Genotyping the local banana landrace groups of East Africa

D. Karamura[1], M. Kitavi[2], M. Nyine[3], D. Ochola[1], W. Ocimati[1], S. Muhangi[4], D. Talengera[4] and E.B. Karamura[1]

[1]Bioversity International, Kampala, Uganda; [2]International Institute of Tropical Agriculture (IITA), Biosciences eastern and central Africa (BecA-ILRI), Nairobi, Kenya; [3]International Institute of Tropical Agriculture (IITA), Kampala, Uganda; [4]National Agricultural Research Organization, Box 7065, Kampala, Uganda.

Abstract

Crop landraces (largely resulting from adaptation and continuous selection by farmers) are more diverse within field populations than modern cultivars (produced by deliberate crossing), yet their distribution has continued to shrink in the past decades. The temporal dynamics of this shrinking is little known. The analysis of genetic variation within and between landraces is essential for making efficient breeding and conservation decisions with the available variability. Seven diploid landraces originally from Tanzania, 37 triploid landraces (24 East African highland bananas (EAHB); 5 'Ilalyi' (AAA genome), and 8 ill-defined types from Tanzania), 6 exotic triploids, and 3 exotic diploids originally from the International Transit Center were genotyped with simple sequence repeat (SSR) markers. This study sought to understand the genetic relationship between the EAHB and the diploid landraces and other banana groups (local triploid landraces, and introduced (exotic) cultivars) so as to decide whether to include the diploids in the breeding scheme of EAHB. Results showed the highest average genetic distance (degree of genomic difference by proportion) within the diploids (0.5666), followed by the hybrid triploids (0.4568) and the lowest within the 'Ilalyi' (0.0748) and the EAHB (0.0827) landraces. The variation within each clone set of EAHB was higher in 'Nakitembe' (0.0948) and 'Musakala' (0.1052). These two are commercial clone sets whose variation may be due to high and long-term selection pressure. In contrast, between the banana groups, the diploid landraces were more distant (highest average genetic distance) from the triploid landraces (0.4351-0.4430) and could thus provide useful breeding traits. On the other side, the triploid landraces had a narrow genetic base which should be broadened. Results did not identify those local east African diploids closest to the EAHB or other local triploids, although local diploids show breeding potential. This could widen the genetic base and probably improve performance of the triploid landraces.

Keywords: diploids, exotic, genetic distance, triploids, variation

INTRODUCTION

East Africa, especially the area stretching from the west to the north of Lake Victoria, is regarded as a secondary center of diversity for bananas (Simmonds, 1966). Here the East African highland (cooking and beer) bananas (EAHB) and other landraces constitute key components to food security, livelihoods and agricultural sustainability. However, many pests and diseases have significantly affected banana cultivation for the past 20 years. As a consequence of these threats, there has been renewed interest in banana breeding since then (Crouch et al., 1997), with the goal of improving cultivars for local consumption in the tropics (Ortiz and Vuylsteke, 1996). Due to difficulties associated with conventional breeding of edible bananas, efforts have been made to collect diploid wild *Musa* relatives and landraces for use in crop improvement. Diploid wild relatives have been found to be vital to banana improvement programs because of their high levels of male and female fertility, and low levels of heterozygosity. However, breeding from wild diploids is lengthy, as it requires several cycles to reduce the undesirable traits that tend to also be inherited by the resultant

progenies. Such hybrids also face challenges of acceptance by farmers and consumers. The use of local East African diploids in the breeding programs of the triploid EAHB would not only be considerably more cost effective, but may also produce the desired result a lot faster. However, not all the East African diploids have been characterized, and their potential resistance to diseases and pests has not been clearly defined, although they seem to be free from a number of diseases.

There are about three types of triploid East African banana landraces: the EAHB, the 'Ilalyi' (AAA genome) in the Pare-Usambara region in Tanzania, and the ill-defined 10-15 astringent triploids found mainly in Tanzania. There are also three groups of diploid East African banana landraces: the Mshare found in the Kilimanjaro in Tanzania, the Muraru in the Gikuyu highlands of Kenya, and 8-10 ill-defined *Musa* AA diploids (NARO, 2012; De Langhe et al., 2001, 2002). Representative samples of each of these landraces are being preserved in the East and Central African regional banana field gene-bank at Mbarara and hence were accessed from here. The ill-defined diploid landraces need to be characterized with respect to how they could be related to triploid EAHB and other East African triploid landraces if they are to be used in breeding. It is hypothetically assumed that there should be a close phylogenetic relationship between East African *Musa* diploids and triploid cultivars since they all occupy the same niche (East African plateau). The main objective of this study was to understand the partitioning of genetic diversity among the main banana landraces of East Africa using the simple sequence repeats (SSR) marker system. It was also the aim of this study to understand the genetic relationships between the EAHB and other triploid landraces ('Ilalyi' and the ill-defined triploids) to the diploid landraces and other bananas that have been introduced in East Africa (exotic) so that a decision can be made of whether to include the local East African diploids in the breeding scheme of EAHB. Knowledge of the genetic base and the genetic relatedness of the ill-defined diploids to the rest of the local triploids would enable the determination of their usefulness to banana crop improvement.

