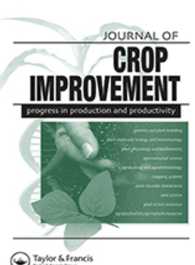


Journal of Crop Improvement

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/wcim20

Assessing the genetic diversity of taro germplasm collection in Vietnam using simple sequence repeat markers

Ba Hoanh Nguyen, Thi Tuoi Le, Thi Bich Huyen Vu & Viet Xuan Nguyen

To cite this article: Ba Hoanh Nguyen, Thi Tuoi Le, Thi Bich Huyen Vu & Viet Xuan Nguyen (2024) Assessing the genetic diversity of taro germplasm collection in Vietnam using simple sequence repeat markers, Journal of Crop Improvement, 38:5, 440-458, DOI: 10.1080/15427528.2024.2361446

To link to this article: https://doi.org/10.1080/15427528.2024.2361446

View supplementary material

Published online: 13 Jun 2024.

Submit your article to this journal

Article views: 85

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=wcim20

JOURNAL OF CROP IMPROVEMENT
2024, VOL. 38, NO. 5, 440–458
https://doi.org/10.1080/15427528.2024.2361446

RESEARCH ARTICLE

Assessing the genetic diversity of taro germplasm collection in Vietnam using simple sequence repeat markers

Ba Hoanh Nguyen[a], Thi Tuoi Le[b], Thi Bich Huyen Vu[b], and Viet Xuan Nguyen[b]

[a]Vinh University, Vinh, Nghe An, Vietnam; [b]Hanoi National University of Education, Hanoi, Vietnam

ABSTRACT

This study evaluated the genetic diversity of 253 taro (*Colocasia esculenta* (L.) Schott) germplasm accessions collected in Vietnam at the molecular level using 20 Simple Sequence Repeat (SSR) markers. High genetic diversity in the taro germplasm accessions were detected with 100% polymorphic SSR loci, and 89 alleles were obtained at the 20 SSR loci with an average of 4.45 alleles per locus, of which 14 markers had high polymorphism (3 alleles/locus). Polymorphic Information Content (PIC) values obtained at 20 SSR loci ranged from 0.31 (Ce0078) to 0.85 (HK34), with an average of 0.65. The genetic similarity coefficient within the accessions ranged from 0.39 to 1.00. All 253 accessions were classified into 2 groups at a genetic similarity level of 0.60; one of the groups (group 2) was divided into 8 sub-groups at a genetic similarity level of 0.64. Some accessions (such as 28279 and SP-19-009; SP-19-061 and SP-19-023; 28278 and SP-19-006) with a high genetic similarity coefficient (1.00) showed a very close genetic relationship, while others (such as 10063 and 11544) with a low genetic similarity coefficient (0.39) showed a fairly distant genetic relationship. The study's results provide a foundation for preserving and breeding taro in Vietnam. Five taro accessions (28211, SP-19-017, T.3578, T.3515, and 10098) carrying specific alleles that could be identified by 6 SSR markers (uq55–112, uq73–164, HK35, Ce0078, HK22, and HK31) were detected. Furthermore, the SSR marker approach showed the potential application in identifying genetic relationships and distinguishing different taro cultivars.

ARTICLE HISTORY
Received 30 January 2024
Accepted 26 May 2024

KEYWORDS
SSR polymorphism; *Colocasia esculenta* (L.) Schott; PIC values; similarity coefficient; specific alleles

Introduction

Taro (*Colocasia esculenta* (L.) Schott) belongs to the Araceae family and is produced mainly for its underground stems which contains a high percentage of starch, minerals, vitamins, and nutrients. Taro with a long history of cultivation is an important staple food in developing countries (such as India, Pacific Islands, and other Southeast Asian countries) (Harun et al. 2022; Temesgen and Retta 2015; Thokchom et al. 2017). According to a statistical

CONTACT Ba Hoanh Nguyen hoanhnb@vinhuni.edu.vn Vinh University, Vinh, Nghe An 43100, Vietnam; Viet Xuan Nguyen vietnx@hnue.edu.vn Hanoi National University of Education, Cau Giay, Hanoi 11300, Vietnam

Supplemental data for this article can be accessed online at https://doi.org/10.1080/15427528.2024.2361446

© 2024 Informa UK Limited, trading as Taylor & Francis Group

database from the Food and Agriculture Organization of the United Nations (FAO), in 2021, global taro production reached 12,396,248.5 tons, in which Africa contributed 77.30% of that output, Asia contributed 18.60%, while Oceania and the Americas contributed 3.40% and 0.70%, respectively (FAOSTAT 2023). In Vietnam, taro is the fourth most important root crop after potato (*Solanum tuberosum* L.), sweet potato (*Ipomoea batatas* (L.) Lam), and cassava (*Manihot esculenta* Crantz). Taro is widely grown in most different ecological regions of Vietnam and has become a precious specialty crop in some localities (such as Son La, Yen Bai) (T. N. H. Nguyen et al. 2010; V. X. Nguyen et al. 2023). Building a DNA database of indigenous taro germplasm collection is crucial for breeders to exploit breeding materials, build core collections, and restore varieties. Genetic diversity studies have shown that distinct gene pools exist in different regions where taro is distributed indicating the need to capture and preserve the genetic diversity of taro. Molecular techniques, such as Simple Sequence Repeat (SSR) genotyping, have been proven effective in characterizing taro genetic diversity and can be used to develop a comprehensive DNA database. This database can be seen as a valuable resource for breeders to select the most divergent parents for breeding programs, develop high-yielding and disease-resistant varieties, and conserve the genetic resources of taro. By using genetic information stored in DNA databases, breeders can make informed decisions in the sustainable management and use of taro genetic resources (Ab Razak et al. 2021; Alam et al. 2019; Devi 2012).

Genetic diversity studies have a very important role in breeding and conservation. In taro, the studies in genetic diversity previously focused on the morphological characteristics (Bammite et al. 2018; Harun et al. 2022), karyotype (Das et al. 2015; V. X. Nguyen et al. 2021) and isozyme variation (V. X. Nguyen, Yoshino, and Tahara 1998; Ochiai et al. 2001). Recently, DNA markers have been widely used to contribute to the classification of the plant and determine the population structure and evolutionary relationships, which provide a basis for building core collections of taro for conservation work (Devi 2012; S. Singh et al. 2012). DNA markers, including Amplified Fragment Length Polymorphism (AFLP) (Kreike, Van Eck, and Lebot 2004; Sharma, Mishra, and Misra 2008a), Random Amplified Polymorphic DNA (RAPD) (Das et al. 2015; Li et al. 2010; Ochiai et al. 2001), Simple Sequence Repeats (SSR)/microsatellite (Alam et al. 2019; Harun et al. 2022; Mace and Godwin 2002; Mezhii et al. 2017; Nunes et al. 2012; Valerie and Edome 2016), Inter-Simple Sequence Repeat (ISSR) (Nair et al. 2018; Okpul et al. 2005), and Single Nucleotide Polymorphism (SNP) (Fufa et al. 2022; Z. Wang et al. 2020) have been applied. Among them, SSR markers have potential applications in genetic diversity in plants, including taro. They can be used for genetic mapping, genetic identity verification, and analysis of genetic relationships between individuals (Valerie and Edome 2016).

SSR markers are highly polymorphic, co-dominant with Mendelian inheritance, and located within genes, which provide valuable information about functional genetic variation and are suitable for studying genetic diversity (Lu, Niu, and Li 2010). In addition, SSR markers have been widely used in plant genome analysis and have been genetically mapped in several important plant species (Victoria, da Maia, and de Oliveira 2011). Therefore, SSR markers are a comprehensive source for identifying the genotypes and genetic diversity of taro (Ab Razak et al. 2021; Hu et al. 2009; Mace and Godwin 2002; Mezhii et al. 2017; D. Singh et al. 2008; Valerie and Edome 2016; L. Wang et al. 2017).

