BIODIVERSITAS
Volume 25, Number 7, July 2024
Pages: 2969-2979

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d250718

Genetic diversity and population structure of Philippine strawberry germplasm based on genome-wide Simple Sequence Repeat markers

VENUS H. ELEC[1,♥], CHARLES ANTHON E. CADORNA[1,♥♥], BELINDA A. TAD-AWAN[2], DARWIN A. BASQUIAL[2], MILAGROS R. DUMASLAN[2], JESSICA D. REY[1,♥♥♥]

[1]Plant Molecular Phylogenetics Laboratory, Institute of Biology, Regidor Street, National Science Complex, University of the Philippines Diliman. Quezon City 1101, National Capital Region, Republic of the Philippines. Tel.: +63-2-89818500 loc. 3727, ♥email: vhelec@up.edu.ph, ♥♥email: cecadorna@up.edu.ph, ♥♥♥email: jdrey@up.edu.ph

[2]College of Agriculture, Benguet State University. La Trinidad, Benguet 2601, Cordillera Administrative Region, Philippines

Manuscript received: 1 March 2024. Revision accepted: 9 July 2024.

Abstract. *Elec VH, Cadorna CAE, Tad-Awan BA, Basquial DA, Dumaslan MR, Rey JD. 2024. Genetic diversity and population structure of Philippine strawberry germplasm based on genome-wide Simple Sequence Repeat markers. Biodiversitas 25: 2969-2979.* Strawberry is a globally recognized fruit due to its health benefits and economic significance. Determining the level of genetic diversity and the relationship between germplasm resources involved in the breeding program is critical for effective crop improvement strategies. This study used 197 Simple Sequence Repeat (SSR) markers to assess the genetic diversity and population structure of 24 strawberry germplasm from the Benguet State University (BSU) collection. In this study, 70% of the markers used were polymorphic with reproducible fragments. A total of 792 alleles were detected, with an average of six per marker. The diversity indices indicated an intermediate level of diversity among the germplasm based on the Jaccard coefficient. Results from cluster analysis generated three groups, separating the Red Milky Way and Summer Princess cultivars from the rest of the germplasm evaluated, while a third cluster was further divided into six subgroups. Furthermore, identified clustering patterns of the evaluated genotypes showed independence from the region of origin, coinciding with Principal Component Analysis (PCA) and population structure analysis results. This study reveals the genetic diversity level and population structure of the strawberry cultivars in the Philippines using SSR markers, facilitating accurate identification and informed parental selection for breeding objectives.

Keywords: Admixture, genetic diversity, geographic origin, Philippine strawberry, SSRs

INTRODUCTION

Strawberries (*Fragaria x ananassa* Duchesne ex Rozier) are economically important fruit crops cultivated worldwide because of their desirable sensory properties and nutritional benefits. Strawberries contain high levels of health-promoting compounds, such as Vitamin C, anthocyanins, and diverse antioxidants, such as flavonoids and folic acid (Porter et al. 2023). In the past decades, there has been substantial growth in the global production of strawberries, reaching over 9.6 million tons in 2022, which reflects a strong consumer demand (FAOSTAT 2023). Similarly, in the Philippines, strawberries have increased prominence as a high-value crop, with the current production situated in the Cordillera Mountain region with the potential to expand to other suitable regions (Padua et al. 2016).

Optimizing strawberry germplasm production under local conditions is critical for the continued growth of the industry in the Philippines. The success of varietal improvement relies on accessing and utilizing the genetic diversity within the current germplasm. Identifying genetic variation within the available germplasm is a foundation for parental selection for breeding programs aimed at developing new cultivars with introgressed adaptive traits (Whitaker et al. 2020). Recently, a comprehensive genetic diversity and structure analysis on the cultivated strawberry collection in the United States was conducted by Zurn et al. (2022). The collection is considered the largest repository of germplasms maintained by the United States Department of Agriculture (USDA) and utilized in several breeding programs across the globe. In Asia, the strawberry germplasm collection in China has been characterized, including its wild relatives (Chen et al. 2018; Guo et al. 2018; Hu et al. 2022). The genetic diversity of strawberries in Korean (Lim et al. 2017) and Japanese germplasms (Wada et al. 2017) have also been characterized and used in developing new cultivars, which have elevated these countries as the top exporters of strawberries in Asia. Similarly, the characterization of strawberry germplasm in different breeding programs from countries outside the temperate regions, such as Pakistan, Indonesia, India, and Brazil, has also been initiated (Bhowal et al. 2019; Kumari 2019; Qarni et al. 2020; Camargo et al. 2022). However, the genetic diversity of strawberry accessions found in Philippine germplasm collection and breeding programs has not yet been fully evaluated. Most cultivated strawberries grown in the Philippines originated from North American, European, and some parts of Asia varieties (Padua et al. 2016). Molecular markers offer an effective tool for characterizing and evaluating the genetic diversity within germplasm collections and differentiating cultivars based on their unique genetic profiles (Zurn et al. 2022). Microsatellites or Simple Sequence Repeats (SSRs) are ideal for identifying cultivars and studying genetic diversity because of their

reproducibility and high level of polymorphism (Testolin et al. 2023; Zhang et al. 2023). SSR markers target short tandem DNA repeat motifs, providing co-dominant and highly multiallelic data (Vieira et al. 2016). Their capacity to capture allelic variation at numerous loci makes them well-suited for discerning relationships and diversity within the germplasm collection. Therefore, SSR has been used for genetic diversity studies of many important crops, such as melon, amaranth, capsicum, and coconut (Zhong et al. 2021; Caro et al. 2022; Delgado and Martín 2023; Zhang et al. 2023). The use of SSR in the characterization of strawberry germplasm was previously conducted by Kaleybar et al. (2018) on strawberry germplasm in Iran. Likewise, the use of SSR markers to characterize strawberry germplasm in their respective breeding programs was utilized by Hilmarsson et al. (2017); Lim et al. (2017); and Biswas (2018). These studies prove SSRs are optimal for the initial molecular analysis of polyploid strawberries when used in sufficient numbers (Aristya et al. 2019; Kim et al. 2019; Xin et al. 2019).

In this regard, genetic characterization and fingerprinting of the germplasm grown in the Philippines using SSR was conducted. Molecular diversity analysis will provide key insights into Philippine strawberries' genetic variability and structure for future breeding programs. These findings will guide breeding initiatives to optimize germplasm for crop improvement and conservation in local conditions.

MATERIALS AND METHODS

Plant materials and study area

A total of 24 cultivated strawberries (*F. ananassa*) were used in this study (Table 1). These germplasm sets have been introduced in the Benguet State University (BSU) breeding program since its establishment in the early 1990s and originated from other regions in the Philippines and other countries through germplasm exchanges and collaborations by farmer scientists and indigenous communities across the globe. The germplasm was maintained in the Horticulture Research and Training Institute through in vitro culture and clonally propagated in the greenhouse. For this experiment, the materials were planted in the fields of BSU, La Trinidad, Benguet, Philippines (16°27′07″N, 120°35′27″E) using a complete randomized design with three replicates.

Procedures

DNA extraction

Ten disease-free, fully expanded trifoliate leaf samples were collected from each germplasm. This included four samples from replicate one, three from replicate two, and three from replicate three. The samples were then transported to the Plant Molecular Phylogenetics Laboratory at the Institute of Biology, University of the Philippines Diliman, Quezon City, Philippines, for total genomic DNA extraction. Total genomic DNA was extracted using the DNeasy® Plant Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's protocol. The concentration and relative purity of the extracted DNA were estimated using NanoDrop Lite (Thermo Fisher Scientific Inc., USA). The integrity of DNA was also assessed through electrophoresis using 1% agarose stained with GelRed (Biotium, USA) in 1x Tris-Acetate-EDTA (TAE) running buffer at 100V for 45 minutes). Genomic DNA was visualized using the Gel Doc EZ Documentation System (Bio-Rad, USA).

Microsatellite genotyping

A total of 197 SSR markers were used in this study, comprising 120 newly developed and 77 previously published markers. Details of each marker are presented in Table 2. Polymerase Chain Reaction (PCR) amplification was conducted with 15 µL reaction volume containing 1 µL genomic DNA, 0.5 µL of the forward primer (10 uM), 0.5 µL of the reverse primer (10 µM), 7.5 uL 2X MyTaq^{TM} HS Red Mix (Bioline, USA), and 5.5 µL nuclease-free water. Amplifications were carried out in a T100^{TM} Thermal Cycler (Bio-Rad, USA) under the following PCR conditions: initial denaturation at 95°C for of 1 min, followed by 35 cycles of denaturation at 95°C for 15 sec, annealing at 60°C for 15 sec, and extension at 72°C for 10 sec, and a final extension step at 72°C for 5 min. The PCR amplification products were resolved by electrophoresis using an 8% non-denaturing polyacrylamide gel in 0.5X Tris-borate EDTA buffer at 100V for 90min in the CBS. Scientific Triple Wide Mini Vertical System^{TM} (CBS.B.S. Scientific Company San Diego, California, USA) was used and visualized using GelRed® Nucleic Acid Gel Stain (Biotium, USA). The gel was then imaged using a Gel Doc EZ Documentation System (Bio-Rad, USA). Gels were scored manually for the presence or absence of bands for each germplasm, and the expected and observed allele sizes were also determined using Bio-Rad software by comparing the detected band to the standard Kb ladder. The data were formatted according to the different programs' input requirements, considering the crop's ploidy level.

