http://www.pjbs.org

PJBS

ISSN 1028-8880

Pakistan Journal of Biological Sciences

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

OPEN ACCESS Pakistan Journal of Biological Sciences

ISSN 1028-8880
DOI: 10.3923/pjbs.2025.121.130

Research Article
Genetic Variation and Phylogenetic Analysis of Strawberry (*Fragaria* spp.) on Yogyakarta and Central Java, Indonesia, Based on *rbc*L DNA Barcoding

[1]Ganies Riza Aristya, [1]Muhammad Sena Budiman, [2]Rina Sri Kasiamdari, [3]Ani Widiastuti and [4]Muhammad Fauzi Arif

[1]Laboratory of Genetics and Breeding, Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[2]Laboratory of Plant Systematics, Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[3]Laboratory of Pests and Plant Diseases, Department of Plant Protection, Faculty of Agriculture, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[4]Laboratory of Biology, Department of Biology, Faculty of Mathematics and Life Sciences, Mulawarman University, Samarinda, 75123, Indonesia

Abstract

Background and Objective: Strawberry (*Fragaria* spp.) is known for producing fruit with high economic value and significant nutritional content. Recently, the growing diversity of cultivated strawberries in Indonesia has made it challenging to distinguish the original characteristics of early ancestors and identify superior traits. The DNA barcoding, mainly through the chloroplast gene *rbc*L, offers a precise and detailed method for this identification. This research aims to reconstruct a phylogenetic tree, analyze genetic variation and determine the haplotype distribution of six strawberry cultivars from Java, particularly Yogyakarta and Central Java, based on the *rbc*L gene. **Materials and Methods:** The *rbc*L gene was amplified using DNA amplification techniques with *rbc*L-F and *rbc*L-R primers. The resulting data were analyzed to construct a phylogenetic tree using ML via IQtree software and BI using MrBayes software. The alignment results were used to determine genetic distances and identify polymorphic sites. This study assessed intraspecific genetic variation by examining h, identifying polymorphic sites, generating a haplotype network using PopART v1.7 and conducting PCoA with GenAlEx 6.503. **Results:** The results showed that the *rbc*L gene was successfully amplified with a length of 1,221 bp after alignment with the GenBank database. Phylogenetic analysis using ML revealed that the six cultivars formed a single clade with a bootstrap value of 97. BI similarly indicated the formation of one clade with a posterior probability value of 1. Haplotype analysis showed that the cultivars 'Californica', 'Knia', 'Mencir', 'Moha' and 'Geolhyang' belonged to the same haplotype group, while the 'Bali × Jumbo' cultivar was placed in a different group. **Conclusion:** Haplotype network analysis and PCoA further indicated that the genetic variation of Indonesian strawberries, as assessed through the *rbc*L gene, is similar to strawberries from the United States and China.

Key words: Bayesian inference, DNA barcoding, genetic distance, maximum-likelihood, *rbc*L

Citation: Aristya, G.R., M.S. Budiman, R.S. Kasiamdari, A. Widiastuti and M.F. Arif, 2025. Genetic variation and phylogenetic analysis of strawberry (*Fragaria* spp.) on Yogyakarta and Central Java, Indonesia, based on *rbc*L DNA barcoding. Pak. J. Biol. Sci., 28: 121-130.

Corresponding Author: Ganies Riza Aristya, Laboratory of Genetics and Breeding, Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, Indonesia

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Strawberry (*Fragaria* spp.) is one of the plants that produce fruits with high economic value and nutritional content[1]. According to FAOSTAT (Food and Agriculture Organization of the United Nations) data in 2020, Global strawberry fruit production amounted to 14 billion US dollars[2]. In Indonesia, strawberry production reached 8,350 ton in 2020, increased to 9,860 ton in 2021 and sharply rose to 28,895 ton in 2022. Asia is the continent that produces the most strawberries[3]. The strawberries successfully cultivated in Indonesia resulted from cultivation abroad, with 23 species and cultivars planted in the highlands of Indonesia[4,5]. These cultivated varieties exhibit varying morphological characteristics, yields, pest resistance and productivity.

Analysis of the diversity of various strawberry cultivars has been widely conducted by comparing their morphology and anatomy. Genetic variation in strawberry plants has also been studied using RAPD and ISSR techniques[1,6]. However, research on analysis relationships inferred by phylogenetic and genetic variation at the gene level in strawberries has not been extensively studied. This resulted in gaps and unclear identities regarding strawberry evolution and domestication lines. Therefore, DNA barcoding analysis is needed because it can examine the level of gene character differences and explain their evolutionary lines. The DNA barcodes as an approach to identifying genetic variation and new species, since having short gene sequences from standard regions of a genome[7]. The commonly used genome is cpDNA because it is highly conserved with low nucleotide substitution rates. The cpDNA is circular, containing two types of RNA, tRNA and rRNA, which play a role in regulating mRNA production and most of the proteins found in the chloroplast[8]. Each subunit forming the photosynthetic protein complex contains genetic information, including one of them, ribulose 1,5-bisphosphate carboxylase oxygenase (*rbc*L). The *rbc*L gene encodes the Rubisco protein[9,10]. The advantage of the *rbc*L gene is capable amplify DNA with a high success rate using one or two types of universal primers[11]. On the whole, the *rbc*L gene had approximately 1400 bp long, providing many characters for phylogenetic studies[8]. The *rbc*L gene has a relatively slow mutation rate compared to other genes used as barcodes for cpDNA. As a result, species tend to exhibit a higher level of similarity. This lower mutation rate is advantageous for conducting detailed examinations of genetic and phylogenetic variations within species[10]. This research concentrates on creating a phylogenetic tree, analyzing genetic diversity and identifying the distribution of haplotypes within six varieties of strawberries from Java, particularly from Yogyakarta and Central Java, using the *rbc*L gene.

