


# Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity

Pavla Christelová<sup>1</sup> · Edmond De Langhe<sup>2</sup> · Eva Hřibová<sup>1</sup> ·  
Jana Čížková<sup>1</sup> · Julie Sardos<sup>3</sup> · Markéta Hušáková<sup>1</sup> ·  
Ines Van den houwe<sup>4</sup> · Agus Sutanto<sup>5</sup> · Angela Kay Kepler<sup>6</sup> ·  
Rony Swennen<sup>2,4,7</sup> · Nicolas Roux<sup>3</sup> · Jaroslav Doležel<sup>1</sup> 

Received: 2 August 2016 / Revised: 1 November 2016 / Accepted: 2 December 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Bananas (*Musa* spp.) are one of the main fruit crops grown worldwide. With the annual production reaching 144 million tons, their production represents an important contribution to the economies of many countries in Asia, Africa, Latin-America and Pacific Islands. Most importantly, bananas are a staple food for millions of people living in the tropics. Unfortunately, sustainable banana production is endangered by various diseases and pests, and the breeding for resistant cultivars relies on a far too small base of genetic variation. Greater diversity needs to be incorporated in breeding, especially of wild species. Such work requires a large and thoroughly characterized germplasm collection, which also is a safe depository of genetic diversity. The largest ex situ *Musa* germplasm collection is kept at the International Transit Centre (ITC) in Leuven (Belgium) and currently comprises

---

Communicated by Daniel Sanchez Mata.

---

This article belongs to the Topical Collection: Ex-situ conservation.

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s10531-016-1273-9](https://doi.org/10.1007/s10531-016-1273-9)) contains supplementary material, which is available to authorized users.

---

✉ Jaroslav Doležel  
dolezel@ueb.cas.cz

<sup>1</sup> Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Šlechtitelů 31, 78371 Olomouc, Czech Republic

<sup>2</sup> Division of Crop Biotechnics, Laboratory of Tropical Crop Improvement, Katholieke Universiteit Leuven, Willem de Croylaan 42, 3001 Heverlee, Belgium

<sup>3</sup> Biodiversity International, Parc Scientifique Agropolis II, 34397 Montpellier, France

<sup>4</sup> Biodiversity International, Willem de Croylaan 42, 3001 Heverlee, Belgium

<sup>5</sup> Indonesian Agricultural Agency for Research and Development, Jl. Ragunan No. 29, Pasar Minggu, Jakarta, Indonesia

<sup>6</sup> Hawaii and Maui Invasive Species Committee, P.O. Box 983, Makawao, HI 96768, USA

<sup>7</sup> International Institute of Tropical Agriculture, P.O. Box 10, Duluti, Arusha, Tanzania

over 1500 accessions. This report summarizes the results of systematic cytological and molecular characterization of the *Musa* ITC collection. By December 2015, 630 accessions have been genotyped. The SSR markers confirmed the previous morphological based classification for 84% of ITC accessions analyzed. The remaining 16% of the genotyped entries may need field verification by taxonomist to decide if the unexpected classification by SSR genotyping was correct. The ploidy level estimation complements the molecular data. The genotyping continues for the entire ITC collection, including newly introduced accessions, to assure that the genotype of each accession is known in the largest global *Musa* gene bank.

**Keywords** Classification · Gene bank · Genetic diversity · ITC collection · Microsatellites · *Musa*

## Introduction

Bananas (*Musa* spp.) are one of the world's major fruit crops that have been cultivated since the dawn of agriculture (Denham et al. 2003). Their annual production has reached 144 million tons (FAOSTAT 2014) and different types of bananas are a staple and nutritious food for hundreds of millions in the tropics and subtropics. Despite the importance of bananas for food security, there is insufficient knowledge and few tools are available to counteract the negative impact of diseases and pests threatening production. Currently, large plantations manage these threats with massive amounts of agrochemicals, creating economically and ecologically unsustainable growing conditions; while smallholders often use improperly characterized local cultivars and mostly cannot afford large investments into agrochemicals.

Breeding new and resistant cultivars is hampered by virtual seed sterility of cultivated clones (Ortiz and Swennen 2014). Most of the modern cultivars are believed to have originated from natural inter- and intraspecific crosses of wild AA and BB diploids ( $2n = 2 \times = 22$ ) derived from *Musa acuminata* Colla (with genome A, D'Hont et al. 2012) and *M. balbisiana* Colla, respectively (with genome B, Davey et al. 2013). *Musa schizocarpa* (S genome) and *M. textilis* (T genome) have also contributed to the origin of some cultivars. Next to the seed sterility, parthenocarpy has played an important role in formation of edible fruits, which were domesticated by early farmers. While *M. acuminata* is thought to have originated in Malaysia or Indonesia (Simmonds 1962; Nasution 1991), the centre of diversity for *M. balbisiana* was designated to India, Myanmar, Thailand and Phillipines (Daniells et al. 2001), to where *M. acuminata* edible diploid and triploid cultivars were brought by humans. This allowed natural hybridization of the two species, resulting in various genomic combinations [diploid AB; triploid ( $2n = 3 \times = 33$ ) AAB, ABB; and less frequent tetraploid ( $2n = 4 \times = 44$ ) AAAB, AABB; Simmonds and Shepherd 1955]. Throughout the years of domestication process, many edible banana varieties have arisen all around SE Asia. Migration of humans in the early days of agriculture have brought bananas to the secondary centers of diversity such as Africa (where AAB Plantains and AAA East African Highland bananas arose) or Pacific Islands (with today's AAB Maoli/Popo'ulu and Iholena cultivars; De Langhe et al. 2009). Despite enormous difficulties, *Musa* breeders have managed to produce successful hybrids (Vuylsteke et al. 1993; Rowe 1998; Ortiz and Swennen 2014). One of the key steps in the *Musa* breeding process is the development of agronomically-improved, disease-resistant

diploid parental lines (Tenkouano et al. 2003), which can be used for producing synthetic tetraploid hybrids (Rowe and Rosales 1996). However, little is known about the exact ancestral genetic basis of domesticated banana cultivars (De Langhe et al. 2010), which hampers the choice of suitable parents for crosses to produce new hybrids with plant and fruit qualities comparable to the currently grown cultivars.

Yet, the genetic diversity of wild banana species is vast, and a rather restricted portion of it seems to have been used during the initial domestication process through their semi-fertile variants (De Langhe et al. 2009). Hence, there is a rich pool of genetic material among the wild *Musa* species (Janssens et al. 2016) that still needs to be described, collected, preserved and used in breeding.

The largest international ex situ collection of banana germplasm (International Transit Centre, ITC) comprising 1518 accessions is maintained by Bioversity International and hosted at the Catholic University in Leuven (KU Leuven), Belgium. The ITC collection has a good overall representation of the groups of cultivated bananas. However, some specific cultivated groups and subgroups from some geographical areas, as well as wild species are still underrepresented in the collection. There are still an estimated 300–400 cultivars and wild specimens known to be missing in the collection, including unique AAB plantains from the Congo-Basin, Fe'i bananas and wild species from the region of Papua New Guinea, diploids from the East African region, *M. balbisiana* diversity from China and India, Callimusa species from Borneo, and wild and cultivated bananas from Myanmar. The efficient management of such collection requires a strategy for thorough characterization of the stored accessions. This is critical for the identification of duplicated and mislabeled accessions, as well as accessions that may have undergone somatic mutations, but particularly needed for efficient selection of germplasm for distribution and use in breeding. Besides the precise morpho-taxonomic classification that, in *Musa*, is based on a set of phenotypic characters (IPGRI-INIBAP/CIRAD 1996) and basic chromosome number, a molecular-level characterization of these accessions has been providing complementary information on diversity and relationships.

Bananas and plantains have traditionally been classified according to Cheesman (1947) into four sections: Eumusa ( $x = 11$ ), Rhodochlamys ( $x = 11$ ), Australimusa ( $x = 10$ ) and Callimusa ( $x = 9, 10$ ), later supplemented with the fifth section Ingentimusa (Argent 1976), containing a single species *M. ingens* ( $x = 7$ ). The recent revision of the sectional classification (Häkkinen 2013), supported by the results obtained using DNA markers and phylogenetic analyses (e.g. Bartoš et al. 2005; Risterucci et al. 2009; Li et al. 2010; Christelová et al. 2011; Hřibová et al. 2011) suggested a merger between Eumusa and Rhodochlamys into the newly established *Musa* section and, similarly, all Australimusa species are covered by the section Callimusa.

DNA markers proved to be useful for characterizing germplasm in various plant species (Chin et al. 1996; Roder et al. 1998; Cregan et al. 1999; Hayden et al. 2007; Jing et al. 2010) and amongst others, microsatellites (simple sequence repeats; SSR) markers gained special popularity thanks to their co-dominant inheritance, good reproducibility and high abundance in plant genomes (Goldstein and Schlotterer 1999; Selkoe and Toonen 2006). In, Christelová et al. (2011) established a standardized procedure for genotyping *Musa* accessions with 19 SSR markers, which were selected from a larger set of SSR markers developed by Crouch et al. (1998), Lagoda et al. (1998) and Hippolyte et al. (2010). This resulted in the establishment of the *Musa* Genotyping Centre (MGC) in Olomouc, Czech Republic, under the auspices of Bioversity International.

The systematic molecular characterization of the ITC collection (including the already genotyped accessions as well as the newly introduced samples) requires a robust

genotyping system which can handle large batches of samples as well as single accessions sent for re-analysis, without compromising the comparability of results. The genotyping technology evolves fast with high-throughput genotyping methods such as DArT or next-generation sequencing based marker systems (Mason et al. 2015), some of which have been applied to *Musa* research (e.g. Risterucci et al. 2009). Yet SSR markers are still used as an effective tool in elucidating the genetic diversity in many plant species, including bananas (Nicolai et al. 2013; Mbanjo et al. 2012; Gross-German and Viruel 2013; Liu et al. 2015; Kitavi et al. 2016). The main advantage of the SSR genotyping approach resides in its capability of systematically adding new information to existing data sets. The newly analyzed accessions can be added to the existing database of SSR profiles, while the optimized and undemanding methodology assures comparability of results gathered over different time points, at a reasonable cost.

