BIODIVERSITAS
Volume 24, Number 7, July 2023
Pages: 4060-4068

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d240744

Molecular evaluation of genetic diversity and relationships of *Musa* cultivars in Thailand using Start Codon Targeted (SCoT) markers

SARUDA NITIWORAKARN, WUTTICHAI PHAE-NGAM, ONGKARN VANIJAJIVA♥
Faculty of Science and Technology, Phranakhon Rajabhat University. Bangkok, 10220, Thailand. Tel./fax.: +66-271-663375,
♥email: vanijajiva@gmail.com

Manuscript received: 24 April 2023. Revision accepted: 23 July 2023.

Abstract. *Nitiworakarn S, Phae-Ngam W, Vanijajiva O. 2023. Molecular evaluation of genetic diversity and relationships of* Musa *cultivars in Thailand using Start Codon Targeted (SCoT) markers. Biodiversitas 24: 4060-4068. Musa* cultivars are mostly obtained from crossing between wild diploid subspecies of *M. acuminata* (A genome) and *M. balbisiana* (B genome), and they reveal numerous levels of ploidy and genomic constitution. The present study, for the first time, investigated the genetic diversity of 90 *Musa* cultivars from 17 localities of Chai Badan district in Lopburi Province, Thailand, using start codon targeted (SCoT) markers. The SCoT technique has shown very high repeatability for characterizing *Musa* cultivars. The result indicated that DNA fingerprints from 30 SCoT primers generated 244 amplicons, among which 238 bands (97.54%) were polymorphic, with a mean of 7.93 polymorphic bands per primer. The average polymorphism information content (PIC) was high, ranging from 0.345 to 0.483, with an average of 0.432. Genetic diversity and its partitioning parameters were calculated and demonstrated that present *Musa* cultivars maintain relatively high genetic diversity. The unweighted pair-group method with an arithmetic cluster analysis, principal coordinate analysis and STRUCTURE analysis results were the same and showed the clear division of the genotypes into two distinct clusters, which were cluster-I consisted of only the AA genomes and cluster-II comprised of ABB and BB genomes. In conclusion, SCoT markers obtained here showed their manipulation in genetic diversity and the relationship of cultivated *Musa*, which would be practical for genetic preservation and continuing breeding programs in Thailand and other regions.

Keywords: Banana, DNA fingerprints, molecular characterization, STRUCTURE analysis

INTRODUCTION

Musa L. is a perennial, monocotyledonous plant genus belonging to the family Musaceae. The genus embraces all the cultivated bananas, which are important worldwide crops (Hinge et al. 2022). The area is the main center of diversity, and most of the *Musa* cultivars are believed to have originated in Southeast Asia (Wahyudi et al. 2022). Currently, they are grown in more than 135 countries providing income to the farming cooperation through local and international commerce (Arora et al. 2018; Dehery et al. 2022), mostly for their fruit, with could be separated into "bananas" and "plantains" (cooking banana) (Marimuthu et al. 2019). Commonly, the cultivated bananas and plantains are grouped according to ploidy and the relative evaluation of wild species *Musa acuminata* (A genome) and *Musa balbisiana* (B genome) in their genome (Christelová et al. 2017). They have contributed ten genomic combinations in different ploidy levels (AA, BB, AB, AAA, AAB, ABB, AAAA, AAAB, AABB, and ABBB) (Davey et al. 2013; Premjet et al. 2022). The present-day classification of *Musa* cultivars has usually been based upon several morphological, phenological, and floral criteria and ploidy level combinations. More than 80 different taxonomically accepted accessions are contained in *the Musa* genus (Carracedo et al. 2021). In Thailand, *Musa* cultivars have been surveyed, comprising approximately 140 varieties with ten wild species and four introduced species (Rotchanapreeda et al. 2016). Based on their morphological descriptors and their similarity to *M. acuminata* and *M. balbisiana*, the genome alignment of *Musa* cultivars and their hybrids can be determined, but the reliability of classification based on phenotypic observations varies depending on the age and developmental stage of *Musa* cultivars, and sensibility of measured characters to the effect of environmental differences. The phenotype of *Musa* species and cultivars can vary depending on the environmental conditions, and the phenotypic variation influences the accuracy of identifying the characteristics or traits. In addition, *Musa* cultivars are widely geographically distributed, accomplished solely by the transportation of vegetative planting materials by anthropological organizations, differentiated by somatic mutation, and multiplied by vegetative reproduction and artificial selection. These showcases made the *Musa* cultivar's genetic background rather complicated.

Understanding the diversity and relationships of genetic resources are required as a foundation for the development and sustainability of cultivated production, and the evaluation of genetic diversity provides effective information to maintain and manipulate genetic resources. Presently, the ability to evaluate genetic diversity and the relationships within and between populations using molecular markers has resulted from the rapid development of molecular genetics, and molecular markers have been extensively applied in the assessment of genetic diversity, particularly in *Musa* cultivars (Rotchanapreeda et al. 2016; Swain et al.

2016; Chang et al. 2017; Silva et al. 2017; Arora et al. 2018; Bawin et al. 2019; Santos et al. 2019; Marimuthu et al. 2019; Wanvisait et al. 2019; Boonsrangsom et al. 2020; Saraswathi et al. 2020; Wahyudi and Rifliyah 2020; Carracedo et al. 2021, González et al. 2021; Mertens et al. 2021). Recently, many new marker techniques have been developed in line with the rapid growth of *Musa* genomic research (Wahyudi et al. 2022; Dehery et al. 2022; Premjet et al. 2022). Among novel DNA-based molecular fingerprints, start codon targeted (SCoT) polymorphism has been presented as a gene-targeted molecular marker classification primarily diverse from random DNA markers, and it normally relies on the short conserved nucleotide region adjacent to the ATG initiation codon of plant genes (Collard and Mackill 2009; Rittirongsakul et al. 2020). Although SCoT markers have been found useful in various plants for genetic diversity and relationship studies because of their high reproducibility and efficiency (Vanijajiva 2020; Rai 2023), only Igwe et al. (2022) have successfully employed SCoT to assess genetic variation in *Musa* cultivars. With this background, the study aimed to explore the genetic diversity of *Musa* cultivars important for Thailand. We applied for the first time SCoT markers for genetically characterizing Thai *Musa* germplasm and hypothesized that SCoT markers could be applied as an effective marker system to study genetic diversity in *Musa* cultivars. Genetically different Thai *Musa* genotypes might benefit from developing additional *Musa* cultivars having higher yield and resistance to biotic and abiotic factors to better meet the demands of plant breeders.

