Hue University Journal of Science: Natural Science
Vol. 134, No. 1S-2, 59–70, 2025 (Special Issue: CBA 2025)

pISSN 1859-1388
eISSN 2615-9678

Phylogenetic tree analysis of six strawberry (*Fragaria spp.*) cultivars using dna barcode

Tran Thi Yen Nhu[1], Nghi Khac Nhu[2], Le Van Thuc[3], Tran Thi Kim Nhu[2], Trinh Ngoc Ai[2]*,

[1] School of Agriculture and Aquaculture, Tra Vinh University, Tra Vinh, Vietnam
[2] Biotechnology Institution, Tra Vinh University, Tra Vinh, Vietnam
[3] Dalat Nuclear Research Institute, Vietnam Atomic Energy Institute, Da Lat, Vietnam

* Correspondence to Trinh Ngoc Ai <ngocai@tvu.edu.vn>

(*Received: 19 June 2025; Revised: 17 August 2025; Accepted: 19 August 2025*)

Abstract. We performed a phylogenetic analysis of six strawberry cultivars using chloroplast *rbc*L, *rpo*C2, and nuclear ribosomal ITS2 sequences. The objectives of this study were to analyze the phylogenetic relationship of 6 strawberry cultivars (Bach Tuyet Trang, Nhat Sky, Ha Lan Hoa Hong, New zealand, Han Goseul and Hana) based on *rbc*L, *rpo*C2, and ITS2 regions. The results compared to Nucleotide BLAST GenBank showed a high level of similarity (97.57-99.15%) for the *rbc*L sequences, 98.40-98.86% for the *rpo*C2 sequences, and 93.32-99.37% for the ITS2 sequences, all of which were similar to *Fragaria x ananassa*. Our study indicated that *rbc*L and nuclear ribosomal ITS2 sequences increased the efficiency of the phylogenetic analysis, while *rpo*C2 sequences did not provide sufficient clarity to confidently resolve the evolutionary history of strawberry cultivars. Phylogenetic analysis using DNA barcode markers (ITS2, *rbc*L and *rpo*C2) through Maximum Parsimony resulted in clades with a high bootstrap value: 99 for HN-HG and bootstrap value of 75 for NZ and HL cultivars. Our findings suggest that DNA barcoding is an efficient tool for identifying the genetic diversity of these six strawberry cultivars and highlights the potential for this study to contribute to the conservation, sustainable genetic resources and breeding program of the *Fragaria* species.

Keywords: DNA barcode, ITS2, phylogenetic relationship, *rbc*L, *rpo*C2, strawberry

1 Introduction

Strawberry (***Fragaria x* ananassa**) is a member of the Rosaceae (Rose) family, subfamily Rosoideae and is grown worldwide due to its versatility and high nutrient content. Most strawberries are cultivated for commercial purposes in over 60 countries, and originated from the hybridization of two local American species (*F. chiloensis X F. virginiana*) [1]. Different strawberries were easily recognizable by their size, fragrance, and red fruit and their exceptional vigor [2], on which basis they were brought into cultivation and breeding [1]. In addition, strawberries are packed with essential nutrients such as manganese, calcium, iron, magnesium and potassium. They are also a significant source of antioxidant compounds, boasting high levels of vitamin C, phenolic constituents and folate [3].

More than 20 species of strawberries were classified with various ploidy levels, and these wild strawberries (octoploid *F x ananassa*) were cultivated in Turkey for a long time, to generate strawberry genetic resources [4]. Global strawberry production occupied about 5% per year in the first two decades of the 20th century [5]. Asia is known as a continent that produces the most strawberries. In Vietnam, Da Lat is considered as a strawberry capital because of its cool climate (temperature around 15-24^{o}C), and a cornerstone of its agricultural economy and tourism. Bach

Tuyet Trang, Hana, Nhat Sky, Han Goseul, Ha Lan Hoa Hong, and New Zealand are strawberry cultivars commonly grown in Da Lat due to their superior fruit quality (large size, attractive color, and appealing flavor), stable yield, and adaptability to the year-round cool climate. They have high commercial value, are easy to market, and meet the demands of premium markets and tourism, thereby contributing to increasing farmers' incomes and enhancing the value of the local agricultural sector.

The strawberry Nhat Sky and Hana (Tochiotome) cultivars originated in Japan and were developed in the late 20th century as improved commercial varieties from Tochigi Prefecture [6, 8]. Hana cultivar was bred from a cross between Kurume No. 49 and Tochinomine, and is renowned for its large fruit size, sweetness, and high yield [7]. Bach Tuyet Trang is a South American cultivar that was once nearly extinct worldwide. In 2010, it was revived as a commercial variety by growers in the Netherlands and Belgium, and was later introduced to Japan for cultivation and improvement [9]. The New Zealand cultivar grown in Da Lat is considered a premium strawberry variety in the region, originating from America and subsequently bred and cultivated widely since the late 18th century [10]. The Netherlands is renowned for its breeding and development of strawberry cultivars, particularly since the 20th century. One notable cultivar is Ha Lan Hoa Hong, which may have been developed in the country. This variety is recognized for its high heat tolerance and deep pink flowers that resemble those of roses (family Rosaceae). [11]. Additionally, Han Goseul cultivar was developed in 2016 by the Highland Agriculture Research Institute [12] in South Korea through a cross between Albion, known for large fruits, and Seolhyang, a continuously flowering variety [13]. As a day-neutral cultivar, Han Goseul can produce fruit year-round under both long- and short-day conditions, with continuous flowering and large fruit production [12, 13].