The maintenance of genebanks of vegetatively propagated plants is expensive and time-demanding. Therefore, it requires a rationalized approach dealing with properly characterized germplasm and ensuring the management of true-to-type germplasm. This is particularly important in banana where somaclonal variation occurs more frequently due to the *in vitro* culture process (Côte et al. 1993; Vuylsteke et al. 1996; Rodrigues et al. 1998). In *Musa*, germplasm is usually conserved as *in vitro* proliferating meristems under limited growth conditions (Van den houwe et al. 1995) complemented with cryopreservation to minimize the risk of contamination or human error during subculturing, as well as to avoid somaclonal variation (Panis et al. 2005; Panis 2009). In fact, Doleželová et al. (2005) identified accessions in the ITC collection that contained plants of different ploidy (mixed ploidy accessions) as well as accessions in which plants comprised cells with different ploidy levels (mixoploid accessions). Although the origin of off-type plants is not understood, chromosome number changes (polyploidy, aneuploidy) due to repeated *in vitro* subcultures might have contributed to this phenomenon. Thus, as a part of the quality management of the MGC, ploidy levels are ascertained via flow cytometry, a method which excels in high throughput and precision (Doležel et al. 1994; Roux et al. 2003; Pillay et al. 2006).

In the present study, we coupled flow-cytometric analysis of ploidy level with the genotyping platform based on SSR markers and analyzed 630 accessions of the genus *Musa* held at the ITC collection including 49 Reference DNA collection samples (<http://www.musanet.org/Musagenotypingcentre/genomicDNA>), as well as 27 samples received from Hawaii to enlarge the representation of individual *Musa* subgroups, and 38 samples collected during the Bioversity International expedition to Indonesia (Sutanto et al. 2016). The current work was undertaken (1) to improve the characterization of accessions held at the ITC collection, (2) to identify problematic accessions at the ITC collection and reduce duplicated entries, (3) to support the *Musa* research and breeding community by facilitating proper evaluation of the available germplasm, and (4) to add new knowledge on *Musa* genetic diversity.

## Materials and methods

### Plant material

Most of the plant material analyzed in this study (Online Resource 1) came from the International Transit Centre (ITC, Leuven, Belgium) and was delivered in batches of

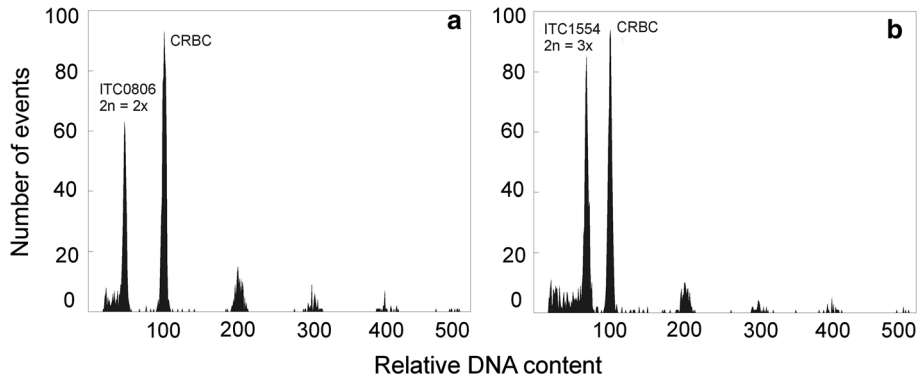
in vitro rooted plantlets of about 50 accessions (five plantlets per accession), or as lyophilized leaf tissues. If fresh plant material was obtained, leaf tissues were lyophilized after ploidy level analysis and kept for further use. In particular situations—if ploidy analysis resulted in ambiguous results (1.2% of measured accessions), the plants were transferred to soil and maintained in a greenhouse for chromosome analysis. Altogether, 695 accessions were genotyped, including 327 diploids, 363 triploids and 5 tetraploids. Apart from the accessions received from ITC and from the *Musa* Reference DNA collection (<http://www.musanet.org/Musagenotypingcentre/genomicDNA>), 27 accessions from Hawaii were provided by Dr. Angela Kay Kepler (Banana Specialist, Hawaii) to enlarge the representation of individual *Musa* subgroups, and 38 accessions were obtained during two *Musa* germplasm collecting expeditions in Indonesia (Sutanto et al. 2016), that are currently being introduced into the ITC collection. The exploration expedition was confined to the East-Indonesian triangle (formed by Halmahera Islands, Sulawesi and Lesser Sunda Islands) and is henceforth referred to as “Indonesian Triangle”. The samples from Hawaii and from the “Indonesian Triangle” expedition were collected as fresh leaves from young banana plants for ploidy level estimation and subsequently lyophilized for further use.

### Ploidy level estimation

Ploidy level of each accession was estimated by flow cytometry according to Doležel et al. (1997, 2007). About 30 mg of a young leaf tissue was chopped with a razor blade in a glass Petri dish containing 500 µl Otto I solution (0.1 M citric acid, 0.5% v/v Tween 20). Crude homogenate was filtered through a 50 µm nylon mesh. Chicken red blood cell nuclei (CRBC), prepared according to Galbraith et al. (1998), were added to the suspension of *Musa* nuclei as an internal reference standard. After 30 min incubation at room temperature, 1 ml Otto II solution (0.4 M Na<sub>2</sub>HPO<sub>4</sub>) (Otto 1990) supplemented with 5 µM DAPI and 3 µl/ml of 2-mercaptoethanol were added. The samples were analyzed using Partec PAS or Sysmex-Partec CyFlow flow cytometers equipped with UV excitation and detectors for DAPI fluorescence. The gain of the instruments was adjusted so that the peak of the CRBC nuclei was positioned approximately at channel 100 on a 512 channel scale. Relative nuclear DNA content of *Musa* accessions was then determined by comparing peak positions of CRBC nuclei and nuclei of the sample (Fig. 1). Every accession was represented by five individual plants and ploidy was estimated in each of them.

### DNA extraction and PCR amplification

Genomic DNA was extracted from lyophilized leaf tissues of young banana plants using a NucleoSpin Plant II kit (Macherey–Nagel, Düren, Germany) following the manufacturer’s recommendations. Each of the accessions received from ITC collection was represented by five individual plantlets. If there were no differences in ploidy among individual plants, genomic DNA was extracted from a pooled sample containing lyophilized tissue from all of them. If plants from the same accession displayed inconsistent results during ploidy analysis, genomic DNA was extracted separately from each plant and individual plants were treated as separate accessions. For the SSR analysis, the pipeline established by Christelová et al. (2011) was used and is illustrated in Fig. 2. Briefly, 19 SSR loci (Lagoda et al. 1998; Crouch et al. 1998; Hippolyte et al. 2010) that are well distributed across ten of the eleven *Musa* genetic linkage groups (Hippolyte et al. 2010) were amplified using a set of M13-tailed specific primers to allow universal labelling with fluorescent dyes. Although new SSR markers are accessible thanks to the *Musa* genome sequence assembly

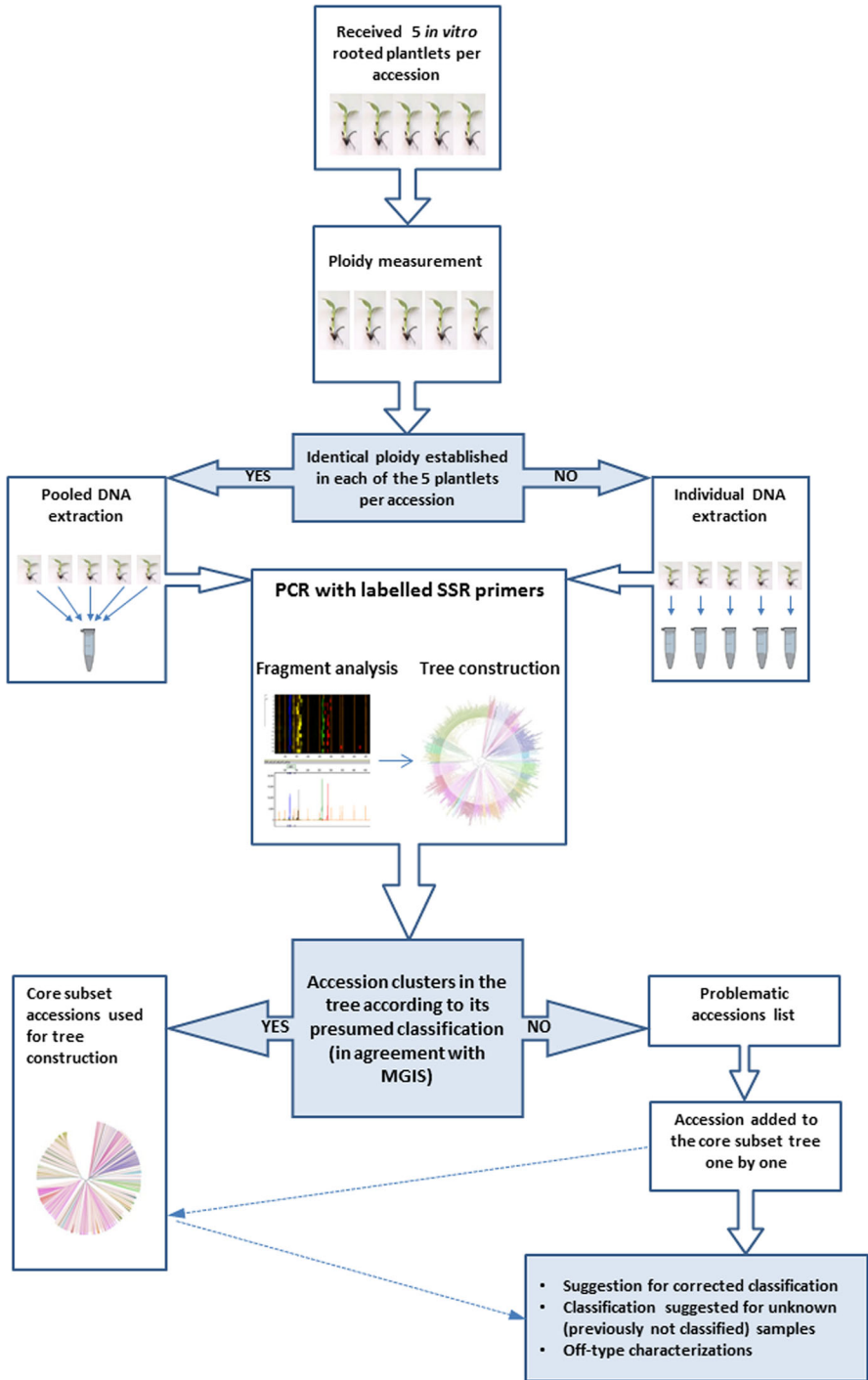


**Fig. 1** Estimation of ploidy level in *Musa* accessions. Histograms of relative nuclear DNA content obtained after simultaneous flow cytometric analysis of DAPI-stained nuclei isolated from fresh leaf tissues of *Musa* and chicken red blood cell nuclei (CRBC). The gain of flow cytometer was adjusted so that the G1 peak of CRBC, which served as an internal reference standard, was positioned on channel 100. Peaks appearing on channels 200, 300, 400 and 500 correspond to doublets, triplets, etc. of CRBC nuclei. Ploidy of *Musa* accessions was determined based on the ratio of G1 peak positions (*Musa* : CRBC), knowing that in diploid, triploid and tetraploid plants, the ratio is  $\sim 0.5$ ,  $0.75$  and  $1$  respectively. **a** Simultaneous analysis of nuclei isolated from a diploid accession (*M. acuminata* ssp. *banksii*, ITC0806) and CRBC. The ratio of G1 peak means was  $\sim 0.5$ . **b** Simultaneous analysis of nuclei isolated from a triploid accession (Itoke, ITC1554) and CRBC. The ratio of G1 peak means was  $\sim 0.75$

