

Genetic diversity among coffee tree progenies Big Coffee VL based on growth traits and production

J.A. Silva¹, S.P. Carvalho², A.T. Bruzi², R.J. Guimarães², L.L. Oliveira¹ and L.C. Simões¹

¹Universidade Federal de Lavras, Lavras, MG, Brasil ²Departamento de Agricultura, Universidade Federal de Lavras, Lavras, MG, Brasil

Corresponding author: J.A. Silva E-mail: josimaraleixo_@hotmail.com

Genet. Mol. Res. 15 (4): gmr15049066 Received August 9, 2016 Accepted September 20, 2016 Published November 21, 2016 DOI http://dx.doi.org/10.4238/gmr15049066

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. In a coffee plantation of a coffee 'Acaiá' cultivar (*Coffea arabica*), on the Midwest of Minas Gerais in Capitólio city, a different kind of coffee tree was found (1989), possibly due to a mutation. It presented larger leaves and grains than those of conventional coffee trees and was named as "Big Coffee VL." The aim of this study was to estimate the genetic diversity of Big Coffee VL progenies cultivated at Universidade Federal de Lavras, by evaluating growth and production traits, based on genetic distances and clusters. The experiment was established in a lattice design with 100 progenies of this coffee tree and 23 repetitions. Traits evaluated were vigor, plant height, stem diameter, node number of plagiotropic branches, pair numbers of plagiotropic branches, and productivity. Genetic divergence was evaluated by multivariate procedures: Mahalanobis generalized distance, clustering

Genetics and Molecular Research 15 (4): gmr15049066

methods, and principal component analysis. Genetic distances were estimated using Mahalanobis distance and presented variations from 0.04 to 18.70. The most similar progenies were P23 and P29 and the most dissimilar progenies were G8 and P14. The progenies were divided into three groups, with P14 present as an isolated group. Thus, it was possible to observe the existence of genetic variability among the progenies of Big Coffee VL, which can be used in breeding programs to increase grain size. Progenies G8 and P14 presented the highest genetic distance, and were the most suitable for future integration of crossings in plant breeding programs.

Key words: *Coffea arabica*; Genetic divergence; Multivariate analysis; Plant breeding

INTRODUCTION

Coffee is one of the most widely consumed beverages in the world. Commercial production of coffee beans of two species, *Coffea arabica* L. and *Coffea canephora* Pierre ex Froehner, represents about 60 and 40% of the total coffee market, respectively. Brazil is the largest producer and exporter of coffee, contributing approximately 30% to the total production (ICO - International Coffee Organization, 2016). Coffees with desirable characters, such as large grains, are requirements needed for specific market niches. Thus, it is important to generate researches in genetic improvement for increased grain size, as genetic diversity study, which is the basis for the improvement.

Genetic variations among individuals of same or different species lead to genetic diversity, which must be ensured for sustainability of breeding programs, and continued availability of agronomically superior cultivars. Genetic diversity has been the subject of several studies in cultivated species, such as triticale (Niedziela et al., 2016), eucalyptus (Moraes et al., 2015), soybean (Ferreira Júnior et al., 2015), and coffee (Guedes et al., 2013; Silva et al., 2015; Rodrigues et al., 2015).

In coffee crops (*C. arabica*), several factors, such as predominance of self-pollination and propagation from seeds, contribute to a low genetic variability. Another factor that contributes to this is the pedigree method for conduction of segregated populations, which is used to obtain the commercial cultivars (Gaspari-Pezzopane et al., 2004; Carvalho and Krug, 2010). In this method, individuals are selected in the segregated generations, maintaining the relationship data between parents and offspring (Bueno et al., 2006).

However, broadening of the genetic variability of *C. arabica* can be achieved through mutations (Krug, 1949), introductions (Sakiyama et al., 2005), and hybridizations (Krug and Carvalho, 1952), thereby, increasing the gains from plant breeding programs for the release of new cultivars.