An assessment of genetic divergences among species based on morphological characteristics can be inaccurate due to the influence of environmental factors. In recent years, molecular studies have gained popularity for providing insights into genetic differences through sequencing. DNA markers are useful tools for detecting variations in living organisms and play a vital role in distinguishing closely related species. Genetic diversity analysis of various strawberry varieties has been widely researched by comparing morphological characteristics and anatomy. Numerous studies have investigated genetic diversity using molecular techniques, including random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), ISSR-EST and simple sequence repeats (SSRs) [14-18]. The DNA barcode as an approach to identifying genetic variation and conservation [19]. This molecular technique is not influenced by developmental stages or environmental factors, as DNA can be easily extracted from all tissues, providing a basis for species identification at the genetic level [20].

The Internal Transcribed Spacer (ITS) region of the 18S-5.8S-26S nuclear ribosomal DNA (nrDNA) has used well as a phylogenetic marker in most groups of flowering plants [21], consists of internal transcribed spacer 1 (ITS1) region and internal transcribed spacer 2 (ITS2) region. The advantages of these regions are biparental inheritance, simplicity, sensitivity, universality, easy PCR amplification, and suitable amplification for phylogenetic analysis at the species or genetic level [22-24]. The ITS2 region is an effective DNA barcode because it is short, easily amplified, and genetically informative. The ITS2 is known as a valuable sequence tag for determining medicinal plants [25-27]. Consequently, it contains a great of

pISSN 1859-1388
eISSN 2615-9678

genetic information and is located in the nuclear region, this leads to use to overcome the issue of failure to amplify the ITS in some species and suitable for identifying low or high level taxa [28-30].

However, the presence of pseudogenes as well as the existence of orthologs and paralogs were considered as problems in evolution [31-33]. The chloroplast genome has been provided as a useful tool in identifying species and studying evolution of species [34, 35]. The chloroplast (cp) is important in various plant cell functions, including photosynthesis, and carbon fixation. Although the plant cp genome is highly conserved in sequence and gene content [36], the loss or mutation of genes can lead to different sequences. The cp genomes are inherited from the maternal parent and no recombination. They are widely used intraspecific and interspecific studies for evaluating genetic diversity, identifying species [37-39] or studying phylogenetic, taxonomy and evolutionary [40]. The slowly evolving *rbc*L gene was conserved to resolve phylogenetic relationships at the lower taxonomic level in Saxifragaceae and Asteraceae species [41,42]. Studies have identified the RNA polymerase C2 (*rpo*C2) gene as a highly variable region, and many studies were successfully tested the phylogenetic relationships in the Poaceae family [43-46]. To my knowledge, research phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus. Although previous studies on *Fragaria* phylogenetic relationships based on DNA sequences, the analysis of the combination between nuclear DNA (ITS2) and chloroplast DNA (*rbc*L + *rpo*C2) are limited. In this study, *rbc*L and *rpo*C2 chloroplast and nuclear DNA (ITS2) genomes representative of the genus *Fragaria* have been sequenced and their sequences compared in order to select a specific resources for practical applications.

2 Materials and Methods

2.1 Plant materials

The leaves of 6 strawberry cultivars including Bach Tuyet Trang (BT), Nhat Sky (NS), Ha Lan Hoa Hong (HL), Newzealand (NZ), Han Goseul (HG) and Hana (HN) were collected in November 2024 from the greenhouse of the Dalat Nuclear Research Institute. In the field, samples were put in plastic bags and kept in cool temperature, transferred to Biotechnology Institution, Tra Vinh University and stored -20^{o}C.

2.2 DNA extraction

Total genomic DNA from 0.1g of young leaf samples were extracted using DNA Dneasy Plant Pro kit (QIAGEN) according to the manufacturer's protocol. DNA samples were tested by nanodrop to assess the quality (Thermo Scientific™ NanoDrop™ One Microvolume UV-Vis spectrophotometer) and diluted to 100ng/µl concentration and then stored at -20^{o}C for further analysis.

2.3 DNA barcoding amplification and sequencing

A polymerase chain reaction (PCR) was performed with barcoding primers (**Table 1**), with 30 µl of the reaction mixture containing 15 µl Phanta flash master mix 1X (Vazyme), 0,6 µl DNA (100ng/µl), 1,2 µl each primer (10uM) and 12 µl of deionized water. PCR amplification was performed in a DNA thermal cycler, which was programmed for initial DNA denaturation at 95^{o}C for 5 mins, followed 40 cycles of 15 seconds denaturation at 95^{o}C, annealing at 55^{o}C – 60^{o}C for 60 seconds, and extension at 72^{o}C for 1 min, with a final extension at 72^{o}C for 10 mins.