MATERIALS AND METHODS

Study sites and sample collection

A total of 90 *Musa* cultivars were collected from 17 localities; Bua Chum (BC), Ban Mai Samakkhi (BM), Chai Badan (CB), Chai Narai (CN), Huai Hin (HH), Khao Laem (KL), Ko Rang (KR), Nikhom Lam Nara (LM), Lam Narai (LN), Muang Khom (MK), Makok Wan (MW), Na Som (NS), Nong Yai To (NY), Sap Takhian (SK), Sila Thip (ST), Tha Din Dam (TD), Tha Manao (TM) of Chai Badan district in Lopburi Province, Thailand (Figure 1). All *Musa* samples were identified as species based on morphological characteristics, according to Simmonds (1957). Entirely they consisted of three different *Musa* species of *M. acuminata* (AA), *Musa x paradisiaca* (ABB) and *M. balbisiana* (BB) (Table 1). All Musa accessions' living collection and voucher specimens were deposited at the Chaibadanpiphat College, Phranakhon Rajabhat University, for future reference.

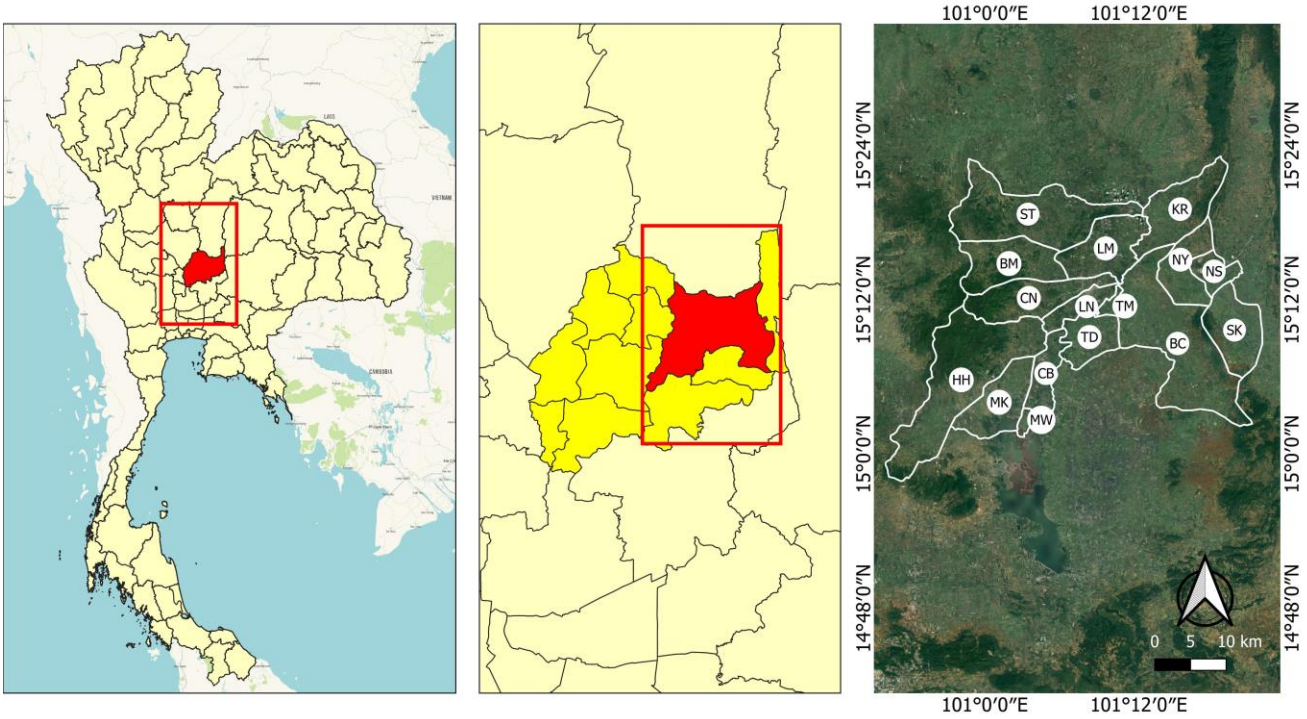

Figure 1. Seventeen collection localities of *Musa* cultivars across the Chai Badan district of Lopburi Province, Thailand

Table 1. List of *Musa* cultivars, vernacular and their collected location used in this study

Species (genome)	Vernacular name (code)	Total samples	Accession codes*
Musa acuminata (AA)	Khai (K)	17	K_LM, K_CN, K_ST, K_HH, K_MK, K_BC, K_TD, K_MW, K_SK, K_NS, K_NY, K_KR, K_TM, K_LN, K_CB, K_BM, K_KL
	Hom (H)	13	H_LM, H_CN, H_ST, H_HH, H_MK, H_TD, H_SK, H_NY, H_KR, H_TM, H_LN, H_CB, H_KL
	Leb Mu Nang (L)	13	L_LM, L_ST, L_HH, L_MK, L_BC, L_TD, L_NS, L_NY, L_KR, L_TM, L_LN, L_CB, L_KL
Musa x paradisiaca (ABB)	Nam Wa (N)	17	N_LM, N_CN, N_ST, N_HH, N_MK, N_BC, N_TD, N_MW, N_SK, N_NS, N_NY, N_KR, N_TM, N_LN, N_CB, N_BM, N_KL
	Hak Muk (M)	13	M_LM, M_CN, M_ST, M_HH, M_MK, M_BC, M_TD, M_MW, M_NY, M_TM, M_LN, M_CB, M_KL
Musa balbisiana (BB)	Tani (T)	17	T_LM, T_CN, T_ST, T_HH, T_MK, T_BC, T_TD, T_MW, T_SK, T_NS, T_NY, T_KR, T_TM, T_LN, T_CB, T_BM, T_KL

Note: Accession codes*: vernacular name_locality

Genomic DNA extraction

Genomic DNA was purified from the leaves of 90 *Musa* accessions using the CTAB method with minor modification (Vanijajiva 2020). The leaf material, about 500 mg, was crushed in a mortar with a pestle. Extraction buffer [(1% (w/v) CTAB, 50 mM Tris-HCl (pH 8), 0.7 M NaCl, 0.1% ß-mercaptoethanol)] 500 µL was included and the solution was incubated at 60°C for 30 min. The homogenate was mixed with 25: 24: 1 phenol: chloroform: isoamyl alcohol (v/v/v) by soft inversion. Afterward, the centrifugation was done at 13,000 rpm for 15 min and the upper aqueous layer was removed to a clean tube. RNA was eliminated by treating with 2.5 µL of the RNase (10 µg/µl) for 30 min at 37°C. DNA extraction with phenol/chloroform/isoamyl alcohol was duplicated again. DNA in the mixture was precipitated with 0.6 volumes of ice-cold isopropanol and washed with 70% ethanol. Subsequently, the DNA was extracted using the CTAB DNA extraction process without RNase. The approach was repeated until the DNA pellet was free of color (two to three times), and the finishing pellet was dissolved in sterile deionized water. DNA quality was determined using Nanodrop Spectrophotometer (Thermo Scientific Nanodrop 1000, USA) at the absorbance ratio of 260 and 280 nm, providing a value of 1.7-1.8, which defines pure DNA preparation. Characteristics of DNA fragments were also electrophoretically examined through 0.8% agarose gel employing 1X TAE buffer. The DNA was kept at -20°C before operating as templates for PCR amplification.

