Advanced Studies in Biology, Vol. 13, 2021, no. 1, 1 – 9
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/asb.2021.91247

Genetic Diversity and Genetic Relationship of Vietnamese *Citrus* Varieties Using Internal Transcribed Spacer Region (ITS)

L.T.T. Trang [1], D.T.T. Ha [2], N.T. Diep [2], L.T. Nghia [1], D.X. Duc [2], N.T. Khoa [2], L.H. Anh [2], K.T. Dung [2], K.H. Trung [2], D.H. Gioi [3], B.T.T. Huong [3], T.D. Khanh [2] and H.T. Hue [1,*]

[1] Plant Resources Center, Hanoi Vietnam, Vietnam Academy of Agricultural Sciences, Vietnam (VAAS)
[2] Agricultural Genetics Institute, Hanoi 122300, Vietnam, Vietnam Academy of Agricultural Sciences, Vietnam (VAAS)
[3]Vietnam National University of Agriculture, Hanoi, Vietnam
[*]Corresponding author

Abstract

Vietnam is considered to be the center of origin of many *Citrus* species. Regional structures of territory, climatic and soil conditions are suitable and favorable for *Citrus* production as well as rich in germplasm resources. In this study, 15 *Citrus* samples were collected in different areas and identified by ITS sequencing. The results showed that the nucleotide sequences and the coverage of the samples were ranged from 95.60%to 99.86% and were 97.0% to 99.0%, respectively, as compared with the published reference samples. Specifically, the C9 sample (Cam Duong Ha Tinh) had a similar level with the reference samples of *JN681155.1 Citrus maxima* (99.72%), which was higher than the homologous sample of *JN681165.1 Citrus sinensis* by 98.46%. The C10 sample (Cam Sanh Bo Ha) was similar to the reference sample *MH721728.1 Citrus reticulata* (99.72%), higher than the reference *JN681150.1 Citrus sinensis* up to 99.45%. Based on the ITS genome sequences, 6/15 *Citrus* samples were accurately identified. The results also indicate that Vietnamese native *Citrus* varieties are diversified in the genus *Citrus* species. Our findings have provided useful information for the accurate identification of some native *Citrus* germplasms in this country.

Keywords: Native *Citrus* germplasm, genetic diversity, ITS sequencing

Introduction

The genus of *Citrus* belongs to the orange subfamily Aurantioideae of the family Rutaceae and reckons in origin from the tropical and subtropical countries and spread to many areas in the world [1]. Some commercially important fruits such as lime (*C.aurantiifolia*), lemon (*C. limon*), grapefruit (*C.paradisi*), sweet orange (*C. sinensis*), mandarin (*C. reticulata*) are included *Citrus* genus [2]. However, many controversial and ambiguous issues related to *Citrus* taxonomy and phylogeny have emerged due to their sexual compatibility among species, apomixis, bud mutation, etc.,[3].

Vietnam is located in central of many *Citrus* arise, along with the segmentation of altitude, terrain which made the ecological subregion are varied, suitable and favorable for specialty *Citrus* producing as well as rich in *Citrus* germplasm resources. Currently, many *Citrus* cultivars including native and imported *Citrus* varieties, are widely grown from north to south of this country. Moreover, *Citrus* fruit has much nutritional value and a high economy that plays a key role in developing sustainable commercial agriculture. According to General Statistical Office [4], there were about 138.000 hectares in the country with an annual output that reached 1.01 million tons, especially some native species widely grown in many areas across this country.

In some recent years, numerous molecular techniques have been developed which are helping to identify plant diversity at the molecular level, provided the basis for the assessment of species conservation value, identification of plant varieties, selection of parents for breeding, and conservation of genetic resources. In *Citrus* species, some previous reports have used various markers such as AFLP, RAPD, SSR, ISSRs and ITS sequences [1, 5-6]. Among them, the internal transcribed spacer (ITS) has been extensively applied for plant systematic to represent the phylogenetic relationships at the different taxonomic levels [7]. ITS region is useful for low-level phylogenetic analysis, comprising infra-generic level, because of its relatively rapid evolution rate [8]. On the other hand, ITS is one of the most comprehensively utilized molecular markers for angiosperm phylogenetic inference and genetic involvement in plants. Typically, ITS1/ITS4 has been broadly used in characteristic molecular studies due to their relatively high variability which has provided molecular evidence to assess the phylogeny of taxonomic groups in numerous plant species.