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

Locus	Primer name	Sequences (5′-3′)	References
1	ITS2-F	ATGCGATACTTGGTGTGAAT	[47]
2	ITS2-R	GACGCTTCTCCAGACTACAAT	
3	*rbc*L-F	ATGTCACCACAAACAGAAAC	[48]
4	*rbc*L-R	TCGCATGTACCTGCAGTAGC	
5	*rpo*C2-F	GGAATTCGAAATTCTCCCGTTT	[49]
6	*rpo*C2-R	AGGGATAATCTAGAGCTTCGAGTTG	

To check the presence or absence of bands, amplified PCR products were electrophoresed using 2% agarose gel (1xTAE buffer and 1 µl 6x gelred loading buffer (ABT)). The gel was run at 100V for 40 mins, and then was visualized by the MultiDoc UVP (UVP GelStudio PLUS, AnalytikJena) system.

The PCR products were sent to Salagene company (https://salagene.com). To obtain the sequence of each region (*rbc*L, *rpo*C2, and ITS2), the forward and reverse sequences were aligned using Geneious [50]. In the searching for the similarities between those sequences and the sequences deposited in the GenBank database, the sequences of this study were analyzed using BLAST (Basic Local Aligment Search Tool) program at https://blast.ncbi.nlm.nih.gov/Blast

2.4 Phylogenetic analysis

The alignment was then exported to Molecular Evolutionary Genetics Analysis (Mega 12) software for phylogenetic tree analysis. The Maximum Parsimony trees were constructed for *rbc*L, *rpo*C2, and ITS2 data using the Kimura 2-parameter model with 1000 bootstrap replications for node supports and bootstrap support was categorized as strong (>85%), moderate (70–85%), weak (50–70%), or poor (<50%) [51].

3 Results and Discussion

3.1 Analysis of qualitative test

The electrophoresis images showed DNA bands in all samples, indicating successful isolation of DNA genomes from leaves of six strawberry plants. The PCR reactions were successfully amplified the *rbc*L, *rpo*C2 and ITS2 regions in all six cultivars, as shown by target bands, the length of the amplified product was from 600bp to 850bp (Fig. 1). The results of amplified target regions indicated the effective recognition of genomic DNA in all samples by the primers.

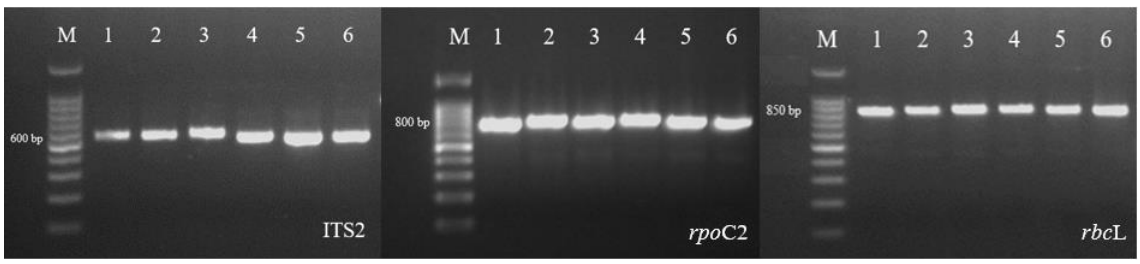

Fig. 1. The gel electrophoresis results confirm successful PCR amplification (clear bands at 600–850 bp for *rbc*L, *rpo*C2, and ITS2 across all six cultivars) Nucleotide of sequencing

pISSN 1859-1388
eISSN 2615-9678

The sequencing results were examined by using Geneious software to detect uncertainties and verify the sequences. Ambiguous nucleotides were excluded, and valid stop codons were authenticated. The high quality of the sequences was demonstrated by the clean alignment and well-resolved phylogenetic trees. In present study, DNA barcoding was used to distinguish between strawberry cultivars. The sequences were stored in Geneious for future alignment.

The length of ITS2 sequences ranged from 855bp to 879bp, following *rbc*L ranged from 456bp to 705bp, and *rpo*C2 sequences was from 369bp to 660bp, respectively. Nucleotide BLAST was applied to check the similarities ranging from 97.57-99.15% with *Fragaria x ananassa* for *rbc*L; 98.40-98.86% with *F. x ananassa* for ITS2 and 93.32-99.37% with *F. x ananassa* for *rpo*C2 (**Table 2**). Sequence similarities between 98-100% generally support classification within the same species, indicating that the specimens tested likely belong to *F. x ananassa* cultivars.

Table 2. Nucleotide sequences were deposited in the NCBI GenBank

Cultivars	Primers	NCBI Accession Number		Cover (%)	Similarity index (%)	Sequencing size (bp)
BT	ITS2	PQ836331	*F. x ananassa*	97	98.86	870
HG		PQ836331	*F. x ananassa*	99	98.64	879
HL		PQ836329	*F. x ananassa*	98	98.61	869
HN		PQ836331	*F. x ananassa*	98	98.40	870
NS		PQ836331	*F. x ananassa*	95	98.63	855
NZ		PQ836331	*F. x ananassa*	99	98.64	862
BT	*rbc*L	JX118093	*F. x ananassa*	95	98.18	686
HG		JX118093	*F. x ananassa*	96	97.99	456
HL		JX118093	*F. x ananassa*	99	99.15	705
HN		JX118093	*F. x ananassa*	96	98.67	691
NS		JX118093	*F. x ananassa*	99	97.86	697
NZ		JX118093	*F. x ananassa*	99	97.57	692
BT	*rpo*C2	ON478181	*F. x ananassa*	96	96.89	486
HG		ON478181	*F. x ananassa*	98	93.32	399
HL		ON478181	*F. x ananassa*	100	94.93	604
HN		ON478181	*F. x ananassa*	95	94.57	369
NS		ON478181	*F. x ananassa*	96	99,37	660
NZ		ON478181	*F. x ananassa*	99	98,41	626

