Acta Scientific AGRICULTURE (ISSN: 2581-365X)

Volume 5 Issue 4 April 2021

Research Article

Morphological Characterization and Genetic Diversity Analysis of wild *Musa* Collections from Garo Hills, Meghalaya by SPAR approach

Anju Hajong[1], Sohini Deb[2], Pronomi M Sangma[1], CP Suresh[1] and Satyawada Rama Rao[2]*

[1]Department of Horticulture, North Eastern Hill University, Tura, Meghalaya, India

[2]Plant Biotechnology Laboratory, Department of Biotechnology and Bioinformatics, North Eastern Hill University, Permanent Campus, Shillong, Meghalaya, India

***Corresponding Author:** Satyawada Rama Rao, Department of Biotechnology and Bioinformatics, Plant Biotechnology Laboratory, North-Eastern Hill University, Shillong, Meghalaya, India.

Received: February 19, 2021

Published: March 06, 2021

Abstract

North east India is a major biodiversity hotspot as it is bestowed with various natural resources. Meghalaya, in particular is known to house many horticultural crops, out of which, the great diversity of *Musa* sp is highly significant. In the present study, 21 genotypes of *Musa* sp. were collected from five distcits of Garo Hills of Meghalaya, India and were analyzed to understand their morphological and genetic variation. The intra-specific relationship prevalent among them was also evaluated. Some important morphological parameters were selected to understand the variation present in the collected genotypes. Further, three single primer based DNA markers *viz.* Random Amplified Polymorphic DNA (RAPD), Inter Simple Sequence Repeats (ISSR) and Directed Amplifications of Mini-satellite DNA (DAMD) were chosen for the diversity analysis. A total of 33 primers (17 RAPD, 10 ISSR, 6 DAMD) were selected, which yielded 207 DNA amplicons. High level of polymorphism was observed among the genotypes. The polymorphic information content (PIC) values of the markers were similar, ranging from 0.34-0.39. The resolving power (Rp) and marker index (MI) in case of both ISSR and DAMD were found to be similar (~6.9 and 2.4 respectively). High level of genetic diversity values were observed in the population genetic parameters, with a significant Gene flow estimates (Nm>1) of 1.7494, high Nm (1.7494) and Gst (0.2223) values. Nei's gene diversity (h) and Shannon's information index (I) values varied between 0.1998-0.2976 and 0.2981-0.4391 respectively, further establishing the high variation among the genotypes. Analysis of molecular variations (AMOVA) also revealed high level of genetic variation within the populations (97%). The dendrogram and the principal component analysis (PCoA) generated based on the Single Primer Amplification Reaction (SPAR) data revealed high intermixing of the wild *Musa* genotypes.

Keywords: *Musa*; Morphology; SPAR; Population Genetics; Garo Hills; Genetic Diversity

Abbreviations

AMOVA: Analysis of Molecular Variance; CTAB: Cetyl-Trimethyl Ammonium Bromide; PCoA: Principal Coordinate Analysis; PCR: Polymerase Chain reaction; RAPD: Random Amplified Polymorphic DNA; ISSR: Inter Simple Sequence Repeats; DAMD: Directed Amplifications of Mini-satellite DNA; SPAR: Single Primer Amplification Reaction; PIC: Polymorphic Information Content; Rp: Resolving power; MI: Marker Index; EMR: Effective Multiplex Ratio; S.Em: Standard Error of Mean; CD: Critical Difference, CV: Coefficient of Variance.

Introduction

Bananas (*Musa* spp.) are monocotyledonous, herbaceous, parthenocarpic plants belonging to the section *Eumusa* under the family Musaceae [24]. It is staple food for millions of people in the developing countries like India and Africa. All the edible banana varieties that are cultivated currently are known to have originated from the two wild species *Musa acuminata* and *M. balbisiana,* through inter and intraspecific hybridization crosses [30]. Wild *Musa* species are largely distributed in tropical and sub-tropical regions, like the North-east India, the Himalayan region, Assam, Arunachal Pradesh, Meghalaya and some parts of Nagaland. Additionally, wild *Musa* species are found in the hilly terrains of Khasi, Jaintia, Naga, Patkai and Garo hills of North-East India, at both lower and higher altitudes [15].

In population of a plant species, variations are known to occur both, genotypically and phenotypically [32]. Though genetic markers are hallmarks to determine the extent of diversity and inter-relationships among them, the phenotypic variations among individuals of a species, may also provide significant insight into their evolutionary patterns [4,9]. Morphological parameters, which are under constant selection pressure, are expected to provide an understanding on how different environmental factors result in different phenotypic characters and demonstrate their role in the evolution of a plant species in question [18].

Molecular marker checks for the presence of the specific nucleotide sequences which encodes a particular trait or character [11]. The PCR based DNA marker techniques provide reliable and authentic genetic information of a species [15]. Random Amplified Polymorphic DNA (RAPD), Inter Simple Sequence Repeats (ISSR) and Directed Amplifications of Mini-satellite DNA (DAMD) are some of the DNA-based marker techniques that has been successfully implemented and widely used to determine the genetic diversity and relationships in several *Musa* sp. from various populations [3,12,15,17,33]. These single primer markers yield many amplicons after PCR, which can be used to study the genetic variation as a multilocus marker system [20]. RAPD uses arbitrary primers and is advantageous being rapid and cost-efficient [22]. Similarly, ISSR technique is considered as a fast, cost-effective, highly discriminative and reliable technique for genetic diversity analysis [25]. DAMD markers are advantageous as they are highly polymorphic [22]. Recently, the single primer amplification reaction (SPAR) method has been established as an efficient tool for genetic diversity analysis in plants [15,16,19,26]. SPAR includes data generated by RAPD, ISSR and DAMD to provide a comprehensive description of the extent of the existing diversity available among the genotypes [2,13,28,29].