In order to understand on the genetic diversity and phylogenetic relationships in Vietnamese native *Citrus* in this country. The objectives of this study were to analyze 15 Vietnamese native *Citrus* collected from the 8 different provinces using ITS sequences and compare with the availably published data on the NCBI-database.

Materials and Methods

Material collection

The young leaf samples of total 15 *Citrus* native varieties were collected in 8 different provinces from north to south areas in Vietnam (Dong Thap, Nghe An, Ha

Tinh, Bac Giang, Quang Ninh, Ha Giang, Hanoi and Cao Bang provinces). The detailed information of the samples as shown in Table 1.

Table 1. List of the samples of *Citrus* varieties used in this study

No	Original name	Collected areas	Code
1	Cam Tay Giang	Tay Giang - Quang Nam province	C1
2	Cam xoan	Lai Vung - Đong Thap province	C2
3	Cam sành	Lai Vung - Đong Thap province	C3
4	Xa doai	N.Quang- N. Dan- Nghe An, province	C4
5	Song con	N.Quang- N. Dan- Nghe An, province	C5
6	Van du	N.Quang- N. Dan- Nghe An, province	C6
7	Cam bu	N.Quang- N. Dan- Nghe An, province	C7
8	Cam giay	N.Quang- N. Dan- Nghe An, province	C8
9	Cam đuong	Huong Son- Ha Tinh, province	C9
10	Cam sanh Bo ha	Bo Ha, Yen The, Bac Giang, province	C10
11	Cam đuong	Quang Ninh, province	C11
12	Cam sap	Quang Ninh, province	C12
13	Cam sanh	Ha Giang, province	C13
14	Cam duong canh	Ha Noi	C14
15	Cam Trung Vuong	Cao Bang, province	C15

Genomic DNA extraction, PCR amplification, ITS sequencing and phylogenetic analysis

A total of DNA extraction was done following the CTAP method with some minor modifications [9]. The yielded DNA products were checked on the agarose gel (1%). The PCR reaction was performed on the Veriti 96 wells Thermal cycler with the nucleotide sequences of ITS1/ITS4 primers: TCCGTAGGTGAACCTGCGG/TCCTCCGCTTATTGATATGC.TCCGTAGGT GAACCTGCGG. The amplification of the primers and PCR program was performed as the method of Trung et al [7]. The DNA of all samples was sent to Apical Scientific (Malaysia) for sequencing.

Statistical Analysis

The sequencing data were compared with the homologous sequence published on the NCBI database and then analyzed by MEGA v6.06 and the neighbor-joining (NJ) to generate a phylogenetic tree.

Results and Discussion

Identification of *Citrus* samples using ITS sequences analysis

In this study, we have used 15 samples of *Citrus* leaves collected from the 8 representative provinces from north to south in Vietnam including Dong Thap Nghe An, Ha Tinh, Bac Giang, Quang Ninh, Hanoi, Cao Bang and Ha Giang provinces. The primer pairs of ITS1/ITS4 were successfully amplified with high quality and

quantity with the appearance of only one band with a length size of approximately 750 bp as shown in Figure 1.

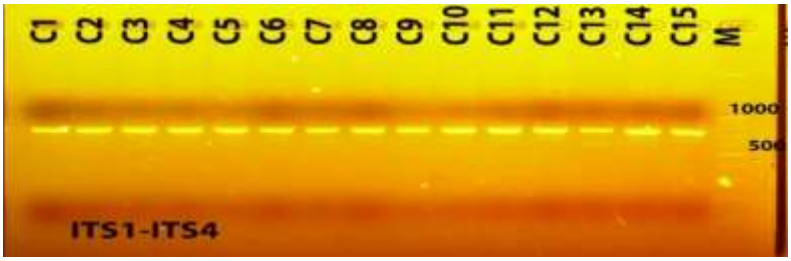

Fig 1. PCR amplification of 15 *Citrus* samples using ITS1/ITS4 primers
M: Standard Marker 100bp; Lane C1-C15: *Citrus* samples

