Genetic dissimilarity and selection of putative mutants of Terra Maranhão plantain cultivar using the Ward-MLM strategy

R.V. Reis¹, E.P. Amorim², V.B.O. Amorim², C.F. Ferreira³, R.K.N. Pestana¹, C.A.S. Ledo², Z. Gonçalves⁴ and A. Borém¹

¹Universidade Federal de Viçosa. Viçosa, MG, Brasil
²Embrapa Mandioca e Fruticultura, Cruz das Almas, BA, Brasil
³Universidade Federal do Recôncavo da Bahia, Cruz das Almas, BA, Brasil

Corresponding author: E.P. Amorim
E-mail: edson.amorim@embrapa.br

Received March 8, 2015
Accepted June 8, 2015
Published November 30, 2015
DOI http://dx.doi.org/10.4238/2015.November.30.11

ABSTRACT. The objective of this study was to evaluate genetic variability and select putative mutants of Terra Maranhão plantain cultivar (AAB genome) subjected to gamma radiation based on agronomic data and inter simple sequence repeat molecular marker profiles using the Ward-MLM strategy. A total of 233 irradiated plants and 41 controls were assessed. The agronomic and molecular data were subjected to the Ward-MLM statistical algorithm in the SAS program. Cluster analysis was performed by the average distance method (UPGMA), based on the distance matrix of the Gower algorithm, and the cophenetic correlation coefficient calculated using the R software. The distance between the putative mutants ranged from 0.321 to 0.524, with an average distance of 0.426, and a cophenetic correlation coefficient of 0.79. Three putative mutants, which were selected based on the best agronomic traits and low height, will undergo further evaluation in the next stages of the banana breeding program at Embrapa. These results describe the first attempt of using combined data of Terra Maranhão plantain cultivar for the purpose of selecting mutants and
assessing genetic variability.

**Key words:** Gamma rays; Ward-MLM; Mutation induction; ISSR markers; Genetic variability

## INTRODUCTION

Cultivation of bananas is an important worldwide agribusiness activity, and has highly significant economic and social implications with an estimated planted area of 5.1 million hectares in 126 countries and an annual production of 106 million tons (FAO, 2014). With an average approximate production of 6.87 million tons in 2013, and a planted area of approximately 526,000 hectares, Brazil stands out as the fifth largest worldwide producer of bananas. Moreover, the Northeast region has a 41% share of national production, employing an estimated 960,000 people in fruit production (IBGE, 2014).

The study of genetic diversity is very important in breeding programs, especially in revealing the genetic variability available to breeders. Molecular markers have been used to determine diversity in many species (Joshi et al., 2000; Gaudeul et al., 2000; Qian et al., 2001; Guasmi et al., 2008).

These markers, especially those utilizing polymerase chain reaction (PCR)-based methods, including amplified fragment length polymorphism, random amplified polymorphic DNA, microsatellites or simple sequence repeats (SSRs), and inter simple sequence repeats (ISSR), have been broadly used to estimate genetic diversity and stability of bananas, and in phylogenetic studies (Creste et al., 2004; Lakshmanan and Venkataramareddy, 2007; Ning et al., 2007; Amorim et al., 2008, 2009a,b; El-Khishin et al., 2009; Miri et al., 2009; Mattos et al., 2010; Opara et al., 2010; Pestana et al., 2011).

ISSR-based technique is one of many techniques that use SSR primers to amplify regions between target sequences. This technique can be used to generate a large number of multilocus markers and may be applied to almost any organism, even those for which there is little or no previous genetic information available (Arcade et al., 2000).

In addition, ISSR markers are useful in genetic studies, especially for clonal detection and to identify closely related individuals (Salimath et al., 1995; de Oliveira et al., 1996). Such markers are dominant and highly reproducible, with the advantage of generating many bands and being widely distributed throughout the eukaryotic genome (Gupta et al., 1994; Fang and Roose, 1997). Each range corresponds to a DNA sequence delimited by two inverted microsatellites.

In this context, the aim of this study was to evaluate genetic variability and select putative mutants of Terra Maranhão plantain cultivar (AAB genome) subjected to gamma radiation based on a set of morphological, agronomic, and molecular data using the Ward-modified location model (MLM) strategy.

## MATERIAL AND METHODS

### Gamma irradiation of in vitro buds

A total of 315 in vitro buds were irradiated using $^{60}$Co in the Center for Nuclear Energy in Agriculture (Centro de Energia Nuclear na Agricultura, CENA), in partnership with Universidade de São Paulo (USP). The dose of gamma irradiation used was 20 Gy, which was the dose previously
selected in a sensitivity test (Pestana et al., 2010). Fifty buds were used as controls, which were sent to CENA/USP, without exposure to $^{60}$Co.