MATERIALS AND METHODS

Sample collection and species identification: This study was carried out from November, 2023 to April, 2024, the sample collection of six strawberry leaf samples were taken from 2 locations (Table 1), UPTD BP2TPH DIY unit Ngipiksari (-7.61343110457552, 110.42648776441945) and Inggit Strawberry Agrotourism (-7.472920285446705, 110.39358789519403). The samples were placed into Ziplock bags, stored in a cool box and transported to the Genetics and Breeding Laboratory, Faculty of Biology, Universitas Gadjah Mada. The samples were kept at a temperature of -20°C for sample preservation. To expand this research, 14 *rbc*L gene sequences were obtained from GenBank, which had been previously documented in countries such as China, the United States, Switzerland and Japan (Table 1). The purpose of including these samples was to improve the comprehension of the genetic variability and connections among diverse strawberry populations on the Island of Java.

DNA extraction and quantitative-qualitative analysis: Genomic DNA was extracted from 50 mg of strawberry leaf tissue using the Genomic DNA Mini Kit (Geneaid) according to the manufacturer's instructions. The purity and concentration of the extracted DNA were then assessed using a Nanodrop spectrophotometer (Thermo Scientific, Loughborough, Leicestershire, United Kingdom) at an A260/280 wavelength,

Table 1: Strawberry (*Fragaria*× *ananassa*) sample collections used in this study

Cultivars	Sample code	Collection time	Location sites	Latitude	Longitude
Californica	CL/S13	December 8, 2023	UPTD BP2TPH DIY, Sleman, Special Region of Yogyakarta	7°36'48.348"S	10°25'35.3568"E
Knia	KN/S14	December 8, 2023			
Mencir	MR/S15	December 28, 2023	Strawberry Garden Inggit 3, Banyuroto Village, Magelang, Central Java	7°28'22.512"S	110°23'36.9168"E
Moha	MK/S16	December 28, 2023			
Geolhyang	GL/S17	December 28, 2023			
Jumbo×Bali	Bali_J/S18	December 28, 2023			

with an optimal purity ratio ranging between 1.8 and 2.0[12]. To evaluate the DNA quality, electrophoresis was performed using an electrophorator (Mupid-exU, Chuo-ku, Tokyo, Japan) to separate and measure the DNA fragments. The samples were mixed with loading dye and loaded into wells containing 0.8% agarose gel. Electrophoresis was conducted at 50 V for 30 min in a 10X TBE buffer solution and the gel was subsequently visualized under ultraviolet light.

DNA amplification, electrophoresis and sequencing of *rbc*L: The *rbc*L gene was amplified using a Biorad Thermal Cycler. One pair of chloroplast primers was used: *rbc*L-F (5'-ATGTCACCACAAACAGAAAC-3') and *rbcL*-R (5'-CTATAAGGTATCCATCGCT-3'). Each PCR reaction had a total volume of 30 µL, including 5 µL of genomic DNA, 12.5 µL of MyTaq HS Red Mix (Bioline), 1 µL of 10 µM primer and 7.5 µL of ultrapure water (ddH_2O). The process of DNA amplification started with an initial denaturation for 5 min at 95°C. Following this, 35 cycles included denaturation at 95°C for 30 sec, annealing at 54.1°C for 45 sec and extension at 45°C for 30 sec. Finally, there was a final extension at 72°C for 5 min. After that, the products were analyzed by electrophoresis on a 0.8% agarose gel, stained with Florosafe dye (Bioline) and run in 10X TBE buffer at 100 V for 30 min. The gel was then visualized under ultraviolet light.

Data analysis

Sequence editing and alignment of *rbc*L: Ambiguous bases in the DNA sequences were manually corrected using GeneStudio Pro v2.2. To confirm the strawberry species, the consensus sequences were compared to existing data in GenBank through the Nucleotide BLAST program. Six *rbc*L sequences from strawberry samples were aligned, yielding a fragment length of 1,221 bp. These *rbc*L sequences were then used for further intraspecific genetic analysis. In constructing the phylogenetic tree, two additional samples were included as outgroups, though this did not alter the fragment length, which remained at 1,221 bp[12].

Phylogenetic and genetic distance analysis: For phylogenetic reconstruction, 20 samples were analyzed, including two outgroup samples from GenBank: *Potentilla sterilis* (KF602156) and *Alchemilla glabra* (KF602207). The phylogenetic tree was created by implementing the ML method and the partition parameters were established using Partition Finder 2.1.1. One thousand bootstrap replicates were utilized in IQtree2. In addition, BI analysis was carried out using the MrBayes v3.2.6 software. Markov Chain Monte Carlo (MCMC) method was utilized for 4 million generations, with sampling conducted every 1,000 generations to estimate the posterior probability distribution. The resulting consensus tree was visualized using FigTree v1.4.4. Genetic distance analysis was conducted using the K-2P model in MEGA11 software.

Genetic variation, haplotype network and Principal Coordinate of Analysis (PCoA): This study analyzed intraspecific genetic variation by examining h, identifying polymorphic sites and generating a haplotype network using PopART v1.7. PCoA was also conducted using GenAlEx 6.503[12].

