Journal of Plant Biochemistry and Biotechnology
https://doi.org/10.1007/s13562-019-00519-x

ORIGINAL ARTICLE

Inter retrotransposon based genetic diversity and phylogenetic analysis among the *Musa* germplasm accessions

Marimuthu Somasundaram Saraswathi[1] · Subbaraya Uma[1] · Sathasivam Ramaraj[1] · Palani Durai[1] · Mohamed M. Mustaffa[1] · Kalaimughilan Kalaiponmani[1] · Arumugam Chandrasekar[1]

Received: 27 April 2018 / Accepted: 5 June 2019

Abstract

In the present investigation, the insertional polymorphisms of retro-elements were studied in the *Musa* germplasm available at ICAR-NRCB field gene bank using IRAP markers. The maximum number of polymorphic bands were produced by the primer pair Nikita and LTR 6150 (48) followed by LTR 6149 and 3′LTR (47) and minimum of 35 bands were produced by the primer pair Sukkula and LTR 6150. The bands produced were scored as 0 (absent) and 1 (present) and the resultant binary data was subjected to diversity analysis. The dendrogram consisted of two major clusters with members of *Eumusa* and *Rhodochlamys* in one indicating their genetic closeness and members of the genus *Ensete* in another cluster. Results of principal coordinate analysis were congruent to those obtained in hierarchial cluster analysis. The molecular markers used in this study could reveal intra and inter-group diversity among the *Musa* germplasm accessions with similarity co-efficient ranging from 0.41 to 0.99. IRAP marker system has performed excellently clustering the accessions based on both genomic and subgroup levels. The entire germplasm was found to be robust with no duplications indicating the diverse group of accessions available at ICAR-NRCB field gene bank. It has also exhibited high polymorphism and hence could be effectively used to detect the genetic relatedness among diverse genome of *Musa*.

Keywords *Musa* · Molecular characterization · Genetic diversity · IRAP markers

Abbreviations

IRAP Inter retrotransposon amplified polymorphism
HCA Hierarchical cluster analysis
PCoA Principal co-ordinates analysis

Introduction

Banana and plantains are the vital source of food and fiber in many parts of the world and the major source of carbohydrates for more than 100 million people (Frison and Sharrock 1998). Currently, India is the largest producer of bananas in the world besides contributing greatly to the economy. Prevalence of varied agro-climatic regions has favored the development of numerous *Musa* varieties in India and therefore it has been recognized as one of the major centers of *Musa* origin and diversity. Genetic sorting

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13562-019-00519-x) contains supplementary material, which is available to authorized users.

✉ Marimuthu Somasundaram Saraswathi
saraswathimse@gmail.com

Subbaraya Uma
umabinit@yahoo.co.in

Sathasivam Ramaraj
ramaraj.newgene@gmail.com

Palani Durai
nkpdurai@yahoo.co.in

Mohamed M. Mustaffa
mmmustaffa@gmail.com

Kalaimughilan Kalaiponmani
kalaiponmani@gmail.com

Arumugam Chandrasekar
chandrugcg_77@yahoo.co.in

[1] ICAR - National Research Centre for Banana, Thogamalai Road, Thayanur Post, Tiruchchirappalli, Tamil Nadu 620 102, India

Published online: 01 July 2019

based on phenotypic data may not reflect the exact genetic diversity due to duplications and genotype x environmental interaction (Hu et al. 2000). Therefore, molecular markers which can analyse the genetic diversity without any influence of the environment (Ghislain et al. 1999) are needed as a complement for the development of a true core collection in banana. Further, robustness of the core collection is likely to increase the evaluation efficiency; reduce the frequency and cost for collection and maintenance (Greene and McFerson 1996).

Retrotransposon are major components of most eukaryotic genomes. They are known to insert themselves into the genome and act as mutagenic agents thereby providing a potential source of gene diversity (Harrison 2000). Thus, it provides a powerful molecular tool to make meaningful inferences about the genetic diversity of any crop. Among the transposable element based markers, IRAP (Inter Retrotransposon Amplified Polymorphism) stands out as a simple and efficient system, requiring only a simple PCR followed by electrophoresis to resolve the genetic diversity (de Queiroz et al. 2014). IRAP makes use of conserved retrotransposon sequences termed LTRs for detection of polymorphism (Kalendar et al. 1999). It is based on the amplification of regions between two neighboring retrotransposon. The polymorphism can thus be used for fingerprinting, diversity studies and linkage maps (Grzebelus 2006). They have been used by Teo et al. (2005) and Nair et al. (2005) only for the genomic classification of banana. Therefore, attempts have been made in the present study to assess the efficiency of IRAP markers in the management of ex situ collections and in addressing the genetic identification, redundancy, variation and phylogenetic relationships prevailing between commercial triploids and their wild relatives available in the *Musa* germplasm at ICAR-NRCB, Trichy. This would further assist in defining the core collection, designing appropriate conservation and breeding strategies.

Materials and methods

Plant materials

The fresh leaf samples (cigar leaf) of the banana accessions used in the study (Supplementary Table S1) were collected from *Musa* germplasm available in the field genebank of ICAR-NRCB, Trichy, India.

DNA extraction

The fresh cigar leaf samples were frozen with liquid nitrogen, ground to a fine powder and the genomic DNA was extracted following CTAB method (Gawel and Jarret 1991) with minor modifications. This genomic DNA was used for PCR amplification.

Inter retrotransposon amplified polymorphism (IRAP) analysis

Genetic diversity of germplasm accessions belonging to varied ploidies and genomic groups was assessed using Long Terminal Repeat (LTR) primers derived from the barley genome (*Hordeum vulgare*) (Kalendar et al. 1999, 2000; Boyko et al. 2002). Primer sequences and sources of retrotransposons are presented in Supplementary Table S2. Amplification was carried out in 20 µl reaction mixture containing 50 ng of genomic DNA, 1X PCR buffer (Genei, India), 5 pmol of each primer, 200 µM each of four dNTP and 1.5 U of Taq polymerase (Genei, India). PCR program consisted of an initial denaturation at 94 °C for 5 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at specific temperature optimized (which varied with the primer combination) for 1 min, followed by extension at 72 °C for 1 min terminated by a final extension 72 °C for 15 min followed by incubation at 4 °C.

Gel electrophoresis and photography

The amplified products of IRAP-PCR were fractionated by gel electrophoresis on 1.5% (w/v) agarose gels using 1X TAE buffer (40 mM Tris–acetate, 1 mM EDTA, pH 8.0) at 75 V. The banding patterns and molecular weight of the amplified fragments were documented in Alpha Innotech Image Analyser (Alpha Innotech Corporation, California, U.S.A).

