

## Article

# Phylogenetic Analysis Based on DNA Barcoding and Genetic Diversity Assessment of *Morinda officinalis* How in Vietnam Inferred by Microsatellites

Thanh Pham <sup>1,\*</sup> , Quynh Thi Nguyen <sup>1</sup> , Duc Minh Tran <sup>2</sup>, Hoi Nguyen <sup>2</sup>, Hung Thai Le <sup>2</sup>,  
Que Thi Hong Hoang <sup>2</sup>, Yen Thi Van <sup>2</sup> and Thang Nam Tran <sup>2,\*</sup> 

<sup>1</sup> Department of Biology, University of Education, Hue University, 34 Le Loi, Hue 530000, Vietnam

<sup>2</sup> Faculty of Forestry, University of Agriculture and Forestry, Hue University, 102 Phung Hung, Hue 530000, Vietnam

\* Correspondence: phamthanh@hueuni.edu.vn (T.P.); tnthang@hueuni.edu.vn (T.N.T.)

**Abstract:** *Morinda officinalis* How is well-known as a valuable medicinal plant found in some regions of Vietnam. This species is mainly used for treating male impotence, irregular menstruation, and rheumatoid arthritis. This study aimed to identify the species of and genetic diversity in three *M. officinalis* populations: one each in Quang Binh (QB), Thua Thien Hue (TTH), and Quang Nam (QN). In this study, four DNA barcoding markers (*ITS1*, *ITS2*, *matK*, and *rbcL*) were used to identify the species and 22 microsatellite markers were applied for population structure and diversity analyses. The results showed that the sequences of gene regions studied in *M. officinalis* had a high similarity (>95%) to the *ITS1*, *ITS2*, *matK*, and *rbcL* sequences of *M. officinalis* on BLAST. Of the four DNA barcoding markers used, *ITS1* and *ITS2* showed higher efficiency in DNA amplification of *M. officinalis*. From this study, 27 GenBank codes were published on BLAST. The results also revealed high levels of genetic diversity in populations. The average observed and expected heterozygosity values were  $H_O = 0.513$  and  $H_E = 0.612$ , respectively. The average  $F_{ST}$  value was 0.206. Analysis of molecular variance (AMOVA) showed 70% variation within populations and 30% among populations. The population structure of *M. officinalis* inferred in STRUCTURE revealed that the optimum number of genetic groups for the admixture model was  $K = 2$ . These findings provided vital background information for future studies in the conservation of *M. officinalis* in both ex situ and in situ plans.

**Keywords:** conservation; genetic variability; *M. officinalis*; population structure



**Citation:** Pham, T.; Nguyen, Q.T.; Tran, D.M.; Nguyen, H.; Le, H.T.; Hoang, Q.T.H.; Van, Y.T.; Tran, T.N. Phylogenetic Analysis Based on DNA Barcoding and Genetic Diversity Assessment of *Morinda officinalis* How in Vietnam Inferred by Microsatellites. *Genes* **2022**, *13*, 1938. <https://doi.org/10.3390/genes13111938>

Academic Editors: Wajid Zaman and Hakim Manghwar

Received: 29 September 2022

Accepted: 21 October 2022

Published: 25 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

*Morinda officinalis* is a perennial vine mainly distributed in tropical and subtropical regions [1]. In Vietnam, *M. officinalis* is found in the wild in provinces such as Cao Bang, Lao Cai, Ha Giang, Quang Binh, Thua Thien Hue, and Quang Tri [2–4].

*M. officinalis* contains various bioactive components and has been used for decades as a tonic and an antirheumatic medicinal herb in some Asian countries [2,5,6]. The root of *M. officinalis* has long been used as a tonic or nutrient supplement for alleviating diseases such as depression, Alzheimer's disease, impotence, osteoporosis, and rheumatoid arthritis [6].

In Vietnam, due to the rapid increase in the demand for medicinal herbs, this plant has been over-exploited, leading to the depletion of raw materials. Additionally, because *M. officinalis* has slow growth and poor regeneration, its natural population has significantly shrunk and become endangered [2,4]. With the natural population reduced by at least 50%, this species has recently been classified as an endangered precious medicinal plant species [7]. Therefore, it is necessary to study the genetic diversity and structure of the natural population in order to conserve it effectively.

Nowadays, phylogenetic studies are necessary to conserve rare medicinal plants. In addition to representing the relationships among species in the tree of life, phylogenetic studies

provide a framework for interdisciplinary investigations in taxonomy, evolutionary biology, biogeography, ecology, and conservation [8]. More recently, phylogenetic approaches based on molecular data have also proven to be an indispensable tool for genome comparisons. These approaches are used to identify genes and regulatory elements, interpret modern and ancient individual genomes, and reconstruct ancestral genomes [9], providing conservationists with background information to make conservation policies efficiently.

DNA barcoding is a universally used and reliable method of identifying plant species and has become a major focus in the fields of biodiversity and conservation. This molecular technique is not influenced by external factors or development stage, and DNA can be easily isolated from all tissues, providing an important basis for species identification at the genetic level [10]. Recently, *rbcL* and *matK* plastid coding genes were recommended as barcodes for plant species and have become the most used markers in flowering plants, as the *rbcL* region is highly suitable for amplification and sequencing [10,11]. Meanwhile, the nuclear *ITS* region includes both the *ITS1* and *ITS2* regions, with relatively strong discrimination power, that serve as complementary barcodes to *matK* and *rbcL* in plants. The four markers (*ITS1*, *ITS2*, *MatK*, and *rbcL*) used in this study have also proven effective in identifying medicinal species [12–14].

Currently, molecular markers help detect variations or polymorphisms that exist among individuals in the population for specific regions of DNA [15]. Among commonly used molecular markers such as AFLP (amplified fragment length polymorphism), RAPD (random amplified polymorphic), SSR (simple sequence repeat), and ISSR (inter simple sequence repeat), simple sequence repeat (SSR) markers are useful tools for research in plant genetics, breeding, and identification of individuals and species due to the allelic sequence diversity. SSRs are widely spread in the genome and have high codominant inheritance, polymorphism, and multiallelic variation [16–19].

Studies on molecular markers have been carried out on *M. officinalis* [5,20]. Liao et al. in particular developed an SSR marker dataset to serve in further research related to this plant.

The SSR marker has the advantage of using only a tiny amount of DNA. In addition, this method involves a more straightforward, faster, and more cost-effective technique than other methods.

In this study, before evaluating the genetic diversity and population structure of *M. officinalis* by using SSR markers, we used DNA barcoding to identify this species in three different provinces in central Vietnam (Quang Binh, Thua Thien Hue, and Quang Nam). The current study aimed to pave the way for protecting wild *M. officinalis* populations.