The wild *Musa* sp. are the progenitors of the present cultivated varieties, and hence their molecular characterization is vital for breeding and conservation purposes [4]. Wild varieties of *Musa* have been reported to have certain favourable characteristics like resistance to diseases, etc. [10,15]. In spite of being bestowed with so many important values, only a few studies have been conducted on the wild *Musa* varieties of Garo Hills, Meghalaya. Molecular characterization studies were carried out on wild *Musa* varieties collected from various locations of Meghalaya [15,16]. However, only a few genotypes from the Garo Hills region was included in that study. Therefore, the characterization of the various wild *Musa* sp. and the detection of their genetic relationship is of utmost importance to devise their biofortification and conservation programmes.

Materials and Methods

Sample collection

Exploration trips were carried out to the various locations of Garo Hills in Meghalaya. A total of 21 genotypes of *Musa* spp. were collected from the 5 districts *viz.,* East Garo Hills, West Garo Hills, North Garo Hills, South Garo Hills and South West Garo Hills (Figure 1) (Table 1). Young and disease-free leaves from the plants were collected, cleaned and stored at -80°C for future experiments.

Figure 1: Geographical locations of wild *Musa* species in five districts of Garo Hills, Meghalaya, India, with numbers indicating the collection sites.

Sl. No.	Collection ID	Place of collection	Latitude	Longitude
	KK-WGH	West Garo Hills (Rongkhon)	N 25°32'10.51"	E90°13'45.02"
	Go-EGH	East Garo Hills (Mandalgre)	N 25°30'67.42"	E90° 23'55.08"
	CT-NGH	North Garo Hills (Mendi)	N 25°35'16.90"	E90° 38'23.61"
	CH-WGH	West Garo Hills (Upper Chandmari)	N 25°30'67.42"	E90° 23'55.82"
	WRL-NGH	North Garo Hills (Resubelpara)	N 25°35'16.91"	E90° 38'23.60"
	TD-EGH	East Garo Hills (Bandigre)	N 25°30'92.06"	E90°22'29.07"
	EB-SWGH	South West Garo Hills (Zikzak)	N 25°26'38.53"	E89°55.67.06"
	EG-SWGH	South West Garo Hills (Phuljuri)	N 25°22'19.04"	E89°52'26.14"
	TG-WGH	West Garo Hills (Daribok)	N 25°29'31.20"	E90°19'26.17"
	MR-EGH	East Garo Hills (Samanda)	N 25°35'07.72"	E90°30'93.82"
	SM-SGH	South Garo Hills (Siju)	N 25°21'09.07"	E90°34'89.90"
	ZN-SGH	South Garo Hills (Nengkong)	N 25°17'85.29"	E90°36'46.27"
	DG-EGH	East Garo Hills (Bansamgre)	N 25°34'12.14"	E90°28'05.18"
	RL-NGH	North Garo Hills (RongmaGitil)	N 25°35'16.90"	E90°38'23.61"
	RG-NGH	North Garo Hills (RongmaGitil)	N 25°51'78.42"	E90°51'98.55"
	MN-SWGH	South West Garo Hills (Borkona)	N 25°20'03.72"	E89°50'02.07"
	ES-WGH	West Garo Hills (Darechekgre)	N 25°32'24.42"	E90°15'47.03"
	AB-SWGH	South West Garo Hills (Mahendra-ganj)	N 25°19'20.71"	E89°51'67.24"
	WB-WGH	West Garo Hills (Sasatgre)	N 25°30'32.51"	E90°20'02.18"
	AS-SWGH	South West Garo Hills (Mahendraganj)	N 25°20'03.16"	E90°50'03.06"
	RB-SGH	South Garo Hills (Rongdong)	N25°22'58'12.17"	E90° 40'70.11"

Table 1: Places of collections of the *Musa* genotypes from Garo Hills, Meghalaya.

Morphological analysis

A few morphological traits of the collected genotypes of *Musa* were described by following the descriptors on morphological and physical traits, provided by the International Network for the Improvement of Banana and Plantain [1] (Table 2, 3).

Sl. No.	Collection ID	Peduncle length (cm)	Fruit pedicel length (mm)	Fruit length (cm)	Fruit weight (g)
1.	EB-SWGH	33.06	11.50	12.66	22.856
2.	MN-SWGH	63.33	17.15	26.16	145.53
3.	AB-SWGH	32.66	11.18	11.66	120.63
4.	AS-SWGH	38.5	17.41	17.26	108.80
5.	EG-SWGH	46.52	18.33	17.26	59.883
6.	SM-SGH	38.26	16.76	17.66	121.62
7.	RB-SGH	40.63	16.83	16.16	146.93
8.	ZN-SGH	40.16	16.50	16.53	118.6
9.	Go-EGH	30.41	16.78	14.45	136.8
10.	TD-EGH	31.46	13.75	13.33	128.2
11.	DG-EGH	28.41	14.21	13.33	155.88
12.	MR-EGH	28.55	17.78	20	111.65
13.	CH-WGH	61.6	9.58	19.33	126.65
14.	WRL-NGH	58.43	8.03	14.33	90.913
15.	RL-NGH	26.93	12.15	11.9	152.43
16.	RG-NGH	54	17.60	23.83	188.22
17.	CT-NGH	36.16	16.36	15.5	117.86
18.	ES-WGH	76.1	15.81	17.83	42.57
19.	WB-WGH	34.51	16.15	17.83	16.903
20.	TG-WGH	33.9	12.25	15.83	46.226
21.	KK-WGH	66.33	16.58	18	217.68
	Mean	33.06	11.5	12.66	22.85
	S.Em±	0.931	0.703	1.087	3.875
	CD (5%)	2.662	2.010	3.108	11.077
	CV (%)	4.1166	8.182	11.270	5.930

Table 2: Metric traits of *Musa* species.