In Vietnam, studies have only focused on preliminary assessments of agro-morphological characteristics of taro (Doan and Nguyen 2019; T. N. H. Nguyen et al. 2010). Also, the genetic diversity of taro in Vietnam at the cytological (chromosomal) (V. X. Nguyen et al. 2021) and molecular levels (using RAPD and SSR markers) has been studied (Dang and Nguyen 2014; V. X. Nguyen, Dang, and Ngo 2015; V. X. Nguyen et al. 2023). However, the application of analysis of genetic diversity in the breeding of taro is still limited. Therefore, studies using SSR polymorphic markers to evaluate the genetic diversity of taro, especially the conserved collections of taro germplasm, are necessary. This can provide a molecular database to select taro germplasm with high genetic polymorphism, thereby establishing a core collection for long-term and effective conservation. This study reports the genetic diversity of a taro germplasm collection in Vietnam with 253 accessions preserved at the Plant Resources Center of Vietnam based on polymorphisms of 20 SSR molecular markers, thereby building a DNA database, evaluating genetic relationships, forming taro potential germplasm for breeding and conservation.

Materials and methods

Plant materials and molecular markers

Two hundred fifty-three accessions of taro germplasm, which were collected from the agro-ecological regions of Vietnam and maintained in an *ex-situ* gene bank at the Plant Resources Center, Vietnam Academy of Agricultural Sciences were used to analyze genetic diversity using SSR primers. Information on taro accessions of Vietnam taro collection used in the study is presented in Table 1 and S1. Twenty SSR primer pairs were selected from the published taro genome database to evaluate the genetic diversity of the taro accessions. Information about the sequence, annealing temperature, and reference source of SSR primers is presented in Table 2.

Table 1. Number of 253 taro (Colocasia esculenta (L.) Schott) accessions collected in the provinces of Vietnam.

Collected provinces	No. of accession	Agro-ecological regions	Agro-ecological subregions
Lai Chau	17	Northern Midland and Mountainous	Northwest
Dien Bien	25	Northern Midland and Mountainous	Northwest
Son La	35	Northern Midland and Mountainous	Northwest
Hoa Binh	14	Northern Midland and Mountainous	Northwest
Cao Bang	16	Northern Midland and Mountainous	Northeast
Lang Son	11	Northern Midland and Mountainous	Northeast
Bac Kan	7	Northern Midland and Mountainous	Northeast
Thai Nguyen	4	Northern Midland and Mountainous	Northeast
Quang Ninh	9	Northern Midland and Mountainous	Northeast
Bac Giang	14	Northern Midland and Mountainous	Northeast
Lao Cai	17	Northern Midland and Mountainous	Northeast
Yen Bai	7	Northern Midland and Mountainous	Northeast
Ha Giang	12	Northern Midland and Mountainous	Northeast
Tuyen Quang	26	Northern Midland and Mountainous	Northeast
Phu Tho	12	Northern Midland and Mountainous	Northeast
Bac Ninh	1	Red River Delta	
Ha Noi	15	Red River Delta	
Ninh Binh	2	Red River Delta	
Others	9	North Central, Central Highlands Southeast	

Genomic DNA extraction

Total DNA was extracted from 500 mg of young leaf tissue of each taro vegetative sample. Taro leaf samples were dried in silica gel, and DNA extraction was performed following the protocol of Sharma, Mishra, and Misra (2008b) with slight modification (added sodium acetate 3 M during the extraction process of chloroform: isoamyl alcohol to remove polysaccharides and proteins). The concentration and quality of DNA samples were quantified with a NanoDrop-2000 spectrophotometer, and DNA samples were adjusted to 50 ng/mL.

SSR-PCR and electrophoresis

Twenty SSR primers were selected for the study based on preliminary assays of amplification and product length polymorphism in taro genotypes (Table 1) (Hu et al. 2009; Mace and Godwin 2002; Mezhii et al. 2017; L. Wang et al. 2017; You et al. 2015). The PCR reaction was carried out in a Veriti 96-well Thermal Cycler. Each 20 µL reaction mix contained 2 µL of 10× PCR buffer, 1.6 µL dNTPs 2.5 mM, 1.4 µL forward and reverse primers (25 ng/µL), 5 µL total DNA (5 ng/µL), 0.1 µL green Taq polymerase (5 U/µL), and 9.9 µL ddH_2O. Thermal cycling conditions for PCR were as follows: 95°C for 5 min; followed by 35 cycles (denaturation at 94°C for 1 min, annealing ranging from 53 to 62°C for 1 min, and extension at 72°C for 1 min), and a final extension step at 72°C for 10 min. The amplification products were

Table 2. Information on the SSR primers used in this study.

S/N	Primer name	Primer sequence (5′→3′)	Annealing temperature (°C)	Reference source
1.	uq55–112	F: CTTTTGTGACATTTGTGGAGC R: CAATAATGGTGGTGGAAGTGG	55	Mace and Godwin (2002)
2.	uq73–164	F: ATGCCAATGGAGGATGGCAG R: CGTCTAGCTTAGGACAACATGC	60	Mace and Godwin (2002)
3.	uq84–207	F: AGGACAAAATAGCATCAGCAC R: CCCATTGGAGAGATAGAGAGAC	62	Mace and Godwin (2002)
4.	uq91–262	F: GTCCAGTGTAGAGAAAAACCAG R: CACAACCAAACATACGGAAAC	55	Mace and Godwin (2002)
5.	AC3	F: AGTGGCATCAATGGAGGA R: CCACTAAACGACGACCCAC	62	Hu et al. (2009)
6.	HK7	F: GTTGTCCGCCTGTGCGTTCT R: CTCTTGGGAATTCTCCGGGTG	55.5	Hu et al. (2009)
7.	HK22	F: ACATCAAACCTCTGGTGGGC R: AGCAATCCTAGCCGAGGTG	62	Hu et al. (2009)
8.	HK31	F: TACCGCCGAGTGCTTATC R: TACGGCTGGAATCAAAGC	62	Hu et al. (2009)
9.	HK34	F: TTACTCCAAACGAGGCAAAC R: CCTTCAAGATGTTACCAAATGC	58	Hu et al. (2009)
10.	HK35	F: TACTAGAACCCCGTCAGTCT R: CGTCGATTTATCAGTGAGC	58	Hu et al. (2009)
11.	HK38	F: AAACGCGGCCAGAAGATC R: GAATAGCGGAACAAGGTAGA	61	(Hu et al. 2009)
12.	CE-EST-SSR8	F: AAGGGGAATTTGGAGAGACG R: GAGGAAGACCCAGGAGAAGG	53	You et al. (2015)
13.	CE-EST-SSR24	F: CATGGAAGGAGACCCCTGTA R: CCAAGCATTGCAGGGACTAT	53	You et al. (2015)
14.	CE-EST-SSR67	F: CGTAGCAAAGTTGGTGTGGA R: GGGAAGAGAGAGGAAGCGTT	55	You et al. (2015)
15.	CE-EST-SSR87	F: ATTAGGGCTTTGGAGGAGGA R: CTTAAAACCCCCGCCTTATC	57	You et al. (2015)
16.	COL-GCC206-122	F: CGTTCAACACAGACCACTAC R: TCCTTTGGAAAGGAGGTCC	58	Mezhii et al. (2017)
17.	COL-GCC208-253	F: TAGAGGGTGGACAGGAG R: CTAGAGGCACTGATGTAAC	57	Mezhii et al. (2017)
18.	Ce0078	F: TAGCATTATTGGATCACCATCCT R: CAAATCTAAAGCTGGGCGTTTAT	53	L. Wang et al. (2017)
19.	Ce0149	F: TCTTGTCCTGTGAAACTGAGGA R: GCATGATAAGCTCTCACGACGTA	57	L. Wang et al. (2017)
20.	Ce0341	F: GTAACTAGGTCACCACTGCAAGG R: GAGCAGTGATCAATACTGGAGC	55	L. Wang et al. (2017)

separated in an 8% denaturing polyacrylamide gel. The gels were stained in 0.5% ethidium bromide and photographed with black and white film under UV light. The molecular size of the DNA fragments was estimated by comparison with the standard DNA ladder (Fermentas).