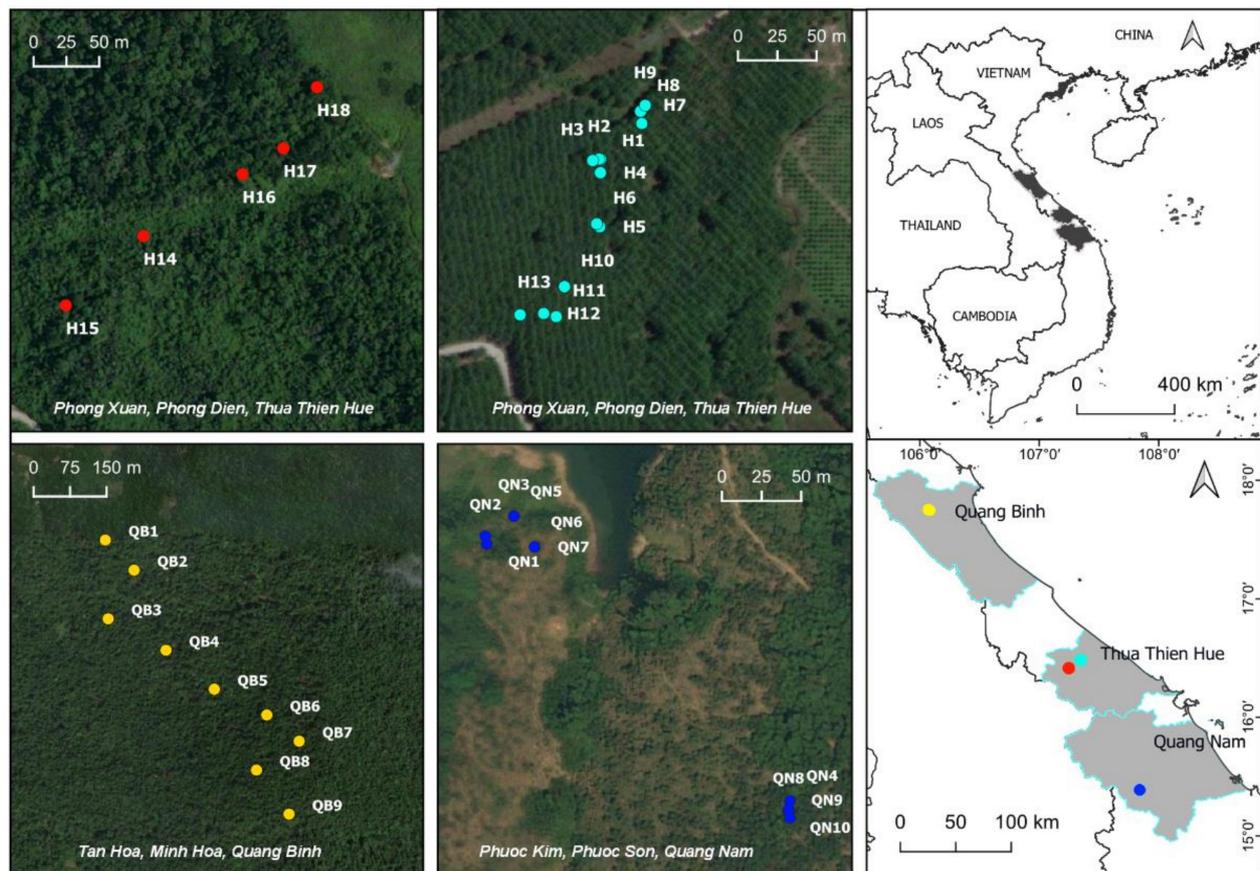
## 2. Materials and Methods

### 2.1. Plant Materials

The leaves of 37 *M. officinalis* trees were randomly collected from 3 populations in the central provinces of Quang Binh, Thua Thien Hue, and Quang Nam in Vietnam. In the field, the samples were placed in plastic bags containing silica gel. Next, the samples were transferred to the Genetic Laboratory of the Biology Department at Hue University of Education and stored at  $-20\text{ }^{\circ}\text{C}$  until DNA extraction. Sampling locations were recorded using a global positioning system (GPS) (Figure 1, Table S1).

### 2.2. DNA Extraction

Genomic DNA was extracted from 100 mg of young leaf samples using a Plant Genomic DNA Extraction Kit (Bioteke Corporation, Beijing, China) according to the manufacturer's protocol. The quality of the extracted DNA was further estimated via 0.8% agarose gel electrophoresis. Safe dye (Phu Sa Biochem, Can Tho, Vietnam) was applied for DNA gel stain. Gel imaging was performed using a UV gel imaging system (Major Science, Saratoga, CA, USA). The isolated DNA was then stored at  $-20\text{ }^{\circ}\text{C}$  until further analysis.



**Figure 1.** Map showing collection sites of three populations of *M. officinalis* in central Vietnam.

### 2.3. DNA Barcoding Amplification and Sequencing

A polymerase chain reaction (PCR) was performed using standard universal plant DNA barcoding primers (Table 1), with 25  $\mu$ L of the reaction mixture containing 2.0  $\mu$ L of template DNA (50 ng), 12.5  $\mu$ L of 2X Taq Master Mix, 0.5  $\mu$ L of each primer (10 pmol) (Table 2), and 9.5  $\mu$ L of deionized water. PCR amplification was carried out with an Aeris™ PCR Aeris Thermal Cycler (21 Changi South Street 1, Singapore). The PCR conditions followed those laid out in the previous publication [8].

**Table 1.** The universal primers for DNA barcoding used in this study.

Locus	Primer Name	Sequences (5'–3')	Reference
<i>ITS1</i>	<i>5a fwd</i>	CCTTATCATTAGAGCAAGGAG	[12]
	<i>4 rev</i>	TCCTCCGCTTATTGATATGC	
<i>ITS2</i>	<i>S2F</i>	ATGCGATACTTGGTGTGAAT	
	<i>S3R</i>	GACGCTTCTCCAGACTACAAT	
<i>rbcL</i>	<i>1f</i>	ATGTCACCACAAACAGAAAC	
	<i>724r</i>	TCCGATGTACCTGCAGTAGC	
<i>matK</i>	<i>390F</i>	CGATCTATTCAATCAATATTC	
	<i>1326R</i>	TCTAGCACACGAAAGTCGAAGT	

**Table 2.** Efficiency of PCR amplification of four potential barcodes in three populations of *M. officinalis*.

Population Codes	Number of Amplified Samples	PCR Efficiency of <i>ITS1</i> (%)	PCR Efficiency of <i>ITS2</i> (%)	PCR Efficiency of <i>MatK</i> (%)	PCR Efficiency of <i>rbcL</i> (%)
QB	3	100.0	100.0	0.0	33.3
TTH	3	100.0	100.0	100.0	66.5
QN	3	100.0	100.0	33.3	66.5
Total	9	100	100	44.44	55.56

To check the presence or absence of bands, amplified PCR products were electrophoresed using 0.8% agarose gel (1 × TAE buffer and 5 µL/mL safe dye). Gel imaging was carried out using a UV gel imaging system (Major Science, Saratoga, CA, USA). Band size of amplified products was determined using a Thermo Scientific GeneRuler 100 bp DNA Ladder (Thermo Fisher Scientific, Waltham, MA, USA).

The PCR products were sent to First Base Laboratories Sdn. Bhd (Taman Serdang Perdana, Seri Kembangan, Selangor, Malaysia) for purification and sequencing service, using the same primers as those used for the PCR.

#### 2.4. Microsatellite Amplification

In all, 37 genomic DNA samples were used in this study. A polymerase chain reaction (PCR) was performed in 25 µL of reaction mixture that contained 2.0 µL of template DNA, 12.5 µL of 2X Taq Master Mix, 0.5 µL of each primer, and 9.5 µL of deionized water.

Liao et al. [20] have described the 22 microsatellite loci used to generate data for the current study (Table S2).

PCR amplification was carried out with an Aeris™ PCR Aeris Thermal Cycler as follows: Initial denaturation was carried out at 95 °C for 3 min. This was followed by 35 cycles of 45 s each at 94 °C for denaturation, 45 s of alignment at the annealing temperature (50–52 °C) for each primer pair, and 45 s of alignment at 72 °C for extension. Finally, 10 min of alignment at 72 °C for the final cycle completed the extension of any remaining products. The samples were kept at 4 °C until they were analyzed.

The amplification products were separated on 8% polyacrylamide gels in 1 × TAE buffer using the Mini Vertical Gel Electrophoresis Apparatus (Major Science, Saratoga, CA, USA) and visualized with safe dye (Phu Sa company). A UV gel imaging system (Major Science, Saratoga, CA, USA) was used to verify the presence of amplified fragments.

#### 2.5. Data Analysis

##### 2.5.1. DNA Barcoding

To obtain the sequence of each region (*ITS1*, *ITS2*, *matK*, and *rbcL*), the forward and reverse sequences were aligned using BioEdit version 7.2.5 software [21]. In searching for the similarities between those sequences and the sequences deposited in the GenBank database, the sequences of this study were analyzed using the BLAST (Basic Local Alignment Search Tool) program at <http://www.ncbi.nlm.nih.gov/BLAST> (accessed on 20 June 2022) [22] (Table S3).

The alignment was then exported to Molecular Evolutionary Genetics Analysis (MEGA-X) software for phylogenetic analysis [23]. The maximum likelihood trees were constructed for *ITS1*, *ITS2*, *matK*, and *rbcL* data separately using the Kimura 2-parameter model [24] with 1000 bootstrap replicates for node supports.

##### 2.5.2. Genetic Diversity

Size of bands was detected by GelAnalyzer 19.1 software ([www.gelanalyzer.com](http://www.gelanalyzer.com)) (accessed on 4 January 2022), by Istvan Lazar Jr. and Istvan Lazar Sr).