RESULTS

Analysis of quantitative test: The solution obtained from the DNA extraction process was clear and colorless and the concentrations and purity of the genomic DNA extracted were summarized in Table 2. Genomic DNA concentrations above 100 ng/µL are considered adequate for successful amplification, with a minimum requirement of 25 ng/µL. Typically, DNA purity values range from 1.8 to 2.0, indicating the absence of contaminants such as polysaccharides, proteins, phenols or RNA molecules. In this investigation, the 'Californica' variety exhibited the highest genomic DNA concentration at 56.4 ng/µL, while the 'Geolhyang' variety had the lowest concentration at 15.2 ng/µL. The lower concentration in 'Geolhyang' may be due to phenolic compounds, overly mature leaf samples or DNA degradation during extraction. Regarding purity, 'Geolhyang' demonstrated the highest value at 1.79, while 'Californica' displayed the lowest at 1.49, falling below the acceptable range, suggesting potential RNA or protein contamination.

Analysis of qualitative test: The DNA analysis of the six types of strawberries showed that all samples had genome sizes larger than 100 kb (Fig. 1), it is reported that strawberry chloroplast genomes are between 155,459 and 155,709 bp. The electrophoresis results confirmed intact DNA bands in all samples, indicating successful isolation of genomic DNA from the leaves of strawberry plants.

The PCR electropherogram successfully amplified the *rbc*L gene in all six cultivars, as indicated by target bands (Fig. 2). The length of the amplified product was about 1,428 bp. The successful amplification of the target region is due to the effective recognition of genomic DNA in the samples by the primers.

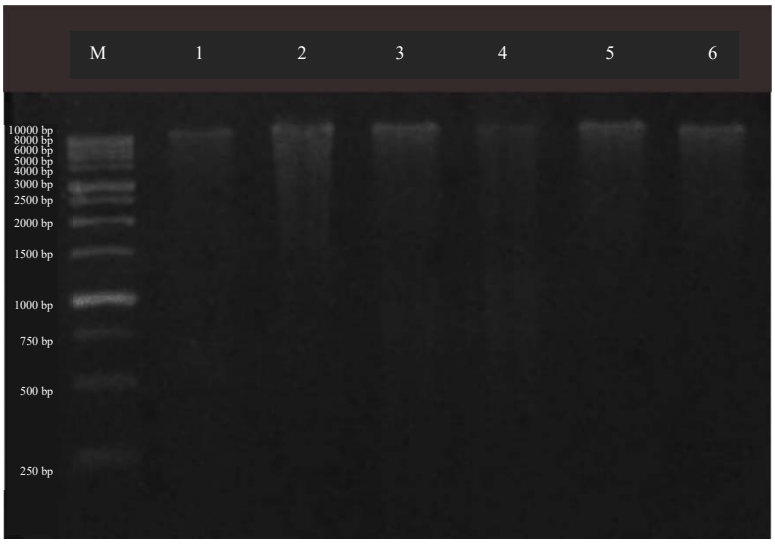

Fig. 1: Gel electrophoresis of the genome of six strawberry cultivars
M: Marker 1 kb, 1: Californica, 2: Knia, 3: Drowning, 4: Moha, 5: Geolhyang and 6: Jumbo×Bali

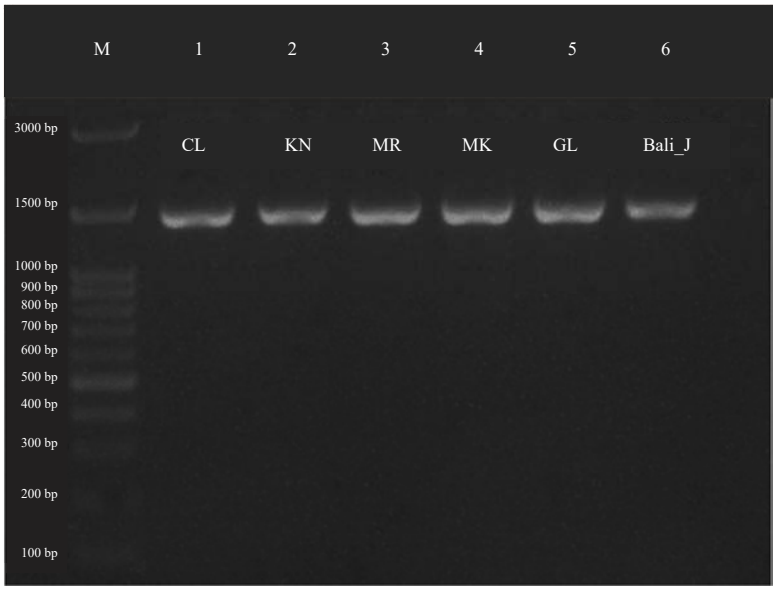

Fig. 2: Gel electrophoresis of target gene *rcb*L of six strawberry cultivars shown 1428 bp
M: Marker 1 kb, CL: Californica, KN: Knia, MR: Mencir, MK: Moha, GL: Geolhyang and Bali_J: Jumbo×Bali

Table 2: DNA concentration results of samples

Cultivars	Concentration (ng/μL)	Purity (A_{260}/A_{280})
Californica	56.4	1.49
Knia	21.9	1.64
Mencir	25.6	1.63
Moha	21.8	1.63
Geolhyang	15.2	1.79
Jumbo×Bali	22.2	1.55