Molecular data analysis

For genetic diversity and phylogenetic relationship determination, all gels were scored in a binary format where the presence of a band was 1 and the

absence was 0. The binary data were analyzed using NTSYSp.c software (version 2.1) to calculate the similarity matrix between pairs of samples and draw tree diagrams of genetic subgroups in genetic resources (Rohlf 2000).

Similarity matrices were calculated from these data based on Jaccard's similarity coefficient (Jaccard 1908): $J_{ij} = a/(n - d)$, where a is the number of DNA bands present in both genetic samples i and j; d is the number of DNA bands that are absent in both genetic samples i and j; n is the total number of DNA bands obtained; J_{ij} is the Jaccard similarity coefficient between two genetic samples i and j.

Polymorphic Information Content (PIC) was used to evaluate the efficiency of primer amplification for each locus (i) calculated according to the formula: $PIC\ (i) = 1 - \sum h_k^2$, where h_k is the frequency of allele k in the population (Mohammadi and Prasanna 2003).

Determination of specific alleles (specific alleles in primer pairs with genetic source samples) was carried out as described by Pervaiz et al. (2009).

Results

SSR polymorphism of taro germplasm collection in Vietnam

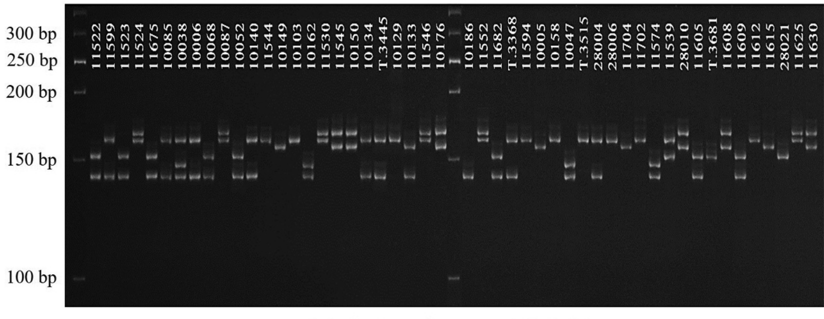
DNA quality is the first important requirement in research on genetic diversity using molecular markers. In this study, the DNA extraction process was optimized for taro leaf samples, and the total DNA of 253 taro accessions was successfully extracted (Figure S1). The quality of the DNA template was detected using a UV spectrophotometer and agarose gel electrophoresis. PCR reactions were performed using the genomic DNA template with the OD from 1.8 to 2.0.

PCR analysis using 20 SSR markers was conducted to screen the genetic diversity of taro germplasm based on polymorphism. PCR-SSR products were electrophoresed on 8% polyacrylamide gel. The results are presented in Figure 1 and Table 3.

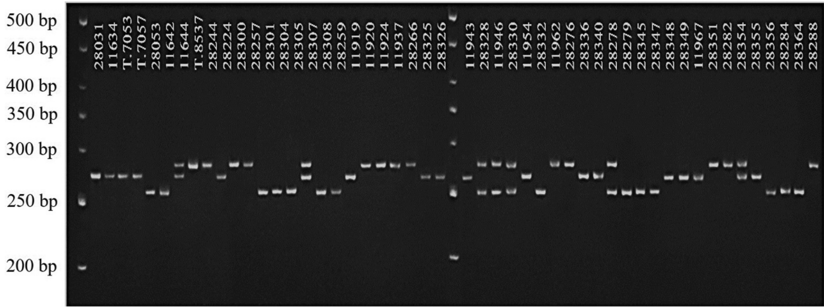
With a total of 5,060 PCR reactions, the product DNA bands were clear, and sharp and showed polymorphism (Figure 1).

The polymorphism data of 20 SSR loci for each accession in the taro germplasm collection (Table S2) showed that the PCR products of the taro samples were DNA bands with sizes ranging from 90 bp to 287 bp. However, the common size of the obtained alleles ranged from 130 bp to 200 bp. At each locus, the size of the obtained alleles fluctuated from 5 bp (locus Ce0341) to 100 bp (locus uq84–207 and locus HK35) (Figure 2).

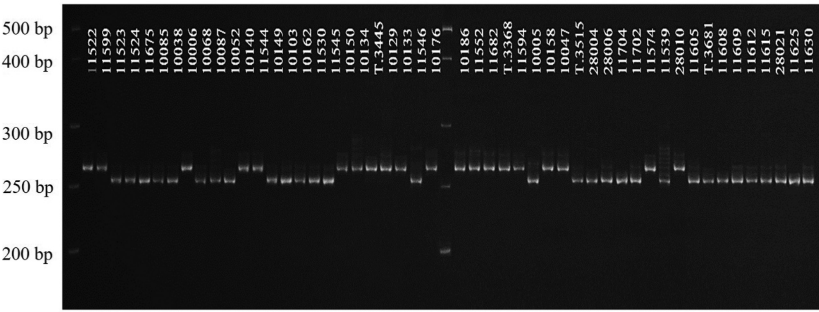
The polymorphism of each SSR primer pair represents the presence and absence of the DNA band of that primer pair in each individual in the population. Data in Table 3 shows that the number of alleles successfully



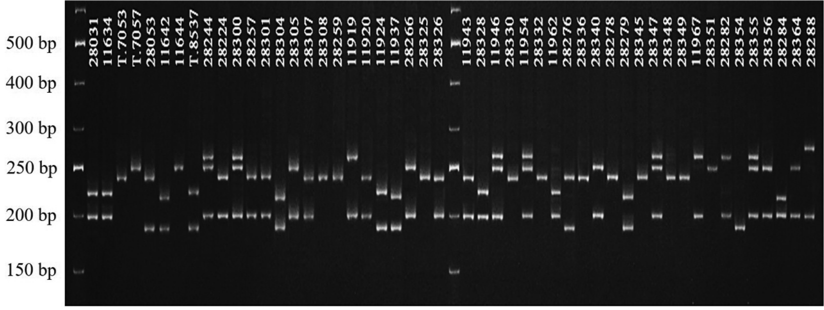
(a) SSR primer uq73-164



(b) SSR primer CE-EST-SSR8



(c) SSR primer COL-GCC208-235



(d) SSR primer HK34

Figure 1. SSR polymorphism detected in taro (*Colocasia esculenta* (L.) Schott) accessions using SSR molecular markers through electrophoresis of PCR-SSR products on 8% polyacrylamide gel.

Table 3. Number of DNA bands obtained and SSR polymorphism information of taro (*Colocasia esculenta* (L.) Schott) germplasm in Vietnam.

