



Full Length Article

Understanding Genetic Diversity in Selected Genotypes of *Coffea canephora* by Seed Attributes

Laísa Zanelato Correia¹, Henrique Duarte Vieira¹, Rosenilda de Souza^{1*}, Alexandre Pio Viana¹, Fábio Luiz Partelli², Eileen Azevedo Santos³ and Amanda Justino Acha⁴

¹Laboratório de Fitotecnia, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes, 28013-602, Brasil

²Centro Universitário Norte do Espírito Santo, Universidade Federal do Espírito Santo, São Mateus, 29932-540, Brasil

³Universidade do Estado do Mato Grosso, Tangará da Serra, 78301-532, Brasil

⁴Centro Universitário Serra dos Órgãos, Teresópolis, 25964-004, Brasil

*For correspondence: rosenilda@pq.uenf.br

Received 27 May 2024; Accepted 22 June 2024; Published 25 September 2024

Abstract

Coffee breeding programs aim at developing superior genotypes, which entail exploring the genetic variability existing in conilon/robusta coffee. This study aimed at examining the variability and estimate genetic param of physical and physiological nature in *Coffea canephora* seeds in order to provide bases for the selection of genotypes for superior seeds quality. Forty-three *C. canephora* genotypes were evaluated for 11 physical and physiological seed traits. Estimates of genetic param, environmental variance, phenotypic variance, genotypic variance, broad-sense heritability, variation index and Pearson's correlation were obtained. The Dissimilarity between genotypes was obtained by generalized Mahalanobis distance. The dendrogram was constructed using unweighted pair group method with arithmetic mean (UPGMA). *C. canephora* genotypes presented high germination potential and exhibited genetic diversity for all seed traits, with estimates of genetic variance greater than environmental variance. Heritability reached values above 96.0% and all genetic variation indices found were higher than unity. The UPGMA hierarchical method grouped the genotypes in three groups. The physical traits (circumference, area and roundness) were those that contributed the most to explaining the dissimilarity between genotypes. Genotypes Alecrim, AP, Bamburral, Z39 and Tardio C were the most divergent. The results revealed the presence of genetic variability between *C. canephora* genotypes and greater genetic than environmental influence on the characters of seeds, indicating that selection for the studied traits showed favorable conditions in terms of immediate gain.

Keywords: *Coffea* spp.; Digital phenotyping; Germination potential; GroundEye® system; Multivariate analysis

Introduction

The coffee belongs to the genus *Coffea* spp, which includes at least 124 species found in tropical regions around the world. Among these species, *C. canephora* Pierre ex Froehner, popularly known as conilon and/or robusta coffee, stands out for its great economic importance in the roasted/ground coffee and solubilization industries as well as for its use in blends with Arabica coffee (Vareltzis *et al.* 2020).

As a diploid species with gametophytic self-incompatibility, conilon coffee reproduces preferentially by allogamy, as a result of highly heterogeneus and heterozygous individuals with high genetic variability (Vázquez *et al.* 2019). This genetic variability makes it possible to identify individuals with different characteristics (Giles *et al.* 2018; Martins *et al.* 2019), which in turn allows

the selection of genotypes for different cultivation purposes, yet always prioritizing productivity (Dubberstein *et al.* 2020; Partelli *et al.* 2020a). Exploring the genetic variability in the conilon coffee population may reveal valuable genetic resources that provide greater success in breeding programs. Moreover, the selection of superior genotypes can contribute favorable alleles to the next generations, ensuring more productive genotypes that better adapt to climatic adversities.

The genetic variability existing between genotypes can be determined by differences based on morphological, physiological, molecular or biochemical characters. In most cases, this diversity can be quantified by multivariate procedures such as generalized Mahalanobis distance (D^2), Euclidean distance, principal components. The most commonly used clustering methods are unweighted pair

To cite this paper: Correia LZ, HD Vieira, R Souza, AP Viana, FL Partelli, EA Santos, AJ Acha (2024). Understanding genetic diversity in selected genotypes of *Coffea canephora* by seed attributes. *Intl J Agric Biol* 32:425–432

© 2024 The Authors. International Journal of Agriculture and Biology published by Friends Science Publishers, Faisalabad, Pakistan

This is an open access article under the terms of the Creative Commons Attribution License, which permits non-commercial use, distribution and reproduction in any medium, provided the original work is properly cited

group with arithmetic mean (UPGMA) and Ward's, among others. Thus, diversity analyses can provide information for the identification of parents whose cross will have a greater heterotic effect, increasing the possibilities of obtaining progeny with superior characters (Fonseca *et al.* 2006).

Researchers have already looked into the genetic diversity of coffee, especially to morphological and physiological characters of plants and fruits (Giles *et al.* 2018; Dubberstein *et al.* 2020; Partelli *et al.* 2021; Schmidt *et al.* 2022). However, there is a need for studies focusing on the physiological quality of seeds to generate early knowledge by testing vigor and germination of genotypes submitted to the selection process. As such, studying the genetic divergence between genotypes to potential of seeds is necessary for the progress of a breeding program.