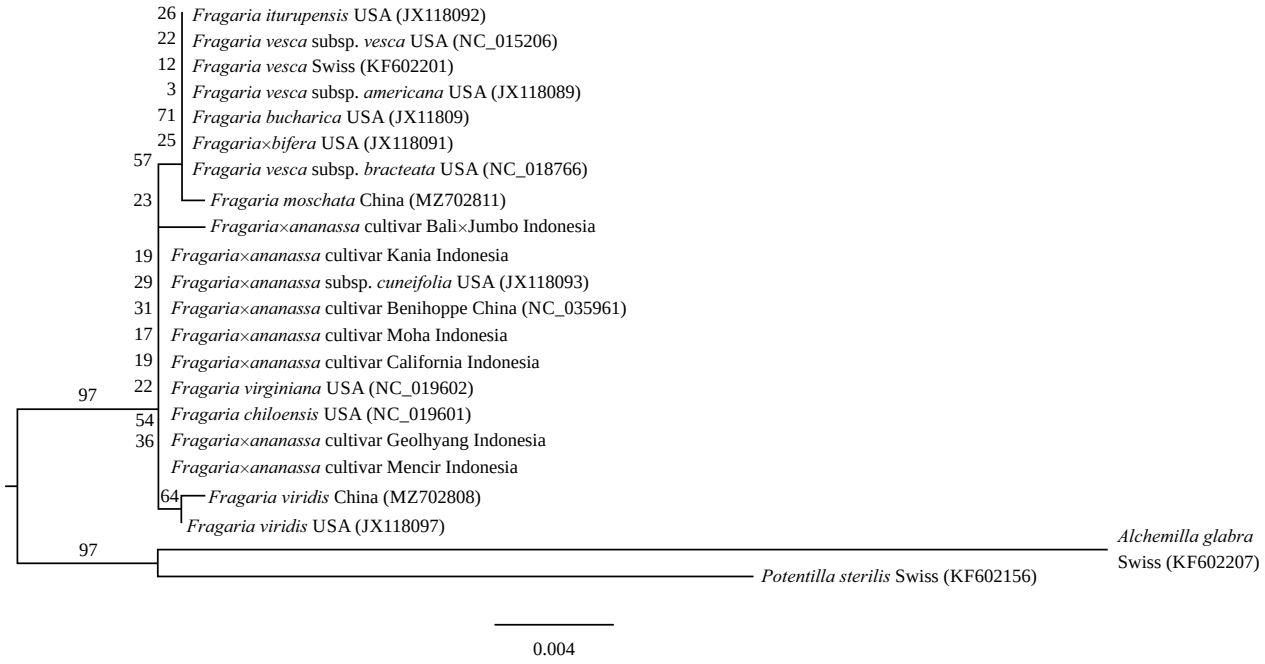

Fig. 3: Phylogenetic tree reconstruction of 20 *rbc*L samples from *Fragaria* spp., members and 2 outgroup sequences *Potentilla sterilis* and *Alchemilla glabra* using IQTree2 program

Result of sequencing: The sequencing results were examined using GeneStudio software to detect uncertainties and verify the sequences. Ambiguous nucleotides were excluded and valid stop codons were authenticated. The absence of discrepancies in the sequence alignment and the successful creation of the phylogenetic tree indicates the high quality of the sequences. In this study, DNA barcoding, which utilizes conserved functional genes to distinguish between species, was implemented. The sequences were stored in fasta format for future alignment.

The length of the *rbc*L sequences from the samples ranged from 1,215-1,317 bp and 1221 bp after alignment. Analysis using Nucleotide BLAST showed sequence similarities ranging from 99.67-100% with *F. virginiana*, *Fragaria* hybrid cultivars, *F. chiloensis*, *F. iturupensis*, *F. moschata* and *F. vesca*. Sequence similarities between 98 and 100% generally support classification within the same species, indicating that the specimens tested likely belong to these species.

Phylogenetic tree and genetic distance: A phylogenetic tree was created using the maximum-likelihood (ML) technique with the Kimura-2-Parameter model and 1,000 bootstrap replicates (Fig. 3). This approach offers insights into evolutionary connections, where the length of branches indicates the level of genetic variation. Longer branches in this analysis represent greater evolutionary distances, while shorter branches indicate closer relationships. The phylogenetic tree illustrated that the *Fragaria* samples from this research formed a large group with other GenBank species, supported by a high bootstrap value of 97.

The reference GenBank samples had a lower bootstrap value of 57, while the tested samples showed similarities ranging from 99.67-100%. The phylogenetic tree displays four distinct groups: One including six cultivar samples with a bootstrap value of 97, a second with various GenBank data with a value of 57, a third with *F. viridis* from China and the USA with a value of 64 and a fourth group with the outgroups, supported by a bootstrap value of 97.

The reconstruction of the Bayesian Inference (BI) phylogenetic tree using the Markov Chain Monte Carlo (MCMC) method was applied for 4 million generations, with sampling every 1,000 generations to estimate the posterior probability distribution in (Fig. 4). Based on the resulting reconstruction of the phylogenetic tree, the six *Fragaria* cultivar samples used are in a large group with posterior probabilities values of 1. A good posterior probability value is >0.98. These results indicate that the six samples are closely related to *Fragaria*× *ananassa* var. *Benihoppe*, *F. chiloensis* and *F. virginiana*.

Genetic distances among *Fragaria* species were measured using the Kimura 2P model (Table 3). The genetic distance among the five sampled cultivars showing perfect

similarity is indicated by ***(0.000), meaning there is no genetic distance among these cultivars, but 'Bali×Jumbo' shows a genetic distance thus implying there is a genetic variation in the sequence.

Polymorphic sites: The alignment of 1,221 bp of the control region sequence revealed the presence of 13 singleton variable sites and two parsimony-informative sites, resulting in 15 mutations across the samples (Table 4). These mutations contributed to the identification of six haplotypes (Table 5). Among the samples, the Bali×Jumbo (S18) cultivar displayed significant genetic variation, with numerous indels and two singleton variable sites distinguishing it from the other cultivars.