To determine the level of genetic variation within a population, the following genetic diversity parameters were calculated: mean number of alleles per locus (A), number of unique alleles, the effective number of alleles (Ne), average observed heterozygosity (Ho),

average expected heterozygosity ( $H_e$ ), and fixation index ( $F_{IS}$ ). All calculations were performed in GenAIEx v.6.5 [25].

A genetic distance matrix of pairwise  $F_{ST}$  values was also used to perform a hierarchical analysis of molecular variance (AMOVA) in GenAIEx v.6.5 [25]. Significance levels were determined using 999 permutations. AMOVA was used to estimate and partition the total variances at two hierarchy levels: within populations and among populations.

Principal coordinate analysis (PCoA) based on the codominant genotypic distance for 37 studied samples of 3 populations of *M. officinalis* was carried out in GenAIEx [25].

To determine the optimal value of the genetic clusters (K), a Bayesian analysis of the population structure was performed with STRUCTURE v.2.3.4 (<https://web.stanford.edu/group/pritchardlab/structure.html>) (accessed on 28 July 2022) [26]. Once the admixture model was set with a correlated allele frequency and ancestry models, 10 separate runs of the number of clusters (K) in the dataset were performed from 1 to 10 for each K value at 500,000 Markov Chain Monte Carlo (MCMC) repetitions and a 100,000 burn-in period. The optimal value of K was detected using Structure Harvester [27] based on the  $\Delta K$  value by Evanno et al. [28].

### 3. Results

#### 3.1. DNA Barcoding

For this study's DNA barcoding amplification process, we selected three samples per population, using four different primers to identify species. The results showed that *ITS1* and *ITS2* markers gave fragments of molecular weight as expected and bands of PCR products were clear for further research in all samples of the three populations, whereas some samples that used *matK* and *rbcL* markers for amplification had negative results and therefore their sequences were not determined.

A total of 27 assembled sequences were obtained and deposited in the GenBank: nine assembled sequences of *ITS1* amplicon, nine assembled sequences of *ITS2* amplicon, four assembled sequences of *matK* amplicon, and five assembled sequences of *rbcL* amplicon. The BLAST tool was used to identify the plant for each sequence, and the closest species in the GenBank was obtained. The results showed that all markers used and plants deposited in this study belonged to the genus *Morinda*. *Gynochthodes officinalis* (synonym of *M. officinalis*) was the most closely related species, with a higher extent than other species of the *Morinda* genus.

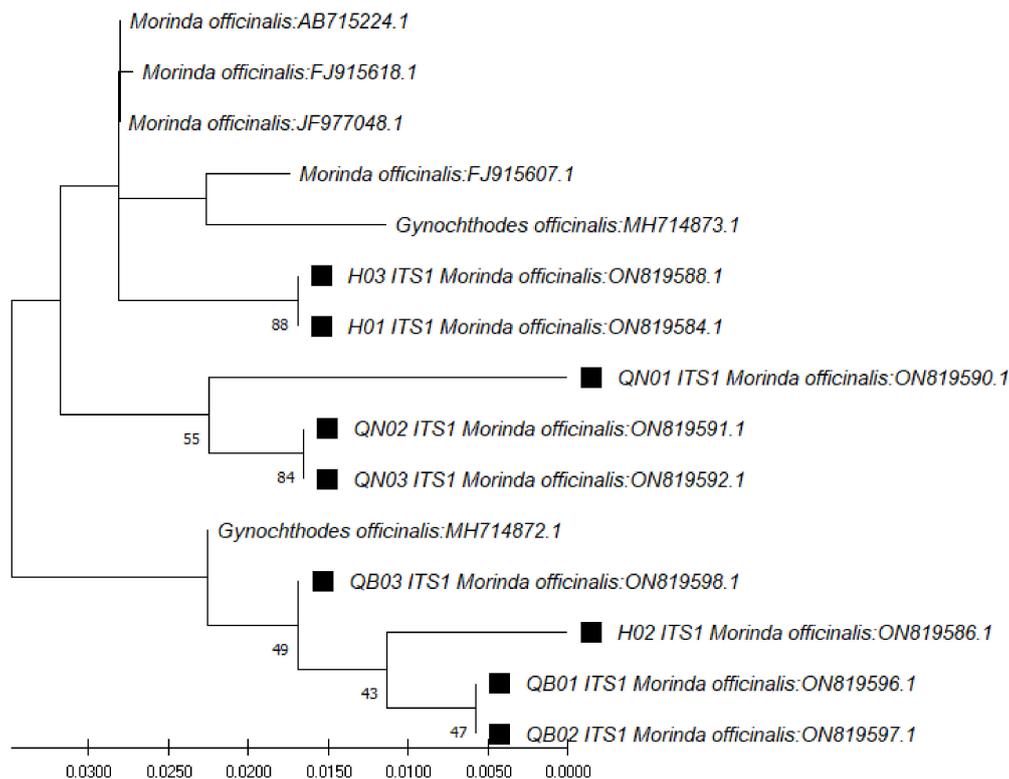
In our study, *ITS1* and *ITS2* were the most effective in terms of the amplification process (100%) (Table 2).

In the case of *ITS1* region sequences (~179 bp), through use of the *ITS1* pair of primers (*ITS 5a* forward and 4 reverse) and the BLAST tool for identification, the closest species was found to be *M. officinalis*, at 98.88% similarity (Table 3). The phylogenetic tree for those sequences was constructed with the closest species (Figure 2). In the case of the *ITS2* region sequences (~235 bp), through use of the *ITS2* pair of primers (*ITS S2F* and *ITS S3R*), the closest species was found to be *Gynochthodes officinalis* (synonym of *M. officinalis*), at 100% similarity (Table 4). The phylogenetic tree for those sequences was constructed with the closest species (Figure 3).

**Table 3.** *M. officinalis* plants in this study and the percentage of similarity to the closest species in the GenBank as per *ITS1* gene sequence.

No.	Species Description	Scientific Name	Aligned <i>ITS1</i> Gene Sequence (bp)	Coverage (%) E	Similarity (%)	Accession Numbers
1	H01- <i>ITS1-M.officinalis</i>	<i>M. officinalis</i>	179 (86bp–264bp)	100	100	ON819584.1 *
2	H03- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (92bp–270bp)	100	100	ON819588.1 *
3	<i>Morinda officinalis</i>	<i>M. officinalis</i>	179 (93bp–271bp)	100	98.88	AB715224.1
4	<i>Morinda officinalis</i> h15	<i>M. officinalis</i>	179 (11bp–189bp)	100	98.88	FJ915618.1
5	<i>Morinda officinalis</i> h4	<i>M. officinalis</i>	179 (11bp–189bp)	100	98.32	FJ915607.1
6	<i>Morinda officinalis</i> voucher	<i>M. officinalis</i>	179 (19bp–197bp)	100	97.21	JF977048.1
7	<i>Gynochthodes officinalis</i>	<i>M. officinalis</i>	179 (73bp–251bp)	100	97.21	MH714873.1
8	QB01- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (85bp–263bp)	100	97.21	ON819596.1 *
9	QN03- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (83bp–261bp)	100	97.21	ON819592.1 *
10	QN02- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (83bp–261bp)	100	97.21	ON819591.1 *
11	QB03- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (86bp–264bp)	100	96.65	ON819598.1 *
12	QB02- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (85bp–263bp)	100	96.65	ON819597.1 *
13	H02- <i>ITS1-M. officinalis</i>	<i>M. officinalis</i>	179 (83bp–261bp)	100	95.53	ON819586.1 *
14	<i>Gynochthodes officinalis</i>	<i>M. officinalis</i>	179 (73bp–251pb)	100	97.21	MH714872.1