completion (D’Hont et al. 2012; Martin et al. 2016), the hitherto used marker set was not enlarged, as the reproducibility and comparability of results gathered until now would then not be assured and a re-start of the whole genotyping effort would be inefficient. The PCR reaction mix contained (in the final volume of 20  $\mu$ l): 10 ng of template genomic DNA, reaction buffer (consisting of 10 mM Tris-HCl (pH 8), KCl 50 mM, 0.1% Triton-X100 and 1.5 mM MgCl<sub>2</sub>), 200  $\mu$ M dNTPs (each), 1 U of Taq polymerase, 8 pmol of the M13-tailed locus specific forward primer, 6 pmol of the fluorescently labeled universal M13 forward primer, 10 pmol of the locus specific reverse primer. The cycling conditions were set as follows; initial denaturation step at 94 °C for 5 min, followed by 35 cycles of denaturation (94 °C/45 s), annealing at temperature corresponding to the locus-specific primer (1 min) and extension (72 °C/1 min). The final extension was allowed for 5 min at 72 °C. Purification of the PCR products was performed by ethanol/sodium acetate precipitation. Two independent PCR reactions were performed in each accession to avoid random errors of allele binning.

### Fragment analysis

Purified PCR products were diluted 40-fold in Hi-Di formamide containing the internal size standard (GeneScan<sup>TM</sup>-500 LIZ size standard; Applied Biosystems, Foster City, CA, USA) and loaded onto a capillary electrophoresis DNA analyzer (ABI 3730xl, Applied Biosystems) after 5 min denaturation (95 °C). Electrophoretic separation and signal detection were carried out with default module settings. In order to reduce the cost and increase throughput of the genotyping platform, samples were multiplexed for electrophoretic separation. Up to fourfold multiplexing was applied by combining four PCR products, labelled with four different fluorescent dyes (6-FAM, VIC, NED and PET) into a single sample for loading. The resulting data were analyzed and called for alleles using



**Fig. 2** SSR genotyping pipeline used in the study for the systematic characterization of *Musa* genebank accessions (ITC collection)



GeneMarker<sup>®</sup> v1.75 (Softgenetics, State College, PA, USA), manually checked and implemented into marker panels (Christelová et al. 2011).

### Genetic diversity analysis using distance-based methods

The extent of genetic diversity among all samples was evaluated using the Nei's genetic distance coefficient calculation (Nei 1973) and subsequent cluster analysis was done using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA; Michener and Sokal 1957). To enable joint analysis of all ploidy levels (2×, 3× and 4×) the genotypic data was converted into binary (coded by 1/0 = presence/absence) and analyzed as a dominant marker's record (Weising et al. 2005). Dendrograms were constructed based on the results of UPGMA analysis and visualized in FigTree v1.4.0 (<http://tree.bio.ed.ac.uk/software/figtree/>). The first dendrogram that comprised all analyzed samples (not shown) was used to identify problematic accessions whose position in the dendrogram did not agree with their current classification and were therefore not used to build the "core subset". The core dataset comprised reliable accessions only, for which the clustering pattern agreed with the classification based on morphological descriptors. The genotyping and evaluating pipeline is shown in Fig. 2. The dissimilarity index threshold of 0.25 was used to assess the grouping of the accessions on the dendrogram. Bootstrap support for individual branches was calculated on 1000 replicates and values above 35% (0.35) were used to confirm the fundamental subclustering pattern. To evaluate the informative power of individual SSR loci, several characteristics were calculated, such as Polymorphism information content (PIC; Botstein et al. 1980), allele number, observed heterozygosity ( $H_o$ , in the diploid dataset) and major allele frequency using Powermarker v3.25 (Liu and Muse 2005).

## Results and discussion

The significant role of genebanks and germplasm collections as essential resource for breeders and researchers is often compromised by misclassification of some accessions. This is more common than expected (Mason et al. 2015), as errors can occur at different stages of handling the material, including the acquisition of new accessions, their delivery to the collection and during maintenance at the genebank, especially if in vitro subculturing is involved. It was shown for a number of plant species that SSR markers are useful to deal with such cases and can help to highlight problematic misclassified or redundant accessions in germplasm collections (e.g. da Cunha et al. 2014; Fjellheim et al. 2015; Roy et al. 2016). The aim of this study was to systematically characterize a substantial portion of the accessions held at the ITC collection and to interpret the genetic structure of the *Musa* germplasm represented by the genotyped accessions. Nearly half of the ITC accessions were genotyped by flow cytometric ploidy estimation and SSR analysis.

### Ploidy level estimation

Flow cytometry was used to estimate the ploidy in all 495 accessions that were received as fresh leaf tissues. The remaining accessions were received either as lyophilized leaf tissues (151) or DNA samples from the Reference DNA collection (49; <http://www.musanet.org/Musagenotyping-centre/genomicDNA>). Ploidy of these later accessions was retrieved from the *Musa* germplasm information system database (MGIS; <http://www.crop-diversity.org/>



mgis/). For a majority of the samples, the expected ploidy was confirmed (Online Resource 1). However, some accessions e.g. Gebi (ITC0877) and Mwitua Pemba (ITC1545) reported to be triploid AAB, were found to be diploid. Moreover, accessions Tongkat Langit Papua (ITC1716) and *M. borneensis* (ITC1531) showed ambiguous results with DNA content between diploid and triploid, or between triploid and tetraploid for accession Sar (ITC0898) and SUP 1 (AAB, Indonesian Triangle mission 2). SSR analysis indicated triploid status by the occurrence of three alleles at some loci. However, this observation does not guarantee three complete chromosome sets, nor it excludes the possibility of tetraploids status. Only chromosome counting would give a definite answer.

Out of 495 accessions analyzed by flow cytometry, two cases of mixoploidy were found. The accession *M. rosea*-hybrid (ITC1598) comprised one mixoploid ( $2\times + 4\times$ ) plant out of the five analyzed plantlets. Mixoploid status ( $3\times + 6\times$ ) was detected in three out of five plants of the accession Malbhog (ITC1631). The remaining two plantlets were either triploid or hexaploid. The presence of the hexaploid plant indicates that a shift to a higher ploidy happened in vitro as hexaploid banana plants do not grow in the field and that the classification of three plants as mixoploid was not an artifact due to accumulation of cells in G2 phase of cell cycle, which is yet to be confirmed in the case of *M. rosea*-hybrid ITC1598. In cases where the five plantlets differed in ploidy levels without showing a mixoploid pattern (ITC1573, ITC1598), analysis of re-sent samples was carried out. No incongruent results were obtained on the re-sent samples, which indicated that the problems detected in the first batch were caused by delivery error or mislabeling. This proved the capability of the genotyping pipeline to detect mislabeled accessions and showed the robustness of the method.

In the latest systematic effort to screen ploidy levels (Doleželová et al. 2005), 1150 entries deposited in the ITC collection by that time were analyzed. Since then, 381 new accessions were introduced into the genebank, out of which 249 accessions (Online Resource 1) had their ploidy level verified by flow cytometry for the first time in this study. For the remaining accessions, repeated analyses provided a valuable feedback for the conservation strategy. As demonstrated in some accessions, e.g., mixoploid ITC1631 and 18 samples listed in Online Resource 2, the ploidy determined in the present work did not agree with previous results.

### Distance-based genetic diversity analysis

The informative power of the 19 SSR loci used in the genetic diversity study (expressed as the Polymorphism Information Content; PIC) ranged between 0.561 and 0.933 with an average value of 0.789 (Table 1). The highest observed heterozygosity within the diploid dataset was observed for the marker mMaCIR01 ( $H_o = 0.623$ ), whereas the locus mMaCIR307 had the lowest  $H_o$  value ( $H_o = 0.250$ ).

The UPGMA analysis of the dataset comprising all ploidy levels together grouped the accessions into clusters, which in general conformed to the morphological traits-based classification into subgroups (in 84% of accessions) as reported by germplasm donors to the collection (Musa Germplasm Information System (MGIS), Bioversity International. Accessed on February 15, 2016 at <http://www.crop-diversity.org/banana/>). However, in 16% of cases, the classification reported did not agree with the clustering observed in the dendrogram. These accessions are listed in Online Resource 2 and the discrepancy between the current classification and the results obtained here are possibly caused by mislabeling or handling error, or due to misclassification before introduction into the genebank. This set of 71 “to be verified” accessions was excluded from the dataset that was used for the

**Table 1** Frequency of the major allele, allele number, heterozygosity and informativeness (PIC) of the 19 microsatellite loci used for genotyping the set of *Musa* accessions

Marker	Major allele frequency	Allele number	Observed heterozygosity	PIC
mMaCIR01	0.113	42	0.623	0.933
mMaCIR03	0.402	14	0.423	0.684
mMaCIR07	0.179	24	0.590	0.894
mMaCIR08	0.418	17	0.520	0.760
mMaCIR13	0.291	29	0.508	0.840
mMaCIR24	0.430	20	0.379	0.742
mMaCIR152	0.395	12	0.188	0.744
mMaCIR214	0.482	11	0.226	0.619
Ma-3-90	0.235	18	0.516	0.884
mMaCIR39	0.316	26	0.583	0.856
mMaCIR40	0.346	31	0.459	0.831
mMaCIR45	0.366	15	0.377	0.749
mMaCIR150	0.386	17	0.429	0.738
mMaCIR164	0.266	37	0.281	0.882
mMaCIR196	0.340	14	0.327	0.801
mMaCIR231	0.308	22	0.537	0.854
mMaCIR260	0.345	17	0.401	0.712
mMaCIR264	0.144	32	0.562	0.912
mMaCIR307	0.584	11	0.250	0.561
Mean	0.334	21.5	0.430	0.789

The marker set used was identical to the panel applied in Christelová et al. (2011). (Original marker references: Hippolyte et al. 2010; Lagoda et al. 1998; Crouch et al. 1998)

final dendrogram construction. Furthermore, 24 accessions for which classification was incomplete or not provided, and eight synthetic hybrid accessions were removed from the dataset, in order to increase the resolution of the “core subset” tree. The remaining 591 accessions, representing only entries with coherent classification and genotyping data, were used to construct a final SSR dendrogram. The representativeness of the subset with regard to individual species/subspecies and subgroups of *Musa* is summarized in Table 2. We considered this subset of accessions as the “core subset” to which any accession from the excluded ones could be added in future and its potential true/more specific classification could be proposed. However, only field verification which is carried out in parallel (Chase et al. 2016) can provide conclusive results (examples included in Online Resource 2).