A total of 15 PCR products of *Citrus* samples were sequenced by ABI 3730X1 (Apical Scientific-Malaysia). The sequences of the samples were analyzed based on the comparison of the published sequences references of NCBI-database using MEGA v6.06. The results showed that the homologous sequences of 15 barcode DNA were ranged from 90.60% to 99.86%, and the coverage was from 97.0% to 99.0% comparing to the published sequences data on NCBI, respectively. It implies that the amplified ITS region sequences of the samples have had a high degree of similarity. Specifically, sample C9 (Ha Tinh Cam Duong) had a similar level with reference sample *JN681155.1 Citrus maxima* (99.72%), which is higher than the similar level with reference sample *JN681165.1 Citrus sinensis* (98.46%). Similarly, sample C10 (Orange Sanh Bo Ha) had a similar level with the reference *MH721728.1 Citrus reticulata* (99.72%) higher than the similarity level with the reference sample and *JN681150.1 Citrus sinensis* (99.45%) (Table2). Our findings are consistent with the previous study of Tina et al [10], who confirmed some *Citrus* accessions belonging to three *Citrus* species included *C.medica, C.mexima and C.reticulata*, and suggested as the "basic" true species.

Table 3 and Figure 1 present the different positions of the nucleotide sequences among the studied samples and the reference samples in detail. There were some deletion and insertion segments occurring in some positions of the genes. Hence, the total number of nucleotides of each sequence are different between the studied samples. The ITS region length of the 15 samples has fluctuated from 179 to 734 nucleotides. Interestingly, the ITS region showed very high morphism of nucleotide sequences with 78 different positions between the studied *Citrus* samples and the reference samples. Typically, 6 *Citrus* samples (varieties) including C1, C8, C9, C10, C12 and C13 have had different nucleotide variation to compare with the other samples and the reference samples such as *JN681155.1 Citrus maxima, JN681165.1 Citrus sinensis, MH721728.1 Citrus reticulata* and *JN681150.1 Citrus sinensis*, respectively (Table 3).

Table 2. Homologous and coverage levels of ITS sequences of the samples compared to the correspondence of the reference sequences of NCBI database

No	Original name	Code	Reference sequences	H (%)	C (%)
1	Cam Tay Giang	C1	*JN681165.1_Citrus_sinensis*	97.23	97
2	Cam xoan	C2	*JN681150.1_Citrus_sinensis*	99.59	98
3	Cam sanh	C3	*JN681150.1_Citrus_sinensis*	99.32	98
4	Xa doai	C4	*JN681150.1_Citrus_sinensis*	99.86	97
5	Song con	C5	*JN681150.1_Citrus_sinensis*	99.45	98
6	Van du	C6	*JN681150.1_Citrus_sinensis*	99.59	98
7	Cam bu	C7	*JN681150.1_Citrus_sinensis*	98.76	98
8	Cam giay	C8	*KY486202.1_Citrus_sinensis*	95.95	99
9	Cam duong Ha Tinh	C9	*JN681165.1_Citrus_sinensis*	98.46	98
			JN681155.1_Citrus_maxima	99.72	98
10	Cam sanh Bo ha	C10	*JN681150.1_Citrus_sinensis*	99.45	97
			MH712728.1_Citrus_reticulata	99.72	97
11	Cam duong Quang Ninh	C11	*JN681165.1_Citrus_sinensis*	98.91	98
12	Cam sap	C12	*KY486202.1_Citrus_sinensis*	94.63	99
13	Cam sanh	C13	*KY486202.1_Citrus_sinensis*	94.60	99
14	Cam duong canh	C14	*JN681150.1_Citrus_sinensis*	98.08	98
15	Cam Trung Vuong	C15	*JN681150.1_Citrus_sinensis*	99.32	98

H: Homologous; C: Coverage

Table 3. Different positions of nucleotide sequences among the samples and references samples

Variety code	Different positions of nucleotide sequences among the samples																								
	250	489	526	546	561	589	597	609	610	618	621	628	629	631	635	642	657	661	688	698	700	704	714	719	727
C1	C	C	T	A	C	-	T	C	C	C	T	A	T	G	C	A	G	G	C	T	C	C	G	T	-
C2	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
C3	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
C4	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	C
C5	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
C6	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
C7	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
C8	.		.	C	.	A	C	.	G	**G**	C	T	G	.	**T**	G	C	**C**	.	**A**	.	**T**	C	.	-
C9	**T**	**T**	C	.	.	-	C	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	.	-
C10	.		C	C	**G**	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	C
C11	.		.	.	.	-	C	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	.	-
C12	.		.	.	.	-	C	G	.	.	G	.	.	.	A	.	C	.	.	.	**A**	.	.	**A**	**A**
C13	.		.	.	.	-	C	.	.	.	G	.	.	.	A	.	C	.	**A**	.	.	.	.	.	-
C14	.		.	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
C15	.		C	C	.	A	C	.	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	-
JN681165.1 Citrus sinensis	.		.	.	.	-	C	.	.	.	.	G	.	.	.		.	.	.	.	.	.	.	.	-
JN681150.1 Citrus sinensis	.		C	C	.	A	C	.	.	.	.	G	.	.	.		.	.	.	.	.	.	.	.	-
JN681155.1 Citrus maxima	T	T	C	.	.	-	C	.	.	.	.	G	.	.	.		.	.	.	.	.	.	.	.	-
MH712728.1 Citrus reticulata	.		C	C	G	A	C	.	.	.	.	G	.	.	.		.	.	.	.	.	.	.	.	-
MF797953.1 Citrus maxima	.		C	.	.	-	C	.	.	.	.	G	.	.	.		.	.	.	.	.	.	.	.	-

