GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

# Descriptor selection for banana accessions based on univariate and multivariate analysis

**L.P. Brandão[1], C.P.F. Souza[1], V.M. Pereira[1], S.O. Silva[1], J.A. Santos-Serejo[2], C.A.S. Ledo[2] and E.P. Amorim[2]**

[1]Universidade Federal do Recôncavo da Bahia, Cruz das Almas, BA, Brasil
[2]Embrapa Cassava and Fruits, Cruz das Almas, BA, Brasil

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

Genet. Mol. Res. 12 (2): 1603-1620 (2013)
Received April 24, 2012
Accepted November 30, 2012
Published May 14, 2013
DOI http://dx.doi.org/10.4238/2013.May.14.1

**ABSTRACT.** Our objective was to establish a minimum number of morphological descriptors for the characterization of banana germplasm and evaluate the efficiency of removal of redundant characters, based on univariate and multivariate statistical analyses. Phenotypic characterization was made of 77 accessions from Bahia, Brazil, using 92 descriptors. The selection of the descriptors was carried out by principal components analysis (quantitative) and by entropy (multicategory). Efficiency of elimination was analyzed by a comparative study between the clusters formed, taking into consideration all 92 descriptors and smaller groups. The selected descriptors were analyzed with the Ward-MLM procedure and a combined matrix formed by the Gower algorithm. We were able to reduce the number of descriptors used for characterizing the banana germplasm (42%). The correlation between the matrices considering the 92 descriptors and the selected ones was 0.82, showing that the reduction in the number of descriptors did not influence estimation of genetic variability between the banana accessions. We conclude that removing these descriptors caused no loss

 

of information, considering the groups formed from pre-established criteria, including subgroup/subspecies.

**Key words:** *Musa* sp; Variability; Morphoagronomic characteristics

## INTRODUCTION

Banana is the second most consumed fruit in Brazil, second only to orange. It is cultured by small rural entrepreneurs and establishes manpower in rural areas because it is a continuous source of income for these farmers. Brazil is the fifth largest producer of bananas. It produced 6.9 million tons in 2010 in an area of approximately 487,000 ha (FAO, 2012).

The expansion of banana crops depends on the development of new cultivars with resistance to major diseases (black leaf streak/yellow Sigatoka and *Fusarium* wilt) and other superior traits primarily aimed at expanding cultivation alternatives for farmers (Amorim et al., 2011). Thus, genetic breeding is crucial for the sustainability of banana agribusiness worldwide.

New cultivars must be registered and receive intellectual protection to qualify for commercial scale production. Protection aims to safeguard technology developers and establish rules of usage rights. To meet this demand, minimum efficient descriptors should be established to facilitate the distinction of new cultivars.

Embrapa Cassava and Fruits has a germplasm collection with 321 accessions obtained through exchange and international collections for carrying out genetic breeding (Santos-Serejo JA, personal communication). Notably, the accessions of the AAB genomic group, in which the most important representatives in Brazil are the cultivars Prata, Pacovan, Prata Anã, Maçã, Mysore, and Terra, occur more frequently (29%), whereas the diploid (AA) and triploid (AAA) groups are represented in the country, respectively, by "Ouro da Mata" and the cultivars Caru Verde, Caru Roxa, São Tomé, Nanica, Nanicão, and Grand Naine. Cultivars in the 2 groups have intermediate frequencies of 26 and 21%, respectively. The groups BB (4%), ABB (8%), AAAB (3%), and AAAA (6%) are less frequently present. Thus, the banana germplasm is well represented and has great potential for use in breeding programs.

Morphoagronomic characterization of the accessions preserved in the Embrapa banana collection is performed by descriptors established by the IPGRI (1996) and Embrapa (Silva et al., 1999). According to Daher (1993), a large number of descriptors may result in the presence of redundant traits because they are always associated with others. Thus, the definition of a minimal set of descriptors reduces the need for collecting data without lowering the reliability of the results (Pereira, 1989).

Principal component analysis is indicated for the identification of descriptors with better capability for discriminating accessions. This analysis also eliminates traits that contribute little to total variation (Cruz et al., 2004). The efficiency of this method has not been tested in banana yet, but reports have appeared in the literature describing the use of principal component analysis as a criterion for the selection of descriptors in various cultures (Daher et al., 1997; Dias et al., 1997; Strapasson et al., 2000; Alves et al., 2003; Oliveira et al., 2006; Oliveira et al., 2012). The effectiveness of principal component analysis has been verified by comparing groups formed by all the descriptors and those selected using various grouping methods (Cury, 1993; Dias et al., 1997; Araújo et al., 2002).

Another tool with the potential to select descriptors - mainly qualitative or multi-category descriptors - is the level of character entropy (H) proposed by Renyi (1961). The greater

 

the entropy of a given descriptor, the greater the number of its phenotypic classes and the more homogeneous the balance between the frequency of accessions in different phenotypic classes (Vieira et al., 2007). The present study aimed to establish a minimum number of morphological descriptors for the characterization of the banana germplasm and evaluate the efficiency of the disposal of redundant traits using univariate and multivariate statistical methods.

## MATERIAL AND METHODS

Seventy-seven accessions from the banana germplasm collection of Embrapa Cassava and Fruits (Table 1) were characterized in Cruz das Almas, Bahia, Brazil, located at 12°40'19"S and 39°06'22"W, 220 m above sea level. The climate is tropical, hot, and humid, with a tropical monsoon to tropical savanna climate, according to Köppen classification, an annual average temperature of 24.5°C, relative humidity of 80%, and average rainfall of 1249.7 mm per year (AGRITEMPO, 2012).

Five clones were characterized in each accession, and each observation was represented by the measurements made in each character. To avoid distortion of the data, the plant evaluation stage was standardized. The plants were evaluated after the occurrence of inflorescence and when the rachis reached approximately 15 cm.

We used 92 morphological descriptors established by the IPGRI (1996) and Embrapa (Silva et al., 1999); 27 were quantitative and 65 were multi-category (Table 2 and Figure 1). The quantitative descriptors were selected via principal component analysis based on the average of each character from the correlation matrix.

The disposal was carried out with 2 procedures: 1) direct selection (Jolliffe, 1972, 1973), which eliminated the characters with the highest weighting coefficient in absolute value (eigenvector) in the principal component with the smallest eigenvalue, starting from the last component and ending with the one with an eigenvalue less than or equal to 0.70; 2) selection with reanalysis (Cury, 1993), in which a new analysis was performed after the disposal of each character, using the remaining characters and examining the correlation coefficients between the character suggested for disposal and the other characters. The final disposal of the characters considered the information that coincided in the 2 methods, eliminating the characters assorted as redundant by both procedures. Pearson's correlation coefficients were estimated among all the characters aiming to assist in the decision to discard certain redundant traits and the completion of the analysis in the selection method with reanalysis.

The selection of multi-categorical descriptors was performed by means of H, proposed by Renyi (1961). The greater the number of its phenotypic classes and the more homogeneous the balance between the frequency of accessions in the various phenotypic classes (Vieira et al., 2007), the greater the entropy of any descriptor. In this study, "low value for H ($\leq$1.00)" and "more than 50% of the accessions classified into one of the descriptor classes" were used as criteria for discarding the descriptor.

The efficiency of the disposal was analyzed through comparative study of the groups formed using the Ward-modified location model (Ward-MLM) algorithm (Franco et al., 1998), considering both the 92 descriptors in total and only the selected descriptors (quantitative and multi-category). Estimates of phenotypic dissimilarity obtained for the 77 banana accessions were carried out only with the descriptors selected using direct methods (Jolliffe, 1972, 1973) with reanalysis (Cury, 1993) and entropy (Renyi, 1961).