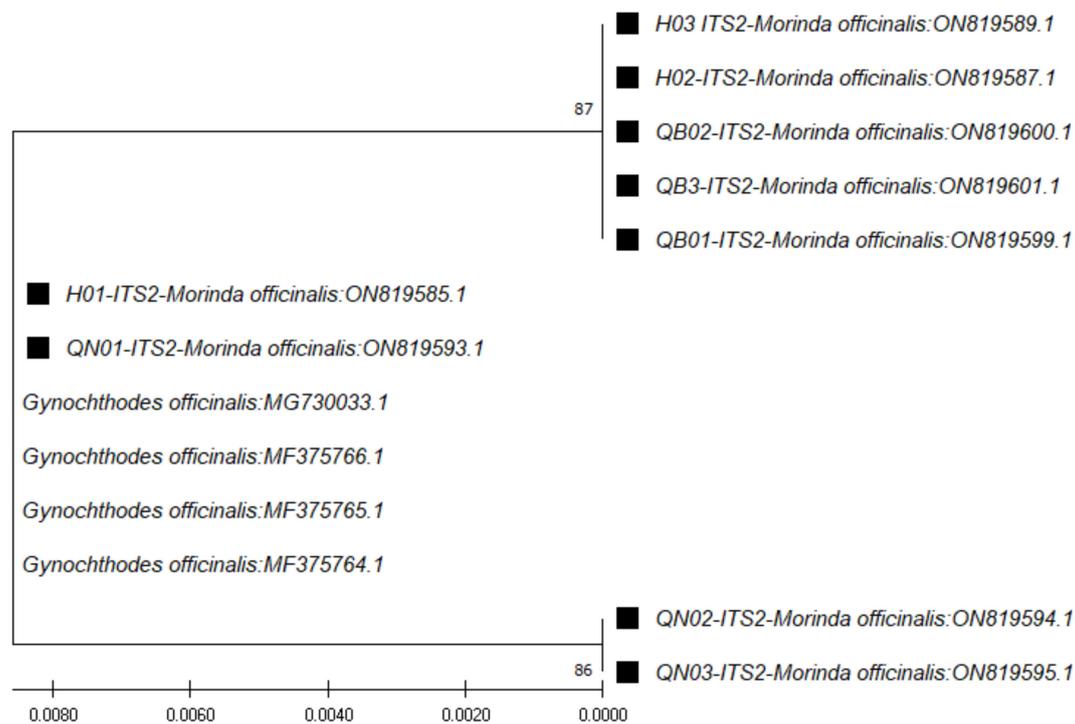
Note: "\*" is the sample in this study.

**Figure 2.** Phylogenetic tree by the maximum likelihood of similarity between *M. officinalis* and the closest species as per *ITS1* region sequences. The samples of our study are marked by black squares next to their names.

**Table 4.** *M. officinalis* plants in this study and the percentage of similarity to the closest species in the GenBank as per *ITS2* gene sequences.

No.	Species Description	Scientific Name	Aligned <i>ITS2</i> Gene Sequence (bp)	Coverage (%) E	Similarity (%)	Accession Numbers
1	H01- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (108bp–342bp)	100	100	ON819585.1 *
2	<i>Gynochthodes officinalis</i>	<i>M. officinalis</i>	235 (53bp–287bp)	100	100	MG730033.1
3	<i>Gynochthodes officinalis</i> B7	<i>M. officinalis</i>	235 (102bp–336bp)	100	100	MF375766.1
4	<i>Gynochthodes officinalis</i> B5	<i>M. officinalis</i>	235 (102bp–336bp)	100	100	MF375765.1
5	<i>Gynochthodes officinalis</i> B3	<i>M. officinalis</i>	235 (105bp–339bp)	100	100	MF375764.1
6	QB01- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (108bp–342bp)	100	99.15	ON819599.1 *
7	QN02- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (106bp–340bp)	100	99.15	ON819594.1 *
8	QB03- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (108bp–342bp)	100	99.15	ON819601.1 *
9	QB02- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (108bp–342bp)	100	99.15	ON819600.1 *
10	QN01- <i>ITS2-M.officinalis</i>	<i>G. officinalis</i>	235 (110bp–344bp)	100	100	ON819593.1 *
11	QN03- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (106bp–340bp)	100	99.15	ON819595.1 *
12	H03- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (75bp–309bp)	100	99.15	ON819589.1 *
13	H02- <i>ITS2-M.officinalis</i>	<i>M. officinalis</i>	235 (71bp–305bp)	100	99.15	ON819587.1 *

Note: "\*" is the sample in this study.

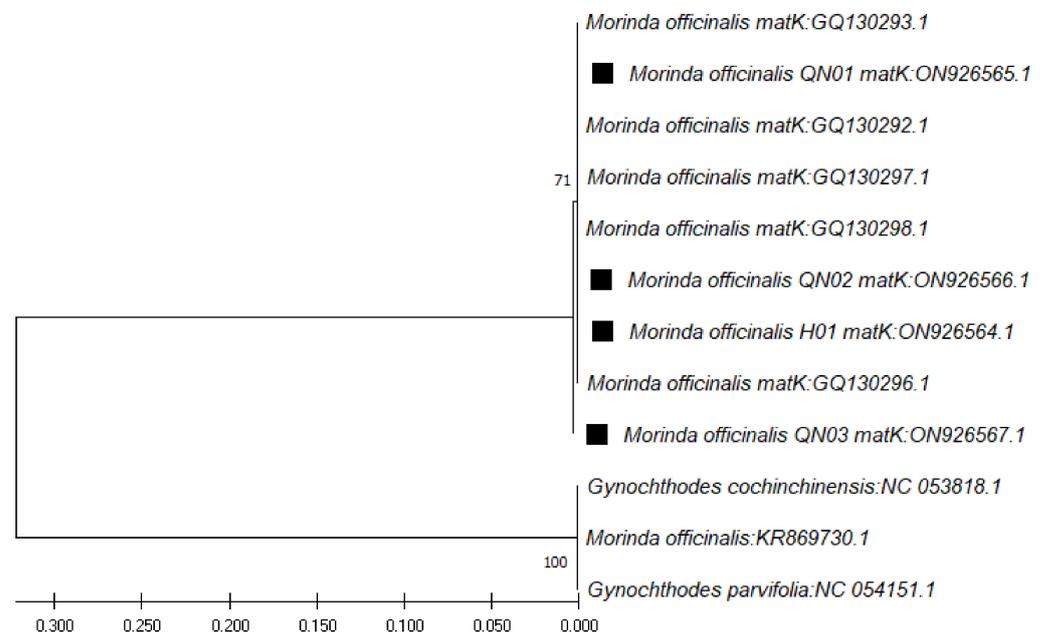
**Figure 3.** Phylogenetic tree by the maximum likelihood of similarity between *M. officinalis* and the closest species as per *ITS2* region sequences. The samples of our study are marked by black squares next to their names.

In the case of the *matK* region sequences (ranging from 861 bp to 899 bp), through use of the *matK* pair of primers (390F and 1326R), the closest species was found to be *M. officinalis*, at 99.89% similarity (Table 5), followed by *Gynochthodes parvifolia*, at 99.78% similarity. The phylogenetic tree for these sequences was constructed with the closest species (Figure 4).

**Table 5.** *M. officinalis* plants in this study and the percentage of similarity with the closest species in the GenBank as per *matK* gene sequences.

No.	Species Description	Scientific Name	Aligned Sequence (bp)	Coverage (%) E	Similarity (%)	Accession Numbers
1	<i>Gynochthodes officinalis</i> _H01_ <i>matK</i>	<i>M. officinalis</i>	899	100	100	ON926564.1 *
2	<i>Morinda officinalis</i>	<i>M. officinalis</i>	897	99.00	99.89	KR869730.1
3	<i>Gynochthodes officinalis</i> _QN02_ <i>matK</i>	<i>M. officinalis</i>	897	99.00	99.89	ON926566.1 *
4	<i>Gynochthodes parvifolia</i>	<i>G. parvifolia</i>	897	99.00	99.78	NC_054151.1
5	<i>Gynochthodes cochinchinensis</i>	<i>M. officinalis</i>	897	99.00	99.78	NC_053818.1
6	<i>Gynochthodes officinalis</i> _QN03_ <i>matK</i>	<i>M. officinalis</i>	897	99.00	99.67	ON926567.1 *
7	<i>Morinda officinalis</i> isolate h7_ <i>matK</i>	<i>M. officinalis</i>	864	96.00	99.88	GQ130298.1 *
8	<i>Morinda officinalis</i> h2_ <i>matK</i>	<i>M. officinalis</i>	864	96.00	99.88	GQ130293.1
9	<i>Morinda officinalis</i> h1_ <i>matK</i>	<i>M. officinalis</i>	864	96.00	99.88	GQ130292.1
10	<i>Morinda officinalis</i> h6_ <i>matK</i>	<i>M. officinalis</i>	864	96.00	99.77	GQ130297.1
11	<i>Morinda officinalis</i> h5	<i>M. officinalis</i>	864	96.00	99.77	GQ130296.1
12	<i>Gynochthodes officinalis</i> _QN01_ <i>matK</i>	<i>M. officinalis</i>	861	95.00	99.77	ON926565.1 *

Note: "\*" is the sample in this study.