The irradiated buds were transferred to a basic MS culture medium, solidified with 2.2 g/L of Phytagel, and supplemented with 30 g sucrose and benzylaminopurine at 3.0 mg/L, pH 5.8. Plantlets were maintained in a growth chamber under a controlled temperature of 27 ± 2°C and a 16-h fluorescent light photoperiod and 8-h dark period. Thereafter, plantlets were subjected to four subcultures performed at 30-day intervals.

Following the subcultures, plants were rooted in MS medium with the addition of 0.25 mg/L naphthaleneacetic acid, and 8 g/L agar was used as a gelling agent, with the pH adjusted to 5.8. Plants were then placed in a growth room and subjected to a 16-h fluorescent light photoperiod with luminous intensity of 40 μmol·m$^{-2}$·s$^{-1}$ and a 8-h dark period with a temperature of 26 ± 2°C, where they remained for a period of 35-40 days.

The already rooted plants were taken to a greenhouse where they were acclimatized in root plugs containing the commercial substrate Plantmax (processed wood bark, expanded vermiculite, granulated carbon, and processed peat enriched with macro- and micronutrients) and Sombrite shade screening, which controls 50% of the luminosity. Irrigation was carried out using an automatic spraying system.

**Field evaluation of mutants**

The experiment was carried out in the experimental field at Embrapa Mandioca e Fruticultura, which presents good agronomic traits and enabled short banana clones of Terra Maranhão plantain cultivar to be selected.

The experiment was established without adopting an experimental design, and the main aim of the study was to select 10% of the best plants from a large group of banana plants. Planting was carried out at a spacing of 3 x 4 m, and basal application and topdressing of fertilizer was performed according to the technical recommendations for the crop. Evaluation and selection were performed with the irradiated and non-irradiated (control) plants that completed the two production cycles in the field.

The following agronomic traits were evaluated during the flowering period of two production cycles: number of days from planting to flowering; pseudostem diameter (cm); plant height (m); number of live leaves at flowering; and the number of suckers. When the bunches were harvested, the following evaluations were made: number of days from planting to harvest and from flowering to harvest; weight of the bunch (kg) and hands (kg); average weight of the fruits (g); number of fruits per bunch; length of the fruit of the second and of the next-to-the-last hand (cm); diameter of the fruit of the second and of the next-to-the-last hand (mm); number of hands and live leaves at harvest; length of the fruit stalk (cm); and diameter of the fruit stalk (mm).

**Selection of mutants for short plants**

The methodology proposed by Resende (2005) was used to select 10% of the shorter plants. Initially, individual classification of the plants was performed for the traits considered to be important for selection, such as shorter plant height, fewer days from planting to flowering, and greater bunch weight. The data obtained for these traits were ranked in increasing order (for plant height and number of days to bunch emergence) and decreasing order (for bunch weight), and the classification number of each plant was obtained.

Thereafter, a selection index based on the ranking was calculated by multiplying the classification number of each plant by the coefficient corresponding to each trait. For plant height, a weight of six was considered most relevant and for the number of days from planting to flowering and bunch weight, a weight of two was considered most relevant. At the end of this process, a final score was obtained for each plant by the following selection index: \[ y = 0.6 \times \text{height classification} + 0.2 \times \text{bunch emergence classification} + 0.2 \times \text{bunch weight classification} \]. In each cycle, the final score was ranked in increasing order, and the final classification was obtained.

The plants selected were those that were among the top 10% classified in both production cycles (Figure 1).

### Table 1

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Ta (°C)</th>
<th>NTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiGA3'C</td>
<td>(GA)8C</td>
<td>48</td>
<td>5</td>
</tr>
<tr>
<td>DiGA3'RC</td>
<td>(GA)8RC</td>
<td>48</td>
<td>6</td>
</tr>
<tr>
<td>DiGA3'YC</td>
<td>(GA)8YC</td>
<td>48</td>
<td>7</td>
</tr>
<tr>
<td>DiGT3'A</td>
<td>A(GT)8</td>
<td>48</td>
<td>8</td>
</tr>
<tr>
<td>DiGT3'CY</td>
<td>CY(GT)8</td>
<td>48</td>
<td>11</td>
</tr>
<tr>
<td>TriCAG5'CR</td>
<td>CR(CAG)5</td>
<td>48</td>
<td>8</td>
</tr>
<tr>
<td>TriAAG3'RC</td>
<td>(AAG)5RC</td>
<td>48</td>
<td>9</td>
</tr>
<tr>
<td>TriAGA3'RC</td>
<td>(AGA)5RC</td>
<td>48</td>
<td>4</td>
</tr>
<tr>
<td>TriTGA3'RC</td>
<td>(TGA)5RC</td>
<td>48</td>
<td>6</td>
</tr>
<tr>
<td>TriTG3'RC</td>
<td>(TGC)5RC</td>
<td>48</td>
<td>5</td>
</tr>
<tr>
<td>TriCTC3'RC</td>
<td>(CTC)5RC</td>
<td>48</td>
<td>8</td>
</tr>
<tr>
<td>TriCTG3'RC</td>
<td>(CTG)5RC</td>
<td>48</td>
<td>8</td>
</tr>
<tr>
<td>TriGAT3'RC</td>
<td>(GAT)5RC</td>
<td>48</td>
<td>5</td>
</tr>
<tr>
<td>TriGAG3'RC</td>
<td>(GAG)5RC</td>
<td>48</td>
<td>6</td>
</tr>
</tbody>
</table>