S/N	SSR locus	Alleles number	Polymorphic alleles	Size of bands (bp)			Common allele frequency (%)	PIC value
				low	high	Most frequent		
1.	uq55–112*	3	3	110	120	115	57.40	0.49
2.	uq73–164*	7	7	143	240	165	22.89	0.81
3.	uq84–207	5	5	130	230	140	37.54	0.74
4.	uq91–262	4	4	240	275	250	40.87	0.67
5.	AC3	4	4	120	165	130	29.17	0.75
6.	HK7	5	5	120	200	120	23.44	0.80
7.	HK22*	6	6	175	210	185	36.21	0.74
8.	HK31*	6	6	116	160	122	28.61	0.79
9.	HK34	8	8	190	275	200	20.83	0.85
10.	HK35*	5	5	180	280	260	37.71	0.71
11.	HK38	4	4	130	140	130	38.50	0.71
12.	CE-EST-SSR8	3	3	260	287	287	35.20	0.67
13.	CE-EST-SSR24	3	3	250	275	260	43.25	0.63
14.	CE-EST-SSR67	4	4	170	185	175	39.76	0.69
15.	CE-EST-SSR87	7	7	130	215	145	39.59	0.76
16.	COL-GCC206-122	2	2	118	125	118	50.18	0.50
17.	COL-GCC208-235	2	2	255	265	255	77.91	0.34
18.	Ce0078*	4	4	130	190	135	81.97	0.31
19.	Ce0149	5	5	90	115	98	51.55	0.67
20.	Ce0341	2	2	175	180	180	78.45	0.34
Average		**4.45**					**43.57**	**0.65**

(*): Locus has the specific allele

amplified with SSR primers is very different between loci. The total number of alleles obtained at the 20 SSR loci was 89 alleles. The number of polymorphic alleles at each locus varied from 2 alleles (locus COL-GCC208-235) to 8 alleles (locus HK34), averaging 4.45 alleles per locus. There were up to 14 SSR markers with polymorphism levels higher than 3 alleles per locus. These markers can be used for molecular breeding of taro germplasm. Among them, the HK34 locus had the highest number of alleles at 8 alleles, followed by 2 loci CE-EST-SSR87 and uq73–164 for 7 alleles, 2 loci HK22 and HK31 for 6 alleles, and 4 loci Ce0149, uq84–207, HK35 and HK7 for 5 alleles. Common allele frequency ranged from 20.83% to 81.97%, locus Ce0078 had the highest common allele frequency of 81.97%; next were Ce0341 and COL-GCC208-235 loci with common allele frequencies of 78.45% and 77.91%, respectively.

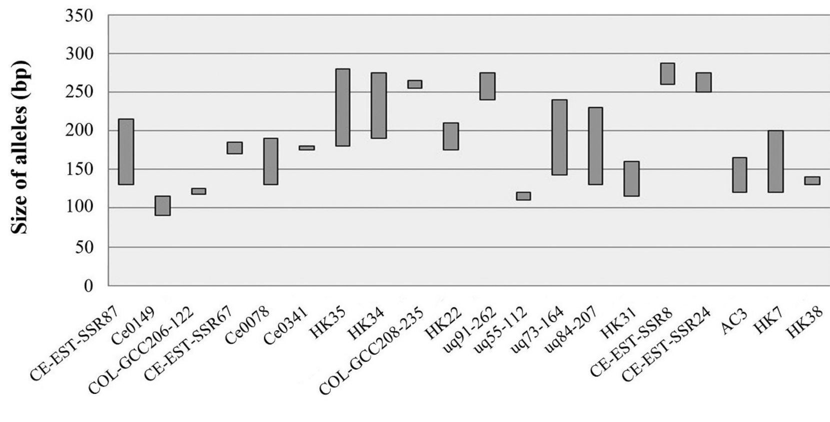


Figure 2. Alleles size fluctuations within each locus and among the 20 SSR loci used in the study. The size of the obtained alleles fluctuates from 5 bp (Ce0341 locus) to 100 bp (uq84–207 and HK35 loci).

The PIC of primers is considered a measure of the genetic diversity of alleles at each SSR locus (DeWoody, Honeycutt, and Skow 1995). Results showed that the PIC value obtained at the 20 SSR loci ranged from 0.31 (Ce0078) to 0.85 (HK34), with an average of 0.65 (Table 3). There were 15 primers with high polymorphism with PIC value ≥ 0.5 (accounting for 75%), of which the highest PIC value at locus HK34 reached 0.85, followed by locus uq73–164, HK7, and HK31 with PIC values reaching 0.81, 0.80, and 0.79, respectively.

Genetic relationships among taro accessions based on SSR markers

Results of the genetic diversity analysis of 253 taro accessions were counted and processed using NTSYS-UPGMA software. After the genetic similarity matrix of the taro accessions was established (Table S3), the genetic grouping tree of the accessions was drawn using NTSYS Tree-Display shown in Figure 3. The UPGMA clustering dendrogram showed that the genetic similarity coefficient of the taro accessions ranged from 0.39 to 1.00. At a genetic similarity level of 0.60, the taro accessions were classified into two groups. Group I included 56 accessions and had a genetic similarity coefficient ranging from 0.65 to 0.96. The highest value of genetic similarity coefficient (0.96) was observed between two accessions 19–1005 and 19–1007 (both were collected in Bac Giang province). In the pair of accessions 28039 and 28041, the genetic similarity coefficient was 0.90. A similarity coefficient of 0.89 was observed in two pairs of accessions 28251 – 28189 and 11549–11566 (all were collected in Ha Noi City). Group II included the