Our results indicate that the SSR markers covered well the genetic diversity of the *Musa* genus as revealed by the clustering pattern on the dendrogram (Fig. 3, Online Resource 3). The dendrogram reveals 48 sets of accessions separated at the Nei’s dissimilarity index threshold of 0.25 (Table 3) with relatively significant bootstrap support for most of them (>35%). For the wild accessions, the sets and some singletons correspond to wild species and subspecies, while at the triploid cultivar level, the sets correspond to subgroups, of which the members are supposed to have derived as somatic mutants from an original domesticate. At the diploid cultivar level, the situation has not yet been clarified, and the sets may be composed of somatic mutants and accessions with different parentage.

**Table 2** Representativeness of the core subset with regard to the species, subspecies and subgroups of Musa

Section	Species/group	Subspecies/subgroup	No. of classified representatives <sup>a</sup>		
Musa	<i>M. acuminata</i>	<i>banksii</i>	30		
		<i>burmannica/burmannicoides</i>	6		
		<i>errans</i>	1		
		<i>malaccensis</i>	12		
		<i>microcarpa</i>	1		
		<i>sumatrana</i>	1		
		<i>truncata</i>	1		
		var. <i>acuminata</i>	5		
		var. <i>tomentosa</i>	3		
	<i>zebrina</i>	2			
	<i>M. balbisiana</i>		18		
	<i>M. itinerans</i>		3		
	<i>M. laterita</i>		2		
	<i>M. mannii</i>		4		
	<i>M. ornata</i>		9		
	<i>M. rosea</i>		3		
	<i>M. rubinea</i>		1		
	<i>M. rubra</i>		1		
	<i>M. schizocarpa</i>		14		
	<i>M. siamensis</i>		3		
	<i>M. velutina</i>		2		
	AA cv.		116		
	AB cv.		9		
	AS cv.		6		
	BxT		1		
	AAA		Ambon	2	
			Cavendish	16	
			Gros Michel	6	
			Ibota	3	
			Lakatan	3	
			Lujugira Mutika	22	
			Orotava	1	
			Red	6	
Rio			2		
Umğ Bir			1		
AAB				Maoli/Popo'ulu	17
				Mysore	10
				Pisang Raja (h)	3
				Plantain	117
	Plantain-like (Iholena, Laknau)	17			
	Pome	19			
	Silk	12			

**Table 2** continued

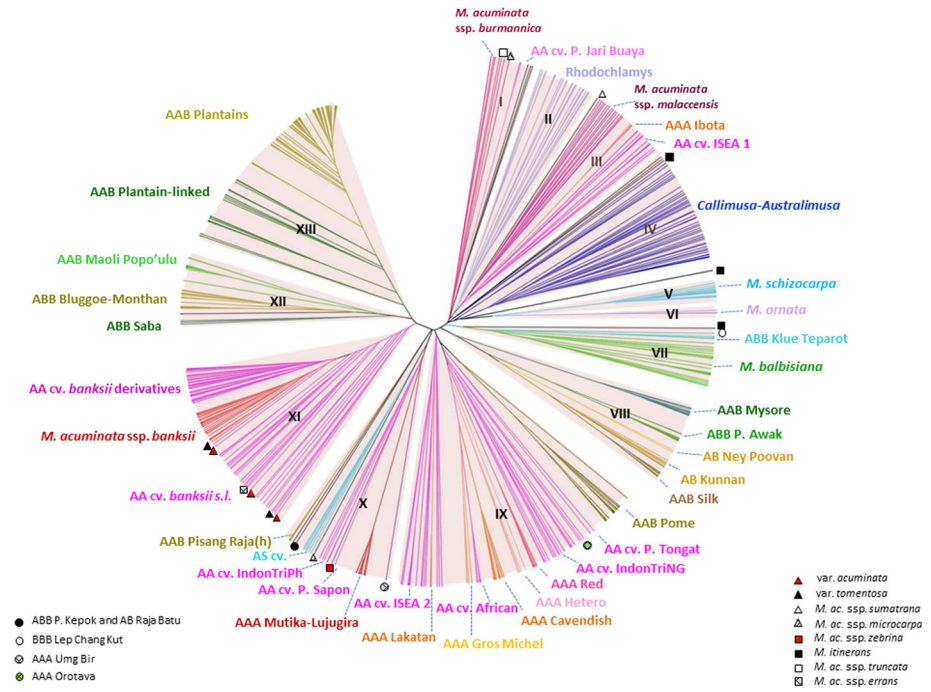
Section	Species/group	Subspecies/subgroup	No. of classified representatives <sup>a</sup>
	ABB	Bluggoe	14
		Bontha	3
		Kluai Teparod	5
		Monthan	7
		Ney Mannan	2
		Pisang Awak	9
		Saba	3
Callimusa	<i>M. barioensis</i>		1
	<i>M. beccarii</i>		3
	<i>M. borneensis</i>		1
	<i>M. campestris</i>		2
	<i>M. coccinea</i>		2
	<i>M. jackeyi</i>		1
	<i>M. maclayi</i>		2
	<i>M. monticola</i>		1
	Fe'i		9
	<i>M. lolodensis</i>		4
	<i>M. lutea</i>		2
	<i>M. peekelii</i>		1
	<i>M. salaccensis</i>		1
	<i>M. textilis</i>		2
	<i>M. troglodytarum</i>		2
	<i>M. violascens</i>		1
	<i>M. viridis</i>		2

<sup>a</sup> Some of the accessions had no subgroup/subspecies classification reported from the donors, but their refined classification was revealed after SSR cluster analysis. Such accessions are included in the total number for each category

The 48 sets of accessions formed 13 clusters as denoted on Fig. 3 and in Table 3. The diversity as revealed by the UPGMA clustering is further discussed by individual groups, according to their classification as wild and cultivated diploids and edible triploid cultivars, respectively.

### Wild *Musa* species diversity

For the ease and clarity of the discussion regarding the diversity within the genus and to illustrate the discussed accessions in all figures, we decided to retain the former classification of genus *Musa* into sections Callimusa, Australimusa, Rhodochlamys and Eumusa. The accessions formerly classified under section Rhodochlamys (cluster II) are now considered as forming one assembly with the ex-section Eumusa, and entitled “*Musa*” (Häkkinen 2013). It may be worthwhile to revisit this naming as it creates confusion with the name of the genus ‘*Musa*’. Similarly, the merger between the former sections Callimusa and Australimusa (cluster IV) in the complex named Callimusa needs further study



**Fig. 3** Diversity tree of the core subset accessions. UPGMA dendrogram was constructed based on the results of SSR analysis. Only accessions clustering in agreement with their previous classification were used for the core subset diversity tree construction. Individual blocks of clustered accessions are named according to Table 3 and differentiated by colored branches. Major clusters are marked I–XIII. Full version of the tree including names of all accessions is shown in Online Resource 3

based on our results. The UPGMA tree (Online Resource 3) illustrates a distinct (and rather geographical) order in the complex with four divisions along the range of accessions: (1) Callimusa in Mainland South East Asia [from ITC1683 to ITC1532]; (2) Australimusa and some Fe'i accessions in Southeast Indonesia [from AMB006 to Fe'i Aiuri S8]; (3) Callimusa in Borneo [from ITC1568 to ITC1516B]; (4) Australimusa in the New Guinea region and the Pacific [other Fe'i, from ITC0614 to ITC0917].

Representatives of *M. ornata* formed a distinct separate block (cluster VI) distantly positioned from the other ex-Rhodochlamys entries (all in cluster II: *M. velutina*, *M. mannii*, *M. rosea*, *M. siamensis*, *M. rubra*, *M. laterita*, *M. sanguinea*). This result may indicate that *M. ornata* is a hybrid species, which was also suggested by Shepherd (1999), who formulated his opinion on *M. ornata* being a naturalized hybrid of *M. velutina* (Rhodochlamys) and *M. flaviflora* (a small species closely resembling *M. acuminata*).

Separate location of three well-represented subspecies of *M. acuminata* (*burmannica*, *malaccensis* and *banksii*) on the SSR tree may have a genetically indicative value. The 31 accessions of ssp. *banksii* form a coherent block in the middle of the cluster XI (Fig. 3) with more or less evident links to diploid AA cultivars. A comparable link for the block of 12 accessions of ssp. *malaccensis* in cluster III can be drawn. The block of seven accessions of ssp. *burmannica* is grouped with a sole and morphologically quite different AA cv. block (Pisang Jari Buayá, cluster I). This observation may indicate participation of this

**Table 3** Table is accompanying the SSR dendrograms (Fig. 3; Online Resource 3) that were constructed from the core subset of accessions

Block of clustered accessions (as denoted on Fig. 3; Online Resource 3)		Boundaries of the blocks of accessions expressed by left and right marginal accession labels (as displayed in Online Resource 3)	Cluster (as displayed on Fig. 3)
Number	Name of the block		
1	<i>M. acuminata</i> ssp. <i>burmannica</i>	Ref22_ITC0283—ITC0249	I
2	<b>AA cv. Pisang Jari Buaya</b>	BL16_1308—ITC0316	I
3*	Unclassified	ITC1192; ITC0610; SUP 1	I
4	Rhodochlamys	ITC1330—ITC1575B	II
5*	<i>M. acuminata</i> ssp. <i>sumatrana</i>	ITC1701	III
6	<i>M. acuminata</i> ssp. <i>malaccensis</i>	ITC0712—Ref52 ITC1511	III
7	AAA Ibota	ITC1707—ITC0662	III
8	<b>AA cv. ISEA 1</b>	ITC1693—ITC1193	III
9	Callimusa—Australimusa complex	ITC1683—ITC0917	IV
10	<i>M. schizocarpa</i>	ITC0599—ITC1208	V
11	<i>M. ornata</i>	BL6_0370—Ref51 ITC0637 + ITC1690,ITC1744	VI
12*	<i>M. itinerans</i>	ITC1571 (+ITC1526, ITC1523 further on)	n/a
13	ABB Klue Teparot	ITC0647—Ref9_ITC0652	VII
14	<i>M. balbisiana</i>	ITC0246—BL1_0080	VII
15	AAB Mysore	ITC1734—ITC1651	VIII
16	ABB P. awak	ITC1719—ITC1731	VIII
17	AB Ney poovan	ITC1635—Ref23_ITC0245	VIII
18	AB kunnan	ITC1752—ITC1751	VIII
19	AAB silk	ITC1550B—ITC1553	VIII
20	AAB pome	Ref46_ITC0243—ITC1640	IX
21	<b>AA cv. P. Tongat</b>	ITC0685—ITC0446	IX
22	AAA/AA Orotava	ITC1646; ITC1694; Ref14_ITC0420	IX
23	<b>AA cv. IndonTriNG</b>	MLU_001—MDO_007	IX
24	AAA red	Dwarf Red S10—SUP 3	IX
25	AAA hetero	ITC0671—Ref19_ITC1064	IX
26	AAA cavendish	ITC0557—ITC1373	IX
27	<b>AA cv. African</b>	ITC1561—ITC1223	IX
28	AAA Gros Michel	ITC1478—ITC1480B	IX
29	<b>AA cv. ISEA 2</b>	ITC0437—ITC1150	IX
30	AAA Lakatan	Ref55_ITC1287—S18	IX
31*	AAA Umg Bir	ITC1743	X