Data analysis

PCR amplified products were scored as dominant markers i.e., as presence (1) or absence of bands (0). The results were analysed using NTYSYS 2.01i. The genetic similarity (GS) matrix was then used in hierarchical cluster analysis using the unweighted pair group method with arithmetic averages (UPGMA) and sequential agglomerative hierarchical and nested (SAHN) clustering methods (NTSYS statistical package, Rohlf 1990) to produce a dendrogram. The reliability and robustness of the dendrogram was tested by bootstrap analysis using DARwin software with one thousand replicates. Molecular data was subjected to PCoA in NTYSYS 2.01i in order to reconfirm the diversity patterns brought out by cluster analysis. Analysis of molecular variance (AMOVA) was performed to determine the percentage share of components of genetic variation. The number of observed alleles (Na), the number of effective alleles (Ne) (Kimura and Crow 1964), mean Nei's genetic diversity index (He) (Nei 1973), and Shannon diversity

index (I) (Lewontin 1972) were determined using Popgene 1.32 (Yeh et al. 1999). Estimate of gene flow (Nm) was calculated as Nm = 0.5 × (1 − Gst)/Gst (McDermott and McDonald 1993), where Gst is the gene differentiation.

Results and discussion

The present day banana breeding efforts are more focused towards the development of superior tetraploids involving commercial triploids and diploid genotypes with favorable traits like disease resistance. Hence establishment of genetic relationships between triploids and diploids will be very much useful in banana breeding programmes and the same has been accomplished in the present study using IRAP markers.

Optimization of PCR conditions for IRAP analysis

The gradient PCR program was adopted to optimize the annealing temperature for various primer combinations and the optimized annealing temperatures are provided in Table 1. The observed products ranged from 300 to 3000 bp.

IRAP diversity

All the primer combinations amplified products resulting in discrete, repeatable amplicons which were considered for the genetic diversity and phylogenetic analysis. The polymorphism exhibited by all the primer combinations was 100% which indicated substantial variation at the DNA level among the 303 test accessions. The number of bands visible reflected the number of different retro-elements situated close enough to one another to allow amplification (Fig. 1). The number of polymorphic bands produced by the primer pair Nikita and LTR 6150 was maximum (48) followed by LTR 6149 and 3′LTR (47) and minimum (35) bands were produced by Sukkula and LTR 6150. The possible reason for the variation could be appearance of new forms of alleles of increased or decreased size due to in-del of the repeat units. This is the first report of IRAP markers being used to investigate the genetic relatedness of unknown *Musa* spp. with respect to the known *Musa* spp.

Population structure and AMOVA

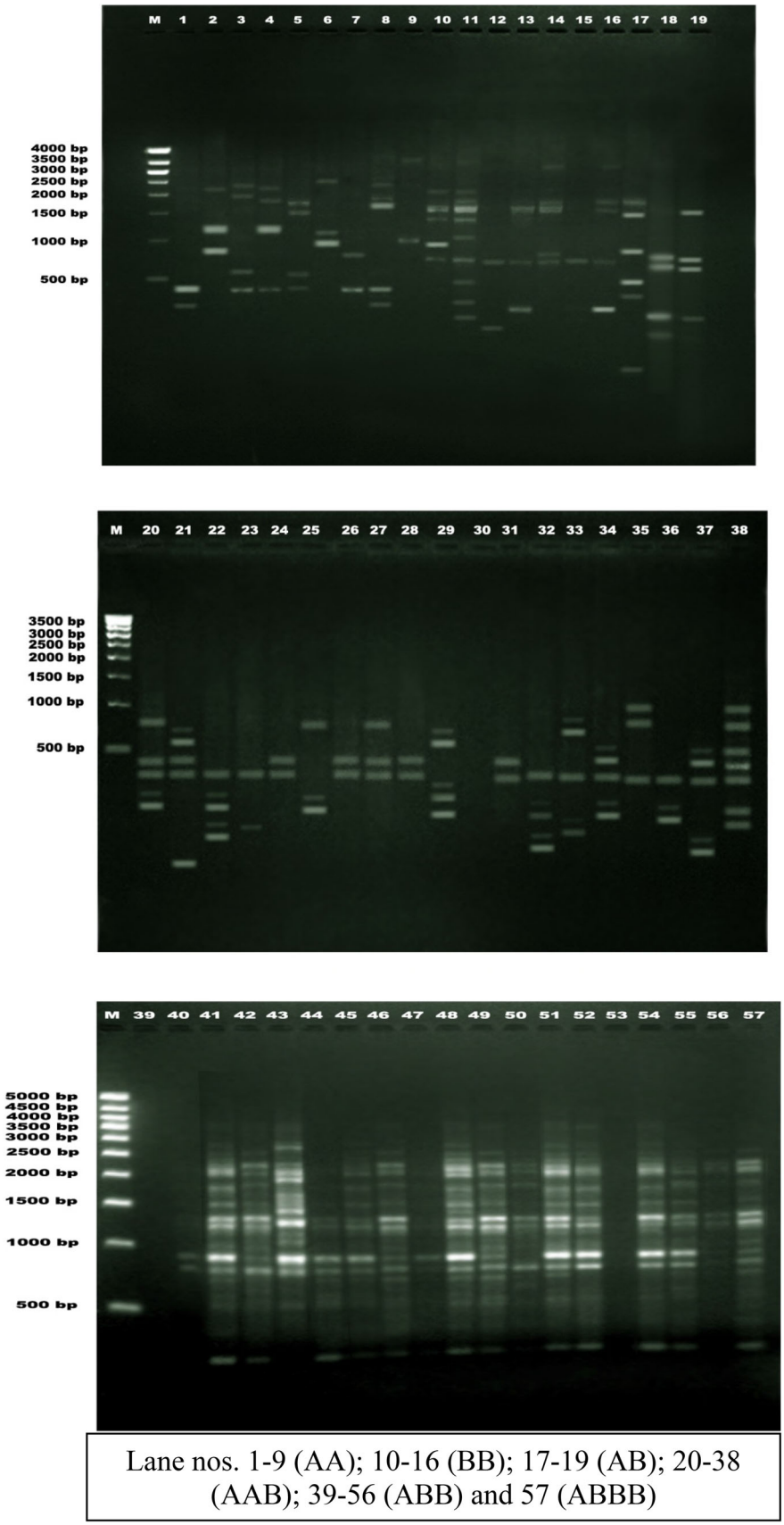
The Nei's gene diversity (*h*) ranged from 0.01 to 0.25 among various genomic groups. The maximum gene diversity was in ABBB (0.25) followed by ABB and AB (0.22). The highest level of genetic diversity detected in ABBB, ABB and AB might be due to their inter-sub-specific hybrid origin with limited fertility and heterozygosity for chromosome structural abnormalities maintained by vegetative propagation (Creste et al. 2003). AAB (1.90) had the highest number of observed alleles (*na*), while ABBB (1.44) had the highest number of effective alleles (*ne*) and Shannon's information index (*I*) (0.38) (Table 2). The coefficient of genetic differentiation (G_{ST}) was 0.429 among the 303 *Musa* cultivars showing a high level of overall difference among the population. This indicated that maximum variation is within the population and minimum among populations (Lamare and Rao 2015) which is quite important for their adaptation to changing environment and long term survival of crop species (Sarabhai et al. 2016). The level of gene flow (*Nm*), which refers to the number of individuals entering into the population in a generation, was found to be 0.23 which is low in nature as per the scheme followed by Kumar et al. (2014). This might be due to low level of gene flow owing to population isolation, lack of suitable pollinator for long distance, terrain nature of the region posing hurdles for seed/pollen dispersal (Slatkin 1987; Padmesh et al. 2012). The results of AMOVA confirmed that there was only 19% variation between the various genomic groups and 81% variation within the various genomic groups.

