FUNCTIONAL PLANT BREEDING JOURNAL
Scientific Journal. ISSN 2595-9433
Volume 3, Number 2, Article n. 3, July/December
D.O.I. http://dx.doi.org/10.35418/2526-4117/v3n2a3
Received: 03/20/2021 - Accepted: 05/17/2021

DIGITAL PHENOTYPING IN INBRED GUAVA LINES: SEED CHARACTERIZATION

Carla Caroline Amaral da Silva[1], Henrique Duarte Vieira[1], Alexandre Pio Viana[1], Eileen Azevedo Santos[1], Mariana Quintas Maitan[1]

1 Universidade Estadual do Norte Fluminense Darcy Ribeiro, Av: Alberto Lamego – 2000, Campos dos Goytacazes-RJ – Brazil.
** Corresponding author: Eileen Azevedo Santos (eileenbiologa@gmail.com).*

Abstract: Guava, *Psidium guajava* L., is a species of economic importance for several countries. The production of new varieties of guava from seeds with good germination and vigor potential has been investigated. In this scenario, the present study characterized and estimated the genetic diversity of seeds of 42 genotypes of guava from an S_2 family based on germination response, vigor tests, and digital phenotyping analysis. The experiment was laid out in a randomized-block design with four replicates. Descriptive statistics were carried out for each of the analyzed variables and multivariate analyses were used to estimate genetic diversity based on Gower's Distance. Digital analysis was performed using the GroundEye S120 system, which extracted color, geometry, texture, and histogram data. In total, three color, two geometry, five texture, four physiological analysis, and 17 histogram variables were used. The physiological variables and those obtained by digital phenotyping were efficient to discriminate the genotypes, indicating that there is genetic variability to be exploited within the guava breeding program. The traits that most contributed to genetic diversity were those related to the histogram. The genotypes belonging to groups I and III had the highest means for germination percentage, thousand-seed weight, and shoot length. Based on the results, two options are viable for conducting the guava breeding program. The most vigorous individuals can be self-pollinated for the development of lines and the superior and most divergent ones from groups I and III can be crossed to exploit heterosis. Thus, crosses between the most vigorous individuals from groups I and III are recommended, as they exhibited higher means for the traits of germination percentage, thousand-seed weight, and shoot length.

Keywords: *Psidium guajava*, genetic diversity, digital image analysis, seed physiological quality.

Introduction

Guava, *Psidium guajava* L., is a species of economic importance for several countries. According to the United Nations Food and Agriculture Organization (FAO, 2019), Brazil is the seventh largest guava producer in the world. In 2018, the country produced 578,608 t of guava fruit. The northeast was the region with the highest production, 293,599 t, followed by the southeast, with 234,817 t (IBGE, 2019). Many products can be generated from it, such as juices and sweets. The guava fruit has high nutritional quality, containing high levels of ascorbic acid, calcium, fibers, lycopene, and vitamins A, B2, B6, C, and E (Lima et al., 2002).

Commercial production of guava seedlings is achieved via cuttings, a method that provides uniform production and increased yields (Hartmann et al., 2002). Commercial seedling production from seeds is not yet exploited due to its variable results in the orchard. Nonetheless, it is an important step in breeding programs as a way to exploit variability within species (Bastos and Ribeiro, 2011).

Although *P. guajava* L. is an allogamous plant, there are reports of self-pollination in the species (Alves and Freitas, 2007). The reduction in types of guava cultivars is worrying, as it can lead to genetic vulnerability of the crop. This situation can be observed in Brazil, where approximately 70% of the guava trees currently grown for industrial processing are cultivar Paluma, due to its efficient rooting capacity (Pereira and Kavati, 2011).

Self-pollination leads to increased homo-zygo-sisty, thereby reducing heterozygosity in the offspring. This can represent an alternative means of obtaining homogeneous fruits in commercial orchards.

The State University of Northern Rio de Janeiro (UENF) has been developing a breeding program with the species *P. guajava* L. The works have yielded promising results, according to Pessanha et al. (2011), Campos et al. (2013), Oliveira et al. (2013), Campos et al. (2016), Quintal et al. (2017), Maitan et al. (2020), Silva et al. (2020), Souza et al. (2020), Ambrosio et al. (2021), and Silva et al. (2021). These good results point to the possibility of obtaining inbred guava families through self-pollination, which will allow the development of varieties (pure lines) as well as the production of hybrids through the exploitation of heterosis. In this respect, studies of seed diversity, morphology, and physiological quality are relevant for the production of genotypes obtained in the breeding program.

Obtaining superior individual sent ails determining the physiological quality of seeds, which can be evaluated and characterized by means of germination, vigor, and stress-resistance tests that will reveal the most or least vigorous genotypes. Another approach is digital seed image analysis, which has been employed for the identification of cultivars, determination of colors, texture, mechanical damage, and classification by size (Venora et al., 2007; Medina et al., 2010; Kara et al., 2013; Pinto et al., 2015; Andrade et al., 2016). The GroundEye® system, developed by the Tbit company, is the only instrument on the national market built specifically for visual analysis of seeds.

Therefore, this study aims to characterize and estimate genetic diversity in 42 genotypes of S_2 inbred guava based on physiological attributes and variables obtained by digital seed phenotyping.

Material and methods

Experiment site

The experiments of germination, vigor, and image analysis were developed in the laboratory of the Seed Production and Technology Section at the State University of Northern Rio de Janeiro (CCTA-UENF), in Campos dos Goytacazes, RJ, Brazil.

Evaluated genotypes

The guava seeds of the S_2 family originated from the 42 most productive genotypes according to the results of previous work by the team at the Plant Breeding Laboratory (CCTA-UENF) (Ambrósio et al., 2021).

The guava fruits were obtained by self-pollination of the S_1 families. The crop is located in the municipality of Itaocara, RJ, Brazil, at the experimental unit of Ilha Barra do Pomba (21º40'S, 42º04'W, 76 m altitude).

Self-pollination was achieved by protecting the flowers by covering, before anthesis. The buds were identified and the fruits were later protected with a Raschel mesh bag.

Once harvested, the fruits were stored in a cold chamber until the moment of seed removal. Auxiliary cutting and pulp-removal tools were used to pulp the fruits and sieves were used to remove all mucilage and fiber. The seeds were rubbed over the steel-mesh sieve under running water.

The removed seeds were left to dry for 48 h at room temperature, on paper towels, in containers. These were turned over 24 h after the start of the drying process for homogeneous drying.

After this process, the seeds were stored in bags identified with the specific genotype.

Fruit collections and genotype assessments were performed in a randomized-block experimental design.

Evaluated traits

Seed physiological quality

Several seed physiological quality traits were evaluated for the different obtained families, namely:

Moisture: determined by oven-drying for 24 h at 105 ± 3 °C (Brasil, 2009). **Thousand-seed weight (TSW)**: obtained according to the Rules for Seed Testing (RAS) (Brasil, 2009). **Germination test**: four replicates of 50 seeds from each of the 42 genotypes were used. The test was set up on a paper roll. The germination chambers were regulated to an alternating temperature of 35-25 °C, with a photoperiod of 8 h of light and 16 h of dark, respectively. The first evaluation took place on the 10th day, and the last one on the 35th day, in which the percentages of normal seedlings, abnormal seedlings, and non-germinated seeds were recorded. **Shoot length (SL) and radicle length**: four rolls of paper were prepared with ten seeds and, at the end of the 35th day, the seedlings were measured with a graduated ruler. **Dry matter weight**: after obtaining the length of the seedlings, they were sectioned to separate the shoots from the radicles and placed in paper bags, which were then oven-dried at 65 °C for 72 h. Once dried, the samples were weighed on an analytical scale to determine the dry matter of shoots (g) and root (g). **Germination speed index**: seeds that produced shoots 1.0-cm long, according to the preliminary tests, were counted on alternate days. This variable was calculated by the formula proposed by Maguire (1962). **Tetrazolium test**: for the seeds that did not germinate in the germination test, the tetrazolium test was performed using a 0.1% solution. The seeds were cut lengthwise and kept 4 h in the dark, immersed in the tetrazolium solution, at a temperature of 30 °C (Masetto et al., 2009). Dead seeds were considered to be those whose color did not change or which deteriorated. **Accelerated aging test:** the seeds were placed uniformly on an aluminum screen inside a germination box with 40 mL of water at the bottom. Subsequently, the germination boxes were subjected to a temperature of 41ºC for 48 h. After this procedure, the test was performed to assess germination as described previously.

Digital seed phenotyping

Each genotype was analyzed in four replicates, for all variables. GroundEye® mini/SAS mini was the equipment used to capture and analyze the seeds (Figure 1).

