



## GENETIC VARIABILITY IN INBRED GUAVA FAMILIES MEDIATED BY AGRONOMIC TRAITS

Willian Krause<sup>1</sup>, Natan Ramos Cavalcante<sup>2\*</sup>, Alexandre Pio Viana<sup>2</sup>,  
 Moises Ambrósio<sup>2</sup>, Eileen Azevedo Santos<sup>2</sup>, Henrique Duarte Vieira<sup>2</sup>

*1 Professor na Universidade do Estado de Mato Grosso (UNEMAT), Tangará da Serra - MT, Brazil;*

*2 Laboratório de Melhoramento Genético Vegetal, Universidade Estadual do Norte Fluminense Darcy Ribeiro (UNEF), Campos dos Goytacazes, RJ, Brazil.*

*\* Corresponding author: Natan Ramos Cavalcante ([natancavalcante2@hotmail.com](mailto:natancavalcante2@hotmail.com)).*

**Abstract:** The aim of this study was to evaluate the existing diversity among genotypes from inbred guava families obtained from a breeding program. To characterize the 61 genotypes, 51 descriptors were evaluated, including 15 quantitative, 16 qualitative, and 20 seed descriptors obtained digitally using the Graundeye<sup>®</sup> image analysis system. Based on the quantitative, qualitative, and seed descriptors, the Ward-MLM strategy was employed and was efficient in detecting divergence among guava genotypes using qualitative and quantitative descriptors. The evaluated population has genetic diversity and genotypes with important agronomic traits that allow the continuity of the breeding program with genetic gains. The descriptors that most contributed to divergence among the evaluated guava genotypes were leaf length/width, fruit length, fruit pulp thickness, leaf length, and fruit pulp weight. The seed descriptors obtained by digital image analysis were useful in estimating genetic diversity in guava.

**Keywords:** *Psidium guajava*, multivariate analysis, guava breeding.

### Introduction

Brazil has a total guava (*Psidium guajava* L.) production area of 22,128 ha that generates an average fruit yield of 26,402 kg ha<sup>-1</sup> (IBGE, 2019). A number of available cultivars have been used by farmers, among which the Paluma, Rica, Sassaoka, Pedro Sato, Cortibel, and Século XXI stand out (Santos et al., 2011). However, approximately 70% of the planted area in Brazil is occupied by cultivar Paluma (Pereira and Kavati, 2011) which demonstrates that this crop is vulnerable due to the little exploitation of genetic variability in the species. For this reason, there is

a need to develop new cultivars that meet the market and producer conditions in order to allow the diversification of guava cultivars in the field featuring different alleles, especially for different types of fruit and resistance to pests and diseases.

In the development of a breeding program, the breeder needs variability in the population not only at the beginning of the program but over its course he or she will also need to check if the population is not losing variability that may prevent future genetic gains and consequently the continuity of the program. Such knowledge allows the breeder to perform the introgression of favorable alleles for traits of interest at any time in

the breeding program when he or she realizes that genetic variability in the population is little or agronomic performance is low. Thus, the studied population needs to have genetic variability and agronomic traits of interest to both the producer and the consumer. This combination of genetic diversity and high mean values will provide the breeding program with productive cultivars.

Guava breeding programs based on the production of segregating populations are not very common, but they hold a great potential to generate new genetic variability for the diverse traits of importance for this species. Narrowed down the first procedures to initiate a guava breeding program, focusing on the production of segregating populations via the formation of full-sibling families, through an investigation of genetic diversity in twenty accession with the use of RAPD markers (Pessanha et al., 2011). Working in the same breeding program, investigated the concept of formation of heterotic groups through genetic diversity studies using artificial neural networks and obtained three groups for the evaluated population (Campos et al., 2016).

Qualitative and quantitative morphological descriptors can be employed to measure genetic variability. The Ward-MLM (Modified Location Model) method, proposed by Franco et al. (1998), has been adopted to use these pieces of information simultaneously in the quantification of the genetic variability of the guava population. This method also allows for the determination of the ideal number of groups for the evaluated population. The technique has already been used successfully by Campos et al. (2013) for the evaluation of genetic diversity in guava genotypes, in which 13 fruit-quality and physical attributes descriptors were utilized. The authors detected the formation of eight groups, with possible and reliable separation of divergent genotypes and genotypes with good agronomic yield.

Given the above considerations, the aim of this study was to evaluate the diversity existing among genotypes originating from inbred guava families based on quantitative, qualitative, and seed descriptors, using the Ward-MLM method.

## Material and methods

The experiment was implemented, in Itaocara, RJ, Brazil (geographical coordinates: 21°40' S and 42°04' W). The soil in the experimental area is classified as a Red-Yellow Argisol. The experiment was set as a randomized-block design with 18 inbred families, three replications, and 10 plants per plot. Families were obtained by self-crossing progenies resulting from controlled biparental crosses (Pessanha et al., 2011). Plants were planted in July 2014, with 4 m spacing between rows and 1.5 m between plants. Liming, fertilization at planting, and top-dressing were performed according to the recommendations of Costa and Costa (2003).