In a breeding program, the genetic variability and high average between genotypes is of paramount importance. With this knowledge, it is possible to improve the efficiency of identification of the best combinations between genotypes, thereby ensuring populations with greater genetic variability. In view of this scenario, the present study was conducted to examine the divergence and estimate the genetic param for physical and physiological characters of *C. canephora* seeds in order to provide bases for the selection of genotypes for superior seeds quality.

Materials and Methods

Genotypes and experimental conditions

The experiment was implemented in the municipality of Nova Venécia – ES, Brazil, in April 2014 (18°39'43" S and 40°25'52" W, 200 asl). The region has an average annual temperature is 23°C, tropical climate, classified as Aw type, according to the Köppen system (Alvares *et al.* 2013). The soil is classified as a Red-Yellow Oxisol (Santos *et al.* 2018). Forty three *Coffea canephora* genotypes were evaluated, including cultivars Emcapa 8111 and Emcapa 8131 (Bragança *et al.* 2001) and the Andean varieties (Partelli *et al.* 2019) Tributum (Giles *et al.* 2018; Partelli *et al.* 2020a) and Monte Pascoal (Partelli *et al.* 2020b). Of these, 42 were derived from clonal propagation by cuttings, and one genotype was obtained from seeds (Partelli *et al.* 2021).

The plants were arranged in a randomized block design with three blocks. Each plot (experimental unit) consisted of seven plants, with a spacing of 3 m between rows and 1 m between plants, which would correspond to a density of 3,333 plants ha⁻¹ (Partelli *et al.* 2021). The management of culture was carried out according to technical recommendations (Partelli *et al.* 2021). Meteorological data from the fruit development period were obtained from the meteorological weather station of Nova Venécia (Fig. 1), located 7 km away from the experimental site (18.695° S, 40.390° W; mean altitude, 156 m asl). The adequate level of water in the soil was maintained with drip

irrigation. The water balance was calculated daily considering crop evapotranspiration, precipitation, and soil water storage (Partelli *et al.* 2021).

The fruits were collected in the field at the cherry stage and sent to the seed laboratory located at the State University of Northern Rio de Janeiro (CCTA-UENF), in Campos dos Goytacazes – RJ, Brazil. The fruits were pulped manually. For drying, the seeds surrounded by the parchment were placed in plastic trays until reaching a moisture content between 14 and 15%, which was determined using a DOLE 500 moisture meter. Subsequently, the seeds were packed in polyethylene bags properly labeled, which were placed inside paper bags (the two packages protect against humidity and exposure to light) and taken to a cold room at 10°C for one day.

Evaluated characters

The seeds of the genotypes were evaluated for percentage of mocha seeds and percentage of flat seeds, physical traits (1000 seed weight (TSW), area, roundness and circumference of the seeds), and physiological traits (germination speed index (GSI), germination percentage, abnormal seedlings, seedlings shoot and root lengths). The experiments were laid out in a completely randomized design with 43 treatments. For the tests, the seeds were used without the endocarp (parchment).

The TSW was determined from eight replicates of 100 seeds of each treatment, following the Rules for Seed Testing (Brasil 2009). The percentage of flat seeds (FS) and mocha seeds (MS) was determined from ten replicates of 100 seeds, randomly collected from the portion of pure seeds. Physical traits, area, roundness and circumference, were evaluated by digital phenotyping using the GroundEye® instrument from four replicates of 50 seeds. The seeds were positioned in an acrylic tray for image capture to obtain geometric data.

The physiological traits, germination speed index, abnormal seedlings, seedlings shoot and root lengths, were evaluated in an experiment with four replicates of 50 seeds, according to the Rules for Seed Testing (Brasil 2009). The seeds were placed in substrate between paper moistened with water at the rate of two parts water (Brasil 2009). Additionally, a 1% CAPTAN® solution was applied on the seeds to avoid possible contamination by microorganisms. The rolls were placed in transparent polyethylene bags and taken to germinators set at a temperature of 30°C with the photoperiod adjusted to 8 am and 4 pm. After 30 days, normal seedlings, abnormal seedlings and non-germinated seeds were counted (Brasil 2009). During the period of conducting the germination test, the GSI was calculated by counting, every two days, the number of germinated seeds. Germinated seeds were considered those with 0.5 cm of developed root system. The GSI was calculated by the formula of Maguire (1962). Shoot and root lengths were measured in 10 random seedlings of each replicate, with a graduated ruler.

Analysis of variance and estimates of genetic param

The obtained data were subjected to analysis of variance by F-test ($p \leq 0.01$) following the random statistical-mathematical model given by equation:

$$Y_{ij} = \mu + G_i + B_j + E_{ij}$$

Where Y_{ij} is the observation regarding effect i in replicate j ; μ is the general constant; G_i is the effect of genotype i ($i = 1, 2, 43$), with $G_i \sim \text{NID}(0, \sigma^2)$; B_j is the effect of block j ($j = 1, 2, 3, 4$), with $B_j \sim \text{NID}(0, \sigma^2)$; and E_{ij} is the experimental error, with $E_{ij} \sim \text{NID}(0, \sigma^2)$. From analysis of variance, estimates of genetic param: environmental variance (σ_e^2), phenotypic variance (σ_p^2), genotypic variance (σ_g^2), heritability (h^2) and genetic variation index ($\hat{I}_v$), were obtained, using the formula proposed by Cruz *et al.* (2014). Genotype means were grouped by the method proposed by Scott Knott (1974) at 5% significance level. Pearson's linear correlation coefficient was estimated for 11 studied variables at 5% significance level.