**Ta** = annealing temperature; **NTB** = total number of bands.

**ISSR primers**

In all, 14 ISSR primers were used. Information regarding their respective sequences, annealing temperatures, and total number of bands are presented in Table 1.
DNA extraction and PCR conditions

Genomic DNA was extracted from young leaves using the CTAB method (Doyle and Doyle, 1990). The quantity and quality of DNA was evaluated through comparative analysis of the samples on 0.8% agarose gel, stained with ethidium bromide, with the samples being diluted in ultrapure water and standardized to a final concentration of 10 ng/mL. For amplification reactions via ISSRs, the samples were used in a final volume of 15 μL, which contained the following reagents: 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 100 mM of each dNTPs (dTTP, dGTP, dCTP), 0.4 mM of each primer, 20 ng genomic DNA, and 1.0 U Taq DNA polymerase (Pharmacia Biotech, USA).

Amplifications were conducted in a Perkin Elmer thermocycler, model 9700, and consisted of 1 cycle at 94°C for 4 min, followed by 35 cycles of 94°C for 40 s, (annealing temperature of 48ºC) for 40 s, 72°C for 1 min, with a final extension at 72°C for 2 min. The amplification products were separated by 2% agarose gel electrophoresis.

Statistical analysis of the combined data

The morphological, agronomic, and molecular data were subjected to the Ward-MLM statistical algorithm (Franco et al., 1998) using the CLUSTER and MLM procedure in the SAS program (2002), considering a total of seven multicategorical traits [pseudostem color, petiole rib color, form of the foliar rosette, color of the suckers, color of the edge of the petiole, opening of the petiole, and leaf position (LP)], 21 quantitative traits [number of days from planting to flowering, number of days from planting to harvest, number of days from flowering to harvest, diameter of the pseudostem (DP, cm); plant height (m), number of leaves up to flowering, number of suckers, bunch weight (kg), hand weight (kg), number of fruits, average weight of fruits (g), length of fruits of the second hand (cm), length of the fruit of the next-to-the-last hand (cm), diameter of fruit from the second hand (mm), diameter of fruit from the next-to-the-last hand (mm), number of hands, number of live leaves at harvest, length of the peduncle (LP, cm), diameter of the peduncle (DP, cm); length of the fruit stalk (cm), diameter of the fruit stalk (cm)], and 14 ISSR primers (96 polymorphic bands) to evaluate genetic diversity. The amplified fragments were evaluated by the presence (1) or absence (0) of bands.

Cluster analysis was performed by UPGMA (unweighted pair-group method with arithmetic mean), using the STATISTICA software produced by StatSoft (Statistica, 2002), based on the distance matrix of the Gower algorithm (Gower, 1971). The cophenetic correlation coefficient between the matrix of genetic similarities and the matrix of the cophenetic values was calculated using the R software (2010) to verify the consistency of clustering. The ideal number of groups was defined according to the criteria of the Pseudo-F (SAS, 2002).

RESULTS AND DISCUSSION

Selected qualitative and quantitative data were used in the second production cycle. The possibility of combining multicategorical, quantitative, and molecular data to identify the number of clusters presents new perspectives in studies of dissimilarity.

The strategy of the Ward-MLM model was first proposed by Franco et al. (1998). This is a useful method that can be used to analyze genetic diversity to allow genealogical comparisons to be made (Barbé et al., 2009). It is efficient at selecting genotypes with minimum loss of
genetic diversity (Crossa and Franco, 2004). Nevertheless, there are few reports regarding this methodology in the literature. This method has been used in crops such as snap beans (Barbé et al., 2009), Capsicum spp (Sudré et al., 2010), tomatoes (Gonçalves et al., 2008), kale (Padilla et al., 2007), and banana Pacovan of the Prata subgroup (Pestana et al., 2011). In these studies, the use of the Gower algorithm proved to be very effective.