In this experiment, 61 guava genotypes (individual plants) were evaluated. For the characterization of the 61 genotypes, 51 descriptors were used, including 15 quantitative; 16 qualitative and also 20 seed descriptors according to the procedures described in Graundeye®. The evaluation period was March to April 2016.

Five young and fully developed leaves and five fruits per plant were sampled for the study of characterization and evaluation of genotypes.

The qualitative descriptors in the young leaves were the presence of anthocyanin and its coloring power. In fully developed leaves, we evaluated the shape, the curvature in the cross section and of the central nervure, ripe fruit color, and shapes of base and tip. In the fruits, we evaluated the final shape of the peduncle, skin color, surface texture, diameter of the calyx cavity relative to that of the fruit, pulp color, and outer pulp thickness in relation to the diameter of the core. We also evaluated branch habit and stem color in young shoots. Quantitative descriptors for fruit were number of fruits, fruit weight, fruit placenta weight, fruit pulp weight, fruit pulp yield, length, diameter, shape index, pulp thickness, and fruit placenta thickness. In the leaf, we evaluated the length, width, length/width ratio, and space between secondary nerves. We also evaluated the stem diameter at 10 cm above the soil, in each genotype.

Additionally, the seed descriptors referring to six color variables, seven geometry variables, and seven texture variables were evaluated. Seed descriptors were measured by digital imaging using the Graundeye® analysis system. Fifty seeds of each genotype were used for image reading. The Graundeye® analysis system provides a spreadsheet with the mean values of each genotype for the color, geometry, and texture variables.

The Ward-MLM (Modified Location Model) method, proposed by Franco et al. (1998), was employed as described by Viana and Resende (2014). First, the Gower (Gower, 1971) index was used to evaluate the quantitative traits, generating the estimate of the dissimilarity index, which ranges from 0 to 1. The dissimilarity was given by:

$$S_{ij} = \frac{\sum_{k=1}^p W_{ijk} \cdot S_{ijk}}{\sum_{k=1}^p W_{ijk}} \quad (Eq. 1)$$

Where: K is the number of variables ( $k = 1, 2, \dots$ );  $p$  = total number of traits;  $i$  and  $j$  = individuals to be compared with respect to trait  $k$ ;  $w_{ijk}$  was a weight used to define the contribution of the  $S_{ijk}$  individuals. In this aspect, when the value of variable  $k$  was absent in one or both individuals,  $w_{ijk} = 0$ ; otherwise,  $w_{ijk} = 1$ .  $S_{ijk}$  = contribution of variable  $k$  for total distance.

If the variable was qualitative,  $S_{ijk}$  was assigned the value 1, when there was positive or negative agreement for trait  $k$  between individuals  $i$  and  $j$ ; otherwise, when the variable was quantitative, the following case was considered:

$$S_{ij} = \frac{|Y_{ik} - Y_{jk}|}{R_k} \quad (Eq. 2)$$

where:  $Y_{ik}$  e  $Y_{jk}$  are the values of variable  $k$  for individuals  $i$  and  $j$ , respectively;  $R_k$  = amplitude of variation of variable  $k$ , assuming values from 0 to 1 or integers between these.

Afterward, the ideal number of groups was defined according to the pseudo F and pseudo  $T^2$  criteria, by Ward's clustering method (Ward, 1963). Based on the optimum number of groups, the hierarchical classification was obtained by Ward's method, which provides the necessary parameters to implement the final step of the MLM model (Crossa and Franco, 2004). The differences between groups, the correlation between variables, and the canonical variable (CV) were analyzed graphically. All analyses were performed using SAS statistical software (SAS, 2000). The diagrams were built using Sigma Plot software version 11.0.