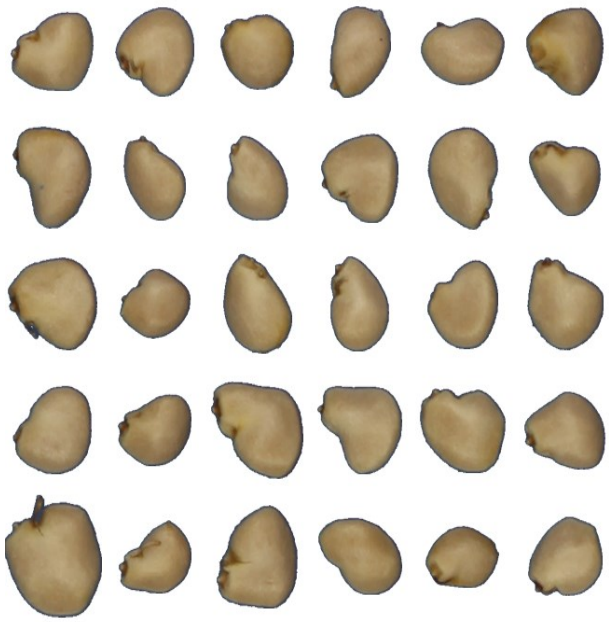

Figure 1. Figure obtained from GroundEye® software with the images of the seeds without the blue background to determine the values of the traits.

The system has a capture module and a software program for analysis. The capture module contains an acrylic tray where 50 seeds were placed for image capture by the high-resolution camera. The software generated data spreadsheets from the digital analysis of the seed images. All traits evaluated by the GroundEye® mini equipment are listed in the user manual. Of these, 51 were related to color, 48 to geometry, 192 to histogram, and 43 to texture.

Description of digital variables

An initial screening was performed to remove redundant and invariant traits. Three colors, two geometries, five textures, 17 histograms, and four physiological quality traits were considered for analysis (Table 1).

Among the traits analyzed by digital image analysis, color was evaluated by three variables. C1measures the points above the Otsu threshold in the color space b, which ranges from yellow to blue. The C2 variable uses the RGB color space to calculate the points above the threshold proposed by Otsu (Otsu thresholding), a technique whereby the image is divided into two distinct classes by calculations of threshold L as the intensity level that maximizes a criterion function η (L), defined as the ratio between the variance between classes and the global variance, i.e. $\eta (L) = \sigma^2 B/\sigma^2 T$, where $\sigma^2 B$ are the variances between classes and σ 2 T are the total variances (Persechino and Albuquerque, 2015). The C3 variable was determined based on the values of the predominant color, obtained by the K-Means method considering the predominant color as the centroid of the largest group found.

Geometry was evaluated by the G1 and G2 variables. G1, which detected the corners termed SUSAN (Smallest Univalue Segment Assimilating Nucleus), is an algorithm that assumes that within a small circular mask, the intensity of brightness belonging to different objects varies little. It also calculates the number of pixels whose brightness is similar to that of the central pixel, or nucleus, of the mask. According to Bay et al. (2006), the SURF algorithm is a detector and descriptor of key points invariant to rotation and scale that is computed much faster. The SURF descriptor detector is based on the Hessian matrix.

Table 1. Codes of the variables used to determine genetic variability in guava seeds by the Ward-MLM method.

Codes	Variables
C1	Below Otsu: CIELab: b [1]
C2	Above do Otsu: Canal azul [1]
C3	Predominant: Blue channel [1]
G1	Number of corners per SUSAN [2]
G2	SURF [2]
T1	Laws: ER [3]
T2	Laws: ES [3]
T11	Run Length: GLD [3]
T12	Run Length: GLNU [3]
T14	Haralick: Variance [3]
H1	Blue: Variance [4]
H2	HSL: Luminance: Variance [4]
H3	HSL: Hue: Minimum index [4]
H4	HSL: Saturation: Variance [4]
H5	LBP: Maximum index [4]
TSW	Thousand-seedweight
H6	LBP: Minimum index [4]
H7	LPQ: Maximum index [4]
H8	LPQ: Minimum index [4]
H9	Luminância: Variance [4]
H10	NDLPQ: Maximum index [4]
H11	NDLPQ: Minimum index [4]
H12	Green: Variance [4]
H13	Red: Variance [4]
H14	YCbCr: Brightness: Variance [4]
H15	YCbCr: Blue intensity: Variance [4]
H16	YCbCr: Intensity of red: Variance [4]
H17	YIQ: Component I: Variance [4]
GPAA	Germination % in the accelerated aging test Envelhecimento Acelerado
GP	Germination %
SL	Shootlength

1. Color; 2. Geometry; 3. Texture; 4. Histogram

The texture of the guava seed was measured by the T1, T2, T11, T12, and T14 variables. The T1 trait extracted edge and ripple data (Figure 2). This technique can evaluate different texture types (level, edge, spot, and ripple) and assessments can be simultaneous. The analyzed images undergo a pre-processing step that reduces the effects of lighting (Setiawan et al., 2015). The T2 variable (texture: Laws: ES [Edge/Spot]) evaluated edges and spots.

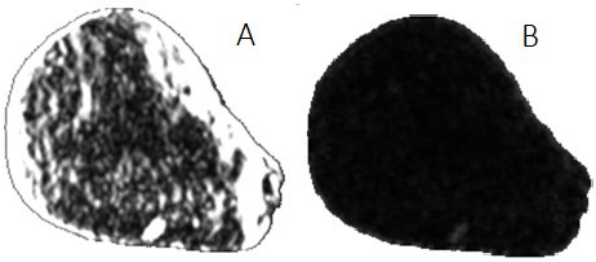

Figure 2. Images of the guava seed generated by the GroundEye® instrument for the texture variables. A - Variable T1 (Laws ER). B - Variable T2 (Laws ES).

T11 and T12 were calculated using the Run Length method, which is a simple form of compression without data loss in which long sequences of repeated values are stored as a single value. This variable measures the distribution of runs over the gray values. In this way, the similarity of the gray level intensity values in the image is measured, where a lower GLN (gray level non-uniformity) value correlates with greater similarity in the intensity value (Galloway, 1975). Finally, the T14 variable measures the homogeneity of the co-occurrence matrix. To describe the textures, Haralick et al. (1973) proposed 14 statistical measures, calculated from the co-occurrence matrix. This method employs a methodology for describing textures based on second-order statistics, defined as traits derived from the calculation of matrices coined co-occurrence matrices. These matrices consist of a tabulation of how many different combinations of gray levels occur in an image in a given direction.

The histogram trait was represented by the H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16and H17variables. H1 indicates the variance of the histogram in the blue channel. The variance provides a measure of the dispersion of the data around the mean. The H2 variable uses the HSL color system to evaluate hue, saturation, and luminosity. Hue determines the type of color, covering all the colors of the spectrum, from red to violet. Saturation evaluates purity, with lower values translating into a grayer image. Luminosity is an attribute of color regardless of its degree of purity, ranging from pure black to white. H3measures the minimum index values, i.e., the index of minimum occurrence of this luminosity in the histogram. Hue values are measures of the average wavelength of the light reflected or emitted, defining the color of the object. H4represents the values pertaining to the variance of the saturation histograms.

The H5 (maximum index) and H6 (minimum index) variables used LBP (Local Binary Pattern), a statistical coding algorithm for surface analysis plotted in the form of a histogram. To generate the histogram, each pixel of the image and its eight neighbors are mapped, resulting in the extraction of several statistics (Ojala et al., 1996). H7 (maximum index) and H8 (minimum index) used the LPQ (Local Phase Quantization) method, which is based on the quantification of the discrete Fourier transform phase. The LPQ code is computed locally, where each point or code is calculated using a neighborhood. From the codes of each point, the LPQ distribution histogram is constructed where texture traits are extracted (Ojansivu and Heikkila, 2008).

H9 was calculated based on the brightness measures of an image, whose extracted values can be used to produce a histogram that can be described by the following equation:

$$L = \sum_{i=0}^{I} . \sum_{j=0}^{J} l(i,j),$$

where i and j represent the number of horizontal and vertical pixels of the image, respectively, and l(i,j) is the relative luminance equation.

The histogram H10 (maximum index) and H11 (minimum index) variables were quantified by the uncorrelated local phase, a variation of the LPQ calculation that was created by the Tbit company, which developed the GroundEye® mini system. NDLPQ does not have a linear correlation, so it does not make a direct correlation between row and column, reducing cross-correlation within a set of signals.

H12 and H13 were determined by variance. H14 (brightness: variance), H15 (blue intensity: variance) and H16 (red intensity: variance) used the YCbCr color system, which is widely employed in digital video and photography systems. Y represents luminance information. Cb and Cr describe the color, where Cb evaluates the blue difference components and Cr the red difference components. YCbCr is not an absolute color space system and cannot be used to encode RGB information. The H17 variable used the

YIQ system, which is very similar to YCbCr. The Y parameter of YIQ is identical to the Y parameter of YCbCr, which is luminance. Parameter i corresponds to a range of color variations from cyan (minimum) to orange (maximum). For the Q parameter, this variation ranges from green to magenta (Lopes, 2013).