**Table 1.** Identification and origin of the 77 banana accessions that were evaluated.

| Code | Accessions | Ploidy | Subgroup/subspecies | Origin |
|---|---|---|---|---|
| 1 | 028003-01[1] | AA | (*Tuugia* x *Calcutta 4*) | Brazil |
| 2 | Abu Perak | ABB | | France |
| 3 | Adimoo | AAB | | New Guinea |
| 4 | Akondro Mainty | AA | | France |
| 5 | Babi Yadefana | AA | | New Guinea |
| 6 | Balbisiana France | BB | *balbisiana* | France |
| 7 | Birmanie | AA | spp *burmanica* | France |
| 8 | Burmannica | AA | spp *burmanica* | Honduras |
| 9 | Butuhan | BB | *balbisiana* | Philippines |
| 10 | Cacambou Naine | ABB | *bluggoe* | Ecuador |
| 11 | Calcutta 4 | AA | spp *burmannicoide* | Jamaica |
| 12 | Canela | AAA | | Brazil |
| 13 | Cici | AA | spp *malaccensis* | Indonisia |
| 14 | D'Angola | AAB | Plátano | Brazil |
| 15 | F3P4 | AA | | Ecuador |
| 16 | FC-0602 | AAB | (*M. balbisiana* x *Buitenzorg*)[3] | Brazil |
| 17 | FHIA 18[2] | AAAB | (Prata Anã x SH3142)[4] | Honduras |
| 18 | Grand Naine | AAA | Cavendish | Brazil |
| 19 | Ice Cream | ABB | | France |
| 20 | Ido 110 | AA | | France |
| 21 | Imperial | AAA | Cavendish | Brazil |
| 22 | Jambi | AA | spp *malaccensis* | Indonisia |
| 23 | Japira | AAAB | (Pacovan x M53)[3] | Brazil |
| 24 | BGB 148 | AAB | | Brazil |
| 25 | Khai | AA | spp *malaccensis* | Tailand |
| 26 | Khi Maeo | AA | | Tailand |
| 27 | Kongo FRF 1259 | AAB | | Brazil |
| 28 | Krasan Saichon | AA | | Tailand |
| 29 | Lidi | AA | | Honduras |
| 30 | Malaccensis | AA | spp *malaccensis* | Honduras |
| 31 | Malbut | AA | | New Guinea |
| 32 | Mambee Thu | AA | spp *banksii* | New Guinea |
| 33 | Mangana | AA | | New Guinea |
| 34 | FHIA 01[2] | AAAB | (Prata Anã x SH3142)[4] | Brazil |
| 35 | Marcatoa | AAA | | New Guinea |
| 36 | Marmelo | ABB | | Brazil |
| 37 | Nam | AAA | | Tailand |
| 38 | NBA 14 | AA | spp *banksii* | New Guinea |
| 39 | NBF 9 | AA | | New Guinea |
| 40 | Niyarma Yik | AA | spp *banksii* | New Guinea |
| 41 | Orotava | AAA | | France |
| 42 | Ouro da Mata | AAAB | pome | Brazil |
| 43 | PA Abssinea | AA | | Tailand |
| 44 | Pisang Kermain | AA | | - |
| 45 | Pa Musore 3 | AA | spp *malaccensis derivada* | Tailand |
| 46 | Pa Patthalung | AA | | Tailand |
| 47 | Pa Rayoung | AA | spp *siamea* | Tailand |
| 48 | Pacovan | AAB | pome | Brazil |
| 49 | Pagatow | AAA | | New Guinea |
| 50 | Pioneira | AAAB | (Prata Anã x Lidi)[3] | Brazil |
| 51 | Pipit | AA | | Indonisia |
| 52 | Prata Anã 2 | AAB | pome | Brazil |
| 53 | Prata Anã 3 | AAB | pome | Brazil |
| 54 | Prata Anã Batico | AAB | pome | Brazil |
| 55 | Prata Anã Rene | AAB | pome | Brazil |
| 56 | Prata Graúda | AAB | pome | Brazil |
| 57 | PV 03-76 | AAAB | (Pacovan x Calcutta 4)[3] | Brazil |
| 58 | Royal (*M. ornata x M. velutina*) | - | *Rhodochlamys* | - |
| 59 | Samura B | AAB | Plantain | Brazil |
| 60 | São Tomé 2 Cachos | AAA | | Brazil |
| 61 | Sowmuk | AA | spp *banksii* | New Guinea |
| 62 | SRI | AAA | | - |

Continued on next page

**Table 1.** Continued.

| Code | Accessions | Ploidy | Subgroup/subspecies | Origin[2] |
|---|---|---|---|---|
| 63 | Tambi | AA | | Brazil |
| 64 | Terrinha | AAB | Plantain | Brazil |
| 65 | Thap Maeo | AAB | | New Guinea |
| 66 | Tomnam | AAB | | Tailand |
| 67 | Thong Dok Mak | AA | | New Guinea |
| 68 | Towolee | AAA | | Hawaii |
| 69 | Tuu Gia | AAB | | New Guinea |
| 70 | Uwati | AA | | New Guinea |
| 71 | Verde | AAB | | Hawaii |
| 72 | Walebo | AAA | | New Guinea |
| 73 | Walha | AAB | pome | France |
| 74 | Wasolay | AAA | | France |
| 75 | Yangambi KM5 | AAA | ibota | France |
| 76 | Yangambi No. 2 | AAB | silk | France |
| 77 | Zebrina | AA | spp *zebrina* | Hawaii |

- = no information; [1,3]hybrids developed by Embrapa, [2,4]hybrids developed by FHIA (Fundación Hondureña de Investigación Agrícola).

The quantitative and multi-category traits selected were analyzed jointly using the Ward-MLM procedure (Franco et al., 1998). The cluster and interactive matrix programming procedures were used to form the groups of accessions. The Ward clustering method was used with the joint matrix obtained from the Gower joint algorithm (Gower, 1971).

To define the optimal number of groups, we considered the procedure indicated in the MLM model, which is based on *pseudo-F* and *pseudo-t*$^2$ statistics. Considering the definition of the optimal number of groups, we obtained a hierarchical classification using the Ward method, which provided the initial value required to program the final step of the MLM model (Crossa and Franco, 2004). All statistical analyses were performed using the SAS software system version 8.1 (SAS Institute, Cary, NC, USA). The dendrogram was obtained using the NTSYS-pc software system (Rohlf, 2000).

## RESULTS AND DISCUSSION

### Phenotypic variation according to the univariate analysis of variance

Significant differences were observed among the 77 accessions for all quantitative morphoagronomic descriptors except fruit pedicel width (FPW) (see Table 2). Plant height (PLH) ranged from 80 cm for accession Royal (AA) to 480 cm for PV 0376 (AAAB), with an average of 235.09 cm. The identification of diploid accessions with short stature is important because these accessions can be used as male parents in crosses aimed at the development of hybrids with low PLH.

The pseudostem diameter (PSD) averaged 15.31 cm and displayed maximum and minimum values of 48.00 cm (diploid "Khai") and 5.30 cm (diploid "Babi Yadefana"), respectively. PSD is an important trait for breeding because it is associated with the capacity to support the fruit bunch.

The number of suckers (NUS) ranged from 0.00 (no seedlings; "Japira", AAAB) to 12.00 (Pioneira, AAAB; "Samura B", AAB), with an average of 3.75. This character is important because the species spreads vegetatively, and replanting is performed by removing seedlings from the field directly or via *in vitro* micropropagation.

Regarding the production components, wide variation was found for each character -