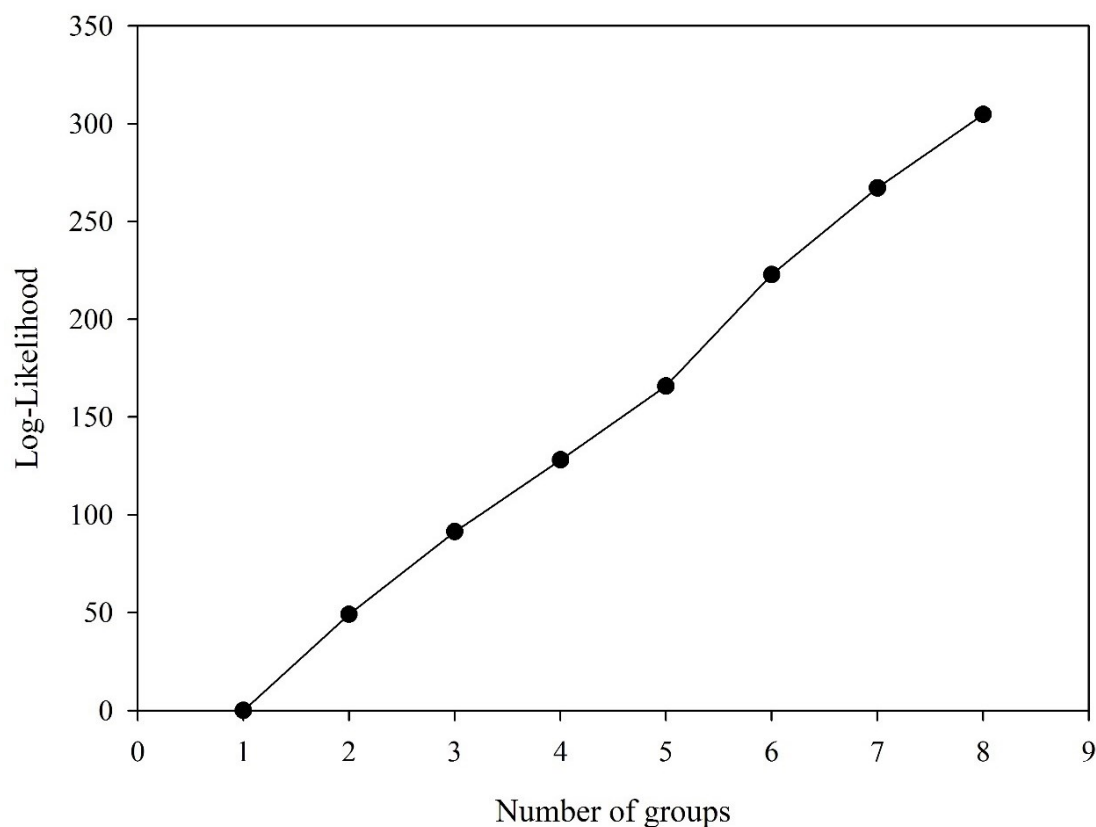
## Results

Three groups were formed according to the pseudo F and pseudo  $t^2$  criteria (Figure 1). The ideal number of groups was found where there was the highest increase in logarithmic function. The highest absolute value found was 42.18.

Groups I, II, and III were formed by 43, 8, and 10 genotypes, respectively. The genotypes were classified into the 49 phenotypic classes of the 16 descriptors evaluated (Table 1).

This demonstrates the wide genetic variability present in the population. According to the qualitative descriptors, 93.5% of the genotypes have pendant branches (Table 1). In the evaluation of young leaves, genotypes with presence of anthocyanin (83.6%) and with medium intensity (52.5%) prevailed.

All genotypes from groups I and II showed presence of anthocyanin in young leaves. Group III comprised only genotypes without anthocyanin. For fully developed leaves, most genotypes had the leaves with an obovate shape (57.4%), with rounded base (63.9%) and tip (83.6%). Additionally, 59% of the genotypes did not have a curved central nervure of the leaves; of those that had it, 65.6% were classified as having a strongly curved central nervure. The predominant color for the leaves was grey-green, as observed in 47.5% of the genotypes, followed by green (26.2%), green-yellow (23%), and dark green (3.3%).



**Figure 1.** Log-likelihood function showing the number of groups formed by the Ward-MLM strategy.

As regards the fruit descriptors, genotypes showed the following final peduncle forms: truncate (54.1%), rounded (26.2%), largely rounded (18.1%), and with neck (1.6%) (Table 1). For skin texture, 93.4% of the genotypes showed smooth texture, while 6.6% had a coarse skin. Fifty-nine percent of genotype had a pale yellow skin, 23% pale green-yellow, 14.7% dark red, and 3.3% orange. For the pulp color variable, genotypes were classified into five categories, ranging from cream to orangish pink, but with a predominance of dark pink, present in 37.7% of genotypes.

For the diameter of the calyx cavity relative to fruit diameter, 63.9% of the genotypes were classified as medium. With regard to the thickness of the outer pulp relative to the core diameter, genotypes were classified into medium (68.8%), thick (21.3%), thin (6.6%), and very thin (3.3%). In the evaluation of the quantitative descriptors, group I showed the largest number of fruits per plant (67.9), followed by groups II (59.3) and III (52.5) (Table 2). However, there was wide variation within the three groups, as observed by the standard deviations of 30.3, 16.6, and 19.7 for groups I, II, and III, respectively.

Variation within groups was also observed for fruit weight, for which group II showed the highest mean value (257.9 g).

As a consequence, group II also had the highest value for placenta weight (45.9 g) and pulp weight (212 g). Fruit pulp yield also had similar values between groups: 79%, 82.2%, and 81.2% for groups I, II, and III, respectively. Group II also showed the highest fruit length (86.5 mm) and fruit diameter (74.4 mm). But for the fruit shape index, values were close between the groups, which averaged 1.1. Groups had similar mean values to each other for the placenta thickness and fruit pulp thickness descriptors, which averaged 40.3 mm and 13.8 mm, respectively. For stem diameter 10 cm above the soil, group II averaged 59.6 mm, followed by group III, with 53.4 mm, and group I, with 50.3 mm. In the analysis of leaf-related descriptors, groups II and III had similar values for length, width, length/width ratio, and space between nerves. Group I had the lowest mean values for length (15 cm) and the highest for width (8.2), resulting in a lower length/width ratio (1.8), i.e., a more rounded leaf (Table 2).