*SEm: Standard Error of Mean, CD: Critical Difference, CV: Coefficient of variance.

Sl. No.	Collection ID	Leaf habit	Pseud-ostemcolour	Pseudostem appearance	Sap colour	Male bud type	Male bud shape	Pulp in fruit	Pulp colour before maturity
1	EB-SWGH	Erect	Green	Dull (Waxy)	Milky	Normal (present)	Lanceolate	With pulp	White
2	RB-SGH	Intermediate	Red	Shiny (Not waxy)	Red-purple	Normal (present)	Ovoid	With pulp	White
3	Go-EGH	Erect	Medium green	Dull (Waxy)	Watery	Normal	Rounded	With pulp	White
4	CT-NGH	Drooping	Dark green	Dull (Waxy)	Milky	Normal	Intermediate	With pulp	White
5	KK-WGH	Intermediate	Dark green	Shiny (Not waxy)	Watery	Normal	Lanceolate	With pulp	White
6	CH-WGH	Drooping	Green	Shiny (Not waxy)	Red-purple	Normal	Intermediate	With pulp	White
7	WRL-NGH	Erect	Green	Shiny (Not waxy)	Red-purple	Normal	Lanceolate	With pulp	White
8	TD-EGH	Intermediate	Green	Dull (Waxy)	Milky	Degenerat-ing before maturity	Intermediate	With pulp	White
9	SM-SGH	Erect	Green-yellow	Shiny (Not waxy)	Watery	Degenerat-ing before maturity	Ovoid	With pulp	White
10	ZN-SGH	Intermediate	Green-red	Dull (Waxy)	Watery	Normal (present)	Rounded	With pulp	White
11	DG-EGH	Erect	Medium green	Dull (Waxy)	Watery	Normal	Rounded	With pulp	White
12	RL-NGH	Drooping	Medium green	Shiny (Not waxy)	Milky	Normal	Ovoid	With pulp	White
13	RG-NGH	Intermediate	Medium green	Shiny (Not waxy)	Milky	Degenerat-ing before maturity	Lanceolate	With pulp	White
14	MN-SWGH	Drooping	Green-yellow	Dull (Waxy)	Watery	Normal	Lanceolate	With pulp	White
15	ES-WGH	Erect	Green	Shiny (Not waxy)	Watery	Normal	Lanceolate	With pulp	White
16	AB-SWGH	Intermediate	Dark green	Dull (Waxy)	Milky	Normal (present)	Intermediate	With pulp	White
17	WB-WGH	Intermediate	Green	Dull (Waxy)	Milky	Normal	Lanceolate	With pulp	White
18	AS-SWGH	Intermediate	Red	Shiny (Not waxy)	Red-purple	Normal (present)	Intermediate	With pulp	White
19	TG-WGH	Erect	Medium green	Dull (Waxy)	Watery	Normal	Rounded	With pulp	White
20	MR-EGH	Drooping	Dark green	Dull (Waxy)	Milky	Normal	Ovoid	With pulp	White
21	EG-SWGH	Intermediate	Dark green	Shiny (Not waxy)	Watery	Normal	Lanceolate	With pulp	White

Table 3: Morphological characters of collected genotypes.

Genomic DNA isolation

Total genomic DNA was isolated from young, disease free leaves of the collected samples of *Musa* spp using CTAB (Cetyltrimethyl ammonium bromide) method with some minor modifications [5]. The quality and quantity of the purified DNA was ascertained by 0.8% (*w/v*) agarose gel electrophoresis and the absorbance ratio (A260/A280) of DNA using Nano-VuePlus™ (GE Healthcare Limited United Kingdom).

PCR optimization and primer selection

Four RAPD kits (OPA, OPC, OPK and OPX) comprising of 20 decamer random primers per kit (total 80 primers) were used, which were procured from Operon Technologies, Alameda, CA, USA. In addition, 36 ISSR primers and 20 DAMD primers obtained from M/S Integrated DNA Technologies (IDT), USA were also used. To optimize the PCR conditions for amplification,varying concentrations of template DNA (20-60 ng), dNTPs (0.1-0.3 mM) and $MgCl_2$ (0–5 mM) were used. Amplification of each of the primers were carried out thrice, and out of the those, 33 (17 RAPD, 10 ISSR, 6 DAMD) primer pairs which produced clear, consistent and reproducible bands were selected for scoring and further analysis.

DNA marker analysis

After the screening of 80 primers from four RAPD kits, 17 primers were found to produce clear and consistent bands. 10 ISSR primers and 6 DAMD primers, which produced clear and reproducible bands were selected. The PCR amplifications were performed as per Lamare and Rao (2015) with certain minor modifications.

Gel electrophoresis

The amplified PCR products were resolved on 2% agarose gel (stained with ethidium bromide) and visualized using a Gel Documentation system (DNA-Minilumi, DNR Bioimaging System, Israel). A DNA ladder of100bp (Thermo Fischer Scientific, USA) was used to estimate the sizes of amplicons generated (Figure 3).