The GA variable can be defined as the surface space in two-dimensional analysis, e.g. the amount of space on the surface of an object (Figure 3). It is calculated by the following formula:

$$A = \sum_{p \epsilon R} 1,$$

where p represents a pixel of the image and R the pixels of the object whose area will be measured.

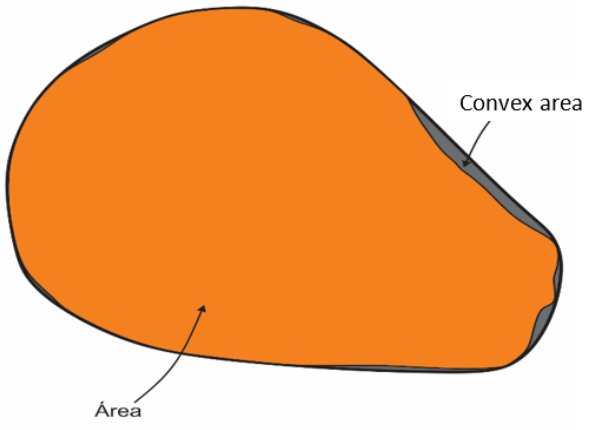

Figure 3. Image provided by the GroundEye® equipment and adapted in CorelDRAW 2018. Measures of length analyzed in guava seeds.

Convex area corresponds to the amount of space (in cm^2) covered by the surface of the convex envelope (or convex closure) of a seed. It is calculated using the following formula:

$$A = \sum_{p \epsilon R'} 1,$$

where p represents a pixel and R' the convex area of the object. Then, a conversion to real size is made from the scale used in the analysis.

Circularity is a circular shape factor that is more sensitive to the elongation of the object and less dependent on the smoothness of the contour. It is assigned 1 for circular objects and less than 1 for objects with other shapes since any other shape having the same maximum diameter has a smaller area. Circularity is calculated using the following formula:

$$C = \frac{4.A}{\pi.MD^2},$$

where A represents the area and MD the maximum diameter.

The diameter of a circumference or circle is any straight line that passes through the center of these figures. Thus, the diameter will be the largest secant line passing through any circumference.

Perimeter is the measure of the contour of a two-dimensional object, that is, the sum of all sides of a geometric figure. Convex perimeter is the measure of the contour of the convex closure of an object.

Maximum diameter is the longest line that passes through the centroid of the seed, in centimeters; and minimum diameter is the shortest line that passes through the centroid.

Sphericity of the shape defines how circular the object is. The closer its value is to 12.56, the closer the object is the shape of a circle. It is calculated using the following formula:

$$C = \frac{\sqrt{(\frac{4}{\pi})A}}{MD},$$

where A represents the area and MD the maximum diameter.

Solidity of the contour defines the object as concave or convex regarding its contour. It is assigned 1 for convex objects and decreases with the presence of concavities. Solidity is more sensitive to the presence of thin and long branches. It is calculated by the following formula:

$$S = \frac{A}{CA},$$

where A represents the area and CA the convex area.

Statistical analysis

Descriptive statistics was performed for each of the variables used. The 31 variables were evaluated based on principal component analysis and the average of the measures taken for each trait, from the correlation matrix, using Genes software (Cruz, 2016). The relative contribution

of the traits to diversity was obtained using the method of Singh (1981). Multivariate analyses were carried out to obtain estimates of genetic divergence of genotypes based on the Gower Distance (Gower, 1971), relative to the 31 evaluated variables. Based on the generated distance matrix, the individuals were clustered by the UPGMA method (Unweighted Pair Group Method with Arithmetic Mean).

Results and discussion

Digital phenotyping means

Table 2 shows the means of 14 variables used to estimate genetic diversity in the 42 studied genotypes, consisting of three colors, two geometry, five texture, and four physiological analysis traits.

The C1 variable (Color: Above Otsu: CIELab: b) averaged 8.65. Its highest value was found in genotype 29 and the lowest in genotype 25. The C2 variable (Color: Above Otsu: Blue channel), averaged 92.87. The maximum value for this trait was 102.30 (genotype 38), and the minimum 79.25 (genotype 34) (Table 2). These results are the average of the blue channel of the RGB color space of the points calculated above the Otsu threshold. C3 (Color: Predominant: Blue Channel) averaged 96.90. The maximum value was 107.62 (genotype 38), and the minimum was 83.76 (genotype 34) (Table 2).

Barros et al. (2014) also used the CIELab system to study the color variability of the wood in the anatomical sections of the species Breu-Vermelha, Taurari-Vermelha and Pequiarana. The forest species showed differences in color, with Breu-Vermelha exhibiting a grayish pink color; Tauari-Vermelho, pinkish-gray; and Pequirana, grayish pink and/or pinkish gray. The b coordinate had a greater influence on the color characterization of the Tauari-Vermelha, and Pequiarana woods.

According to the method proposed by Singh (1981), the contribution of the 31 traits ranged from 1.57 to 4.13%. Of these, 16 contributed 50.40% to genetic diversity, the rest contributing 49.6% (Table 3). The traits that most contributed were related to histogram H5 and H6, 4.10 and 4.13%, respectively, and a texture-related variable (T1), 4.12%. Then, the color variables C2 and C3 contributed 3.75 and 3.70%, respectively; and, finally, a geometry trait (G2) contributed 3.78%.

Krause et al. (2017) evaluated 61 genotypes from inbred guava families and found that the geometry traits were those that most contributed to diversity, using the strategy of 30% of the evaluated descriptors for color, texture, and geometry. In the 10% strategy, the contributing variables were Below Otsu: green band; Below Otsu: CIELab: L and CIELab: distribution of a. The C1, C2 and C3 variables did not contribute in any of the adopted strategies.

Fachi et al. (2019) used digital phenotyping in seeds of 98 full-sib families of passion fruit and observed that the geometry descriptors were those that most contributed to genetic divergence between families. The C1, C2 and C3 variables also did not contribute to the calculation of genetic diversity in seeds of *Passiflora edulis*.

In this study, G1 (number of corners per SUSAN) was one of the geometry variables selected to calculate diversity, which averaged 64.84. Its maximum value was 99.91 (genotype 24) and the minimum was 46.60 (genotype 12). The G2 variable (geometry: SURF [speeded up robust features]) averaged 4.24, with a maximum value of 7.83 (genotype 26) and a minimum of 2.39 (genotype 17) (Table 2).

The average T1 was 109356.15, with the highest value (131586.46) found in genotype 38 and the lowest (95350.37) in individual 25. The T2 variable averaged 95971.36, with the maximum value (115617.34) found in genotype 38 and the minimum (83524.99) in genotype 25. T11 averaged 778.81; the highest and lowest values found for this variable were 1363.51 and 579.14, which were found in genotypes 24 and 4, respectively (Table 2).

The gray level difference method evaluates the texture features that describe the size and highlights of the image (Bharathi and Subashini, 2013). T12 averaged 658.30, ranging from 1074.85 (genotype 24) to 494.65 (genotype 22). The T14 variable averaged 623.10, with a maximum of 814.96 (genotype 38) and a minimum of 508.61 (genotype 28) (Table 2).

Table2. Mean values of color, geometry, texture and physiological analysis traits evaluated in the 42 genotypes.