**Table 1.** Phenotypic classes of plant, leaf, and fruit descriptors and number of inbred guava families per group formed by the Ward-MLM method.

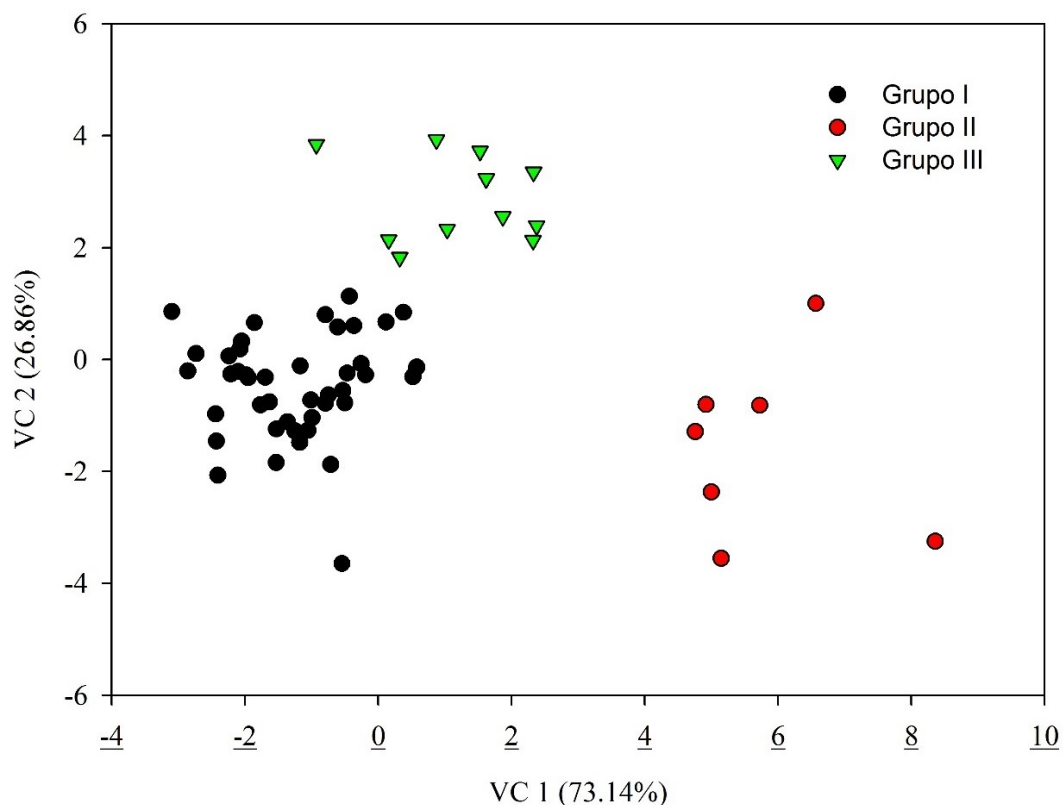
Descriptor	Classification	Group			Total (%)
		I (43)	II (8)	III (10)	
Tree: branch habit	Erect	2	-	2	6.5
	Pendant	41	8	8	93.5
Young shoot: stem color	Green	16	-	8	39.3
	Green yellow	17	7	2	42.6
	Dark red	10	1	-	18.1
Young leaf: anthocyanin	Absent	-	-	10	16.4
	Present	43	8	-	83.6
Young leaf: anthocyanin coloring power	Weak	21	8	-	47.5
	Medium	22	-	-	52.5
Fully developed leaf: shape	Ovate	5	-	-	8.2
	Obovate	28	-	7	57.4
	Obovo lanceolate	3	1	-	6.6
	Oblong	7	7	3	27.8
Fully developed leaf: curvature in the cross section	Medium	11	6	4	34.4
	Strong	32	2	6	65.6
Fully developed leaf: curvature of central nervure	Absent	23	8	5	59.0
	Present	20	-	5	41.0
Fully developed leaf: ripe fruit color	Green-yellow	7	5	2	23.0
	Grey-green	26	-	3	47.5
	Green	9	2	5	26.2
	Dark green	1	1	-	3.3
Fully developed leaf: base shape	Obtuse	6	-	1	11.5
	Rounded	26	7	6	63.9
	Cordate	11	1	3	24.6
Fully developed leaf: tip shape	Apiculate	1	-	-	1.6
	Obtuse	1	5	3	14.8
	Rounded	41	3	7	83.6
Fruit: final peduncle shape	Largely rounded	9	1	1	18.1
	Rounded	13	1	2	26.2
	Truncate	20	6	7	54.1
	With neck	1	-	-	1.6
Fruit: skin color	Pale green-yellow	12	1	1	23.0
	Pale yellow	24	6	6	59.0
	Dark yellow	5	1	3	14.7
	Orange	2	-	-	3.3
Fruit: surface texture	Smooth	41	7	9	93.4
	Coarse	2	1	1	6.6
Fruit: diameter of calyx cavity relative to fruit diameter	Small	3	3	6	19.7
	Medium	34	4	1	63.9
	Large	6	1	3	16.4
Fruit: pulp color	Cream	3	4	3	16.4
	Pale pink	11	2	1	23.0
	Pink	10	1	2	21.3
	Dark pink	18	1	4	37.7
	Orangish pink	1	-	-	1.6
Fruit: outer pulp thickness relative to core diameter	Very thin	2	-	-	3.3
	Thin	4	-	-	6.6
	Medium	33	2	7	68.8
	Thick	4	6	3	21.3