Data analysis

The genetic diversity of 24 strawberry germplasm was characterized using various parameters calculated using the R 4.3.1 statistical software (R Core Team 2020) using the poppr (v2.9.0), pegas, and polysat packages (Paradis 2010; Clark and Jasieniuk 2011; Kamvar et al. 2014). These parameters included the observed number of alleles, Simpson's index, Nei's gene diversity and evenness, genotypic richness (the number of multilocus genotypes observed per population-MLG); genotypic diversity (percentage of polymorphism detected by each population)-%P; Shannon-Wiener Index of MLG diversity per population (Simpson 1949; Nei 1978; Goudet 2005).

Similarly, data were analyzed using the Unweighted Pair Group Method with Arithmetic mean (UPGMA; http://genomes.urv.cat/UPGMA) algorithm utilizing the Jaccard coefficient to calculate the genetic similarities among germplasm. Subsequently, the output data in Newick format was used with FigTree v1 4.4 (http://tree.bio.ed.ac.uk/software/figtree) to generate a dendrogram. Principal Component Analysis (PCA) was performed using PAST v4.14 (Hammer et al. 2001) software with a bootstrap frequency of n=1000.

Table 1. List of genotypes used in the study and source of origin grouped according to country and continent

Code no.	Germplasm	Cultivar code	Samples	Ploidy	Origin
V1	Sweet Charlie (Argentina)	SCA	10	Octoploid	SA_Argentina, South America
V2	Camarosa	CMA	10	Octoploid	US_USA, North America (USA)
V3	Earlibrite	ERB	10	Octoploid	US_California, North America (USA)
V4	Festival	FTV	10	Octoploid	US_Florida, North America (USA)
V5	Honeoye	HNY	10	Octoploid	US_New York, North America (USA)
V6	Missionary	MSR	10	Octoploid	US_California, North America (USA)
V7	Sweet Charlie (Washington)	SCW	10	Octoploid	US_Washington, North America (USA)
V8	Sweet Charlie California	SCC	10	Octoploid	US_California, North America (USA)
V9	Sweet Charlie Hawaii	SCH	10	Octoploid	US_Hawaii, North America (USA)
V10	Sweet Charlie Texas	SCT	10	Octoploid	US_Florida, North America (USA)
V11	Whitney	WTY	10	Octoploid	US_California, North America (USA)
V12	Winterdawn	WTD	10	Octoploid	US_Florida, North America (USA)
V13	Agsapa	AGP	10	Octoploid	AS_Philippines, Asia
V14	139 (OPG)	OPG	10	Octoploid	AS_Philippines, Asia
V15	Benehoppe	BNP	10	Octoploid	AS_Japan, Asia
V16	Cavite Japan	CVJ	10	Octoploid	AS_Japan, Asia
V17	Japan Strawberry	JPS	10	Octoploid	AS_Japan, Asia
V18	Korean Variety	KRV	10	Octoploid	AS_Korea, Asia
V19	Pistro	PST	10	Octoploid	AS_Japan, Asia
V20	Red Milky way	RMW	10	Octoploid	AS_China, Asia
V21	Snow White	SNW	10	Octoploid	AS_Japan, Asia
V22	Summer Princess	SMP	10	Octoploid	AS_Japan, Asia
V23	Sweet Angel Bohol	SAB	10	Octoploid	AS_Japan, Asia
V24	White strawberry	WTS	10	Octoploid	AS_Japan, Asia

Therefore, to investigate the population structure and admixture, Bayesian model-based clustering was performed using the STRUCTURE v.2.3.4 program (Pritchard et al. 2000) to infer the appropriate clusters (K), with the following criteria: burn-in of 10,000, run length of 150,000, and assumed admixture model and correlated allele frequencies. Ten STRUCTURE runs were performed in population (K) settings from 1 to 10. The optimal K was determined using the Evanno method based on the change in log probabilities between successive K values using the web-based software STRUCTURE HARVESTER v.0.6.94 (Earl and VonHoldt 2012). The analysis permitted an inference of population structure and identification of admixed individuals in the sample set and was calculated using CLUMPP v. 1.1.2 (Evanno et al. 2005).

RESULTS AND DISCUSSION

Genetic diversity analysis

A total of 197 genome-wide SSR markers composed of 120 newly developed markers (Elec et al. 2023) and 77 published SSR markers (Salinas et al. 2017; Biswas 2018) were evaluated. From the 120 developed SSRs, 98 (82%) were polymorphic with scorable alleles, while the remaining 22 SSRs were identified as monomorphic. In contrast, among the published SSR markers, 33 (43%) produced unambiguous polymorphic amplicons, 26 did not amplify, and 18 were monomorphic. A total of 131 polymorphic markers were utilized in the genotyping of the strawberry germplasm, which identified a total of 792 alleles. The number of alleles per SSR locus ranged from two to eight, with an average of 6 alleles (Table 2). The frequency of SSR markers having 8 alleles is 36% (48 markers), and 6 alleles by 19.8%, and 28 markers (21%) that generated 2 alleles per locus.

Genetic diversity was computed using the Simpson Index (1949), Nei's Gene Diversity (1978), and Evenness (Pielou 1975) (Table 2). Values of Simpson's index and Nei's distance generated similar values except for MK31, MK39, MK47, MK51, MK62, PM34, and PM38. The genetic distance among the 131 markers ranges from 0.20 to 0.8. The highest diversity index was observed from MK42 with a value of 0.88, while the lowest was from Marker73 (0.20). The average diversity index observed is 0.78 for Nei's and Simpson's indices. In addition, it can also be observed that 62% of the markers have detected a diversity index higher than 0.80, and only 4.58% of the markers have detected a diversity index lower than 0.5. Evenness ranges from 0.59 to 1, with an average of 0.943. Therefore, 10 markers generated an evenness value of 1, which includes MK 26, MK40, MK42, MK46, MK49, MK54, MK81, MK82, FA27 and FA30. The lowest evenness detected is from marker MK73, with a value of 0.59.