*G	C1	C2	C3	G1	G2	T1	T2
1	9.11	96.62	100.14	63.88	3.89	107634.78	93902.92
2	9.75	89.07	93.35	69.15	7.18	122138.11	107176.98
3	9.63	94.11	97.48	64.38	4.78	116851.82	103115.31
4	7.38	97.86	101.03	60.63	3.09	99652.27	87336.29
5	9.80	94.42	98.97	63.89	3.98	109508.43	96222.91
6	7.39	97.29	101.56	65.90	3.57	109300.37	96829.62
7	9.70	89.96	93.49	62.14	4.10	106597.80	93607.91
8	8.03	90.90	95.46	56.93	3.43	104371.00	92441.99
9	10.76	87.64	92.12	61.06	5.13	114758.24	101618.70
10	9.99	94.81	99.54	49.57	2.78	102541.07	91863.18
11	7.51	100.35	103.47	73.87	2.99	110247.84	96021.54
12	9.23	90.67	93.46	46.60	2.55	10028.70	89148.98
13	10.93	96.44	102.54	60.92	5.00	116543.01	103253.30
14	6.48	90.27	94.40	56.86	2.91	104348.06	91420.71
15	8.39	85.92	89.10	49.12	3.49	97041.61	85893.37
16	7.80	90.47	94.95	80.81	4.92	118677.09	103591.98
17	7.05	89.65	92.51	62.45	2.39	97910.40	84981.52
18	7.19	100.29	104.13	66.44	3.14	107836.54	94354.68
19	8.45	96.59	100.05	71.22	4.00	107069.36	92506.66
20	8.57	100.27	104.04	74.02	4.12	117517.20	102325.67
21	12.21	83.54	88.63	48.49	4.15	102608.21	90747.23
22	6.76	94.61	98.91	58.46	2.59	101072.40	89722.90
23	8.11	92.43	96.23	61.13	3.13	101151.63	88985.34
24	7.80	98.70	104.05	99.91	6.50	131472.85	114225.83
25	5.57	85.64	89.31	54.12	3.17	95350.37	83524.99
26	9.72	84.76	88.19	79.20	7.83	126110.91	109105.77
27	7.18	95.52	99.78	61.93	2.74	97871.20	85377.49
28	6.80	88.62	92.05	62.02	3.52	104908.99	92581.42
29	14.33	85.95	89.58	53.50	5.75	105176.60	92621.61
30	7.28	88.21	92.38	60.24	3.66	103955.39	90860.36
31	8.63	90.57	93.16	72.67	5.32	113946.88	99614.89
32	9.01	98.45	103.67	68.98	4.29	111075.10	97190.33
33	7.56	101.63	106.26	67.99	4.08	107857.73	94097.33
34	10.82	79.25	83.76	56.65	4.16	98480.03	86497.29
35	7.38	91.07	95.12	60.18	3.92	99575.44	86844.91
36	9.37	84.09	88.18	53.11	5.28	112681.93	10000.62
37	9.18	95.49	99.71	65.21	4.12	106742.89	94061.23
38	9.86	102.30	107.62	81.35	6.85	131586.46	115617.34
39	9.17	99.21	102.87	79.80	4.28	120307.01	104639.75
40	7.85	98.75	102.53	65.23	4.22	111997.85	99323.38
41	7.26	92.63	97.42	83.93	6.93	123116.92	106930.23
42	8.18	95.51	98.78	69.27	4.15	115337.85	100612.74
Average	**8.65**	**92.87**	**96.90**	**64.84**	**4.24**	**109356.15**	**95971.36**
Maximum	**14.33**	**102.30**	**107.62**	**99.91**	**7.83**	**131586.46**	**115617.34**
Minimum	**5.57**	**79.25**	**83.76**	**46.60**	**2.39**	**95350.37**	**83524.99**

**G = genotypes; (C1) Color: Below Otsu: CIELab: b; (C2) Color: Above Otsu: Blue channel; (C3) Color: Predominant: Blue channel; (G1) Geometry: Number of corners per SUSAN; (G2) Geometry: SURF; (T1) Texture: Laws: ER; (T2) Texture: Laws: ES; (T11) res Texture: Run Length: GLD; (T12) Texture: Run Length: GLNU; (T14) Texture: Haralick: Variance; (GPAA) Germination % in the accelerated aging test; (GP) Germination %; (SL) Shoot length (cm); (TSW) Thousand-seed weight (g). Variables in bold are indices.

Cont. (tab. 2)

*G	T11	T12	T14	GPAA	GP	SL	TSH
1	660.54	565.51	624.96	95.50	98.50	3.34	13.59
2	1054.76	919.57	591.55	90.00	38.50	2.43	16.05
3	799.65	665.81	679.49	91.00	98.50	3.24	16.82
4	579.14	495.91	663.07	74.00	95.00	2.46	10.10
5	682.91	584.77	652.30	94.50	100.00	2.97	12.54
6	706.90	591.34	646.65	97.00	91.00	2.96	12.90
7	682.94	582.44	631.41	95.50	99.00	2.60	12.95
8	676.26	572.23	615.93	94.50	96.50	3.25	11.66
9	942.07	794.06	592.58	53.50	81.50	2.53	10.73
10	611.42	508.91	718.30	94.50	93.50	2.95	11.20
11	747.23	616.10	626.42	97.50	98.00	2.71	15.43
12	628.88	532.68	575.84	93.50	93.50	2.74	9.56
13	768.87	631.46	764.64	93.50	24.50	2.30	11.73
14	789.58	642.43	537.75	100.00	95.50	1.77	10.83
15	639.11	553.91	516.20	84.50	100.00	2.98	9.79
16	1056.65	866.69	576.23	32.00	97.00	3.23	15.61
17	681.73	589.72	539.27	94.50	91.50	2.60	11.26
18	720.90	603.59	635.07	100.00	93.00	2.47	11.72
19	724.65	624.81	632.74	90.00	96.00	3.28	13.48
20	789.33	664.53	680.92	96.50	100.00	3.47	16.04
21	743.63	639.69	583.75	98.00	100.00	2.52	11.30
22	588.62	494.65	649.93	95.00	100.00	2.60	9.12
23	632.12	541.45	624.21	84.00	100.00	2.69	10.36
24	1363.51	1074.85	624.46	98.50	92.50	3.73	17.01
25	603.15	531.99	545.91	53.50	100.00	2.97	7.52
26	1096.75	960.22	577.91	77.50	34.50	1.17	21.11
27	617.76	516.30	635.54	91.50	99.00	3.48	9.65
28	839.65	691.68	508.61	95.50	59.00	1.37	12.00
29	776.24	680.56	619.06	68.50	96.00	3.18	13.23
30	788.73	670.09	551.78	69.00	90.00	3.05	11.86
31	875.32	738.73	605.28	98.00	91.00	3.09	16.97
32	744.36	627.75	707.48	100.00	100.00	3.56	14.25
33	655.41	559.46	752.11	99.50	100.00	3.41	11.58
34	766.03	660.25	530.62	91.50	99.50	2.42	11.63
35	653.00	553.62	600.92	45.50	84.50	1.82	8.77
36	793.05	683.69	572.29	95.00	87.50	1.68	14.40
37	700.43	607.66	666.69	97.50	95.50	2.94	12.69
38	1014.04	822.87	814.96	95.50	100.00	2.98	16.18
39	857.95	721.37	646.85	89.00	94.50	2.63	18.21
40	776.75	648.75	647.04	92.50	97.00	2.99	15.39
41	1018.26	896.43	616.55	94.00	93.50	4.08	17.03
42	861.69	720.26	586.76	96.50	99.50	3.50	16.98
Average	**778.81**	**658.30**	**623.10**	**87.80**	**90.35**	**2.81**	**13.12**
Maximum	**1363.51**	**1074.85**	**814.96**	**100.00**	**100.00**	**4.08**	**21.11**
Minimum	**579.14**	**494.65**	**508.61**	**32.00**	**24.50**	**1.17**	**7.52**

**G = genotypes; (C1) Color: Below Otsu: CIELab: b; (C2) Color: Above Otsu: Blue channel; (C3) Color: Predominant: Blue channel; (G1) Geometry: Number of corners per SUSAN; (G2) Geometry: SURF; (T1) Texture: Laws: ER; (T2) Texture: Laws: ES; (T11) res Texture: Run Length: GLD; (T12) Texture: Run Length: GLNU; (T14) Texture: Haralick: Variance; (GPAA) Germination % in the accelerated aging test; (GP) Germination %; (SL) Shoot length (cm); (TSW) Thousand-seed weight (g). Variables in bold are indices.

Table 3. Relative contribution of 31 physical and physiological variables of *Psidium guajava* seeds to genetic divergence by the Singh method.

Variable	S_j	Value (%)
C1	61.45	2.27
C2	101.36	3.75
C3	100.09	3.70
G1	69.75	2.58
G2	102.09	3.780
H1	93.20	3.451
H2	82.90	3.070
H3	79.50	2.94
H4	97.06	3.59
H5	110.96	4.109
H6	111.57	4.13
H7	75.56	2.79
H8	67.36	2.49
H9	79.97	2.96
H10	87.44	3.23
H11	42.39	1.57
H12	80.50	2.98
H13	93.15	3.45
H14	81.34	3.012
H15	72.34	2.67
H16	105.26	3.89
H17	97.490	3.61
T1	111.41	4.12
T2	103.61	3.83
T11	74.53	2.760
T12	89.22	3.30
T14	79.05	2.92
GPAA	93.81	3.47
GP	95.87	3.550
SL	77.00	2.851
TSH	82.84	3.068

(C1) Color: Below Otsu: CIELab: b; (C2) Color: Above Otsu: Blue channel; (C3) Color: Predominant: Blue channel; (G1) Geometry: Number of corners per SUSAN; (G2) Geometry: SURF; (T1) Texture: Laws: ER; (T2) Texture: Laws: ES; (T11) res Texture: Run Length: GLD; (T12) Texture: Run Length: GLNU; (T14) Texture: Haralick: Variance; (H1) Histogram: Blue: Variance; (H2) Histogram: HSL: Luminance: Variance; (H3) Histogram: HSL: Hue: Minimum index; (H4) Histogram: HSL: Saturation: Variance; (H5) Histogram: LBP: Maximum index; (H6) Histogram: LBP: Minimum index; (H7) Histogram: LPQ: Maximum index; (H8) Histogram: LPQ: Minimum index; (H9) Histogram: Luminance: Variance; (H10) Histogram: NDLPQ: Maximum index; (H11) Histogram: NDLPQ: Minimum index; (H12) Histogram: Green: Variance; (H13) Histogram: Red: Variance; (H14) Histogram: YCbCr: Brightness: Variance; (H15) Histogram: YCbCr: Blue intensity: Variance; (H16) Histogram: YCbCr: Redintensity: Variance; (H17) Histogram: YIQ: Component I: Variance; (GPAA) Germination % in the accelerated aging test; (GP) germination %; (SL) Shoot length (cm); (TSW) Thousand-seed weight (g).