The majority of cultivars of *C. arabica* used in Brazil are genetically related, having originated from ancestral cultivars, such as Typica, Bourbon, and Sumatra, which have undergone spontaneous mutations and natural and artificial crossings (Mendes and Guimarães, 1998; Aguiar et al., 2004; Vieira et al., 2010).

Usage of genetic variability is an important strategy for the sustainability of breeding programs. Hence, in 1989, on a coffee plantation of Acaiá cultivar (*C. arabica*) in the Midwest

Genetics and Molecular Research 15 (4): gmr15049066

of Minas Gerais (Capitólio), a coffee tree different from others was found, possibly due to a mutation; this tree presented larger leaves and grains than those of conventional coffee plants, and was named "Big Coffee VL". Subsequently, these coffee progenies were cultivated in Piumhí, Minas Gerais. Seeds collected from these progenies were used in this study, and have been maintained by the Department of Agriculture of Universidade Federal de Lavras (UFLA). These progenies have a great potential for use in breeding programs. According to Ferreira et al. (2005), increasing the grain size is one of the aims in coffee breeding.

Obtaining coffee beans with high sieves gives an added value to the product and such beans are preferred by some importing countries. In a cooperative case study of coffee, Oliveira (1997) suggested that a standard coffee for export must be classified by its grain size, besides the quality of the beverage prepared from it.

In this context, the objective of this study was to estimate the genetic diversity of Big Coffee VL progenies cultivated in UFLA, starting from the diversity in growth and production characters, based on genetic distances.

MATERIAL AND METHODS

The experiment was conducted in the Agriculture Department of UFLA, Lavras, Minas Gerais, Brazil, located at 910 m in altitude, and lying on 21°14"06'S and 45°00"00'W.

This study was initiated in 2012, with 100 progenies of Big Coffee VL procured from Piumhí, Minas Gerais. Of these, 32 progenies were classified as "Grande" (G5, G6, ..., G36), 36 were classified as "Médio" (M1, M2, ..., M36), and 32 were classified as "Pequeno" (P5, P6,..., P36). This classification was based on the grain size. Planting was done in February, 2016, and culturing practices recommended for the culture of coffee were followed (Ribeiro et al., 1999). The space adopted was $3.5 \ge 0.9$ m. The design was a $10 \ge 10$ lattice with 23 repetitions, with a total of 2300 plots, in which each plant composed a plot.

The growth characters evaluated were: vigor (VIG), with notes from 1 to 5, where the most vigorous plants were note 5; stem diameter (SD), measured using a precision caliper (mm) at the base of plant's stem; plant height (PH) measured with the aid of a graduated ruler (m); pair number of plagiotropic branches (PNP); node number of plagiotropic branches, counted taking a branch from the third part of the plant to perform the count. Grain productivity (PROD) was measured using a precision scale during the harvest time (kg/plant). The evaluations were conducted annually. This study used the data for growth traits and production obtained from the second harvest in 2015.

The above-mentioned variables were analyzed by analysis of variance to evaluate the existence of genetic variability among the progenies. Thereafter, multivariate analysis was carried out for the genetic diversity study among the progenies through genetic distances. Genetic distances were estimated based on Mahalanobis distance and Euclidean distance was determined by principal components analysis as described by Cruz et al. (2014).

For a better understanding, distance clustering was performed using a hierarchical method, unweighted pair-group method using arithmetic averages (UPGMA) or the average linkage, which does not allow characterization of the dissimilarity for extreme values between individuals (Cruz et al., 2014). The cutoff point of the dendrogram was determined by the method described by Mojena (1977), which allowed the determination of the number of groups used for clustering by principal components.

Therefore, to evaluate the dendrogram clustering consistency, cophenetic correlation

Genetics and Molecular Research 15 (4): gmr15049066

J.A. Silva et al.

was carried out. The higher the value of cophenetic correlation coefficient (CCC), the lower would be the distortion caused by the grouping individuals in a dendrogram (Cruz et al., 2014).

Statistical analysis for the estimation of genetic distances and for the clustering of Big Coffee VL progenies was performed using the programs, Genes (Cruz, 2013) and R (R DEVELOPMENT CORE Team, 2015).