**Table 3** continued

Block of clustered accessions (as denoted on Fig. 3; Online Resource 3)		Boundaries of the blocks of accessions expressed by left and right marginal accession labels (as displayed in Online Resource 3)	Cluster (as displayed on Fig. 3)
Number	Name of the block		
32	AAA Mutika-Lujugira	ITC1760—ITC1557	X
33*	<b>AA cv. P. Sapon</b>	ITC0679	X
34	<i>M. acuminata</i> ssp. <i>zebrina</i>	ITC0728—Ref37_ITC1177	X
35	<b>AA cv. IndonTriPh</b>	ITC0434—MLU_13	X
36*	<i>M. acuminata</i> ssp. <i>microcarpa</i>	Ref 7_ITC0253	X
37	AxS	ITC0859—ITC0822	X
38*	ABB P. Kepok	SUM 001	X
39*	AB Raja Batu	MLU 015	X
40	AAB P. Raja(h)	ITC1703—Ref60_ITC0587	X
41	<b>AA cv. banksii sensu lato</b>	Ref38_ITC1187—AMB 004	XI
42	<i>M. acuminata</i> ssp. <i>banksii</i>	ITC0937—ITC0616	XI
43	<b>AA cv. banksii derivatives</b>	ITC0818—ITC0949	XI
44	ABB Saba	'saba' S14—ITC1745	XII
45	ABB bluggoe-monthan	ITC1607—ITC1636	XII
46	AAB Maoli-Popo'ulu	ITC1327—Ref27_ITC0335	XII
47	AAB plantain-linked	ITC0719—ITC0875	XIII
48	AAB plantain	ITC1491—ITC1714	XIII

The table summarizes the groups of clustered accession and indicates boundaries of *Musa* accession blocks that are referred to in the text. The ten AA cv. blocks are highlighted. Asterisk marks groups with bootstrap support values lower than 0.35

subspecies in the origin of AA cultivars to a much lesser extent, and conforms to Simmond's (1962) findings that the subspecies does not seem to possess parthenocarpic genes.

The weakly represented East-Indonesian *acuminata* taxa, var. *tomentosa* and var. *acuminata* (Nasution 1991), as well as ssp. *errans* are spread over the “*banksii sensu lato*”-block in cluster XI, and are discussed further below.

Proportion of shared SSR-alleles and the proximity between ssp. *zebrina*, and African “AAA Mutika-Lujugira block” within the cluster X supports the previous reports on the contribution of *zebrina* and *banksii* subspecies to the formation of this triploid subgroup (Carreel et al. 2002; Boonruangrod et al. 2008; Risterucci et al. 2009; Hippolyte et al. 2012) and based on our results, the *zebrina* contribution is predominant.



The position of the singletons ssp. *microcarpa*, and ssp. *sumatrana* (Fig. 3, white and grey triangles) in the dendrogram seems unreliable with rather low bootstrap values for corresponding branches (0.04 and 0.12, respectively). Increasing the sample representation for each of the subspecies might be needed to produce a more robust position on the tree.

As shown in a number of molecular studies on bananas (e.g. Hippolyte et al. 2012; de Jesus et al. 2013), the intraspecific diversity of *M. balbisiana* is low compared to *M. acuminata*. This is reflected in our work by clustering all *balbisiana* accessions in a single coherent block (Fig. 3, cluster VII) that contains morphologically distinct BB forms from Mainland and Island SE Asia ('Butuhan', 'P. Klutuk' and 'Pisang batu') intermingled with BBs of South Asian origin. Although genetic diversity within *M. balbisiana* has been reported (Ude et al. 2002; Ge et al. 2005; Swangpol et al. 2007; Wang et al. 2007), it is believed that the ITC collection does not cover enough *balbisiana* diversity to trace the ancestral lineages of triploid cultivars comprising the B-genome (Hippolyte et al. 2012).

Another wild species, *M. itinerans*, ubiquitous from NE India in the west to Taiwan in the east (Häkkinen et al. 2008) is morphologically a very distinct species with characteristic long rhizomes, which was, however, not reflected by the cluster analysis. Indeed, the three representatives of this species were scattered over the dendrogram instead of being grouped in one block (black boxes on Fig. 3), which could be attributed to underrepresentation of the sample in the present dataset or potential mislabeling. On the other hand, it may also indicate the limitations of the 19 SSR marker set used. The selected 19 SSR markers proved to work very well for verifying classification of diploid and triploid cultivars subgroups (as described further below), and for some *M. acuminata* subspecies (see above). However, there seem to be limitations when it comes to certain species/subspecies, such as *M. itinerans*, *M. acuminata* ssp. *errans* and others (see Table 2), which are rather underrepresented in the dataset. As compared to the result of the first study where the platform was applied (Christelová et al. 2011), the resolution has greatly improved by enlarging the dataset, which is clearly demonstrated by the AA cv. clusters, or triploid subgroups diversification (see further below). Only future results comprising the whole diversity of the germplasm available for genotyping would give clear answer and show the borders for the application of the platform.

## Diploid cultivars

SSR analysis classified the large group of edible AA accessions into ten sets (highlighted in Table 3) with distinct positions on the SSR tree (Fig. 3). This not only helps in classification of these accessions, but also provides insights into the origin and putative progenitors of some triploid cultivars. The clustering reflects geographical origins as well. While the set of Island SE-Asian accessions (AA cv. ISEA 1) is connected to the *malaccensis* subgroup and the dessert AAA 'Ibota' subgroup in cluster III, the second ISEA block (AA cv. ISEA 2) co-localizes with a large spectrum of dessert AAAs within the cluster IX, suggesting that some of its members may have played a role in the constitution of cultivated AAAs. The great majority of these AA cvs. have sweet taste and are usually eaten raw.

The AA cv. African set in cluster IX contains AA cultivars found in East Africa and nearby islands of the Indian Ocean. It has been proposed that progenitors of this set (which most probably originated in the area between New Guinea, Borneo and Java and was brought to the East African coast through human migration; Perrier et al. 2009) were the diploid parents of the subgroups AAA 'Cavendish' and AAA 'Gros Michel' (Hippolyte et al. 2012). Our results support this notion. Similarly, the small group of Tongat AA cv.

may be the source of another diploid ancestor of AAA ‘Cavendish’ or ‘Gros Michel’ (Hippolyte et al. 2012).

The AA cv. IndonTriNG group of cultivars, in cluster IX, originated in New Guinea and within the Triangle shaped by the islands in Eastern Indonesia Halmahera (NE), Sulawesi (NW) and Sunda (South), that was recently explored for wild and edible bananas (Sutanto et al. 2016). The set contains a number of new accessions, of which some could have played a role in the formation of less known AAA cultivars. Three specimens from the Indonesian Triangle expedition clustered in a block of AA cultivars, which contains three Philippine cultivars (therefore called AA cv. IndonTriPh block in the cluster X). Its members produce bunches that vaguely resemble those of the AAB ‘Plantain’ (e.g. Roa Cakalang, strongly resembling the French plantain bunch) and ‘Plantain-linked’ (strikingly for Guyod and GabaGaba Putih; Sutanto et al. 2016).

Diploid AA cultivars that share the *banksii* background appear into two sets on the tree (Fig. 3, cluster XI). The first of them called the AA cv. *banksii* derivatives is on the same clade as the wild *M. acuminata* ssp. *banksii* and seems to contain direct/pure derivatives of the subspecies, all coming from New Guinea. The second block is termed AA cv. ‘*banksii* sensu lato’ and contains several wild AAs which are geographically proximal to ssp. *banksii* and ssp. *errans* in the Philippines, var. *acuminata* in Western New Guinea and Maluku and var. *tomentosa* in Sulawesi. The possibility that AA cultivars from this set are domesticated products/hybrids of these wild *acuminatas* with *banksii* deserves to be investigated. The AA cv. ‘Pisang Jari Buaya’ set (in cluster I) is separated from other edible AAs. This cultivar is used in breeding as a well-known source of resistance to pests i.e. burrowing nematodes (Wehunt et al. 1978; Marin et al. 1998; Ray 2002), resistance to Fusarium tropical race 4 disease and tolerance to black Sigatoka (Rowe 1998). However, based on our SSR analysis, its genetic affinity to wild and edible AA is unclear and deserves special attention. The AA cv. ‘Pisang Sapon’ singleton appears to be remotely linked to subspecies *zebrina* only, which could point to its being an edible derivative of *zebrina* rather than an inter-subspecies hybrid. However, low statistical support for this particular clade does not allow for conclusive remarks. The interspecific AS (*acuminata* x *schizocarpa*) hybrids are grouped on a separate clade within cluster X, closely linked to the *banksii*-dominant cluster XI (Fig. 3). *M. schizocarpa* (Cluster V) is sympatric with *M. acuminata* ssp. *banksii* (Cluster XI) in New Guinea and hybrids have been observed, but these are probably transient F1’s, except if maintained by vegetative propagation. Indeed, several members of this subcluster show the *banksii* morphology except for the split fruit peels typical for *schizocarpa* (Daniells et al. 2001).

Indian AB cultivars ‘Ney Poovan’ and ‘Kunnan’ were clearly differentiated in this work by SSR analysis, but no conclusion could be reached on the contribution of particular A genomes to the origin of this subgroup. On the other hand, three AB cultivars, ‘Muku Bugis’, ‘Mu’u Seribu’ and ‘Mu’u Pundi’ (collected in the Indonesian Triangle expedition; Sutanto et al. 2016) were located in the ‘Plantain-linked’ set of cluster XIII (see Online Resource 3), hence at a remote distance from the previous ABs ‘Ney Poovan’ and ‘Kunnan’. Therefore we wonder whether they may have delivered the B-genome to AA cultivars to form AABs such as ‘Laknau’ and ‘Iholena’.