Dissimilarity between genotypes: Genetic diversity between genotypes was assessed by the multivariate cluster analysis technique applied to the genotype means. To test the efficiency of each hierarchical clustering method, the cophenetic correlation coefficient was calculated, where higher coefficients indicate greater efficiency of the method in question. Generalized Mahalanobis distance (D^2) was adopted as the dissimilarity measure. The clustering dendrogram was obtained from the method unweighted pair group method with arithmetic mean (UPGMA). Data were analyzed using the computational resources of Genes software (Cruz 2016).

Relative importance of characters: The relative importance of the eight characters was calculated using the method proposed by Singh (1981), which is based on the partition of the total estimates of D^2_{ii} distances, considering all possible pairs of individuals, for the part due to each character.

Results

Analysis of variance revealed significant differences between the genotypes ($P \leq 0.01$) for all variables evaluated (Table 1). The experimental coefficient of variation ranged from 0.97 (roundness) to 31.16% (abnormal seedlings) (Table 1).

With respect to the physiological seed traits, the genotypes presented averages ranging from 0 to 100% for germination, of 3.19 to 7.52 for germination speed index (GSI), of 0 to 93% for abnormal seedlings, of 0.0 to 4.72 cm for shoot length, and of 0.0 to 4.73 cm for root length (Table 2). Genotype Z21 reached the highest abnormal seedlings (93.0%), followed by genotypes Ouro

Negro (85.5%) and 700 (73.5%). As regards the shoot length variable, genotypes CH1, Emcapa 02 and 122 achieved the greatest growth, differing statistically from the others. For root length, genotypes AD1, Beira Rio 8, Peneirão, Z39, Z35, Z38, Tardio C, Verdim D and Sementes exhibited the greatest growth.

In the analysis of physical seed traits, the genotypes presented averages ranging from 0.44 to 0.77 cm² for area; of 0.64 to 0.81 cm for roundness; and of 2.63 to 8.43 cm for circumference. The TSW varied from 172.30 to 348.62 g. The percentage of mocha seeds and flat seeds varied from 3.1 to 40.8% and of 60.0 to 96.9%, in this order (Table 2). In terms of seed area, genotype Beira Rio 8 showed the highest value (Table 2). Genotypes Graudão HP and P1 had the highest TSW values. As for percentage flat seeds, genotypes 700, Imbigudinho, Z36, Emcapa 02, LB1 and 122 stood out. Lastly, for percentage mocha seeds, genotypes Bicudo and Z29 had the highest values.

In the estimates of genetic param, genotypic variance (σ_g^2) was higher than the environmental variance for all variables. All variables achieved broad-sense heritability (h^2) estimates above 96%. In the same parameter, TSW reached the highest estimate (Table 1).

The variation index ($\hat{I}_v$) reached values greater than unity, ranging from 2.73 (root length) to 17.98 (TSW). In the same data sequence, seed circumference also stood out with a $\hat{I}_v$ of 15.42 (Table 1).

Correlation analysis detected positive and negative correlations between characters at a 1% probability. There was a strong negative correlation between germination percentage and the percentage of abnormal seedlings ($r = -0.994$) and between flat and mocha seeds ($r = -1.0$). In addition, there was a moderate positive correlation ($r = 0.673$) between germination percentage and shoot length. The percentage of abnormal seedlings showed a moderate negative correlation with shoot length ($r = -0.689$) and root length ($r = -0.603$). There was a strong positive correlation ($r = 0.961$) between seed area and TSW (Table 2).

Of 11 analyzed variables in the present study, the eight that presented variations between genotypes were selected for dissimilarity analysis, being: germination, GSI, abnormal seedlings, shoot length, root length; area, roundness and circumference seeds.

Genetic divergence analysis resulted in the formation of three groups by taking an arbitrary cut-off point, with cophenetic correlation coefficient of 0.95 (Fig. 2). Group I was the largest, including 81% of the studied genotypes. The group reached an average germination of 92.4%, GSI of 5.10 and root length of 3.29 cm, standing out positively among the clusters. Seed circumference, in turn, was the smallest among the groups (3.12 cm). Group II was composed of three genotypes, which had an average AS of 84.0%, germination of 10.2%, GSI of 4.20, shoot length 1.26 cm and root length 0.94 cm. The genotypes in this cluster had the highest means for area (0.68 cm²) and roundness (0.77 cm). Group III was composed of five

Table 1: Estimates of mean squares and genetic parameters of physical and physiological seeds traits of 43 genotypes of *C. canephora*