Table 3: P-distance of six strawberry cultivars from Indonesia and a strawberry cultivar from GenBank

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2	***																			
3	***	***																		
4	***	***	***																	
5	***	***	***	***																
6	0.002	0.002	0.002	0.002	0.002															
7	***	***	***	***	***	0.002														
8	***	***	***	***	***	0.002	***													
9	***	***	***	***	***	0.002	***	***												
10	***	***	***	***	***	0.002	***	***	***											
11	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001										
12	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	0.002									
13	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	***	0.002								
14	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	***	0.002	***							
15	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	***	0.002	***	***						
16	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	***	0.002	***	***	***					
17	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	***	0.002	***	***	***	***				
18	0.001	0.001	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.001	***	0.002	***	***	***	***	***			
19	0.002	0.002	0.002	0.002	0.002	0.003	0.002	0.002	0.002	0.002	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.001		
20	0.002	0.002	0.002	0.002	0.002	0.004	0.002	0.002	0.002	0.002	0.003	0.001	0.003	0.003	0.003	0.003	0.003	0.003	0.003	

***Means no genetic distance (0,000)

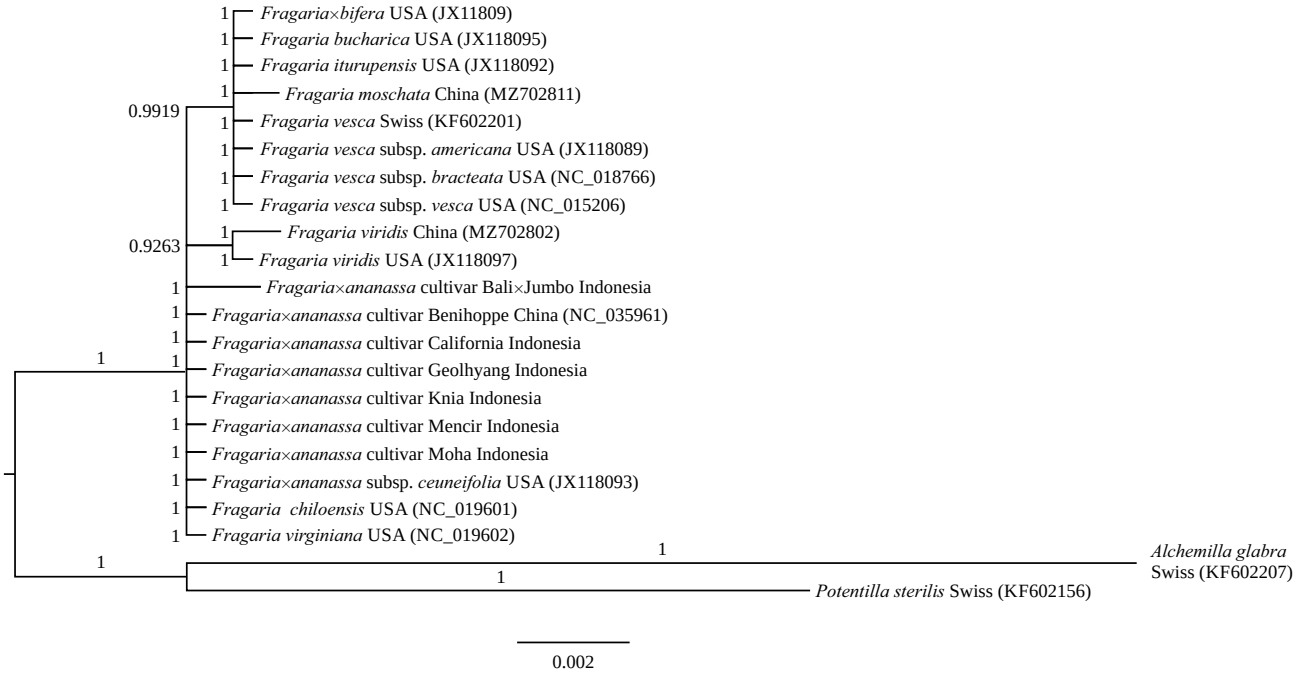

Fig. 4: Phylogenetic tree reconstruction of 20 *rbc*L samples from *Fragaria* spp., members and 2 outgroup sequences *Potentilla sterilis* and *Alchemilla glabra* using the MrBayes program

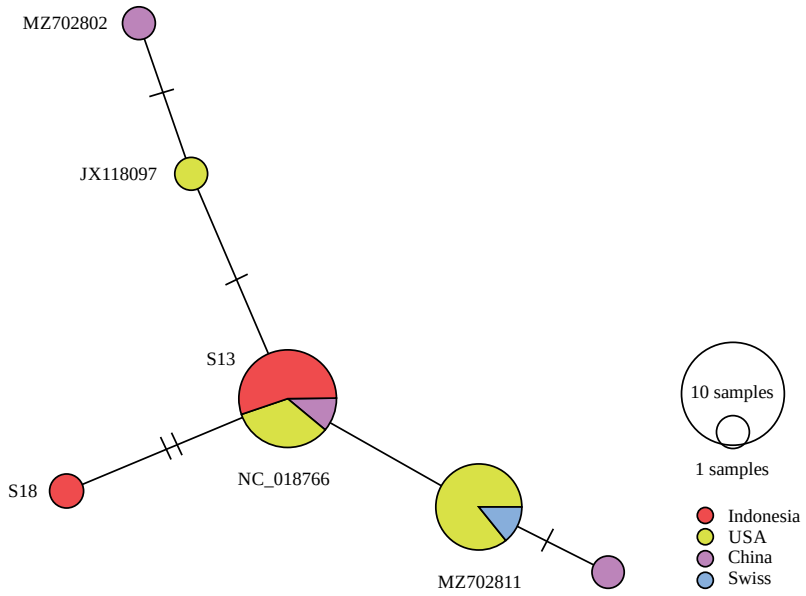

Fig. 5: Median-joining haplotype network of 6 strawberry samples and references from GenBank based on the *rbc*L gene sequence (1221 bp)