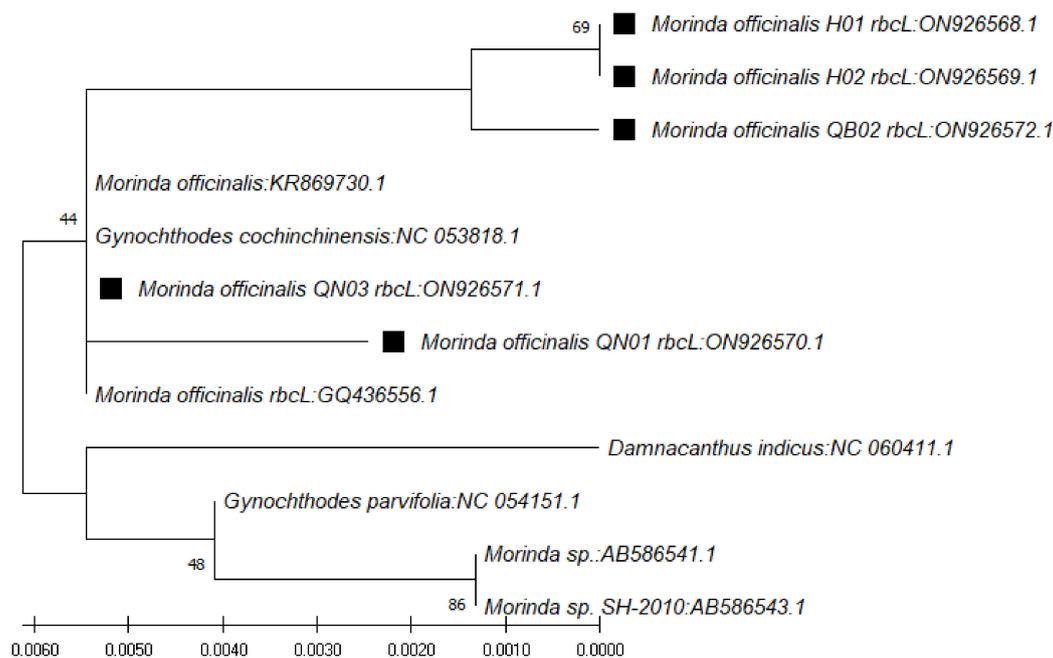
**Figure 4.** Phylogenetic tree by maximum likelihood of similarity between *M. officinalis* and the closest species as per *matK* region sequences. The samples of our study are marked by black squares next to their names.

In use of the *rbcl* gene, the BLAST result on NCBI showed that the nucleotide sequences obtained were highly similar to those of species of *Morinda officinalis* (accession numbers KR869730.1, NC\_053818.1, and GQ436556.1), *G. parvifolia* (accession number NC\_054151.1), and *Morinda* sp. SH-2010 (accession numbers AB586541.1 and AB586543.1), with similarity ranging from 99.31% to 100% (Table 6). The phylogenetic tree for those sequences was constructed with the closest species (Figure 5).

**Table 6.** *M. officinalis* plants in this study and percentage of similarity with the closest species in the GenBank base as per *rbcL* gene sequences.

No.	Species Description	Scientific Name	Aligned Sequence (bp)	Coverage (%) E	Similarity (%)	Accession Numbers
1	<i>Gynochthodes officinalis</i> _H01_ <i>rbcL</i>	<i>M. officinalis</i>	738	100.00	100	ON926568.1 *
2	<i>Gynochthodes officinalis</i> _QB02_ <i>rbcL</i>	<i>M. officinalis</i>	738	99.00	99.73	ON926572.1 *
3	<i>Morinda officinalis</i>	<i>M. officinalis</i>	737	99.00	99.46	KR869730.1
4	<i>Gynochthodes cochinchinensis</i>	<i>M. officinalis</i>	737	99.00	99.46	NC_053818.1
5	<i>Gynochthodes parvifolia</i>	<i>G. parvifolia</i>	737	99.00	99.32	NC_054151.1
6	<i>Gynochthodes officinalis</i> _H02_ <i>rbcL</i>	<i>M. officinalis</i>	713	96.00	100.00	ON926569.1 *
7	<i>Morinda</i> sp.	<i>Morinda</i> sp.	737	97.00	99.31	AB586541.1
8	<i>Morinda</i> sp.	<i>Morinda</i> sp.	721	97.00	99.31	AB586543.1
9	<i>Morinda officinalis</i>	<i>M. officinalis</i>	703	95.00	99.86	GQ436556.1
10	<i>Gynochthodes officinalis</i> _QN03_ <i>rbcL</i>	<i>M. officinalis</i>	670	90.00	99.85	ON926571.1 *
11	<i>Gynochthodes officinalis</i> _QN01_ <i>rbcL</i>	<i>M. officinalis</i>	670	90.00	99.55	ON926570.1 *

Note: "\*" is the sample in this study.

**Figure 5.** Phylogenetic tree by maximum likelihood of similarity between *M. officinalis* and the closest species as per *rbcL* region sequences. The samples of our study are marked by black squares next to their names.

### 3.2. Nucleotide Components

In terms of the occurrence of each type of nucleotide in the *ITS1* and *ITS2* gene region, cysteine (C) accounted for the highest proportion, ranging from 35.3% to 38.9%, followed by guanine (G), which ranged from 31.1% to 34.0%. Thymine (uracil) was observed with the lowest percentage (ranging from 10.0% to 14.5%) (Table 7). For *matK* and *rbcL*, however, the percentage of Thymine (uracil) was higher than that of the remaining nucleotides (ranging from 28.1% to 36.9%) (Table 8).

**Table 7.** Nucleotide components of the *ITS1* and *ITS2* gene regions of nine *M. officinalis* samples.

Sam.	Compute Nucleotide Composition (%)											
	<i>ITS1</i>						<i>ITS2</i>					
	T(U)	C	A	G	C + G	Total	T(U)	C	A	G	C + G	Total
H01	12.3	35.8	21.2	30.7	66.5	179	14.5	35.7	15.7	34.0	69.7	235
H02	11.2	36.9	20.1	31.8	68.7	179	14.0	36.2	16.2	33.6	69.8	235
H03	12.3	35.8	21.2	30.7	66.5	179	14.0	36.2	16.2	33.6	69.8	235
QN01	10.0	38.9	20.0	31.1	70	180	14.5	35.7	15.7	34.0	69.3	235
QN02	11.7	35.8	20.7	31.8	66.8	179	14.9	35.3	16.2	33.6	68.9	235
QN03	12.3	36.3	20.1	31.3	67.6	179	14.9	35.3	16.2	33.6	68.9	235
QB1	11.7	36.3	20.7	31.3	67.6	179	14.0	36.2	16.2	33.6	69.8	235
QB2	12.3	36.3	20.1	31.3	67.6	179	14.0	36.2	16.2	33.6	69.8	235
QB3	12.3	36.3	20.1	31.3	67.6	179	14.0	36.2	16.2	33.6	69.8	235
Avg.	11.8	36.5	20.5	31.3	68.0	179.1	14.3	35.9	16.1	33.7	69.6	235

**Table 8.** Nucleotide components of the *matK* and *rbcL* gene regions of four *M. officinalis* samples.

Sam.	Compute Nucleotide Composition (%)											
	<i>matK</i>						<i>rbcL</i>					
	T(U)	C	A	G	C + G	Total	T(U)	C	A	G	C + G	Total
H01	36.5	17.8	28.8	16.9	34.7	899	28.6	20.3	28.3	22.8	43.1	738
H02	-	-	-	-	-	-	28.1	20.6	28.6	22.7	43.3	713
QN01	36.9	18.2	27.8	17.1	35.3	861	28.5	20.7	27.5	23.3	44.0	670
QN02	36.5	17.8	28.7	17.1	34.9	897	-	-	-	-	-	-
QN03	36.5	18.1	28.4	17.1	35.2	897	28.8	20.7	27.5	23.0	43.7	670
QB2	-	-	-	-	-	-	28.5	20.5	28.2	22.8	43.3	737
Avg.	36.6	18.0	28.4	17.0	35	888.5	28.5	20.6	28.0	22.9	43.5	705.6

The (G + C) percentage was the highest, with an average of 69.6% in the *ITS2* gene region (Tables 7 and 8).