RESULTS

Measures of genetic distances, estimated by Mahalanobis distance between the pairs of coffee plant progenies of Big Coffee VL (<u>Table S1</u>), had magnitudes ranging from 0.04 to 18.70. Progenies P23 and P29 were the most similar and progenies G8 and P14 were the most dissimilar. Progeny P14 was involved in the bigger genetic dissimilarities that were observed. It is important to highlight that the progenies used in this study represent the third generation of the original plant found in Capitólio (Minas Gerais), a fact that should be considered for determining the variation of genetic distances.

These progenies were classified in "Grande" (G), "Médio" (M), and "Pequeno" (P). For progenies "G", the magnitude of the genetic distances were from 0.07 to 3.30, respectively for pairs G23 and G30, as well for G12 and G30. The variation for the class "M" was from 0.06 to 3.22 for the pairs M16 and M18, as well as for M34 and M5. Lastly, for the class "P," such variation was from 0.04 to 10.6 for pairs P23 and P29, respectively, as well as for P14 and P26.

The hierarchical grouping by UPGMA method used the measure of genetic dissimilarity by Mahalanobis distance (Figure 1), which allowed the formation of three different groups of progenies, considering the cutoff point of Mojena method at 20% of the maximum fusion level, and a CCC of 0.80.

The first (I) group was composed of progenies P23, P29, M17, M4, G17, G9, M30, M5, G14, P9, P22, P34, P16, P11, P25, P5, P31, P6, P8, P7, P20, P12, M14, G10, G12, M9, M10, G35, M29, M27, P13, M22, M8, G18, M3, M28, P26, M20, P27, G16, M2, M7, G36, G33, P24, P10, P18, P36, P28, P15, P32, P35, P17, P33, P30, and P19. The second (II) group was composed of progenies M36, G25, M1, G27, M21, G30, G23, G20, M31, M34, G8, M19, G29, M11, G31, M18, M16, M35, M26, M6, G34, G26, G7, G32, M32, G5, M23, G6, G19, G24, G28, M12, M33, M24, G21, M25, G11, P21, G15, G13, M15, M13, and G22. The third (III) group was composed only of P14 progeny.



Figure 1. Hierarchical grouping of 100 Big Coffee VL progenies using the UPGMA method based on the matrix of Mahalanobis distance. The cutoff point of the Mojena method was considered at 20% of the maximum fusion level.

Genetics and Molecular Research 15 (4): gmr15049066

The principal component analysis (Table 1) showed that the first two principal components resulted in 73.04% of the contained variance in the progenies. The principal component 1 (PC1) and principal component 2 (PC2) contributed 45.43 and 27.61% to this variance, respectively. In this context, Teixeira et al. (2013) studied the principal components in the morphological characters of *C. arabica* and found a contribution of 63.8 and 16.2% for PC1 and PC2 components, respectively.

Table 1. Summary of the principal component (PC) analysis with their eigenvectors and eigenvalues, from growth traits and production of 100 Big Coffee VL progenies.

Components	Eigenvectors					Eigenvalues			
	VIG	PH	SD	NNP	PNP	PROD	Eigenvalue	% Variance	% Accumulated
PC1	0.48	0.42	0.45	0.43	0.30	0.33	2.73	45.43	45.43
PC2	0.01	-0.47	-0.41	0.07	0.58	0.52	1.66	27.61	73.05
PC3	-0.48	-0.19	0.05	0.84	-0.17	-0.05	0.63	10.53	83.58
PC4	-0.43	-0.05	0.48	-0.03	-0.25	0.65	0.48	8.04	91.62
PC5	0.58	-0.47	-0.05	0.13	-0.62	0.02	0.36	6.04	97.66
PC6	0.07	-0.59	0.63	-0.09	0.32	-0.04	0.14	2.33	100.00

VIG = vigor; PH = plant height; SD = stem diameter; NNP = node numbers of plagiotropic branches; PNP = pair numbers of plagiotropic branches; PROD = productivity.