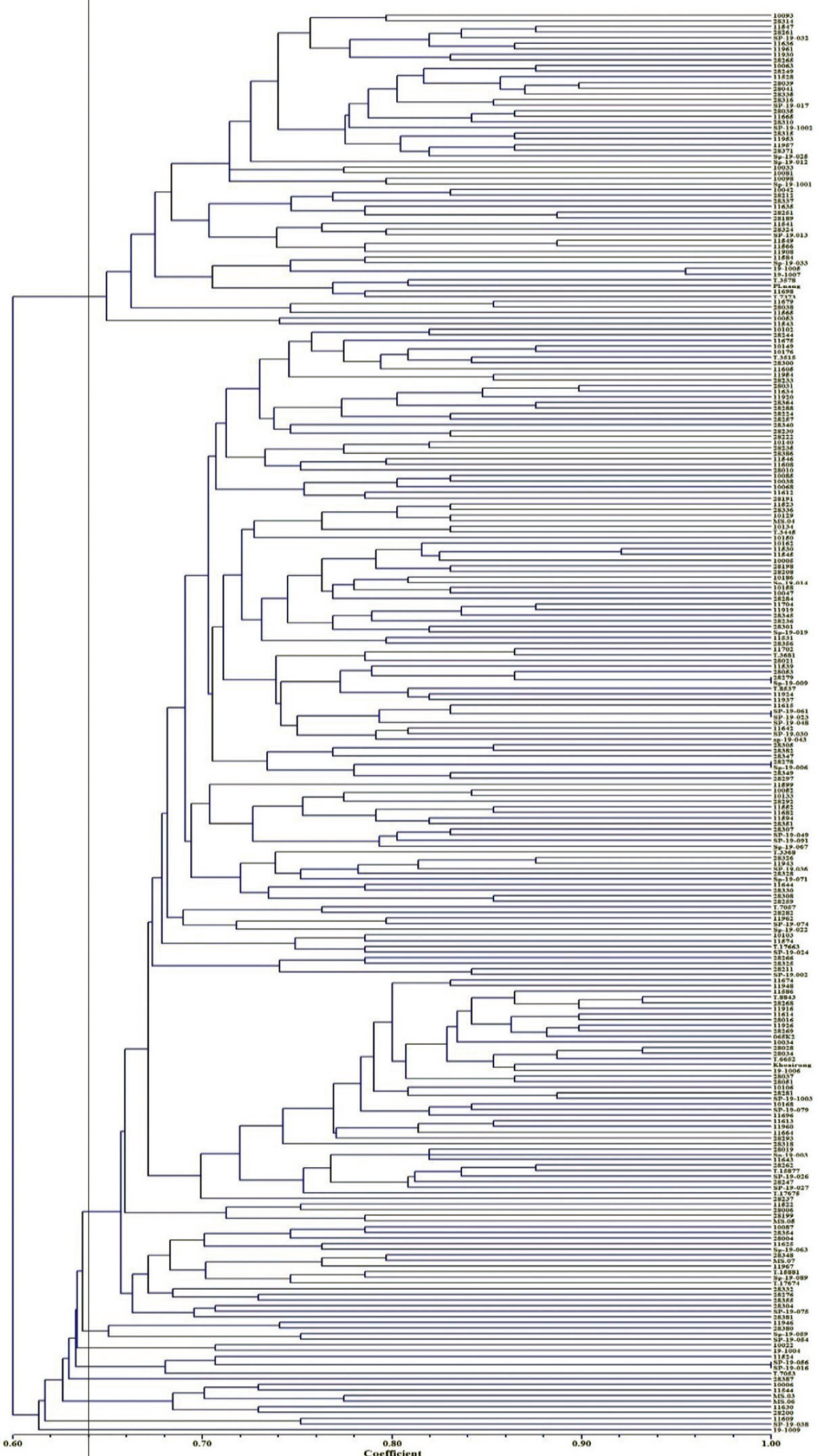


Figure 3. UPGMA clustering dendrogram and genetic relationships in taro (*Colocasia esculenta* (L.) Schott) accessions based on SSR polymorphism data. Taro accessions are denoted by the numeric part of the registration number (Table S1). UPGMA cluster analysis showing the diversity and relationship among 253 taro accessions based on 89 alleles generated by 20 SSR markers.

remaining 197 accessions. At a genetic similarity level of 0.64, group II was classified into eight sub-groups, of which accession 28387 and accession 19–1009 split into separate sub-group V and sub-group VIII at the genetic similarity coefficient of 0.63 and 0.615, respectively. The remaining sub-groups had genetic similarity coefficient ranging from 0.63 to 1.00. Sub-group I was the largest group with 177 accessions. In particular, three pairs of accessions that included 28279 and SP-19-009, SP-19-061 and SP-19-023, and 28278 and SP-19-006, exhibited a genetic similarity coefficient of 1.00. All three pairs of accessions were identical DNA band positions (alleles) at all 20 SSR loci surveyed, indicating that these accessions are very closely related genetically or that there had been genetic overlap. Sub-group II included four accessions (11946, 28380, SP-19-059, and SP-19-054). Sub-group III included two accessions (10022 and 19–1004). Sub-group IV included four accessions (11524, SP-19-056, SP-19-016, and T.7053). In particular, two other accessions, SP-19-056 and SP-19-016, had a similarity coefficient of 1.0, indicating the duplicate samples observed during the sample collection process. Sub-group VI included six accessions (10006, 11544, MS.03, MS.06, 11630, and 28200) and sub-group VII included two accessions (11609 and SP-19-038).

Specific SSR alleles identify taro germplasm collection

An allele that is observed in only one or a few of the total samples is considered rare (defined as equal to or less than a frequency of 0.05 in the total samples) (Pervaiz et al. 2009). These rare alleles have become specific alleles to the samples and the identification of specific alleles can accurately identify genetic resources based on the appearance of this DNA band. Out of a total of 20 SSR markers used in this study, there were six markers with specific alleles appearing in five taro accessions (Table 4 and Figure 4).

The 28211 accession was identified by the uq55–112 marker (with a DNA band of 120 bp), whereas the SP-19-017 accession was identified by the uq73–164 (240 bp DNA band), the T.3578 accession was identified by the HK35 (180 bp DNA band), the T.3515 accession was identified by the Ce0078 (190 bp DNA band), and the 10098 accession was identified by two markers (HK22 with a 210-bp DNA band and HK31 with a 160-bp DNA band).

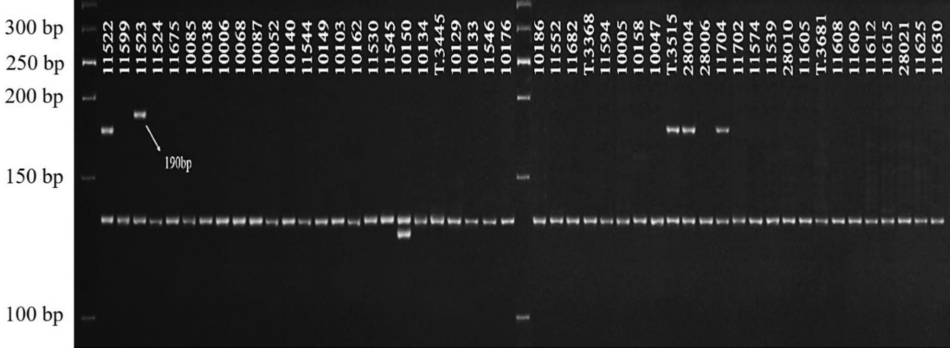
Discussion

SSR polymorphism in genetic diversity analysis of taro germplasm

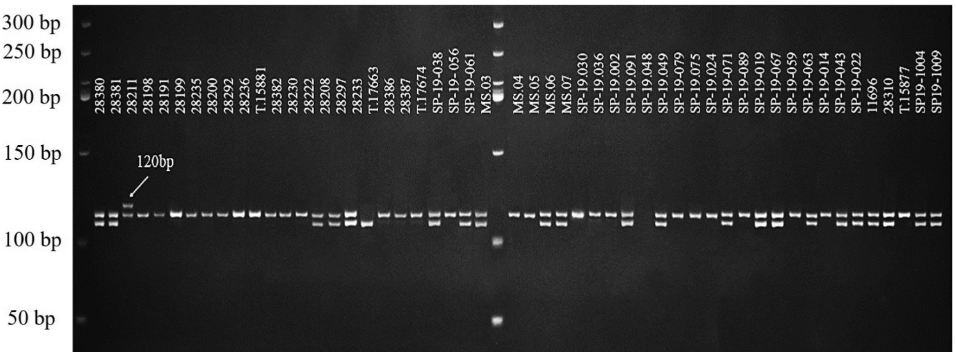
The use of SSR markers to evaluate the genetic diversity of taro varieties has been reported for several decades. In 2002, a set of 16 polymorphic SSR markers specific to taro were developed by Mace and Godwin (2002), which

Table 4. List of SSR loci with specific alleles and taro (*Colocasia esculenta* (L.) Schott) accessions carrying specific alleles.

S/N	SSR locus	Alleles number/locus	Alleles size range (bp)	Number of specific allele	Size of specific allele (bp)	Accession No. with specific allele
1.	uq55–112	3	110–120	1	120	28211
2.	uq73–164	7	143–240	1	240	SP-19-017
3.	HK22	6	175–210	1	210	10098
4.	HK31	6	116–160	1	160	10098
5.	HK35	5	180–280	1	180	T.3578
6.	Ce0078	4	130–190	1	190	T.3515



(a) SSR primer Ce0078



(b) SSR primer uq55-112

Figure 4. The characteristic allele band with size 190 bp of marker Ce0078 observed in taro accession 11523 (a) and the characteristic allele band with size 120 bp of marker uq55–112 observed in taro accession 28211 (b).

were used to analyze taro genetic diversity collected from many different regions of the world (Alam et al. 2019; Mace et al. 2006; Nunes et al. 2012; D. Singh et al. 2008; Valerie and Edome 2016). Several years later, a set of 11

new polymorphic microsatellite markers were isolated from taro (Hu et al. 2009) that were used for the evaluation of genetic diversity of 26 accessions of taro and 101 accessions of yautia (*Xanthosoma mafaffa*) in Togo, West Africa (Bammite et al. 2018). You et al. (2015) and Wang et al. (2017) developed microsatellite markers for taro using transcriptome data and used these markers to evaluate the genetic diversity of taro accessions in China. Also, Mezhii et al. (2017) used 28 microsatellite markers to analyze 50 accessions of taro collected from 11 districts of Nagaland, India.