### Triploid cultivars

The UPGMA clustering analysis separated the commonly recognized accessions into distinct clusters scattered over the tree with blocks of triploid edible accessions located among clusters of wild and cultivated diploids (Fig. 3, Online Resource 3). The clustering

pattern reflects the proportion of shared SSR-alleles between the accessions. Thus, the probable ancestral relationships among diploid progenitors and edible triploids can be hypothesized. Using a similar set of SSR markers, Hippolyte et al. (2012) showed the African 'Mlali' subgroup to be the closest 2n donor for 'Cavendish' and 'Gros Michel' subgroups. Our results confirm this concept as illustrated by the position of the block of African AA cv. (Fig. 3, cluster IX) within the clade of 'Cavendish' and 'Gros Michel' subgroups. At a broader scope, this clade is loosely linked with the AAB 'Pome' subgroup that was also proposed to share the eastern ancestry with the diploid African cv. 'Mlali' (Perrier et al. 2011). Moreover, *M. acuminata* 'Pisang Pipit', that was proposed to be related with the n-gamete donor of Cavendish bananas (Hippolyte et al. 2012) is part of the cluster, supporting its potential role in the ancestral relationships.

Similarly, our results support the conclusion of Hippolyte et al. (2012), that the AAA 'Ibota' subgroup (cluster III) originated from *malaccensis* subspecies (but missing the *banksii* contribution; Perrier et al. 2009) and further substantiate it with AA cultivars from Island SE Asia (AA cv. ISEA 1) as potential contributors to the formation of this triploid subgroup. Another AAA subgroup 'Lakatan' also called 'Pisang Berangan' (cluster IX) suggests the contribution of ISEA 2 edible AA cultivars. The poorly defined AAA 'Orotava' seems to be loosely connected with Vietnamese AA cvs 'Ayam' and 'Chuoi La Rung'.

While for the AAA subgroups 'Pisang Ambon', 'Leite' and 'Rio' (AAA Hetero on Fig. 3, cluster IX) no clear donors of the 2n and n-gamete were identified previously, our results suggest that their putative progenitors may come from the AA cvs. from the Indonesian Triangle including New Guinea (AA cv. IndonTriNG, Fig. 3). Moreover, our analysis proved that the AAA 'Pisang Ambon' subgroup (cluster IX; Online Resource 3 for details) is genetically distinct from the 'Gros Michel' subgroup, while the opposite has been suggested based on its morphological appearance and from the analysis of mitochondrial and plastid DNA (Carreel et al. 2002). Although for the 'Red' subgroup a *malaccensis* background was proposed (Perrier et al. 2009), our results (cluster IX, Fig. 3) do not support this assumption.

The four morphologically distinctive clone-sets of the African AAA 'Lujugira/Mutika' subgroup (Nfuuka, Musakala, Nakabululu, Nakitembe; Karamura 1998; Karamura and Pickersgill 1999) form a single coherent block in the cluster X of the SSR dendrogram. The existence of a previously assumed fifth clone set (Mbidde), containing only bananas for beer making, has recently been abandoned (Kitavi et al. 2016), its cultivars being morphologically and genetically indistinguishable from one or other of the aforementioned clone sets. Our results are in line with Kitavi et al. (2016) findings that the four clone sets are part of a single subgroup stemming from a single original AAA hybrid that has probably been formed in eastern Indonesia (Perrier et al. 2011) and diversified subsequently by multiple somatic mutations. A similar conclusion can be made for the subgroup of AAB 'African Plantains' that is subdivided into three clone types: Horn, False Horn and French, based on inflorescence morphology (De Langhe et al. 2005). Their phenotypic differentiation is not reflected by SSR analysis (cluster XIII, Fig. 3). The fact that these clone types cannot be separated into different subsets on the tree despite their striking morphological differentiation (Vuylsteke et al. 1996; Daniells et al. 2001) seems to confirm the fact that the differences are due to epigenetic changes, or resulted from somatic mutations occurring along the centuries of cultivation (Simmonds 1966; De Langhe et al. 1995; Crouch et al. 2000; Noyer et al. 2005).

The same cluster XIII contains subgroups 'Iholena' and the Philippine 'Laknau', with a bunch morphology reminding that of AAB Plantains, which are hence assembled under the

name AAB Plantain-linked. The ‘Iholena’ originated in mainland New Guinea (Perrier et al. 2009; Kagy et al. 2016), but its place in the ‘Plantain-linked’ set with accessions from eastern Indonesia points to a broader area of origin. Because of the proved *banksii* genetic background of both A genomes in the plantain AAB subgroup (Lebot et al. 1993; Carreel et al. 2002; Kagy et al. 2016), it is proposed that the Plantain-linked cultivars have the same A-background.

Interestingly, the AAB subgroup ‘Maoli-Popo’ulu’ is grouped together with the ABB subgroup ‘Saba’ and the ABB ‘Bluggoe-Monthan’ accessions within the cluster XII. This could point to the presence of the same A (presumably *banksii*) or B genome across all these accessions and indicates a need for more detailed investigation. Although the two components of the ‘Bluggoe-Monthan’ complex are considered to belong to different subgroups, their members are part of a single set on the tree. We therefore adopt the term ‘complex’ rather than ‘subgroup’. No genetic ground was obtained to differentiate between ‘Maoli’ and ‘Popo’ulu’ clone sets through the SSR analysis.

Similarly the cluster VIII contains a heterogeneous set of different genome-combinations. Representatives of the ABB triploid subgroup ‘Pisang Awak’ and the Indian AAB ‘Silk’ subgroup were reported to share the *malaccensis* A-genome (Perrier et al. 2009), but this conclusion is weakly supported by our results. Interestingly, two supposedly Silk accessions—verified triploids cultivated in East Africa—are on the same branch as AB ‘Kunnan’. The possible genotypic relation between ‘Silk’ and ‘Kunnan’ warrants further examination.

‘Pisang Raja(h)’ is the only AAB set in cluster X and its position may point to the dominant presence of *banksii* in its A genomes. However, the name ‘Raja’ is frequently applied to cultivars with different morphology as a mark of superior quality characteristics, hence a much larger number of candidate ‘P. Raja(h)’ cultivars should be analyzed to examine this and exclude potential mislabeling.

## Conclusions

Recently, plant genetic diversity has been frequently assessed with high-throughput next-generation sequencing methods, such as Genotyping by Sequencing (GBS) or SNP-based DNA chips (e.g. DArT or DArTSeq). Although such methods are powerful and offer low costs per data-point, they may not be suitable for long-term projects in which entries need to be analyzed in batches, or even one by one over extended periods of time. The comparability of results gathered at different time points cannot be assured unless the whole set of accessions is analyzed again, which may not be economical. In contrast, SSR genotyping allows processing even single samples over time. Moreover, with each new accession, new alleles may be added into the dataset thereby increasing the resolution of the analysis and improving the classification. The results of the present large-scale molecular characterization effort are essential for the management and conservation of the global *Musa* germplasm collection and will facilitate its efficient use by the banana research and breeding community. The genotyping platform used in this work identified potentially problematic accessions for which field or other verification were deemed necessary and it has also helped in proposing subgroup classification for corrections. Coupled to the flow cytometric estimation of ploidy level, this genotyping system presents the new standard for molecular characterization of *Musa* genebank accessions. Moreover, the results of the study provide a detailed picture of genetic diversity of available *Musa*

germplasm and lay a solid basis for a more focused molecular and morphological evaluation of certain, until now less studied groups.

**Acknowledgements** We thank Marie Seifertová for her excellent technical assistance and Ludvík Urda and Alžběta Němečková for assisting with the flow cytometry measurements. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic (Grant award LO1204 from the National Program of Sustainability I and awards INGO II No. LG12021 and INGO II No. LG15017). The authors would furthermore like to thank all donors who supported this work through their contributions to the CGIAR Fund (<http://www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2/>), and in particular to the CGIAR Research Program for Managing and Sustaining Crop Collections (Genebanks-CRP). We also thank Vincent Johnson for his kind review and editing before submission. We dedicate this article to the late Dr. Pierre Jean Laurent Lagoda who has pioneered the use of SSR markers in *Musa* genetics.

#### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Argent GCG (1976) The wild bananas of Papua New Guinea. *Notes R Bot Gard Edinburgh* 35(1):77–114
- Bartoš J, Alkhimova O, Doleželová M, De Langhe E, Doležel J (2005) Nuclear genome size and genomic distribution of ribosomal DNA in *Musa* and *Ensete* (Musaceae): taxonomic implications. *Cytogenet Genome Res* 109(1–3):50–57. doi:[10.1159/000082381](https://doi.org/10.1159/000082381)
- Boonruangrod R, Desai D, Fluch S, Berenyi M, Burg K (2008) Identification of cytoplasmic ancestor gene-pools of *Musa acuminata* Colla and *Musa balbisiana* Colla and their hybrids by chloroplast and mitochondrial haplotyping. *Theor Appl Genet* 118(1):43–55. doi:[10.1007/s00122-008-0875-3](https://doi.org/10.1007/s00122-008-0875-3)
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am J Hum Genet* 32(3):314–331
- Carreel F, Gonzalez de Leon D, Lagoda P, Lanaud C, Jenny C, Horry J, du Montcel HT (2002) Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome* 45(4):679–692. doi:[10.1139/G02-033](https://doi.org/10.1139/G02-033)
- Chase R, Sardos J, Ruas M, Van den Houwe I, Roux N, Hribova E, Dolezel J (2016) The field verification activity: a cooperative approach to the management of the global *Musa* in vitro collection at the International Transit Centre. *Acta Hort* 1114:61–66. doi:[10.17660/ActaHortic.2016.1114.8](https://doi.org/10.17660/ActaHortic.2016.1114.8)
- Cheesman EE (1947) Classification of the bananas. The genus *Ensete* Horan and the genus *Musa* L. *Kew Bull* 2(2):97–117
- Chin EC, Senior ML, Shu H, Smith JSC (1996) Maize simple repetitive DNA sequences: abundance and allele variation. *Genome* 38:866–873. doi:[10.1139/g96-109](https://doi.org/10.1139/g96-109)
- Christelová P, Valárik M, Hřibová E, Van den Houwe I, Channelière S, Roux N, Doležel J (2011) A platform for efficient genotyping in *Musa* using microsatellite markers. *AoB PLANTS*. doi:[10.1093/aobpla/plr024](https://doi.org/10.1093/aobpla/plr024)
- Côte FX, Sandoval JA, Marie P, Auboiron E (1993) Variation in micropropagated bananas and plantains: literature survey. *Fruits* 48(1):15–22
- Cregan PB, Jervik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VonToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated molecular genetic linkage map of soybean genome. *Crop Sci* 39:1464–1490
- Crouch HK, Crouch JH, Jarret RL, Cregan PB, Ortiz R (1998) Segregation of microsatellite loci in haploid and diploid gametes of *Musa*. *Crop Sci* 38:211–217. doi:[10.2135/cropsci1998.0011183X003800010035x](https://doi.org/10.2135/cropsci1998.0011183X003800010035x)
- Crouch HK, Crouch JH, Madsen S, Vuylsteke DR, Ortiz R (2000) Comparative analysis of phenotypic and genotypic diversity among plantain landraces (*Musa* spp., AAB group). *Theor Appl Genet* 101(7):1056–1065. doi:[10.1007/s001220051580](https://doi.org/10.1007/s001220051580)