Cluster analysis of the test accessions

All the banana accessions (303) were characterized using IRAP markers. This divided the banana accessions into two major clusters and the cluster composition is given in Table 3. Cluster I included two sections of the genus *Musa* namely *Eumusa* and *Rhodochlamys* while cluster II consisted of the genus *Ensete* (Supplementary Figure 1). The

Table 1 IRAP primer combination and their annealing temperatures

Primers	Annealing temperature (°C)	No. of polymorphic bands produced
Sukkula and LTR 6149	45.5	37
Sukkula and LTR 6150	48.4	35
Sukkula and 5′LTR2	43.2	38
LTR 6149 and Nikita	45.5	36
LTR 6149 and 3′LTR	45.5	47
Nikita and LTR 6150	45.5	48



Lane nos. 1-9 (AA); 10-16 (BB); 17-19 (AB); 20-38 (AAB); 39-56 (ABB) and 57 (ABBB)

Fig. 1 Polymorphism patterns for 57 *Musa* accessions belonging to various genomic groups generated by IRAP primer pair Sukkula + LTR 6149

Table 2 Genetic diversity parameters within each of the genomic groups of genera *Musa* and *Ensete*

Population	N	na	ne	He	i
AA	23	1.77	1.25	0.17	0.29
BB	21	1.62	1.22	0.14	0.16
AB	24	1.83	1.34	0.22	0.35
AAA	23	1.76	1.29	0.19	0.30
AAB	93	1.90	1.31	0.21	0.34
ABB	100	1.89	1.33	0.22	0.35
AAAB	4	1.38	1.22	0.13	0.20
ABBB	7	1.68	1.44	0.25	0.38
Ensete	3	1.02	1.02	0.01	0.01
Rhodochlamys	4	1.12	1.09	0.05	0.07

na observed number of alleles, *ne* effective number of alleles (Kimura and Crow 1964), *He* Nei's (1973) gene diversity, *i* Shannon's information index (Lewontin 1972)

genotypes were mixed up in all clusters except for few cases where clusters were dominated by genotypes of a particular genome or geographical origin as reported by Changadeya et al. (2012). Lack of clustering based on geographical origin could also be due to introductions followed by naturalization of cultivars in areas away from their initial origin. No specific correlation could be established between genetic similarities and geographical origin which reflects the human intervention in the dispersion of *Musa* accessions (Creste et al. 2004). The dendrogram obtained has proved the genetic closeness of the two sections namely *Eumusa* and *Rhodochlamys.* Basically these two sections have the same basic chromosome number 11 and the present results corroborate with the earlier phylogenetic studies of Simmonds and Shepherd, (1955), Wong et al. (2002), Doležel and Bartos, (2005), Liu et al. (2010), Li et al. (2010), Christelova et al. (2011). In the present study, three accessions of the genus *Ensete* have grouped together in cluster II and proved their distinctiveness from others.

Cluster I

Cluster I had two subclusters namely Ia (16 microclusters) and Ib (two microclusters) with 58% similarities.

Sub cluster 1a Microcluster 1 of sub cluster Ia consisted of nine *Musa acuminata* diploid land races in one cluster with 90% similarities ($P = 96\%$). Parthenocarpic *Musa acuminata* landraces viz, Matti (0182), Namarai (0185), Sanna Chenkadali (0201), Anaikomban (0208) and Siguzani (1836) also clustered with *M. acuminata* ssp. *burmannica* and *M. acuminata* ssp. *burmaniccoides* with 92 and 91% similarities respectively. The south Indian origin of the landraces and wild species might have contributed for their close genetic proximity. Matti is phenotypically closer to *M. acuminata* ssp. *burmannica* and has assumed commercial significance as therapeutic banana (Uma et al. 2014). The present results are in conformity with the earlier phylogenetic studies of Durai et al. (2013) using SSR markers suggesting that Matti might have evolved from *M. acuminata* ssp. *burmannica.* The common geographical origin of Chengdawt (1030) and *M. acuminata* (Assam wild, 1712) i.e., Assam, could be one of the reasons for $> 88\%$ similarities registered between them. Similarly, eleven *Musa balbisiana* types grouped with more than 92% similarities. Bhimkol and Athiakol grouped with more than 95% similarities ($P = 53\%$) as reported earlier by Uma et al. (2006) however *M. balbisiana* (1353) type from Andaman and Nicobar islands grouped separately. This indicates that geographical origin could be one of the key factors influencing the genetic makeup. This is in line with the earlier reports of Uma et al. (2006) in their diversity studies on *M. balbisiana* using RAPD markers. Three AB diploids viz., Aktoman (0053), Adukkan (0114) and Kodappanillakunnan (0174) clustered with members of BB with 90% similarities. Genetic closeness of AB genotypes to *M. balbisiana* indicates their possible ancestry during the evolution of AB/BA genome. This is contradicting the results of Ravishankar et al. (2012) who reported that AB cultivars are more closely related to AA than BB genotypes using SSR markers.

Despite their different subgroups, AAA types like Manjahaji (0017), Tulsimanohar (0039) and Thellachakkarakeli (0166) ($P = 56\%$) grouped together in *microcluster 2 of subcluster Ia* and other 19 accessions belonging to varied subgroups like Mysore, Plantain, Silk and Pome grouped with 92–98% similarities owing to their same genomic group "AAB". In *microcluster 3 of sub cluster Ia* despite their different geographical origins, ploidy and sub groups, Singhalaji (0137) and FHIA-02 (1673) grouped together with more than 90% similarities. This might because of a common progenitor either *M. acuminata* or *M. balbisiana.* However this needs reconfirmation. *Microcluster 4 of sub cluster Ia* had members of Pome, Silk and Mysore subgroups grouped separately indicating that IRAP markers were able to cluster the accessions based on their subgroup as reported earlier. Morphologically Ladan small and Padathi (0537) are very similar except for blunt fruit tip and the present study also indicated that they are genetically closer with 96% similarities. Indian hybrids, H-1 (0209) and H-2 (0210) released by KAU, clustered together with 95% similarities ($P = 64\%$). Further these two hybrids shared 96% similarities despite their different subgroups namely Pome (H1) and Mysore (H2) respectively and this could be attributed to their common male parent Pisang Lilin (AA).