Data scoring and analysis

The banding profile were scored against the presence or absence of a DNA band and denoted as '1' or '0', respectively. Faint bands that could not be scored clearly were not considered for analysis. The molecular sizes of the amplicons were estimated by using 100bp DNA ladder. Discriminatory power of each of the markers (RAPD, ISSR and DAMD) was calculated by three parameters viz. polymorphic information content (PIC), resolving power (RP), and marker index (MI). PIC measures the discriminatory ability and the informativeness of a marker and is calculated [27]. Resolving power (RP) which detects the level of variations between individual genotypes was estimated [23]. Further, marker Index (MI) which may provide an estimate of marker utility was calculated [6]. The dendrograms were generated using the pair-wise distance data by the software DARwin 6.0.21 [21]. POPGENE version 1.32 was used to generate the Nei's genetic diversity (h), Shannon index (I) and the diversity among the populations (G_{ST}) for the analysis of population genetics. Free-Tree 0.9.1.5 software was used to construct Nei's (1978) unbiased genetic diversity dendrogram [6]. AMOVA (Analysis of Molecular Variance) and PCoA (Principal Co-ordinate Analysis) results were generated using GenAlEx software, version 6.503 [9,19].

Results

Morphological characteristics

Data pertaining to various parameters is represented in table 2. The average measurements of the peduncle, fruit pedicel, fruit besides fruit weight were recorded as 33.06cm, 11.5mm, 12.66cm and 22.85gm respectively. The appearance of pseudostem were mostly waxy and its colour varied from dark/bluish green to light green. However, a striking difference was observed in genotypes RB-SGH and AS-SWGH, which had red pseudostem. In all the genotypes pulp was present and it was invariably white in colour. A Principal Component Analysis (PCoA) was also performed using the data obtained from the physical parameters (Figure 2).

Figure 2: PCo Analysis of five populations comprising of 21 accessions of *Musa* sp.from five districts of Garo Hills (Colour Coded) generated by using the various physical parameters. The numbers indicate the collected samples (Table 2).

RAPD analysis

The 17 primers with reliable banding patterns, produced a total of 92 amplicons, ranging from 10-100bp. A very high percentage of polymorphism was observed. The Polymorphic Information Content (PIC) value ranged from 0.14 (OPA-4) to 0.48 (OPX-3, OPC-7), having a mean value of 0.35 per primer. The average RP value was found to be 3.83 with OPA-13 andOPK-2 having the highest (8.00) and lowest (0.57)values, respectively. The Marker Index (MI) for RAPD was calculated to be 1.90.

ISSR Analysis

The 10 selected ISSR primers generated a total of 63 amplicons and showed a high degree of polymorphism. The PIC values ranged from 0.30 (ISSR-2) to 0.48 (ISSR-17899A), with an average value of 0.39. The RP value was highest for ISSR-17899B (12.19) and lowest for ISSR-P8 (3.05) and an average value at 6.89. The MI for ISSR was found to be 2.46.

DAMD analysis

Six DAMD primers were selected after thorough screening,which generated 46 ampliconsin total. The percentage of polymorphic bands was found to be 96.67%. The PIC values ranged from 0.25 (D-8) to 0.44 (D-4) with an average of 0.34. The average RP value was found to be 6.90, with the highest in D-8 (14.57) and lowest in D-11(2.29). The MI for DAMD was found to be 2.49.

SPAR analysis

The gel profiles obtained after DNA amplification using the three markers revealed polymorphism at various levels and independent of each other (Figure 3). A combined analysis (SPAR) was performed as well for a better understanding of the diversity in the collected samples of *Musa* sp. The 33 primers which were selected yielded a total of 207 amplicons. The polymorphism percentage was found to be 98.89%. The mean values of PIC, Rp and MI of the three markers were found to be 0.36, 5.87 and 2.28 respectively (Table 4).

Figure 3: Banding profiles obtained with (a)RAPD primer, (b) ISSR primer and (c)DAMD primer with all 21 individuals of *Musa* sp.

Sl. No.	Primer	Sequences (5'-3')	TB	PB	MB	PPB	PIC	Rp	MI
RAPD									
1	OPK-10	GTGCAACGTG	5	5	0	100	0.45	3.52381	1.9
2	OPA-4	AATCGGGCTG	8	8	0	100	0.14	1.428571	
3	OPA-2	TGCCGAGCTG	7	7	0	100	0.42	7.714286	
4	OPC-5	GATGACCGCC	6	6	0	100	0.22	1.52381	
5	OPA-13	CAGCACCCAC	7	7	0	100	0.38	8.0	
6	OPK-8	GAACACTGGG	8	8	0	100	0.42	4.952381	
7	OPC-11	AAAGCTGCGG	7	7	0	100	0.48	6.761905	
8	OPC-9	CTCACCGTCC	4	4	0	100	0.47	4.190476	
9	OPX-3	TGGCGCAGTG	3	3	0	100	0.29	1.142857	
10	OPC-12	TGTCATCCCC	5	5	0	100	0.43	3.52381	
11	OPK-2	GTCTCCGCAA	6	6	0	100	0.09	0.571429	
12	OPK-6	CACCTTTCCC	5	5	0	100	0.34	2.190476	