### 3.3. Genetic Diversity and Population Structure

A total of 140 different alleles were detected from 22 microsatellite markers in *M. officinalis* in the three surveyed populations, with sizes ranging from 106 bp to 328 bp.

In this study, of all 22 loci, the locus MO05 showed the highest number of alleles in all populations ( $N_a = 9$ ), while the locus MO90 showed the lowest number of alleles, with  $N_a = 2.333$ . The  $F_{IS}$  values of the loci ranged between  $-0.287$  and  $0.965$ , with an average of  $0.216$ . Half of the loci (MO02, MO04, MO05, MO12, MO26, MO30, MO39, MO41, MO60, MO63, and MO96) showed positive  $F_{IS}$  values (Table 9). The  $F_{IT}$  values of the loci ranged between  $-0.086$  and  $0.978$ .

All the loci showed positive  $F_{ST}$  values ranging between  $0.127$  and  $0.554$ , with an average of  $0.262$ . This result indicated a relatively high genetic differentiation at the locus level in all populations (Table 9).

Table 10 presents the genetic diversity of *M. officinalis* populations. The allele numbers ( $N_a$ ) ranged from 2 to 8 in the QB population, from 3 to 12 in the TTH population, and from 0 to 10 in the QN population. Specifically, the TTH population had the highest number of alleles, with an average of  $6.455$ , whereas the population with the lowest number of alleles ( $3.955$ ) was the QB population (Table 10). The average percentage of polymorphic loci was  $98.48\%$ , of which two populations, QB and TTH, reached the absolute value ( $100\%$ ).

**Table 9.** Genetic diversity characteristics of 22 microsatellite loci of *M. officinalis*.

Locus	Na	Ne	Ho	He	F <sub>IS</sub>	F <sub>IT</sub>	F <sub>ST</sub>	Nm
MO02	4.667	3.494	0.893	0.698	−0.279	−0.086	0.151	1.405
MO04	4.333	2.902	0.652	0.633	−0.030	0.222	0.244	0.774
MO05	9.000	4.280	0.796	0.733	−0.086	0.109	0.179	1.144
MO12	5.667	3.820	0.815	0.727	−0.121	0.021	0.127	1.718
MO19	7.667	5.277	0.619	0.782	0.209	0.322	0.144	1.491
MO26	6.000	4.061	0.889	0.739	−0.204	−0.015	0.157	1.343
MO30	7.333	5.525	0.963	0.783	−0.230	−0.044	0.152	1.400
MO38	4.000	2.628	0.144	0.574	0.749	0.828	0.318	0.537
MO39	4.000	3.421	0.785	0.669	−0.174	0.071	0.209	0.948
MO41	3.000	2.184	0.689	0.539	−0.278	0.174	0.354	0.457
MO43	2.667	2.142	0.019	0.531	0.965	0.978	0.370	0.425
MO47	3.333	2.091	0.304	0.466	0.349	0.586	0.364	0.437
MO53	3.333	1.892	0.167	0.393	0.576	0.698	0.287	0.621
MO57	7.667	4.660	0.456	0.747	0.391	0.488	0.160	1.311
MO60	5.667	3.501	0.381	0.704	0.458	0.556	0.180	1.139
MO61	5.333	2.863	0.337	0.579	0.418	0.556	0.237	0.805
MO63	4.000	2.333	0.726	0.564	−0.287	0.053	0.265	0.695
MO88	5.000	2.917	0.433	0.652	0.336	0.461	0.189	1.074
MO89	5.333	2.540	0.144	0.425	0.660	0.818	0.463	0.290
MO90	2.333	1.764	0.085	0.396	0.785	0.892	0.498	0.252
MO94	3.333	1.834	0.093	0.405	0.771	0.898	0.554	0.202
MO96	4.667	4.076	0.907	0.736	−0.233	−0.038	0.159	1.327
Mean ± SE					0.216 ± 0.090	0.389 ± 0.077	0.262 ± 0.027	0.900 ± 0.098

**Table 10.** Genetic diversity estimates of three populations of *M. officinalis*.

Pop.	N	Na	Np	Ne	P%	Ho	He	F <sub>ST</sub>
QB	9	3.955	1.818	2.567	100.00	0.556	0.554	0.085
TTH	18	6.455	4.364	3.786	100.00	0.394	0.683	0.451
QN	10	4.364	2.318	3.221	95.45	0.591	0.601	0.076
Mean ± SE		4.924 ± 0.296		3.191 ± 0.197	98.48 ± 1.52	0.513 ± 0.047	0.612 ± 0.023	0.206 ± 0.067

The highest total number of private alleles (Np = 4.364) was observed in the TTH population, followed by the QN population (Np = 2.318) and the QB population (Np = 1.818). Similarly, of the three populations, the effective allele numbers of the TTH population were the highest at 3.786, compared to 3.221 and 2.567 for the QN and QB populations, respectively.

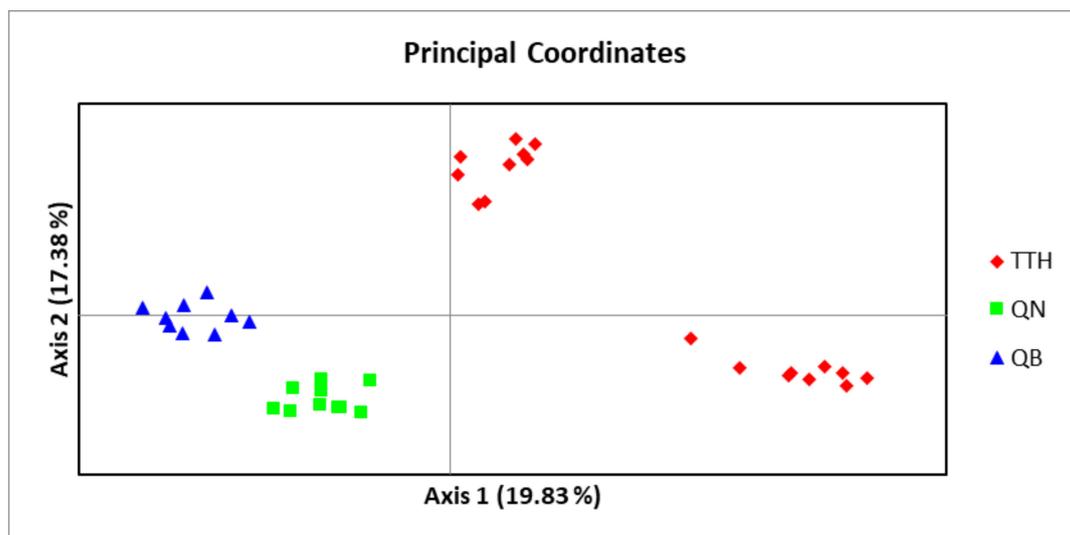
Observed heterozygosity (Ho) and expected heterozygosity (He) averaged 0.513 and 0.612, respectively. The QN population had the highest expected heterozygosity (0.591) of the three populations, whereas the value of the observed heterozygosity of that population was 0.601. The fixation index (F<sub>ST</sub>) was positive for all populations (F<sub>ST</sub> = 0.206).

The results of the AMOVA pointed to higher levels of variation within populations (70%) than between populations (30%) (Table 11), indicating a relatively low genetic divergence among the three studied populations (TTH, QB, QN). However, within populations, individuals were found to have high genetic differentiation.

