PAB
Pesquisa Agropecuária Brasileira

Genetics/ Original Article

ISSN 1678-3921

Journal homepage: www.embrapa.br/pab

For manuscript submission and journal contents, access: www.scielo.br/pab

Genetic variability in conilon coffee related to grain attributes in an irrigated crop in the Cerrado

Abstract – The objective of this work was to quantify the genetic variability of 213 conilon coffee accessions of the Robusta Tropical cultivar, based on chemical characteristics related to the quality of green coffee beans, as well as to identify promising accessions for the breeding program of irrigated conilon coffee in the Brazilian Cerrado. The chemical characteristics evaluated were: protein and caffeine contents, total soluble solids and total lipids, ether extract, pH, and total titratable acidity. The data were subjected to the principal component analysis and cluster analysis based on the similarities observed within the first two principal components using the minimum variance method (Wards) and, as a measure of similarity, the Euclidean distance. The three main components explained 72.64% of the total variation of the data. All characteristics, except pH, were correlated with the first three components. It was possible to separate the genotypes in three clusters, according to the similarities observed in the behavior of the variables. The evaluated accessions present genetic variability regarding the assessed quality characteristics of green coffee beans, and CPAC 160 and CPAC 32 are the most promising for the breeding program of conilon coffee for cultivation under irrigation in the Cerrado.

Index terms: *Coffea canephora*, chemical composition, genetic diversity, multivariate analysis.

Variabilidade genética em café conilon relativa aos atributos dos grãos em cultivo irrigado no Cerrado

Resumo – O objetivo deste trabalho foi quantificar a variabilidade genética de 213 acessos de café conilon, da cultivar Robusta Tropical, quanto às características químicas de qualidade dos grãos de café crus, bem como identificar acessos promissores para o programa de melhoramento do café conilon irrigado no Cerrado. As características químicas avaliadas foram: teores de proteína e cafeína, sólidos solúveis e lipídeos totais, extrato etéreo, pH e acidez titulável total. Os dados foram submetidos à análise de componentes principais e à análise de cluster com base nas semelhanças observadas nos dois primeiros componentes principais, com uso do método de variância mínima (Wards) e da distância Euclidiana, como medida da dissimilaridade. Os três principais componentes explicaram 72,64% da variação total dos dados. Todas as características, exceto o pH, foram correlacionadas com os três primeiros componentes. Foi possível separar os genótipos em três grupos de acordo com as semelhanças no comportamento das variáveis. Os acessos avaliados apresentam variabilidade genética em relação às características de qualidade de grãos avaliadas, e o CPAC 160 e o CPAC 32 são os mais promissores o programa de melhoramento do café conilon irrigado no Cerrado.

Termos para indexação: *Coffea canephora*, composição química, diversidade genética, análise multivariada.

Felipe Augusto Alves Brige[1 ✉], Sonia Maria Costa Celestino[2], Renato Fernando Amabile[2], Marcelo Fagioli[1], Francisco Marcos dos Santos Delvico[2], Ana Paula Leite Montalvão[3] and Pedro Ivo Aquino Leite Sala[1]

[1] Universidade de Brasília, Instituto Central de Ciências, Faculdade de Agronomia e Medicina Veterinária, CEP 70910-970 Brasília, DF, Brazil. E-mail: felipebrige@gmail.com, mfagioli@unb.br, pedroivo.sala@gmail.com

[2] Embrapa Cerrados, Rodovia BR-020, Km 18, CEP 73310-970 Planaltina, DF, Brazil. E-mail: sonia.celestino@embrapa.br, renato.amabile@embrapa.br, francisco.delvico@embrapa.br

[3] Universitat Politecnica de Valencia, Camino de Vera, s/nº, C.P 46022 Valencia, Spain. E-mail: anapaulambrbsb@gmail.com

✉ Corresponding author

Received
December 4, 2017

Accepted
May 13, 2019

How to cite
BRIGE, F.A.A.; CELESTINO, S.M.C.; AMABILE, R.F.; FAGIOLI, M.; DELVICO, F.M. dos S.; MONTALVÃO, A.P.L.; SALA, P.I.A.L. Genetic variability in conilon coffee related to grain attributes in an irrigated crop in the Cerrado. **Pesquisa Agropecuária Brasileira**, v.54, e00358, 2019. DOI: https://doi.org/10.1590/S1678-3921.pab2019.v54.00358.

Introduction

Coffea sp. is among introduced crops in the Cerrado that have been shown a good performance. Besides the suitability to the cultivation of arabica (*Coffea arabica* L.), the Cerrado shows also potential areas for growing conilon coffee (*Coffea canephora* Pierre ex Froehner).

Cerrado's climate allows of the irrigation management for the synchronization of flowering, which leads to a uniform maturation of fruit and harvest with a higher percentage of cherry fruit. In addition, the harvest coincides with the dry season, which avoids unwanted fermentations and results in better coffee quality (Guerra et al., 2005; Bonomo et al., 2008; Fernandes et al., 2012).

Few agricultural products have their price tied to quality; however, coffee beans from regions and genetics known for producing high-quality beverages have great commercial value (Teuber, 2010). Coffee quality can be determined by factors such as the chemical composition of coffee beans, which influences the use of the final product. The formation of chemical compounds, during seed development, is highly affected by environmental and genetic factors (Joët et al., 2010; Scholz et al., 2011; Cheng et al., 2016; Zaidan et al., 2017).

In Brazil, it is believed that conilon coffee shows low quality inherent to the species that is impossible to be improved. But the favorable environment and the technology allied to its great genetic diversity and adaptability (Fonseca et al., 2013) suggest that a selection of higher-quality materials adapted to this environment can be made.