Size of the circles indicates the number of samples. Lines between circles show how the samples are related and shorter lines show where changes have occurred between the samples. Each color represents the location of the sample in this study

Table 4: Twenty polymorphic sites of strawberry cultivar based on *rbc*L gene

	5	3	4	5	6	9	1	1	1	1	1	1	1	1	1
		6	4	7	3	1	2	2	2	2	2	2	2	2	2
			6	3	8	4	0	1	1	1	1	1	1	1	2
Sample							8	0	1	2	3	4	5	6	0
JX118089	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
JX118091	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
JX118092	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
JX118093	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
JX118095	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
JX118097	A	C	C	T	C	A	T	A	G	G	G	T	A	A	C
KF602201	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
MZ702802	A	C	T	T	C	A	T	A	G	G	G	T	A	A	C
MZ702811	A	C	C	G	T	C	T	A	G	G	G	T	A	A	C
NC_015206	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
NC_018766	A	C	C	G	T	A	T	A	G	G	G	T	A	A	C
NC_019601	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
NC_019602	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
NC_035961	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
S13-California	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
S14-Knia	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
S15-Mencir	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
S16-Moha	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
S17-Geolhyang	A	C	C	T	T	A	T	A	G	G	G	T	A	A	C
S18-Bali×Jumbo	-	-	C	T	T	A	C	-	-	-	-	-	-	-	T

Data on the number of nucleic acids that have experienced mutations, is indicated in the second column of the first row. This data is read in a vertical format. The subsequent row in the same column lists various nucleotides, which will appear distinct when aligned

Analysis of haplotype and PCoA: Haplotype analysis classified the 20 samples into six groups, with five cultivars-'Californica' (S13), 'Knia' (S14), 'Mencir' (S15), 'Moha' (S16) and 'Geolhyang' (S17)-belonging to haplotype group H4, while 'Bali×Jumbo' (S18) was assigned to haplotype group H6 (Fig. 5). The analysis reveals that nearly all haplotypes are grouped closely together in the coordinate axis, except haplotype H6, which is distinctly separate from the others. The PCoA revealed that Indonesian strawberry cultivars share the same genetic flow of the *rbc*L gene as strawberries from the United States and China (Fig. 6).

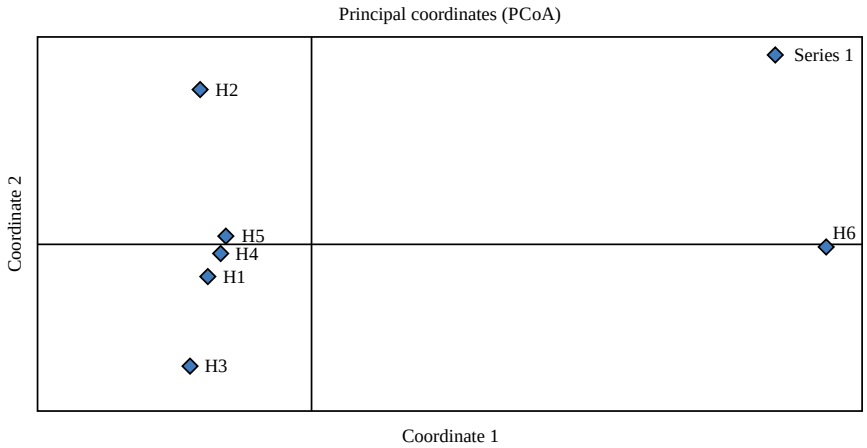

Fig. 6: PCoA of 6 strawberry cultivars based on the *rbc*L gene (1221 bp)

Table 5: Haplotype of twenty samples based on *rbc*L

Haplotype	Number of samples	Sample codes
H1	7	JX118089
		JX118091
		JX118092
		JX118095
		KF602201
		NC_015206
		NC_018766
H2	1	MZ702811
H3	1	JX118097
H4	9	JX118093
		NC_019601
		NC_019602
		NC_035961
		S13
		S14
		S15
		S16
		S17
H5	1	MZ702802
H6	1	S18

S13: Californica, S14: Knia, S15: Mencir, S16: Moha, S17: Geolhyang and S18: Jumbo×Bali. The Haplotype codes in the first column, highlight the presence of six distinct haplotypes. To gain a deeper understanding of the variations among these haplotypes, it is essential to reference Table 4 as well, which outlines the specific differences

DISCUSSION

The results of the DNA extraction show that while some samples, particularly 'Geolhyang', had lower-than-desired concentrations of genomic DNA, the overall purity across several cultivars could have been better. Low concentrations and impurities, such as DNA amplification, can negatively impact subsequent analyses[13,14]. This highlights the need to optimize extraction protocols, especially for samples with high levels of interfering compounds like phenolics[15,16]. The suboptimal purity observed suggests potential contamination, which could compromise downstream processes such as sequencing[17,18].