**Table 2.** Means and standard deviation of the variables for each of the three groups formed by the Ward-MLM method in 61 genotypes of inbred guava families.

Descriptor	Group			Overall mean
	I (43)	II (8)	III (10)	
Number of fruits (n)	67.9±30.3	59.3±16.6	52.5±19.7	64.1±27.7
Fruit weight (g)	177.2±42.8	257.9±55.2	175.2±46.3	186.1±50.7
Fruit placenta weight (g)	37.0±9.5	45.9±10.9	32.5±8.3	37.2±9.8
Fruit pulp weight (g)	140.2±35.7	212.0±45.2	142.7±39.4	148.9±43.0
Fruit pulp yield (%)	79.0±3.0	82.2±1.6	81.2±2.7	79.8±3.0
Fruit length (mm)	73.8±6.0	86.5±9.3	76.6±6.7	75.8±7.7
Fruit diameter (mm)	67.0±6.4	74.4±4.8	65.2±7.0	67.5±6.7
Fruit shape index	1.1±0.1	1.2±0.1	1.2±0.1	1.1±0.1
Fruit pulp thickness (mm)	13.1±1.8	16.7±2.7	14.5±2.1	13.8±2.3
Fruit placenta thickness (mm)	40.9±5.3	40.5±3.6	38.1±4.7	40.3±5.0
Stalk diameter 10 cm above the soil (mm)	50.3±6.3	59.6±4.0	53.4±3.6	52.0±6.4
Leaf length (cm)	15.0±0.8	16.2±0.7	16.4±0.5	15.4±1.0
Leaf width (cm)	8.2±0.6	7.7±0.4	7.9±0.9	8.1±0.7
Leaf length/width ratio	1.8±0.1	2.1±0.1	2.1±0.2	1.9±0.2
Space between leaf secondary nerves (cm)	1.1±0.1	1.2±0.2	1.3±0.1	1.1±0.2

In the analysis based on the canonical variables (CV), the first two CVs obtained by the Ward-MLM explained 100% of the total variation, with 73.2% for CV1 and 26.8% for CV2 (Figure 2). This high value indicates that a two-dimensional graph is adequate for visualizing the

relationship between groups and between the genotypes within groups. The shortest difference was detected between groups I and III (17.27), while the distance between groups I and II (50.79) was the highest (Table 3), as also shown in the previous graphic analysis (Figure 2).

**Figure 2.** Dispersion of the first two canonical variables (CV1 and CV2) representing the formation of three groups by the Ward-MLM strategy, based on 15 quantitative, 16 qualitative, and 20 seed descriptors in inbred guava families.

**Table 3.** Distance between groups formed by the Ward-MLM strategy based on 15 quantitative, 16 qualitative, and 20 seed descriptors in 61 genotypes of inbred guava families.

Group	II (8)	III (10)
I (43)	50.79	17.27
II (8)	-	40.45

The quantitative descriptors that most contributed to quantifying the genetic diversity existing between genotypes, i.e., those which showed the highest correlations with the first canonical variables, were leaf length/width (0.574), leaf length (0.656), fruit pulp thickness (0.554), leaf length (0.540), and fruit pulp pass (0.537) (Table 4).