Table 3 Cluster composition

Cluster number	Section	Subcluster number	Microcluster members
Ia	Eumusa and Rhodochlamys	1	0182 (AA), 0185 (AA), 0201 (AA), 0208 (AA), 1836 (AA), 1631 (AA), 0642 (AA), 1030 (AA), 1712 (AA), 0007 (BB), 0011 (BB), 0167 (BB), 1168 (BB), 1184 (BB), 0446 (BB), 0444 (BB), 1914 (BB), 0508 (BB), 1186 (BB), 1353 (BB), 0053 (AB), 0114 (AB), 0174 (AB)
		2	0017 (AAA), 0039 (AAA), 0166 (AAA), 0110 (AAB), 0731 (AAB), 0025 (AAB), 0138 (AAB), 0045 (AAB), 0100 (AAB), 0118 (AAB), 0701 (AAB), 0001 (AAB), 0355 (AAB), 0445 (AAB), 0447 (AAB), 0618 (AAB), 0312 (AAB), 0501 (AAB), 0125 (AAB), 0128 (AAB), 0207 (AAB), 0733 (AAB)
		3	1673 (AAAB), 0137 (ABB)
		4	0006 (AAB), 0365 (AAB), 0008 (AAB), 0798 (AAB), 0043 (AAB), 0048 (AAB), 0537 (AAB), 0036 (AAB), 0202 (AAB), 0240 (AAB), 0241 (AAB), 0209 (AAB), 0210 (AAB), 0376 (AAB)
		5	0004 (ABB), 0026 (ABB), 0094 (ABB), 0034 (ABB), 0085 (ABB), 0086 (ABB), 0087 (ABB), 0103 (ABB), 0106 (ABB), 0493 (ABB), 0427 (ABB), 0795 (ABB), 0050 (ABB), 0231 (ABB), 0063 (ABB), 0102 (ABB), 0265 (ABB),0189 (ABB), 0228 (ABB)
		6	1621 (AAA), 1683 (AA), 0579 (AAB), 0637 (AAB), 0635 (AA), 0398 (ABB), 0634 (ABB)
		7	1717 (AA), 1715 (AA), 1710 (AA), 0555 (BB), 0597 (BB), 0113 (AB), 0623 (AB), 0717 (AB), 0388 (AB), 0012 (AAA), 1653 (AAA), 0071 (AAA), 0161 (AAA), 0111 (AAA), 0081 (AAA), 0500 (AAA), 0612 (AAA)
		8	0064 (AA), 1183 (AA), 1731 (AA), 1019 (AA), Andaman (BB), 0018 (BB), 0047 (BB), 1182 (BB), 0449 (BB), 0234 (AB), 0369 (AB), 0186 (AB), 0482 (AB), 1913 (BB), 1912 (BB)
		9	0067 (BB), 0204 (ABB), 0117 (ABB), 0803 (ABB), 0415 (ABB), 0492 (ABB), 0188 (AB), 0361 (AB), 0737 (AB), 0177 (AAB), 0078 (AAB), 0519(AAB), 0619 (AAB)
		10	0165 (AAA), 0370 (AAA), 0580 (AAA), 0200 (AAA), 1419 (AAA), 1065 (AAA), 0057 (AAB), 0403 (ABB), 0700 (AAB), 0023 (AAB), 0190 (AAB), 0242 (AAB), 0489 (AAB), 0499 (AAB), 0160 (AAB), 0015 (AAB), 0192 (AAB), 0294 (AAB)
		11	0626 (AAAB), 0483 (AAB), 0059 (ABB), 0089 (ABB), 0354 (ABB), 0421 (ABB), 0453 (ABB), 0732 (ABB), 0411 (ABB), 0016 (ABB), 0163 (ABB), 0429 (ABB), 0112 (ABB), 0090 (ABB), 0435 (ABB), 0479 (ABB)
		12	0215 (AAB), 0290 (AAB), 0735 (AAB), 0028 (AAB), 0495 (AAB), 0364 (AAB), 0296 (AAB), 0615 (AAB), 0702 (AAB), 0197 (AAB), 0491 (AAB), 0297 (AAB), 0367 (AAB), 0122 (AAB), 0428 (AAB), 1005 (AAB), 0211 (AAB), 0362 (AAB)
		13	0164 (AAB), 0191 (AAB), 0256 (AAB), 0107 (AB), 0153 (AB), 0699 (AB), 0346 (ABB), 0799 (ABB), 0344 (ABB), 0129 (ABB), 0338 (ABB), 0804 (ABB), 0494 (ABB), 0065 (ABB), 0132 (ABB), 0511 (AB), 0178 (AB), 0458 (AB)
		14	0084 (ABB), 0288 (ABB), 0552 (ABB), 0091 (ABB), 0124 (ABB), 0518 (ABB), 0251 (ABB), 0734 (ABB), 0227 (ABB), 0291 (ABB), 0173 (ABB), 0736 (ABB), 0193 (ABB), 0481 (ABB), 0347 (ABB), 0794 (ABB), 0442 (ABB), 0088 (ABB)
		15	0639 (AAA), 0641 (AA), 0640 (AA), 0646 (AAB), 0656 (AA), 0723 (AAA), 0628 (AAAA), 1642 (AAAB), 1670 (AAAB)
		16	0121 (ABB), 0097 (ABB), 0530 (ABB), 0221 (ABB), 0130 (ABB), 0553 (ABB), 0413 (ABB), 1260 (*M.ornata*), 1745 (*M.velutina* hybrid), 1718 (*M.velutina*), 1376 (*M.laterita*), 0134 (ABB), 0171 (ABB), 0271 (ABB), 0366 (ABB), 0246 (ABB), 0247 (ABB), 0450 (ABB), 0426 (ABB), 0430 (ABB), 0525 (ABB), 0490 (ABB)
		17	0253 (ABBB), 0471 (ABBB), 0254 (ABBB), 0400 (ABBB)
Ib	Eumusa	1	0021 (AA), 2074 (AA), 0009 (AAA), 0032 (AAA), 0147 (AB), 0439 (AB), 0486 (AB), 0387 (AAB), 0212 (AAB), 0488 (AAB), 0093 (AAB), 0614 (AAB), 0243 (AAB), 0275 (AAB), 0397 (AAB), 1066 (AAB), 0360 (AAB), 0410 (AAB), 0497 (AAB), 0554 (AAB), 0135 (AAB), 0721 (AAB), 0187 (AAB), 0249 (AAB), 0259 (AAB), 0280 (AAB), 0374 (AAB), 0014 (AAB), 0712 (AAB)
		2	0333 (ABB), 0416 (ABB), 0172 (ABB), 0123 (ABB), 0096 (ABB), 0584 (ABB), 0079 (ABB), 0075 (ABBB), 0157 (ABB), 0417 (ABB), 0293 (ABB), 0115 (ABB), 0538 (ABB), 0055 (ABB), 0407 (ABBB), 0390 (ABBB)
II	*Ensete*	1	1915 (*E. superbum*-dwarf), 1267 (*E. superbum*), 1018 (*E. glaucum*)