3.2 Molecular characters of the *Fragaria* species

The sequence characteristics and parsimony-based tree statistics of the tree ITS2 and cpDNA regions were shown in Table 3. The aligned sequences originated from all cpDNA regions and the ITS2 revealed differences in the sequence length between the *Fragaria* cultivars. The maximum-parsimony (MP) analysis showed the number of variable and parsimony informative sites. The length of *rpo*C2 region is the smallest one among the molecular markers, compared to *rbc*L and ITS2 regions. The percentage of informative characters in *rbc*L region is 1.89%; following ITS2 region (1.58%); *rpo*C2 region (1.55%); cpDNA+ITS2 (2.85%) and *rbc*L+ *rpo*C2 region (4.7%). A total 908-912 characters were generated when combining cpDNA (*rbc*L+ *rpo*C2) and ITS2 sequences, and 26 characters were parsimony- informative. The heuristic search produced 161 steps with 784 of constant and 871 of parsimony – uninformative. Maximum parsimony analysis resulted in CI of 0.931; RI of 0.592; and RC of 0.551.

Table 3. Sequence characteristics and tree statistic of the cpDNA and ITS2 regions from the maximum-parsimony (MP) analysis

Characteristics	cpDNA		nrDNA ITS2	Combined cpDNA	Combined cpDNA and ITS2
	*rbc*L	*rpo*C2			
LAS (bp)	710	664	508	720	912
TL	64	89	45	174	161
PICs*	11(1.89%)	8(1.55%)	8(1.58%)	27(4.70%)	26(2.85%)
CI	0.922	0.955	0.978	0.983	0.931
RI	0.545	0.500	0.875	0.903	0.592
RC	0.502	0.477	0.856	0.888	0.551
Constant	652	548	470	529	784
Parsimony-uninformative	692	616	495	660	871

LAS: length of aligned sequences; TL: tree length; PICs: parsimony-informative characters; CI: consistency index; RI: retention index; RC: rescaling consistency index.

3.3 Phylogenetic tree analysis

Many researches indicated that DNA barcoding technology is used to classify the known or unknown species of berry fruit products. It provides much accurate information regarding species which should be recognized. In this study, six strawberry cultivars were identified through molecular markers (ITS2 and cpDNA regions) which further helped in classification, identification and evaluation of *Fragaria x ananassa* species. A phylogenetic tree was created by using the Maximum Parsimony (MP) with 1,000 bootstrap replicates. This approach offers insights into evolutionary connections, where the length of branchs indicates the level of genetic variation.

For ITS2 marker, although the phylogenetic tree analysis had lower steps (TL = 45 steps), consistency index (CI=0.978), retention index (RI=0.875), rescaling consistency index (RC=0.856) has the highest values. The phylogenetic tree was divided into three clades: clade I was divided into 2 sub-clades. Sub-clade I.1 included HG and HL

pISSN 1859-1388
eISSN 2615-9678

cultivars with strong bootstrap (86%) and NZ specie was placed together in clade I. NS and BT species were placed in clade II with a moderate bootstrap (82%), while HN specie was placed out of clade I and clade II as it showed diversity (Fig. 2). Previous studies showed the ribosomal DNA ITS2 region was used as a tag to identify species, has gained a lot of attention recently [52]. Kress at al [55] indicated that ITS2 region is a powerful tool in identifying Rosaceae family and successfully identified 78 and 100% of them at the species and genus levels, respectively. [25] Proposed the use of the ITS2 locus as a universal barcode for all major plant taxa in traditional herbal medicine. Indeed, ITS2 has already been suggested as a suitable marker applicable for taxonomic classification and phylogenetic reconstruction in eukaryotes by many researchers because of its advantages such as short nuclear region, strong universality, low intraspecific variance [53, 54, 49].

The phylogenetic tree constructed using the *rbc*L marker through the maximum parsimony method, with bootstrap support, revealed 64 steps and had the following indices: CI = 0.922; RI = 0.545; RC = 0.502 (Table 3). The phylogenetic tree from *rbc*L sequences showed that BT, NH and NS species in cluster I formed clearly distinctive clades with a weak bootstrap 58%. Within this group, BT and NH are shown to be the most closely related, forming a subclade. NS is positioned as a sister taxon to the BT-NH clade. Clade II comprised NZ and HG species, which are depicted as being very closely related (strong bootstrap 95%). The taxon HL is clearly distinct from the others, representing a more distant relative within these species (Fig. 2). Previous studies indicated that the sequence data from *rbc*L gene is used for studying phylogenetic relationships among other plant genes because it has some advantages such as the slow rate of evolutionary changes and shows the least amount of divergence among plastid genes in flowering plants [56, 57]. Reddy and Li eat al [58, 59] showed this region is highly suitable for resolving relationships between genera and species, and no difficulties of alignment.