SV	DF	Mean square										
		GSI	G	AS	SL	RL	Area	RD	Circ	TSW	FS	MS
B	3	0.77	10.60	8.34	0.13	0.13	0.00	0.00	0.01	0.10	13.11	13.11
G	42	4.94**	1937.49**	1775.57**	3.22**	4.55**	0.03**	0.01**	8.13**	88.80**	425.03**	425.03**
R	126	0.05	15.18	11.36	0.08	0.15	0.00	0.00	0.01	0.07	9.00	9.00
Total	171											
CV _e (%)		4.35	4.52	31.16	8.80	12.33	2.03	0.97	2.54	1.06	3.65	16.66
Mean		5.06	86.2	10.8	3.13	3.12	0.58	0.73	3.63	246.21	82.10	17.9
Amplitude		3.19 – 7.52	0.0 – 100.0	0.0 – 93.0	0.0 - 4.72	0.0 – 4.73	0.44 – 0.77	0.64 – 0.81	2.63 – 8.43	172.3 – 348.6	60.0 – 96.9	3.1 – 40.8
Genetic parameters												
σ_p^2		1.24	484.37	443.89	0.81	1.14	0.01	0.00	2.03	22.20	106.26	106.26
σ_e^2		0.01	3.79	2.84	0.02	0.04	0.00	0.00	0.00	0.02	2.25	2.25
σ_g^2		1.22	480.58	441.05	0.79	1.10	0.01	0.00	2.03	22.18	104.01	104.01
h^2 (%)		99.01	99.22	99.36	97.64	96.75	99.46	99.4	99.89	99.92	97.88	97.88
$\hat{I}_v$		5.03	5.63	6.23	3.22	2.73	6.76	6.47	15.42	17.98	3.40	3.40

SV = source of variation; DF = degrees of freedom; GSI = germination speed index (%); G = germination (%); AS = abnormal seedlings (%); SL = shoot length (cm); RL = root length (cm); Area = (cm²); RD = Roundness (cm); Circ = Circumference (cm); TSW = 1000-seed weight (g); FS = flat seeds (%); MS = mocha seeds (%); CV_e (%) = coefficient of genetic variation; σ_p^2 = phenotypic variance; σ_e^2 = environmental variance; σ_g^2 = genetic variance; h^2 = heritability; $\hat{I}_v$ = variation index. ** - significant at 1% probability by the F test

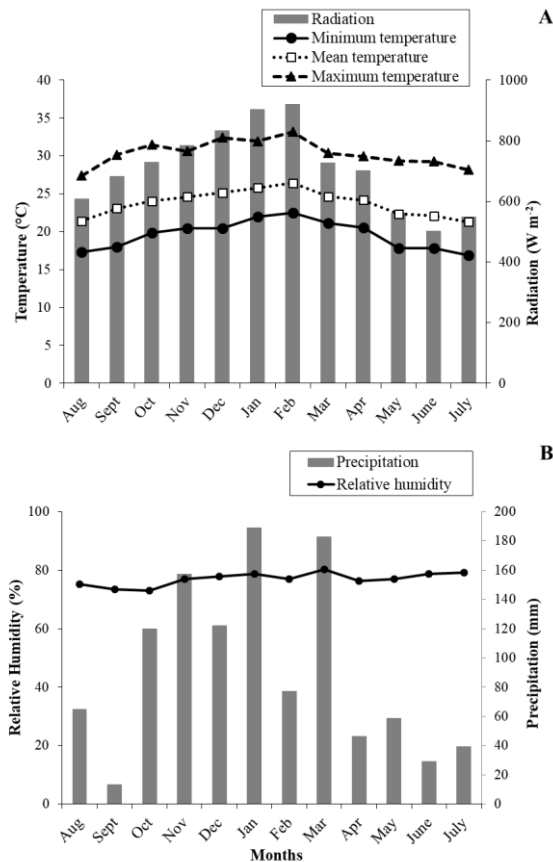


Fig. 1: Monthly means of (A) solar radiation and minimum, mean and maximum air temperature and (B) relative air humidity and precipitation for harvest (August 2019 to July 2020), obtained at the meteorological station of Nova Venécia, Espírito Santo, Brazil

genotypes, with an average of 88.6% for germination and 8.8% for abnormal seedlings. This group exhibited the highest mean for shoot length (3.33 cm) and circumference (7.41 cm) but had the lowest area (0.54 cm²) and roundness (0.72 cm) among the groups.