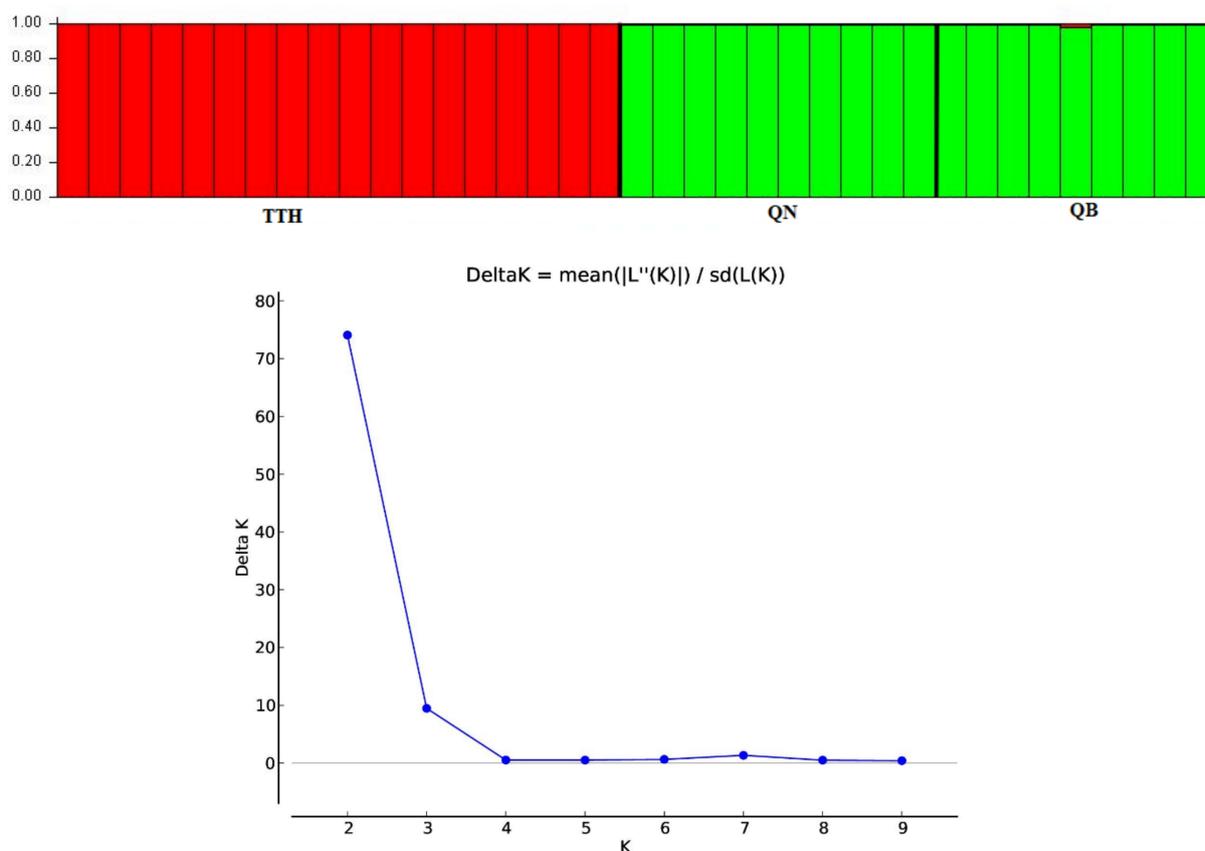
**Table 11.** Analysis of molecular variance (AMOVA) for different groups of *M. officinalis* How.

Source	df	SS	MS	Est. Var.	%
Among populations	2	158.697	79.348	3.089	30%
Within populations	71	512.222	7.214	7.214	70%
Total	73	670.919		10.303	100%
Stat	Value	P(rand >= data)			
Fst	0.300	0.001			

An admixture model was performed to determine the group population of the 37 individuals of *M. officinalis* based on Bayesian analysis using the STRUCTURE program. The population structure of *M. officinalis* inferred in STRUCTURE revealed that the optimum number of genetic groups for the admixture model was  $K = 2$ , with a delta  $K$  value of 74.084117. After a PCA analysis, the three populations were classified into two distinct clusters.  $K = 2$ , thus, can be considered to be the most suitable number of clusters for the structure of the populations in this study (Figures 6 and 7). The QB and QN populations were in one cluster (green cluster), whereas the TTH population was in the remaining cluster (red cluster) (Figure 7).



**Figure 6.** Principal coordinate analysis (PCoA) based on codominant genotypic distance for 37 studied samples of three populations of *M. officinalis* ( TTH: TTH population; QN: QN population; QB: QB population).



**Figure 7.** Delta K distribution graph and bar plot of the admixture assignment for three *M. officinalis* populations to clusters ( $K = 2$ ; highest  $\Delta K$  value = 74.084117) based on Bayesian analysis.

#### 4. Discussion

DNA barcoding is a novel approach for identifying and classifying species based on the nucleotide diversity of conserved sequences. Recently, many studies have indicated that DNA markers *ITS1*, *ITS2*, *matK*, and *rbcL* are highly effective in identifying medicinal plants at species and genus levels. Therefore, this study was conducted to investigate the efficiency of those markers in barcoding *M. officinalis* plant species for the first time to identify the best marker for this valuable plant.

In previous studies, the *matK* gene and (*ITS*) region genetic markers have been proven highly efficient in distinguishing plant species and therefore identified as potential candidates for barcoding plants [29]. This study demonstrated that the *ITS1* and *ITS2* regions of *M. officinalis* were amplified more effectively compared to the other two markers (*matK* and *rbcL*).

Sequence analysis using the above four primers showed that the samples studied in this experiment belonged to the genus *Morinda* and the most closely related species was *Gynochthodes officinalis* (synonym of *M. officinalis*). In the present study, the *ITS2* marker showed the highest efficiency in terms of amplification and species identification. The higher capacity of *ITS2* in terms of phylogenetic reconstruction has been proven in previous studies [12,13,29].

The origin and ecology of a species are often expressed through its genetic diversity [30]. Many previous studies have proven a positive correlation between population genetic diversity, population size, and geographic distribution range [30,31]. Species with broad distribution and a large population size generally maintain high genetic diversity in comparison to species with narrow distribution and small population size.

Heterozygosity can reflect the genetic variation in natural populations and is considered a measure of genetic diversity. The higher the heterozygosity in a population is, the more the genetic variability in it is. *M. officinalis* had high genetic diversity ( $H_o = 0.513$  and

$H_e = 0.612$ ) in this study compared to in Luo's publication for the same species ( $H_o = 0.3436$  and  $H_e = 0.2881$ ) [5].

$F_{ST}$  is considered one of the measures of genetic variation in populations. An  $F_{ST}$  value of higher than 0.15 can indicate significant differentiation in a population [32]. Therefore, remarkable diversity was observed in the three populations in this study (Table 10). The  $F_{ST}$  value ranged from 0.076 (QN population) to 0.451 (TTH population), with an average  $F_{ST}$  value of 0.206. This result was higher than that in the study of *P. vietnamensis* ( $F_{ST} = 0.13$ ) [19] and lower than those in the studies of *Cinnamomum balansae* ( $F_{ST} = 0.601$ ) [26] and *Pulsatilla patens* (L.) ( $F_{ST} = 0.22$ ) [33] and ( $F_{ST} = 0.13$ ) using microsatellite markers.

The results of the AMOVA also revealed that 30% of the total variation was found among subpopulations, while the rest (70%) was within subpopulations. These results indicated a relatively low genetic differentiation among the three populations and a high genetic divergence within those populations. Genetic variation among populations was highly influenced by gene flow, selection, genetic drift, and other factors.

A Bayesian analysis in STRUCTURE showed two different groups of genetically mixed individuals of *M. officinalis*. Even though the populations were separated by large geographical distances, the majority of individuals from the QB and QN populations shared the same ancestral origin.

## 5. Conclusions

The study proved the *ITS1* region and the *ITS2* region as reliable markers in the bar-coding of *M. officinalis*. In addition, the results confirmed that *M. officinalis* had low genetic differentiation among the three studied populations and high genetic divergence within those three populations. The three populations of *M. officinalis* were divided into genetic groups on the basis of the codominant genotypic distance. Our study of *M. officinalis* in terms of genetic diversity and population structure is intended to make a large contribution to the conservation of this species.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/genes13111938/s1>: Table S1. Geographical characteristics of the 37 samples of *M. officinalis* used in this study; Table S2. The primers used for PCR and sequencing in this study; Table S3. The samples were deposited in the GenBank and their accession numbers.

**Author Contributions:** T.P.: conceptualization, methodology, investigation, writing—original draft, revision. Q.T.N.: methodology, investigation, writing—original draft, revision. D.M.T., H.N., H.T.L., Q.T.H.H. and Y.T.V.: collecting samples; T.N.T.: conceptualization, writing—review, revision and editing, project administration. All authors have read and agreed to the published version of the manuscript.