Torres et al. (2019) evaluated genetic diversity in a segregating population of passion fruit based on seed morphological and physiological descriptors and found that the texture variables were those that most contributed to genetic diversity. Similarly, the T1 variable (Laws: ER) was one of the variables that most contributed to genetic diversity, in this study.

The results obtained in this study and those described by Torres et al. (2019), Fachi et al. (2019), and Krause et al (2017) clearly demonstrate that the contribution of each variable to determine genetic divergence can vary according to species, genotype, and family.

As for the physiological traits, the average germination percentage in the accelerated aging test (GPAA) was 87.80%. Genotypes 14, 18, and 32 achieved 100% germination. Genotype 16 showed the lowest germination percentage (32%) (Table 2). According to Negreiros and Perez (2004), factors inherent to the seed such as differences in the level of vigor, moisture, conditions of the mother plant, and seed production location are important when evaluating the development of seedlings subjected to accelerated aging. Genotype 16 showed sensitivity to stress as a result of accelerated aging, which reduced its germination potential.

Germination percentage (GP) averaged 90.35%. Individuals 5, 15, 20, 21, 22, 23, 25, 32, 33 and 38 showed 100% germination. However, genotype 13 had a GP of only 24.50% (Table 2). This genotype also showed the highest percentage of seeds killed by the tetrazolium test, reinforcing the germination result. This difference in germination potential between genotypes can be explained by the different genetic constitutions of the evaluated accessions. As stated by Singh and Soni (1974), the guava seedcoat is impermeable to water and/or gases and may exhibit low germination. Thus, the low germination of genotype 13 may have been influenced by seed coat genetics.

The use of alternating temperatures can produce small cracks due to the expansion and retraction of the seed coat, thereby facilitating the passage of water into the seed. In the analysis of the GPAA variable, genotype 13 showed good

germination results after accelerated aging and so did genotype 2, which can be explained by the temperature difference that favored the rupture of the seed coat (Table 2). Fanti and Perez (1999) stated that the seed germination behavior can vary widely according to the type of substrate, physico-chemical factors, aeration, structure, water-holding capacity, among others. This variation can favor or hinder the germination process and seedling development. The genetic factor also influences the results of seed vigor and germination.

Alves et al. (2015) evaluated the germination of guava seeds at different temperatures and in different substrates. The highest average GP found by those authors was 97%, at alternating temperatures on paper roll, similar to the percentage found in this study.

Shoot length averaged 2.81 cm, ranging from 4.08 cm (genotype 41) to 1.17 cm (genotype 26). Genotype 41 showed a TSW of 17.03g, whereas genotype26 had a TSW of21.11g (Table 2). These results clearly show that genetic inheritance influences the vigor of genotypes and that higher TSW values will not be decisive for the choice of the genotypes with longer shoots.

Average TSW was 13.12 g. The highest weight (21.11 g) was found in genotype 26 and the lowest (7.52 g) in genotype 25 (Table 2). Genotype 26 also showed low germination, which may have resulted in low shoot dry matter. Seed vigor is a reflection of a set of traits that determine its physiological potential. Thus, the response of a seed can be more efficiently estimated by combining these data.

Cardoso et al. (2009) analyzed the physiological quality of papaya seeds from a germplasm bank and found an average TSW of 17.10 g and a GP of 78.19%. These results disagree with those obtained in this study for both variables, which averaged 13.12g and 90.35, respectively. These findings indicate that heavier seeds do not necessarily have higher germination rates, as was the case of genotype 26.

Among the 17 histogram variables used in this study, H5 (Histogram: LBP: maximum index) averaged 209.05, with values ranging from 252.30 (genotype 24) to 154.65 (genotype 4) (Table 4). This variable contributed 4.10% to genetic diversity.

The H6 variable (Histogram LBP: minimum index), in turn, averaged 47.76. Its maximum value was 66.23 (genotype 26) and its minimum was 39.07 (genotype 27), in relation to the rates of minimum occurrence in the histogram (Table 4).

Lei et al. (2019) used the RGB, LAB, YCbCr, YIQ and HSV color models under different lighting conditions to recognize ripe pomegranate fruits. The Cr component of the YCbCr model showed the best image, and the ideal segmentation was the threshold with recognition at 0.048s during the day, with 90.3% accuracy in the recognition of ripe pomegranates.

The use of histograms is important because they allow the extraction of important statistical attributes. To determine genetic diversity in guava seeds, there was a predominance of histograms. Among these, the variance values stood out, possibly due to their representative dispersions in relation to the means.

Genetic dissimilarity between S2 families.

The analysis of genetic divergence based on the seed's physical and physiological traits resulted in the formation of four groups (Figure 4). Genotypes with greater similarity were clustered within the groups. Group I was composed of 22 genotypes, representing the largest group; Group II was formed by 12 genotypes; group III contained six individuals; and group IV was composed of only two genotypes (Figure 4).

Campos et al. (2013) examined the genetic divergence of 138 guava genotypes obtained from controlled biparental crosses. Eight groups were formed by the Ward-MLM method, based on morphological, agronomic and physico-chemical analyses. According to Gonçalves et al. (2009), the number of groups can vary according to species, number of accessions and number and type of descriptors. Krause et al. (2017) evaluated 61 genotypes of *Psidium guajava* L. via digital seed analysis, using six clustering strategies, and three groups were formed in all of them, by the Ward-MLM method.

Table 4. Mean values of histogram traits evaluated in the 42 genotypes.

*G	H1	H2	H3	H4	H5	H6	H7	H8
1	239.21	494.06	1.13	611.42	204.23	43.59	148.59	104.74
2	222.59	445.87	3.21	727.17	240.60	66.04	121.51	98.66
3	252.88	520.74	4.28	656.86	223.80	49.09	136.40	90.45
4	238.35	514.50	4.17	648.93	154.65	40.04	126.99	85.25
5	236.52	495.71	3.86	676.74	197.40	42.42	170.14	84.47
6	230.84	498.83	4.02	614.81	193.50	43.64	134.81	98.14
7	219.36	476.06	4.00	706.59	202.50	46.28	138.70	95.86
8	209.57	468.06	5.33	702.21	208.43	43.28	136.46	83.64
9	237.41	464.68	3.77	781.69	219.60	53.00	144.01	92.07
10	266.24	554.70	5.59	676.26	164.85	40.90	127.52	95.49
11	227.11	489.93	1.39	554.44	220.50	45.95	114.03	101.30
12	182.75	432.63	1.30	687.71	203.70	41.29	134.83	90.58
13	313.35	596.54	1.88	600.89	211.67	46.99	117.66	91.75
14	160.87	407.69	1.36	736.60	222.08	43.66	153.67	88.44
15	165.00	394.69	3.01	759.83	192.90	40.90	133.50	83.75
16	180.99	435.38	3.53	651.29	244.50	55.89	137.89	89.66
17	158.96	411.30	3.55	670.14	217.73	44.12	130.31	91.48
18	213.86	490.86	2.29	567.71	200.33	42.58	136.19	92.32
19	225.25	493.21	3.97	599.56	204.53	45.64	134.13	97.05
20	251.51	529.67	2.37	584.78	206.25	48.20	120.58	98.82
21	232.60	451.53	2.70	868.49	195.38	46.66	118.75	94.54
22	203.36	485.74	2.84	709.68	172.73	39.43	121.76	79.10
23	198.84	471.05	5.53	659.31	175.50	41.76	128.11	86.81
24	223.59	487.43	2.81	548.19	252.30	63.42	136.75	109.83
25	154.35	393.25	4.17	843.80	199.73	40.77	113.47	67.94
26	230.33	442.21	1.47	811.40	244.88	66.23	138.98	104.64
27	195.85	481.63	3.09	621.13	167.80	39.07	124.45	81.07
28	155.32	385.00	3.49	700.81	226.20	49.19	139.80	92.94
29	323.63	496.98	5.40	935.39	179.63	51.33	122.58	88.45
30	186.65	425.82	5.99	686.94	216.45	46.55	123.48	100.35
31	214.59	464.06	8.70	658.95	235.43	53.60	137.68	95.10
32	305.39	564.27	5.89	630.62	195.68	47.70	123.48	98.85
33	279.26	580.25	5.78	593.76	182.33	43.30	129.32	97.36
34	195.05	411.65	2.39	837.16	216.53	49.39	162.08	87.87
35	172.99	446.63	4.98	723.10	194.18	41.35	130.45	82.31
36	204.26	431.63	4.79	792.76	234.30	52.42	126.59	91.80
37	280.28	527.82	3.57	695.10	179.11	44.27	148.72	94.24
38	347.79	636.67	1.25	600.12	237.15	54.83	134.14	103.98
39	233.57	501.20	2.53	542.02	229.80	54.89	132.31	99.04
40	239.92	511.72	3.30	543.15	224.63	48.30	123.09	93.82
41	226.85	482.85	3.40	624.96	243.08	58.50	125.75	107.53
42	218.88	462.00	1.40	587.42	249.90	49.30	126.56	100.89
Average	**225.14**	**479.92**	**3.56**	**676.90**	**209.20**	**47.76**	**132.53**	**93.15**
Maximum	**347.79**	**636.67**	**8.70**	**935.39**	**252.30**	**66.23**	**170.14**	**109.83**
Minimum	**154.35**	**385.00**	**1.13**	**542.02**	**154.65**	**39.07**	**113.47**	**67.94**