The genotypes Alecrim, AP, Bamburral, Z39 and Tardio

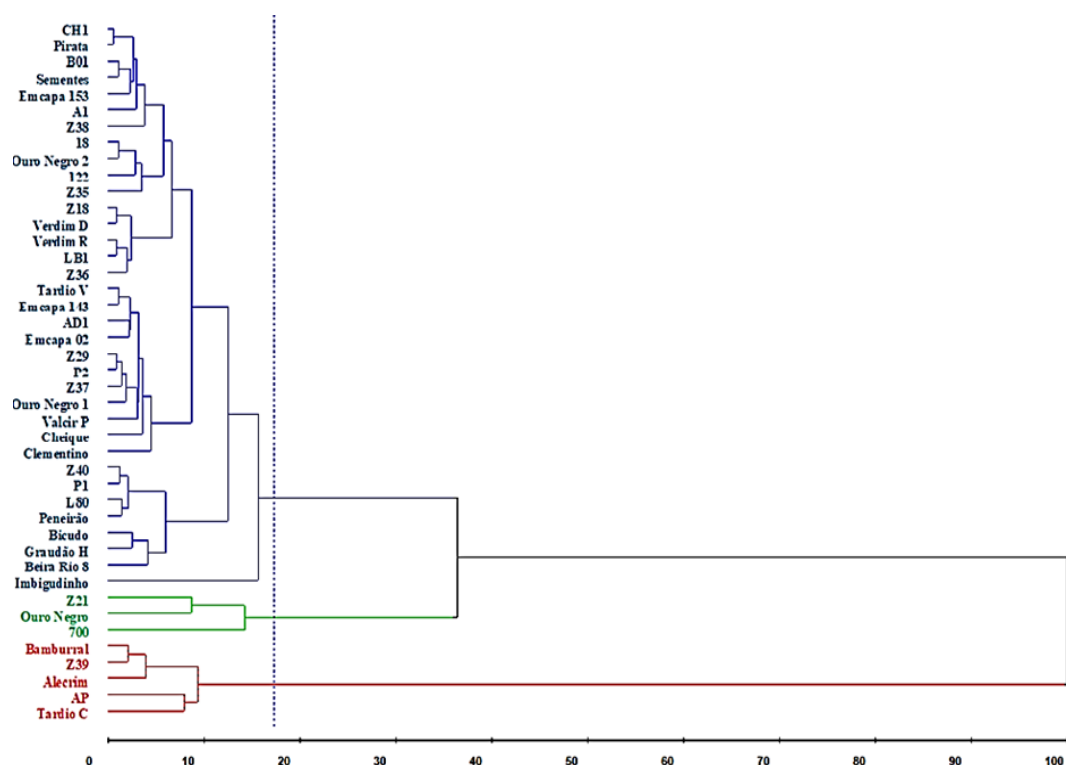
C, belonging to group III, were the most genetically distant from the other genotypes (Fig. 2). These had the largest seed circumference, for which genotype Tardio C stood out with the highest value (Table 2).

In Fig. 3, the genotypes were arranged on the map according to their geographic location of origin as well as

Table 2: Estimates of phenotypic correlation coefficients between seeds traits of 43 genotypes of *C. Canephora*

Variables	G	AS	SL	RL	Area	RD	Circ	FS	MS	TSW
GSI	0.377*	-0.334*	0.336*	-0.015 ^{ns}	-0.263 ^{ns}	-0.045 ^{ns}	-0.110 ^{ns}	0.383 ^{ns}	-0.383*	-0.245 ^{ns}
G		-0.994**	0.673**	0.591**	-0.348*	-0.215 ^{ns}	0.001 ^{ns}	-0.040 ^{ns}	0.038 ^{ns}	-0.351*
AS			-0.689**	-0.603**	0.343*	0.221 ^{ns}	0.005 ^{ns}	0.059 ^{ns}	-0.058 ^{ns}	0.350*
SL				0.546**	-0.332*	0.001 ^{ns}	0.043 ^{ns}	0.138 ^{ns}	-0.138 ^{ns}	-0.319*
RL					-0.174 ^{ns}	-0.305*	0.042 ^{ns}	0.043 ^{ns}	-0.040 ^{ns}	-0.183 ^{ns}
Area						-0.298*	-0.032 ^{ns}	-0.069 ^{ns}	0.069 ^{ns}	0.961**
RD							-0.112 ^{ns}	-0.060 ^{ns}	0.060 ^{ns}	-0.164 ^{ns}
Circ								-0.186 ^{ns}	0.186 ^{ns}	-0.088 ^{ns}
FS									-1.000**	-0.096 ^{ns}
MS										0.104 ^{ns}

GSI = germination speed index; G = germination percentage; AS = percentage of abnormal seedlings; SL = shoot length; RL = root length; RD = Roundness (cm); Circ = Circumference (cm); FS = percentage of flat seeds; MS = percentage of mocha seeds; TSW = thousand-seed weight. The symbols * and ** represent significance at 5% and 1% probability, by the F test, respectively. ^{ns} = Not significant by the F test

**Fig. 2:** Dendrogram obtained by the Mahalanobis distance (D^2) (unweighted pair group method with arithmetic mean clustering method) for eight physical and physiological traits of seeds of 43 *C. canephora* genotypes. Cophenetic correlation coefficient = 0.95

identified by groups formed in the dendrogram. Some genotypes from group I are located in the municipality of Marilândia, followed by Vila Valério and Nova Venécia, and the others in Jaguaré, Aracruz, São Mateus, Rio Bananal, Sooretama, Linhares, Itabela-BA and Teixeira de Freitas-BA. For the other groups, the genotypes are located in the municipalities of São Mateus, Nova Venécia and Aracruz and Pinheiros, located in the north of Espírito Santo.

In the study of genetic divergence, the relative importance of the eight variables showed that the character that most contributed to dissimilarity was seed circumference (61.3%), a seed geometric character. The character with the lowest contribution was germination (0.61%) (Fig. 4).

Discussion

The results, both the means of the genotypes and the genetic param, as well as the high values identified for σ_g^2 , indicate the presence of genetic variability between the *C. canephora* genotypes in the present study. The analysis of the breakdown of variance components identified σ_g^2 greater than σ_e^2 of all the variables studied, shows that the existing variance is the result of the genetic difference between the genotypes, allowing us to infer that the selection of superior genotypes for seed traits can be maximized.

Analyzing the same genotypes as in the present study,

Sousa *et al.* (2022) also identified the presence of genetic variability between genotypes based on data from microsatellite markers. For seed traits in *Passiflora* genotypes, Rodrigues *et al.* (2023a) identified σ_g^2 higher than the environmental one for all the characters.