SCoT-PCR amplification

Thirty SCoT primers used in the current study were considered according to Collard and Mackill (2009) and were primarily screened for analysis (Table 2). PCR was performed using a Thermohybaid Px2 (Roche Molecular Systems, Inc., USA). PCR was adjusted for testing the SCoT method. All PCR reactions were performed within a total volume of 20 µL. PCR reaction mixtures contained PCR buffer (Promega; 20 mM Tris-HCl (pH 8.4), 50 mM KCl), two mM $MgCl_2$, 0.24 mM of each deoxyribonucleotide triphosphates, 0.5 U of Taq polymerase (Promega), and 0.8 µm of primer. Each reaction contained 50 ng of template DNA. A typical PCR cycle was used: an initial denaturation step at 94°C for 3 min, followed by 35 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 2 min; the final extension at 72°C was held for 5 min. The SCoT products were all considered by agarose (1.8% w/v) gel electrophoresis at 150 A for 30 minutes in 0.04 M TAE (Tris–acetate 0.001 M-EDTA) buffer pH 8. The gels were dyed with ethidium bromide (10 mg/mL). The gels were viewed and photographed by Bio-Imaging System (Syngene, Genegenuis). To determine SCoT profiles, the size of each DNA band was inferred by comparison with a 100 bp DNA ladder (Promega), used as a molecular weight marker (M). Polymorphisms at all loci were confirmed by three repeating tests for each primer at different times.

Data analysis

The presence of a band of individual cultivars was scored as 1 and 0 for the absent band. The fragment size of the DNA band was estimated using a 1 Kb DNA ladder. The scoring of binary data was then investigated to determine the most suitable primer for amplification, using four parameters, i.e., TAB: total amplified bands; NPB: number of polymorphic bands; PPB: percentage of polymorphic bands; PIC: polymorphism information content. The PIC value for each SCoT marker was calculated as proposed by Roldan-Ruiz et al. (2000), as $PIC_i=2f(1-f_i)$, where PIC_i is the polymorphic data content of marker i, f_i is the occurrence of the marker bands present, and 1-f_i is the frequency of absent marker bands. The PIC was averaged over all the bands for each primer (Wahyudi and Rifliyah 2020). In addition, the basic parameters of genetic diversity were also analyzed using GenAlEx 6.5 software (Peakall and Smouse 2012). The parameters include the number of alleles (Na), the effective number of alleles (Ne), Shannon index (I), Nei's genetic diversity (He), unbiased He (uHe), and percentage of polymorphism (PPB).

Clustering analyses were performed using Palaeontological Statistics (PAST) 3.0 software based on Ward's method of hierarchical clustering (minimum variance method) with Euclidean dissimilarity index, bootstraps 1000 (Hammer et al. 2001). This matrix

generated a dendrogram based on the hierarchical cluster analysis, which is based on the average linkage between groups, i.e., the unweighed pair-group method algorithm (UPGMA). A principal component analysis (PCoA) was organized using a genetic similarity matrix obtained from the basis of Jaccard's coefficients of similarities set.

The genetic structure was established using model-based cluster analysis (STRUCTURE v. 2.3.4). The number of groupings (K) was estimated to be every ten runs for every population, which varied from 1 to 10, characterized by a set of distinctive allele frequencies at each locus, and the individuals were sited in K clusters. Using this method, Markov chain Monte Carlo (MCMC) posterior probabilities were approximated. The MCMC chains were run with a 10,000-iteration burn-in period, followed by 100,000 iterations using a model allowing admixture and correlated allele frequencies (Vanijajiva and Pornpongrungrueng 2020). The most anticipated value for K was predicted with Evanno's ΔK method (Evanno et al. 2005) using STRUCTURE HARVESTER (Earl and van Holdt 2012).

RESULTS AND DISCUSSION

Musa cultivars classification

In Thailand, commercializing fresh fruits and related products of *Musa* cultivars accounts for a significant percentage of domestic income in productive regions, particularly in Lopburi Province. Moreover, the biodiversity and genetic evidence of locally cultivated *Musa* is not well recognized, particularly in the Lopburi region, since the current *Musa* cultivars are restricted to a few selected varieties chosen by the consumers, and the presence of inadequate diversity in them makes them prone to diseases and pests. Consequently, characterization and systematic assessment of *Musa* assets of the locality will assist in providing strategies for management and sustainable utilization of the *Musa* germplasm. Hence, the sampling procedure for this study was carefully carried out in the most frequent germplasm cultivated in all 17 localities of Chai Badan district in Lopburi Province, Thailand (Figure 1). The identifications of *Musa* cultivated across Chai Badan district, an agricultural standing area of Lopburi Province, were based on morphological characteristics and ploidy information (Simmonds 1957). All 90 *Musa* accessions were classified into three species of *M. acuminata*, *M. x paradisiaca* and *M. balbisiana*. The *M. acuminata* species were divided into three subgroups, including 17 samples of the Khai variety (Figure 2A-2C), 13 samples of the Hom variety (2D-2F) and 13 accessions Leb Mu Nang variety (2G-2I), for *M. x paradisiaca* species was separated into two subgroups containing Nam Wa variety (Figure 2J-2L) and Hak Muk variety (2M-2O). Lastly, *M. balbisiana* species was observed only in the Tani variety comprising 17 samples (Figure 2P-2R).

SCoT polymorphism analysis

The documentation, preservation, and collection of genotypes belonging to *Musa* cultivars deserve particular attention to reduce genetic loss and to present the practice of genotypes in further breeding programs (Koutouleas et al. 2023). Discovering the proper molecular markers for genetic fingerprinting is an important step in genetic estimation. For this explanation, different molecular markers are used and evaluated in the genetic fingerprinting of reasonably significant plant species (Dar et al. 2019; Boban et al. 2022; Zhang et al. 2023). This also holds accuracy for *Musa* cultivars. The molecular markers can be attempted for their utility in cultivar diversity and revealing genetic resemblance (Lamare and Rao 2015; Hinge et al. 2022; Premjet et al. 2022).