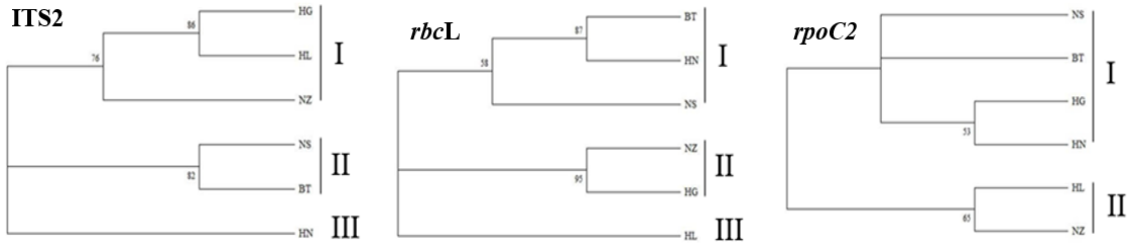

Fig. 2. Phylogenetic tree reconstruction of six samples from ITS2, *rpo*C2 and *rbc*L regions using Maximum Parsimony technique with 1,000 bootstrap replicates

For *rpo*C2 marker, the phylogenetic tree constructed through the maximum parsimony method by using the bootstrap mode had 89 steps, CI = 0.955; RI = 0.500; RC = 0.477 (Table 3). The phylogenetic tree analysis is divided into two primary clades. The first clade included NS, BT, HG and HN taxa. In this clade, the relationships between HG and HN were solved with weak bootstrap value (53%). It indicated that data are not reliable group these two species together. The HL and NZ clade is a stronger relationship than the HG-HN pairing, with a weak to moderate bootstrap values (65). The analysis of the *rpo*C2 gene suggests a specific set of evolutionary relationships. It is shown that HL and NZ grouped together and separately to HG and HN together. However, the statistical support for these groupings is low, particularly for the HG and HN pair (53) (Fig. 2). Our results are different with previous researches on *rpo*C2 [60, 61]. Lin et al. [60] indicated that horizontal gene transfer of cp genes from *Haloxylon ammodendron* species to

Holoparasite *Cistanche deserticola* species through the phylogenetic evidence that was built base on *rpo*C2 sequences. Their phylogenetic tree showed a very high bootstrap value, especially is under species level (almost 100%). Study of Gomolińska et al. [61] also indicated the same high bootstrap value of phylogenetic tree base on *rpo*C2 gene sequences. This suggests that the *rpo*C2 gene alone may not provide sufficient information to confidently resolve the evolutionary history of strawberry cultivars. This failure may be due to incomplete lineage sorting or the absence of a "barcode grap", and related to large effective population size as well as a slow species evolutionary rate [62]. Moreover, inadequate sampling within a genus rich in species and the absence of agreement on suggested markers also lead to errors in identifying closely related species [63].

3.4 Phylogenetic tree based on ITS2 and cpDNA data

The use of the *rbc*L combination with *rpo*C2 increased the efficiency of the barcode analysis. The phylogenetic tree constructed through the maximum parsimony contained 174 steps of tree length, a consistency index of 0.983; retention index of 0.903; rescaling consistency index of 0.888 (Table 3). In the phylogenetic tree, HG, HN, BT and HL species were part of clade I, while NS and NZ species were placed in clade II and clade II, respectively (Fig. 3)

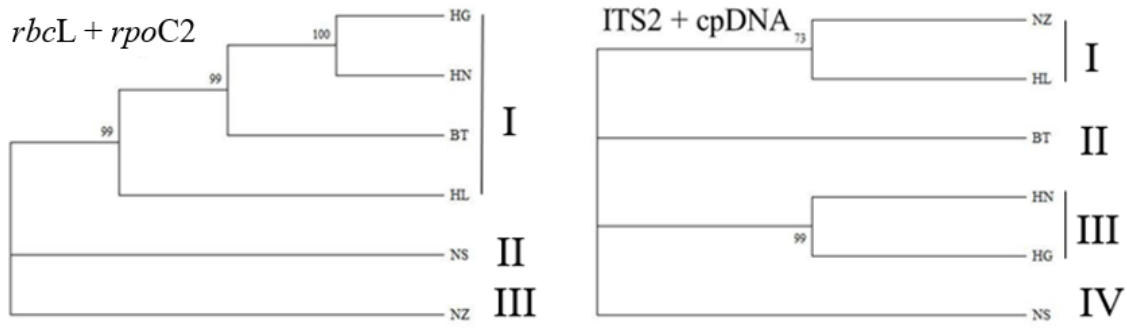

Fig. 3. Phylogenetic tree reconstruction of six samples from ITS2, *rpo*C2 and *rbc*L genes using Maximum Parsimony technique with 1,000 bootstrap replicates