* G = genotypes; (H1) Histogram: Blue: Variance; (H2) Histogram: HSL: Luminosity: Variance; (H3) Histogram: HSL: Hue: Minimum index; (H4) Histogram: HSL: Saturation: Variance; (H5) Histogram: LBP: Maximum index; (H6) Histogram: LBP: Minimum index; (H7) Histogram: LPQ: Maximum index; (H8) Histogram: LPQ: Minimum index; (H9) Histogram: Luminance: Variance; (H10) Histogram: NDLPQ: Maximum index; (H11) Histogram: NDLPQ: Minimum index; (H12) Histogram: Green: Variance; (H13) Histogram: Red: Variance; (H14) Histogram: YCbCr: Brightness: Variance; (H15) Histogram: YCbCr: Blue intensity: Variance; (H16) Histogram: YCbCr: Red intensity: Variance; (H17) Histogram: YIQ: Component I: Variance. All variables in bold are indices.

Cont. (tab. 4)

*G	H9	H10	H11	H12	H13	H14	H15	H16	H17
1	669.49	138.27	132.38	651.24	1049.45	669.70	88.03	34.64	99.72
2	629.17	109.02	118.53	617.72	988.22	626.03	95.38	35.66	104.46
3	720.40	117.41	121.76	706.86	1104.03	717.53	94.90	34.58	102.53
4	710.25	113.41	106.17	692.67	1117.04	709.18	96.28	37.41	108.17
5	698.13	135.44	110.71	684.62	1091.06	695.09	101.74	36.43	108.65
6	690.99	109.60	119.42	674.59	1084.05	689.81	94.54	35.59	104.37
7	671.33	116.03	107.74	656.30	1067.29	668.93	99.80	37.61	110.00
8	657.59	100.99	110.27	639.30	1065.92	656.23	100.09	39.01	112.67
9	631.86	109.96	109.90	609.48	1037.01	631.71	100.88	40.01	114.55
10	762.45	98.30	110.76	744.88	1185.89	760.30	103.08	37.97	112.27
11	674.10	108.62	116.08	658.02	1044.76	673.32	87.00	32.19	95.25
12	614.13	111.75	108.22	593.25	1018.95	613.76	98.87	37.92	110.42
13	807.76	105.50	110.93	797.00	1175.18	803.07	90.57	30.74	93.83
14	584.56	118.34	103.62	558.95	1025.46	585.49	107.75	43.19	123.44
15	555.18	116.18	92.15	527.76	982.06	557.04	101.56	43.39	120.83
16	619.52	118.58	105.25	600.79	1019.30	618.14	98.03	37.01	108.23
17	582.43	108.30	112.38	557.55	999.40	583.00	97.51	39.52	112.55
18	682.89	128.53	109.60	665.27	1075.49	682.20	91.73	33.96	100.54
19	677.98	110.81	118.07	661.17	1058.88	677.19	88.44	33.57	98.12
20	727.66	119.11	111.38	712.37	1115.31	725.77	92.58	33.62	100.02
21	625.50	108.93	104.65	602.43	1047.64	625.02	110.46	43.27	124.49
22	697.73	121.98	103.06	678.55	1142.09	696.41	112.80	40.46	121.09
23	667.15	109.32	100.41	648.97	1078.23	665.45	99.84	37.74	110.38
24	674.39	115.81	108.76	658.86	1048.22	672.71	90.28	32.03	95.53
25	589.18	115.52	79.01	567.11	1039.45	587.87	119.92	46.20	134.10
26	614.38	125.38	114.42	598.20	989.45	612.41	98.56	38.31	110.51
27	683.91	102.92	104.33	666.59	1098.00	682.80	99.66	36.50	108.24
28	549.89	117.52	103.46	526.34	956.97	550.39	99.05	40.28	114.46
29	658.26	100.11	110.91	644.72	1011.54	655.21	99.48	40.39	113.83
30	592.00	107.82	104.66	569.10	991.99	592.19	94.18	39.09	110.04
31	644.68	112.25	116.68	628.26	1021.70	642.28	91.03	35.74	102.69
32	753.55	114.43	111.23	740.27	1119.22	751.37	88.37	33.14	97.08
33	802.32	100.12	109.58	790.08	1198.72	798.03	97.14	33.65	101.86
34	570.08	125.95	111.20	546.12	972.62	570.60	99.26	40.65	114.85
35	647.24	109.65	99.12	626.31	1090.07	645.37	109.40	42.45	122.85
36	607.69	104.62	114.23	585.40	1023.42	606.88	106.16	42.42	121.29
37	709.32	103.26	114.82	693.48	1082.11	707.66	89.52	35.57	101.86
38	863.91	111.42	107.85	855.61	1223.15	855.26	92.10	30.35	93.76
39	692.41	113.26	111.13	679.09	1052.44	689.92	85.57	30.69	91.85
40	689.75	120.04	109.92	671.79	1058.63	689.18	80.47	31.89	91.70
41	658.09	102.86	111.75	639.02	1044.44	658.03	89.45	35.29	101.48
42	627.72	119.96	119.59	609.44	993.51	628.27	84.26	33.33	95.89
Average	**666.36**	**113.27**	**109.67**	**648.47**	**1061.63**	**664.92**	**96.80**	**36.99**	**107.63**
Maximum	**863.91**	**138.27**	**132.38**	**855.61**	**1223.15**	**855.26**	**119.92**	**46.20**	**134.10**
Minimum	**549.89**	**98.30**	**79.01**	**526.34**	**956.97**	**550.39**	**80.47**	**30.35**	**91.70**

* G = genotypes; (H1) Histogram: Blue: Variance; (H2) Histogram: HSL: Luminosity: Variance; (H3) Histogram: HSL: Hue: Minimum index; (H4) Histogram: HSL: Saturation: Variance; (H5) Histogram: LBP: Maximum index; (H6) Histogram: LBP: Minimum index; (H7) Histogram: LPQ: Maximum index; (H8) Histogram: LPQ: Minimum index; (H9) Histogram: Luminance: Variance; (H10) Histogram: NDLPQ: Maximum index; (H11) Histogram: NDLPQ: Minimum index; (H12) Histogram: Green: Variance; (H13) Histogram: Red: Variance; (H14) Histogram: YCbCr: Brightness: Variance; (H15) Histogram: YCbCr: Blue intensity: Variance; (H16) Histogram: YCbCr: Red intensity: Variance; (H17) Histogram: YIQ: Component I: Variance. All variables in bold are indices.