- da Cunha CP, Resende FV, Zucchi MI, Pinheiro JB (2014) SSR-based genetic diversity and structure of garlic accessions from Brazil. *Genetica* 142(5):419–431. doi:[10.1007/s10709-014-9786-1](https://doi.org/10.1007/s10709-014-9786-1)
- Daniells J, Jenny C, Karamura D, Tomekpe K (2001) Musalogue: Diversity in the genus *Musa*. A catalogue of *Musa* germplasm. IPGRI, Rome
- Davey MW, Gudimella R, Harikrishna JA, Sin LW, Khalid N, Keulamans J (2013) A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. *BMC Genom* 14:683. doi:[10.1186/1471-2164-14-683](https://doi.org/10.1186/1471-2164-14-683)
- De Jesus ON, Silva SDE, Amorim EP, Ferreira CF, de Campos JMS, Silva GD, Figueira A (2013) Genetic diversity and population structure of *Musa* accessions in ex situ conservation. *BMC Plant Biol* 13:41. doi:[10.1186/1471-2229-13-41](https://doi.org/10.1186/1471-2229-13-41)
- De Langhe E, Swennen R, Vuylsteke D (1995) Plantain in the Early Bantu World. Sutton J, editor. *Azania XXIX–XXX 1994–1995*. Special volume on ‘The Growth of Farming Communities in Africa from the Equator Southwards. Proceedings of a conference of the British Institute in Eastern Africa. Cambridge. 4–8 July 1994. pp. 147–160
- De Langhe E, Pillay M, Tenkouano A, Swennen R (2005) Integrating morphological and molecular taxonomy in *Musa*: the African plantains (*Musa* spp. AAB group). *Plant Syst Evol* 255(34):225–236. doi:[10.1007/s00606-005-0346-0](https://doi.org/10.1007/s00606-005-0346-0)
- De Langhe E, Vrydaghs L, de Maret P, Perrier X, Denham T (2009) Why bananas matter: an introduction to the history of banana domestication. *Ethnobot Res Appl* 7:165–178. doi:[10.17348/era.7.0.165-177](https://doi.org/10.17348/era.7.0.165-177)
- De Langhe E, Hřibová E, Carpentier S, Doležel J, Swennen R (2010) Did backcrossing contribute to the origin of hybrid edible bananas? *Ann Bot* 106(6):849–857. doi:[10.1093/aob/mcq187](https://doi.org/10.1093/aob/mcq187)
- Denham TP, Haberler SG, Lentfer C, Fullagar R, Field J, Therin M, Porch N, Winsborough B (2003) Origins of agriculture at Kuk Swamp in the highlands of New Guinea. *Science* 301(5630):189–193. doi:[10.1126/science.1085255](https://doi.org/10.1126/science.1085255)
- D’Hont A, Denoëuf F, Aury JM et al (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488(7410):213–217. doi:[10.1038/nature11241](https://doi.org/10.1038/nature11241)
- Doležel J, Doleželová M, Novák FJ (1994) Flow cytometric estimation of nuclear-DNA amount in diploid bananas (*Musa acuminata* and *Musa balbisiana*). *Biol Plantarum* 36(3):351–357. doi:[10.1007/BF02920930](https://doi.org/10.1007/BF02920930)
- Doležel J, Lysák MA, Doleželová M, Roux N (1997) Use of flow cytometry for rapid ploidy determination in *Musa* species. *InfoMusa* 6:6–9
- Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. *Nat Protoc* 2(9):2233–2244. doi:[10.1038/nprot.2007.310](https://doi.org/10.1038/nprot.2007.310)
- Doleželová M, Doležel J, van Den Houwe I, Roux N, Swennen R (2005) Focus on the *Musa* collection: ploidy levels revealed. *InfoMusa* 14:34–36
- FAOSTAT. <http://faostat3.fao.org/home/index.html>
- Fjellheim S, Tanhuanpaa P, Marum P, Manninen O, Rognli OA (2015) Phenotypic or molecular diversity screening for conservation of genetic resources? An Example from a Genebank Collection of the Temperate Forage Grass Timothy. *Crop Sci* 55(4):1646–1659. doi:[10.2135/cropsci2014.12.0825](https://doi.org/10.2135/cropsci2014.12.0825)
- Galbraith DW, Doležel J, Lambert G, Macas J (1998) Nuclear DNA content and ploidy analyses in higher plants. In: Robinson JP (ed) *Current protocols in cytometry*. Wiley, New York, pp 7.6.1–7.6.22
- Ge XJ, Liu MH, Wang WK, Schaal BA, Chiang TY (2005) Population structure of wild bananas, *Musa balbisiana*, in China determined by SSR fingerprinting and cpDNA PCR-RFLP. *Mol Ecol* 14(4):933–944. doi:[10.1111/j.1365-294X.2005.02467.x](https://doi.org/10.1111/j.1365-294X.2005.02467.x)
- Goldstein DB, Schlotterer C (1999) *Microsatellites: evolution and applications*. Oxford University Press, Oxford
- Gross-German E, Viruel MA (2013) Molecular characterization of avocado germplasm with a new set of SSR and EST-SSR markers: genetic diversity, population structure, and identification of race-specific markers in a group of cultivated genotypes. *Tree Genet Genomes* 9(2):539–555. doi:[10.1007/s11295-012-0577-5](https://doi.org/10.1007/s11295-012-0577-5)
- Häkkinen M (2013) Reappraisal of sectional taxonomy in *Musa* (Musaceae). *Taxon* 62(4):809–813
- Häkkinen M, Hong W, Ge XJ (2008) *Musa itinerans* (Musaceae) and its intraspecific taxa in China. *Novon* 18(1):50–60. doi:[10.3417/2006162](https://doi.org/10.3417/2006162)
- Hayden MJ, Nguyen TM, Waterman A, McMichael GL, Chalmers KJ (2007) Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. *Mol Breeding* 21(3):271–281. doi:[10.1007/s11032-007-9127-5](https://doi.org/10.1007/s11032-007-9127-5)
- Hippolyte I, Bakry F, Seguin M, Gardes L, Rivallan R, Risterucci AM, Jenny C, Perrier X, Carreel F, Argout X, Piffanelli P, Khan IA, Miller RNG, Pappas GJ, Mbéguié-A-Mbéguié D, Matsumoto T, De Bernardinis V, Huttner E, Kilian A, Baurens FC, D’Hont A, Cotel F, Courtois B, Glaszmann JC (2010)



- A saturated SSR/DARt linkage map of *Musa acuminata* addressing genome rearrangements among bananas. *BMC Plant Biol* 10:65. doi:[10.1186/1471-2229-10-65](https://doi.org/10.1186/1471-2229-10-65)
- Hippolyte I, Jenny C, Gardes L, Bakry F, Rivallan R, Pomies V, Cubry P, Tomekpe K, Risterucci AM, Roux N, Rouard M, Arnaud E, Kolesnikova-Allen M, Perrier X (2012) Foundation characteristics of edible *Musa* triploids revealed from allelic distribution of SSR markers. *Ann Bot* 109(5):937–951. doi:[10.1093/aob/mcs010](https://doi.org/10.1093/aob/mcs010)
- Hřibová E, Čížková J, Christelová P, Taudien S, De Langhe E, Doležel J (2011) The ITS1-5.8S-ITS2 sequence region in the Musaceae: structure, diversity and use in molecular phylogeny. *PLoS ONE* 6(3):e17863. doi:[10.1371/journal.pone.0017863](https://doi.org/10.1371/journal.pone.0017863)
- [IPGRI-INIBAP/CIRAD] International Plant Genetic Resources Institute-International Network for the Improvement of Banana and Plantain/Centre de Coopération internationale en recherche agronomique pour le développement (1996) Description for Banana (*Musa* spp.). Int. Network for the Improvement of Banana and Plantain, Montpellier, France; Centre de coopération int. en recherche agronomique pour le développement, Montpellier, France; IPGRI Press, Rome
- Janssens SB, Vandeloock F, De Langhe E, Verstraete B, Smets E, Van den Houwe I, Swennen R (2016) Evolutionary dynamics and biogeography of Musaceae reveal a correlation between the diversification of the banana family and the geological and climatic history of Southeast Asia. *New Phytol* 210(4):1453–1465. doi:[10.1111/nph.13856](https://doi.org/10.1111/nph.13856)
- Jing R, Vershinin A, Grzebyta J, Shaw P, Smýkal P, Marshall D, Ambrose MJ, Ellis TH, Flavell AJ (2010) The genetic diversity and evolution of field pea (*Pisum*) studied by high throughput retrotransposon based insertion polymorphism (RBIP) marker analysis. *BMC Evol Biol* 10:44. doi:[10.1186/1471-2148-10-44](https://doi.org/10.1186/1471-2148-10-44)
- Kagy V, Wong M, Vandenbroucke H, Jenny C, Dubois C, Ollivier A, Cardi C, Mournet P, Tuia V, Roux N, Doležel J, Perrier X (2016) Traditional Banana Diversity in Oceania: an endangered heritage. *PLoS ONE* 11(3):e0151208. doi:[10.1371/journal.pone.0151208](https://doi.org/10.1371/journal.pone.0151208)
- Karamura D (1998) Numerical taxonomic studies of the East African highland bananas (*Musa* AAA—East Africa) in Uganda. INIBAP, Montpellier
- Karamura D, Pickersgill B (1999) A classification of the clones of East African Highland bananas (*Musa*) found in Uganda. *Plant Genetic Resources Newsletter* 119:1–6
- Kitavi M, Downing T, Lorenzen J, Karamura D, Onyango M, Nyine M, Ferguson M, Spillane C (2016) The triploid East African Highland Banana (EAHB) gene pool is genetically uniform arising from a single ancestral clone that underwent population expansion by vegetative propagation. *Theor Appl Genet* 129(3):547–561. doi:[10.1007/s00122-015-2647-1](https://doi.org/10.1007/s00122-015-2647-1)
- Lagoda PJL, Noyer JL, Dambier D, Baurens FC, Grapin A, Lanaud C (1998) Sequence tagged microsatellite site (STMS) markers in the Musaceae. *Mol Ecol* 7(5):659–663
- Lebot V, Aradhya KM, Manshardt R, Meilleur B (1993) Genetic relationships among cultivated bananas and plantains from Asia and the Pacific. *Euphytica* 67(3):163–175. doi:[10.1007/BF00040618](https://doi.org/10.1007/BF00040618)
- Li LF, Häkkinen M, Yuan YM, Hao G, Ge XJ (2010) Molecular phylogeny and systematics of the banana family (Musaceae) inferred from multiple nuclear and chloroplast DNA fragments, with a special reference to the genus *Musa*. *Mol Phylogenet Evol* 57(1):1–10. doi:[10.1016/j.ympev.2010.06.021](https://doi.org/10.1016/j.ympev.2010.06.021)
- Liu K, Muse SV (2005) PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics* 21(9):2128–2129. doi:[10.1093/bioinformatics/bti282](https://doi.org/10.1093/bioinformatics/bti282)
- Liu QW, Song Y, Liu L, Zhang MY, Sun JM, Zhang SL, Wuuu J (2015) Genetic diversity and population structure of pear (*Pyrus* spp.) collections revealed by a set of core genome-wide SSR markers. *Tree Genet Genomes*. 11(6):128. doi:[10.1007/s11295-015-0953-z](https://doi.org/10.1007/s11295-015-0953-z)
- Marin DH, Sutton TB, Barker KR, Kaplan DT, Opperman CH (1998) Burrowing nematode resistance of black sigatoka resistant hybrids. *Nematropica* 28(2):241–247
- Martin G, Baurens FC, Droc G, Rouard M, Cenci A, Hastie A, Doležel J, Aury JM, Alberti A, Carreel F, D’Hont A (2016) Improvement of the banana “*Musa acuminata*” reference sequence using NGS data and semi-automated bioinformatics methods. *BMC Genom* 17:243. doi:[10.1186/s12864-016-2579-4](https://doi.org/10.1186/s12864-016-2579-4)
- Mason AS, Zhang J, Tollenaere R, Teuber PV, Dalton-Morgan J, Hu L, Yan GJ, Edwards D, Redden R, Batley J (2015) High-throughput genotyping for species identification and diversity assessment in germplasm collections. *Mol Ecol Res* 15(5):1091–1101. doi:[10.1111/1755-0998.12379](https://doi.org/10.1111/1755-0998.12379)
- Mbanjo EGN, Tchoumboungang F, Mouelle AS, Oben JE, Nyine M, Dochez C, Ferguson ME, Lorenzen J (2012) Molecular marker-based genetic linkage map of a diploid banana population (*Musa acuminata* Colla). *Euphytica* 188(3):369–386. doi:[10.1007/s10681-012-0693-1](https://doi.org/10.1007/s10681-012-0693-1)
- Michener CD, Sokal RR (1957) A quantitative approach to a problem of classification. *Evolution* 11(2):130–162. doi:[10.2307/2406046](https://doi.org/10.2307/2406046)