Several different techniques of DNA fingerprinting have been developed during the last decade, each with specific advantages and disadvantages. SSR or microsatellite polymorphism has been known to be a potentially powerful technique of DNA fingerprinting and has been successful in the amplification of tandem repeat sequences, known to be polymorphic and widespread in plant genomes (Cregan, 1992). It was therefore considered to be more efficient in establishing the available diversity in the East African landraces in this study.

MATERIALS AND METHODS

In this study, 53 accessions were genotyped with SSR markers. The genotyped materials included 7 diploid landraces originally from Tanzania, 37 triploid landraces (24 East African highland bananas (EAHB); 5 'Ilalyi', 8 ill-defined types from Tanzania), 6 exotic triploids (one plantain, Mysore, Cavendish, Gros Michel, Red and Sukali Ndizi), and 3 exotic diploids originally from the International Transit Center (ITC). The ploidy level of most of these materials had already been determined by flow cytometry or their ploidy level was known through literature (NARO, 2012; De Langhe et al., 2001, 2002). All the materials were accessed from the East African regional collection except two of the exotic diploids (zebrina and banksii) which were accessed from the International Transit Center, Belgium (Table 1). Sixty-four SSR markers were used to assess the level of genetic variability or relatedness within and between these materials. The 64 polymorphic SSR primer pairs were chosen from a pool of over 300 previously optimized SSR primer pairs developed by CIRAD/IITA (Kolesnikova and Kouassi, 2006; Mbanjo et al., 2009).

Table 1. *Musa* accessions used in the study.

Sample ID	Accession name	Genome group	Category
BU01	Halahala	AA	Diploid landrace
BU02	Mjenga	AA	Diploid landrace
BU03	Mlambichi	AA	Diploid landrace
BU04	Paka	AA	Diploid landrace
BU05	Ndyali	AA	Diploid landrace
BU06	Mwitupemba	AA	Diploid landrace
BU07	Muvubo	AAA	Triploid landrace (EAHB)
BU08	Namunwe	AAA	Triploid landrace (EAHB)
BU09	Mukazi-alanda	AAA	Triploid landrace (EAHB)
BU10	Musakala	AAA	Triploid landrace (EAHB)
BU11	Mpologoma	AAA	Triploid landrace (EAHB)
BU12	Enyoya	AAA	Triploid landrace (EAHB)
BU13	Kibuzi	AAA	Triploid landrace (EAHB)
BU14	Nakyetengu	AAA	Triploid landrace (EAHB)
BU15	Keitabunyonyi	AAA	Triploid landrace (EAHB)
BU16	Nakabululu	AAA	Triploid landrace (EAHB)
BU17	Nakasabira	AAA	Triploid landrace (EAHB)
BU18	Salalugazi	AAA	Triploid landrace (EAHB)
BU19	Namaliga	AAA	Triploid landrace (EAHB)
BU20	Ingoromora	AAA	Triploid landrace (EAHB)
BU21	Mbwazirume	AAA	Triploid landrace (EAHB)
BU22	Nakitembe	AAA	Triploid landrace (EAHB)
BU23	Oruhuna	AAA	Triploid landrace (EAHB)
BU24	Nabuyobo	AAA	Triploid landrace (EAHB)
BU25	Enzirabushera	AAA	Triploid landrace (EAHB)
BU26	Nakinyika	AAA	Triploid landrace (EAHB)
BU27	Siira	AAA	Triploid landrace (EAHB)
BU28	Nfuuka	AAA	Triploid landrace (EAHB)
BU29	Nante	AAA	Triploid landrace (EAHB)
BU30	Enyamashari	AAA	Triploid landrace (EAHB)
BU31	Suu	AAA	Triploid landrace('Ilalyi')
BU32	Kitarasa	AAA	Triploid landrace(Ilalyi)
BU33	Haahaa	AAA	Triploid landrace('Ilalyi')
BU34	Ilalyi	AAA	Triploid landrace('Ilalyi')
BU35	Mlema	AAA	Triploid landrace('Ilalyi')
BU36	Ntindi II	AAA	Triploid landrace (ill-defined)
BU37	Ntindi I	AAA	Diploid landrace
BU38	Ntebwe	AAA	Triploid landrace (ill-defined)
BU39	Bura	AAA	Triploid landrace (ill-defined)
BU40	Kikundi	AAA	Triploid landrace (ill-defined)
BU41	Luholele	AAA	Triploid landrace (ill-defined)
BU42	Diana	AAA	Triploid landrace (ill-defined)
BU43	Kitombo	AAA	Triploid landrace (ill-defined)
BU44	Munyamimbwa	AAA	Exotic triploid (ill- defined)
BU45	Poyo	AAA	Exotic triploid (Cavendish)
BU46	Kabila	AAB	Exotic triploid (plantain)
BU47	Mysore	AAB	Exotic triploid
BU48	Mzungu Mwekundu	AAA	Exotic triploid (Red)
BU49	Gros michel	AAA	Exotic triploid
BU50	Sukali ndizi	AAB	Exotic triploid
BU51	Calcutta 4	AA	Exotic diploid (ITC 0249)
MC52	Banksii	AA	Exotic diploid (ITC 0896)
MC53	Zebrina	AA	Exotic diploid (ITC 0728)