Despite these challenges, the successful isolation of genomic DNA and the subsequent amplification of the *rbc*L gene in all samples demonstrate the reliability of the selected primers and PCR conditions[19-21]. The clear DNA bands observed during electrophoresis confirm the robustness of the extraction and amplification procedures[22]. Proper primer selection is crucial for ensuring successful amplification, particularly for target regions like the *rbc*L gene, which is essential in phylogenetic studies[23,24].

The high sequence similarities between the tested strawberry samples and known *Fragaria* species support the hypothesis that these cultivars share significant genetic traits with other strawberry species[5,25]. The absence of alignment gaps and valid stop codons further indicates the high integrity of the DNA sequences[26]. The DNA barcoding remains an effective tool for species identification and in this study, the *rbc*L gene proved valuable for confirming species-level classification for these strawberry cultivars.

Identifying polymorphic sites within the control region highlights the genetic variability among the tested strawberry cultivars. The genetic divergence observed in 'Bali×Jumbo' suggests the influence of selective breeding, leading to unique allelic patterns. Detecting haplotypes and genetic variation provides valuable insights for future breeding programs to develop improved strawberry varieties[27,28].

Additionally, the haplotype and Principal PCoA suggest that the genetic flow of strawberries in Indonesia has been shaped by hybridization with cultivars from the United States and China. The genetic divergence observed in 'Bali×Jumbo' may result from selective breeding efforts to enhance cultivar traits[29]. These findings underscore the importance of genetic analysis in understanding the evolutionary history and breeding dynamics of strawberry cultivars in Indonesia.

CONCLUSION

The six strawberry cultivars (*Fragaria*×*ananassa*) 'Californica', 'Knia', 'Mencir', 'Moha', 'Geolhyang' and 'Bali×Jumbo' showed amplification of the *rbc*L gene with a fragment length of 1221 bp after alignment. Phylogenetic analysis using Maximum-Likelihood (ML) by IQtree2 showed that the six cultivars formed a large group with a bootstrap value of 97, while according to Bayesian Inference (BI), they also formed a large group with a posterior probability (pp) value of 1. The haplotype values indicated that the 'Californica', 'Knia', 'Mencir', 'Moha' and 'Geolhyang' cultivars belong to the same haplotype group, while the 'Bali×Jumbo' cultivars belong to a different group. Haplotype network analysis and (PCoA) indicated that strawberries found in Indonesia share the same genetic flow of *rbc*L as those from the United States and China.

SIGNIFICANCE STATEMENT

This study allows precise identification of local strawberry species because many may not have been genetically characterized. This will enhance the understanding of regional biodiversity. The genetic data generated can be used as data for agricultural practices, aiding in the development of strawberry varieties with improved disease resistance, climate adaptability and fruit quality. The use of the *rbc*L gene as a barcode in this research demonstrates the reliability of DNA barcoding as a powerful tool for species identification. The unique contribution of this study lies in its focus on the genetic diversity of strawberries in a largely underexplored region, addressing a critical gap in scientific knowledge and agricultural advancement in Southeast Asia.

ACKNOWLEDGMENTS

The author expresses gratitude to the Government of the Republic of Indonesia through the Ministry of Education, Culture, Research and Technology for funding the Fundamental Research Grant No: 3128/UN1/DITLIT/Dit-Lit/PT.01.03/2023. The researcher also extends thanks to all parties involved in the preparation of this article, including the supervising professors, colleagues and the entire extended family of the Laboratory of Genetics and Breeding, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta.

REFERENCES

1. Latifatul Azizah, U.D., F. Yulianti, A.L. Adiredjo and Sitawati, 2019. Genetic relationship analysis of strawberry (*Fragaria* Sp) germplasm based on morphology character and random amplified polymorphic DNA (RAPD) [In Indonesian]. Plantropica J. Agric. Sci., 4: 77-85.
2. Cayambe, J., M. Heredia-R, E. Torres, L. Puhl and B. Torres *et al*., 2023. Evaluation of sustainability in strawberry crops production under greenhouse and open-field systems in the Andes. Int. J. Agric. Sustainability, Vol. 21. 10.1080/14735903.2023.2255449.
3. Zacharaki, A.K., J.M. Monaghan, J.R. Bromley and L.H. Vickers, 2024. Opportunities and challenges for strawberry cultivation in urban food production systems. Plants People Planet, 6: 611-621.
4. Arif, M.F., G.R. Aristya and R.S. Kasiamdari, 2019. Genetic diversity of strawberry cultivars in Banyuroto, Magelang, Indonesia based on cleaved amplified polymorphic sequence. Biodiversitas J. Biol. Diversity, 20: 1721-1728.
5. Aristya, G.R., R.S. Kasiamdari, R. Setyoningrum and B. Larasati, 2019. Genetic variations of strawberry cultivars of *Fragaria* x *ananassa* and *Fragaria vesca* based on RAPD. Biodiversitas J. Biol. Diversity, 20: 770-775.
6. Corrêa, J.V.W., G.G. Weber, A.R. Zeist, J.T.V. de Resende and P.R. da-Silva, 2021. ISSR analysis reveals high genetic variation in strawberry three-way hybrids developed for tropical regions. Plant Mol. Biol. Rep., 39: 566-576.
7. Zhang, H.X. and M.L. Zhang, 2017. Spatial patterns of species diversity and phylogenetic structure of plant communities in the Tianshan Mountains, Arid Central Asia. Front. Plant Sci., Vol. 8. 10.3389/fpls.2017.02134.
8. Nurhasanah, Sundari and N. Papuangan, 2019. Amplification and analysis of rbcL gene (ribulose-1,5-bisphosphate carboxylase) of clove in Ternate Island. IOP Conf. Ser.: Earth Environ. Sci., Vol. 276. 10.1088/1755-1315/276/1/012061.
9. Berry, J.O., C.M. Mure and P. Yerramsetty, 2016. Regulation of Rubisco gene expression in C_4 plants. Curr. Opin. Plant Biol., 31: 23-28.
10. Ihwan, D.R.T. Sari, L. Hakim, C. Retnaningdyah and E.L. Arumingtiyas, 2021. *In silico* analysis of rbcl protein and D-Ribulose 1,5-bisphosphate bond. JSMARTech: J. Smart Bioprospect. Technol., 2: 65-68.
11. Taariwuan, M.B., J. Ngangi, Y. Mokosuli and S. Gedoan, 2021. DNA barcoding of Dalugha (*Cyrtosperma merkusii*) in Talaud Islands and North Minahasa based on rbcL gene [In Indonesian]. J. Bios Logos, 11: 134-138.
12. Jin, X., H. Du, C. Zhu, H. Wan and F. Liu *et al*., 2023. Haplotype-resolved genomes of wild octoploid progenitors illuminate genomic diversifications from wild relatives to cultivated strawberry. Nat. Plants, 9: 1252-1266.