In this study, based on the ITS sequences, together with results of ITS1/ITS4 primers analyses, it is possible to accurately distinguish 6 samples of *Citrus* samples included C1 (Cam Tay Giang), C8 (Cam Giay), C9 (Cam Duong Ha Tinh), C10 (Cam Sanh Bo Ha), C12 (Cam Sap), C13 (Cam Sanh), respectively. Specifically, the sample C1 (Cam Tay Giang) has had a single nucleotide substitution at position 597 and changed from C to T compared with 5 reference varieties published on the NCBI database. On the other hand, this *Citrus* variety (C1) can be accurately identified by this difference. The sample C8 (Cam Giay) has had nucleotide substitution at the positions 618, 635, 661, 698, 704 and 714, respectively. These results may help confirm the correct scientific names of these species samples. Particularly note that the sample C1 (Cam Tay Giang) has had a single nucleotide substitution at position 597, which changed from C to T compared to the 5 published reference varieties on the NCBI database. Hence, this *Citrus* variety (Cam Tay Giang) can be accurately identified by this difference. Similarly, the nucleotide substitutions of sample C8 (Cam Giay) were found at positions of 618, 635, 661, 698, 704, and 714, respectively. Especially, 5 out of 6 nucleotide substitutions at the positions 618, 635, 661, 698 and 704 were replaced C to G, C to T, G to C, T to A and G to C, respectively.

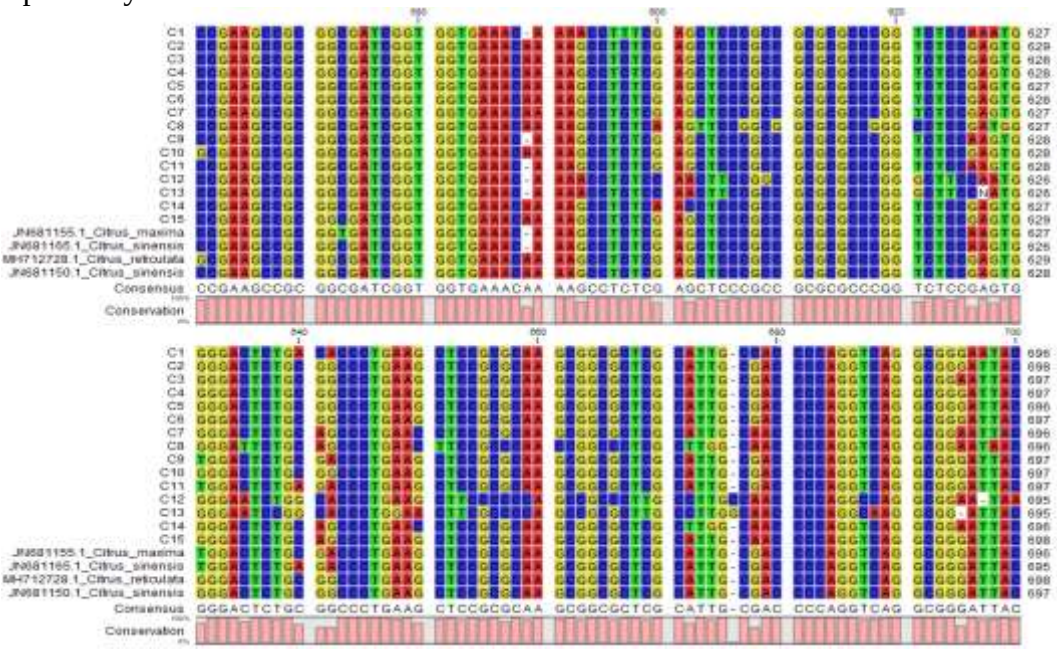

Fig 1. Comparison of the nucleotide sequences of ITS regions of the studied samples with the published reference samples using ITS1/ITS4

Based on these differences, the C8 is possible to accurately identify among the other *Citrus* varieties. At positions of 250 and 489, only the sample C9 (Cam Duong Ha Tinh) has the replacement of C to T nucleotides, which was similar to the reference sample *JN681155.1 Citrus maxima* and was different from the other *Citrus* varieties in this study. Also, we have found that the sample C10 (Cam Sanh Bo Ha) has nucleotide replacement at position 561 and additional nucleotide at 727.