Table 2. Characterized SSR primers with their diversity indices

Locus code	Forward sequence	Reverse sequence	No. of alleles	Simpson's Index	Nei's gene diversity	Evenness
MK1	CCATTTGATCCCAAGCAACT	AGGCCGATTGTGTTTAGCAG	8	0.86	0.86	0.94
MK2	CCCAGCTGATGGATGTTGTA	TGGAATCAAAGGGAACTGATG	6	0.83	0.83	0.99
MK3	CCATCCATGGCTTCTCTGTT	GGTGCTTCAGTTTTGGGATG	8	0.87	0.87	0.99
MK4	CATGGTGAATGCAACTGGAA	AAGAGCGAGGCAAAATCAAA	7	0.84	0.84	0.96
MK5	AGGTGCTCCTTGAACGTAGG	CGACATGAAGAACGATCGAG	8	0.86	0.86	0.96
MK6	GCATAGTACCACGCCCTGTC	TCATTATTCGGCTCCTCCTC	8	0.86	0.86	0.93
MK8	TTGGATTCCCATCTTCATCC	GAAGCCTAAGGAAGCGAGGT	8	0.79	0.79	0.78
MK9	GGTAATTATGGCGGAAACCA	CACCTGCGATTGTAGAGCAA	6	0.81	0.81	0.94
MK10	ATCTCTTTTGACGCGCACTT	AGCAGAGTTCATGGGGTTTG	2	0.26	0.26	0.66
MK11	GCAAACTTGCAAACAATGGA	ATTGAGGCCCAAATCAAGG	5	0.77	0.77	0.92
MK12	TCTGCTTGGGTGCTTCCTAT	GGGAGATCATAGCCTTGAAAGA	5	0.78	0.78	0.94
MK13	TCAGAATCCGATGCAAAACA	ACCAAACCCAACCAAAATGA	7	0.85	0.85	0.98
MK14	CATGGCAAGGAAGGAAGAAA	GCCTCGCTCTGATCTTCAAC	5	0.79	0.79	0.96
MK15	GTGGAAGTGATCGGAGGTGT	ACCTCATCCTCCCCAATTCT	4	0.73	0.73	0.96
MK16	TTGCCCCAATTAACAAAAGG	AGAAGATCCTCAAGCCCACA	2	0.46	0.46	0.93
MK17	CATACCGACTCCGAAACCAT	TCAAAATCCAAACGCCTCTT	8	0.86	0.86	0.96
MK18	CAGCATGCAAAACTCCTTCA	TCTTGGGTCCCTTCTTGTTG	6	0.83	0.83	0.98
MK19	TGGCCTTTAAATTGGTCGAT	CGTGAATGGCCTATTTGGTT	8	0.86	0.86	0.97
MK20	TGAATGACCTTTCGCAACAA	CCCTGCAGTCATGTTGGTTA	6	0.82	0.82	0.96
MK25	TCTTCCTCCTCATCCGAGAA	TGACCCCAAAATGAAGAAGC	5	0.80	0.80	0.98
MK26	GGTCCAAATAAAAGAGCAACCA	TGGGCAAGAAATAAGGGACA	4	0.75	0.75	1.00
MK27	TCACTTACCATGAGCCGTTG	TGGGCAAATCTTGGGTTAAG	6	0.81	0.81	0.96
MK28	GAGATTTTGAAGGCCCTGTG	TTCTTTGGCTTTTGCGATCT	8	0.84	0.84	0.92
MK29	TGTTGCTGGGGAATACAAGA	CTGACCAACATCACCCTCCT	8	0.86	0.86	0.94
Mk30	CTGACCAACATCACCCTCCT	TGTTGCTGGGGAATACAAGA	8	0.85	0.85	0.94
MK31	TTTGCTATGATGCGGAGTTG	TTTCTCCGGTGGTGATTAGG	3	0.61	0.62	0.91
MK32	GAGAGTGGATCGCCTTCTTG	GTTGACTGTTGGAGGGGTGT	8	0.84	0.84	0.92
MK33	GTTGACTGTTGGAGGGGTGT	GAGAGTGGATCGCCTTCTTG	8	0.84	0.84	0.92
MK34	GGCGCTAGACAAGTCCTTTG	GCGGTACTTGGAGCATGAAT	8	0.85	0.85	0.93
MK35	GAAGGGTGGAGTGAATTGGA	TTGGAAGGAAGCTGAATTGC	6	0.83	0.83	0.99
MK36	TTGAATGCCTTAACGATCACC	AACGCAGTTTTGAGTGCCTTA	7	0.84	0.84	0.95
MK37	AGCAAGGGTTAACATTCCAGT	TAGCTTTCCAGCAAACATGC	7	0.84	0.84	0.97
MK38	TAGCTTTCCAGCAAACATGC	AGCAAGGGTTAACATTCCAGT	7	0.84	0.84	0.97
MK39	ACCCACTTTATGCACCGTTT	CAAGAGGTGTTGGTGGCTCT	8	0.86	0.87	0.97
MK40	AACATTCCTGGCCAATTCAC	TGCTTATCATTGCTGCATCC	6	0.83	0.83	1.00
MK41	GACTCCGCCTATGTTTTGGA	GCCGGCAGCCATATATAGAA	8	0.86	0.86	0.94
MK42	TGACCAGACCCACACAAGAA	TTGGATTTGCAGTCCATCAA	8	0.88	0.88	1.00
MK44	AGGAAAGGGTCAAACCTCGT	AAAGTGAGGGACACGGACAC	6	0.81	0.81	0.96
MK45	CCTACAATGGCTCGAGGAAA	TATTTCTCCAAACCCCACCA	6	0.81	0.81	0.95
MK46	TATTTCTCCAAACCCCACCA	CCTACAATGGCTCGAGGAAA	8	0.87	0.87	1.00
MK47	AATCATTCTCCCCTCCCTTTT	CAAGCAGCATCACAATCACC	6	0.82	0.83	0.97
MK48	GGCGCTTAGAAAACGTGAAG	TCTCAGGAGAAGCCATGGAG	3	0.66	0.66	0.99
MK49	TCTCAGGAGAAGCCATGGAG	GGCGCTTAGAAAACGTGAAG	5	0.80	0.80	1.00
MK50	AATTCGATCGGAAAGTGGTG	TCGGCTAAAGTGCTTGTCCT	8	0.84	0.84	0.91
MK51	AGGATGACGGCGAGATATTG	TGTTGTAGCAGTTGCCCTTG	5	0.79	0.80	0.99
MK52	GGGTGGAAATAGCCTCATCC	AAATGGGTAAAGGAGGCAAGA	5	0.73	0.73	0.84
MK53	AAATGGGTAAAGGAGGCAAGA	GGGTGGAAATAGCCTCATCC	3	0.65	0.65	0.97
MK54	TCGTTTGGGGTTGTTTCTTT	CCTGCATGAGAAGGTTAGGG	6	0.83	0.83	1.00
MK55	AACGCAGTTTTGAGTGCCTTA	AACGCAGTTTTGAGTGCCTTA	6	0.83	0.83	0.99
MK56	CCTGCTCACTGTCTCCATCA	TGCGATACTTGATCCCACAA	8	0.87	0.87	0.97
MK57	TAGATCGCAACGAGTGCAAC	ACGGAAGCAGAGGAAGCATA	6	0.83	0.83	0.98
MK58	CGTCAAAATGCTTTGGAAGA	GAGCGGGTGATTTAGGACAA	6	0.80	0.80	0.92
MK59	ACCGCTTCTGGGTTTCTTTT	TACGTGAGCTTCGAACATGG	6	0.83	0.83	0.98
MK60	TTTTGATGAATGTGGCTTGG	CACTGGCACACAAAGTAGCAA	6	0.81	0.81	0.96
MK62	TTTGCACGCTGATTGTGACT	TGCATCTTGACTCCATTCCA	4	0.72	0.73	0.94
MK63	TGAAACATAGGCTCTCATTTGC	CGTCTCTCTTGGGGCACTTA	6	0.81	0.81	0.95
MK64	AAACGCCCTAACAACAAACG	GACGCACGAACTCAAGAAAA	4	0.74	0.74	0.98
MK65	GTCGGCAGAGGAATTTACGA	GGTGCACATTCGACCAATTA	5	0.80	0.80	0.99
MK66	TCGATCGTGCATTTCTTGAG	GGATTTTCGCAGTCGTCTTT	6	0.83	0.83	0.99
MK67	GGATTTTCGCAGTCGTCTTT	TCGATCGTGCATTTCTTGAG	7	0.85	0.85	0.98
MK68	CCACACAACACAGAACCAGAG	CCCACTTGCTGTAATGGTTG	2	0.38	0.38	0.79
MK69	TATCGCAATCCTTCACCACA	CTCCTCTTCCATCCAGGACA	2	0.49	0.49	0.99
MK71	CTTGCCAAAACGCTTGGTAT	CAGGACGAAGTGCTCATCAA	8	0.86	0.86	0.96
MK72	CAGGACGAAGTGCTCATCAA	CTTGCCAAAACGCTTGGTAT	8	0.86	0.86	0.94
MK73	CTCCAATCCCTTAGCATCCA	CTTGCTTTGCCTTCACTTCC	2	0.20	0.20	0.59
MK74	ACATGACCCTCCGGATACTG	TGAAACTGGGAGCTTCTGCT	5	0.78	0.78	0.96