The present study is the first report showing significant genetic polymorphism amongst various *Musa* accessions using SCoT markers in Thailand. To examine uncertainty, the SCoT method can detect genetic variability at the cultivar level; our collection was analyzed with all thirty selected SCoT primers. In these experiments, clearly different band patterns were detected for all cultivars, hence indicating SCoT method can easily detect genetic variation at the cultivar level. A total of 244 amplicons were obtained, among which 238 bands were polymorphic, with a mean of 7.93 polymorphic bands per primer. The length of the bands ranged from 50 to 1500 bp. The polymorphism percentage was 97.54%. Among the 30 SCoT primers, SCoT12 amplified a maximum of 14 bands, and the polymorphism percentage was 100%. The number of bands amplified by the SCoT15 primer was at least 4, and the polymorphism percentage was 100%. Also, the informativeness of each primer was evaluated using PIC criteria. Guo et al. (2014) stated that the dominant markers had a maximum PIC value of 0.5 and higher PIC values attributed to high polymorphism. The closer the PIC value to 0.5, the more useful the primer is to be benefited in the analysis of genetic diversity (Marimuthu et al. 2019; Wahyudi et al. 2022). This present study showed the highest PIC value was 0.457, obtained from SCoT4, while the lowest was 0.345, obtained from SCoT21, with an average PIC value of 0.432 per primer (Table 2). The results obtained in the present study proved that the SCoT marker is more effective. Thus, this study exposed that SCoT markers are appropriate, efficient markers for showing within cultivar/ grouping genetic variability for *Musa* cultivars.

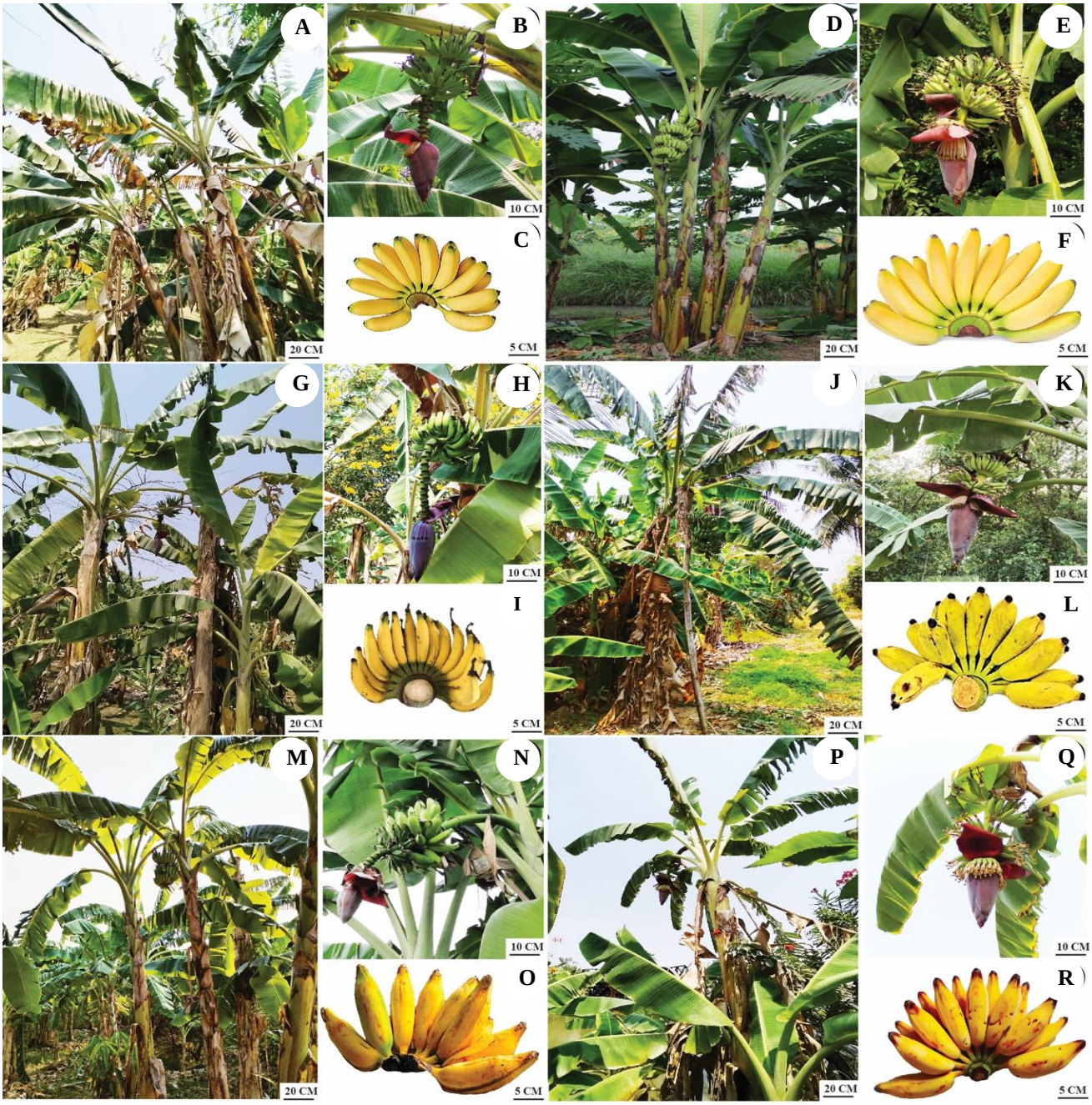

Figure 2. Morphological characteristics of present *Musa* cultivars. A-C: Khai variety; D-F: Hom variety; G-I: Leb Mu Nang variety; J-K: Nam Wa variety; M-O: Hak Muk variety; P-R: Tani variety

Genetic diversity and cluster analysis

Breeding approaches to enhance additional varieties having high yield, high nutritional amount, and resistance to biotic and abiotic stress are essential to compensate for food deficiencies (Mekonnen et al. 2022; Biswas et al. 2023). *Musa* cultivars breeding can be problematic and difficult (Singh et al. 2021; Hinge et al. 2022). The situation was considered impractical essentially because of considerations such as vegetative propagation, the existence of polyploidy level, low fertility of clones, and time required to produce an improved cultivar, preventing fast continue in breeding, particularly now that the problem of climate change is globally threatening harvest produces. Genetic studies as a tool in breeding programs for the description of germplasm have been performed with several molecular marker systems. In addition, evaluating genetic variety and genetic factors can instantly assist the finding of *Musa* accessions for assortment and assimilation into breeding programs for *Musa* cultivar's improvement and management (Rotchanapreeda et al. 2016; Dehery et al. 2022).

