK4	1.3173	0.9738			
4	BK3	0.0053	0.0122	1.3740		
5	BK1	0.0058	0.0047	1.4444	0.0143	

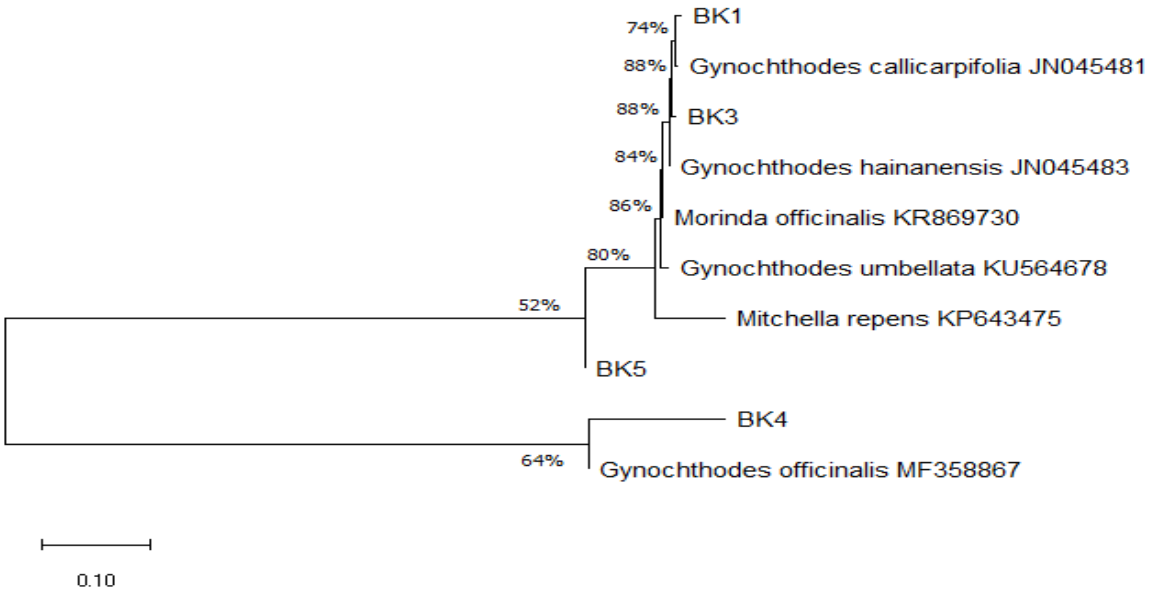

Figure 9: The neighbour-joining tree subspecies on trnH-psbA

On a sub-tree based on *trnH-psbA* gene, BK4 was completely separated from *G. officinalis* KR869730 and grouped with *G. officinalis* MF 358867 with 64% bootstrap. The results also suggest that the gene is highly variable and unstable at the species level, which is suitable for subspecies evaluation or for further understanding the influence of geography. In our previous studies, we have been successfully identified genetic diversity and genetic relationship of some medicinal plants by using molecular markers including *ITS*, *mat*K, *rbcL* and *trnH-psbA* [12-16].

Conclusions

By use of the *ITS* region which is possible to identify *M. officinalis* and its genetic relationship. Chloroplast genes such as *matK* have the potential to make pharmaceutical identification barcode because chloroplast is small in size and less likely to be damaged as gene regions in chromosomes. The results showed that the *matK* and *ITS* genes regions are very similar in the four *M. officinalis* samples, suggesting that they may stem from ordinary roots in a common origin. The *rbcL* and *trnH-psbA* sequences begin to accumulate mutations and on the *trnH-psbA*, as a sample has separated from the group. In this case, the *rbcL* and *trnH-psbA* sequences are useful to analyze the variation under species and geographic impacts. The combination of these genes is suitable for the identification *Baiji tian* at species and sub-species level.

References

[1]. Editorial Committee of Flora of China, Chinese Academy of Sciences (2015). Flora of China. Science Press, Beijing, pp. 35-37.

[2]. Olatunde, O.Z., Yang, Y., Yong, J., and Lu, C. (2018). Progress of the components and biological activities of *Morinda officinalis* How. Biolmed. Res. Rev, 2(3): 1-10.

[3]. Zhang, J.H., Xin, H.L., Xu, Y.M., Shen, Y., He, Y.Q., Yeh, H., Lin, B., Song, H.T., Liu, J., Yang, H.Y., Qin, L.P., Zhang, Q.Y. (2018). *Morinda officinalis* How- A comprehensive review of tradition uses phytochemistry and pharmacology. J.Ethnopharma, 213: 230-255.

[4]. Saludes, J.P., Garson, M.J., Franzblau, S.G. (2002). Antitubercular constituents from the hexane fraction of *Morinda eitrifolia* Linn (Rubiaeeae). Phytother. Res. 16, 683–685.

[5]. Yanine, C.B., Vaillant, F., Perez, A.M., et al., 2006. The noni fruit *(Morinda citrifolia* L.): a review of agricultural research, nutritionaland therapeutic properties. Food Compos. Anal. 19, 643–645.

[6]. Internet link: https://lahien.com/san-pham/ba-kich-tim/

[7]. Hien, N.T.T (2015). Research on genetic diversity of *Morinda officinalis* How. In Northern Vietnam, Master Thesis, Hanoi University of Pharmacy

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

[9]. Philippe, C., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J., Chase, M.W. (2002). Molecular phylogenetics of caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, atpB, and matK DNA sequences. Am. J. Bot, 89(1): 132–144. 2002.

[10]. Kress J, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding trnHpsbA spacer region. PLoS One, 6:1-10.

[11]. Levin, R.A., Wagner, W.L., Hoch, P.C, et al. (2003) Family-level relationships of Onagraceae based on chloroplast *rbcL* and ndhF data. Am.J.Bot, 90:107-115.

[12]. Dong, D.V., Ngoan, M.T., Diep, P.N., Khoa, N.T., Trung, K.H., Khanh, T.D. (2019). Genetic diversity and genetic relationship of Camellia plants collected in Van Giang district, Hung Yen province, Vietnam. J. Sci. Eng Res, 6(10): 16-23.

[13]. Tam, T.V., Dong, D.V., Tin, N.V et al (2019). Genetic diversity and identification of the local podocarpus samples by molecular markers at mountainous areas of Yen Tu, Quang Ninh province, Vietnam. J. Sci. Eng Res, 6(10): 24-32.

[14]. Trang, D.H., Dong, D.V., Chung, B.H., Gioi, D.H., Khanh, T.D. (2018). Identification of Vietnamese Ochna integerrima (lour.) Merr species based on ribosomal DNA internal transcribed space sequence. Inter. Let. Nat. Sci, 68:9-18.

[15]. Tam, T.V., Dong, D.V., Lam, D.V., Anh, T.H., Van, N.D., Trung, D.M., Duong, T.D., Trung, K.H., Duong, N.T., Khanh, T.D. (2018). Identification of Vietnamese native Rhododendron and Enkianthus species of Ericaceae family based on ribosomal DNA internal transcribed spacer (ITS) sequences. J. Horticul. Plant Res, 2:33-42.