**Table 4.** Canonical variables of 35 plant, fruit, and seed descriptors in 61 genotypes of inbred guava families.

Descriptor	Canonical variable (CV)	
	CV1	CV2
Number of fruits (un)	-0.159	-0.206
Fruit weight (g)	0.502	-0.276
Fruit placenta weight (g)	0.240	-0.356
Fruit pulp weight (g)	0.537	-0.244
Fruit pulp yield (%)	0.411	0.170
Fruit length (mm)	0.565	-0.102
Fruit diameter (mm)	0.320	-0.306
Fruit shape index	0.313	0.317
Fruit pulp thickness (mm)	0.554	0.022
Fruit placenta thickness (mm)	-0.078	-0.240
Stem diameter 10 cm above the soil (mm)	0.515	-0.012
Leaf length (cm)	0.540	0.460
Leaf width (cm)	-0.270	-0.060
Leaf length: width ratio	0.574	0.355
Space between leaf secondary nerves (cm)	0.226	0.389
Color: Below Otsu: CIELab: b	0.154	0.019
Color: Above Otsu: CIELab: b	0.116	0.233
Color: CIELab: a	0.153	0.120
Color: CIELab: Dispersion	-0.074	0.196
Color: CIELab: Dispersion of a	-0.108	0.185
Color: Medium: Green channel	-0.084	-0.176
Geometry: Shape complexity	-0.059	-0.220
Geometry: Compression	0.202	-0.238
Geometry: Convex distances	-0.190	-0.058
Geometry: Sphericity of the modified shape	-0.170	-0.280
Geometry: Modification ratio	0.245	-0.107
Geometry: Circumscribed rectangles: average circumference (cm)	-0.208	-0.209
Geometry: Elliptical variance	-0.152	0.294
Texture: Haralick - Medium	-0.042	-0.173
Texture: Laws - ES	-0.200	-0.249
Texture: Laws - LE	-0.002	-0.037
Texture: Laws - SR	-0.195	-0.238
Texture: Run Length - GLD	-0.070	-0.321
Texture: Run Length - GLNU	-0.134	-0.249
Texture: SFM - Measurement of modified frequency space	-0.300	0.025

Evaluating only the contribution among the 20 seed descriptors assessed, we observed that of the Five most important ones, three were related to the seed geometry (geometry: modification ratio; geometry: circumscribed rectangles: average circumference; and geometry: compression) and two to texture (texture: SFM - Measurement of modified frequency space; and texture: Laws – LE).

## Discussion

Guava breeding programs developed based on the generation of segregating populations and their subsequent evaluation are still rare, which is also because of the difficulty to conduct these strategies, coupled with the insufficient longevity of orchards for proper phenotypic assessments. On the basis of the previously discussed strategies, we chose for the production of inbred families as a means of fixation of potential alleles for plant- and fruit-related traits, as we observed that the inbreeding process seemingly does not severely affect the guava, as noted by aspects related to inbreeding depression. Based on this assumption, it is of great importance to evaluate this partial inbreeding on the many attributes of importance for the breeding of guava, since the new genetic variability can be added to the ongoing program.

The likelihood function analysis can define more accurate criteria in the formation of divergent groups, resulting in the determination of less subjective groups (Barbé et al., 2010). Campos Quantified the divergence among 138 guava genotypes obtained from controlled biparental crosses and found eight groups with a 67.51 increase (Campos et al., 2013). Their result was higher than the three groups formed in our study (Figure 1).

However, the number of groups can vary according to species, number of accessions, and number and type of descriptors (Gonçalves et al., 2009). Characterized 70 accessions of 11 species of the genus *Passiflora*: *P. edulis*, *P. mucronata*, *P. setacea*, *P. pentagona*, *P. caerulea*, *P. gibertii*, *P. cincinnata*, *P. suberosa*, *P. micropetala*, *P. alata*, and *P. coccinea*, obtained from the germplasm bank of Embrapa Cassava and

Tropical Fruits and from the *Passiflora* collection of State University of Northern Rio de Janeiro (UENF) (Paiva et al., 2014). These authors observed the formation of five groups, with a logarithmic function increase of 97.94. The number of genotypes used in the present study impacted the formation of groups, since not all plants from the obtained families were also evaluated for the many attributes measured.

The use of morphological descriptors allows the breeder to know genotypes and estimate the genetic variability, which is an essential step in plant breeding programs. In this regard, in the evaluation of qualitative descriptors, the pendant branch habit, which occurred in the majority of genotypes of this population (Table 1), is beneficial in that it provides a more open canopy where branches will have a lateral growth, reducing the plant height. This is important for the application of pesticides, and especially pruning and fruit handling. In this way, we can predict that the evaluated population will provide new genotypes with the potential for use in the program under development.

The fruit descriptors appear to be important in the choice of genotypes aiming at fresh consumption because they are related to factors that depreciate the product in case they are not attractive to consumers. For the domestic consumption, there is a preference for the guava with a dark-pink pulp, whereas the external market prefers the white pulp. In this evaluated population, the most part of genotypes had a dark-pink pulp (37.7%), demonstrating the potential of the population.

The diameter of the calyx cavity relative to the fruit diameter and the outer pulp thickness relative to the core diameter are related to pulp yield. In this way, the lower the diameter of the calyx cavity in relation to that of the fruit and the thicker the outer pulp in relation to the core diameter, the better. The classification found for the majority of genotypes was 'medium'. However, the population has sufficient variability to permit the selection of genotypes with a low cavity diameter and a thick pulp. But we also observed that groups II and III showed

pulp yields of 82.2% and 81.2%, respectively (Table 2). Fruit yield values higher than 80% are acceptable for the selection of guava fruits (Campos et al., 2013). These results suggest that inbreeding depression seems not to affect the guava breeding process; genotypes can be obtained through this process, and traits with high means and genetic variability allow for the continuity of the evaluation and selection of plants.