13	OPK-4	CCGCCCAAAC	2	2	0	100	0.25	0.666667	
14	OPC-7	GTCCCGACGA	7	7	0	100	0.48	6.47619	
15	OPC-19	GTTGCCAGCC	4	4	0	100	0.35	2.095238	
16	OPA-3	AGTCAGCCAC	8	8	0	100	0.44	7.714286	
17	OPC-8	TGGACCGGTG	6	6	0	100	0.33	2.571429	
Total:			98	98	0				
Average						100	0.35	3.83	
ISSR									
1	I-17898A	CACACACACACAAC	7	7	0	100	0.47	6.666667	2.46
2	I-17899B	CACACACACACAGG	8	8	0	100	0.34	12.19048	
3	I-P3	AGAGAGAGAGAGAGAGTG	6	6	0	100	0.38	8.666667	
4	I-P6	CCACCACCACCACCA	6	6	0	100	0.33	6.952381	
5	I-P8	CACCCACCACCACCA	5	5	0	100	0.41	3.047619	
6	I-17898B	CACACACACACAGT	6	6	0	100	0.36	3.52381	
7	I-17899A	CACACACACACAAG	8	8	0	100	0.48	8.380952	
8	I-5	CACCACCACGC	6	6	0	100	0.38	6.666667	
9	I-2	CTCTCTCTCTCTCTCTAC	8	8	0	100	0.3	9.047619	
10	I-3	TCCTCCTCCTCCTCCAC	3	3	0	100	0.46	3.714286	
Total:			63	63	0				
Average						100	0.39	6.89	
DAMD									
1	D-8	AATGTGGGCAAGCTGGTGGT	10	8	2	80	0.25	14.57143	2.49
2	D-10	GGACAAGAAGAGGATGTGGA	6	6	0	100	0.39	7.619048	
3	D-1	ATCCAAGGTCCGAGACAACC	9	9	0	100	0.33	4.571429	
4	D-2	GTGTGCGATCAGTTGCTGGG	5	5	0	100	0.36	2.761905	
5	D-4	AGGACTCGATAACAGGCTCC	9	9	0	100	0.44	9.619048	
6	D-11	TACACGTCTCGATCTACAGG	7	7	0	100	0.27	2.285714	
Total:			46	44	2				
Average						96.67	0.34	6.90	
			SPAR						
	SPAR		207	205	2	98.89	0.36	5.87	2.28

Table 4: RAPD, ISSR, and DAMD primers used for amplification. TB, total band; PB, polymorphic band; MB monomorphic band; PPB, percentage polymorphic band; PIC, polymorphic information content; RP, resolving power.

Cluster analysis

RAPD analysis

The dendrogram was generated using Neighbour-joining method and it consisted of three major clusters (Figure 4). Cluster I has two sub-divisions into IA and IB and a total of 9 samples. The genotype WGH-TGb at Cluster I, is observed to have the shortest branch, thereby denoting its probable ancestry. Six collections were retained at Cluster II where a mixed clustering is observed. The collections from SWGH demonstrate close similarity at IB. SWGH-ABb genotype in Cluster III, appears to have the longest branch and therefore can possibly be most recently evolved. It is followed shortly by the WGH-WBb genotype in the evolutionary history. All the three clusters show notable bootstrap values, signifying reliability of the tree generated. Cluster III comprises of samples representing all the collection sites. The genotypes SWGH-ASb and SGH-RBa showed the highest bootstrap value (71%) denoting a strong relationship among them.

Figure 4: Cluster analysis of RAPD data for 21 genotypes of *Musa* from Garo Hills, Meghalaya. The NJ tree was generated with 100 replicate bootstrap analysis.

ISSR analysis

The ISSR dendrogram separates the collections into two major clusters. Cluster I, comprises of 13 samples with two noticeable sub-clusters IA and IB (Figure 5). IA shows distinct similarity among collections from the same sampling sites, NGH and SWGH, respectively. Interestingly, IB on the other hand comprises of 3 samples from unrelated sampling locations.Cluster II, demonstrates mixed clustering of collections with a significant bootstrap value. The appearance of SWGH-ABb collection in Cluster III, as the longest branch represents its recent position in the evolutionary time frame.A striking feature of Cluster III was the appearance of the genotype, SGH-ZNa with the shortest branch which signified its ancestral position.

Figure 5: Cluster analysis of ISSR data for 21 genotypes of *Musa* from Garo Hills, Meghalaya. The NJ tree was generated with 100 replicate bootstrap analysis.

DAMD analysis

The DAMD dendrogram distinguishes the 21 collections into 2 distinct clusters, I and II (Figure 6). Cluster I comprises of 9 genotypes demonstrating similarity from different sampling sites. There is a mixed clustering of collections in both the clades. Cluster II, has two further sub-divisions, IIA and IIB.The genotype SWGH-ABb appears to be the most recent as it has the longest branch. EGH-MRa appears as a distinct outgroup with the shortest branch length.

Figure 6: Cluster analysis of DAMD data for 21 genotypes of *Musa* from Garo Hills, Meghalaya. The NJ tree was generated with 100 replicate bootstrap analysis.

SPAR analysis

Two distinct clusters, I and II are formed in the SPAR dendrogram (Figure 7). An overall mixed clustering was observed among the collections. Eight samples, from random sampling sites are assembled in Cluster I**.** The genotype SWGH-ABb has the longest branch, followed shortly by WGH-WBb**.** The same cluster also has the genotype SGH-RBa, with the shortest branch, which might be an indication of its ancestral position.Cluster II groups 12 genotypes, from different collection sites. The genotype, WGH-TGb appears as an outgroup. Both the clusters have significant values of bootstrap confidence, with the genotypes EGH-MRa and SGH-SMb having the highest bootstrap value (81%).

Figure 7: Cluster analysis of SPAR data for 21 genotypes of *Musa* from Garo Hills, Meghalaya. The NJ tree was generated with 100 replicate bootstrap analysis.