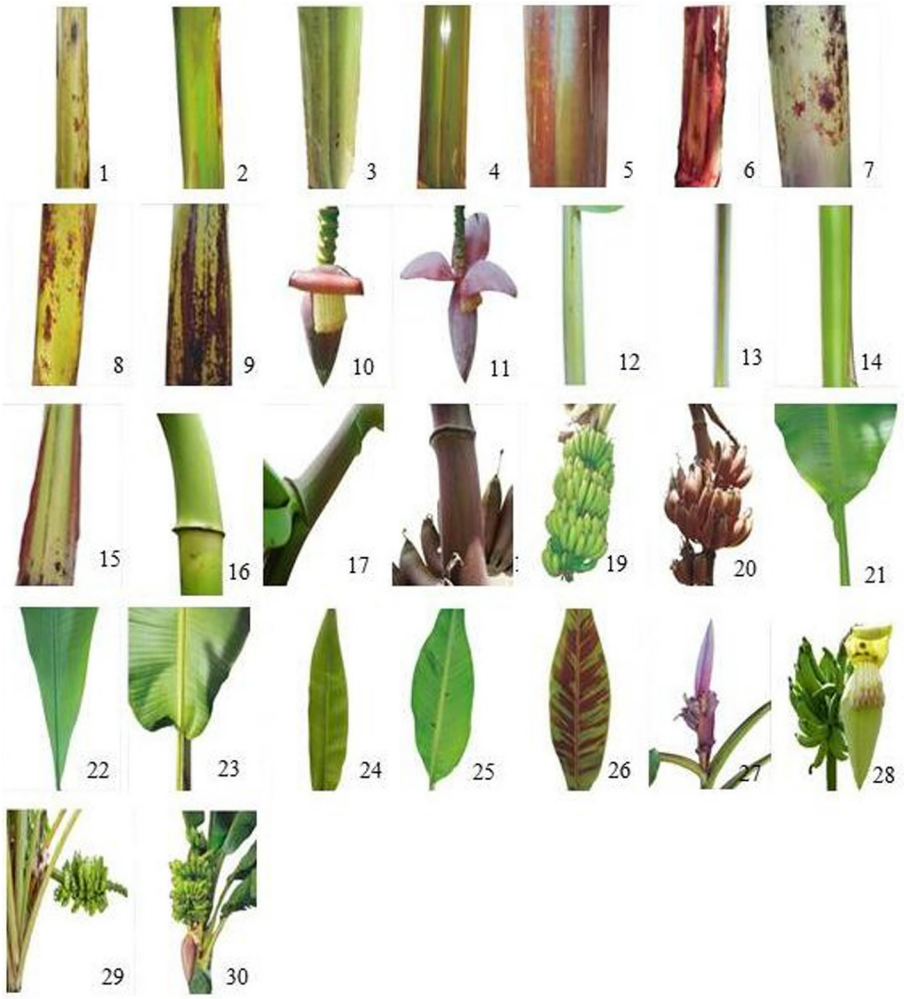


**Figure 1.** Multy-category descriptors for the characterization of banana germplasm at Embrapa Cassava and Fruits. Pseudostem color (1. green-yellow, 2. light green, 3. medium green, 4. dark green, 5. green-red, and 6. red); predominant underlying color of the pseudostem (7. light brown, 8. dark brown and 9. black); bract behavior before falling (10. revolute and 11. not revolute); petiole margins (12. winged and undulating, 13. winged, 14. not clasping the pseudostem, winged and clasping the pseudostem, and 15. not winged and clasping the pseudostem); peduncule color (15. green-yellow, 16. tinted with red, and 18. brown); bunch shape (19. cylindrical and 20. asymmetric - bunch axis is nearly straight); shape of leaf blade base (21. one side rounded, one pointed, 22. both sides pointed, and 23. both sides rounded); blotches at the leaf blade of the sukers (24. absent, 25. little, and 26. mean); bunch position (27. hanging vertically, 28. slightly angled, 29. horizontal, and 30. hanging at angle 45°).

mainly for number of hands per bunch (NHB; 1 to 12 bunches), number of fingers per bunch (NFB; 5 to 24 fruits), fruit length (FRL; 3.63 to 27.75 cm), and bunch length (BUL; 9 to 90 cm). The NHB is of great interest for producers and of fundamental importance for banana genetic breeding because bunches are the commercial unit used. In addition, an increased NHB may increase the weight of the bunch, a character that expresses genotype productivity (Silva et al., 2002).

**Table 2.** Summary of the analysis of variance based on the F-test, average, minimum, maximum, and coefficient of variation (CV) for the quantitative characteristics of banana accessions.

| Quantitative | Abbreviations. | Mean square | | Average | Minimum | Maximum | CV (%) |
|---|---|---|---|---|---|---|---|
| | | Accession | Error | | | | |
| Plant height (cm) | PLH | 11104.89** | 603.82 | 235.09 | 80.00 | 480.00 | 10.45 |
| Pseudostem diameter (cm) | PSD | 68.14** | 6.00 | 15.31 | 5.30 | 48.00 | 15.99 |
| Crown (cm) | CRO | 7.59** | 1.96 | 5.87 | 1.25 | 11.23 | 23.85 |
| Number of suckers | NUS | 10.25** | 1.84 | 3.75 | 0.00 | 12.00 | 36.19 |
| Petiole length (cm) | PEL | 242.21** | 46.85 | 47.20 | 20.00 | 82.00 | 14.49 |
| Petiole diameter (cm) | PED | 1.62** | 0.26 | 3.74 | 1.40 | 6.17 | 13.72 |
| Leaf blade length (cm) | LBL | 2504.88** | 390.51 | 169.94 | 74.50 | 249.00 | 11.62 |
| Leaf blade width (cm) | LBW | 201.51** | 28.14 | 49.76 | 23.33 | 81.00 | 10.66 |
| Stalk length (cm) | STL | 889.48** | 50.74 | 41.16 | 8.00 | 107.00 | 17.30 |
| Stalk diameter (cm) | STD | 2.71** | 0.26 | 4.23 | 1.85 | 6.90 | 12.19 |
| Internode length of the bunch (cm) | ILB | 4.31** | 0.77 | 5.62 | 2.20 | 10.00 | 15.69 |
| Number of hands per bunch | NHB | 5.92** | 0.98 | 6.01 | 1.00 | 12.00 | 16.46 |
| Raquis diameter (cm) | RAD | 0.37** | 0.07 | 2.13 | 1.00 | 3.50 | 12.55 |
| Bract scars on rachis (cm) | BSR | 0.09** | 0.00 | 0.73 | 0.32 | 1.30 | 12.97 |
| Male bud length (cm) | MBL | 71.95** | 10.45 | 19.98 | 8.50 | 35.50 | 16.17 |
| Male bud diameter (cm) | MBD | 11.40** | 0.89 | 7.10 | 2.10 | 15.50 | 13.31 |
| Number of fingers per bunch | NFB | 20.89** | 4.71 | 13.87 | 5.00 | 24.00 | 15.65 |
| Fruit length (cm) | FRL | 36.905** | 2.69 | 12.58 | 3.63 | 27.75 | 13.04 |
| Bunch length (cm) | BUL | 404.76** | 79.36 | 37.17 | 9.00 | 90.00 | 23.96 |
| Radial calibration of the finger (cm) | RCF | 2.36** | 0.26 | 3.16 | 1.15 | 6.95 | 16.30 |
| Bunch diameter (cm) | BUD | 241.79** | 16.25 | 25.12 | 5.43 | 56.00 | 16.05 |
| Lateral calibration of the finger (cm) | LCF | 1.78** | 0.10 | 3.00 | 0.95 | 5.34 | 10.51 |
| Fruit peel thickness (cm) | FPT | 0.01** | 0.01 | 0.20 | 0.10 | 0.43 | 15.33 |
| Fruit pedicel width (cm) | FPW | 0.19$^{ns}$ | 0.19 | 0.94 | 0.26 | 2.11 | 52.66 |
| Fruit pedicel length (cm) | FPL | 1.48** | 0.07 | 1.56 | 0.39 | 3.90 | 17.37 |
| Fruit apex length (cm) | FAL | 0.84** | 0.04 | 1.04 | 0.20 | 3.56 | 20.95 |
| Presence of seed | PSE | 3.45** | 0.01 | 1.80 | 1.00 | 4.00 | 6.14 |

**Significant at 1%; ns = non-significant.

Accessions from the Embrapa germplasm collection with values above the average for agronomic characters (except PLH and NUS) have the potential for use in breeding programs. Variation was observed for these characteristics, which allows the identification and use of accessions directly in banana breeding focused on diploids or the development of secondary triploid and tetraploid hybrids. It is important to stress that the variation detected between genotypes allows estimations of genetic variability between accessions.

## Selection of quantitative morphoagronomic descriptors

Table 3 shows estimates of the eigenvalues associated with major components and their respective relative and cumulative variances obtained for the 27 quantitative morphological characters. The first 2 principal components explained 55.01% of the total variation accumulated. The relative variances and their respective percentages show that much of the variation was concentrated up to the 17th principal component, accounting for 96.31% of all variation available in the germplasm collection. Variance distribution is associated with the nature and number of characters used in the analysis, and it is concentrated in the first principal components only when few descriptors are used (Pereira et al., 1992).

Using the direct method proposed by Jolliffe (1972, 1973), we chose the variable lateral calibration of the finger (LCF) first for disposal, as it presented the highest weighting in the module with the last principal component (-0.644). The characters for disposal that followed were stalk di-

**Table 3.** Estimates of the eigenvalues associated with the principal components and their accumulated relative variances obtained from 27 quantitative descriptors that were evaluated in 77 banana accessions.

| Component | Eigenvalues | % Relative | % Accumulated |
|---|---|---|---|
| 1 | 12.3957 | 45.91 | 45.91 |
| 2 | 2.4566 | 9.10 | 55.01 |
| 3 | 1.8930 | 7.01 | 62.02 |
| 4 | 1.7238 | 6.38 | 68.40 |
| 5 | 1.4681 | 5.44 | 73.84 |
| 6 | 1.0128 | 3.75 | 77.59 |
| 7 | 0.7938 | 2.94 | 80.53 |
| 8 | 0.6821 | 2.53 | 83.06 |
| 9 | 0.6473 | 2.40 | 85.46 |
| 10 | 0.5313 | 1.97 | 87.42 |
| 11 | 0.4602 | 1.70 | 89.13 |
| 12 | 0.4342 | 1.61 | 90.74 |
| 13 | 0.3499 | 1.30 | 92.03 |
| 14 | 0.3306 | 1.22 | 93.26 |
| 15 | 0.2955 | 1.09 | 94.35 |
| 16 | 0.2846 | 1.05 | 95.41 |
| 17 | 0.2450 | 0.91 | 96.31 |
| 18 | 0.1790 | 0.66 | 96.98 |
| 19 | 0.1646 | 0.61 | 97.59 |
| 20 | 0.1506 | 0.56 | 98.14 |
| 21 | 0.1282 | 0.47 | 98.62 |
| 22 | 0.1121 | 0.42 | 99.03 |
| 23 | 0.0813 | 0.30 | 99.33 |
| 24 | 0.0674 | 0.25 | 99.58 |
| 25 | 0.0496 | 0.18 | 99.77 |
| 26 | 0.0414 | 0.15 | 99.92 |
| 27 | 0.0212 | 0.08 | 100.00 |

ameter (STD), leaf blade length (LBL), and male bud diameter (MBD), whose highest eigenvalues in the module occurred in principal components 26, 25, and 24, respectively (Table 4). The direct method considered 20 characters redundant according to the following sequence of disposal: LCF, STD, LBL, MBD, fruit peel thickness (FPT), petiole diameter (PED), stalk length (STL), radial calibration of the finger (RCF), crown (CRO), raquis diameter (RAD), fruit pedicel length (FPL), petiole length, presence of seed (PSE), FPW, FRL, fruit apex length (FAL), bunch diameter (BUD), PSD, NHB, and NFB. This procedure can be considered drastic, because it eliminated 20 of the 27 quantitative morphological characters used as descriptors in banana.