The advancement of SCoT molecular markers in agronomically significant cultures supplies knowledge on genetic diversity, which is of great value when these

genetic resources are used in breeding and germplasm conservation programs. To estimate genetic diversity among *Musa* accessions from Thailand. The genetic parameters were calculated for all the 90 *Musa* cultivars amplified with SCoT primers (Table 2). Unbiased expected heterozygosity (uHe) ranged from 0.100±0.011 (BB group) to 0.216±0.014 (AA group). A similar pattern was observed for Shannon's information index (I), with the highest value of 0.314±0.019 observed in the AA group and the lowest value of 0.147±0.016 observed in the BB group. Moreover, the observed number of alleles (Na) ranged from 0.852±0.053 in the BB group to 1.471±0.042 in the AA genome. The effective number of alleles (Ne) ranged from 1.165±0.020 (BB group) to 1.379±0.026 (AA group). From the SCoT, the overall mean values of Na, Ne, He, uHe, and I and their respective standard deviations across the diverse genomic constitutions were 1.180±0.030, 1.297±0.014, 0.168±0.008, 0.171±0.009 and 0.247±0.011, respectively.

Consequently, the genetic diversity statistics identified in variable genomic (ploidy) groups were ranked: AA>ABB>BB from high to low (Table 3). The high intensity of genetic diversity in the AA cultivars group in this analysis may be due to the remoteness and speciation that conducts the progression of traits. Furthermore, variations in the genetic constitution and ecological conditions will lead to diverse adaptation patterns and other genetic traits that assist taxa existence. Selection and vegetative reproduction are also influential factors that may affect high variation in *Musa* cultivars. Species with plenty of genetic adaptation can tolerate ecological strain for a long instance (Hapsari et al. 2018). In addition, these results strongly agreed with the previously conducted using SCoT analyses by Igwe et al. (2022). Hence, the significant potential of genetic diversity is current within the group of studied germplasm, which can be utilized for better breeding of *Musa* cultivars in the future.

Table 2. The sequence of SCoT primers and polymorphic SCoT bands and polymorphism of *Musa* cultivars

Primer	Sequences (5'–3')	Size (bp)	TAB	NPB	PPB	PIC
SCoT1	CAACAATGGCTACCACCA	150-1,000	7	7	100	0.453
SCoT2	CAACAATGGCTACCACCC	150-900	7	7	100	0.439
SCoT3	CAACAATGGCTACCACCG	50-1,000	8	8	100	0.451
SCoT4	CAACAATGGCTACCACCT	300-700	5	5	100	0.457
SCoT5	CAACAATGGCTACCACGA	100-1,000	5	5	100	0.394
SCoT6	CAACAATGGCTACCACGC	100-1,000	5	5	100	0.435
SCoT7	CAACAATGGCTACCACGG	100-1,200	8	8	100	0.464
SCoT8	CAACAATGGCTACCACGT	400-1,200	7	7	100	0.483
SCoT9	CAACAATGGCTACCAGCA	200-1,000	9	9	100	0.452
SCoT10	CAACAATGGCTACCAGCC	50-1,200	13	13	100	0.47
SCoT11	AAGCAATGGCTACCACCA	150-1,300	7	7	100	0.459
SCoT12	ACGACATGGCGACCAACG	50-1,200	14	14	100	0.469
SCoT13	ACGACATGGCGACCATCG	200-1,000	7	7	100	0.433
SCoT14	ACGACATGGCGACCACGC	50-700	5	5	100	0.433
SCoT15	ACGACATGGCGACCGCGA	100-500	4	4	100	0.394
SCoT16	ACCATGGCTACCACCGAC	50-1,000	8	7	87.5	0.363
SCoT17	ACCATGGCTACCACCGAG	50-1,200	10	10	100	0.445
SCoT18	ACCATGGCTACCACCGCC	200-1,200	8	7	87.5	0.374
SCoT19	ACCATGGCTACCACCGGC	100-1,000	10	10	100	0.451
SCoT20	ACCATGGCTACCACCGCG	300-1,500	12	11	91.67	0.347
SCoT21	ACGACATGGCGACCCACA	100-1,300	13	11	84.62	0.345
SCoT22	AACCATGGCTACCACCAC	300-1,000	7	7	100	0.453
SCoT23	CACCATGGCTACCACCAG	100-1,000	7	7	100	0.43
SCoT24	CACCATGGCTACCACCAT	200-1,000	8	8	100	0.474
SCoT25	ACCATGGCTACCACCGGG	300-1,000	10	9	90	0.414
SCoT26	ACCATGGCTACCACCGTC	200-1,300	5	5	100	0.448
SCoT27	ACCATGGCTACCACCGTG	100-1,200	10	10	100	0.427
SCoT28	CCATGGCTACCACCGCCA	100-1,000	6	6	100	0.433
SCoT29	CCATGGCTACCACCGGCC	400-1,200	8	8	100	0.44
SCoT30	CCATGGCTACCACCGGCG	100-1,000	11	11	100	0.42
	Total		244	238	97.54	0.432

Note: TAB: total amplified bands; NPB: number of polymorphic bands; PPB: percentage of polymorphic bands; PIC: polymorphism information content

Table 3. Summary of genetic diversity parameters

Genome Group	N	NPB	PPB	Na	Ne	He	uHe	I
AA	43	136	55.74	1.471±0.042	1.379±0.026	0.214±0.014	0.216±0.014	0.314±0.019
ABB	30	114	46.72	1.217±0.053	1.349±0.026	0.193±0.014	0.196±0.014	0.280±0.020
BB	17	69	28.28	0.852±0.053	1.165±0.020	0.097±0.011	0.100±0.011	0.147±0.016
Mean	30	106.33	43.58	1.180±0.030	1.297±0.014	0.168±0.008	0.171±0.009	0.247±0.011

Note: N: number of samples; NPB: number of polymorphic bands; PPB: percentage of polymorphic bands; Na: number of different alleles; Ne: number of effective alleles; He: expected heterozygosity; uHe: unbiased expected heterozygosity; I: Shannon's information index