Diversity analysis

The 5 populations of *Musa* sp. were evaluated to estimate their gene flow. SPAR data was used to generate the various parameters to ascertain genetic diversity of the samples (Table 5). The Na (observed number of alleles) values ranged from 1.5362-1.7729. The Ne (effective number of alleles) were in the range of 1.3386-1.5165. Number of polymorphic loci (n) ranged from 111-160, while percentage of polymorphism among the populations ranged between 53.62%-77.29%. Nei's gene diversity (h) and Shannon's information index (I) values varied between 0.1998-0.2976 and 0.2981-0.4391, respectively. Gene flow (Nm) and the diversity among populations (G_{ST}) were found to be 1.7494 and 0.2223 respectively.

Nei's unbiased measures (Table 6) were also calculated to interpret the Genetic identity and Genetic distance amongst the populations. Nei's Maximum (0.9402) was accorded between populations P2(WGH) and P4 (EGH) while Minimum (0.8252) genetic identity was observed among populations P1 (NGH) and P3 (SGH). Nei's maximum genetic distance(0.1922) was recorded between populations P3 (SGH) and P1 (NGH) and minimum values (0.0617) were observed for populations P4 (EGH) and P2 (WGH). The AMOVA data generated shows a high level of genetic variation within populations (93%) and a little variation among the populations (3%) (Table 7). An UPGMA dendrogram based on the Nei's (1978) unbiasedgenetic diversity analysis was also generated (Figure 8) to ascertain the relationship among the five *Musa* populations of Garo Hills (P1 = North Garo Hills, P2 = West Garo Hills, P3 =South Garo Hills, P4 =East Garo Hills, P5= South-West Garo Hills).

Discussion

Musa genotypes were studied to understand their morphological variations, majority of which showed significant differences among them. However, a few parameters like bract arrangement, the presence of pulp and its colour showed no difference among the genotypes. The PCoA graph generated, based on the studied metric traits of the genotypes (Figure 2) further confirmed the presence of high variation and inter-mixing among the *Musa* genotypes, from different populations. Apparently no reports about the quantum and range of variation in morphological traits of *Musa* collections of Garo Hills are available. Though Lamare and Rao (2015), observed and reported significant variation in DNA marker profiles, but did not record the morphological variations in Garo Hill populations. A critical appraisal of genetic variation has been undertaken to define the existing morphological variability.

Several technologies involving DNA-based molecular markers are known, which have proved to be powerful tools to unravel the genetic intricacies of plants as well as aid in the study of species inter-relationships of different populations [7,14,31]. Several studies have reported the reliability of molecular markers like RAPD, ISSR, DAMD in assessing the genetic information within individuals of a species [6,7,26]. In this study,these three markers were used individually as well as in combination (SPAR) to have a better and clearer understanding of the genetic diversity [15,26] among the collections from Garo Hills of Meghalaya, India. All the markers used have showed significant level of polymorphism. RAPD and ISSR have generated an exceptionally high level of polymorphism (100%), followed by DAMD at 96.67%. Parameters for evaluation of discriminatory efficacy of the markers were found to be significant for all the three markers. However, the data generated by com-

Population	N	Na ± SD	Ne ± SD	I ± SD	h ± SD	n	P (%)	G_{ST}	Nm
P1	4	1.6280 ± 0.4845	1.4129 ± 0.3851	0.3524 ± 0.2883	0.2381 ± 0.2025	130	62.80		
P2	5	1.7488 ± 0.4348	1.3947 ± 0.3356	0.3710 ± 0.2473	0.2419 ± 0.1751	155	74.88		
P3	3	1.5362 ± 0.4999	1.3386 ± 0.3673	0.2981 ± 0.2873	0.1998 ± 0.1981	111	53.62	0.2223	1.7494
P4	5	1.7391 ± 0.4402	1.4422 ± 0.3547	0.3937 ± 0.2607	0.2622 ± 0.1848	153	73.91		
P5	4	1.7729 ± 0.4199	1.5165 ± 0.3612	0.4391 ± 0.2591	0.2976 ± 0.1849	160	77.29		

Table 5: Parameters of genetic diversity ascertained from pooled data of RAPD, DAMD and ISSR markers among five populations of *Musa* sp.;N = sample Size , Na = observed number of alleles , Ne = effective number of alleles , h = Nei's gene diversity , I = Shannon's Information index , n = number of polymorphic loci , P(%) = percentage of polymorphism , G_{ST} = diversity among populations , Nm = Gene flow , SD = standard deviation.

P1 = North Garo Hills, P2 = West Garo Hills, P3 =South Garo Hills, P4 =East Garo Hills, P5= South- West Garo Hills

Population ID	P1	P2	P3	P4	P5
P1	****	0.8775	0.8252	0.8782	0.8956
P2	0.1307	****	0.8344	0.9402	0.9174
P3	0.1922	0.1810	****	0.8795	0.8643
P4	0.1299	0.0617	0.1284	****	0.9225
P5	0.1103	0.0862	0.1458	0.0806	****

Table 6: Nei's Unbiased Measures of Genetic Identity and Genetic distance.

Nei's genetic identity (above diagonal) and genetic distance (below diagonal); (P1 = North Garo Hills, P2 = West Garo Hills, P3 = South Garo Hills, P4 = East Garo Hills, P5= South- West Garo Hills).The bold values above diagonal which signifies Nei's Maximum (0.9402) and Minimum (0.8252) genetic identity, while the values in bold below the diagonal represent the Nei's Maximum (0.1922) and Minimum (0.0617) genetic distance between the five populations.

Source	Df	SS	MS	Est. Var.	%
Among Pops	4	159.460	39.865	2.322	7%
Within Pops	16	483.017	30.189	30.189	93%
Total	20	642.476		32.511	100%

Table 7: Analysis of Molecular Variance (AMOVA) based upon collective data of SPAR (RAPD, ISSR and DAMD) markers for five populations of *Musa* sp.