To investigate correlations between the multicategorical, quantitative, and binary data, a preliminary study was carried out, in which the distance of the matrices including the multicategorical and quantitative data, and the distance matrix of the binary data using the simple coincidence index were calculated. The correlation between these two matrices was quite low at $r = 0.0218$, which shows that the data may be analyzed in a combined manner, and that most of the variability observed may be attributed to genetic factors.

Subsequently, the genetic variability of 21 putative mutants and two controls was analyzed using 21 quantitative traits, seven multicategorical traits, and 14 ISSR markers, which gave rise to 96 polymorphic bands by the Ward-MLM method (Figure 2).

![Dendrogram](image_url)

**Figure 2.** Dendrogram constructed with 21 putative mutants of Terra Maranhão plantain cultivar and two controls through combined data using the Gower algorithm and the Ward-MLM strategy.
The distance between putative mutants ranged from 0.321 to 0.524, with a mean distance of 0.426 and a cophenetic correlation coefficient of 0.79 - a value considered adequate in these studies (Vaz Patto et al., 2004; Pestana et al., 2011). The nearest genotypes were those of Tests 22 and 23, which had a genetic distance of 0.321, and the most distant genotypes were Irra 09 and 11, with a distance of 0.524. These results show that there is variability among the putative mutants, which may be exploited, and that this variability is mainly due to the effect of gamma radiation. Somaclonal variation and subcultures do not normally lead to genetic changes in a large number of clones from the same genotype.

According to the classification performed in the two production cycles relating to plant height, bunch weight, and number of days from planting to flowering, three putative mutants were selected (Figure 3): Irra 12, Irra 14, and Irra 16, which showed the greatest values for bunch weight when compared to the controls. The genotype Irra 16 stood out with a bunch weight of 29 kg, which was an increase of 6.37 kg in comparison to that exhibited by the controls. With regard to the number of days from planting to flowering, the putative mutants selected, Irra 12, Irra 14, and Irra 16, showed lower values in comparison to the mean values of the controls. The genotype Irra 16 began flowering 400 days after planting, 42.33 days earlier than the mean of the controls, which began flowering 442.33 days after planting. A similar result was observed in a study by Pestana et al. (2011). Those authors used the Prata type banana, Pacovan, and after the second production cycle, they selected four genotypes, most notably Pacovan 40, which flowered 44 days earlier than did the controls, and it was also among those that showed the highest bunch weight.

The height of the three plants selected among the 23 evaluated after the second production cycle was less than the mean value of the controls. The lowest height among the plants selected
was 3.65 m in the genotype Irra 16, with a difference of 3.2% in comparison to the mean value of the controls (mean height = 4.02 ± 0.29 m). In addition, this putative mutant, Irra 16, flowered 42.33 days earlier than the mean value for the controls and was also among those with highest bunch weight (Figure 3).

A similar result was observed in a study conducted by Bermúdez et al. (2000). Those authors, working with banana FHIA-21, reported 4.78% variation in the first production cycle and reported that the most altered traits were plant height and the number of fruits. For ISSR analysis using 14 primers, 21 putative mutants, and two controls, a total of 96 polymorphic bands was generated, with a mean of 6.86 bands per primer. The greatest number of bands was observed with the primer TriAAG3’RC (11 bands) and the least number with the primer TriTGC3’RC (four bands).

The Pseudo-F criteria showed that the optimum number of groups is 2-3 (Figure 4). Adopting the formation of three groups, the lowest genetic distance, 0.321, was detected between the two controls. This shows that the controls are closely related and are in the same group (G2) in the dendrogram. They may therefore serve as a good baseline to be used in comparison with the irradiated material because these variations are attributed to environmental effects.

Of the three mutants selected, Irra 12 was placed in group 1 (G1), and Irra 14 and Irra 16 were placed in group 2 (G2). In this case, the use of combined data allowed the reliable study of genetic variability, and shows that in the putative Terra type banana mutants selected, there is variability that might be utilized.

These three putative mutants will be subjected to field evaluations in the next stages of the banana plant-breeding program of Embrapa Mandioca e Fruticultura for the purpose of introducing a new variety that is earlier to flower, more productive, and shorter in height. These results describe the first attempt of using combined data in Terra Maranhão plantain cultivar for...
the purpose of selecting putative mutants and evaluating genetic variability, and show that there is genetic variability in the putative mutants selected that may be used in genetic breeding programs with a view towards obtaining earlier and more productive plants of reduced height.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors thank the Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) for providing financial support and for granting a doctoral scholarship to the first author.

REFERENCES