MK75	CCCTCAATTAAGCCACCATC	TAGTGTGCCACCCATTCTCA	5	0.75	0.75	0.90
MK76	GCATGATGTTGATGCGTACC	TGGTGGTCAGGTGAAGATGA	3	0.54	0.54	0.88
MK77	TTCCAATTTTGATGGCTTCC	TTTGGTGCATTCCATGCTTA	8	0.86	0.86	0.95
MK78	TCGTCAAGTGCCATGAAGAG	GCCAAAGTCGAAGAGCAAAG	4	0.69	0.69	0.88
MK79	AGAGTGAAGGGAACGGTGTG	TTTTCCCAACGGCTAGTTTG	2	0.42	0.42	0.85
MK80	GGTGGGAGTGTCATTGTGTG	CGAGGCAGAGAGGAAGAAGA	4	0.73	0.73	0.97
MK81	AGATTGTTCGGCATTCTTGG	GCAGCAATTTCAAAACCACA	3	0.67	0.67	1.00
MK82	ATCAGATTGAGTGGCGCAGT	GCCTGTCCGGTCTCATGTAT	2	0.50	0.50	1.00
FA1	TACAATCACCGGACGTCTCA	GATGGCTCATCGTTTGGATT	6	0.80	0.80	0.93
FA2	AAGTCATGCTTGAAGGCACA	TCGGAGGATAGGATCAAACG	7	0.85	0.85	0.98
FA3	CATCCCTACTCCCAGTTCCA	TGCACGTGGACTTTCTTGAG	7	0.81	0.81	0.89
FA4	GGCCGTCTTCTTTCCCTTTA	ATTGGTTCTTGGCTGTTTGG	5	0.77	0.77	0.95
FA5	CCTCTTGGATTTGGATTGGA	GTGTTGGCTGGGATTCTTGT	4	0.75	0.75	0.99
FA6	GTGTTGGCTGGGATTCTTGT	CCTCTTGGATTTGGATTGGA	6	0.81	0.81	0.96
FA7	GAGTGACCCACAAGTGCAAA	CGGATGGATAGATTGGATGG	5	0.75	0.75	0.93
FA8	CTTGTTTCCCCATGTCATCA	AATTCGATCGGTTGGCATTA	3	0.54	0.54	0.83
FA9	ATCGTCCTTTCGGTGATGAG	AGCTCGGAACTGCAATCAAT	4	0.68	0.68	0.94
FA10	AGCTCGGAACTGCAATCAAT	ATCGTCCTTTCGGTGATGAG	4	0.68	0.68	0.94
FA11	AAGCATGTTTTCCCTTGCAC	TGCTCATTAAGTTCACTGTGTGG	8	0.86	0.86	0.94
FA12	GCAGGTCTGTGCTGAACTTTT	TCCCTTGCACTGAAGAATTG	6	0.83	0.83	0.98
FA14	AGGCTTCGTGGTAAGAGCAA	GGTCCAATGCCAGAAGAAAT	6	0.80	0.80	0.94
FA16	TCCTGGGTTTCTGAGATTCG	TGTGCGACCATGCTACTGTT	3	0.57	0.57	0.82
FA19	GCTTCCCTATTTCCCAGACC	AAGGTTGGTCCTCCTCCACT	8	0.86	0.86	0.95
FA20	GCTTCCCTATTTCCCAGACC	TGAGGGGTGGTCGTCTCTAC	8	0.87	0.87	0.98
FA21	AAGGTTGGTCCTCCTCCACT	GCTTCCCTATTTCCCAGACC	7	0.85	0.85	0.96
FA22	ACCATCCAATGAGGATGAGC	GCCACATCTTTACCCTCTGC	5	0.76	0.76	0.90
FA23	AAGAAAGCGTGAAGGGCATA	TTGCCGTTCCTCTACAATCA	3	0.65	0.65	0.96
FA25	GTTTCCGCCTGTGAAAATGT	TGTTGATGGCGGTGTAGAAA	3	0.62	0.62	0.91
FA26	TGGCACATTTAGCAATGGAA	CCGGAATTGTAATGGTGGAC	5	0.79	0.79	0.97
FA27	GCAGAGTGGGGTAAGTCCAA	AAGTGGAAAGACCTGGATGG	2	0.50	0.50	1.00
FA28	TTCTTGGCAACCGTCTCTTT	TGCCTCATTTCAGCACATTC	8	0.87	0.87	0.97
FA30	TGGTTCCAAAACAACACCAA	ACTTGACCCCAGCTTCGTC	3	0.67	0.67	1.00
PM1	GCGAAGCATAACTGGCAGTATCTG	GCGGGCCTAGGTGATCTTGGA	8	0.85	0.85	0.94
PM2	AGGAGAAGACCGGCTGTGTA	TGCCTATAGCTGTGGCTGTG	8	0.86	0.86	0.95
PM3	CAGTTTGCCCAACAACAAGG	TTGATGGCAACAAATCACG	8	0.86	0.86	0.96
PM4	CCAAATTCAAATTCCTCTTTCC	GCCGAAAAACTCAAACTACCC	8	0.85	0.85	0.93
PM5	GCAACAAAGGAGGTTAGAGTCG	TGGTGAGTGCTCATTGTT	8	0.86	0.86	0.97
PM6	TGGAAACATTCTTACATAGCCAAA	CAGACGAGTCCTTCATGTGC	8	0.87	0.87	0.97
PM7	GAGCCTGCTACGCTTTTCTATG	CCTCTGATTCGATGATTTGCT	8	0.86	0.86	0.95
PM8	ACCGACAGCTGAGTTAGAGGAG	AGTCATAGGACCCCACTTCAAA	8	0.86	0.86	0.98
PM9	GCGAGGCGATCATGGAGAGA	GCGTTTCCTACGTCCCAATAAATC	8	0.87	0.87	0.99
PM10	GCGTCAGCCGTAGTGATGTAGCAG	GCGCCAGCCCCTCAAATATC	6	0.81	0.81	0.96
PM11	GCGGTCGCATTGAGTTGG AGGATA	GCGTAGCCAAACACCGATCTACC	8	0.86	0.86	0.94
PM12	GCGGAACCAAGCCAATAAGATG	GCGACCACGACAGTTTCTCACTCT	3	0.66	0.66	0.98
PM13	GCGGGCTGTCCACACTCCTTTCT	GCGATGCGTAAGTCTCTTCAAATA	8	0.87	0.87	0.97
PM15	CAAGCAATCCAACAGCTCAA	ACGCCTCTAAGCACTTCCTG	4	0.73	0.73	0.96
PM16	TGGGATCTGCTTAGGCTTTT	AAGCCACTTTTTACCCCTCAA	8	0.84	0.84	0.88
PM18	AAATCCTGTTCCTGCCAGTG	TGGTGACGTATTGGGTGATG	7	0.85	0.85	0.97
PM19	CGACCCAGCGACTACATTG	ACTTTAACCGCCACCAACTG	6	0.82	0.82	0.96
PM21	TTCGAAGATTGGAGAAGAAAGG	AAGCCACTTTTTACCCCTCAA	8	0.84	0.84	0.90
PM26	TCCGGTGACGAATCTAAAGG	GAAGAACAAGCACCACCACA	8	0.87	0.87	0.98
PM27	AGTCTAGGCTGCTTGGGTTG	CCAAGGGAAGAACAGACATGA	8	0.87	0.87	0.97
PM28	GCGGACCCATAGCACACTGTTGAC	GCGCCTTCCCTTGATACAACTTAC	8	0.87	0.87	0.96
PM29	GCGGGTGCTTAGGTTTTCACAACT	GCGCAAGTGGTATTTAAGGGTTAG	6	0.79	0.79	0.90
PM30	GCGCGCATAAGGCAACAAAG	GCGAATGGCAATGACATCTTCTCT	8	0.86	0.86	0.96
PM31	TGGTTTGCCGGTAGCAAATAGCAGCA	TGACACACACTCTCTCTGTCTGATCCCT	8	0.84	0.84	0.92
PM32	TGAACTGGTCCATCGGTGCTGAAA	TGATCACACAATACGCATTACCAAGCCT	4	0.72	0.72	0.94
PM33	GCAGTGCTACATCGACTCAGGTCCAA	ACCAAGGAAGTGCCGAAGTGGGTTT	8	0.86	0.86	0.94
PM34	AGGTGTCCAAAGAGGGTTGCTGTAGA	TCCCTCTCCCAATAACCCTTTGCTTC	7	0.82	0.83	0.92
PM37	CAATTTCCCGCCAAAAGTAA	GTTTGTTGGAGCTTCGAGCAAGTT	8	0.84	0.84	0.90
PM38	TCCTCTCCTTCTTTCCCTTCA	CGAGATCTCCCGAGACTGAG	8	0.84	0.85	0.94
PM50	TGAGCCCCAAGCCTAAAGTA	ACTGTTTTGCTTCCCCACAC	8	0.87	0.87	0.99
PM53	ATGGAACCTGCTCTCGACAT	TCGAAGAAGTCGTCAGGGAT	5	0.72	0.72	0.80
PM58	GTTTTGCTTGGAGGTGTAAAGG	GCTGCTGCTCTCTTGTAATGTG	6	0.83	0.83	0.98
PM61	ATGGAGAAGTTGCGTTTGG	CGTTCCAGACAAAAGAAGCA	7	0.84	0.84	0.93
Mean			6	0.78	0.78	0.95

Table 3. Genetic diversity parameters of the 24 Philippine strawberry germplasm grouped based on the region of origin

Pop	N	MLG	eMLG	SE	H	G	lambda	E.5	hexp	Ia	rbarD
South America	10	1	1	0	0	1	0	NaN	0.729	NaN	NaN
US	110	11	6.940	0.996	2.400	11.000	0.909	1.000	0.768	16.040	0.155
Asia (AS)	120	12	7.150	1.013	2.480	12.000	0.917	1.000	0.775	14.570	0.137
Total	240	24	8.45	0.978	3.18	24	0.958	1	0.774	9.74	0.0896

Strawberries are grown and propagated through clones or runners. This study used ten plants per cultivar, representing individuals per population. Additionally, the germplasm was grouped according to their region of origin to detect if clustering would be congruent to their geographical location before introduction in the Philippines. The diversity metrics present in the samples of each cultivar within and among the geographic origin are presented in Table 3. The total number of samples tested is 240, representing the 24 cultivars with 10 samples each (N). There are 24 Multilocus Genotypes (MLG) observed and 8.45 expected Multilocus Genotypes (eMLG) based on rarefaction. The Shannon-Wiener Index of MLG diversity (H) is 3.18, while Stoddart and Taylor's is 24. The Simpson's Index (lambda) generated is 0.958, Nei's unbiased gene diversity is 0.774, and the association index is 9.47. Results indicate that samples within cultivars produce similar alleles, and variation was absent within a specific population or cultivar. Having 24 MLGs indicates that samples within a population are real clone representatives and are genetically pure or similar to a cultivar. These results suggest a genetically diverse population with a high degree of variation in multilocus genotypes based on several indices when considered as one population (H=3.18) but show a moderate level of diversity when assessed based on continent of origin- either US or Asia (AS) with diversity of H=2.4.