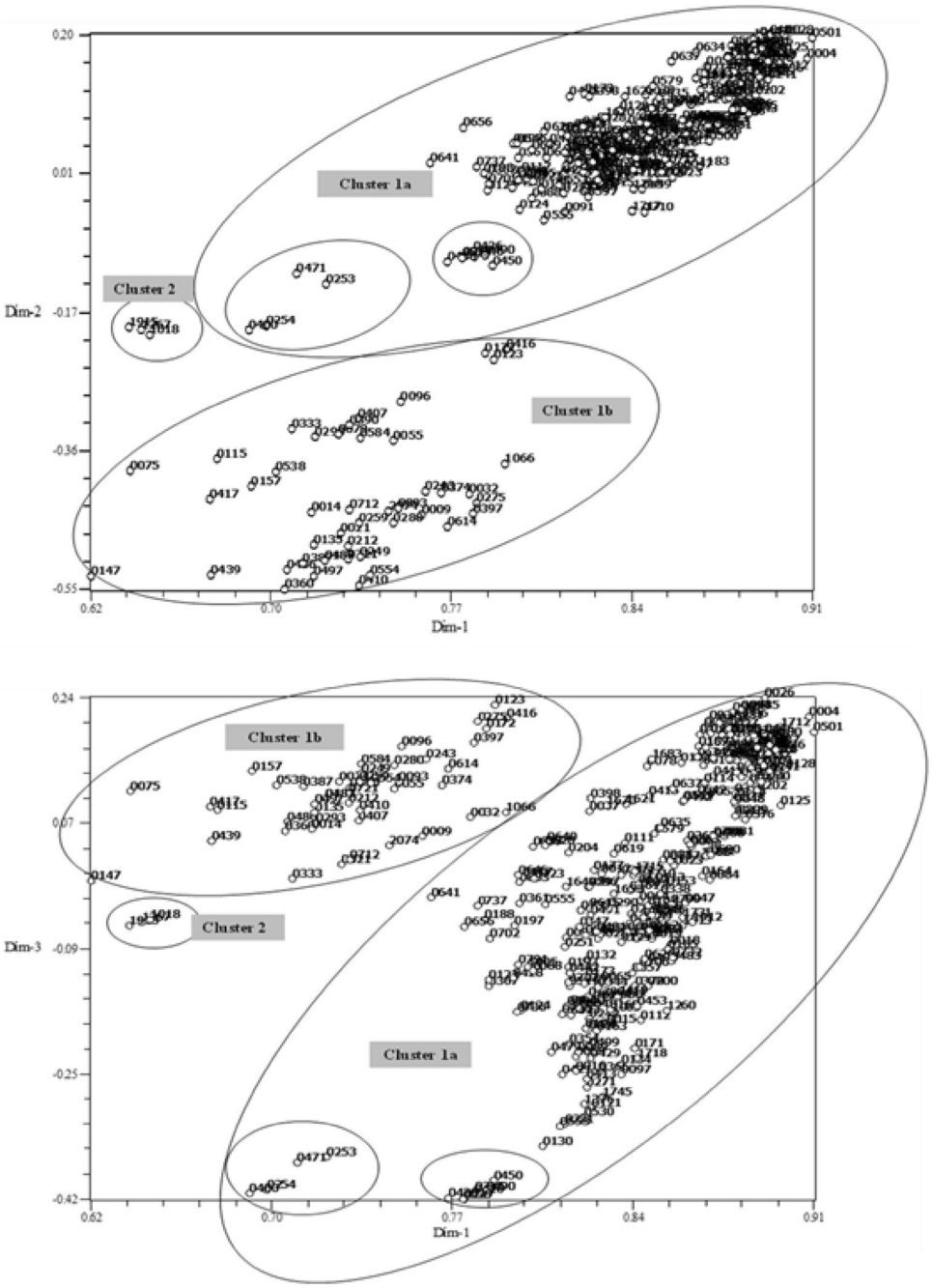


Fig. 2 Two dimensional PCoA based on IRAP data

In microcluster 6 of sub cluster Ia cv.Rose (1621) grouped with Gros Michel (1683), suggesting that one common *M. acuminata* subspecies might have contributed for their evolution. Three AAB accessions namely Pisang Seribu (0579), Pisang Nangka (0637) and Pisang Mas (0635) grouped together with 88% similarities (P = 96%). All the three accessions are basically from Malaysia and resistant to Sigatoka leaf spot diseases which might have contributed for this clustering pattern. Although ABB accessions namely Alukhel (0398) and Burro Cemsa (0634) grouped in this microcluster (P = 98%), they have proved their difference in genomic group by exhibiting 15% dissimilarities with other members. *Microcluster 7 of sub cluster* Ia consisted of three newly collected *M. acuminata* wild forms namely Bhalukpong wild (1717), *M.acuminata*-Assam (1715) and another *M. acuminata* type (1710 labeled as GP-15) which shared about 90% similarities (P = 100%). These three wild accessions have been collected from north eastern India and this suggested that the same sub sp. of *M. acuminata* might have contributed for their evolution as confirmed through morpho-taxonomic characterization (Durai et al. 2013). Although

morphologically they exhibited similarities, 10% dissimilarities have been recorded at the genotypic level indicating that there is no strong correlation between genetic dissimilarity and phenotypic resemblances (Ude et al. 2003; Crouch et al. 1999). Among the Cavendish clones, the medium tall types namely Harichal (0081), Peddapacha Aratti (0500) and Shrimanthi (0612) clustered together as reported by Saraswathi et al. (2011) in her previous study using IRAP markers. Elakki Bale (0113) of Karnataka, Ney Poovan (0623) of Tamil Nadu and Rasakadali (0717) of Kerala have been recorded as synonyms by Rekha et al. (2001) and Uma and Sathiamoorthy (2002). The present study has also confirmed their synonymity, by grouping in one cluster with 96% similarities (P = 88%). AB members namely Valiya Kunnan (0386) and Ney Poovan (P = 99%) clustered with *M. balbisiana* clones like Bhimkol (0597) and Elavazhai (0555). The unique features of *M. balbisiana* clones like robustness and sturdy stature might have contributed for the evolution of Valiya Kunnan having sturdy petiole unlike other AB's and their resistance to leaf spot diseases. Grouping pattern of *microcluster 8 of sub cluster Ia* was mainly based on their genomic group. Four AB accessions namely Valiya Kunnan (0234), Narmine (0369), Nattu Poovan (0186) and Padali Moongil (0482) grouped together (P = 97%).