Figure 8: UPGMA (clustering based on Nei's (1978) unbiased genetic diversity) dendogram highlighting inter-relationship among 21 genotypes of *Musa* sp. from five districts of Garo Hills (P1 = West Garo Hills, P2 = South-West Garo Hills, P3 = East Garo Hills, P4 = South Garo Hills, P5= North Garo Hills), Meghalaya (India).

bination of all the markers in SPAR analysis might be considered to be more reliable and meaningful to understand the genetic variation in the genus *Musa.*

The dendrograms gave the information regarding the inter-relations and evolutions of the various collected genotypes. Unlike the cultivated *Musa* varieties, wild *Musa* species are predominantly outbreeders, which maybe responsible for the inter-mixing of genotypes from various populations as observed from the dendrograms in this study. This may also result in the increasing varia-

tions within the populations [8,15,34]. The dendrogram generated by SPAR data revealed that the genotype SWGH-ABb has the longest branch and thereby may be very recently evolved**.** In contrast, genotypes SGH-RBa and WGH-TGb appeared to have the shortest branches, indicating that they evolved very early in the evolutionary timeline. The considerable gene flow among the genotypes is further established by the high Nm value (1.7494) obtained from the population diversity data.

The population P5 (SWGH) appears to be distinct from the other populations as it has the highest number of polymorphic loci(n), and subsequently the highest percentage polymorphism of 77.29%. This distinction of P5 is also reinforced in the dendrograms generated by individual as well as SPAR approach. Population, P3 from SGH, is observed to have the lowest number of polymorphic loci and percentage polymorphism of 111 and 52.62, respectively. This is in tandem with the observations of ISSR and SPAR dendrograms which demonstrated the SGH population as eldest and ancestral in evolution. P5 can hence be utilized to devise breeding and crop improvement programs for *Musa* sp. In an earlier study, *Musa* genotypes were collected from various regions of Meghalaya and variations existant among genotypes *vis-à-vis* altitudes they inhabit [15]. In this study, however, the genotypes were collected from various areas of Garo Hills only and hence there was no significant variation in the altitudes of the collected samples.

All the parameters for measurement of population genetics indicated noteworthy genetic variation and gene flow. Gene flow estimates are classified as low (if Nm<1), moderate (if Nm>1) and high (if Nm>4) [13]. In this study, value of Nm was found to be 1.7494 (Nm>1), which suggests moderate gene flow among the populations of *Musa*. This appreciable rate of gene flow might be a reason for the high level of intermixing of the geno types. The AMOVA values generated revealed higher genetic variation within populations (97%), as compared to variations among the populations. This type of high variation within the populations have been widely reported in case of outbreeders [8,15,34]. This variation is further established by the G_{ST} value (0.2223), which is an important parameter for determination of genetic diversity among the populations. The PCoA analysis of molecular marker data also showed high intermixing of the genotypes (Figure 9).

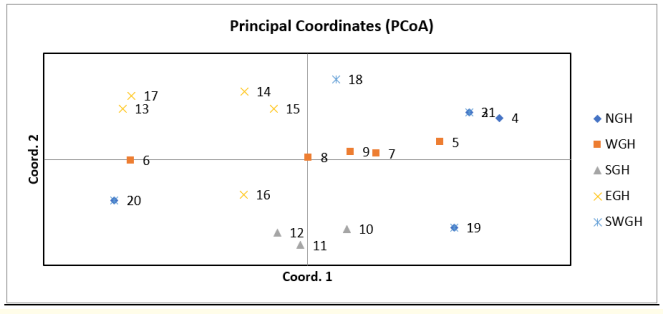

Figure 9: PCo Analysis of five populations comprising of 21 accessions of *Musa* sp. from five districts of Garo Hills (colour coded) generated by using the various physical parameters. The numbers indicate the collected samples (Table 1).

Conclusion

The study concludes that the data obtained from morphological and molecular studies, shows the prevalence of significant amount of variation among the various *Musa* genotypes of Garo Hills. The morphological parameters and the PCo analysis indicated variation as well as intermixing of the genotypes. The molecular markers effectively provided information regarding the genetic diversity, genetic identity and distance, and the phylogenetic relationships among the various collected genotypes of *Musa* sp. The concatenated data (SPAR) provided with a comprehensive and clearer image of the species inter-relationships and their diversity. The data generated can be used to devise effective breeding programs for preserving and improving the wild *Musa* genotypes which are available in the Garo hills of Meghalaya.

Author Contribution Statement

SRR and CPS planned and designed the research work. Survey and procurement of the plant materials were done by AH and PS. AH, SD and PS conducted the laboratory work. SD and SRR generated and analyzed the data. SD, AH, SRR and CPS interpreted the data and prepared the manuscript accordingly.

Acknowledgement

The authors extend their heartfelt gratitude to the Head, Department of Biotechnology and Bioinformatics, and Head, Department of Horticulture, North Eastern Hill University, Shillong for providing the facilities. The authors are thankful to Dr. Harish and Dr. Devendra Biswal, for help in finalising the manuscript. The authors are also thankful to the Ministry of Tribal Affairs, National Fellowship and Scholarship for higher studies of ST students (Awardee no.- 201718-NFST-MEG-00963) for funding the research work.

Declaration of Competing Interest

The authors declare no known competing interest.