In the disposal carried out by selection with reanalysis (Cury, 1993), only nine characters were indicated. From the last descriptor eliminated (NFB), the characters (LCF, STD, LBL, MBD, FPT, PED, STL, RCF, RAD, FPL, FRL, bract scars on rachis, NUS, leaf blade width, PLH, male bud length, internode length of the bunch, BUL) started to break the preestablished norms, showing significant correlation with a variable already discarded (Table 5).

Based on the simultaneous analysis of the 2 procedures, 9 traits were coincident and were part of the final disposal - namely, CRO, FPL, PSE, FPW, FAL, BUD, PSD, NHB, NFB. This decision reduced the rigidity of selection and minimized possible errors in the disposal procedure, also reducing 33% of the characters evaluated and, consequently, the costs and labor necessary for evaluation and characterization.

The analysis of the 2-disposal procedures demonstrated that direct selection was less consistent, because it eliminated 20 of the 27 quantitative morphological descriptors considered important in the characterization of the banana germplasm, including descriptors used to evalu-

**Table 4.** Estimates of the weighting coefficients associated with the principal components with eigenvalues less than 0.70 and identification of the quantitative descriptors to be discarded in each component (in bold) for direct selection in 77 banana accessions.

| | Descriptor | Principal components | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 27 | 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 |
| 1 | PLH | 0.153 | 0.302 | -0.276 | 0.172 | 0.108 | -0.006 | -0.353 | -0.065 | -0.053 | 0.147 | 0.045 | -0.209 | -0.215 | 0.155 | -0.071 | -0.255 | -0.131 | 0.054 | -0.046 | -0.190 |
| 2 | **PSD** | -0.071 | 0.018 | -0.020 | -0.015 | -0.216 | 0.094 | -0.018 | -0.121 | -0.221 | 0.011 | -0.018 | 0.253 | 0.004 | 0.047 | -0.027 | 0.141 | 0.366 | **-0.300** | - | - |
| 3 | **CRO** | 0.094 | -0.189 | -0.027 | -0.010 | -0.013 | -0.252 | 0.163 | -0.216 | **0.444** | - | - | - | - | - | - | - | - | - | - | - |
| 4 | NUS | 0.004 | 0.088 | 0.071 | -0.017 | 0.060 | 0.009 | -0.100 | -0.042 | -0.141 | 0.181 | 0.311 | -0.239 | -0.107 | -0.072 | -0.075 | 0.220 | -0.327 | -0.082 | -0.130 | 0.051 |
| 5 | **PEL** | -0.003 | 0.118 | -0.075 | -0.104 | -0.113 | 0.310 | 0.059 | -0.001 | -0.121 | 0.127 | -0.063 | **0.413** | - | - | - | - | - | - | - | - |
| 6 | **PED** | 0.078 | 0.431 | 0.117 | -0.025 | 0.214 | **-0.459** | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 7 | **LBL** | -0.244 | -0.461 | **0.484** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 8 | LBW | 0.034 | -0.073 | -0.023 | -0.198 | 0.040 | -0.209 | 0.267 | -0.048 | -0.017 | 0.119 | 0.007 | -0.224 | 0.303 | -0.251 | -0.242 | 0.366 | -0.246 | 0.270 | -0.170 | 0.085 |
| 9 | **STL** | -0.107 | 0.025 | -0.052 | 0.019 | 0.061 | -0.060 | **0.411** | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 10 | **STD** | 0.202 | **-0.510** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 11 | ILB | -0.029 | 0.105 | -0.214 | -0.198 | -0.265 | -0.261 | -0.012 | 0.128 | 0.156 | -0.285 | 0.296 | 0.250 | 0.011 | 0.323 | -0.222 | 0.071 | 0.005 | 0.232 | 0.010 | 0.129 |
| 12 | **NHB** | -0.075 | 0.198 | -0.096 | -0.247 | -0.369 | 0.098 | -0.017 | -0.028 | 0.222 | -0.305 | 0.088 | -0.158 | -0.039 | -0.001 | 0.018 | 0.007 | -0.003 | -0.201 | **0.376** | - |
| 13 | **RAD** | -0.052 | 0.081 | 0.230 | 0.133 | 0.222 | 0.020 | -0.354 | -0.217 | -0.033 | **-0.411** | - | - | - | - | - | - | - | - | - | - |
| 14 | BSR | 0.035 | 0.120 | 0.056 | 0.025 | -0.099 | 0.273 | 0.257 | -0.332 | 0.224 | -0.081 | -0.268 | 0.044 | -0.302 | 0.049 | -0.159 | -0.090 | -0.268 | 0.156 | -0.306 | -0.238 |
| 15 | MBL | 0.016 | 0.129 | 0.049 | 0.524 | -0.285 | -0.147 | 0.148 | 0.177 | -0.003 | 0.023 | 0.360 | -0.004 | 0.111 | 0.045 | 0.374 | 0.045 | -0.121 | 0.134 | -0.116 | -0.103 |
| 16 | **MBD** | -0.018 | -0.113 | -0.054 | **-0.553** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 17 | **NFB** | -0.026 | 0.099 | -0.112 | 0.024 | -0.136 | 0.234 | 0.114 | -0.110 | -0.075 | 0.114 | -0.030 | -0.007 | 0.184 | 0.038 | 0.210 | -0.334 | -0.018 | 0.062 | -0.052 | **0.582** |
| 18 | **FRL** | -0.007 | 0.098 | 0.071 | -0.053 | 0.210 | 0.084 | -0.336 | -0.130 | -0.003 | -0.315 | -0.099 | 0.278 | 0.065 | -0.210 | **0.545** | - | - | - | - | - |
| 19 | BUL | 0.192 | -0.199 | 0.125 | 0.331 | 0.129 | -0.009 | 0.017 | -0.334 | 0.257 | 0.309 | 0.036 | 0.282 | -0.004 | 0.312 | -0.129 | 0.183 | 0.036 | -0.050 | 0.352 | 0.131 |
| 20 | **RCF** | 0.596 | 0.046 | 0.270 | -0.121 | 0.023 | 0.378 | 0.150 | **0.374** | - | - | - | - | - | - | - | - | - | - | - | - |
| 21 | **BUD** | -0.037 | 0.040 | -0.051 | 0.084 | 0.059 | 0.183 | 0.047 | 0.169 | 0.189 | -0.026 | -0.098 | -0.246 | -0.377 | -0.183 | 0.053 | 0.124 | **0.449** | - | - | - |
| 22 | **LCF** | **-0.644** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 23 | **FPT** | -0.030 | 0.054 | 0.122 | -0.053 | **-0.388** | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 24 | **FPW** | -0.054 | 0.011 | -0.081 | -0.069 | 0.081 | -0.016 | 0.130 | 0.007 | -0.030 | 0.082 | 0.016 | -0.216 | 0.101 | **0.340** | - | - | - | - | - | - |
| 25 | **FPL** | -0.114 | -0.121 | 0.167 | -0.196 | -0.306 | 0.046 | -0.166 | 0.050 | -0.192 | 0.146 | **-0.433** | - | - | - | - | - | - | - | - | - |
| 26 | **FAL** | 0.070 | -0.039 | -0.158 | -0.056 | 0.237 | -0.132 | 0.217 | -0.095 | -0.153 | -0.221 | 0.067 | 0.242 | 0.103 | -0.004 | -0.132 | **-0.454** | - | - | - | - |
| 27 | **PSE** | 0.040 | -0.001 | -0.025 | -0.017 | -0.009 | -0.127 | 0.140 | 0.044 | 0.199 | -0.152 | -0.114 | 0.035 | **-0.402** | - | - | - | - | - | - | - |

For abbreviations, see Table 2.