Microcluster 9 of sub cluster Ia clustering was mainly based on the ploidy and genomic composition. Manguthamang (0067) which was morphotaxonomically classified earlier under "BB" type, has grouped with ABB members. Similar results have been obtained through flow cytometry (Uma et al. 2005) and molecular analysis. However, this result needs further confirmation.

Among the six *Musa acuminata* triploids which clustered together in *microcluster 10 of sub cluster Ia*, four are Cavendish members namely Dwarf Cavendish (0165), Gandevi selection (0370), Grand Naine (0580) and Robusta (0200) which are genetically closer while Manoranjitham (1419) and Pacha Kappa (1065) have distanced from other AAA members proving their morphological uniqueness. There are 11 bispecific triploid (AAB) cultivars belonging to Silk, Pome and Mysore sub groups which have clustered in one group with minimum differences based on their sub group. Only one ABB triploid accession Bainsa has grouped with Silk subgroup which needs further investigation. The possible reason for clustering of AAB (Chinali) and AAAB (FHIA 01) and 14 ABB accessions with 81% similarities (P = 98%) as a *microcluster 11 of 1a* could be due to the fact that all of them produce dessert bananas though they belonged to varied genomic and sub groups.

Eighteen accessions belonging to sub groups viz, Pome, Mysore, Silk and Plantain grouped in *microcluster 12 of sub cluster* Ia with 81% similarities (P = 100%). Members of the Silk subgroup grouped together as reported previously by Uma et al. (2004) and grouping of Pome and Mysore subgroups in one cluster is in conformity with the findings of Rekha et al. (2001). Further plantain types clustered together with P = 71%. Similar grouping was also reported by Carreel et al. (2002) based on RFLP studies using chloroplast and mitochondrial DNA. In *microcluster 13 of sub cluster Ia* Nendra Kunnan (0107), Agnishwar (0153) and Poovilla Chundan (0699) joined with ABB genotypes with more than 90% similarities, which indicate that these particular AB and ABB genotypes might have common ancestry as reported earlier by Onyago et al. (2010). Eighteen accessions belonging to Monthan, Bluggoe, Bontha, Pisang Awak and Peyan sub groups clustered together in this microcluster with 80% similarities. All the Pisang Awak members clustered together. Peyan is a unique member grouped with members of Pisang Awak and exhibited its closeness with Pisang Awak types. Irrespective of the morphological variations, members of Monthan, Bluggoe and Bontha sub groups clustered together as reported earlier by Uma et al. (2005). This is in conformity with the earlier reports that phenotypical variations need not always be expressed at the genotypic level (Saraswathi 2007). *Microcluster 14 of sub cluster Ia* had AA members of diverse geographic origin, Asia and Africa suggesting the possible gene flow across the centres of origin. It also comprises of three FHIA hybrids namely FHIA 18, 21 and 23 belonging to Pome, Plantain and Cavendish subgroups respectively. No relationship could be established between the AAs and FHIA hybrids of this cluster since the ancestry of pollen parents used in FHIA breeding programme is not very clear. However, this clustering needs reconfirmation.

Microcluster 15 of sub cluster Ia included 18 accessions of *Eumusa* all of which belong to ABB and 4 accessions of Rhodochlamys viz., *M. ornata*, *M. velutina, M. velutina* natural variant and *M. laterita* (P = 98%). Grouping pattern proved that sections, Eumusa and Rhodochlamys are genetically closer and share the same basic chromosome number 11. Similar results have been reported by Wong et al. (2002) who emphasized their grouping in a common section. There are four tetraploids grouped in *microcluster 16 of sub cluster Ia* with 85% similarities (P = 100%) which might be attributed to their same genomic group (ABBB) and ploidy nature (tetraploidy). Morphotaxonomic characterization has inferred Neyvannan Sawai and Hybrid Sawai to be synonyms while IRAP was able to distinguish them by exhibiting 1–2% dissimilarities. Similarly it was able to discriminate Klue Teparod and Foconah by exhibiting 10% dissimilarities. IRAP clustered Sawai separately in cluster Ib despite its morphological closeness to Neyvannan Sawai and Hybrid Sawai. The possible assumption is that they are the natural hybrids of Sawai and IRAP could distinguish the parents and their derivatives.

Sub cluster Ib

The evolution of cultivated bananas is not simply by hybridization between *M. acuminata* and *M. balbisiana* (Valmayor et al. 2000). But occasional occurrence of meiosis, recombination and fertilization are also the factors leading to diversity in cultivated bananas (Jesus et al. 2013). This sub cluster included 29 accessions belonging to the Pome and Silk subgroups sharing 87% similarities. Similarly all members of plantain subgroup namely Mysore Eathen (0387) and Attunendran (0488) grouped together with $P = 100\%$. Pure *M. acuminata* clustered together ($P = 100\%$) while bispecific diploids AB clustered together with a bootstrap value of $P = 100\%$. Hatidat, one of the AA diploid land races clustered with other *acuminata* triploids.

Microcluster 2 of sub cluster Ib consisted of 16 accessions of "B" genome rich triploids and tetraploids ($P = 53\%$). Triploids formed two clusters ($P = 100\%$) based on their sub groups viz, cooking types and dessert types (Pisang Awak). Among the four tetraploids tested, Bhat Manohar, a natural tetraploid clustered with Pisang Awak sub-group which included Karpuravalli types with boot strap value of $P = 95\%$. Morphotaxonomical classification also placed this accession under Pisang Awak subgroup (Uma and Sathiamoorthy 2002). Interestingly all the cooking types clustered together irrespective of their ploidy.

Cluster II

Cluster II included 3 members of *Ensete* sharing 98% similarities ($P = 100\%$). Both species namely *Ensete superbum* and *Ensete glaucum* clustered separately and they shared 53% similarity with *Eumusa* and *Rhodochlamys* members in cluster I. This indicated that *Ensete* differed from *Musa* both phenotypically and genotypically as reported earlier by Ude et al. (2002).