However, the replacement C to G at position 561 was similar to the reference sample *MH721728.1 Citrus reticulata*. Therefore, it needs to refer to this reference sample to identify this *Citrus* variety among the others precisely.

For the sample C12 (Cam Sap), we have found the replacements of the nucleotides included C, C, T to A at the position of 635, 700, and 719, respectively, and the additional nucleotide A at position 727. Hence, due to these different positions at 700, 719 and 727, the sample C12 can be distinguished among other varieties by using these nucleotides' differences. The sample C13 (Cam Sanh) was found that the nucleotide substitution of C to A at the position of 635 and 688. Nevertheless, in this study, 03 samples including C12, Q4 and Q9 have had the same nucleotide replacements. Therefore, by identifying the sample C13, the different position at 688 can be used to accurately distinguish this sample. Currently, most *Citrus* species are widely accepted to derive from hybridization. In this study, some samples showed high homologous sequence with *C. reticulata, C. maxima and C. sinensis*, that implying they might be derived from the parent of sweet orange [11] which was a hybrid of pummelo and mandarin [12].

Analysis of ITS regions of DNA sequences among the samples

Based on the analysis of ITS regions among the samples, they have been grouped into 4 clusters (Figure 2).

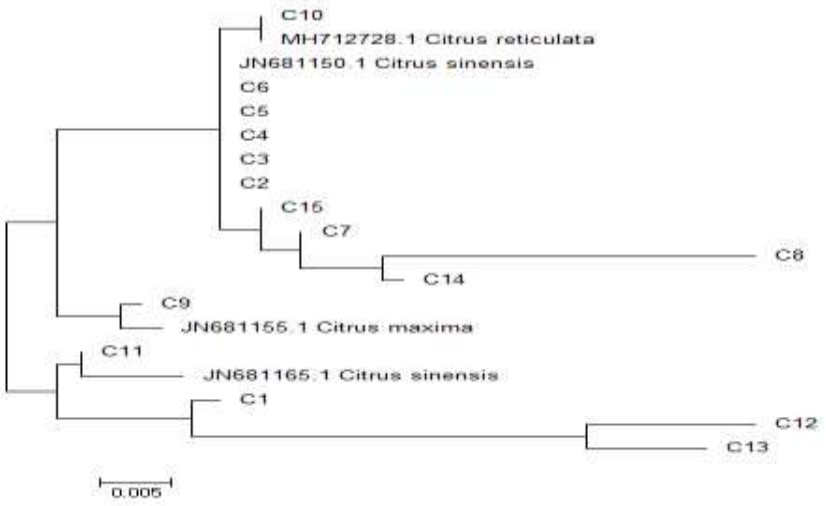

Fig 2: A genetic relation tree generated among the 15 *Citrus* samples

The first cluster included 10 *Citrus* samples (C2, C3, C4, C5, C6, C7, C8, C10, C14, and C15), respectively. In which the sample C10 has had the same branch as the published reference sample *MH712728.1_Citrus_reticulata*. Similarly, the nucleotide sequence at position 561, only C10 sample has had a replacement from C to G substitution, which was similar to the published reference *MH712728.1_Citrus_reticulata*. Nine (09) samples were the same branch as the reference sample *JN681150.1_Citrus_sinensis*.

The second cluster included only C9 sample (Cam Duong Ha Tinh) and the reference sample *N681155.1_Citrus_maxima*. It disclosed that the nucleotide sequences at the position 250 and 489 were changed from C to T, which was similar to the published reference sample *JN681155.1 Citrus maxima*.

The third cluster consisted of 4 samples C1, C11, C12 and C13 were the same as the reference sample *JN681165.1_Citrus_sinensis.* Undoubtedly, much effort on using molecular markers including ITS sequences, has been universally utilized to identify numerous *Citrus* varieties and the genetic relationships of this genus and distinguish the native germplasm and wild species of *Citrus* [13, 7, 6, 10]. However, in this study, further works should be applied multiple advanced molecular markers such as *matK, rbcL, PsbA-trnHtrnL-F* to more precisely classify and identify *Citrus* species for a better understanding of genetic diversity and its distribution in this country.