Hierarchical clustering analysis

A dendrogram using UPGMA analysis was constructed based on the corresponding genetic similarity coefficient among the tested strawberry germplasm (Figure 1 and Table 4). Within the octoploid germplasm, the Red Milky Way (RMW) clustered from the rest with an average similarity index of 0.57. This germplasm was believed to have originated from China. RMW is most similar to a Korean variety with a 0.615 similarity index and least similar to Pistro (PST) from Japan with a similarity index of 0.509 (Table 4). RMW is designated as cluster 1 (Figure 1). Another cultivar clustered out from the rest is Summer Princess (SMP), which originated from Japan and was designated cluster 2. The most similar to the SMP is the Camarosa (CMA) cultivar introduced from the US with a similarity index of 0.710, while distantly similar to the Japanese cultivar (JPS) with an index of 0.563 (Table 4). The remaining 22 cultivars are grouped as cluster 3 (Figure 1), with an average similarity index of similarity of 0.697 (Table 4). Cluster 3 was further subdivided into six subgroups. Missionary (MSR) clustered away from the rest of the cultivars and was designated cluster III-F. Subcluster III-A comprises the Japanese strawberry (JPS) and the Korean cultivar (KRV). Subcluster III-B is comprised of Camarosa (CMA) and Erlibright (ERB) from the USA and Cavite Japan (CVJ), and Snow White (SNW) from Asia. Subcluster III-C consisted of Sweet Angel Bohol (SAB), White Strawberry (WST), Agsapa (AGP), and Pistro (PST) from Asian origin and Whitney (WTY) which originated from North America (Figure 1). SubclusterIII-D is represented by the Sweet Charlie (SC) cultivars. The most similar is Sweet Charlie Washington (SCW), with an index of 0.89, and the least similar to Sweet Charlie California (SCC), with 0.81 (Table 4). Among the genotyped SC, the most similar is SC California (SCC) with SC Hawaii (SCH) based on the value of 0.914, while the least similar is SCC with SC Washington (SCW). The average similarity among the Sweet Charlie cultivars is 0.838 while Sweet Charlie originating from North America (US) is 0.834 (Table 4). The open-pollinated variety (OPG) originating from the Philippines has the highest similarity with Winterdawn (WTD) from the US, as indicated by the similarity index value of 0.80. The fifth subcluster is III-E comprising of the Festival (FTV) cultivars, Honeoye (HNY), and Benehoppe (BNP) (Figure 1). Generally, the cultivars were grouped into one with only two outliers, both from Asian origin.

Principal Component Analysis

The SSR data were subjected to PCA to obtain an alternative view of the phylogenetic relationship among germplasm (Figure 2). PCA results coincided with results from cluster analysis (Figure 1). The first principal component explained 12.75% of the total marker variation, while the second principal component accounted for 9.06%. Subsequent principal components each explained less than 8% of the total variation. Most of the germplasm clustered into one group, while Red Milky Way and Summer Princes, originating from Asia, deviated from other germplasm similar to the results obtained using UPGMA analysis. Based on the PCA and UPGMA analyses, it was evident that there was no geographical distribution pattern was observed indicating a high genetic crossing and gene flow among individuals of different populations.

Population structure analysis

Population **s**tructure analysis was conducted to determine the genetic structure of *F. ananassa* germplasm. This analysis determines the most probable number of potential clusters (K) by computing the log-likelihood of the data for each K value. With the 24 germplasm genotyped in this study, the optimal K value generated was 3 (Figure 3.A). Structure analysis revealed that the cultivars available in the Philippine germplasm collection are all admixtures regardless of the countries where they were introduced (Figure 3.B).

Table 4. Jaccard similarity coefficient of the 24 Philippine strawberry germplasm

	SCA	SCW	BNP	HNY	FTV	WTY	SAB	MSR	RMW	JPS	KRV	WTS	PST	AGP	SCT	SMP	CMA	SCH	SCC	CVJ	ERB	SNW	WTD	OPG
SCA																								
SCW	0.886																							
BNP	0.714	0.729																						
HNY	0.751	0.759	0.759																					
FTV	0.725	0.717	0.678	0.745																				
WTY	0.732	0.736	0.712	0.748	0.728																			
SAB	0.737	0.734	0.700	0.736	0.723	0.780																		
MSR	0.675	0.678	0.668	0.711	0.662	0.683	0.668																	
RMW	0.565	0.579	0.583	0.575	0.549	0.533	0.564	0.611																
JPS	0.655	0.641	0.632	0.640	0.635	0.647	0.727	0.662	0.599															
KRV	0.680	0.654	0.668	0.652	0.633	0.640	0.671	0.644	0.615	0.700														
WTS	0.730	0.710	0.694	0.712	0.697	0.735	0.745	0.640	0.556	0.683	0.690													
PST	0.713	0.695	0.654	0.683	0.677	0.718	0.728	0.661	0.509	0.647	0.642	0.764												
AGP	0.736	0.726	0.712	0.723	0.672	0.724	0.741	0.683	0.593	0.664	0.714	0.776	0.748											
SCT	0.831	0.800	0.678	0.714	0.690	0.713	0.743	0.687	0.566	0.675	0.658	0.721	0.716	0.742										
SMP	0.602	0.612	0.585	0.612	0.596	0.590	0.596	0.590	0.583	0.563	0.586	0.619	0.613	0.649	0.632									
CMA	0.691	0.672	0.637	0.657	0.655	0.685	0.686	0.649	0.577	0.627	0.671	0.707	0.685	0.726	0.716	0.710								
SCH	0.838	0.802	0.689	0.724	0.727	0.721	0.746	0.696	0.569	0.672	0.674	0.731	0.727	0.750	0.870	0.635	0.769							
SCC	0.814	0.785	0.669	0.700	0.696	0.692	0.716	0.668	0.568	0.648	0.663	0.702	0.697	0.716	0.836	0.625	0.728	0.914						
CVJ	0.676	0.699	0.698	0.729	0.669	0.705	0.720	0.658	0.604	0.646	0.681	0.721	0.682	0.734	0.690	0.673	0.747	0.716	0.716					
ERB	0.734	0.706	0.712	0.743	0.703	0.724	0.734	0.682	0.574	0.656	0.688	0.713	0.699	0.734	0.698	0.651	0.721	0.725	0.706	0.730				
SNW	0.681	0.672	0.681	0.661	0.642	0.684	0.713	0.615	0.548	0.643	0.648	0.707	0.668	0.686	0.654	0.611	0.672	0.677	0.659	0.696	0.751			
WTD	0.751	0.723	0.686	0.720	0.695	0.745	0.709	0.659	0.569	0.619	0.652	0.731	0.703	0.730	0.734	0.614	0.689	0.764	0.733	0.690	0.707	0.665		
OPG	0.753	0.736	0.664	0.705	0.684	0.670	0.694	0.645	0.548	0.614	0.636	0.683	0.680	0.709	0.722	0.578	0.643	0.741	0.718	0.674	0.686	0.655	0.798	

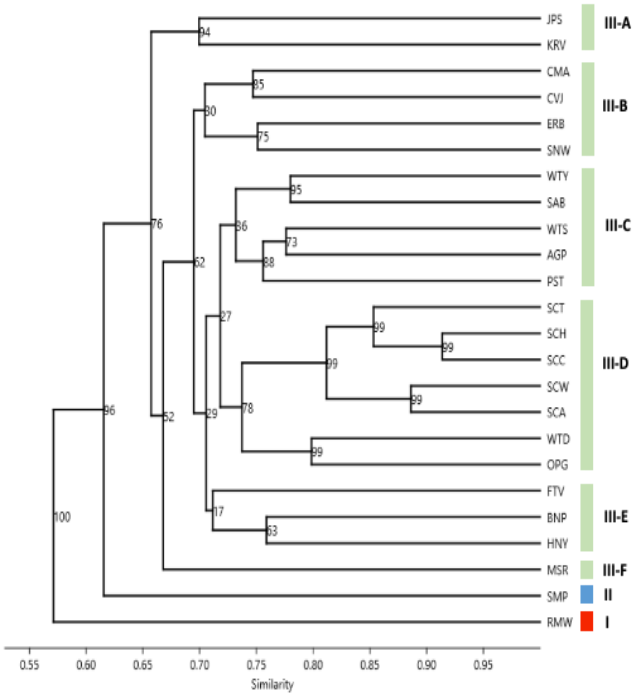

Figure 1. Dendrogram depicting the classification of 24 strawberry germplasm constructed using UPGMA analysis based on 131 SSR markers. The clusters are marked on the right side of the dendrogram. The scale at the bottom is the Jaccard's coefficient of similarity. Bootstrap values are given as nodal values along the branches

Discussion

There is currently limited information regarding the genetic characterization of the strawberry germplasm and cultivars available in the Philippines. Previous studies conducted adaptability trials that involved morphological characterization focusing on traits related to survival, growth, agronomic, and yield characteristics (Almorado et al. 2019; Mainem et al. 2021; Taylaran et al. 2023). The genetic information, focusing on the diversity and population structure of strawberry germplasm generated in this study, can assist breeders further in optimizing the development of varieties for adaption among farmers. This study used genome-wide SSR markers to assess the diversity of the strawberry germplasm available in the BSU germplasm collection.