Principal coordinates analysis

To visualize the similarity or dissimilarity among groups or individual genotypes, principal coordinate analysis (PCoA) was done using *NTYSYS 2.01i* (Fig. 2). Members of *Ensete* and four tetraploids with ABBB genome formed a distinct cluster. Seven accessions of the genome ABB formed a distinct cluster and it was genetically closer to ABBB. The first three coordinates obtained during PCoA, explained 25.63% of the total variation. The first axis (eigen vector) displayed 12.03% of the variance, while the second and the third axes showed 7.85% and 5.73% of the variance respectively.

The PCoA further confirmed the positions and grouping of genotypes. Similar to the results of hierarchical cluster analysis (HCA), *Ensete* stood as a separate cluster and *Eumusa* and *Rhodochlamys* members intermingled with each other indicating their genetic closeness as reported by Wong et al. (2002) and Durai et al. (2011). Though both the sections are phenotypically far apart, their common chromosome number (n = 11) might have clustered them together (Durai et al. 2011). The present study also confirmed that *M. acuminata* and *M. balbisiana* are the progenitors of the present day bananas. In the present study, seven cooking bananas clustered separately while the rest were scattered and found in subclusters 1a and 1b. Silk and Pome (AAB) which are relatively starchy grouped closer to ABB members in subcluster 1b as reported by Osuji et al. (1997).

IRAP, though a dominant marker system, performed excellently both at genomic and subgroup levels with no synonyms indicating the robustness of the *Musa* collection at NRCB, Trichy. The current study placed sections *Eumusa* and *Rhodochlamys* in one cluster proving their genetic closeness and *Ensete in a* separate cluster. They also exhibited high polymorphism indicating their potential in detecting the genetic relatedness among diverse genome components of *Musa*. Results of PCoA were congruent to those obtained in HCA and morphotaxonomic characterization.

Acknowledgements We would like to thank the Indian Council of Agricultural Research (ICAR) for the financial support, Director, ICAR - NRCB, Trichy for the support and infrastructure facilities provided for the above study. Thanks to Mr. K. P. Sajith, Ph.D. scholar for the Adobe Photoshop work.

Compliance with ethical standards

Conflict of interest The authors declare that there are no conflicts of interest.

References

Boyko E, Kalendar R, Korzun V, Gill B, Schulman AH (2002) Combined mapping of *Aegilops tauschii* by retrotransposon, microsatellite and gene markers. Plant Mol Biol 48:767–790

Carreel F, Gonzalez de Leon D, Lagoda P, Lanaud C, Jenny C, Horry J, Tezenas du Montcel H (2002) Ascertaining maternal and paternal lineage within Musa by chloroplast and mitochondrial DNA RFLP analyses. Genome 45:679–692. https://doi.org/10.1139/g02-033

Changadeya W, Kaunda E, Ambali AJD (2012) Molecular characterization of Musa L. cultivars cultivated in Malawi using microsatellite markers. Afr J Biotechnol 11(18):4140–4157. https://doi.org/10.5897/ajb10.1738

Christelova P, Valarik M, Hribova E, Langhe ED, Dolezel J (2011) A multigene sequence-based phylogeny of the Musaceae (banana)

family. BMC Evol Biol 11:103. https://doi.org/10.1186/1471-2148-11-103

Creste S, Tulmann Neto A, Silva SO, Figueira A (2003) Genetic characterization of banana cultivars (Musa spp) from Brazil using microsatellite markers. Euphytica 132:259–268. https://doi.org/10.1023/A:1025047421843

Creste S, Tulmann-Neto A, Vencovsky R, Silva S, Figueira A (2004) Genetic diversity of Musa diploid and triploid accessions from the Brazillian banana breeding program estimated by microsatellite markers. Genet Resour Crop Evol 51:723–733. https://doi.org/10.1023/B:GRES.0000034578.37951.c4

Crouch JH, Crouch HK, Constandt H, Van Gysel A, Breyne P, Montagu MV, Jarret RL, Ortiz R (1999) Comparison of PCR-based molecular marker analyses of Musa breeding populations. Mol Breed 5:233–244. https://doi.org/10.1023/A:1009649521009

de Queiroz CB, Santana MF, Da Silva GF, Mizubuti ESG, de Araujo EF, de Queiroz MV (2014) Use of the IRAP marker to study genetic variability in *Pseudocercospora fijiensis* populations. Curr Microbiol 68:358–364. https://doi.org/10.1007/s00284-013-0454-y

Doležel J, Bartos J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot 95:99–110. https://doi.org/10.1093/aob/mci005

Durai P, Uma S, Saraswathi MS, Jayabalan N, Mustaffa MM (2011) Intersectional relationship between Eumusa and Rhodochlamys of the genus Musa using morphotaxonamy and microsatellite markers. Acta Hortic (ISHS) 897:267–270. https://doi.org/10.17660/ActaHortic.2011.897.31

Durai P, Uma S, Saraswathi MS, Jayabalan N, Mustaffa MM (2013) Assessment of phylogenetic lineage of landraces (AA) and wild *Musa acuminata* Colla. through morphotaxonomic traits and microsatellite markers. Indian J Hortic 70(4):463–468

Frison E, Sharrock S (1998) The economic, social and nutritional importance of banana in the world. In: Picq C, Fouré C, Frison EA (eds) Bananas and food security, 10–14 November 1998, pp 19–20

Gawel NJ, Jarret RL (1991) A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. Plant Mol Biol Rep 9:262–266

Ghislain M, Zhang D, Fajardo D, Huaman Z, Hijimans RJ (1999) Marker-assisted sampling of the cultivated Andean potato *Solanum phureja* collection using RAPD markers. Genet Resour Crop Evol 46:547–555. https://doi.org/10.1023/A:1008724007888

Greene SL, McFerson JR (1996) Conservation of Lotus genetic resources: status of the U.S. collection. In: Beuselinck PR, Roberts CA (eds) First international lotus symposium, 22–24 March 1994. University of Missouri, Columbia, pp 39–42

Grzebelus D (2006) Transposons insertion polymorphism as a new source of molecular markers. J Fruit Ornam Plant Res 14:21–29

Harrison JS (2000) Comparative genome organisation in plants: from sequence and markers to chromatin and chromosomes. Plant Cell 12:617–636. https://doi.org/10.1105/tpc.12.5.617

Hu J, Zhu J, Xu HM (2000) Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theor Appl Genet 101:264–268. https://doi.org/10.1007/s001220051478

Jesus ON, de Silva SO, Amorim EP, Ferreira CF, Campos JMS, Silva G, Figueira A (2013) Genetic diversity and population structure of Musa accession in ex situ conservation. BMC Biol 13(1):41. https://doi.org/10.1186/1471-2229-13-41

Kalendar R, Grab T, Regina M, Suoniemi A, Schulman AH (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theor Appl Genet 98:704–711. https://doi.org/10.1007/s001220051124

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp micro-climatic divergence. Proc Natl Acad Sci USA 97:6603–6607. https://doi.org/10.1073/pnas.110587497