The average number of fruits per group was low (Table 2). However, the average weight of the fruits was 155.2 g (group III) to 257.9 g (group II), the latter being considered large fruit. These high values for fruit weight are a reflection of the first production, when plants had not yet expressed their full production potential and thus bore larger fruits. In general, the fruit weight tends to decrease as commercial production is stabilized, when the plant typically produces more fruits. The average fruit weight is an important trait, since heavier fruits are also larger and thus more attractive to the consumer. The selection criteria in breeding programs are fruit with an average weight of 100 to 300 g, with few seeds, and a diameter of 7 cm or more. In this population, group II met these criteria, weighing 257.9 g and with a diameter of 7.4 cm (Table 2) (Teixeira et al., 2012; Campos et al., 2013).

The observation of fruit length and diameter as separate parameters has little importance for the characterization of fruit from the guava genotypes. However, the ratio between these variables is highly representative, because it indicates the fruit shape, and the closer it is to 1, the rounder the fruit is. Pyriform or ovate fruits (shape index greater than 1) can be destined for fresh consumption, and those with a rounded shape (shape index close to 1) are more suitable for industrialization (Fachi et al., 2018).

Leaf length and leaf width are important traits in that they are correlated with the plant leaf area. The leaf area has a direct relationship with the use of solar energy, which is transformed into chemical energy during photosynthesis. Thus, the leaf area has a correlation with the photosynthetic and transpiration activities, reflecting the plant's ability to intercept

radiations and perform gas exchanges; hence its use as indicative of agricultural crop yields (Francisco et al., 2014). Group III showed the largest leaf size (length  $\times$  width) (Table 2).

In the analysis of canonical variables in this study, the first two variables explained 100% of the total variation (Figure 2). If the first two canonical variables provide estimates greater than 80%, a satisfactory interpretation of variability among accessions can be achieved, as occurred in this study, which in turn allowed the two-dimensional graphic representation (Figure 2). In a study conducted by Campos et al. (2013), who evaluated 138 guava accessions, the authors found that the first two variables were responsible for only 61.79%, with CV1 responsible for 39.12% and CV2 for 22.68%.

Therefore, a two-dimensional graph would not capitalize a good proportion of the total variability, thereby requiring the introduction of the third variable, which was responsible for 19.50%. Hence, the three-dimensional representation was the most suitable for the representation of the data set, and the sum of the three variables accounted for 81.30% of the total variation. Used the same statistical method to estimate the genetic variability among 25 coffee genotypes and had 100% of the total variation also explained by the first two canonical variables (Rodrigues et al., 2016). The graphic analysis of CV also showed greater distancing between groups I and II and greater approximation between groups I and III, as shown in Table 3. These results agree with the obtained distances proposed by Franco et al. (1998).

In a breeding program, the use of parents with high genetic diversity is recommended for the maintenance of superior segregating individuals in subsequent generations. However, it is also important to choose an individual with superior traits for plant, flower, and fruit, and, lastly, the analysis of its behavior for a later recommendation of use *per se*. Thus, the cross between divergent genotypes and desirable agronomic traits should be performed. In this regard, because genotypes II and II displayed important traits for the genetic breeding of guava, crosses between genotypes from group II and

genotypes from group III are recommended so as to obtain superior individuals.

With respect to the importance of the descriptors for genetic divergence, in this study, we observed that the most important ones were leaf length/width ratio, fruit length, fruit pulp thickness, leaf length, and fruit pulp weight (Table 4). Found that the highest correlations of variables with the first canonical variable were pulp yield and TSS/TA ratio (Campos et al., 2013). Although the seed descriptors are not considered the most important in contribution to genetic diversity, they were useful in estimating divergence, because after their discard there was a change in grouping (Krause et al., 2017).

## Conclusion

The Ward-MLM strategy was efficient in detecting divergence among the guava genotypes using qualitative and quantitative descriptors. The evaluated population has genetic diversity and genotypes with important agronomic traits, allowing the continuity of the breeding program with genetic gains. Leaf length/width, fruit length, fruit pulp thickness, leaf length, and fruit pulp weight were the descriptors that most contributed to divergence among the evaluated guava genotypes. Seed descriptors obtained by digital image analysis were useful in the estimate of genetic diversity of guava.