Molecular markers in genetic studies

With the complete genome sequencing of *F. ananassa*, more SSR markers have been developed and utilized (Edger et al. 2020; Elec et al. 2023). SSRs are abundant, highly polymorphic molecular markers distributed throughout the genome. Their co-dominant nature makes them highly informative and provides an effective tool for studying the genetic diversity of several crops (Clark and Schreier 2017). The relatively high mean allele number highlights the ability of the SSR markers to detect multi-allelic variation across the strawberry accessions. The subset of markers generating six to eight alleles is highly informative for diversity analysis. Markers amplifying only two alleles are less polymorphic but may still provide value for pairwise germplasm discrimination when combined with markers from other loci. As detected in this study, the number of alleles ranging from one to eight is comparable to previous genetic studies (Lim et al. 2017; Biswas 2018; Bhowal et al. 2019). However, these observations are lower compared to other studies conducted by Kim et al. (2019) and Sharma et al. (2021), with ranges from 13 to as high as 42). Variations in allele numbers detected across these studies may be attributable to the number and type of genotypes characterized, SSR primers, PCR conditions, and genotyping

equipment. The evaluation of diversity in Philippine strawberry germplasm revealed substantial allelic variation at each locus assessed.

The genetic diversity of a population is crucial for its ability to adapt to different environments. Greater genetic diversity improves adaptability and resilience to environmental changes (Pincot et al. 2021; Fan and Whitaker 2023). In our study, we calculated diversity indices for 24 strawberry germplasm genotypes, including the Shannon-Wiener index (3.18), Simpson's index (0.958), and Nei's gene diversity (0.774). These indices indicate a moderate level of genetic diversity. Comparatively, other studies on strawberry germplasm collections from different countries have reported higher diversity levels. For instance, Njuguna et al. (2013) analyzed 268 North American and European germplasm using 34 SSRs and found a Shannon-Wiener index of 3.87 and Nei's gene diversity of 0.81, higher than our study's values of 3.18 and 0.774, respectively. These higher values reflect the broader cultivation regions sampled in their study. Similarly, in a study on 9 farmer-developed strawberry genotypes from Brazil and 20 cultivars from various global regions, Camargo et al. (2022) used 128 ISSR molecular markers; they found a high level of variability, with a Nei's gene diversity value of 0.33. Mezzetti et al. (2018) analyzed 179 strawberry accessions from central Europe using 19 SSRs. They reported higher diversity indices, including a Shannon-Wiener index of 4.1 and Nei's gene diversity of 0.83. Once again, the larger sample size contributed to increased diversity indices compared to our study's 24 Philippine germplasm samples. Furthermore, previous research has indicated low genetic diversity, with 71% of rare alleles being lost in heirloom varieties due to a small number of founders (Fan and Whitaker 2023). On the contrary, differences observed in strawberries can be attributed to geographical distance (Lim et al. 2017). Although our diversity indices may be slightly lower than other global collections, our study shows that the 24 germplasm samples from the Philippines still exhibit moderate genetic variation. This intermediate level of diversity present in these germplasm provides a balance that allows breeders to carefully select and utilize the genetic variability present in the elite cultivars and, at the same time, avoid genetic uniformity that could result in widespread crop failures if a particular pathogen or stress affects a genetically uniform crops.

Clustering analysis

The UPGMA cluster analysis based on the SSR genetic similarity matrix provided insights into the genetic relationships among the strawberry accessions. In this study, the groupings within the *F. ananassa* germplasm were separated at a 60% level of similarity. However, within the cultivated octoploid accessions, only Red Milky Way from China and Summer Princess from Japan separated from the other cultivars. This unique grouping indicates that these two cultivars are genetically distinct and may be attributed to their geographical origins. However, the other germplasm grown in the Philippines, which represents the other genotypes from Asia and America, did not cluster according to their geographical origin. Similarly, the studies by Zareii et al. (2021) and Hu et al. (2022) showed that the genotypes of the same origin did not necessarily group in the same cluster. However, further research is needed to explain these differences; it can be hypothesized that the genotypes' variability and response contributed to the clustering patterns. The close grouping of genotypes using SSR markers suggests close genetic relationships and frequent gene flow. These indicate genotypes adapted to the tropical environment (Bhowal et al. 2019). The current study provided contrastive results presented by previous diversity studies of strawberries (Lim et al. 2017; Zurn et al. 2022) concerning the correlation of the genetic distance with their geographic location. One possible reason is the continuous interbreeding and introgression among the current genotypes. In addition, adaptation to new environments can alter the genetics compared to the original source population, which dismantles distinct genetic signatures related to geographic locations.

In addition to cluster analysis, Principal Component Analysis (PCA) was utilized to examine genetic relationships among the strawberry accessions in an alternative multidimensional framework since PCA transforms the original variable space into new uncorrelated principal component axes that successively maximize variance (Clark and Schreier 2017). In the structure analysis, the germplasm was pre-grouped according to the origin of the introduction: South America (Sweet Charlie-Argentina), North America, and Asia. The clustering pattern largely mirrored the relationships observed in the UPGMA dendrogram (Figure 1). Consistent with its unique pedigree history, the cultivars Red Milky Way introduced from China occupied an outlier position on the PCA plot. This verifies its genetically distinct status based on the allele frequencies captured by the first two principal components. The Summer Princess cultivar from Japan also occupied a discrete position in the PCA space, confirming its divergence detected by clustering analysis (Cockerton et al. 2021). The remaining accessions formed a large central cluster, with internal structure reflecting the sub-groupings revealed in the dendrogram. The Sweet Charlie cultivars clustered tightly, similar to the Festival breeding lines.

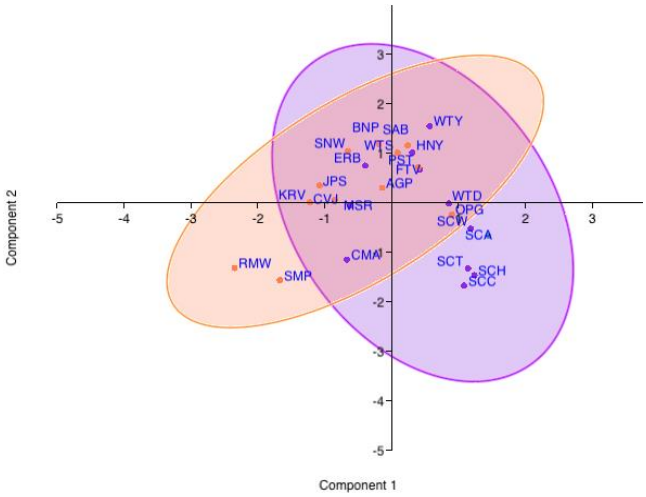

Figure 2. Principal Component Analysis (PCA) biplot evaluated using PAST (v4.14) software with a bootstrap frequency of n = 1000 conducted on 24 strawberry germplasm based on 131 SSR markers

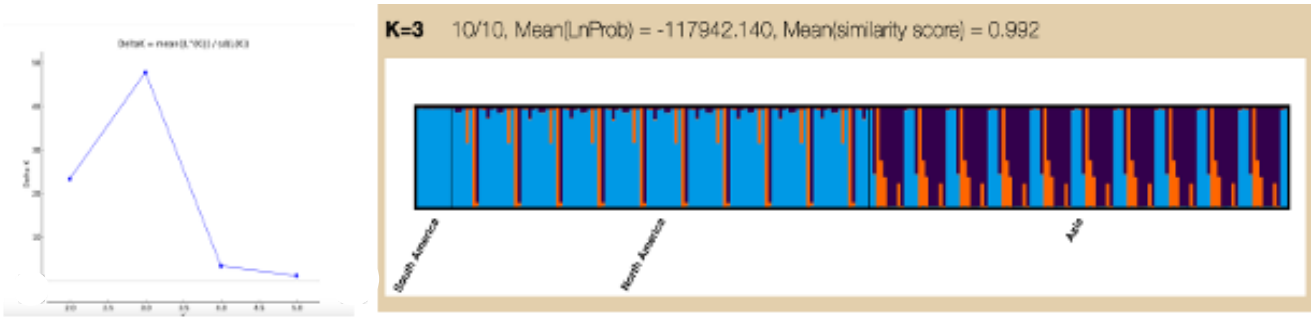

Figure 3. Graphical representation of estimation of the best subpopulation numbers based on the: A. Appropriate K value; B. Image representing the population structure for K=3 for the 240 strawberry samples representing 24 germplasm tested. Each colored bar represents one test object according to the group to which it belongs

Population structure analysis

The population genetic structure reveals the distribution pattern of genetic diversity within and among the population. Analysis of population structure using STRUCTURE and STRUCTURE HARVESTER detected three subpopulations corresponding to geographic origin when K=3. However, observations showed that the Philippine germplasm collections are admixtures regardless of the subpopulation identified. The PCA generated validated the results from UPGMA clustering wherein groupings were not based solely on geographic origin. Strawberries originating from Asia and America before the breeding program facilitated by the BSU have adapted well to the Cordillera environment over time, resulting in admixture between materials from different source regions. The strawberry germplasm in the Philippines may have experienced genetic drift. This could arise from factors like introducing a limited number of strawberry varieties from Asia and America, maintaining small, isolated populations, or restricted gene flow among different growing regions. Consequently, these random processes can lead to genetic differentiation and loss of diversity within subpopulations over time. Genetic bottlenecks might have also influenced the genetic makeup of the Philippine strawberry germplasm. Such bottlenecks could have occurred during the adaptation phase to the Cordillera environment, where intense selection pressures favored specific genotypes. Although initial genetic differentiation reflects geographic origins, the germplasm has subsequently undergone integration and adaptation within the Philippine environment (Bird et al. 2021).