Conclusions

In summary, by applying the ITS sequencing and comparing the published references, we have successfully distinguished 15 different *Citrus* samples collected in the different regions in Vietnam. Based on the ITS genome sequences 6/15 *Citrus* samples varieties were accurately identified. The results also indicate that Vietnamese native *Citrus* varieties are diversified due to their species – adaptation locations -dependent. Our findings may provide useful information for *Citrus* identification and breeding programs in this country.

References

[1] H.J. Webber, History and development of the *Citrus* industry, p.1–39, in: W. Reuther, H.J. Webber, and L.D. Batchelor (eds.). *The Citrus industry*,Vol.1, History, world distribution, botany, and varieties, *Univ. California, Div. Agr.Sci., Berkeley, CA* (1967).

[2] K.H. Trung, T.D. Khanh, Assessment of genetic diversity of native pumelo (*Citrus maxima* Merr) by utilizing SSR markers, *Int. J. Sci. Res.*, **9** (2017), 881-884.

[3] E. Nicolosi, Z.N. Deng, A. Gentile, S.L. Malfa, G. Continella, E. Tribulato, *Citrus* phylogeny and genetic origin of important species as investigated by molecular markers, *Theor. Appl. Genet*,, **100** (2000), 1155–1166. https://doi.org/10.1007/s001220051419

[4] GSO-General Statistical Office, annual reports, 2011.

[5] N.A. Barkley, M.L. Roose, R.Krueger, C.T. Federici, Assessing genetic diversity and population structure in a *Citrus* germplasm collection utilizing simple sequence repeat markers (SSRs), *Theor. Appl. Genet.*, **112** (2006), 1519-1531. https://doi.org/10.1007/s00122-006-0255-9

[6] M.K. Biswas, L. Chai, M.H. Mar, X. Zhang, X.X. Deng, Comparative analysis of genetic diversity in *Citrus* germplasm collection using AFLP, SSAP, SAMPL and SSR markers, *Sci. Hort.*, **129** (2011), 798-803. https://doi.org/10.1016/j.scienta.2011.06.015

[7] K.H. Trung, T.D. Khanh, L.H. Ham, T.D. Duong, N.T. Khoa, Molecular phylogeny of the endangered Vietnamese *Paphiopedilum* species based on the Internal Transcribed Spacer of nuclear ribosomal DNA, *Adv. Stud. Biol*, **5** (7) (2013), 337-346. https://doi.org/10.12988/asb.2013.3315

[8] C.C. Tsai, S.C. Huang, C.H. Chou, Molecular phylogeny of Phalaenopsis Blume (Orchidaceae) based on the internal transcribed spacer of the nuclear ribosomal DNA, *Plant Sys. Evol.,* **256** (2005), 1-16. https://doi.org/10.1007/s00606-005-0356-y

[9] J.J. Hoyle, J.L. Doyle, A rapid DNA isolation procedure for small qualities of fresh leaf tissue, *Phytochem. Bull.,* **19** (1987) 11–15.

[10] T. Kyndt, T.N. Dung, P. Goetghebeur, H.T. Toan, G. Gheysen, Analysis of ITS of the rDNA to infer phylogentic relationships among Vietnamese *Citrus* accessions, *Genet. Resour. Crop. Evol.*, **57** (2010), 183-192. https://doi.org/10.1007/s10722-009-9460-0

[11] T. Handa, Y. Ishizawa, C. Oogaki, Phylogenetic study of fraction I protein in the genus *Citrus* and its close related genera, *Jpn. J. Genet.,* **61** (1986), 15–24. https://doi.org/10.1266/jjg.61.15

[12] X. Xi, and R. Xie, The origin of cultivated *Citrus* as inferred from internal transcribed spacer and chloroplast DNA sequence and amplified fragment length polymorphism fingerprints, *J. Amer. Soc. Hort. Sci.,* **135** (2010), 341-350. https://doi.org/10.21273/jashs.135.4.341

[13] A. Uzun, T. Yesiloglu, Genetic diversity of *Citrus*, in: Genetic diversity in plants, Book Chapter published 14 Mar 2012 in *Genetic Diversity in Plants*. https://doi.org/10.5772/32885

Received: December 3, 2020; Published: January 8, 2021