**Table 5.** Estimates of Pearson correlation coefficients between selected and discarded quantitative descriptors evaluated in 77 banana accessions.

| Descriptors selected | Descriptors discarded | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | NFB | NHB | PSD | BUD | FAL | PFW | PSE | FPL | CRO |
| BUL | 0.18 | 0.56** | 0.55** | 0.59** | 0.39** | 0.37** | -0.55** | 0.41 | 0.30* |
| ILB | 0.05 | -0.17 | 0.39** | 0.54 | 0.15 | 0.23* | -0.48** | 0.13 | 0.14 |
| MBL | 0.01 | 0.36** | 0.41 | 0.49 | 0.27** | 0.26** | -0.17* | 0.10** | 0.28* |
| PLH | 0.04 | 0.25* | 0.47** | 0.45** | 0.39** | 0.41 | -0.26* | 0.04** | 0.03** |
| LBW | 0.22* | 0.47** | 0.26** | 0.44** | 0.28* | 0.35** | -0.36** | 0.13** | 0.14** |
| NUS | 0.16 | 0.12 | 0.31** | 0.30** | 0.19 | 0.08* | -0.24* | 0.16 | 0.49** |
| BSR | 0.06 | 0.27* | 0.45** | 0.43** | 0.32** | 0.41** | -0.53** | 0.08 | 0.25* |
| FRL | 0.02 | 0.25* | 0.51** | 0.59** | 0.40** | 0.47** | -0.52** | 0.05** | 0.39** |
| FPL | -0.19 | 0.08 | 0.40** | 0.65** | 0.58** | 0.32** | -0.31** | 0.26* | 0.49 |
| RAD | 0.25* | 0.30** | 0.38** | 0.42** | 0.27* | 0.26** | -0.17 | 0.41** | 0.33** |
| RCF | -0.02 | 0.18 | 0.59 | 0.63** | 0.43** | 0.44** | -0.48** | 0.16 | 0.54** |
| STL | 0.01 | 0.14 | 0.43** | 0.57** | 0.20 | 0.24* | -0.45** | 0.20** | 0.55** |
| PED | 0.42** | 0.57** | 0.58** | 0.47** | 0.22* | 0.28* | -0.53** | 0.20** | 0.38** |
| FPT | -0.12 | 0.11 | 0.35** | 0.58** | 0.54** | 0.53** | -0.34** | 0.19 | 0.32** |
| MBD | 0.19 | 0.36** | 0.44** | 0.44** | 0.26** | 0.19 | -0.29* | 0.26* | 0.26* |
| LBL | 0.40** | 0.53** | 0.51** | 0.45** | 0.33** | 0.45** | -0.39** | 0.40** | 0.47** |
| STD | 0.28* | 0.52** | 0.55** | 0.53** | 0.31** | 0.39** | -0.52** | 0.22** | 0.35** |
| LCF | 0.58 | 0.22 | 0.58** | 0.60** | 0.46** | 0.39** | -0.48** | 0.18 | 0.54** |

*,**Significant at 5 and 1%, respectively. For abbreviations, see Table 2.

ate the production of fruits, such as NFB and NHB. However, the selection with reanalysis was more appropriate, although it also suggested the disposal of the descriptors NFB and NHB.

Regarding estimates of the Pearson correlation, between the set of redundant descriptors and the set of 18 selected, we observed that the disposal revealed no significant loss of information because the redundant characteristics exhibited high binding to at least one descriptor selected (see Table 5). Furthermore, the 2 descriptors of fruits disposed in this study, NFB and NHB, are correlated with other descriptors selected (LCF, BUL, PED) and therefore should cause no loss of information.

## Selection of multi-category morphoagronomic descriptors

The percentage frequency of each category and H of the characters were evaluated for the selection of 65 multi-category descriptors using the coefficient of entropy of Renyi (1961). The criteria adopted for the disposal of a particular descriptor were "low value for H (≤1.00)" and "more than 50% of the accessions classified in one of the descriptor classes".

Table 6 shows the multi-category descriptors, phenotypic classes, percentage frequency of the accessions in each class, and H. The combination of the information and the low values of H (≤1.00) along with the frequency of accessions in the same class within a certain descriptor (>50%) suggested the disposal of 33 traits: DES, PCP, BDP, IAP, NAP, CIS, CLS, BLS, SLB, WUS, RPO, floral remains and bracts, male bud shape, CBS, MBA, BBF, MBL, CTC, CTA, LCT, RTL, FTC, FAP, SSH, OAN, OSH, BUH, TSF, IFP, MPF, AFP, FFF, and FWP (as in Table 6).

The descriptors floral remains and bracts, male bud shape, and mature fruit peel color were maintained, even with H ≤ 1.00, because they allow the differentiation between ploidies and banana subgroups. Thus, we selected 30 multi-category morphoagronomic descriptors for disposal, a reduction of approximately 46%.

**Table 6.** Multi-category descriptors evaluated, phenotypical classes, frequency, and entropy (H).

| Descriptors | Phenotypical classes | Frequency | H |
|---|---|---|---|
| Development of suckers (DES) | 1. Taller than parent plant | 96.72 | 0.18 |
| | 2. More than 3/4 of the height of the parent plant | 1.64 | |
| | 3. Between 1/4 and 3/4 of the height of the parent plant | 0.55 | |
| | 4. Inhibited | 1.09 | |
| Pseudostem wax (PSW) | 1. Very waxy | 27.87 | 1.16 |
| | 2. Moderately | 46.99 | |
| | 3. Very few wax | 21.86 | |
| Predominant underlying color of the pseudostem (PCP) | 1. Light brown | 28.42 | 0.94 |
| | 2. Dark brown | 58.47 | |
| | 3. Black | 13.11 | |
| Blotches density of the pseudostem (BDP) | 1. Continuous | 6.01 | 0.87 |
| | 2. High | 1.64 | |
| | 3. Fuzzy | 71.58 | |
| | 4. Discrete | 18.58 | |
| | 5. Low | 2.19 | |
| | 6. Very low | 5.66 | |
| Intensity of pigmentation on pseudostem (IAP) | 1. Intense | 7.65 | 0.80 |
| | 2. Mean | 21.31 | |
| | 3. Weak | 70.49 | |
| Narrowing of the pseudostem (NAP) | 1. Intense | 9.84 | 0.94 |
| | 2. Mean | 50.27 | |
| | 3. Weak | 39.89 | |
| Pseudostem color (PSC) | 1. Green-yellow | 7.65 | 1.19 |
| | 2. Light green | 63.39 | |
| | 3. Medium green | 9.84 | |
| | 4. Dark green | 2.19 | |
| | 5. Green-red | 12.57 | |
| | 6. Red | 4.37 | |
| Color of the inner surface of the sheath (CIS) | 1. Purple | 0.55 | 0.96 |
| | 2. Red | 2.73 | |
| | 3. Pink | 10.93 | |
| | 4. Pale | 62.84 | |
| | 5. Green | 16.39 | |
| Leaf habit (LEH) | 1. Erect | 58.47 | 1.00 |
| | 2. Intermediate | 25.14 | |
| | 3. Drooping | 15.30 | |
| Petiole canal (PTC) | 1. Open with margins spreading | 27.32 | 1.01 |
| | 2. Wide with erect margins | 57.38 | |
| | 3. Straight with erect margins | 13.66 | |
| | 4. Margins curved inward | 1.64 | |
| Petiole margins (PMA) | 1. Winged and undulating | 6.56 | 1.19 |
| | 2. Winged and not clasping the pseudostem | 74.32 | |
| | 3. Winged and clasping the pseudostem | 16.94 | |
| | 4. Not winged and clasping the pseudostem | 2.19 | |
| Scarious petiole margin at basis (SPM) | 1. Absent | 4.90 | 0.78 |
| | 2. Little | 25.10 | |
| | 3. Mean | 47.12 | |
| | 4. Much | 22.88 | |
| Petiole margins (PMA) | 1. Purple | 44.81 | 1.11 |
| | 2. Red-pink | 30.05 | |
| | 3. Green | 1.09 | |
| | 4. Brown | 24.04 | |
| Color of leaf lower surface (CLS) | 1. Green-yellow | 55.19 | 0.87 |
| | 2. Dark green | 39.34 | |
| Blotches at the leaf blade of the sukers (BLS) | 1. Absent | 80.87 | 0.56 |
| | 2. Little | 16.94 | |
| | 3. Mean | 2.19 | |
| Shape of leaf blade base (SLB) | 1. Both sides rounded | 57.92 | 0.94 |
| | 2. One side rounded, one pointed | 13.11 | |
| | 3. Both sides pointed | 28.96 | |
| Wax on leaf upper surface (WUS) | 1. Very little or no visible sign of wax | 96.17 | 0.16 |
| | 2. Few wax | 3.83 | |