Diversity and implications of breeding

The BSU breeding program started in 1999 and has generated advanced breeding lines from several crosses. From then, superior genotypes with better yield, high sugar content, and resistance to pests and diseases have already been developed and released, such as Agsapa and Baguio Pierre. Therefore, to continuously increase the genetic gain from breeding programs, it is important to optimize the genetic diversity present in the current population. In this study, we genetically characterize only 24 germplasm, a number considered to be a limited source of founders or genetic base that would serve as parentals and falls short based on previous recommendations regarding the maintenance of the optimum population size for any trait conservation for specific purposes and its utilization for crop improvement (Whitaker et al. 2020; Jyoti et al. 2023). In addition, the admixture observed in the germplasm pool may suggest continuous introgression within the germplasm resulting from continuous planting of these materials (Hardigan et al. 2021). These could potentially result in deleterious effects and loss of fitness of the population due to recombination between undesirable genes from outcross populations (Zurn et al. 2022).

The low to intermediate genetic variability among the current strawberry accessions can be rapidly recovered and improved through introgression from wild relatives and contemporary improved cultivars (Whitaker et al. 2020; Zurn et al. 2022). This also requires screening a small number of progenies as the genetics of elite germplasms are preserved. However, it is limited by the amount of diversity that can be incorporated in each designed backcross. The SSR genotyping successfully elucidated genetic relationships and distance among this strawberry germplasm set, demonstrating their usefulness for diversity studies. Ongoing efforts to characterize and enhance diversity will further strengthen the BSU breeding programs. Information on diversity and population structure will guide them in selecting parents that could utilize heterosis and select against with incompatible genotypes and with potential inbreeding depression results. Selecting genotypes with high level of diversity could mean that there might be traits in one of the parents that could have negative tradeoff in another trait. Also, yield and quality are controlled by several genes that go beyond the diversity of the germplasm.

In conclusion, this study successfully characterized the genetic diversity and population structure of strawberry germplasm available in the BSU Horticulture Research Department. Genome-wide SSR markers revealed high allelic variations among the strawberry germplasm, which could be utilized for future breeding programs. The study's parameters also showed a moderate diversity level among the 24 strawberry germplasm evaluated. Furthermore, clustering patterns of these germplasm were independent of their country of origin. These findings can assist in accurate germplasm identification and certification during the

propagation in nurseries and help plant breeders select parentals for different breeding objectives.

ACKNOWLEDGEMENTS

This research was supported by the Office of the Vice President for Academic Affairs-Emerging Inter-Disciplinary Research Program (OVPAA-EIDRP-C09-12), Philippines. The authors gratefully acknowledge the Benguet State University (BSU) HORTI collaborators for providing the strawberry germplasm for molecular characterization and the Plant Molecular Phylogenetics Laboratory group of the Institute of Biology, College of Science, for providing the necessary laboratory materials.

REFERENCES

Almorado RC. 2019. Adaptability of strawberry (*Fragaria ananassa* Duch) in Lanao Del Sur, Philippines. Intl J Humanit Soc Sci 11 (3): 25-32. DOI: 10.26803/ijhss.11.3.4.

Aristya GA, Kasiamdari RS, Setyoningrum R, Larasati B. 2019. Genetic variations of strawberry cultivars of *Fragaria* × *ananassa* and *Fragaria vesca* based on RAPD. Biodiversitas 20 (3): 770-775. DOI: 10.13057/biodiv/d200322.

Bhowal RR, Hossain MM, Kayesh E, Hasan M. 2019. Morphological and molecular characterization of tropical strawberry. Plant Tissue Cult Biotechnol 29 (2): 267-276. DOI: 10.3329/ptcb.v29i2.44515.

Bird KA, Hardigan MA, Ragsdale AP, Knapp SJ, VanBuren R, Edger PP. 2021. Diversification, spread, and admixture of octoploid strawberry in the western hemisphere. Am J Bot 108 (11): 2269-2281. DOI: 10.1002/ajb2.1776.

Biswas A. 2018. Characterization of strawberry (*Fragaria ananassa*) by genotyping and phenotyping. [Dissertation]. Delaware State University ProQuest Dissertations Publishing, US.

Camargo LKP, Pilati L, Zchonski FL, de Resende JTV, Da-Silva PR. 2022. Genetic diversity of Brazilian farmers-made strawberry genotypes and their relationship with commercial cultivars. Genet Resour Crop Evol 69: 1879-1888. DOI: 10.1007/s10722-022-01350-5.

Caro RES, Cagayan J, Gardoce RR, Manohar ANC, Canama-Salinas AO, Rivera RL, Lantican DV, Galvez HF, Reaño CE. 2022. Mining and validation of novel Simple Sequence Repeat (SSR) markers derived from coconut (*Cocos nucifera* L.) genome assembly. J Genet Eng Biotechnol 20 (1): 71. DOI: 10.1186/s43141-022-00354-z.

Chen J, Zheng Y-P, Guo W-J, Yang Z-M, Zhuang Z-H, Zhu B-Y. 2018. Identification of major and introduced strawberry cultivars in Fujian using SSR molecular markers. Fujian J Agric Sci 33 (2): 150-153. DOI: 10.19303/j.issn.1008-0384.2018.02.008.

Clark LV, Jasieniuk M. 2011. POLYSAT: An R package for polyploid microsatellite analysis. Mol Ecol Resour 11 (3): 562-566. DOI: 10.1111/j.1755-0998.2011.02985.x.

Clark LV, Schreier AD. 2017. Resolving microsatellite genotype ambiguity in populations of allopolyploid and diploidized autopolyploid organisms using negative correlations between allelic variables. Mol Ecol Resour 17 (5): 1090-1103. DOI: 10.1111/1755-0998.12639.

Cockerton HM, Karlström A, Johnson AW, Li B, Stavridou E, Hopson KJ, Whitehouse AB, Harrison RJ. 2021. Genomic informed breeding strategies for strawberry yield and fruit quality traits. Front Plant Sci 12: 724847. DOI: 10.3389/fpls.2021.724847.

Delgado H, Martín JP. 2023. Genetic diversity of black amaranth (*Amaranthus quitensis* Kunth) landraces of Ecuadorian highlands: Association genotypes - color morphotypes. Agriculture 13 (1): 34. DOI: 10.3390/agriculture13010034.

Earl DA, vonHoldt BM. 2012. Structure Harvester: A website and program for visualizing structure output and implementing the Evanno method. Conserv Genet Resour 4: 359-361. DOI: 10.1007/s12686-011-9548-7.

Edger PP, McKain MR, Yocca AE, Knapp SJ, Qiao Q, Zhang T. 2020. Reply to: Revisiting the origin of octoploid strawberry. Nat Genet 52 (1): 5-7. DOI: 10.1038/s41588-019-0544-2.

Elec VH, Cadorna CAE, Rey JD. 2023. In silico mining and validation of Simple Sequence Repeat (SSR) markers derived from the *Fragaria* x *ananassa* genome assembly. Sabrao J Breed Genet 55 (5): 1573-1586. DOI: 10.54910/sabrao2023.55.5.11.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: A simulation study. Mol Ecol 14 (8): 2611-2620. DOI: 10.1111/j.1365-294X.2005.02553.x.

Fan Z, Whitaker VM. 2023. Genomic signatures of strawberry domestication and diversification. Plant Cell 36 (5): 1622-1636. DOI: 10.1093/plcell/koad314.

FAOSTAT. 2023. UN food and agricultural organization statistical databases. https://www.fao.org/faostat/en/#data/QV.

Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchial F-statistics. Mol Ecol Notes 5 (1): 184-186. DOI: 10.1111/j.1471-8286.2004.00828.x.

Guo R, Xue L, Luo G, Zhang T, Lei J. 2018. Investigation and taxonomy of wild *Fragaria* resources in Tibet, China. Genet Resour Crop Evol 65: 405-415. DOI: 10.1007/s10722-017-0541-1.

Hammer O, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4 (1): 1-9. http://palaeo-electronica.org/2001_1/past/issue1_01.htm.