Kimura M, Crow JK (1964) The number of alleles that can be maintained in a finite population. Genetics 49:725–738

Kumar A, Mishra P, Singh SC, Sundaresan V (2014) Efficiency of ISSR and RAPD markers in genetic divergence analysis and conservation management of *Justicia adhatoda* L., a medicinal plant. Plant Syst Evol 300:1409–1420

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 3:349–358

Lewontin RC (1972) The genetic basis of evolutionary change. Columbia University Press, New York

Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 34:816–834. https://doi.org/10.1002/gepi.20533

Liu M, Plunkett GM, Lowry PP (2010) Fruit anatomy provides structural synapomorphies to help define *Myodocarpaceae* (Apiales). Syst Bot 35:1–7

McDermott JM, McDonald BA (1993) Gene flow in plant pathosystems. Annu Rev Phytopathol 31:353–373

Nair AS, Teo CH, Schwarzacher T, Harrison PH (2005) Genome classification of banana cultivars from South India using IRAP markers. Euphytica 144:285–290. https://doi.org/10.1007/s10681-005-7321-2

Nei M (1973) Analysis of genetic diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Onyago M, Haymer D, Keeley S, Manshardt R (2010) Analysis of genetic diversity and relationships in East African 'Apple banana' (AAB genome and 'Muraru' (AA genome) dessert bananas using microsatellite markers. In: Dubois T et al (eds) Proceedings of the IC on bananas and plantain in Africa. Acta horticulturae, vol 879, pp 623–632. https://doi.org/10.17660/actahortic.2010.879.69

Osuji JO, Okoli BE, Vuylsteke D, Ortiz R (1997) Multivariate pattern of quantitative trait variation in triploid banana and plantain cultivars. Sci Hortic 71:197–202. https://doi.org/10.1016/S0304-4238(97)00101-5

Padmesh P, Mukunthakumar S, Vineesh PS, Skaria R, Hari Kumar K, Krishnan PN (2012) Exploring wild genetic resources of *Musa acuminata* Colla distributed in the humid forests of southern Western Ghats of peninsular India using ISSR markers. Plant Cell Rep 31:1591–1601. https://doi.org/10.1007/s00299-012-1273-5

Ravishankar KV, Vidhya L, Cyriac A, Rekha A, Goel R, Singh NK, Sharma TR (2012) Development of SSR markers based on a survey of genomic sequences and their molecular analysis in banana (Musa spp.). J Hortic Sci Biotechnol 87:84–88. https://doi.org/10.1080/14620316.2012.11512835

Rekha A, Ravishankar KV, Anand L, Hiremath SC (2001) Genetic and genomic diversity in banana (Musa species and cultivars) based on D2 analysis and RAPD markers. InfoMusa 10(2):29–34

Rohlf FJ (1990) NTSYS-PC. Numerical taxonomy and multivariate analysis system. Exeter Software, New York

Sarabhai MN, Abdul RB, Kurien AM (2016) Levels of diversity and population structure present in banana cultivars distributed in various genome groups. Int J Plant Breeed Genet 3(2):177–188

Saraswathi MS (2007) Morpho-molecular characterization and marker based screening for BSV integrants in B-genome rich Musa germplasm. Ph.D thesis submitted to Horticultural College and Research Institute, TNAU, Coimbatore

Saraswathi MS, Uma S, Prasanya Selvam K, Ramaraj S, Durai P, Mustaffa MM (2011) Assessing the robustness of IRAP and RAPD marker systems to study intra group diversity among Cavendish (AAA) clones of banana. J Hortic Sci Biotechnol 86(1):7–12. https://doi.org/10.1080/14620316.2011.11512717

Simmonds NW, Shepherd K (1955) The taxonomy and origins of the cultivated bananas. Bot J Linn Soc 55:302–312

Slatkin M (1987) Gene flow and the geographic structure of populations. Science 236:787–792

Teo CH, Tan SH, Ho CL, Faridah QZ, Othman YR, Heslop-Harrison JS, Kalendar R, Schulman AH (2005) Genome constitution and classification using retrotransposon-based markers in the orphan crop banana. J Plant Biol 48(1):96–105

Ude G, Pillay M, Nwakanma D, Tenkouano A (2002) Analysis of genetic diversity and sectional relationships in Musa using AFLP markers. Theor Appl Genet 104(8):1239–1245. https://doi.org/10.1007/s00122-001-0802-3

Ude G, Ogundiwin E, Pillay M, Tenkouano A (2003) Genetic diversity in an African plantain core collection using AFLP and RAPD markers. Theor Appl Genet 107(2):248–255. https://doi.org/10.1007/s00122-003-1246-8

Uma S, Sathiamoorthy S (2002) Names and synonyms of bananas and plantains of India. National Research Centre for Banana, Trichy

Uma S, Sudha S, Saraswathi MS, Manikavasagam M, Selvarajan R, Durai P, Sathiamoorthy S, Siva SA (2004) Analysis of genetic diversity and phylogenetic relationship among indigenous and exotic Silk (AAB) group of bananas using RAPD markers. J Hortic Sci Biotechnol 79(4):523–527. https://doi.org/10.1080/14620316.2004.11511799

Uma S, Siva SA, Saraswathi MS, Durai P, Sharma TVRS, Singh DB, Selvarajan R, Sathiamoorthy S (2005) Studies on the origin and diversification of Indian wild banana (*Musa balbisiana*) using arbitrarily amplified DNA markers. J Hortic Sci Biotechnol 80(5):575–580. https://doi.org/10.1080/14620316.2005.11511980

Uma S, Siva SA, Saraswathi MS, Manickavasagam M, Durai P, Selvarajan R, Sathiamoorthy S (2006) Variation and intraspecific relationship Indian wild *Musa balbisiana* (BB) population as evidenced by randomly amplified polymorphic DNA. Genet Resour Crop Evol 53(2):349–355. https://doi.org/10.1007/s10722-004-0576-y

Uma S, Saraswathi MS, Backiyarani S, Durai P, Mustaffa MM (2014) Banana-Less exploited landraces in India. Technical bulletin no. 21. ICAR-National Research Centre for Banana, Tiruchirapalli

Valmayor RV, Jamaluddin SH, Silayoi B, Kusumo S, Danh LD, Pascua OC, Espino RRC (2000) Banana cultivar names and synonyms in South-east Asia. International Network for the Improvement of Banana and Plantain-Asia and the Pacific Office, Laguna

Wong C, Kiew R, Argent G, Set O, Lee SK, Gan YY (2002) Assessment of the validity of the sections in Musa using AFLP. Ann Bot 90:231–238

Yeh FC, Yang R, Boyle T (1999) POPGENE version 1.31, Microsoft Window-based freeware for population genetics analysis. Quick user guide, University of Alberta