The maximum parsimony tree was built based on the combination between ITS2 region and cpDNA (*rbc*L+ *rpo*C2) divided into four clades (Fig. 3). A closely related group including NZ and HL formed a distinct group with moderate bootstrap values (73%), and a very high genetic relationship was found in HN and HG with strong bootstrap (99%). The BT clade appeared to be on a separate evolutionary line from the NZ-HL and HN-HG-NS groups, while the NS clade is distinct from all the other groups. This result indicated that the phylogenetic relationships among six different plant have diversified into at least four distinct genetic groups, with two of those groups containing pairs of more closely related populations. The use of both nuclear (ITS2) and chloroplast DNA (*rbc*L and *rpo*C2) has made genetic relationship between selected strawberry cultivars become more apparent. The similar results were reported that ITS is more successful in genetic diversity analyses of more asunder populations, in contrasts to ISSR-PCR and RAPD-PCR methods, ITS method provided more clear data in regional genetic discrimination of populations [64].

pISSN 1859-1388
eISSN 2615-9678

4 Conclusion

The results of this study showed that DNA barcoding is an effective technique for identifyingand resolving phylogenetic relationships among six strawberry cultivars. The genetic diversity of selected cultivars showed similarities with the data available in GenBank. Specifically, the *rbc*L and ITS2 regions were found to be highly suitable for assessing genetic diversity at both the genus and species level. The combination of *rbc*L and *rpo*C2 along with all used DNA barcode markers used (ITS2, *rbc*L and *rpo*C2) increased the efficiency of the phylogenetic analysis. Our results also indicated that there are two pairs of strawberry cultivars (HN-HG and NZ-HL) that are closely related genetically consistently clustering in the same clade across almost all phylogenetic analyses. However, combining more DNA barcoding markers and increasing the number of collected samples could further clarify the phylogenetic relationships within the *Fragaria* species. This information will be valuable for the development and utilization of ornamental strawberries, as well as for the selection and breeding of new cultivars.

Acknowlegdment

The study were funded by Tra Vinh University under the contract number 61/2024/HĐ.HĐKH& ĐT-ĐHTV.

Declarations

Conflict of interest all authors have no interests to declare.

References

1. Hancock JF, Strawberries. Wallingford: CABI Publishing; 1999
2. Darrow G. The strawberry. History, breeding and physiology. New York: Holt, Rinehart, Winston; 1966. xvi + 447 pp. p.
3. Häkkinen SH, Törrönen AR. Content of flavonols and selected phenolic acids in strawberries and Vaccinium species: influence of cultivar, cultivation site and technique. Food Research International. 2000;33(6):517-24.
4. Ozturk E S, Cekic C. The genetic relationship between some strawberry cultivars and F1 population derived from these cultivars. International Society for Horticultural Science (ISHS). 2020.
5. Jofre-y-Garfias AE, Dávalos-González PA, Aguilar-García R, Rodríguez-Guillén A. The Genetic Diversity of Strawberry Species, the Underutilized Gene Pool and the Need for Cultivars with New Quality and Agronomic Attributes. In: Kafkas NEY, Oğuz İ, editors. Recent Studies on Strawberries. London: IntechOpen; 2022.
6. Specialty Produce. Tochiotome strawberries information and facts [Internet]. San Diego (CA): Specialty Produce [cited 2025 Aug 14]. Available from: https://www.specialtyproduce.com/produce/Tochiotome_Strawberries_15084.php
7. Eats.jp. Tochiotome was registered as a variety in 1996, born from Kurume No. 49 × Tochi no Mine [Internet]. [cited 2025 Aug 14]. Available from: https://eats.jp/en/foods/10323
8. Amazon Japan. Tochigi Prefecture Skyberry Strawberry [Internet]. [cited 2025 Aug 14]. Available from: https://www.amazon.co.jp/-/en/Tochigi-Prefecture-Skyberry-Strawberry-Deluxe/dp/B0CZ45NPS8
9. Dalat Tourism Promotion Center. Snow White strawberry – The most delicious and rare strawberry in the world [Internet]. Dalat: Dalat.vn; c2024 [cited 2025 Aug 14]. Available from: https://dalat.vn/vi/detailnews/?id=news_187&t=snow-white-strawberry-the-most-delicious-and-rare-strawberry-in-the-world
10. EU-Vietnam Business Network. New Zealand strawberries – Origin, provenance & wonderful health benefits [Internet]. 2024 [cited 2025 Aug 14]. Available from: https://evbn.org/dau-tay-new-zealand-1678662139
11. Rowe K. 7 strawberry varieties with charming pink flowers [Internet]. Epic Gardening; 2025 Apr 29 [cited 2025 Aug 14]. Available from:

https://www.epicgardening.com/strawberry-pink-flowers/

12. Highland Agriculture Research Institute. Development of off-season strawberry varieties Goseul and Miha [Internet]. National Institute of Crop Science; 2024 [cited 2025 Aug 14]. Available from: https://www.nics.go.kr/u/700001490.do
13. Lee J, Suh J, Nam J, Hong S, Kim S, Shon H, et al. Breeding of New Day-neutral Strawberry 'Goseul'. Journal of the Korean Society of International Agricultue. 2020;32:42-6.
14. Azizah UDL, Yulianti F, Adiredjo AL, Sitawati D. Genetic relationship analysis of strawberry (Fragaria sp.) based on morphology character and Random Amplified Polymorphic DNA (RAPD). Plantropica Journal of Agricultural Science. 2019;4(1):77-85.
15. Corrêa, Jessica Vanessa Wosniak, Weber, G. G., Zeist, A. R., de Resende, J. T. V., & Da-Silva, P. R. . ISSR analysis reveals high genetic variation in strawberry three-way hybrids developed for tropical regions. Plant Molecular Biology Reporter, 2021,39(3):566-576.
16. Debnath SC. Development of ISSR markers for genetic diversity studies in Vaccinium angustifolium. Nordic Journal of Botany. 2009;27 (2):141-148.
17. Lim SH, Lee JY, Lee HJ, Park KH, Kim DS, Min SR. The genetic diversity among strawberry breeding resources based on SSRs. Scientia Agricola, 2017;74(3):226-234.
18. Clark JR, Stafne ET, Hall HK, Finn CE. Blackberry breeding and genetics. Plant breeding reviews, 2007;29:19-152.
19. Zhang HX, Zhang ML. Spatial patterns of species diversity and phylogenetic structure of plant communities in the Tianshan Mountains, arid Central Asia. Frontiers in Plant Science. 2017;8: 2134.
20. Yu N, Wei YL, Zhang X, Zhu N, Wang YL, Zhu Y, et al. Barcode ITS2: a useful tool for identifying Trachelospermum jasminoides and a good monitor for medicine market. Scientific Reports. 2017;7(1): 5037.
21. Baldwin BG. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Molecular phylogenetics and evolution. 1992;1(1):3-16.
22. Hamby RK, Zimmer EA. Ribosomal RNA as a phylogenetic tool in plant systematics. In: Molecular systematics of plants. Boston: Springer US; 1992. p. 50-91.
23. Álvarez I, Wendel JF, Jonathan F. Ribosomal ITS sequences and plant phylogenetic inference. Molecular phylogenetics and evolution, 2003;29(3):417-434.
24. Ansari S, Solouki M, Fakheri B, Fazeli-Nasab B, Mahdinezhad N. Assessment of molecular diversity of internal transcribed spacer region in some lines and landrace of Persian Clover (*Trifolium resupinatum* L.). Potravin Slovak J Food Sci, 2018;12(1):657-666.
25. Chen S, Yao H, Han J, Liu C, Song J, Shi L, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PloS one. 2010;5(1):e8613.
26. Yao H, Song J, Liu C, Luo K, Han J, Li Y, et al. Use of ITS2 region as the universal DNA barcode for plants and animals. PloS one. 2010;5(10):e13102.
27. Han J, Zhu Y, Chen X, Liao B, Yao H, Song J, et al. The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. BioMed research international 2013;2013(1):741476.
28. Petit RJ, Excoffier L. Gene flow and species delimitation. Trends in Ecology & evolution. 2009;24(7):386-393.
29. Naciri Y, Caetano S, Salamin N. Plant DNA barcodes and the influence of gene flow. Molecular Ecology Resources. 2012;12(4):575-580.
30. Braukmann TWA, Kuzmina ML, Sills J, Zakharov EV, Hebert PDN. Testing the efficacy of DNA barcodes for identifying the vascular plants of Canada. PLoS One. 2017;12(1):e0169515.
31. Soltis DE, Bell CD, Kim S, Soltis PS. Origin and early evolution of angiosperms. Annals of the New York Academy of Sciences. 2008;1133(1):3-25.
32. Feliner GN, Rosselló JA. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Molecular Phylogenetics and Evolution. 2007;44(2):911-919.
33. Bailey PC, Martin C, Toledo-Ortiz G, Quail PH, Huq E, Heim MA, et al. Update on the basic helix-loop-helix transcription factor gene family in Arabidopsis thaliana. The Plant Cell. 2003;15(11):2497-2502.
34. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: the coding *rbc*L gene

pISSN 1859-1388
eISSN 2615-9678

complements the non-coding trnH-psbA spacer region. PLoS One. 2007;2(6):e508.

35. Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, et al. DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences USA. 2008;105(8):2923-8.

36. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biology. 2009;7(1):84.

37. Qiao J, Cai M, Yan G, Wang N, Li F, Chen B, et al. High-throughput multiplex cpDNA resequencing clarifies the genetic diversity and genetic relationships among Brassica napus, Brassica rapa and Brassica oleracea. Plant Biotechnology Journal. 2016;14(1):409-418

38. Jiao LC, Lu Y, He T, Li JN, Yin YF. A strategy for developing high-resolution DNA barcodes for species discrimination of wood specimens using the complete chloroplast genome of three Pterocarpus species. Planta. 2019;250:95-104.

39. Liu ZF, Ma H, Ci XQ, Li L, Song Y, Liu B, et al. Can plastid genome sequencing be used for species identification in Lauraceae? Botanical Journal of the Linnean Society. 2021;197(1):1-14.

40. Jheng C-F, Chen T-C, Lin J-Y, Chen T-C, Wu W-L, Chang C-C. The comparative chloroplast genomic analysis of photosynthetic orchids and developing DNA markers to distinguish Phalaenopsis orchids. Plant Science. 2012;190:62-73.