Amplification procedures, data generation and analysis

PCR reactions were performed on a Gene-Amp PCR system 9700 using the 384-well thermocycler (Applied Biosystems, Foster City, CA, USA) in a total reaction of volume of 10 µL. The PCR reactions contained 1× standard *Taq* buffer with $MgCl_2$, 0.2 mM dNTP mix, 0.5 units μL^{-1} *Taq* polymerase (New England Biolabs), 30 ng μL^{-1} DNA template and PET, VIC, FAM or NED fluorescent SSR primer (0.15 mM primer mix of directly labelled primer or M13 primer mix with 0.03 µM tailed forward primer, 0.27 µM universal fluorescent labelled primer and 0.3 µM reverse primer). The PCR amplification profile followed: initial denaturation at 95°C for 3 min; 40 cycles of 30 s at 95°C, 1 min at annealing temperature (primer pair specific), and 2 min at 72°C; with a final extension of 20 min at 72°C. Successful amplifications were confirmed on 2.0% agarose gel stained with GelRed™ run against a classical 100 bp DNA ladder (Promega) in 1× TBE buffer and visualized under UV light. A standardized platform for molecular characterization developed for *Musa* germplasm by Christelová et al. (2011) was followed and automated capillary electrophoretic separation with internal standard (GeneScan™-500 LIZ size standard, Applied Biosystems) was used for detection of the amplified DNA fragments. To minimize the cost of genotyping, the PCR products were multiplexed (based on the dye and expected size of the fragment) prior to the separation and loaded onto the automatic 96-capillary ABI 3730xl DNA Analyser. Electrophoretic separation and signal detection was carried out with default module settings. The resulting data were then analyzed and alleles called using Gene mapper v4.0 software (Applied Biosystems Foster City, CA). Allele sizing and calling was done as described in the user's manual and alleles were scored manually as fragment sizes in base pairs.

Allelobin software (Prasanth et al., 2006) was used for adjusting allele size inconsistencies for the SSR markers that occur when size-calling alleles. SSR allelic data was converted to binary data matrix, where the alleles present in a locus were replaced by "1" and those absent were replaced by "0" using ALS binary software (Prasanth and Chandra, 2006). The efficiency of SSR primers to detect polymorphisms among the accessions was assessed by calculating the discriminatory power (Polymorphic Information Content; PIC). PIC is the relative discriminatory value of a locus which measures the information content as a function of a marker system's ability to distinguish between genotypes (Weir, 1990).

The genetic distances within and among the accessions were calculated based on Nei (1973) genetic distance measure of Power Marker software package Version 3.25 (Liu and Muse, 2005). Principal coordinate analysis (PCoA) was conducted with NTSYS-pc software package v2.3.3 (Rohlf, 2001), based on the simple-matching (SM) coefficient of Sokal and Michener (1958). A pair-wise dissimilarity matrix was generated from the single data following modalities dissimilarity index method of Roger-Tanimoto (DARwin V5.0; Perrier and Jacquemoud-Collet, 2006). The dissimilarity matrix was used to run cluster analysis based on unweighted Neighbour Joining (NJ) of Darwin V5.0. Bootstrap analysis, with 10^4 replications was performed for determining confidence limits in clusters produced by NJ algorithm.