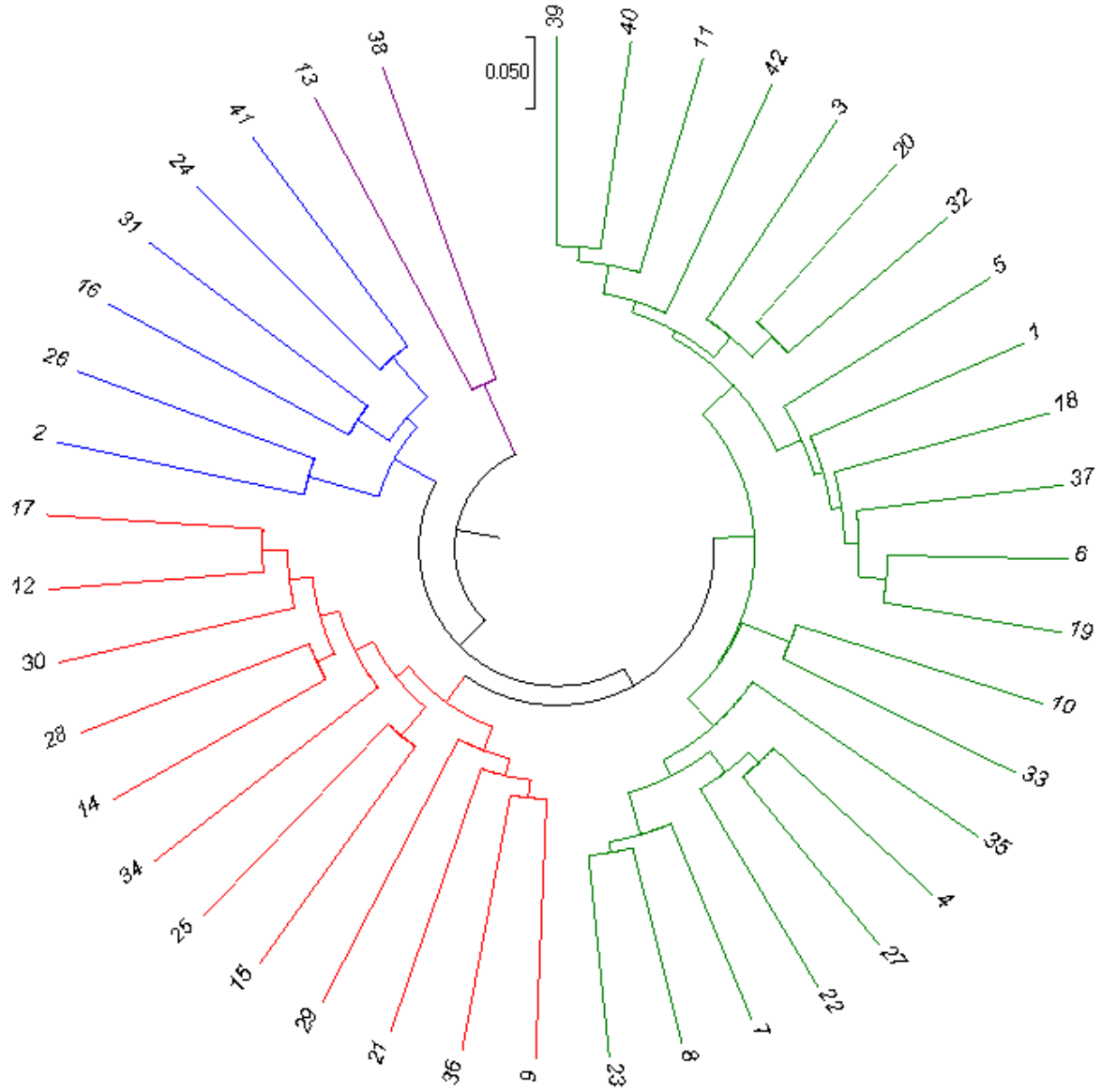

Figure 4. Dendrogram of genetic dissimilarity between 42 *Psidium guajava* genotypes, obtained by the UPGMA method, based on seed physical and physiological variables.

The genotypes allocated to group I obtained the highest mean for the following variables: C2, C3, T14, GPAA, H1, H2, H9, H11, H12, H13, and H14. These individuals showed the highest germination percentages under the evaluated conditions and, therefore, can be considered the most vigorous. Accordingly, these genotypes can tolerate more severe stresses in adverse conditions in the field and better withstand storage (Lopes et al., 2010). The accelerated aging test subjects the seeds to high temperature and humidity conditions, causing stress before they are taken to the germination test.

Group I also showed the highest means for the variables of GP and H3 (Histogram: HSL: Hue: minimum index). Therefore, these genotypes can be considered those that would have the best germination results in the field. Nevertheless, this group exhibited the lowest average TSW. Germination may have been more effective in the smaller seeds because of the larger contact surface with the substrate and, consequently, greater water absorption. Water is the factor that exerts the greatest influence on germination since its absorption allows the tissues to be rehydrated.

As a result of this action, respiration and all metabolic activities that provide energy and nutrients for the resumption of embryonic growth are intensified. The entry of water causes the seed

to enlarge, facilitating the rupture of the coat and, consequently, the emergence of the root hypocotyl (Carvalho and Nakagawa, 2012).

The individuals that comprised group II showed the highest means for six variables, namely, C1, H4, H7, H15, H16, and H17. The genotypes of group III, in turn, exhibited the highest means for 12 variables: G1, G2, T1, T2, T11, T12, SL, TSW, H5, H6, H8, and H10. In contrast, the genotypes in this group had the lowest values for GP. Seeds are considered reservoir organs, as they contain all the material necessary for the formation of future plants. In general, larger or denser seeds are those which were better nourished during their development, possessing larger amounts of reserve. However, in certain situations, larger seeds may not be the most vigorous (Carvalho and Nakagawa, 2012).

Genotypes 13 and 38, which belong to group 4, exhibited the highest histogram means. Among the physiological variables, these individuals showed the highest means for GPAA, and only individual 13 had alow GP (Table 3).

If one aims at crossing between the most vigorous groups, the cross between groups I and III would be indicated, given their higher means for the traits of GP, TSW and SL.

Conclusion

The traits that most contributed to genetic diversity were related to histogram.

There were the formation of four divergent groups. The genotypes allocated to group I obtained the highest mean for germination percentages.

Genotype 38, from group IV, can be recommended for future crosses with individuals from groups I and III, as it showed higher means for physiological traits and was the most divergent in relation to the individuals in the other groups.

References

ALVES, C.Z.; SILVA, J.B.; CÂNDIDO, A.C.S. 2015. Metodologia para condução de teste de germinação em sementes de goiaba. **Revista Ciência Agronômica**, 46(3):615-621.

ALVES, J.E.; FREITAS, B. M. 2007. Requisitos de polinização da goiabeira. **Ciência Rural**, 37:1281-1286.

AMBROSIO, M.; VIANA A.P.; RIBEIRO, R.M.; PREISIGKE, S.C.; CAVALCANTE, N.R.; SILVA, F.A.; TORRES, G.X.; SOUSA, C.M.B. 2021. Genotypic superiority of Psidium Guajava S1 families using mixed modeling for truncated and simultaneous selection. **Scientia Agricola,** 78(2):1-9.

ANDRADE, D.B.; OLIVEIRA, A.S.; PINTO, C.A.G.; PIRES, R.M.O.; OLIVEIRA, A.S.; SILVA, M.A.; CARVALHO, M.L.M. 2016. Detection of green seeds in soybean lots by the seed analysis system (SAS). **International Journal of Current Research**, 8:26462-26465.

BARROS, S.V. DOS S.;, MUNIZ, G.I.B; MATOS, J.L.M. 2014. Caracterização colorimétrica das madeiras de três espécies florestais da amazônia. **Cerne**, 20:337-342.

BASTOS, D.C.; RIBEIRO, J.M. 2011. **Produção de mudas de goiabeira**. Embrapa Semiárido. Comunicado técnico (CPATSA), 148:1-3.

BAY, H.; TUYTELAARS, T.; VAN, G.L. 2006. **SURF: Speed eduprobust features**. *In* 'European conference on computer vision'. pp. 404-417. Springer, Berlin, Heibelberg.

BHARATHI, P.T.; SUBASHINI, P. 2013. Texture feature extraction of infrared river ice images using second order spatial statistics. **International Journal of Computer, Electrical, Automation and Information Engineering**, 7:195-204.

BRAZIL. Ministério da Agricultura, Pecuária e Abastecimento. 2009. Regras para Análise de Sementes. Brasília: Ministério da Agricultura e Abastecimento. Available at: https://www.gov.br/agricultura/pt-br/assuntos/insumos-agropecuarios/arquivos-publicacoes-insumos/2946_regras_analise__sementes.pdf. Accessed Jul 15 2020.

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**, 38:151-157.

CAMPOS, B.M.; VIANA, A.P.; QUINTAL, S.S.R.; GONÇALVES, L.S.A.; PESSANHA, P.G.O. 2013. Quantification of the genetic divergence among guava accessions using Ward-MLM strategy. **Revista Brasileira de Fruticultura**, 35:571-578.

CARDOSO, D.L.; SILVA, R.F.; PEREIRA, M.G.; VIANA, A.P.; ARAÚJO, E.F. 2009. Diversidade genética e parâmetros genéticos relacionados à qualidade fisiológica de sementes em germoplasma de mamoeiro. **Revista Ceres**, 56:572-579.