In the present study, we used 20 SSR markers to evaluate the genetic diversity of 253 accessions in the taro germplasm collection preserved at the Plant Resources Center, Vietnam Academy of Agricultural Sciences. Of these 4, 7, 4, 2, and 3 primer pairs presented in Table 2 were developed by Mace and Godwin (2002), Hu et al. (2009), You et al. (2015), Mezhii et al. (2017), and Wang et al. (2017), respectively. The results showed a high level of genetic variation in the taro accessions used with 100% polymorphic SSR loci. Specifically, 89 alleles were detected from 20 SSR loci, with an average of 4.45 alleles per locus, of which 14 markers have high polymorphism with over 3 alleles per locus. The PIC values obtained at 20 SSR loci ranged from 0.31 (Ce0078) to 0.85 (HK34), with an average of 0.65. There are 15 primers with high polymorphism with PIC values ≥ 0.5 (accounting for 75%), of which the highest PIC values at locus HK34 reached 0.85 (Table 3).

It can be seen from the above findings that there were similarities in the studies when using the same SSR markers to evaluate the genetic diversity of taro, showing the stability of the SSR markers used. The significantly different results among publications demonstrate the high genetic diversity of taro varieties grown in different eco-regions worldwide. These results proved that SSR markers are increasingly becoming the most common polymorphic genetic markers and are being widely used in plant genetic identification.

SSR primers with PIC values greater than or equal to 0.50 indicate high discrimination ability in terms of polymorphism rates (DeWoody, Honeycutt, and Skow 1995). Thus, the diversity of alleles in taro germplasm in our study is quite high and significantly higher than some previous results when using SSR markers to evaluate the genetic diversity of taro (Ab Razak et al. 2021; Bammite et al. 2018; Harun et al. 2022; D. Singh et al. 2008; Valerie and Edome 2016; L. Wang et al. 2017). However, our diversity value is lower than those reported by Nunes et al. (2012) and Alam et al. (2019).

The results of assessing the genetic diversity of taro germplasm collection showed that the lowest genetic similarity coefficient was 0.39, the highest was 1.00, and taro germplasm could be divided into 2 groups (9 sub-groups). This proved that the taro germplasm collection researched was very diverse (genetic differences between taro samples are quite large from 0 to 61%) and showed that the SSR markers were more effective than the RAPD markers in determining genetic relationships between groups as a basis for

classification, identifying groups with hybrid advantage, and identifying genetic resources for conservation and breeding (V. X. Nguyen et al. 2023).

Identification of plant genetic resources through rare SSR alleles

Rare alleles (or specific alleles) are alleles that have a low frequency in a population. These alleles are not commonly found and are present in only a small proportion of individuals. The rarity of these alleles can be determined by analyzing the distribution of alleles in large samples and calculating the average number of rare alleles per locus (Raychaudhuri 2011).

Specific SSR alleles in plants can be identified and characterized using various molecular markers such as SSR markers. These markers are effective tools for certifying plant genetic resources and identifying genes that determine valuable traits. For example, unique alleles were found in the SSR marker loci of sunflower (*Helianthus annuus* L.) lines, which differed in length, number of repeat units, and presence of nucleotide substitutions (Karabitsina et al. 2022). Similarly, Snow et al. (2010) developed species-specific SSR markers to identify early- and later-generation hybrids in cattail (*Typha latifolia* L.) populations, allowing for a better understanding of the prevalence and effects of hybridization.

In a study determining the genetic diversity of Asian rice (*Oryza sativa* L.) cultivars using SSR markers, Pervaiz et al. (2009) reported that a total of 32 (22%) rare alleles were observed in 12 of the 32 SSR loci, with an average of 2.7 rare alleles per locus and considered an allele to be rare when it was observed in only one or two of the thirty-five cultivars (defined as a frequency of 0.05 in the total samples). A core collection with 2,199 accessions of taro germplasm from 10 countries in Oceania: Papua New Guinea, Solomon Islands, Vanuatu, New Caledonia, Fiji, Palau, Niue, Tonga, Cook Islands, and Samoa has been developed based on phenotypic and molecular characterization by Mace et al. (2006). In this report, rare alleles were identified in the taro from the Solomon Islands province of Choiseul which were not observed in any of the other collections (Mace et al. 2006).

In our study, six rare alleles were detected from six SSR markers out of 20 SSR markers used. Specifically, four DNA bands with sizes of 120 bp, 240 bp, 180 bp, and 190 bp were detected from four SSR markers uq55–112, uq73–164, HK35, and Ce0078 that are specific to four taro accessions 28211, SP-19-017, T.3578, and T.3515, respectively. In particular, two DNA bands with sizes of 210 bp and 160 bp were detected from two SSR markers, HK22 and HK31, respectively, which were specific to the taro accession 10098. Five taro accessions carrying this specific allele were collected in the mountainous provinces of Vietnam with different sampling locations and different names of local taro varieties, suggesting a hypothesis that long-term selection of taro varieties in localities has created variations in different taro genetic resources.

In addition, our study, on assessing genetic diversity in the Northern taro collection using RAPD markers showed the appearance of two rare alleles with sizes of 1,400 bp and 2,480 bp from two RAPD markers OPN-11 and OPN-19, respectively. From which, the two taro genetic resources SP-19-002 and SP-19-063, respectively, can be identified through these two specific alleles (V. X. Nguyen et al. 2023). The above results have opened up promising prospects in applying rare alleles obtained from the amplification of specific DNA regions by molecular markers, especially SSR markers that can identify plant genetic resources, including taro.

Conclusion

In the present study, a DNA polymorphism database was built using 20 polymorphic SSR markers for 253 accessions of taro germplasm collection in Vietnam. Genetic diversity analysis recorded very high SSR polymorphism. A total of 89 alleles were obtained with several alleles at each SSR locus ranging from 2 to 8, with an average of 4.45 alleles per locus. The PIC values at 20 SSR loci ranged from 0.31 to 0.85, with an average of 0.65. Nine subgroups of taro genetic resources were divided at a genetic similarity level of 0.64. Six unique molecular markers were discovered to accurately identify 5 taro germplasm in the Vietnam taro germplasm collection. This report once again proved that SSR markers are suitable for collecting DNA polymorphism information, distinguishing between genotypes, studying phylogenetic relationships, and identifying sources of variation for breeding and conservation. Therefore, this is the basis for the orientation of building a taro core collection in Vietnam in the future.

Acknowledgements

We would like to thank the Ministry of Science and Technology, Vietnam for sponsoring funds to implement the project "*Research, evaluation and development of taro genetic resources in Northern Vietnam*" code NVQG-2019/ĐT.05 for the period 2019 - 2023.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Program for Conservation and Sustainable Use of Genetic Resources to 2025, orientation to 2030 of the Ministry of Science and Technology, Vietnam under "Research, evaluation and development of taro genetic resources in North of Vietnam" project, code number: [NVQG-2019/ĐT.05].

Data availability statement

All data that support the findings reported in this study will be available from the corresponding author upon reasonable request.

Author contributions

BHN and XVN designed the experiments, analyzed the data, and wrote the manuscript. TTL and TBHV performed the experiments. All the authors read and approved the final manuscript.