Hardigan MA, Feldmann MJ, Pincot DDA et al. 2021. Blueprint for phasing and assembling the genomes of heterozygous polyploids: Application to the octoploid genome of strawberry. BioRxiv Preprint 2021: 1-12. DOI: 10.1101/2021.11.03.467115.

Hilmarsson HS, Hytönen T, Isobe S, Göransson M, Toivainen T, Hallsson JH. 2017. Population genetic analysis of a global collection of *Fragaria vesca* using microsatellite markers. PLoS One 12 (8): e0183384. DOI: 10.1371/journal.pone.0183384.

Hu Y, Feng C, Yang L, Edger PP, Kang M. 2022. Genomic population structure and local adaptation of the wild strawberry *Fragaria nilgerrensis*. Hortic Res 9: uhab059. DOI: 10.1093/hr/uhab059.

Jyoti DE, Dubey RK, Sagar V, Verma RK, Singh PM, Behera TK. Vegetable peas (*Pisum sativum* L.) diversity: An analysis of available elite germplasm resources with relevance to crop improvement. 2023. Spanish J Agric Res 21: e0701-e0701. DOI: 10.5424/sjar/2023212-19457

Kaleybar BS, Nematzadeh GA, Ghasemi Y, Hamidreza S, Petroudi H. 2018. Assessment of genetic diversity and fingerprinting of strawberry genotypes using inter simple sequence repeat marker. Hortic Intl J 2 (5): 264-269. DOI: 10.15406/hij.2018.02.00062.

Kamvar ZN, Tabima JF, Grünwald NJ. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2: e281. DOI: 10.7717/peerj.281.

Kim HJ, Lee JN, Cho KS, Won HS, Suh JT. 2019. Genetic diversity and population structure analysis of ever-bearing and june-bearing strawberry cultivars using SSR markers. Korean J Hortic Sci Technol 37 (1): 108-118. DOI: 10.12972/kjhst.20190010.

Kumari S. 2019. Molecular characterization of different strawberry (*Fragaria* x *ananassa* Duch.) cultivars growing under mid hill conditions of Himachal Pradesh. Intl J Agric Environ Biotechnol 12 (4): 331-337. DOI: 10.30954/0974-1712.12.2019.6.

Lim S, Lee J, Lee HJ, Park K-H, Kim D-S, Min SR, Jang WS, Kim TI, Kim H. 2017. The genetic diversity among strawberry breeding resources based on SSRs. Sci Agric 74 (3): 226-234. DOI: 10.1590/1678-992X-2016-0046.

Mainem CAT, Agustin RQ, Duza GM. 2021. Growth and yield adaptability of selected varieties of strawberry (*Fragaria x ananassa*) under BPSU Abucay Bataan condition, Philippines. J Biol Environ Sci 18 (2): 23-30.

Mezzetti B, Giampieri F, Zhang Y-T, Zhong C-F. 2018. Status of strawberry breeding programs and cultivation systems in Europe and the rest of the world. J Berry Res 8 (3): 205-221. DOI: 10.3233/JBR-180314.

Nei M. 1978. Estimation of average heterozygosity ang genetic distance from a small number of individuals. Genetics 89 (3): 583-590. DOI: 10.1093/genetics/89.3.583.

Njuguna W, Liston A, Cronn R, Ashman T-L, Bassil N. 2013. Insights into phylogeny, sex function, and age of Fragaria based on whole chloroplast genome sequencing. Mol Phylogenet Evol 66 (1): 17-29. DOI: 10.1016/j.ympev.2012.08.026.

Padua DP, Ligat CS, Kilakil CM. 2016. Breeding strawberry for organic production. Proceedings of the ISTF2016 International Symposium of Tropical Fruits: Towards Sustainable fruit producton and global food Security. Davao City, Philippines, 23 Aug-3 Sept 2016. DOI: https://itfnet.org/istf2016/ConcurrentPaperSlide/4%20-%20Breeding%20Strawberry%20for%20Organic%20Production.pdf

Paradis E. 2010. Pegas: An R package for population genetics with an integrated-modular approach. Bioinformatics 26 (3): 419-420. DOI: 10.1093/bioinformatics/btp696.

Pielou EC. 1975. Ecological Diversity. John Wiley & Sons, New York.

Pincot DDA, Ledda M, Feldmann MJ, Hardigan MA, Poorten TJ, Runcie DE, Heffelfinger C, Dellaporta SL, Cole GS, Knapp SJ. 2021. Social network analysis of the genealogy of strawberry: Retracing the wild roots of heirloom and modern cultivars. G3: Genes, Genomes, Genetics 11 (3): jkab015. DOI: 10.1093/g3journal/jkab015.

Porter M, Fan Z, Lee S, Whitaker VM. 2023. Strawberry breeding for improved flavor. Crop Sci 63: 1949-1963. DOI: 10.1002/csc2.21012.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. 2000. Association mapping in structured populations. Am J Hum Genet 67 (1): 170-181. DOI: 10.1086/302959.

Qarni A, Muhammad K, Wahab A, Ali A, Khizar C, Ullah I, Kazmi A, Sultana T, Hameed A, Younas M, Rahimi M. 2022. Molecular characterization of wild and cultivated strawberry (*Fragaria* × *ananassa*) through DNA barcode markers. Genet Res 2022: 9249561. DOI: 10.1155/2022/9249561.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.r-project.org/.

Salinas NR, Zurn JD, Mathey M, Mookerjee S, Denoyes B, Perrotte J, Potier A, Finn CE, Hancock JF, Stewart P, Bassil NV. 2017. Validation of molecular markers associated with perpetual flowering in octoploid *Fragaria* germplasm. Mol Breed 37: 70. DOI: 10.1007/s11032-017-0672-2.

Sharma S, Kaur R, Kumar K, Prasad H. 2021. Genetic variability in strawberry (*Fragaria x ananassa* Duch.) cultivars assessed by morphological traits and EST-SSR markers of *Rubus ellipticus*. Indian J Biotechnol 20: 81-90.

Simpson EH. 1949. Measurement of diversity. Nature 163: 688. DOI: 10.1038/163688a0.

Taylaran RD, Gonzaga Jr AB, Lantud JMM, Silverio BGS. 2023. Runner production of strawberry (*Fragaria x ananassa* Duch) production under greenhouse and open field conditions in Misamis Oriental, Philippines. IOP Conf Ser: Earth Environ Sci 1208: 012028. DOI: 10.1088/1755-1315/1208/1/012028.

Testolin R, Messina R, Cipriani G, De Mori G. 2023. SSR-based DNA fingerprinting of fruit crops. Crop Sci 63 (2): 390-459. DOI: 10.1002/csc2.20896.

Vieira MLC, Santini L, Diniz AL, de Freitas Munhoz C. 2016. Microsatellite markers: What they mean and why they are so useful. Genet Mol Biol 39 (3): 312-328. DOI: 10.1590/1678-4685-GMB-2016-0027.

Wada T, Oku K, Nagano S et al. 2017. Development and characterization of a strawberry MAGIC population derived from crosses with six strawberry cultivars. Breed Sci 67: 370-381. DOI: 10.1270/jsbbs.17009.

Whitaker VM, Knapp SJ, Hardigan MA, Edger PP, Slovin JP, Bassil NV, Hytönen T, Mackenzie KK, Lee S, Jung S, Main D, Barbey CR, Verma S. 2020. A roadmap for research in octoploid strawberry. Hortic Res 7: 33. DOI: 10.1038/s41438-020-0252-1.

Xin Y, Fang X, Wang S, Tong J, Lai W, Wang J, Yu H. 2019. Comparison of Simple Sequence Repeat (SSR) and sequence related amplified polymorphism (SRAP) markers for genetic diversity analysis in strawberry. J Zhejiang Univ Agric Life Sci 45 (3): 278-287. DOI: 10.3785/j.issn.1008-9209.2018.10.261.

Zareei E, Karami F, Aryal R, Saed-Moucheshi A. 2021. Genetic analysis and cross associations of strawberry (*Fragaria* × *ananassa*) genotypes and morpho-biochemical features: A practical breeding program. Research Square. DOI: 10.21203/rs.3.rs-403398/v1.

Zhang J, Yang J, Lv Y, Zhang X, Xia C, Zhao H, Wen C. 2023. Genetic diversity analysis and variety identification using SSR and SNP markers in melon. BMC Plant Biol 23 (1): 39. DOI: 10.1186/s12870-023-04056-7.

Zhong Y, Cheng Y, Ruan M, Ye Q, Wang R, Yao Z, Zhou G, Liu J, Yu J, Wan H. 2021. High-throughput SSR marker development and the analysis of genetic diversity in *Capsicum frutescens*. Horticulturae 7 (7): 187. DOI: 10.3390/horticulturae7070187.

Zurn JD, Hummer KE, Bassil NV. 2022. Exploring the diversity and genetic structure of the US.S. National Cultivated Strawberry Collection. Hortic Res 9: uhac125. DOI: 10.1093/hr/uhac125.