- Nasution RE (1991) Taxonomic study of the *Musa acuminata* Colla with its intraspecific taxa in Indonesia. *Memoirs of Tokyo University of Agriculture* 32:1–122
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70(12):3321–3323. doi:[10.1073/pnas.70.12.3321](https://doi.org/10.1073/pnas.70.12.3321)
- Nicolai M, Cantet M, Lefebvre V, Sage-Palloix AM, Palloix A (2013) Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types. *Genet Resour Crop Evol* 60(8):2375–2390. doi:[10.1007/s10722-013-0006-0](https://doi.org/10.1007/s10722-013-0006-0)
- Noyer JL, Causse S, Tomekpe K, Bouet A, Baurens FC (2005) A new image of plantain diversity assessed by SSR, AFLP and MSAP markers. *Genetica* 124(1):61–69. doi:[10.1007/s10709-004-7319-z](https://doi.org/10.1007/s10709-004-7319-z)
- Ortiz R, Swennen R (2014) From crossbreeding to biotechnology-facilitated improvement of banana and plantain. *Biotechnol Adv* 32(1):158–169. doi:[10.1016/j.biotechadv.2013.09.010](https://doi.org/10.1016/j.biotechadv.2013.09.010)
- Otto F (1990) DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In: Crissman HA, Darzynkiewicz Z, editors. *Methods in Cell Biology*. Academic Press, New York, pp. 33:105–110
- Panis B (2009) Cryopreservation of *Musa* germplasm. In: Engelmann F, Benson E (eds) *Technical guidelines No 9, 2nd edn*. Bioversity International, Montpellier
- Panis B, Piette B, Swennen R (2005) Droplet vitrification of apical meristems: a cryopreservation protocol applicable to all Musaceae. *Plant Sci* 168(1):45–55. doi:[10.1016/j.plantsci.2004.07.022](https://doi.org/10.1016/j.plantsci.2004.07.022)
- Perrier X, Bakry F, Carreel F, Jenny C, Horry JP, Lebot V, Hippolyte I (2009) Combining biological approaches to shed light on the evolution of edible bananas. *Ethnobot Res Appl* 7:199–216. doi:[10.17348/era.7.0.199-216](https://doi.org/10.17348/era.7.0.199-216)
- Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, Bakry F, Carreel F, Hippolyte I, Horry JP, Jenny C, Lebot V, Risterucci AM, Tomekpe K, Doutrelepont H, Ball T, Manwaring J, de Maret P, Denham T (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc Natl Acad Sci USA* 108(28):11311–11318. doi:[10.1073/pnas.1102001108](https://doi.org/10.1073/pnas.1102001108)
- Pillay M, Oguniwin E, Tenkouano A, Dolezel J (2006) Ploidy and genome composition of *Musa* germplasm at the International Institute of Tropical Agriculture (IITA). *Afr J Biotechnol* 5(13):1224–1232
- Ray PK (2002) *Breeding tropical and subtropical fruits*. Springer Verlag, Berlin
- Risterucci AM, Hippolyte I, Perrier X, Xia L, Caig V, Evers M, Huttner E, Kilian A, Glaszmann JC (2009) Development and assessment of diversity arrays technology for high-throughput DNA analyses in *Musa*. *Theor Appl Genet* 119(6):1093–1103. doi:[10.1007/s00122-009-1111-5](https://doi.org/10.1007/s00122-009-1111-5)
- Roder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of wheat. *Genetics* 149(4):2007–2023
- Rodrigues PHV, Tulmano Neto A, Cassieri Neto P, Mendes BMJ (1998) Influence of the number of subcultures on somaclonal variation in micropropagated Nanição (*Musa* spp., AAA group). *Acta Hort* 490:469–474. doi:[10.17660/ActaHortic.1998.490.49](https://doi.org/10.17660/ActaHortic.1998.490.49)
- Roux N, Toloza A, Radecki Z, Zapata-Arias FJ, Doležel J (2003) Rapid detection of aneuploidy in *Musa* using flow cytometry. *Plant Cell Rep* 21(5):483–490. doi:[10.1007/s00299-002-0512-6](https://doi.org/10.1007/s00299-002-0512-6)
- Rowe PR (1998) A banana breeder's response to "The global programme for *Musa* improvement (PRO-MUSA)", article published by the CGIAR. *InfoMusa* 7(1):1–4
- Rowe PR, Rosales FE (1996) Bananas and plantains. In: Janick J, Moore JN (eds) *Fruit breeding, vol I., Tree and tropical fruits* Wiley, New York, pp 167–211
- Roy S, Mamdi BC, Mawkhleng B, Banerjee A, Yadav RM, Misra AK, Bansal KC (2016) Genetic diversity and structure in hill rice (*Oryza sativa* L.) landraces from the North-Eastern Himalayas of India. *BMC Genet* 14:107. doi:[10.1186/s12863-016-0414-1](https://doi.org/10.1186/s12863-016-0414-1)
- Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol Lett* 9(5):615–629. doi:[10.1111/j.1461-0248.2006.00889.x](https://doi.org/10.1111/j.1461-0248.2006.00889.x)
- Shepherd K (1999) Cytogenetics of the genus *Musa*. International Network for the Improvement of Banana and Plantain, Montpellier
- Simmonds NW (1962) *The evolution of the bananas*. Longmans, London
- Simmonds NW (1966) *Bananas*. Longmans, London
- Simmonds NW, Shepherd K (1955) The taxonomy and origin of the cultivated bananas. *Bot J Linn Soc* 55(359):302–312. doi:[10.1111/j.1095-8339.1955.tb00015.x](https://doi.org/10.1111/j.1095-8339.1955.tb00015.x)
- Sutanto A, Edison HS, Riska Nasution F, Hermanto C, Čížková J, Hřibová E, Doležel J, Roux N, Horry JP, Daniells JW, De Langhe E (2016) Collecting banana diversity in Eastern Indonesia. *Acta Hort* 1114:19–25. doi:[10.17660/ActaHortic.2016.1114.3](https://doi.org/10.17660/ActaHortic.2016.1114.3)
- Swangpol S, Volkaert H, Sotto RC, Seelanan T (2007) Utility of selected non-coding chloroplast DNA sequences for lineage assessment of *Musa* interspecific hybrids. *J Biochem Mol Biol* 40(4):577–587

- Tenkouano A, Vuylsteke D, Okoro J, Makumbi D, Swennen R, Ortiz R (2003) Diploid banana hybrids TMB25105–1 and TMB29128–3 with good combining ability, resistance to Black Sigatoka and nematodes. *Hortic Sci* 38(3):468–472
- Ude G, Pillay M, Nwakanma D, Tenkouano A (2002) Genetic diversity in *Musa acuminata* Colla and *Musa balbisiana* Colla and some of their natural hybrids using AFLP Markers. *Theor Appl Genet* 104(8):1246–1252. doi:[10.1007/s00122-002-0914-4](https://doi.org/10.1007/s00122-002-0914-4)
- Van den Houwe I, De Smet K, Montcel HTD, Swennen R (1995) Variability in storage potential of banana shoot cultures under medium term storage conditions. *Plant Cell Tiss Org* 42(3):269–274
- Vuylsteke D, Swennen R, Ortiz R (1993) Development and performance of Black Sigatoka-resistant tetraploid hybrids of plantain (*Musa* spp., AAB group). *Euphytica* 65(1):33–42. doi:[10.1007/BF00022197](https://doi.org/10.1007/BF00022197)
- Vuylsteke D, Swennen R, De Langhe E (1996) Field performance of somaclonal variants of plantain (*Musa* spp., AAB Group). *J Am Soc Hortic Sci* 121(1):42–46
- Wang XL, Chiang TY, Roux N, Hao G, Ge XJ (2007) Genetic diversity of wild banana (*Musa balbisiana* Colla) in China as revealed by AFLP markers. *Genet Res Crop Evol* 54(5):1125–1132. doi:[10.1007/s10722-006-9004-9](https://doi.org/10.1007/s10722-006-9004-9)
- Wehunt EJ, Hutchinson DJ, Edwards DI (1978) Reaction of banana cultivars to burrowing nematode (*Radopholus similis*). *J Nematol* 10(4):368–370
- Weising K, Nybom H, Wollf K, Kahl G (2005) DNA fingerprinting in plants. Principles, methods, and applications, 2nd edn. Taylor & Francis, London