## References

- BARBÉ, T.C.; AMARAL JÚNIOR, A.T.; GONÇALVES, L.S.A.; RODRIGUES, R.; SCAPIM, C.A. 2010. Association between advanced generations and genealogy in inbred lines of snap bean by the Ward-Modified Location Model. **Euphytica**. 173(3): 337–343.
- CAMPOS, B.M.; VIANA, A.P.; QUINTAL, S.S.R.; GONÇALVES, L.S.A.; PESSANHA, G.O.P. 2013. Quantification of the genetic divergence among guava accessions using Ward-MLM strategy. **Revista Brasileira de Fruticultura**, 35(2): 571-578.
- CAMPOS, B.M.; VIANA, A.P.; QUINTAL, S.S.R.; BARBOSA, C.D.; DAHER, R.F. 2016. Heterotic group formation in *Psidium guajava* L. by artificial neural network and discriminant analysis. **Revista Brasileira de Fruticultura**, 38 (1):151-157.
- COSTA, A.F.S.; COSTA, A.N. 2003. Tecnologias para produção de goiaba. Vitória, ES: Incaper, 341p.
- Crossa, J; Franco, J 2004. Statistical methods for classifying genotypes. **Euphytica**, 137(1): 19-37.
- FACHI, L.R.; GARBUGIO, E.L.S.; FERREIRA, A.F.N.; MACHADO, R.F.C.; KRAUSE, W. 2018. Qualidade e correlação dos parâmetros físicos e químicos dos frutos de cultivares de goiaba. **Scientific Electronic Archives**, 11(4): 36-40.
- FRANCISCO, J.P.; DIOTTO, A.V.; FOLEGATTI, M.V.; DA SILVA, L.D.B.; PIEDADE, S.M.D.S. 2014. Estimativa da área foliar do abacaxizeiro cv. Vitória por meio de relações alométricas. **Revista Brasileira de Fruticultura**, 36(2): 285-293.
- FRANCO, J.; CROSSA, J.; VILLASEÑOR, J.; TABA, S.; EBERHART, S.A. 1998. Classifying genetic resources by categorical and continuous variables. **Crop Science**, 38(6): 1688-1696.
- GONÇALVES, L.S.A.; RODRIGUES, R.; AMARAL JÚNIOR, A.T.; KARASAWA, M.; SUDRÉ, C.P. 2009. Heirloom tomato gene bank: Assessing genetic divergence based on morphological, agronomic and molecular data using a Ward-modified location model. **Genetics and molecular research**, 8(1): 364-74.

- GOWER, J.C. 1971. A general coefficient of similarity and some of its properties. **Biometrics**, Washington, 27(4): 857- 874.
- IBGE - Instituto Brasileiro De Geografia E Estatística. 2021. **Produção agrícola municipal**. Available at: <https://sidra.ibge.gov.br/tabela/1613#resultado>. Accessed on may 27, 2021.
- KRAUSE, W.; VIANA, A.P.; CAVALCANTE, N.R.; AMBRÓSIO, M.; SANTOS, E.A.; VIEIRA, H.D. 2017. Digital phenotyping for quantification of genetic diversity in inbred guava (*Psidium guajava*) families. **Genetics and Molecular Research**, 16(1): 1-11.
- PAIVA, C.L.; VIANA, A.P.; SANTOS, E.A.; SILVA, R.N.O.; OLIVEIRA, E.J. 2014. Diversidade genética de espécies do gênero *Passiflora* com o uso da estratégia Ward-MLM. **Revista Brasileira de Fruticultura**, 36(2): 381-390.
- PEREIRA, F.M.; KAVATI, R. 2011. Contribution of Brazilian scientific research in developing some of subtropical fruit. **Revista Brasileira de Fruticultura**, 33(1): 92-108.
- PESSANHA, P.G.O.; VIANA, A.P.; JUNIOR, A.T.A.; SOUZA, R.M.; TEIXEIRA, M.C.; PEREIRA, M.G. 2011. Avaliação da diversidade genética em acessos de *Psidium* spp. via marcadores RAPD. **Revista Brasileira de Fruticultura**, 33(1):129–136.
- RODRIGUES, W.P.; VIEIRA, H.D.; TEODORO, P.E.; PARTELLI, F.L.; BARBOSA, D.H. 2016. Assessment of genetic divergence among coffee genotypes by Ward-MLM procedure in association with mixed models. **Genetics and Molecular Research**, 15(2): 1-7.
- SANTOS, C.A.F.; CORRÊA, L.C.; COSTA, S.R. 2011. Genetic divergence among *Psidium* accessions based on biochemical and agronomic variables. **Crop Breeding and Applied Biotechnology**, 11(2): 149-156.
- SAS INSTITUTE: Statistical Analysis System: **User's Guide**. 2000, SAS, Cary.
- TEIXEIRA, D.H.L.; OLIVEIRA, M.D.S.P.; GONÇALVES, F.M.A.; NUNES, J.A.R. 2012. Índices de seleção no aprimoramento simultâneo dos componentes da produção de frutos em açaizeiro. **Pesquisa Agropecuária Brasileira**. 47(2): 237-243.
- VIANA, A.P.; RESENDE, M.D.V. 2014. **Genética quantitativa do melhoramento de fruteiras**. Rio de Janeiro, RJ: Interciência. 282 p.
- WARD, J.H. 1963. Hierarchical grouping to optimize an objective function. **Journal of the American Statistical Association**, 58(301): 236-244.