RESULTS AND DISCUSSION

Results of Principal Coordinate Analysis (PCoA) provided genetic similarities within and between the banana accessions (Figure 1) and had the first two axes accounting for 48% (Eigen values) of the similarities observed. Within the different banana groups in this study, the diploids combined had the highest average genetic distance of 0.5666, with the ill-defined diploids alone having an average genetic distance of 0.5178. In contrast, the lowest average genetic distances were observed within the 'Ilalyi' (0.0748), one of the triploid landraces and the EAHB (0.0827) (Table 2, Figure 1). These observations suggest that the diploids (known and ill-defined) were more genetically diverse, while the 'Ilalyi' and the EAHB were not genetically distinct from each other. On the other hand, when the ill-defined East African diploids were compared with the 'Ilalyi' and EAHB, respectively, genetic distances of 0.4430 and 0.4350, were realized (Table 3). This suggests that these diploids are far different from the 'Ilalyi' and EAHB and could potentially be used in breeding programs

for improving the diversity within the triploid EAHB.

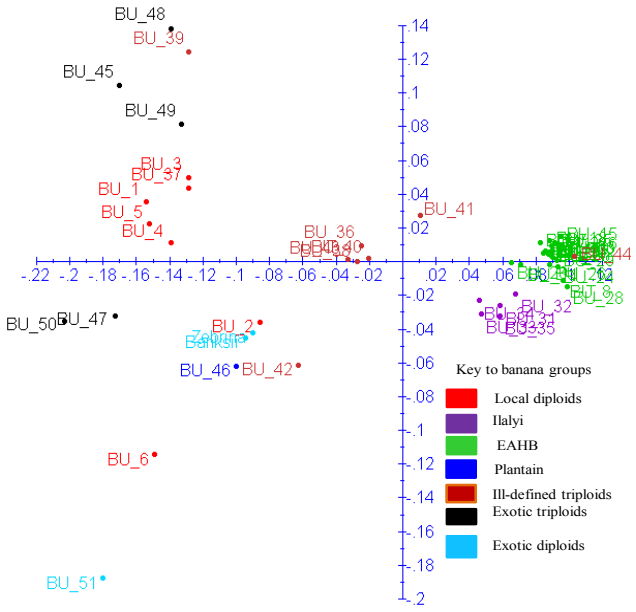

Figure 1. Principal Coordinate Analysis (PCoA) showing the genetic similarities among the 53 *Musa* samples.

Table 2. Average genetic distance within different banana landrace groups in East Africa.

No.	*Musa* groups	Average genetic distance
1	All diploids (AA)	0.5666
2	Ill-defined diploids	0.5178
3	Exotic triploids	0.4083
	---------Exotic hybrid (AAB) triploids	0.4568
	---------Exotic (AAA) triploids	0.3599
4	Ill-defined triploids	0.3425
5	East African highland bananas (clone sets)	0.0827
	- Musakala	0.1052
	- Nakitembe	0.0948
	- Nfuuka	0.0788
	- Nakabululu	0.0770
6	'Ilalyi'	0.0748

Table 3. Average genetic distance between diploid landraces and different triploid groups.

No.	Triploid *Musa* groups	Average genetic distances
1	Exotic triploids	0.1095
2	---------Exotic hybrid (AAB) triploids	0.1090
3	---------Exotic (AAA) triploids	0.1599
4	Ill-defined triploids	0.1743
5	East African highland bananas	0.4351
6	'Ilalyi'	0.4430

The PCoA (Figure 1) showing genetic similarities based on the markers among the 53 accessions reveals that the EAHB constitute a tight cluster indicating that they are genetically uniform. The EAHB are also closely associated with the 'Ilalyi' but not to the ill-defined triploids. BU44 which was considered to be among the ill-defined triploids in this study clustered with the EAHB. The local diploids do not form a tight cluster as they seem to be diverse and they are quite close to the exotic triploids. BU39 ('Bura'), one of the ill-defined triploid landraces, clustered with BU48, one of the exotic triploids, while BU42 ('Diana'), another ill-defined triploid, clustered with BU51, an exotic diploid. Further investigation to clarify the relationship between Bura and one of the exotic triploid cultivars would be useful. Although the 'Ilalyi' seem to be close to the EAHB, it is evident that they form their own distinct sub-cluster from the EAHB (Figure 1). PCoA results clearly show that there is no genetic link between EAHB, 'Illalyi' and the diploid landraces, meaning that these diploids are not possibly genetically related to the EAHB.