13. Galal-Khallaf, A., A. Abdelbaset-Donya, W. Hamza and K. Mohammed-Geba, 2021. Molecular tools for assuring human health and environment-friendly frozen shellfish products in the United Arab Emirates markets. Food Chem.: Mol. Sci., Vol. 3. 10.1016/j.fochms.2021.100028.
14. Pelt-Verkuil, E., A. Belkum and J.P. Hays, 2008. Principles and Technical Aspects of PCR Amplification. 1st Edn., Springer, Dordrecht, Netherlands, ISBN: 978-1-4020-6241-4, Pages: 330.
15. Kuffel, A., A. Gray and N.N. Daeid, 2021. Impact of metal ions on PCR inhibition and RT-PCR efficiency. Int. J. Legal Med., 135: 63-72.
16. Rezadoost, M.H., M. Kordrostami and H.H. Kumleh, 2016. An efficient protocol for isolation of inhibitor-free nucleic acids even from recalcitrant plants. 3 Biotech, Vol. 6. 10.1007/s13205-016-0375-0.
17. Aboul-Maaty, N.A.F. and H.A.S. Oraby, 2019. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. Bull. Nat. Res. Cent., Vol. 43. 10.1186/s42269-019-0066-1.
18. Park, G., J.K. Park, S.H. Shin, H.J. Jeon and N.K.D. Kim *et al*., 2017. Characterization of background noise in capture-based targeted sequencing data. Genome Biol., Vol. 18. 10.1186/s13059-017-1275-2.
19. Lucena-Aguilar, G., A.M. Sánchez-López, C. Barberán-Aceituno, J.A. Carrillo-Ávila, J.A. López-Guerrero and R. Aguilar-Quesada, 2016. DNA source selection for downstream applications based on DNA quality indicators analysis. Biopreserv. Biobanking, 14: 264-270.
20. Meiklejohn, K.A., M.L. Jackson, L.A. Stern and J.M. Robertson, 2018. A protocol for obtaining DNA barcodes from plant and insect fragments isolated from forensic-type soils. Int. J. Legal Med., 132: 1515-1526.
21. Skuza, L., I. Szućko, E. Filip and A. Adamczyk, 2019. DNA barcoding in selected species and subspecies of rye (*Secale*) using three chloroplast loci (*matK*, *rbcL*, *trnH-psbA*). Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 47: 54-62.
22. Aguilar-Rabiela, A.E., E.M. Hernández-Cooper, J.A. Otero and B. Vergara-Porras, 2020. Modeling the release of curcumin from microparticles of poly(hydroxybutyrate) [PHB]. Int. J. Biol. Macromol., 144: 47-52.
23. Lee, P.Y., J. Costumbrado, C.Y. Hsu and Y.H. Kim, 2012. Agarose gel electrophoresis for the separation of DNA fragments. J. Visualized Exp., Vol. 62. 10.3791/3923.
24. Kress, W.J., K.J. Wurdack, E.A. Zimmer, L.A. Weigt and D.H. Janzen, 2005. Use of DNA barcodes to identify flowering plants. Proc. Natl. Acad. Sci. U.S.A., 102: 8369-8374.
25. Ford, C.S., K.L. Ayres, N. Toomey, N. Haider and J. van Alphen Stahl *et al*., 2009. Selection of candidate coding DNA barcoding regions for use on land plants. Bot. J. Linn. Soc., 159: 1-11.
26. Folta, K.M. and C.R. Barbey, 2019. The strawberry genome: A complicated past and promising future. Hortic. Res., Vol. 6. 10.1038/s41438-019-0181-z.
27. Seo, T.K., B.D. Redelings and J.L. Thorne, 2022. Correlations between alignment gaps and nucleotide substitution or amino acid replacement. Proc. Natl. Acad. Sci. U.S.A., Vol. 119. 10.1073/pnas.2204435119.
28. Fan, Z. and V.M. Whitaker, 2024. Genomic signatures of strawberry domestication and diversification. Plant Cell, 36: 1622-1636.
29. Zhang, X., Q. Wu, L. Lan, D. Peng and H. Guan *et al*., 2024. Haplotype-resolved genome assembly of the diploid *Rosa chinensis* provides insight into the mechanisms underlying key ornamental traits. Mol. Hortic., Vol. 4. 10.1186/s43897-024-00088-1.