Continued on next page

**Table 6.** Continued.

| Descriptors | Phenotypical classes | Frequency | H |
|---|---|---|---|
| Wax on leaf lower surface (WLS) | 1. Very little or no visible sign of wax | 42.08 | 1.22 |
| | 2. Few wax | 27.32 | |
| | 3. Moderately waxy | 25.68 | |
| | 4. Very waxy | 4.92 | |
| Peduncule color (STC) | 1. Green-yellow | 20.77 | 1.79 |
| | 2. Medium green | 18.03 | |
| | 3. Green | 13.11 | |
| | 4. Dark green | 3.83 | |
| | 5. Tinted with brown | 18.03 | |
| | 6. Tinted with red | 22.95 | |
| | 7. Other | 3.28 | |
| Peduncle hairiness (PHA) | 1. Hairless | 27.32 | 1.28 |
| | 2. Slightly hairy | 38.8 | |
| | 3. Very hairy, short hairs | 25.68 | |
| | 4. Very hairy, long hairs | 8.20 | |
| Bunch position (BPO) | 1. Hanging vertically | 9.29 | 1.08 |
| | 2. Slightly angled | 13.66 | |
| | 3. Horizontal | 11.48 | |
| | 4. Hanging at angle 45° | 64.48 | |
| Rachis position (RPO) | 1. Falling vertically | 67.21 | 0.95 |
| | 2. At an angle | 20.22 | |
| | 3. With a curve | 1.64 | |
| | 4. Horizontal | 4.37 | |
| | 5. Erect | 6.56 | |
| Floral remains and bracts (FRB) | 1. Absent | 78.69 | 0.67 |
| | 2. Little | 10.93 | |
| | 3. Mean | 10.38 | |
| | 4. Much | 8.20 | |
| Rachis color (RCO) | 1. Dark green | 8.74 | 1.32 |
| | 2. Green with other colors in the youth share | 30.05 | |
| | 3. Green with other colors in the cushions | 43.72 | |
| | 4. Green | 14.75 | |
| Male bud shape (MBS) | 1. Male bud shape | 25.68 | 0.90 |
| | 2. Lanceolate | 2.19 | |
| | 3. Ovoid | 7.10 | |
| | 4. Intermediate | 65.03 | |
| Curvature below the shoulder of the heart (CBS) | 1. Convex | 1.64 | 0.37 |
| | 2. No curve | 8.20 | |
| | 3. Concave | 90.16 | |
| Male bud apex shape (MBA) | 1. Pointed | 66.12 | 0.37 |
| | 2. Slightly pointed | 33.88 | |
| Bract imbrication (BIM) | 1. Old bracts overlap at apex of bud | 34.97 | 1.26 |
| | 2. Young bracts slightly overlap | 38.25 | |
| | 3. Young bracts greatly overlap | 26.78 | |
| Bract apex shape (BAS) | 1. Pointed | 37.70 | 1.22 |
| | 2. Slightly pointed | 39.34 | |
| | 3. Intermediate | 18.03 | |
| | 4. Obtuse | 3.83 | |
| | 5. Obtuse and split | 1.09 | |
| Bract base shape (FBS) | 1. Small shoulder | 22.95 | 1.07 |
| | 2. Medium | 40.44 | |
| | 3. Large shoulder | 36.61 | |
| Bract behavior before falling (BBF) | 1. Revolute | 87.98 | 0.37 |
| | 2. Not revolute | 12.02 | |
| Male bract lifting (MBL) | 1. Not lifting from male bud (bracts are persistent) | 9.84 | 0.61 |
| | 2. Lifting 2 or more at a time | 5.46 | |
| | 3. Lifting 1 at a time | 84.70 | |
| Wax on the bract (WBR) | 1. Very little or no visible sign of wax | 21.31 | 1.27 |
| | 2. Very few wax | 46.45 | |
| | 3. Moderately waxy | 18.58 | |
| | 4. Very waxy | 13.66 | |

Continued on next page

**Table 6.** Continued.

| Descriptors | Phenotypical classes | Frequency | H |
|---|---|---|---|
| Color of the bract external face (CBE) | 1. Orange-red | 3.28 | 1.63 |
| | 2. Red-purple | 3.28 | |
| | 3. Violeta café | 20.22 | |
| | 4. Pink-purple | 33.33 | |
| | 5. Red | 24.04 | |
| | 6. Purple | 13.66 | |
| | 7. Green | 1.09 | |
| | 8. Other | 1.09 | |
| Color of the bract internal face (CBI) | 1. Light red | 25.68 | 1.76 |
| | 2. Opaque red | 14.75 | |
| | 3. Dark red | 33.33 | |
| | 4. Pink | 1.09 | |
| | 5. Ivory | 7.65 | |
| | 6. Purple violet | 10.93 | |
| | 7. Purple | 0.55 | |
| | 8. Dark purple | 2.19 | |
| | 9. Yellow-green | 0.55 | |
| | 10. = Bract color | 3.28 | |
| Pollen (PLL) | 1. Absent | 27.87 | 1.29 |
| | 2. Little | 41.53 | |
| | 3. Mean | 13.66 | |
| | 4. Much | 16.94 | |
| Compound tepal basic color (CTC) | 1. White | 6.56 | 0.59 |
| | 2. Cream | 84.70 | |
| | 3. Yellow | 2.73 | |
| Compound tepal anthocyanin (CTA) | 1. Very few or no visible sign of pigmentation | 74.86 | 0.62 |
| | 2. Very few or no visible sign of pigmentation | 12.57 | |
| | 3. Presence of pink | 12.57 | |
| Lobe color of compound tepal (LCT) | 1. Orange | 1.09 | 0.88 |
| | 2. Orange-yellow | 7.65 | |
| | 3. Yellow | 65.03 | |
| | 4. Light yellow | 26.23 | |
| Relationship of the free tepal from the perigon (RTL) | 1. More than half of perigony | 27.32 | 0.81 |
| | 2. = to half of perigony | 66.12 | |
| | 3. Less than half of perigony | 6.56 | |
| Free tepal color (FTC) | 1. Translucent white | 83.06 | 0.55 |
| | 2. Opaque white | 3.83 | |
| | 3. Tinted with pink | 13.11 | |
| Free tepal apex shape (FTS) | 1. Rectangular | 36.61 | 1.07 |
| | 2. Oval | 21.86 | |
| | 3. Rounded | 41.53 | |
| Wrinkle traverses close to the tepal apex (WTA) | 1. Absent | 32.24 | 1.26 |
| | 2. Weak | 37.16 | |
| | 3. Mean | 23.5 | |
| | 4. Strong | 7.10 | |
| FTS form of the tepal apex (FAP) | 1. Narrow | 71.04 | 0.74 |
| | 2. Wide | 23.5 | |
| Filament color (FIC) | 1. White | 10.93 | 1.05 |
| | 2. Cream | 68.31 | |
| | 3. Light yellow | 9.29 | |
| | 4. Opaque yellow | 5.46 | |
| | 5. Black | 6.01 | |
| Anther color (ACO) | 1. White | 4.92 | 1.60 |
| | 2. Brown/rusty brown | 19.67 | |
| | 3. Cream | 19.67 | |
| | 4. Yellow | 4.37 | |
| | 5. Pink/pink-purple | 40.98 | |
| | 6. Red | 4.37 | |
| | 7. Black (anthers aborted) | 6.01 | |
| Stigma color (SCO) | 1. Orange | 8.20 | 1.24 |
| | 2. Orange-yellow | 31.69 | |
| | 3. Light yellow | 49.18 | |
| | 4. Pallid | 5.46 | |
| | 5. Other | 5.46 | |