Moreover, the contribution of each variable to the components, namely, the set of associated eigenvector is shown in Table 1. As seen for PC1, the highest values, whether positive or negative, for VIG and SD, were 0.48 and 0.45, respectively. In PC2, higher values were consistently obtained for PNP (0.58), PROD (0.52), and PH (-0.47).

The distribution of 100 progenies of Big Coffee VL determined through the graphic dispersion (Figure 2) was in concordance with the hierarchical UPGMA method, based on Euclidean distance, estimated from the PCs and corroborated with the dendrogram results. Despite the formation of groups with different progenies, the most dissimilar progenies (P14 and G8) remained in different groups, even considering the most dissimilar progenies in the class "P", "M", and "G."



Figure 2. Graphic dispersion of the first and second principal components (PC1 and PC2) derived from growth traits and production, and grouping of 100 Big Coffee VL progenies according to the hierarchical method, UPGMA, based on Euclidean distance, and estimated from the Principal Components Analysis.

Genetics and Molecular Research 15 (4): gmr15049066

J.A. Silva et al.

DISCUSSION

Estimation of genetic Mahalanobis distance can be found in studies such as that of Fonseca et al. (2006), who found magnitudes ranging from 0.67 to 87.74 in conilon coffee clones. In this context, Guedes et al. (2013) found magnitudes ranging from 0.52 to 84.90 in coffee germplasm of Maragogipe. These authors pointed out the wide genetic diversity observed in these studies, a fact that should be considered for the progenies of Big Coffee VL.

The crossings of the most dissimilar individuals are of great interest in plant breeding programs; however, this does not guarantee the heterosis, because the superior performance also depends on the allelic complementarity. Moreover, in the parental selection, the selection of superior individuals in subsequent generations, especially their performance and allelic complementarity, should be taken in account (Dias and Kageyama, 1997; Souza et al., 2005; Guedes et al., 2013.).

The CCC, a measure that evaluates the consistency of groups, was found to be 0.80 (Cruz et al., 2014). A similar value for this coefficient was also found by Moura et al. (2015) in a group of 20 *C. arabica* cultivars with a CCC of 0.85.

It was observed that the most dissimilar progenies, G8 and P14, were situated in different groups indicating a greater genetic divergence between them. In other genetic diversity studies on coffee trees, the formation of different groups was also observed using the UPGMA method (Ivoglo et al., 2008; Guedes et al., 2013; Moura et al., 2015).

It is important to highlight that among the evaluated traits, the grain size was not determined; therefore, there was no tendency to previously classify the progenies in the same group as "P", "M", and "G." In addition, the most dissimilar progenies, considering such a classification, were divided into different groups, such as G12 and G30, M34 and M5, and P14 and P26.

Multivariate analysis, such as principal component analysis, is interesting due to the possibility of discarding the characteristics that contribute less to genotype discrimination (Pereira et al., 2010). The characters that contribute more to the components are those that have higher weightings in the major eigenvalue components, and those that contribute less are those showing greater eigenvectors in less eigenvalue components (Ferreira, 2011; Cruz et al., 2014).

The candidate characters discarded in this study were SD and PH (Table 1), where these were the main contributors to the PC having lower eigenvalues. In this context, Teixeira et al. (2013) selected the characters that contributed most to the discrimination of coffee genotypes and pointed out that smaller the number of variables evaluated, more efficient was the plant breeding program, thereby, maximizing the gains and reducing the labor costs.

The graphics associated with the groupings are of great importance in breeding programs aimed to obtain heterosis, since they assist as an indicator of the parent in the groups or different subgroups to be included in crossings (Fonseca et al., 2006).

Despite the genetic diversity found among the progenies of Big Coffee VL, generally, only the superior progenies are of greater interest to breeders in association with the most important characteristics and show sufficient divergence to produce variability in the segregating populations. The breeding programs for coffee seek an ideotype, which its performance covers; among other traits, these include high production and increased grain size, with the last one being an inherent characteristic of Maragogipe group (Ferreira et al., 2005; Guedes et al. 2013), and a main feature of Big Coffee VL.