Bibliography

1. Anonymous. "Descriptors for Banana (Musa spp.)". INIBAB Montpellier (1996).
2. Bhattacharya E., *et al*. "Single primer amplification reaction methods reveal exotic and indigenous mulberry varieties are similarly diverse". *Journal of Biosciences* 30.5 (2005): 669-677.
3. Brinemugha BE., *et al*. "Assessment of genetic diversity in plantain (Musa paradisiaca L.) using RAPD marker". *Nigerian Journal of Biotechnology* 35 (2018): 99-104.
4. Brown A., *et al*. "Bananas and Plantains (Musa spp.)". Springer (2017): 219-240.
5. Doyle JJ and Doyle JL. "Isolation of plant DNA from fresh tissue". *Focus* 12 (1990): 39-40.
6. Goswami B., *et al*. "Genetic diversity, population structure and gene flow pattern among populations of LasiurussindicusHenr.-an endemic, C4 grass of Indian Thar desert". *Plant Gene* 21 (2020): 100206.
7. Hazarika TK., *et al*. "Genetic variability and phylogenetic relationships studies of genus Citrus L. with the application of molecular markers". *Genetic Resources and Crop Evolution* 61 (2014): 1441-1454.
8. Hogbin PM., *et al*. "Evaluation of the contribution of genetic research to the management of the endangered plant Zieriaprostrata". *Conservation Biology* 13 (1999): 514-522.
9. Hoyos-Leyva JD., *et al*. "Physical, morphological characterization and evaluation of pasting curves of Musa spp". *Acta Agronomy* 61 (2012): 214-229.
10. Javed MA., *et al*. "Study of resistance of Musa acuminata to Fusariumoxysporum using RAPD markers". *Biologia Plantarum* 48 (2004): 93-99.
11. Kameswara RN. "Biotechnology for plant resources conservation and use". Principles of seed handling in Genebanks Training course, Kampla, Uganda (2004).
12. Khatri A., *et al*. "Use of RAPD for the assessment of genetic diversity among exotic and commercial banana clones". *Pakistan Journal of Botany* 41 (2009): 2995-2999.
13. Kumar S., *et al*. "Genetic diversity assessment of Jatrophacurcas L. germplasm from Northeast India". *Biomass Bioenergy* 35 (2011): 3063-3070.
14. Kumar A., *et al*. "Efficiency of ISSR and RAPD markers in genetic divergence analysis and conservation management of Justiciaadhatoda L., a medicinal plant". *Plant Systematics and Evolution* 300 (2014): 1409-1420.
15. Lamare A and Rao SR. "Efficacy of RAPD, ISSR and DAMD markers in assessment of genetic variability and population structure of wild Musa acuminata colla". *Physiology and Molecular Biology of Plants* 21 (2015): 349-358.
16. Lamare A., *et al*. "Analysis of intraspecific genetic variation in Musa balbisiana Colla from Meghalaya as revealed by Single Primer Amplification Reaction approach". *Nucleus* 59 (2016): 25-34.
17. Lu Y., *et al*. "Molecular assessment of genetic identity and genetic stability in banana cultivars (Musa spp.) from China using ISSR markers". *Australian Journal of Crop Science* 5 (2011): 25.
18. MØller AP and Merilä J. "Analysis and interpretation of long-term studies investigating responses to climate change". *Advances in Ecological Research* 35 (2004): 111-130.
19. Nair S., *et al*. "Genetic diversity analysis of Leptadeniapyrotechnica in Jodhpur region of India". *Gene Reports* 10 (2018): 157-161.
20. Ng WL and Tan SG. "Inter-simple sequence repeat (ISSR) markers: are we doing it right". *ASMS* 9 (2015): 30-39.
21. Perrier X. "DARwin software" (2006).
22. Pillay M., *et al*. "Analysis of genetic diversity and relationships in East African banana germplasm". *Theoretical and Applied Genetics* 102 (2001): 965-970.
23. Prevost A., *et al*. "A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars". *Theoretical and Applied Genetics* 98 (1999) 107-112.
24. Purseglove JW. "Monocotyledons (tropical crops S)". *Monocotyledons* (Tropical Crops S) (1972).
25. Reddy MP., *et al*. "Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding". *Euphytica* 128 (2002): 9-17.

26. Sangeeta S., *et al*. "Analysis of genetic diversity among papaya cultivars using Single Primer Amplification Reaction (SPAR) methods". *Journal of Horticultural Science and Biotechnology* 80 (2005): 291-296.

27. Sehgal D., *et al*. "Assaying polymorphism at DNA level for genetic diversity diagnostics of the safflower (Carthamustinctorius L.) world germplasm resources". *Genetica* 135 (2009): 457-470.

28. Sharma SK., *et al*. "Single primer amplification reaction (SPAR) reveals intra-specific natural variation in Prosopis cineraria (L.) Druce". *Trees* 24 (2010): 855-864.

29. Sharma SK., *et al*. "Single primer amplification reaction (SPAR) reveals inter-and intra-specific natural genetic variation in five species of Cymbidium (Orchidaceae)". *Gene* 483 (2011): 54-62.

30. Simmonds NW and Shepherd K. "The taxonomy and origins of the cultivated bananas". *Journal of the Linnean Society of London, Botany* 55 (1955) 302-312.

31. Srilekha V. "Genetic diversity and Molecular characterization of few citrus species in Visakhapatnam by RAPD markers". *Journal of Integral Sciences* (2018): 22-27.

32. Thorpe RS. "Geographic variation. The Oxford Encyclopedia of Evolution (2002).

33. Venkatachalam L., *et al*. "The use of genetic markers for detecting DNA polymorphism, genotype identification and phylogenetic relationships among banana cultivars". *Molecular Phylogenetics and Evolution* 47 (2008): 974-985.

34. Zawko G., *et al*. "Conservation genetics of the rare and endangered Leucopogon obtectus (Ericaceae)". *Molecular Ecology* 10 (2001): 2389-2396.