Continued on next page

**Table 6.** Continued.

| Descriptors | Phenotypical classes | Frequency | H |
|---|---|---|---|
| Style shape (SSH) | 1. Straight | 79.78 | 0.71 |
| | 2. Curved under stigma | 6.01 | |
| | 3. Curved at the base | 9.84 | |
| | 4. Curved twice | 4.37 | |
| Stigma shape (STH) | 1. Rounded | 21.86 | 1.26 |
| | 2. Spatulate | 44.81 | |
| | 3. Slightly lobulated | 24.04 | |
| | 4. Strongly lobed | 9.29 | |
| Ovary pigmentation (OAN) | 1. Absent | 73.22 | 0.58 |
| | 2. Present | 26.78 | |
| Ovary shape (OSH) | 1. Straight | 25.68 | 0.57 |
| | 2. Arched | 74.32 | |
| Yellow sigatoka (YES) | 1. No symptoms | 42.08 | 1.52 |
| | 2. Symptoms on 1-10% | 3.83 | |
| | 3. Symptoms on 11-30% | 11.48 | |
| | 4. Symptoms on 31-50% | 19.13 | |
| | 5. Symptoms on 51-70% | 5.46 | |
| | 6. Symptoms on over 70% of the leaf | 18.03 | |
| Fruit shape (FRS) | 1. Straight (or slightly curved) | 48.63 | |
| | 2. Straight in the distal part | 42.62 | |
| | 3. Curved (sharp curve) | 7.10 | 0.97 |
| | 4. Curved in 'S' shape (double curvature) | 1.64 | |
| Bunch shape (BUH) | 1. Cylindrical | 58.47 | 0.83 |
| | 2. Asymmetric - bunch axis is nearly straight | 36.61 | |
| | 3. With a curve in the bunch axis | 4.92 | |
| Transverse section of fruit (TSF) | 1. Pronounced ridges | 28.96 | 0.91 |
| | 2. Slightly ridged | 60.11 | |
| | 3. Rounded | 10.93 | |
| Fruit apex (FRA) | 1. Pointed | 9.29 | 1.03 |
| | 2. Lengthily pointed | 67.21 | |
| | 3. Blunt-tipped | 4.92 | |
| | 4. Bottle-necked | 15.85 | |
| | 5. Rounded | 2.73 | |
| Remains of flower relicts at fruit apex (RFR) | 1. Without any floral relicts | 46.45 | 1.05 |
| | 2. Persistent style | 20.22 | |
| | 3. Base of the style prominent | 33.33 | |
| Immature fruit peel color (IFP) | 1. Dark green | 15.85 | 0.56 |
| | 2. Light green | 81.42 | |
| | 3. Yellow | 2.73 | |
| Pulp color before maturity (PCM) | 1. White | 41.53 | 1.29 |
| | 2. Yellow | 27.32 | |
| | 3. Cream | 13.11 | |
| | 4. Orange | 18.03 | |
| Mature fruit peel color (MFP) | 1. Opaque yellow | 0.55 | |
| | 2. Yellow | 83.61 | |
| | 3. Green and pink, red or purple | 3.28 | 0.62 |
| | 4. Pink, red or purple | 2.73 | |
| | 5. Green | 9.84 | |
| Pulp color at maturity (PCM) | 1. White | 15.3 | 1.58 |
| | 2. Opaque white | 14.21 | |
| | 3. Cream | 27.87 | |
| | 4. Yellow | 20.22 | |
| | 5. Orange | 22.4 | |
| Adherence of the fruit peel (AFP) | 1. Adherence of the fruit peel | 73.22 | 0.58 |
| | 2. Fruit does not peel easily | 26.78 | |
| Fruits fall from hands (FFF) | 1. Deciduous | 60.66 | 0.92 |
| | 2. Intermediary | 25.68 | |
| | 3. Persistent | 13.66 | |

Continued on next page

**Table 6.** Continued.

| Descriptors | Phenotypical classes | Frequency | H |
|---|---|---|---|
| Firmness of flesh with peel (FFP) | 1. Not evaluated (presence of seeds) | 16.39 | 1.30 |
| | 2. Flimsy | 32.24 | |
| | 3. Consistently | 37.70 | |
| | 4. Very consistent | 13.66 | |
| Firmness of flesh without peel (FWP) | 1. Not evaluated (presence of seeds) | 16.39 | 0.92 |
| Consistency of the pulp (without peel) (FWP) | 2. Flimsy | 2.73 | |
| | 3. Consistently | 68.31 | |
| | 4. Very consistent | 12.57 | |

Classes adapted according IPGRI (1996), Silva et al. (1999).

## Efficiency of disposal

We estimated the correlation between the matrices obtained from the 92 descriptors in total and the 53 selected, which corresponded to a reduction of 42% in the number of descriptors evaluated. The correlation achieved was 0.83 ($P \leq 0.01$), which demonstrated that the reduced number of descriptors had no effect on the study of genetic variability among the accessions of banana (data not shown).

Dias et al. (1997) characterized cacao clones and verified small changes in the formation of groups using original and remaining characters and noticed the efficiency of the selection methodology after reanalysis for the disposal of characters. Araújo et al. (2002) also analyzed the efficiency of disposal based on the formation of groups in fruits of cupuaçu tree clones. The authors found little change in the number and composition of the groups. Oliveira et al. (2006) observed similar behavior when describing accessions of açaí palm, in which the number of groups formed was higher when only the selected descriptors were used.

## Phenotypic diversity considering the morphoagronomic descriptors selected

Figure 2 presents the dendrogram of genetic dissimilarity among the 77 accessions of banana considering the joint analysis of 53 quantitative and multi-category morphoagronomic descriptors selected and carried out using the Ward-MLM procedure. With *pseudo-F* and *pseudo-t*$^2$ statistics considered, the ideal number of groups was 3: G1, formed by 24 accessions, including 18 AA diploids, 2 BB diploids, one each of triploid AAB and AAA, 1 AAAB tetraploid, and 1 accession of the subspecies *rhodochlamys*; G2, formed by 22 genotypes, including 1 AA diploid, triploids AAA (1), AAB (12), and ABB (4), and 4 AAAB tetraploids; and G3, formed by 31 accessions including 13 AA diploids, triploids AAA (11) and AAB (6), and 1 AAAB tetraploid (see Figure 2).

The values of genetic divergence among the accessions ranged from 0.37 ("Prata Anã Batico" and "Prata Anã 2") to 0.89 ("Royal" and "Akondro Mainty"), with an average of 0.70. We were unable to group the accessions exclusively according to their ploidy, subgroup/subspecies, or origin. However, some accessions grouped together because they have a high degree of relatedness (see Table 1).

In the first group (G1), the wild diploids of subspecies *banksii* grouped together ("Mambee Thu", "NBA 14", and "Sowmuk", and "Nyarma Yik"), which may suggest the exchange of alleles through natural mating because they all originate in New Guinea. These

results corroborate those of Jesus (2010), who have used simple-sequence repeat markers to define groups of genetic similarity among accessions.

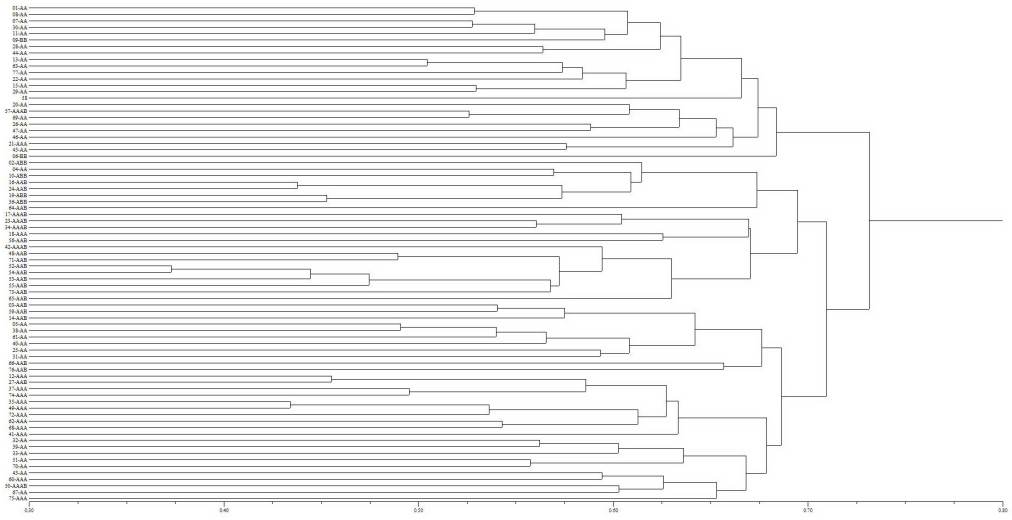

**Figure 2.** Dendrogram constructed by the Ward-MLM method using the genetic distances from 53 morphoagronomic descriptors from 77 banana accessions.

The plantains "Terrinha - G1" and "D'Angola - G2" were separated into different groups because they presented high genetic dissimilarity. These genotypes were evaluated under conditions in the State of Bahia for agronomic performance. The results revealed that they differ in a number of characters, including the NHB and NFB, in addition to being different types ("Terrinha" - Horn type and "D'Angola" - False Horn type) (Faria et al., 2010). Conversely, "D'Angola" and "Samura B" (False Horn type) were grouped together, agreeing with their type. The accessions "Wasolay", "Nam", "Towolee", and "Marcatoa" were grouped in G1. These accessions showed similar traits for PLH, PSD, diameter and weight of fruit, weight of the rachis, number of bunches, number of fruits, and weight of the bunch (Mattos et al., 2010).