Based on SCoT bands, genetic distances among the 90 accessions were calculated, and a dendrogram was constructed by the UPGMA method (Figure 3A). Dendrogram analysis revealed two major clusters from SCoT information. This variation in the number of subclusters obtained from the SCoT markers could be due to the efficiency of the SCoT markers in targeting different loci, which in turn resulted in the continuation of some accessions in separate groups instead of positioning with their respective larger groups (Igwe et al. 2022; Rai 2023). Most of the genomic groups were also well resolved. For instance, in cluster I, *M. acuminata*, the diploid AA genome dominated the group and indicated greater genetic differences, which among 43 *M. acuminata* accessions was separated into three major subgroup agreements with morphology consisting Khai, Hom and Leb Mu Nang varieties. While the major cluster II can be diversified into two subgroups, each subgroup resolved the accessions based primarily on genomic constitutions. The first major subcluster in cluster II comprised generally *M. balbisiana* (BB) agreement with ploidy evidence comprised only the Tani variety. The second main subgroup in cluster II consisted of *Musa x paradisiaca* (ABB) involving Nam Wa and Hak Muk varieties. This similar hetero-genomic grouping was reported in *Musa* accessions analyzed with several molecular markers (Rotchanapreeda et al. 2016; Chang et al. 2017; Arora et al. 2018; Wanvisait et al. 2019; Boonsrangsom et al. 2020; González et al. 2021). Besides, associations among 90 *Musa* cultivars were also resolved by principal coordinate analysis (PCoA) (Figure 3B). In the diagram generated by PCoA, two main clusters were shown, revealing a similar cluster result as that in the dendrogram; evidently, I and II in the PCoA plot are similar to the agreements in cluster I and II. According to a previous study, *Musa* accessions with AA, ABB, and BB genomes can be divided into different subgroups.

Nevertheless, the present investigation could not cluster the genotypes according to their locality source. The genotypes were probably mixed up in all the clusters irrespective of their locality origin. The lack of clear clustering based on locality origins among the *Musa* cultivars can be due to introductions followed by the movement of cultivars in areas away from their identified origins. Despite the few resources of SCoT markers constructed compared with other molecular analyses yielded more polymorphic data and were more effective in identifying the genetic differences among *Musa* accessions. This indicates the SCoT marker's ability to recognize genetic diversity at both the species and genus levels.

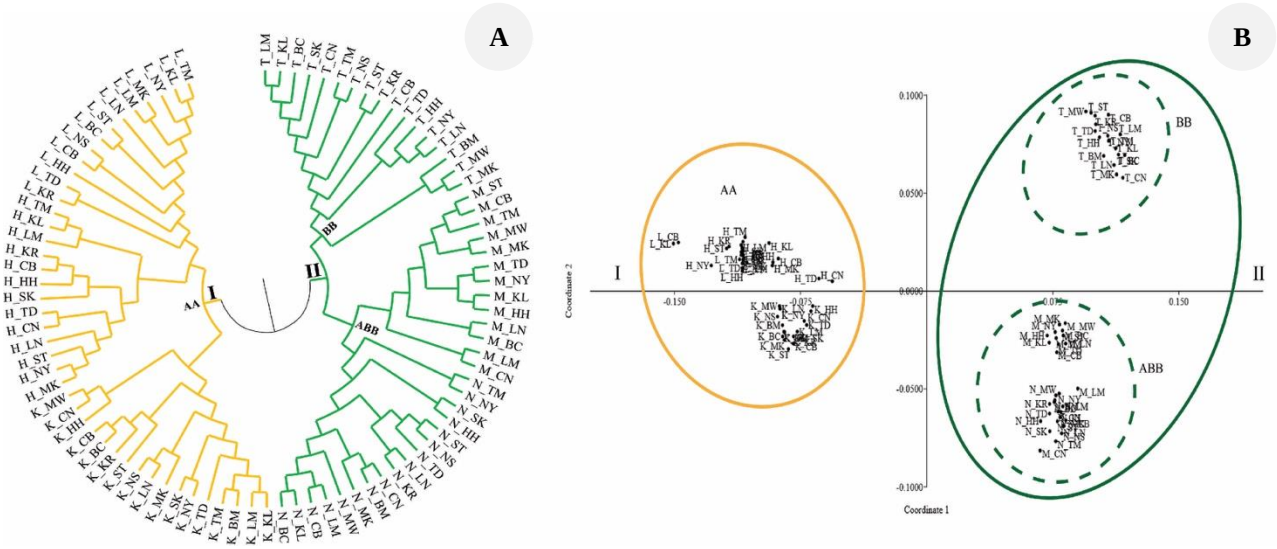

Figure 3. Relationship of 90 *Musa* cultivars with SCoT analysis. A. Dendrogram analysis B. Principal coordinate analysis. Accession codes correspond to samples listed in Table 1

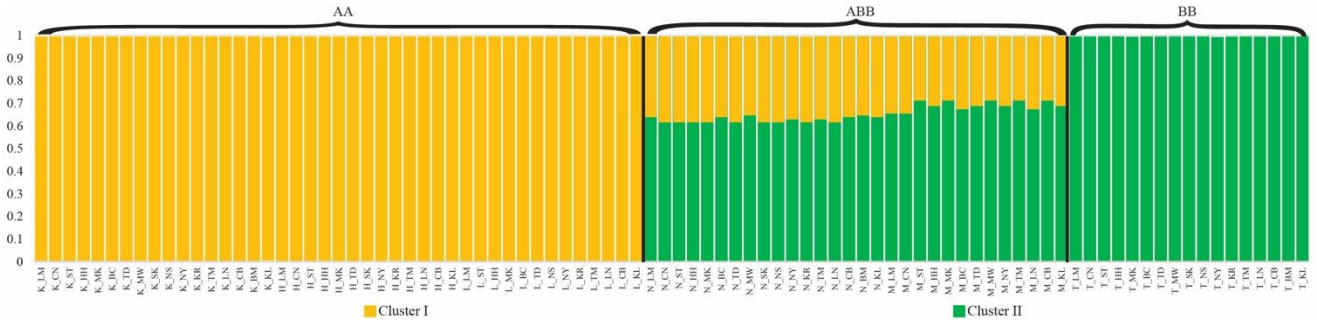

Figure 4. Structure-based clustering of 90 Musa genotypes using 30 SCoT markers. Accession codes correspond to samples listed in Table 1

The Bayesian clustering model performed by the STRUCTURE software, shown in Figure 4, clustered *Musa* accessions into two clusters, supported by the greatest ΔK value at K=2. This K was correlated with dendrogram and PCoA analyses, in which each K represented the A and B genome clusters. The first cluster included only AA genome comprising Khai, Hom and Leb Mu Nang cultivars. The second cluster contained the rest of the cultivars with mostly genotypes of ABB and BB. In addition, the STRUCTURE calculation inferred that *Musa* cultivars separated into two major clusters and most possibly more reflect a continuous genetic gradation or admixture of diploid BB group and triploid ABB groups. The outcomes from STRUCTURE analyses agreed regarding the relatedness of genomic information. This observation is also in concordant with various genetic structure and diversity studies were conducted on Thai *Musa* genetic resources and associated according to the genotypes of cultivars (Rotchanapreeda et al. 2016; Zozimo et al. 2018; Wanvisait et al. 2019; Boonsrangsom et al. 2020; Premjet et al. 2022). The results proposed that SCoT markers were effective in genomic discrimination. It was proved that these performances would bring more efficiency and precision to genetic/breeding programs. Findings herein will suggest the protection of the valuable diversity of genetic resources and germplasm management to generate enhanced populations and improve breeding applications in Thai *Musa* cultivars, particularly in the Chai Badan district of Lopburi Province soon.