Genetics and Molecular Research 15 (4): gmr15049066

CONCLUSIONS

In conclusion, it was possible to observe the existence of genetic variability among the progenies of Big Coffee VL, which can be utilized in breeding programs to increase the grain size. The coffee tree progenies of Big Coffee VL were more genetically distant and may be promising in future artificial crossings. Progenies G8 and P14 had the highest genetic distance. Moreover, the characters stem diameter and plant height were the candidates that were determined to be discarded from the subsequent evaluations.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG), and Núcleo de Estudos em Melhoramento e Clonagem (NEMEC/UFLA).

REFERENCES

- Aguiar ATE, Guerreiro Filho O, Maluf MP, Gallo PB, et al. (2004). Caracterização de cultivares de *Coffea arábica* mediante utilização de descritores mínimos. *Bragantia* 63: 179-192. <u>http://dx.doi.org/10.1590/S0006-87052004000200003</u>
- Bueno LCS, Mendes ANG and Carvalho SP (2006). Melhoramento Genético de Plantas: Princípios e Procedimentos. UFLA, Lavras.
- Carvalho A and Krug CA (2010). Biologia da flor do cafeeiro Coffea arabica L. Cienc. Cult. 62: 35-38.
- Cruz CD (2013). GENES a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276. <u>http://dx.doi.org/10.4025/actasciagron.v35i3.21251</u>
- Cruz CD, Carneiro PCS and Regazzi AJ (2014). Modelos biométricos aplicados ao melhoramento genético. v.2, 3rd ed. UFV, Viçosa.
- Dias LAS and Kageyama PY (1997). Multivariate genetic divergence and hybrid performance of cacao (*Theobroma cacao* L.). *Braz. J. Genet* 20: 63-70. http://dx.doi.org/10.1590/S0100-84551997000100012
- Ferreira A, Cecon PR, Cruz CD, Ferrão RG, et al. (2005). Seleção simultânea de Coffea canephora por meio da combinação de análise de fatores e índices de seleção. *Pesqui. Agropecu. Bras.* 40: 1189-1195. <u>http://dx.doi.org/10.1590/S0100-204X2005001200005</u>
- Ferreira DF (2011). Estatística Multivariada. 2nd edn. UFLA, Lavras.
- Ferreira Júnior JA, Unêda-Trevisoli SH, Espíndola SMCG, Vianna VF, et al. (2015). Diversidade genética em linhagens avançadas de soja oriundas de cruzamentos biparentais, quádruplos e óctuplos. *Rev. Cienc. Agron* 46: 339-351.
- Fonseca AFA, Sediyama T, Cruz CD, Sakaiyama NS, et al. (2006). Divergência genética em café conilon. *Pesqui.* Agropecu. Bras. 41: 599-605. http://dx.doi.org/10.1590/S0100-204X2006000400008
- Gaspari-Pezzopane C, Medina Filho HP and Bordignon R (2004). Variabilidade genética do rendimento intrínseco de grãos em germoplasma de Coffea. Bragantia 63: 39-54. http://dx.doi.org/10.1590/S0006-87052004000100005
- Guedes JM, Vilela DJM, Rezende JC, Silva FL, et al. (2013). Divergência genética entre cafeeiros do germoplasma Maragogipe. Bragantia 72: 127-132. <u>http://dx.doi.org/10.1590/S0006-87052013000200003</u>
- ICO International Coffee Organization (2016). Total production 2015. Available at: [http://www.ico.org/trade_statistics.asp]. Ivoglo MG, Fazuoli LC, de Oliveira ACB, Gallo PB, et al. (2008). Divergência genética entre progênies de café robusta.
- Bragantia 67: 823-831. http://dx.doi.org/10.1590/S0006-87052008000400003
- Krug CA (1949). Mutações em Coffea arabica L. Bragantia 9: 1-10.
- Krug CA and Carvalho A (1952). Melhoramento do caffeiro V: Melhoramento por hibridação. *Bragantia* 12: 141-152. http://dx.doi.org/10.1590/S0006-87051952000200003

Mendes ANG and Guimarães RJ (1998). Genética e Melhoramento do Cafeeiro. UFLA/FAEPE, Lavras.