The high values of genotypic variance resulted in high magnitude heritability estimates, especially for TSW and circumference, which allowed us to infer that selection based on seed traits will provide significant genetic gains. Rodrigues *et al.* (2023a) obtained heritability estimates ranging from 92.20 to 99.70% for seed traits in *Passiflora* genotypes.

All variables reached variation indices ($\hat{I}_v$) greater than 1.0 (Table 1). $\hat{I}_v$ greater than 1.0, with higher values indicating greater genetic than environmental influence on the characters. Together with heritability, this index is used as an indicator to facilitate the selection process in breeding (Rodrigues *et al.* 2012). High heritability values suggest great possibilities of genetic gain, and $\hat{I}_v$ indicates the chance of success in the selection of superior genotypes, in addition to facilitating the choice of the most appropriate method, using simpler breeding methods (Rodrigues *et al.* 2012). All the characters studied, in the present study, presented high values $\hat{I}_v$, i.e., selection for these characters is under the most favorable conditions in terms of immediate genetic gains.

The values obtained for the genetic param show a favorable situation for selection regarding the germination potential of *C. canephora* genotypes. The results suggest that selection will be effective, especially when considering both physiological variables and seed geometry characteristics.

In the context of maximizing genetic gains and reducing the number of variables analyzed, breeders make use of the knowledge of correlations between variables, including phenotypic correlations, that enable the evaluating the association between different characters. This information, coupled with heritability estimates for each variable, ensures faster progress in breeding.

In the present study, the negative correlation between germination and abnormal seedlings demonstrated that the values are inversely proportional (Table 2). Germination percentage has 99.22% heritability; thus, better responses was obtained by estimating this character (Table 1). In addition, there was also a moderate positive correlation between germination and shoot length, indicating that it is possible to produce genotypes with a greater number of normal seedlings and greater shoot development.

The perfect positive correlation between seed area and TSW demonstrated that larger the seed, the greater is its weight (Table 2). However, in this study, these variables did not influence seed physiological quality, since the genotypes with the largest seed area and TSW did not stand out with the highest percentage of normal seedlings. Nonetheless, genotype Graudão HP reached the highest TSW and had a germination above 95.0% (Table 2). Therefore, the highlighted variables do not influence seed vigor or germination. Phenotypic correlation studies are of great

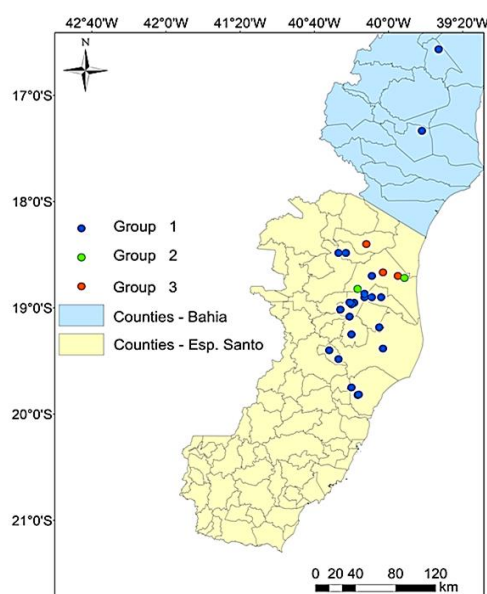


Fig. 3: Collection sites of the 43 *C. canephora* genotypes evaluated and the groups formed by seed traits

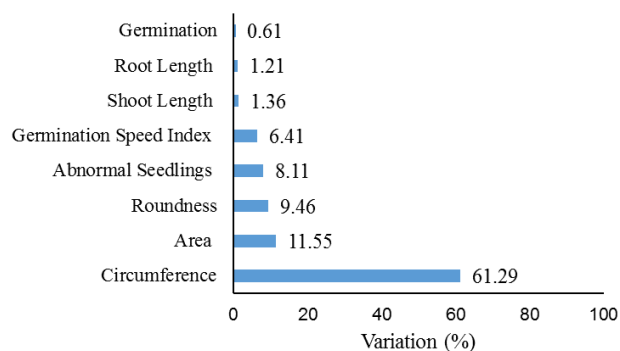


Fig. 4: Percentage of the relative contribution of eight physical and physiological seed traits to the dissimilarity of 43 *C. canephora* genotypes

relevance for plant improvement of *Coffea* spp., allowing, with some caution, to make use of indirect selection to obtain more productive genotypes with high vigor seeds.

Seeking understanding variability in selected genotypes *C. canephora*, the study of dissimilarity, based on seed traits, classified the genotypes into three groups. The high cophenetic correlation coefficient (0.95) indicates a very good fit between the dissimilarity matrix and the plot (Fig. 2). Sousa *et al.* (2022), while studying the same genotypes as in the present study, identified the formation of six groups by cluster analysis based on data from microsatellite markers. The dendrogram obtained by Sousa *et al.* (2022) did not group genotypes in the same way as the dendrogram based on seed characters. The exceptions were Emcapa 143 and Tardio V genotypes and the Verdum D and Z36 genotypes, which remained together. Such results reinforce the need to

consider the evaluation of seed characters in a *C. canephora* breeding program, since analysis based on genetic markers does not necessarily group the genotypes that are most similar in terms of seed vigor.