The fruits of the subgroup Prata are characterized by mild aroma, sweetness, slight acidity, and digestibility. They are especially appreciated in northeast Brazil (Moreira, 1987). With the exception of "Pioneira" and "PV 03-76", all accessions of this subgroup ("FHIA 18", "FHIA 01", "Ouro da Mata", "Pacovan", "Prata Anã 2", "Prata Anã 3", "Prata Anã Batico", "Prata Anã Rene", "Prata Graúda", and "Walha") grouped in G2. Similar results were observed by Jesus (2010) during the genotyping of accessions from the subgroup Prata using simple-sequence repeat markers. Similarly, the female parent of the accession "Japira" ("Pacovan") was also grouped in G2. All the accessions with the ABB genome were classified in this group ("Cacambou Naine", "Marmelo", "Abu Perak", "Ice Cream"); "Ice Cream" is a synonymy of "Abu Perak" (Silva et al., 1999).

The wild diploids "Birmanie" and "Burmannica", which belong to subspecies *burmannica*, grouped in G1. Diploids with the BB genome ("Butuhan" and "Balbisiana France")

were also classified in this group, which occurred with "PV 03-76" ("Pacovan" x "Calcutta 4") and the improved diploid 028003-01, a hybrid between "Calcutta 4" and "Tuu Gia". The diploids "Khi Maeo" and "PA Phatthalung" grouped together, agreeing with results obtained by Jesus (2010), who estimated the genetic composition of these accessions using the mixture model. The subspecies complex *burmannica/burmannicoides/siamea* originated in northeast India, Burma, southeastern China, and Thailand and is considered genetically close to the subspecies *malaccensis* (Malay Peninsula), which explains the fact that they are grouped in G3 (Jesus, 2010).

According to Cury (1993), during the disposal of descriptors, some information may be lost. Considering the results obtained in this study, we can infer that these losses were minimal, given that the groups formed resulted from pre-established criteria, such as subgroup/subspecies or ploidy. The reduced number of descriptors capable of discriminating accessions of banana should reduce the time, labor, and cost of evaluating banana germplasm collections.

Wide genetic variability occurs in the agronomic characteristics of the 77 accessions of banana from the collection of Embrapa Cassava and Fruits. The cluster method using the Ward-MLM strategy appropriately classified and grouped the accessions of banana and elucidated their genetic relationships. The disposal of 42% of the descriptors caused no loss of information, and it can reduce costs and boost the management of banana germplasm collections.

## ACKNOWLEDGMENTS

We thank Conselho Nacional de Desenvolvimento Científico e Tecnológico for funding this project, and Fundação de Amparo à Pesquisa do Estado da Bahia for granting a scholarship to L.P. Brandão.

## REFERENCES

AGRITEMPO (2012). Sistema de Monitoramento Agrometeorológico. Available at [http://www.agritempo.gov.br/agroclima/sumario]. Accessed March 6, 2012.

Alves RM, Garcia AAF, Cruz ED and Figueira A (2003). Selecão de descritores botânico-agronômicos para caracterização de germoplasma de cupuaçuzeiro. *Pesq. Agropec. Bras.* 38: 807-818.

Amorim EP, Amorim VBO, Silva SO and Pillay M (2011). Quality Improvement of Cultivated Musa. In: Banana Breeding: Progress and Challenges (Pillay M and Tenkouano A, eds.). CRC Press, New York, 252-280.

Araújo DG, Carvalho SP and Alves RM (2002). Divergência genética entre clones de cupuaçuzeiro (*Theobroma grandiflorum Willd* ex *Spreng Schum*). *Cienc. Agrotec.* 26: 13-21.

Crossa J and Franco J (2004). Statistical methods for classifying genotypes. *Euphytica* 137: 19-37.

Cruz CD, Regazzi JA and Carneiro PCS (2004). Modelos Biométricos Aplicados ao Melhoramento Genético. Editora UFV, Viçosa.

Cury R (1993). Dinâmica Evolutiva e Caracterização de Germoplasma de Mandioca (*Manihot esculenta*, Crantz) na Agricultura Autóctone do Sul do Estado de São Paulo. Master's thesis, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

Daher RF (1993). Diversidade Morfológica e Isoenzimática em Capim-Elefante (*Pennisetum purpureum* Schum.). Master's thesis, Universidade Federal de Viçosa, Viçosa.

Daher RF, Moraes CF and Cruz CD (1997). Seleção de caracteres morfológicos discriminantes em capim-elefante (*Pennisetum purpureum* Schum.). *Rev. Bras. Zootec.* 26: 247-253.

Dias LAS, Kageyama PY and Castro GCT (1997). Divergência genética multivariada na preservação de germoplasma de cacau (*Theobroma cacao* L.). *Agrotrópica* 9: 29-40.

FAO (2012). Agriculture. Available at [http://faostat.fao.org/default.aspx]. Accessed March 2, 2012.

Faria HC, Donato SLR, Pereira MCT and Silva SO (2010). Avaliação fitotecnica de bananeiras tipo terra sob irrigação em

condições semi-áridas. *Ciênc. Agrotec.* 34: 830-836.
Franco J, Crossa J, Villasenõr J, Taba S, et al. (1998). Classifying genetic resources by categorical and continuous variables. *Crop Sci.* 38: 1688-1696.
Gower JC (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-874.
IPGRI (1996). Descriptors for Banana (*Musa* spp.). IPGRI, Rome. Available at [http://bananas.bioversityinternational.org/files/files/pdf/publications/descriptors_en.pdf]. Accessed March 2, 2012.
Jesus ON (2010). Caracterização Molecular de Acessos de Bananeira do Banco Ativo de Germoplasma da Embrapa. Doctoral thesis, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.
Jolliffe IT (1972). Discarding variables in a principal component analysis. I: artificial data. *J. Royal Stat. Soc. Series C - Appl. Stat.* 21: 160-173.
Jolliffe IT (1973). Discarding variables in a principal component analysis. II: Real data. *J. Royal Stat. Soc. Series C - Appl. Stat.* 22: 21-31.
Mattos LA, Amorim EP, Cohen KO, Amorim TB, et al. (2010). Agronomic, physical and chemical characterization of banana fruits. *Crop Breed. Appl. Biotecnol.* 10: 225-231.
Moreira RS (1987). Banana Teoria e Prática de Cultivo de Cultivo. Fundação Cargil, Campinas.
Oliveira EJ, Dias NLP and Dantas JLL (2012). Selection of morpho-agronomic Descriptors for characterization of papaya cultivars. *Euphytica* 185: 253-265.
Oliveira MSP, Ferreira DF and Santos JB (2006). Seleção de descritores para caracterização de germoplasma de açaizeiro para produção de frutos. *Pesq. Agropec. Bras.* 41: 1133-1140.
Pereira AV (1989). Utilização de Análise Multivariada na Caracterização de Germoplasma de Mandioca (*Manihot esculenta* Crantz). Doctoral thesis, Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.
Pereira AV, Vencovsky R and Cruz CD (1992). Selection of botanical and agronomical descriptors for the characterization of cassava (*Manihot esculenta* Crantz.) germplasm. *Rev. Bras. Genet.* 15: 115-124.
Renyi A (1961). On Measures of Entropy and Information. Fourth Berkeley Symposium, 547: 561. Available at [http://digitalassets.lib.berkeley.edu/math/ucb/text/math_s4_v1_article-27.pdf]. Accessed March 2, 2012.
Rohlf FJ (2000). NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Version 2.1. Exeter Software, New York.
Silva SO, Carvalho PCL, Shepard K, Alves EJ, et al. (1999). Catálogo de Germoplasma de Bananeira (*Musa* spp.). Embrapa Mandioca e Fruticultura, Cruz das Almas.
Silva SO, Flores JCO and Lima Neto FP (2002). Avaliação de cultivares e híbridos de bananeira em quatro ciclos de produção. *Pesq. Agropec. Bras.* 37: 1567-1574.
Strapasson E, Vencovsky R and Batista LAR (2000). Seleção de descritores na caracterização de germoplasma de *Paspalum* sp. por meio de componentes principais. *Rev. Bras. Zoot.* 29: 373-381.
Vieira EA, Fialho JF and Faleiro FG (2007). Variabilidade Genética do Banco Ativo de Germoplasma de Mandioca do Cerrado por Meio de Descritores Morfológicos. Embrapa Cerrados, Planaltina. Available at [www.cpac.embrapa.br/download/1229/r]. Accessed March 2, 2012.