References

Ab Razak, S., M. N. Ghazalli, N. H. E. Azman, A. M. Abd Majid, and S. N. Ismail. 2021. "RAD Sequencing for the Development of Microsatellite Markers for Identification of Malaysian Taro Cultivars." *Biotechnology & Biotechnological Equipment* 35 (1): 1284–1290. https://doi.org/10.1080/13102818.2021.1969278.

Alam, K. S., M. Barooah, D. B. Phookan, P. Kalita, M. C. Talukdar, and S. Luikham. 2019. "SSR Marker-Based Molecular Characterization of Some Upland Taro (*Colocasia Esculenta* L. Schott) Cultivars of North-East India." *International Journal of Current Microbiology and Applied Sciences* 8 (6): 2310–2320. https://doi.org/10.20546/ijcmas.2019.806.274.

Bammite, D., P. J. Matthews, D. Y. Dagnon, A. Agbogan, K. Odah, A. Dansi, and K. Tozo. 2018. "Agro Morphological Characterization of Taro (*Colocasia esculenta*) and Yautia (*Xanthosoma mafaffa*) in Togo, West Africa." *African Journal of Agricultural Research* 13 (18): 934–945. https://doi.org/10.5897/AJAR2018.13043.

Dang, T. T. M., and X. V. Nguyen. 2014. "Assessment of Genetic Diversity of Taro (*Colocasia Esculenta* (L.) Schott) Using SSR Markers." *Vietnam Journal of Agriculture and Rural Development* 16:10–16. https://sti.vista.gov.vn/file_DuLieu/dataTLKHCN//CVv201/2014/CVv201S162014010.pdf. (in Vietnamese).

Das, A. B., A. Das, C. Pradhan, and S. K. Naskar. 2015. "Genotypic Variations of ten Indian Cultivars of *Colocasia Esculenta* Var. *Antiquorom* Schott. Evident by Chromosomal and RAPD Markers." *Caryologia* 68 (1): 44–54. https://doi.org/10.1080/00087114.2015.1013335.

Devi, A. A. 2012. "Genetic Diversity Analysis in Taro Using Molecular Markers. An Overview." *Journal of Root Crops* 38 (1): 15–25.

DeWoody, J. A., R. L. Honeycutt, and L. C. Skow. 1995. "Microsatellite Markers in White-Tailed Deer." *The Journal of Heredity* 86 (4): 317–319. https://doi.org/10.1093/oxfordjournals.jhered.a111593.

Doan, D. L., and H. P. Nguyen. 2019. "Assessing Some Agro-Biology Characteristics of Taro Varieties in the North Western Rigion." *Journal of Natural Science and Technology* 14:82–89. https://113165166110/81/bitstream/DL/134679/9093/1/CVv446S142019082.pdf. (in Vietnamese)

FAOSTAT. 2023. "FAO Statistical Database." Accessed June 20, 2023. https://www.fao.org/faostat/en/#data/QCL.

Fufa, T. W., W. G. Abtew, C. O. Amadi, and H. O. Oselebe. 2022. "DArTseq SNP-Based Genetic Diversity and Population Structure Studies Among Taro (*Colocasia Esculenta* (L.) Schott) Accessions Sourced from Nigeria and Vanuatu." *PLOS ONE* 17 (11): e0269302. https://doi.org/10.1371/journal.pone.0269302.

Harun, A. R., F. Ahmad, S. Salleh, and M. N. Bin Norngaini. 2022. "Identification of Genetic Diversity Among Mutant Taro (*Colocasia Esculenta* L. Cv Wangi) Using

Agro-Morphological Trait and Simple Sequence Repeats (SSR) Molecular Markers." *Journal of Experimental Biology and Agricultural Sciences* 10 (2): 359–368. https://doi.org/10.18006/2022.10(2).359.368.

Hu, K., X. F. Huang, W. D. Ke, and Y. Ding. 2009. "Characterization of 11 New Microsatellite Loci in Taro (*Colocasia Esculenta*." *Molecular ecology resources* 9 (2): 582–584. https://doi.org/10.1111/j.1755-0998.2008.02441.x.

Jaccard, P. 1908. "Nouvelles researches sur la distribution floral." *Bulletin de la Société Vaudoise Sciences Naturelles* 44:223–270. https://doi.org/10.5169/seals-268384.

Karabitsina, Y. I., N. V. Alpatieva, E. B. Kusnetsova, V. A. Gavrilova, N. V. Titov, E. E. Radchenko, and I. N. Anisimova. 2022. "Polymorphism of Microsatellite Markers Linked with Rf1 and Pl5/Pl8 Loci in Sunflower *Helianthus annuus* L." *Plant Biotechnology and Breeding* 4 (4): 5–14. https://doi.org/10.30901/2658-6266-2021-4-o1.

Kreike, C. M., H. J. Van Eck, and V. Lebot. 2004. "Genetic Diversity of Taro, *Colocasia Esculenta* (L.) Schott, in Southeast Asia and the Pacific." *Theoretical and Applied Genetics* 109:761–768. https://doi.org/10.1007/s00122-004-1691-z.

Li, M., R. Li, M. Zhou, P. Eyzaguirre, and W. G. Ayad. 2010. "Genetic Diversity Assessment of Taro Collections in China Using RAPD Assays." In *The Global Diversity of Taro: Enthnobotany and Conservation*, edited by R. V. Ramanatha, P. J. Matthews, P. B. Eyzaguirre, and D. Hunter, 92–97. Rome, Italy: Bioversity International.

Lu, D., N. Niu, and Y. Y. Li. 2010. "Application of SSR Marker in Plant Genome Research." *Shenyang Shifan Daxue Xuebao* 28 (1): 83–85.

Mace, E. S., and I. D. Godwin. 2002. "Development and Characterization of Polymorphic Microsatellite Markers in Taro (*Colocasia Esculenta*)." *Genome* 45 (5): 823–832. https://doi.org/10.1139/G02-045.

Mace, E. S., P. N. Mathur, L. Izquierdo, D. Hunter, M. B. Taylor, D. R. Singh, I. H. DeLacy, G. V. H. Jackson, and I. D. Godwin. 2006. "Rationalization of Taro Germplasm Collections in the Pacific Island Region Using Simple Sequence Repeat (SSR) Markers." *Plant Genetic Resources* 4 (3): 210–220. https://doi.org/10.1079/PGR2006125.

Mezhii, T. L., S. Changkija, A. Pattanayak, H. P. Chaturvedi, S. V. Devi, and P. R. Kole. 2017. "Genetic Characterization of Locally Cultivated Taro Germplasm from Eleven District of Nagaland, India." *International Journal of Current Microbiology and Applied Sciences* 6 (8): 3338–3348. https://doi.org/10.20546/ijcmas.2017.607.398.

Mohammadi, S. A., and B. M. Prasanna. 2003. "Analysis of Genetic Diversity in Crop Plants-Salient Statistical Tools and Considerations." *Crop Science* 43 (4): 1235–1248. https://doi.org/10.2135/cropsci2003.1235.

Nair, U. M. A., A. A. Devi, S. S. Veena, B. S. P. Krishnan, and R. S. Arya. 2018. "Genetic Diversity Analysis of Leaf Blight Resistant and Susceptible Taro [*Colocasia Esculenta* (L.) Schott] Genotypes Using ISSR Markers." *Journal of Root Crops* 44 (2): 44–52.

Nguyen, V. X., T. T. M. Dang, and T. A. Ngo. 2015. "Genetic Diversity and Phylogenetic Relationships Among *Colocasia* Species (Araceae) in Vietnam Using SSR Markers." *Aroideana* 38E (1): 177–185.