As seen in Fig. 2, each group had a defining character. Group I achieved higher GSI, germination and root length. Group II stood out for the geometric characters of seed area and roundness. Genotypes Alecrim, AP, Bamburral, Z39 and Tardio C, belonging to group III, were the most genetically distant from the others. These genotypes achieved greater shoot length and seed circumference, which is a character of greater contribution to diversity (Fig. 4). According to Valeriano *et al.* (2019), individual variations in seed biometrics may occur within the same species depending on environmental conditions and genetic characteristics during its development. As stated above, studies investigating seed morphology contribute to the knowledge of species, allowing for correlations to be made between physical and physiological characters. These correlations, in turn, can be used as indicators for studies on vigor, drying and storage, ultimately contributing to the improvement of the species.

Added to the study of genetic divergence, knowing of the contribution of variables to the dissimilarity of genotypes is fundamental step in the identification of superior genotypes. Based on the method of Singh (1981), due to its relative importance, the character that most contributed to divergence between the genotypes was seed circumference, and the character with the smallest contribution was germination (Fig. 4). Fachi *et al.* (2019) described similar results, in which the geometry variables contributed to quantifying the genetic diversity between full-sib families (*Passiflora edulis*). Despite low contribution to germination, its discard is not suggested as it has important contributions to the discrimination of genetic divergence in the population. As identified in the present study, Rodrigues *et al.* (2023b), working with seed characters in *Passiflora*, also identified the characters shoot length, root length, GSI as those that contribute least to the genetic diversion among the genotypes, and the characters circumference as one of those that most contributes to dissimilarity.

By establishing correlation between the dissimilarity grouping and the origin of the genotypes studied, it is interesting to observe a dispersion pattern. The genotypes of groups 2 and 3, close in the dendrogram, have the same origin, the northern region of Espírito Santo. While for group 1, which is larger, the genotypes are spread across the states of Espírito Santo and Bahia (Fig. 3). However, it is noteworthy that all genotypes were collected in the regions (Fig. 3) and replanted in the Nova Venécia region, where the fruits were harvested for the present study. The first *C. canephora* crops introduced in the state of Espírito Santo were established in the municipalities of Cariacica, Cachoeiro de Itapemirim and Itabapoana and later taken to the northern region of the state (Partelli 2016). Initially, there was a greater concentration of crops in the southern region, which later expanded to the northern region

of the state (Ferrão *et al.* 2017). Thus, it can be inferred that the genotypes, especially those from the south of Bahia, were taken from Espírito Santo.

The results obtained and the discussions highlighted in the present work support the importance of taking into account the germination potential of seeds when evaluating genetic variability and in the identification of the best coffee genotypes. It is important to emphasize that the good germination performance of the seeds provides good initial performance in the field, resulting in dominant plants with high agronomic performance. Therefore, in *C. canephora* breeding program must select genotypes considering not only agronomic characteristics related to the plant and the fruit, but also characters related to seed vigor, such as seed weight and circumference, germination speed index, weight seedling dryness.

Conclusion

There is high genetic divergence between the genotypes of *C. canephora*, and the phenotypic variation identified is the result of the genetic difference between them. The variables achieved heritability values greater than 96.0% and variation index above unity, indicating that selection for all studied characters shows favorable conditions in terms of immediate gain. Three groups were formed by the UPGMA hierarchical method, and the genotypes Alecrim, Bamburral, Z39, AP and, Tardio C are the most divergent. The seed circumference contributed the most to clarifying the dissimilarity between genotypes.

Acknowledgements

The authors thank the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro for financial support to carry out the project.

Author Contributions

LZC, HDV, APV and FLP contributed to the study's conception and planned the experiments, HDV, APB and FLP resources; supervision, project administration; funding acquisition, LZC conducted the experiments, LZC, EAZ, RS and AJA analyzed the data and interpreted the results, LZC and RS made the write up and made illustrations, RS carried out the writing-review & editing.

Conflicts of Interest

All authors declare no conflict of interest.

Data Availability

The data shown in this research will be available upon request to the corresponding author.