**Funding:** The Ministry of Education and Training under grant number B2021-DHH-18.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Acknowledgments:** This research was funded by the Ministry of Education and Training under grant number B2021-DHH-18 and the Strong Research Group Program of Hue University.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bin, W.Y.; Guo, W.J.; Jian, Z.C.; Ting, H.; Ping, Q.L.; Zhong, W.J.; Yan, Z.Q.; Wu, Y.-B.; Zheng, C.-J.; Han, T.; et al. Quantitative and chemical profiles analysis of the root of *Morinda officinalis* based on reversed-phase high performance liquid chromatography combined with chemometrics methods. *J. Med. Plants Res.* **2013**, *7*, 2249–2258. [CrossRef]
2. Hieu, N.C.; Hong, T.L.; Oanh, N.L.; Nguyen, D.T.; Hoa, N.T. Effects of different plant densities and fertilizers on root rot disease of indian mulberry (*Morinda officinalis* How.) in Thai Nguyen. *TNU J. Sci. Technol.* **2019**, *202*, 199–204. (In Vietnamese)
3. The, H.T.; Thao, N.T.P.; Thao, N.T.; Thuy, N.T. In vitro Propagation of *Morinda officinalis* How. *J. Sci. Devel.* **2013**, *11*, 285–292. (In Vietnamese)

4. Hieu, H.D.; Ha, C.T.T.; Ngoc, P.B.; Nhan, D.L.; Huong, N.T.T.; Ha, C.H. Study on genetic diversity in *Morinda officinalis* F. C. How populations in Quang Ninh using ISSR marker. *Acad. J. Biol.* **2016**, *38*, 89–95. [[CrossRef](#)]
5. Luo, Z.; Chen, Z.; Liu, M.; Yang, L.; Zhao, Z.; Yang, D.; Ding, P. Phenotypic, chemical component and molecular assessment of genetic diversity and population structure of *Morinda officinalis* germplasm. *BMC Genom.* **2022**, *23*, 605. [[CrossRef](#)]
6. Zhang, J.H.; Xin, H.L.; Xu, Y.M.; Shen, Y.; He, Y.Q.; Lin, B.; Song, H.T.; Yang, H.Y.; Qin, L.P.; Zhang, Q.Y.; et al. *Morinda officinalis* How.—A comprehensive review of traditional uses, phytochemistry and pharmacology. *J. Ethnopharmacol.* **2018**, *213*, 230–255. [[CrossRef](#)]
7. Thinh, N.Q.; Linh, P.T.; Trang, N.T.H.; Huy, Đ.Q. Study of conserving the medicinal plant resources in Phu Luong district, Thai Nguyen province, Vietnam. *TNU J. Sci. Technol.* **2019**, *194*, 47–52. (In Vietnamese)
8. Zaman, W.; Ye, J.F.; Ahmad, M.; Saqib, S.; Shinwari, Z.K.; Chen, Z. Phylogenetic exploration of traditional Chinese medicinal plants: A case study on Lamiaceae (angiosperms). *Pak. J. Bot.* **2022**, *54*. [[CrossRef](#)]
9. Yang, Z.; Rannala, B. Molecular phylogenetics: Principles and practice. *Nat. Rev. Genet.* **2012**, *13*, 303–314. [[CrossRef](#)]
10. Al-Juhani, W.S.; Khalik, K.N. Identification and molecular study of medicinal *Plectranthus* species (Lamiaceae) from Saudi Arabia using plastid DNA regions and ITS2 of the nrDNA gene. *J. King Saud Univ. Sci.* **2021**, *33*, 101452. [[CrossRef](#)]
11. Li, M.; Cao, H.; But, P.P.; Shaw, P.C. Identification of herbal medicinal materials using DNA barcodes. *J. Syst. Evol.* **2011**, *49*, 271–283. [[CrossRef](#)]
12. Chen, S.; Yao, H.; Han, J.; Liu, C.; Song, J.; Shi, L.; Zhu, Y.; Ma, X.; Gao, T.; Pang, X.; et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE* **2010**, *5*, e8613. [[CrossRef](#)]
13. Wang, X.; Zheng, S.; Liu, Y.; Han, J. Chinese Herbal Medicines (CHM) Original article ITS2, a Better DNA Barcode than ITS in Identification of Species in *Artemisia* L. *Chin. Herb. Med.* **2016**, *8*, 352–358. [[CrossRef](#)]
14. Wattoo, J.I.; Saleem, M.Z.; Shahzad, M.S.; Arif, A.; Hameed, A.; Saleem, M.A. DNA Barcoding: Amplification and sequence analysis of *rbcl* and *matK* genome regions in three divergent plant species. *Int. J. Adv. Life Sci.* **2016**, *4*, 3–7.
15. Avinash, M.; Anurag, K.S.; Gaur, R.K. Molecular markers: Tool for genetic analysis. *Anim. Biotechnol.* **2014**, *1*, 289–305.
16. Kalia, R.K.; Rai, M.K.; Kalia, S.; Singh, R.; Dhawan, A.K. Microsatellite markers: An overview of the recent progress in plants. *Euphytica* **2011**, *177*, 309–334. [[CrossRef](#)]
17. Li, Y.C.; Korol, A.B.; Fahima, T.; Nevo, E. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* **2004**, *21*, 991–1007. [[CrossRef](#)]
18. Vieira, M.L.C.; Santini, L.; Diniz, A.L.; Munhoz, C.D.F. Microsatellite markers: What they mean and why they are so useful. *Genet. Mol. Biol.* **2016**, *39*, 312–328. [[CrossRef](#)]
19. Vu, D.D.; Shah, S.N.M.; Pham, M.P.; Bui, V.T.; Nguyen, M.T.; Nguyen, T.P.T. De novo assembly and Transcriptome characterization of an endemic species of Vietnam, *Panax vietnamensis* Ha et Grushv., including the development of EST-SSR markers for population genetics. *BMC Plant Biol.* **2020**, *20*, 358. [[CrossRef](#)]
20. Liao, B.; Lee, S.; Meng, K.; Yin, Q.; Huang, C.; Fan, Q.; Liao, W.; Chen, S. Characterization and novel Est-SSR marker development of an important Chinese medicinal plant, *Morinda officinalis* How (Rubiaceae) Boyong. *Biotechnol. Biotechnol. Equip.* **2019**, *33*, 1311–1318. [[CrossRef](#)]
21. Hall, T.A. BIOEDIT: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.
22. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)] [[PubMed](#)]
23. Kumar, S.; Stecher, G.; Li, M.; Nkya, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
24. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **1980**, *16*, 111–120. [[CrossRef](#)] [[PubMed](#)]
25. Peakall, R.; Smouse, P.E. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)] [[PubMed](#)]
26. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [[CrossRef](#)]
27. Earl, D.A.; VonHoldt, B.M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
28. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)]
29. Yu, J.; Wu, X.; Liu, C.; Newmaster, S.; Ragupathy, S.; Kress, W.J. Progress in the use of DNA barcodes in the identification and classification of medicinal plants. *Ecotoxicol. Environ. Saf.* **2021**, *208*, 111691. [[CrossRef](#)]
30. Cui, B.; Vu, D.D.; Vu, D.G.; Bui, T.T.X.; Rahman, S.U.; Pham, M.P.; Nguyen, M.T.; Shah, S.N.M.; Tran, V.H. Genetic diversity and population structure of *Cinnamomum balansae* Lecomte inferred by microsatellites. *Open Life Sci.* **2022**, *17*, 323–332. [[CrossRef](#)]
31. Vandewoestijne, S.; Schtickzelle, N.; Baguette, M. Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biol.* **2008**, *6*, 46. [[CrossRef](#)]

32. Luo, Z.; Brock, J.; Dyer, J.M.; Kutchan, T.; Schachtman, D.; Augustin, M.; Ge, Y.; Fahlgren, N.; Abdel-Haleem, H. Genetic diversity and population structure of a *Camelina sativa* spring panel. *Front. Plant Sci.* **2019**, *10*, 184. [[CrossRef](#)]
33. Szczecińska, M.; Sramko, G.; Wołosz, K.; Sawicki, J. Genetic diversity and population structure of the rare and endangered plant species *Pulsatilla patens* (L.) Mill in East Central Europe. *PLoS ONE* **2016**, *11*, e0151730. [[CrossRef](#)] [[PubMed](#)]