Nguyen, T. N. H., V. V. Nguyen, L. C. Vu, and M. S. Prana. 2010. "Taro Germplasm Collection in Vietnam." In *The Global Diversity of Taro: Enthnobotany and Conservation*, edited by R. V. Ramanatha, P. J. Matthews, P. B. Eyzaguirre, and D. Hunter, 60–68. Rome, Italy: Bioversity International.

Nguyen, V. X., T. V. A. Pham, T. Q. H. Nguyen, T. T. M. Le, T. B. H. Vu, and T. T. Le. 2021. "Study on Chromosome Number and Karyotype in the North Taro Germplasm Preserved at the Plant Resources Center - Vietnamese Academy of Agricultural Sciences." *HNUE Journal of Science (Natural Sciences)* 66 (4F): 144–151. https://doi.org/10.18173/2354-1059.2021-0077.

Nguyen, V. X., T. B. H. Vu, T. T. Le, and T. V. A. Pham. 2023. "Assessment of Genetic Diversity in Collection of North Taro Accessions Using RAPD Markers." *Vietnam Journal of Agriculture and Rural Development* 4:13–21. http://tsdt.tapchinongnghiep.vn/Epaper/files/library/site-3/20230417/2.pdf. (in Vietnamese).

Nguyen, V. X., H. Yoshino, and M. Tahara. 1998. "Genetic Control of Four Enzymes in Diploid Taro, *Colocasia Esculenta* (L.) Schott." *Japanese Journal of Breeding* 48 (3): 273–280.

Nunes, R. S. C., F. R. Pinhati, L. P. Golinelli, T. N. H. RebouçRebouçAs, V. M. F. Paschoalin, and J. T. da Silva. 2012. "Polymorphic Microsatellites of Analysis in Cultivars of Taro." *Horticultura Brasileira* 30 (1): 106–111. https://doi.org/10.1590/S0102-05362012000100018.

Ochiai, T., V. X. Nguyen, M. Tahara, and H. Yoshino. 2001. "Geographical Differentiation of Asian Taro, *Colacasia Esculenta* (L.) Schott, Detected by RAPD and Isozyme Analyses." *Euphytica* 122:219–234. https://doi.org/10.1023/A:1012967922502.

Okpul, T., E. S. Mace, I. D. Godwin, D. Singh, and M. E. Wagih. 2005. "Evaluation of Variability Among Breeding Lines and Cultivars of Taro (*Colocasia esculenta*) in Papua New Guinea Using ISSR Fingerprinting and Agro-Morphological Characterization." *Plant Genetic Resources Newsletter* 143:8–16. https://www.academia.edu/37478143/Evaluation_of_variability_among_breeding_lines_and_cultivars_of_taro_Colocasia_esculenta_in_Papua_New_Guinea_using_ISSR_fingerprinting_and_agro_morphological_characterization.

Pervaiz, Z. H., M. A. Rabbani, S. R. Pearce, and S. A. Malik. 2009. "Determination of Genetic Variability of Asian Rice (*Oryza Sativa* L.) Varieties Using Microsatellite Markers." *African Journal of Biotechnology* 8 (21): 5641–5651. https://doi.org/10.5897/AJB09.827.

Raychaudhuri, S. 2011. "Mapping Rare and Common Causal Alleles for Complex Human Diseases." *Cell* 147 (1): 57–69. https://doi.org/10.1016/j.cell.2011.09.011.

Rohlf, F. J. 2000. *NTSYS-Pc: Numerical Taxonomy and Multivariate Analysis System, Version 2.1*. New York: Exeter Software.

Sharma, K., A. K. Mishra, and R. S. Misra. 2008a. "Analysis of AFLP Variation of Taro Population and Markers Associated with Leaf Blight Resistance Gene." *Academic Journal of Plant Sciences* 1 (3): 42–48.

Sharma, K., A. K. Mishra, and R. S. Misra. 2008b. "A Simple and Efficient Method for Extraction of Genomic DNA from Tropical Tuber Crops." *African Journal of Biotechnology* 7 (8): 1018–1022.

Singh, D., E. S. Mace, I. D. Godwin, P. N. Mathur, T. Okpul, M. Taylor, D. Hunter, R. Kambuou, V. R. Rao, and G. Jackson. 2008. "Assessment and Rationalization of Genetic Diversity of Papua New Guinea Taro (*Colocasia esculenta*) Using SSR DNA Fingerprinting." *Genetic Resources and Crop Evolution* 55:811–822. https://doi.org/10.1007/s10722-007-9286-6.

Singh, S., D. R. Singh, F. Faseela, N. Kumar, V. Damodaran, and R. C. Srivastava. 2012. "Diversity of 21 Taro (*Colocasia Esculenta* (L.) Schott) Accessions of Andaman Islands." *Genetic Resources and Crop Evolution* 59:821–829. https://doi.org/10.1007/s10722-011-9721-6.

Snow, A. A., S. E. Travis, R. Wildová, T. Fér, P. M. Sweeney, J. E. Marburger, S. Windels, B. Kubátová, D. E. Goldberg, and E. Mutegi. 2010. "Species-Specific SSR Alleles for Studies of Hybrid Cattails (*Typha latifolia* × *T. angustifolia*; Typhaceae) in North America." *American Journal of Botany* 97 (12): 2061–2067. https://doi.org/10.3732/ajb.1000187.

Temesgen, M., and N. Retta. 2015. "Nutritional Potential, Health and Food Security Benefits of Taro *Colocasia Esculenta* (L.): A Review." *Food Science and Quality Management* 36:23–30. https://www.academia.edu/35223835/Nutritional_Potential_Health_and_Food_Security_Benefits_of_Taro_Colocasia_Esculenta_L_A_Review.

Thokchom, M., L. S. Devi, R. K. Thirumdasu, A. K. B. Devi, and K. H. James. 2017. "Growth, Physiological Studies and Yield of Taro (*Colocasia Esculenta* Schott) Cv. Mukhi Pan As Influenced by Intercropping and Row Pattern Under Manipur Condition." *International Journal of Current Microbiology and Applied Sciences* 7 (5): 925–931. https://doi.org/10.20546/ijcmas.2018.705.113.

Valerie, A. P. P., and P. A. Edome. 2016. "Genetic Diversity Analysis of Kenyan Taro [*Colocasia Esculenta* (L.) Schott] Accessions Using SSR Markers." *Sky Journal of Agricultural Research* 5 (5): 076–086.

Victoria, F. C., L. C. da Maia, and A. C. de Oliveira. 2011. "*In Silico* Comparative Analysis of SSR Markers in Plants." *BMC Plant Biology* 11:1–15. https://doi.org/10.1186/1471-2229-11-15.

Wang, Z., Y. Sun, X. Huang, F. Li, Y. Liu, H. Zhu, Z. Liu, and W. Ke. 2020. "Genetic Diversity and Population Structure of Eddoe Taro in China Using Genome-Wide SNP Markers." *PeerJ* 8:e10485. https://doi.org/10.7717/peerj.10485.

Wang, L., J. Yin, P. Zhang, X. Han, W. Guo, and C. Li. 2017. "*De Novo* Assembly and Characterization of Transcriptome and Microsatellite Marker Development for Taro (*Colocasia Esculenta* (L.) Schott." *International Journal of Genetics and Molecular Biology* 9 (5): 26–36. https://doi.org/10.5897/IJGMB2017.0155.

You, Y., D. Liu, H. Liu, X. Zheng, Y. Diao, X. Huang, and Z. Hu. 2015. "Development and Characterisation of EST-SSR Markers by Transcriptome Sequencing in Taro (*Colocasia Esculenta* (L.) Schott." *Molecular Breeding* 35:1–11. https://doi.org/10.1007/s11032-015-0307-4.