Ethics Approval

Not applicable to this research

References

- Alvares CA, JL Stape, PC Sentelhas, JL Gonçalves, G Sparovek (2013). Köppen's climate classification map for Brazil. *Meteorol Z* 22:711–728
- Bragança SM, CHSD Carvalho, AFAD Fonseca, RG Ferrão (2001). Variedades clonais de café Conilon para o Estado do Espírito Santo. *Pesq Agropec Bras* 36:765–770
- Brasil (2009). *Regras Para Análise de Sementes*. Ministério da Agricultura, Pecuária e Abastecimento, Brasília, Brasil
- Cruz CD (2016). Genes Software – extended and integrated with the R, Matlab and Selegen. *Acta Sci Agron* 38:547–552
- Cruz CD, PCS Cameiro, AJ Regazzi (2014). *Modelos Biométricos Aplicados ao Melhoramento Genético*, 6th edn. UFV, Viçosa, Brasil
- Dubberstein D, FL Partelli, JHS Guilhen, WP Rodrigues, JC Ramalho, AI Ribeiro-Barros (2020). Biometric traits as a tool for the identification and breeding of *Coffea canephora* genotypes. *Genet Mol Res* 19:gmr18541
- Fachi LR, W Krause, HD Vieira, DV Araújo, PB Luz, AP Viana (2019). Digital image analysis to quantify genetic divergence in passion fruit (*Passiflora edulis*) seeds. *Gen Mol Res* 18:gmr16039955
- Ferrão RG, AFA Fonseca, MAG Ferrão, LH Muner (2017). *Café Conilon*, 2th edn. Incaper, Espírito Santo, Brasil
- Fonseca AFAD, T Sediya, CD Cruz, NS Sakaiyama, MAG Ferrão, RG Ferrão, SM Bragança (2006). Divergência genética em café Conilon. *Pesq Agropec Bras* 41:599–605
- Giles JAD, FL Partelli, A Ferreira, JP Rodrigues, G Oliosi, FH Silva (2018). Genetic diversity of promising 'conilon' coffee clones based on morpho-agronomic variables. *Anais da Acad Bras de Ciências* 90:2437–2446
- Maguire JD (1962). Speed of germination aid in selection and evaluation for seedling emergence and vigor. *Crop Sci* 2:176–177
- Martins MQ, FL Partelli, A Ferreira, CdeO Bernardes, A Golynski, HD Vieira, MSM Freitas, JC Ramalho (2019). Genetic variability on nutrient contents in *Coffea canephora* genotypes cultivated at 850 m of altitude in two crop seasons. *Funct Plant Breed J* 1:1–12
- Partelli FL (2016). *História de VILA VALÉRIO: Colonização, Desenvolvimento e Café conilon*. Vila Valério, Espírito Santo, Brasil
- Partelli FL, A Golynski, A Ferreira, MQ Martins, AL Mauri, JC Ramalho, HD Vieira (2019). Andina-first clonal cultivar of high-altitude conilon coffee. *Crop Breed Appl Biotechnol* 19:476–480
- Partelli FL, JAD Giles, G Oliosi, AM Covre, A Ferreira, VM Rodrigues (2020a). Tributun: A coffee cultivar developed in partnership with farmers. *Crop Breed Appl Biotechnol* 20:e30002025
- Partelli FL, AM Covre, G Oliosi, DT Covre (2020b). Monte Pascoal: Primeira Cultivar para Bahia. In: *Café conilon: Desafios e oportunidades*, pp:15–22. Partelli FL, A Campanharo (Eds.). CAUFES, Espírito Santo, Brasil
- Partelli FL, G Oliosi, JR Dalazen, CA Silva, HD Vieira, MC Espindula (2021). Proportion of ripe fruit weight and volume to green coffee: Differences in 43 genotypes of *Coffea canephora*. *Agron J* 113:1050–1057
- Rodrigues CA, HD Vieira, R Souza, DS Mendes, AP Viana (2023a). Genetic gain in *Passiflora* seed traits from recurrent selection among full-sib families. *Crop Breed Appl Biotechnol* 23:e44302319
- Rodrigues CA, HD Vieira, R Souza, DS Mendes, AP Viana (2023b). Genetic divergence for seed traits in *Passiflora* full-sib families under recurrent selection for resistance to the cowpea aphid-borne mosaic virus. *Bragantia* 82:e20230063
- Rodrigues WN, MA Tomaz, RG Ferrão, MAG Ferrão, AFAD Fonseca, FDD Miranda (2012). Estimativa de parâmetros genéticos de grupos de clones de café conilon. *Coffee Sci* 7:177–186
- Santos HG, PKT Jacomine, LHC Anjos, VA Oliveira, JF Lumberreras, MR Coelho, JA Almeida, JC Araujo Filho, JB Oliveira, T Cunha (2018). *Sistema Brasileiro de Classificação de Solos*, 5th edn. Embrapa, Brasília, Brasil
- Schmidt R, LOE Silva, A Ferreira, I Gontijo, RJ Guimarães, JC Ramalho, FL Partelli (2022). Variability of root system size and distribution among *Coffea canephora* genotypes. *Agronomy* 12:647
- Scott AJ, M Knott (1974). Cluster analysis method for grouping means in the analysis of variance. *Biometrics* 30:507–512
- Singh D (1981). The relative importance of characters affecting genetic divergence. *Ind J Genet Plant Breed* 41:237–245
- Sousa PGF, HD Vieira, EA Santos, AP Viana, MSB Boecheat, FL Partelli (2022). *Coffea canephora*: Heterotic crosses indicated by molecular approach. *Plants* 11:3023
- Valeriano FR, MC Nery, NAVD Pinto, AR Melo Campos, AS Oliveira, CMT Fialho (2019). Morfologia de sementes de gergelim. *Acta Iguazu* 8:23–36
- Vareltzis P, I Gargali, S Kiroglou, M Zeleskidou (2020). Production of instant coffee from cold brewed coffee process characteristics and optimization. *Food Sci Appl Biotechnol* 3:39–46
- Vázquez N, H López-Fernández, CP Vieira, F Fdez-Riverola, J Vieira, M Reboiro-Jato (2019). BDBM 10: A desktop application for efficient retrieval and processing of high-quality sequence data and application to the identification of the putative *Coffea* s-locus. *Interdiscip Sci Comput Life Sci* 11:57–67