CARVALHO, N.M.; NAKAGAWA, J. 2012. **Sementes: ciência, tecnologia e produção**. 5. ed. Jaboticabal: Funep, 590p.

CRUZ, C.D. 2016. GenesSoftware - ampliado e integrado ao R, Matlab e Selegen. **Acta Scientiarum Agronomy**, 38:547-552.

FACHI, L.R.; KRAUSE, W.; VIEIRA, H.D.; ARAÚJO, D.V.; LUZ, P.B.; VIANA A.P. 2019. Digital image analysis to quantify genetic divergence in passion fruit (*Passiflora edulis*) seeds. **Genetics and Molecular Research**, 2:1-12.

FANTI, S.C.; PEREZ, S.C.J.G.A. 1999. Influência do substrato e do envelhecimento acelerado na germinação de olho-de-dragão (*Adenanthera pavonina* L. – Fabaceae). **Revista Brasileira de Sementes**, 21:135-141.

FAO - Food and Agriculture Organization of the United Nations. 2019. Available at http://www.fao.org/faostat/en/#data/QC/. Accessed on Dez 01, 2019.

GALLOWAY, W.E. 1975. **Process framework for describing the morphologic and stratigraphic evolution of deltaic depositional system**. In Broussard, M. L., (ed). Deltas: Models for exploration. Houston Geological Society, 87- 98.

GONCALVES, L.S.; RODRIGUES, R.; AMARAL JUNIOR, A.T.; KARASAWA, M. 2009. Heirloom tomato gene bank: assessing genetic divergence based on morphological, agronomic and molecular data using a Ward-modified location model. **Genetic and Molecular Research**.8:364-374.

GOWER, J. 1971. "A general coefficient of similarity and some of its properties". **Biometrics**. 27:857-872.

HARALICK, R.M., SHANMUGAM K., DINSTEIN I. 1973. **Textural Features for Image Classification**; IEEE Transaction on Systems, Man and Cybernetics, 3: 610-621.

HARTMANN, H.T.; KESTER, D.E . 2002**. Propagação de plantas: princípios e práticas**. (7. ed.). Nova Jersey: Prentice-Hall, 880p.

IBGE - Instituto Brasileiro de Geografia e Estatística. Rio de Janeiro. 2019. Available at https://sidra.ibge.gov.br/tabela/5457#resultado. Accessed on Dez 01, 2019.

KARA, M.,; SAYINCI, B.; ELKOCA, E.; ÖZTÜRK, İ.; OZMEN, T.B. 2013. Seed size and shape analysis of registered common bean (*Phaseolus vulgaris* L.) cultivars in turkey using digital photography. **Journal of Agricultural Sciences**, 19: 219-234.

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. **Genetic and Molecular Research**, 16:1-11.

LEI, X.; OUYANG, H.; XU, L. 2019. Mature pomegranate recognition methods in natural environments using machine vision. **Ciência Rural,** 49:1-7.

LIMA, M.A.C.; ASSIS, J.S.; NETO, L.G. 2002. Caracterização dos frutos de goiabeiras e seleção de cultivares na região do sub-médio São Francisco. **Revista Brasileira de Fruticultura,** 24:273-276.

LOPES, M.M.; SADER, R.; PAIVA, A.S.; FERNANDES, A.C. 2010 Teste de envelhecimento acelerado em sementes de quiabo. **Bioscience Journal**, 26:491-501.

MAGUIRE, J. D. 1962. Speed of Germination-Aid in Selection and Evaluation for Seedling Emergence and Vigor. **Crop Science**, 2:176-177. https://doi.org/10.2135/cropsci1962.0011183X000200020033x.

MAITAN, M.Q.; VIANA, A.P.; VIEIRA, H.D.; SILVA, C.C.A.; RODRIGUES, D.L.; LEAL, L.S.S. 2020. Physiological responses of seeds from full-sib guava families to different substrate temperatures. **Revista Brasileira de Fruticultura,** 42:1-8.

MASETTO, T. E.; DAVIDE, A. C.; FARIA, J. M. R.; SILVA, E. A. A.; REZENDE, R. K. S. 2009. Avaliação da qualidade de sementes de Eugenia pleurantha (MYRTACEAE) pelos testes de germinação e tetrazólio. **Agrarian**, 2:33-46.

MEDINA, W.; SKURTYSB, O.; AGUILERA, J.M. 2010.Study on image analysis application for identification quinoa seeds (*Chenopodium quinoa Willd)* geographical provenance. **Journal of Food Science and Technology**, 43: 238-246.

NEGREIROS, G.F.; PEREZ, S.C.J.G.A. 2004. Resposta fisiológica de sementes de palmeiras ao envelhecimento acelerado. **Pesquisa Agropecuária Brasileira**, 39:391-396.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. 1996. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, 29:51-59.

OJANSIVU, V., HEIKKILÄ, J. 2008.Blur insensitive texture classification using local phase quantization. In: **International conference on image and signal processing**, Springer, Berlin, Heidelberg, 236-243p.

OLIVEIRA, F.T.; HAFLE, M. O.; MENDONÇA, V.; MOREIRA, J.N.; MENDONÇA, L.F.M. 2013. Fontes e proporções de materiais orgânicos na germinação de sementes e crescimento de plantas jovens de goiabeira. **Revista Brasileira de Fruticultura,** 35:866-874.

PEREIRA, F.M.; KAVATI, R. 2011.Contribution of Brazilian scientific research in developing some of subtropical fruit. **Revista Brasileira de Fruticultura**, 92-108.

PERSECHINO, A.; ALBUQUERQUE, M.P. 2015. **Processamento de imagens: conceitos fundamentais**. Monografias do Centro Brasileiro Pesquisas Físicas (CBPF). 1-41.

PESSANHA, P.G.O.; VIANA, A.P.; AMARAL JÚNIOR, A.T.; SOUZA, R.M.; TEIXEIRA, M.C., PEREIRA; M.G. 2011. Avaliação da diversidade genética em acessos de *Psidum Spp*. via marcadores RAPD. **Revista Brasileira de Fruticultura**, 33:129-136.

PINTO, C.A.G.; CARVALHO, M.L.M.; ANDRADE, D.B.; LEITE, E.R.; CHALFOUN, I. 2015. Image analysis in the evaluation of the physiological potential of maize seeds. **Revista Ciência Agronômica**, 46:319-328.

QUINTAL, S.S.R.; VIANA, A.P.; CAMPOS, B.M.; VIVAS, M.; AMARAL JÚNIOR, A.T. 2017. Selection via mixed models in segregating guava families based on yield and quality traits. **Revista Brasileira de Fruticultura**, 39:1-8.

SETIAWAN, A. S.; WESLEY, J.; PURNAMA, Y. 2015. Mammogram classification using law's texture energy measure and neural networks. **Procedia Computer Science**, 59:92-97.

SILVA, F. A.; VIANA, A.P.; CORRÊA, C. C. G.; CARVALHO, B. M.; SOUSA, C. M. B.; AMARAL, B.D.; AMBRÓSIO, M.; GLÓRIA, L. S. 2020. Impact of Bayesian Inference on the Selection of *Psidium guajava*. **Scientific Reports**, 10:1999-2008.

SILVA, F.A.; CORREA, C.C.G.; CARVALHO, B.M.; VIANA, A.P.; PREISIGKE, S.C.; AMARAL JÚNIOR, A.T. 2021. Novel approach to the selection of *Psidium guajava* genotypes using latent traits to bypass multicollinearity. **Scientia Agricola**, 78:5-13.

SINGH, S., SONI, S.L. 1974. Effect of water and acid soaking periods seed germination in guava. **Punjab Horticulture Journal**, 14:122-124.

SINGH, D. 1981. The relative importance of characters affecting genetic divergence. **The Indian Journal of Genetics e Plant Breeding**, 41:237-245.

SOUZA, C.M.B; RIBEIRO, R.M.; VIANA, A.P.; CAVALCANTE, N.R.; SILVA, F.A.; AMBRÓSIO, M.; AMARAL JÚNIOR, A.T. 2020. Guava breeding via full-sib family selection: conducting selection cycle and divergence between parents and families. **Crop Breeding and Applied Biotechnology**, 20:1-9.

TORRES, G.X.; VIANA, A.P.; VIEIRA, H.D.; RODRIGUES, D.L.; SANTOS, V.O. 2019. Contribution of seed traits to the genetic diversity of a segregating population of *Passiflora* spp. **Chilean Journal of Agricultural Research**, 79:288-295.

VENORA, G.; GRILLO, O.; SHAHIN, M.A.; SYMONS, S.J. 2007.Identification of Sicilian landraces and Canadian cultivars of lentil using an image analysis system. **Food Research International,** 40:161-166.