Genetics and Molecular Research 15 (4): gmr15049066

J.A. Silva et al.

- Mojena R (1977). Hierarchical grouping methods and stopping rules: an evaluation. *Comput. J.* 20: 359-363. <u>http://dx.doi.org/10.1093/comjnl/20.4.359</u>
- Moraes CB, Carvalho EV, Zimback L, Luz OSL, et al. (2015). Variabilidade genética em progênies de meios-irmãos de eucaliptos para tolerância ao frio. *Rev. Arvore* 39: 1047-1054. <u>http://dx.doi.org/10.1590/0100-67622015000600007</u>
- Moura WM, Soares YJB, Amaral Júnior AT, Lima PC, et al. (2015). Genetic diversity in arabica coffee grown in potassiumconstrained environment. *Cienc. Agrotec.* 39: 23-31. <u>http://dx.doi.org/10.1590/S1413-70542015000100003</u>
- Niedziela A, Orłowska R, Machczyńska J and Bednarek PT (2016). The genetic diversity of triticale genotypes involved in Polish breeding programs. *Springerplus* 5: 355. <u>http://dx.doi.org/10.1186/s40064-016-1997-8</u>
- Oliveira LH (1997). S.I.G. para a gestão de custos por qualidade: estudo de caso em uma cooperativa de café. *Rev. Adm. Contemp.* 1: 97-119. http://dx.doi.org/10.1590/S1415-65551997000300006
- Pereira MC, Chalfoun SM, Carvalho GR and Savian TV (2010). Multivariate analysis of sensory characteristics of coffee grains (*Coffea arabica L.*) in the region of upper Paranaiba. *Acta Sci. Agron.* 32: 635-641.
- R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing (2015). Available at: [http://www.R-project.org]. Accessed December 20, 2015.
- Ribeiro AC, Guimarães PTG and Alvarez VVH (1999). Recomendação para o uso de corretivos e fertilizantes em Minas Gerais: 5ª aproximação. Comissão de Fertilidade de Solo do Estado de Minas Gerais, Viçosa.
- Rodrigues WN, Tomaz MA, Ferrão MAG, Ferrão RG, et al. (2015). Diversity among genotypes of conilon coffee selected in Espírito Santo state. *Biosci. J.* 31: 1643-1650. <u>http://dx.doi.org/10.14393/BJ-v31n6a2015-26042</u>
- Sakiyama NS, Pereira AA, Moura WM and Zambolim L (2005). Melhoramento do café arábica. In: Melhoramento de espécies cultivadas (Borém A, ed). UFV, Viçosa, 203-223.
- Silva FL, Baffa DCF, Oliveira ACB, Pereira AA, et al. (2013). Integração de dados quantitativos e multicategóricos na determinação da divergência genética entre acessos de cafeeiro. *Bragantia* 72: 224-229. <u>http://dx.doi.org/10.1590/ brag.2013.039</u>
- Souza FF, Queiroz MA and Dias RCS (2005). Divergência genética entre linhagens de melância. Hortic. Bras. 23: 179-183. http://dx.doi.org/10.1590/S0102-05362005000200003
- Teixeira AL, Gonçalves FMA, Rezende JC, Rocha RB, et al. (2013). Principal component analysis on morphological traits in juvenile stage arabica coffee. *Coffee Sci.* 8: 205-210.
- Vieira ESN, Pinho EV, Carvalho MGG, Esselink DG, et al. (2010). Development of microsatellite markers for identifying Brazilian Coffea arabica varieties. *Genet. Mol. Biol.* 33: 507-514. <u>http://dx.doi.org/10.1590/S1415-47572010005000055</u>

Supplementary material

Table S1. Dissimilarity matrix between 100 Big Coffee VL progenies based on Mahalanobis distance.

Genetics and Molecular Research 15 (4): gmr15049066