In conclusion, determining the degree and distribution of genetic diversity is crucial for future studies of the breeding material to be improved. In addition, accurate and rapid genotype identification is important, especially in vegetatively propagated plant species, such as *Musa* cultivars. In this study, SCoT markers, which are very effective markers for detecting *Musa* plants' genetic diversity, have been evaluated. This is the first study to determine genetic diversity and relationships among *Musa* cultivars in Thailand using SCoT markers. Findings herein will lead to the defense of the beneficial diversity of genetic resources and germplasm management to generate enhanced populations and improve breeding applications in *Musa* studies. Thus, the identification and molecular characterization of common *Musa* cultivars in the Chai Badan district of Lopburi Province, Thailand, can be done by improving more effective and reliable markers. The evidence generated in this study can be utilized for projecting future conservation strategies and sustainable exploitation of cultivated *Musa* gene pools available in the region.

ACKNOWLEDGEMENTS

This work was supported by the Fundamental Fund (65A140000020) of Phranakhon Rajabhat University and Thailand Science Research and Innovation. The author is grateful to all participants for supplying the plant materials and thanks to Nakanate Ngampak for assistance during fieldwork.

REFERENCES

Arora V, Kapoor N, Fatma S, Jaiswal S, Iquebal MA, Rai A, Kumar D. 2018. BanSatDB, a whole-genome-based database of putative and experimentally validated microsatellite markers of three *Musa* species. Crop J 6: 642-650. DOI: 10.1016/j.cj.2018.01.005.

Bawin Y, Panis B, Abeele SV, Li Z, Sardos J, Paofa J, Ge XJ, Mertens A, Honnay O, Janssens SB. 2019. Genetic diversity and core subset selection in ex-situ seed collections of the banana crop wild relative *Musa balbisiana*. Plant Genet Res 17 (6): 536-544. DOI: 10.1017/S1479262119000376.

Biswas MK, Patil A, Sunkad G. 2023. Enhancing legume cultivars through agronomy, breeding, and genetics. Agronomy 13 (4): 1035. DOI: 10.3390/agronomy13041035.

Boban S, Maurya S, Jha, Z. 2022. DNA fingerprinting: An overview on genetic diversity studies in the botanical taxa of Indian Bamboo. Genet Resour Crop Evol 69 (2): 469-498. DOI: 10.1007/s10722-021-01280-8.

Boonsrangsom T, Phetnin B, Ratanasut K, Sujipuli K. 2020. Assessment of genetic diversity among *Musa* cultivars based on sequence-related amplified polymorphism technique. Naresuan Univ J Sci Technol 28 (2): 52-61. DOI: 10.14456/nujst.2020.15.

Carracedo MG, Tejera-Pérez H, Ferrer MH, Arias DJ, Pérez JAP. 2021. Comparative assessment of microsatellite and retrotransposon-based markers for genetic characterization of commercial banana cultivars (*Musa* spp.). Plant Breed 140 (5): 968-980. DOI: 10.1111/pbr.12960.

Chang SF, Chang YL, Yen YF, Miyajima I, Huang KL. 2017. Genetic diversity and sectional relationships from an amplified fragment length polymorphism analysis of Taiwan bananas. J Fac Agric Kyushu Univ 62 (1): 31-37. DOI: 10.5109/1798138.

Collard BC, Mackill DJ. 2009. Start codon targeted (SCoT) polymorphism: A simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol 27 (1): 86. DOI: 10.1007/s11105-008-0060-5.

Christelová P, de Langhe E, Hřibová E, Čížková J, Sardos J, Hušáková M, van den Houwe I, Sutanto A, Kepler AK, Swennen R, Roux N, Doležel J. 2017. Molecular and cytological characterization of the

global *Musa* germplasm collection provides insights into the treasure of banana diversity. Biodivers Conserv 26: 801-824. DOI: 10.1007/s10531-016-1273-9.

Dar AA, Mahajan R, Sharma S. 2019. Molecular markers for characterization and conservation of plant genetic resources. Indian J Agric Sci 89 (11): 1755-1763. DOI: 10.56093/ijas.v89i11.95286.

Davey MW, Gudimella R, Harikrishna JA, Sin LW, Khalid N, Keulmens J. 2013. A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. BMC Genome 14: 683. DOI: 10.1186/1471-2164-14-683.

Dehery SK, Das AB. 2022. Genetic diversity of twelve triploid bananas and plantains under section Eumusa as evident by chromosome morphology and SSR markers. Nucleus 65 (1): 35-48. DOI: 10.1007/s13237-021-00364-1.

Earl AD, von Holdt BM. 2012. Structure Harvester: A website and program for visualizing structure output and implementing the Evanno method. Conserv Genet Resour 4 (2): 359-361. DOI: 10.1007/s12686-011-9548-7.

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.

González CM, Tejera-Pérez H, Hernández FM, Jiménez AD, Pérez PJA. 2021. Comparative assessment of microsatellite and retrotransposon-based markers for genetic characterization of commercial banana cultivars (*Musa* spp.). Plant Breed 140 (5): 968-980. DOI: 10.1111/pbr.12960.

Guo DL, Guo M X, Hou XG, Zhang, GH. 2014. Molecular diversity analysis of grape varieties based on iPBS markers. Biochem Syst Ecol 52: 27-32. DOI: 10.1016/j.bse.2013.10.008.

Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4: 1-9.

Hapsari L, Azrianingsih R, Arumingtyas EL. 2018. Genetic variability and relationship of banana cultivars (*Musa* L.) from East Java, Indonesia based on the Internal Transcribed Spacer region nrDNA sequences. J Trop Biol Conserv 15: 101-120. DOI: 10.51200/jtbc.v15i0.1482.

Hinge VR, Shaikh IM, Chavhan RL, Deshmukh AS, Shelake RM, Ghuge SA, Dethe1 AM, Suprasanna P, Kadam, US. 2022. Assessment of genetic diversity and volatile content of commercially grown banana (*Musa* spp.) cultivars. Sci Rep 12 (1): 7979. DOI: 10.1038/s41598-022-11992-1.

Igwe DO, Ihearahu OC, Osano AA, Acquaah G, Ude GN. 2022. Assessment of genetic diversity of *Musa* species accessions with variable genomes using ISSR and SCoT markers. Genet Resour Crop Evol 6 9(1): 49-70. DOI: 10.1007/s10722-021-01202-8.