The results show that genetic diversity within and between the local triploid landraces is low and thus having a dangerously narrow genetic base (Tables 2 and 4, Figure 1). This makes them prone to genetic erosion due to potential biotic/or abiotic shocks. The genetic bases of local triploids need therefore to be broadened. However, there is a clear-cut distinction between the East African diploids and local triploids (the EAHB, the 'Ilalyi' and the local ill-defined triploids) based on the genetic distances. Results of this work failed to indicate the nearest local east African diploids to the EAHB or other local triploids, although local diploids could be now useful for breeding. This will not only raise the genetic base but will also probably improve the yield and quality performance of the triploid landraces.

Table 4. Average genetic distance between the East African Highland bananas and other triploid local landraces.

No.	*Musa* group	Average genetic distance
1	'Ilalyi'	0.0079
2	Ill-defined triploids after removing BU39 and BU42 (since these clustered with exotic triploids)	0.0640

ACKNOWLEDGEMENT

The authors thank Bioversity International and the CGIAR Research Program on Roots, Tubers and Banana (RTB) for the financial support.

Literature cited

Christelová, P., Valárik, M., Hřibová, E., De Langhe, E., and Doležel, J. (2011). A multi gene sequence-based phylogeny of the *Musaceae* (banana) family. BMC Evol. Biol. *11* (*1*), 103 http://dx.doi.org/10.1186/1471-2148-11-103. PubMed

Cregan, P.B. (1992). Simple sequence repeat DNA length polymorphisms. Probe *2*, 18–22.

Crouch, J., Crouch, K., Ortiz, R., and Jarret, L. (1997). Microsatellite markers for molecular breeding of *Musa.* InfoMusa *6*, 5–6.

De Langhe, E., Karamura, D., and Mbwana, A.S.S. (2001). Tanzania *Musa* Expedition (Kampala, Uganda: INIBAP).

De Langhe, E., Karamura, D., and Mbwana, A.S.S. (2002). Tanzania *Musa* Expedition. Follow up. A Report of the visit to the Field Collection at Tengeru (Kampala, Uganda: INIBAP).

Kolesnikova, A.M., and Kouassi, K.S. (2006). Development of reference molecular marker kits to analyze diversity of germplasm for the year 1 GCP crops (*Musa*). In Final Project report, p.26. http://gcpcr.grinfo.net/files/cr_files/gcpcr_file825.doc (Accessed on July 10, 2014).

Liu, K., and Muse, S.V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics *21* (*9*), 2128–2129 http://dx.doi.org/10.1093/bioinformatics/bti282. PubMed

Mbanjo, E.G.N., Tchoumbougnang, F., Nyine, M., Dochez, C., and Lorenzen, J. (2009). An EST-based SSR genomic resource for bananas. Paper presented at: Plant and Animal Genomes XVII Conference (San Diego, USA).

NARO. (2012). Annual Technical report 'Novel Approaches to the Improvement of Banana Production in Eastern

Africa - the Application of Biotechnological Methodologies'. p.13–14.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. *70* (*12*), 3321–3323 http://dx.doi.org/10.1073/pnas.70.12.3321. PubMed

Ortiz, R., and Vuylsteke, D. (1996). Recent advances in *Musa* genetics, breeding and biotechnology. Plant Breeding Abstracts *66*, 1355–1363.

Perrier, X., and Jacquemoud-Collet, J.P. (2006). DARwin software. http://darwin.cirad.fr/.

Prasanth, V.P., and Chandra, S. (2006). *ALS-Binary*: AC Program for converting allele size in microsatellite markers to 0-1 (*Binary* Data). Biometrics (Patancheru, India: ICRISAT). http://www.icrisat.org/bt-software-d-als.htm.

Prasanth, V.P., Chandra, S., Jayashree, B., and Hoisington, D. (2006). AlleloBin - A program for allele binning of microsatellite markers based on the algorithm of Idury and Cardon (1997) (ICRISAT). http://www.researchgate.net/publication/270648184.

Rohlf, F.J. (2001). NTSYS-pc Numeral Taxonomy and Multivariate Analysis System, Version 2.10j (New York, USA: Exeter Publications).

Simmonds, N.W. (1966). Bananas, 2nd edn, Tropical Agricultural Series (New York, USA: Longman).

Sokal, R.R., and Michener, C.D. (1958). A statistical method for evaluating systematic relationships. Univ. Kans. Sci. Bull. *38*, 1409–1438.

Weir, B.S. (1990). Genetic Data Analysis II: Methods for Discrete Population Genetic Data (Sunderland, UK: Sinauer Associates Inc).