41. Kim KJ, Jansen RK, Wallace RS, Michaels HJ, Palmer JD. Phylogenetic implications of *rbc*L sequence variation in the Asteraceae. Annals of the Missouri Botanical Garden. 1992;79:428-445.

42. Soltis DE, Kuzoff RK, Conti E, Gornall R, Ferguson K. *mat*K and *rbc*L gene sequence data indicate that *Saxifraga* (Saxifragaceae) is polyphyletic. American Journal of Botany. 1996;83(3):371-82.

43. Cummings MP, King LM, Kellogg EA. Slipped-strand mispairing in a plastid gene: *rpo*C2 in grasses (Poaceae). Molecular Biology and Evolution. 1994;11(1):1-8.

44. Barker NH, Linder P, Harley E. Sequences of the grass-specific insert in the chloroplast *rpo*C2 gene elucidate generic relationships of the Arundinoideae (Poaceae). Systematic Botany. 1999;23(3):327-336.

45. Duvall MR, Doebley JF. Restriction site variation in the chloroplast genome of *Sorghum* (Poaceae). Systematic Botany. 1990;15(3):472-480.

46. Moon JC, Kim JH, Jang CS. Development of multiplex PCR for species-specific identification of the Poaceae family based on chloroplast gene, rpoC2. Applied Biological Chemistry. 2016;59(2):201-207.

47. Gu W, Song J, Cao Y, Song J, Sun Q, Yao H, et al. Application of the ITS2 region for barcoding medicinal plants of Selaginellaceae in Pteridophyta. PLoS One. 2013;8(6):e67818.

48. Song J, Yao H, Li Y, Yao H, Li Y, Li X, et al. Authentication of the family Polygonaceae in Chinese pharmacopoeia by DNA barcoding technique. Journal of Ethnopharmacology. 2009;124 3):434-439.

49. Njuguna W, Liston A, Cronn R, Ashman TL, Bassil N. Insights into phylogeny, sex function and age of Fragaria based on whole chloroplast genome sequencing. Molecular Phylogenetics and Evolution. 2013;66(1):17-29.

50. Karp A, Institute IPGR. Molecular Tools in Plant Genetic Resources Conservation: A Guide to the Technologies: IPGRI; 1997.

51. Kress WJ, Prince LM, Williams KJ. The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. American Journal of Botany. 2002;89(10):1682-1696.

52. Potter D, Gao F, Bortiri PE, Oh SH, Baggett S. Phylogenetic relationships in Rosaceae inferred from chloroplast *mat*K and *trn*L-*trn*F nucleotide sequence data. Plant Systematics and Evolution. 2002;231(1-4):77-89.

53. Mosa KA, Gairola S, Jamdade R, El-Keblawy A, Al Shaer KI, Al Harthi EK, et al. The promise of molecular and genomic techniques for biodiversity research and DNA barcoding of the Arabian Peninsula Flora. Frontiers in Plant Science. 2019;Volume 9 - 2018.

54. Raclariu AC, Heinrich M, Ichim MC, De Boer H. Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication. Phytochemical Analysis. 2018;29(2):123-128.

55. Pang X, Song J, Zhu Y, Xu H, Huang L. Chen S. Applying plant DNA barcodes for Rosaceae species identification. Cladistics. 2011;27(2):165-170.

56. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences. 2005;102(23):8369-8374.

57. Ahmed I, Biggs PJ, Matthews PJ, Collins LJ, Hendy MD, Lockhart PJ. Mutational dynamics of aroid chloroplast genomes. Genome Biology and Evolution. 2012;4(12):1316-1323.

58. Reddy BU. Cladistic analyses of a few members of Cucurbitaceae using *rbc*L nucleotide and amino acid sequences. International Journal of Bioinformatics Research. 2009;1:58-64

59. Li X, Zhang T-C, Qiao Q, Ren Z, Zhao J, Yonezawa T, et al. Complete Chloroplast Genome Sequence of Holoparasite *Cistanche deserticola* (Orobanchaceae) Reveals Gene Loss and Horizontal Gene Transfer from Its Host *Haloxylon ammodendron* (Chenopodiaceae). PLOS ONE. 2013;8(3):e58747.

60. Lin C-S, Chen JJW, Huang Y-T, Chan M-T, Daniell H, Chang W-J, et al. The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. Scientific Reports. Scientific Reports. 2015;5(1):9040.

61. Gomolińska AM, Szczecińska M, Sawicki J, Krawczyk K, Szkudlarz P. Phylogenetic analysis of selected representatives of the genus Erica based on the genes encoding the DNA-dependent RNA polymerase I. Biodiversity: Research and Conservation. 2017;46:1-18.

62. van Velzen R, Weitschek E, Felici G, Bakker FT. DNA Barcoding of Recently Diverged Species: Relative Performance of Matching Methods. PLOS ONE. 2012;7(1):e30490.

63. Ran J-H, Wang P-P, Zhao H-J, Wang X-Q. A Test of Seven Candidate Barcode Regions from the Plastome in Picea (Pinaceae). Journal of Integrative Plant Biology. 2010;52(12):1109-26.

64. Poyraz I. Comparison of ITS, RAPD and ISSR from DNA-based genetic diversity techniques. Comptes Rendus. Biologies. 2016;339(5-6):171-8.