Koutouleas A, Collinge DB, Ræbild A. 2023. Alternative plant protection strategies for tomorrow's coffee. Plant Pathol J 72 (3): 409-429. DOI: 10.1111/ppa.13676.

Lamare A, Rao SR. 2015. Efficacy of RAPD, ISSR and DAMD markers in assessment of genetic variability and population structure of wild *Musa acuminata* Colla. Physiol Mol Biol Plants 21: 349-358. DOI: 10.1007/s12298-015-0295-1.

Marimuthu SS, Subbaraya U, Durairajan SG, Rajendran S, Gopalakrishnan J, Hameed S, Palani BD, Suthanthiram B. 2019. Comparison of two different electrophoretic methods in studying the genetic diversity among plantains (*Musa* spp.) using ISSR markers. Electrophoresis 40 (9): 1265-1272. DOI: 10.1002/elps.201800456.

Mekonnen TW, Gerrano AS, Mbuma NW, Labuschagne MT. 2022. Breeding of vegetable cowpea for nutrition and climate resilience in Sub-Saharan Africa: Progress, opportunities, and challenges. Plants 11 (12): 1583. DOI: 10.3390/plants11121583.

Mertens A, Bawin Y, Vanden AS., Kallow S, Toan VD, Thi LL, Vu TD, Swennen R, Vandelook F, Panis B, Janssens SB. 2021. Genetic diversity and structure of *Musa balbisiana* populations in Vietnam and its implications for the conservation of banana crop wild relatives. PLoS ONE 16: e0253255. DOI: 10.1371/journal.pone.0253255.

Peakall R, Smouse PE. 2012. GENALEX6.5: Genetic analysis in EXCEL. Population genetic software for teaching and research an update. Bioinform 28: 2537-2539. DOI: 10.1093/bioinformatics/bts460.

Premjet D, Boonsrangsom T, Sujipuli K, Rattanasut K, Kongbungkerd A, Premjet, S. 2022. Morphological and molecular characterizations of *Musa* (ABB) 'Mali-Ong'in Thailand. Biology 11 (10): 1429. DOI: 10.3390/biology11101429.

Rai MK. 2023. Start codon targeted (SCoT) polymorphism marker in plant genome analysis: current status and prospects. Planta 257 (2): 34. DOI: 10.1007/s00425-023-04067-6.

Rittirongsakul K, Srinual A, Vanijajiva O. 2020. Anatomical features and SCoT profiles provide new insight into phenotypic plasticity in the halophyte *Suaeda maritima* in Thailand. Biodiversitas 21 (3): 1082-1090. DOI: 10.13057/biodiv/d210331.

Roldan-Ruiz I, Dendauw J, Bockstaele EV, Depicker A, Loose MD. 2000. AFLP marker reveal high polymorphic rates in ryegrasses (*Lollium* spp.). Mol Breed 6: 125-134. DOI: 10.1023/A:1009680614564.

Rotchanapreeda T, Wongniam S, Swangpol SC, Chareonsap PP, Sukkaewmanee N, Somana, J. 2016. Development of SSR markers from *Musa balbisiana* for genetic diversity analysis among Thai bananas. Plant Syst Evol 302: 739-761. DOI: 10.1007/s00606-015-1274-2.

Santos TTC, Amorim VBO, Santos-Serejo JA, Ledo CAS, Haddad F, Ferreira CF, Amorim EP. 2019. Genetic variability among autotetraploid populations of banana plants derived from wild diploids through chromosome doubling using SSR and molecular markers based on retrotransposons. Mol Breed 39 (7): 95-107. DOI: 10.1007/s11032-019-0996-1.

Saraswathi MS, Uma S, Ramaraj S, Durai P, Mustaffa MM, Kalaiponmani K, Chandrasekar A. 2020. Inter retrotransposon based genetic diversity and phylogenetic analysis among the *Musa* germplasm accessions. J Plant Biochem Biotechnol 29: 114-124. DOI: 10.1007/s13562-019-00519-x.

Simmonds NW. 1957. Classification of the bananas. Kew Bulletin 11: 463-489.

Singh WA, Singh NS, Devi EJ, Handique PJ, Devi HS. 2021. Collection and characterization of banana gene pools (*Musa* spp.) in Manipur (NE India) using PCR–RFLP and RAPD and ISSR markers. Braz J Bot 44 (3): 671-684. DOI: 10.1007/s40415-021-00722-y.

Silva AVC, Nascimento ALS, Vitória MF, Rabbani ARC, Soares ANR, Lédo AS. 2017. Diversity and genetic stability in banana genotypes in a breeding program using inter simple sequence repeats (ISSR) markers. Genet Mol Res 16: 1-9. DOI: 10.4238/gmr16019402.

Swain S, Das S, Munsi PS, Lenka PC, Rout GR, Swain D. 2016. Molecular diversity study on dessert banana genotypes (*Musa* spp.) from Odisha using ISSR markers. Intl J Agric Environ Biotechnol 9 (4): 513-523. DOI: 10.5958/2230-732X.2016.00068.1.

Vanijajiva O. 2020. Start codon targeted (SCoT) polymorphism reveals genetic diversity of *Manilkara* in Thailand. Biodiversitas 21 (2): 666-673. DOI: 10.13057/biodiv/d210232.

Vanijajiva O, Pornpongrungrueng P. 2020. Inter-primer binding site (iPBS) markers reveal the population genetic diversity and structure of tropical climbing *Cissampelopsis* (Asteraceae) in Thailand. Biodiversitas 21: 3919-3928. DOI: 10.13057/biodiv/d21090.

Wahyudi D, Nursita DC, Hapsari L. 2022. Genetic diversity among and within genome groups of banana cultivars based on ISSR markers. Intl J Agric Biol 2 (6): 366-374. DOI: 10.17957/IJAB/15.1990.

Wahyudi D, Rifliyah K. 2020. Genome evaluation of banana cultivars based on morphological character and Inter-Simple Sequence Repeat (ISSR) molecular marker. Biodiversitas 21 (7): 2982-2990. DOI: 10.13057/biodiv/d210715.

Wanvisait P, Pinthong K, Tankrathok A, Sukkapan P, Srisamoot N. 2019. Assessing genetic diversity of some banana cultivars using inter simple sequence repeats (ISSR) markers. Khon Kaen Agr J 47: 1449-1458.

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: 39. DOI: 10.1186/s12870-023-04056-7.

Zozimo ROB, Ratanasut K, Boonsrangsom T, Sujipuli K. 2018. Assessment of genetic diversity among Thai banana cultivars (*Musa* spp.) based on RAPD and SRAP markers. Intl J Biosci 12 (4): 172-180. DOI: 10.12692/ijb/12.4.172-180.