The study of conilon coffee in an irrigated system in the Cerrado is still incipient, and information on its behavior in this environment is still scarce. Therefore, the introduction of cultivars and progenies of *Coffea canephora* in the Cerrado edaphoclimatic environment is an opportunity to increase the knowledge of its genetic variability.

The study of the genetic variability is essential in a breeding program, since crossings involving genetically divergent parents may induce greater heterotic effects and variability in the segregating generations, allowing of the building of populations with a wider genetic base and providing greater genetic gains with the further selection cycles (Cecon, 2008; Cruz et al., 2012).

The principal component analysis (PCA) is a multivariate method used to determine the genetic variability in populations that allows to identify the characters which explain most of the existing variability. PCA permits inferring which characters represent the largest fraction of the found variability, as it enables the removal of those characters showing few contributions to the genetic divergence of the studied population, and reduces the number of traits to be analyzed (Cruz et al., 2012).

The objective of this work was to quantify the genetic variability of 213 conilon coffee accessions of the Robusta Tropical cultivar, based on chemical characteristics related to the quality of green coffee beans, as well as to identify promising accessions for the breeding program of irrigated conilon coffee in the Brazilian Cerrado.

Materials and Methods

The work was carried out at the laboratory of food science technology of Embrapa Cerrados. Green coffee beans of 213 conilon genotypes – of a working collection originated from open natural crosses of an 'Emcaper 8151 – Robusta Tropical' population – were evaluated in an experimental field of Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural do Espírito Santo – Incaper.

The collection was planted in April 2009, Planaltina (15°35'57"S, 47°42'30"W, at 1.030 m altitude), DF, Brazil, in the experimental field of Embrapa Cerrados. at 3.5 m spacing between rows and 1.0 m between plants, in a Latossolo Vermelho-Amarelo distrófico, i.e., Typic Haplustox, under an irrigation system by central pivot.

The irrigation management was based on the Cerrado irrigation monitoring program proposed by Embrapa Cerrados (2009); in order to standardize the flowering, water stress management was used as suggested by Guerra et al. (2005).

To prepare the soil prior to planting, liming, and plastering were performed with 2 Mg ha^{-1} dolomitic limestone, and 2 Mg ha^{-1} agricultural gypsum, respectively. According to the soil analysis, during the planting 58 g P_2O_5 were added by pit, and maintenance fertilization was executed annually with 450 kg ha^{-1} N, 450 kg ha^{-1} K_2O, and 300 kg ha^{-1} P_2O_5.

Coffee genotypes were classified according to their ripening cycle, defined in days between the return of irrigation after water stress, and cherry phase of fruit defined by the following groups based on the scale developed by Pezzopane et al. (2003): very early (243–255 days); early (256–267 days); medium (268–280 days); and semilate (281–293 days).

For the purposes of this study, the harvest was carried out manually during June and July 2014, by collecting from each plant only fruit in the cherry phase, which were immediately subjected to a natural processing separately, in a conventional yard, until they reached 11% humidity. After peeling, green coffee beans were grounded and passed through a 20 mesh sieve, and the samples were stored in closed glass containers covered with aluminum to protected them against light until the analysis was performed. Caffeine content was determined based on the method suggested by Instituto Adolfo Lutz (Zenebon et al., 2008); and protein content was determined by the Kjeldhal method (Wiles et al., 1998). Total soluble solids (TSS) and ether extract (EE), as well as pH and total titratable acidity were determined according to AOAC (Helrich, 1990). These analysis were performed in triplicates, in a completely randomized design.

Protein and caffeine contents and ether extract were predicted by near-infrared spectrometry (NIR), by reflectance in the spectral range between 1,108 and 2,492.8 nm with the aid of ISIscan version 2.85 (Infrasoft International LLC, State College, PA, USA), without replicates. The NIR refers to a nondestructive, extremely simple and quick method, that does not require reagents or dilutions, and that is based on the principle of emission of electromagnetic radiation (Siesler et al., 2008). It is employed to analyze organic components such protein and caffeine content in coffee (Barbin et al., 2014; Scholz et al., 2014), and in many other crops (Jones et al., 2012; Plans et al., 2012; Santos et al., 2012; Pinheiro et al., 2013; Pojić & Mastilović, 2013).

PCA was used to quantify the genetic variability with the standardized data and the dissimilarity distance. In order to assess the significance of a principal component it is necessary to verify its eigenvalue. If this eigenvalue is higher than 1.0, then, in theory, the associated component has inherently more information than would a single isolated variable have. All principal components with eigenvalue greater than 1.0 should be subjected to interpretation. There are reasons to believe that any principal component (PC) is significant if it explains some percentage of the total variability in the dataset (Iezzoni & Pritts, 1991).

Another applicability of the principal component analysis is to support the identification of variables which provide greater or lesser contributions to the accumulated variability. The characters with higher contribution are the ones that show higher eigenvectors in the components with higher eigenvalues; however, the ones with lower contributions are the those with the higher eigenvectors in the components with lower eigenvalues, in absolute values. Thus, when collecting a large number of variables, it is possible that some of them are less relevant when discriminating the evaluated materials, which allows of their disposal (Mardia et al., 1979; Cruz et al., 2012).

To define the clustering of genotypes, a cluster analysis was executed, based on the similarities observed within the first two principal components, using the minimum variance method (Wards); and, as a measure of similarity, the Euclidean distance was chosen. All statistical analyses were performed using the free software R version 3.2.3 (R Core Team, 2015).

Results and Discussion

The most correlated variables with PC1 (29.08%) were protein (0.781) and caffeine (0.729), whereas acidity (0.795) and total soluble solids (0.738) were more correlated with PC2 (27.82%) and ether extract (0.879) with PC3 (15.75%) (Table 1). This means that these characters show a greater variability within the studied population (Iezzoni & Pritts, 1991). The most correlated trait with PC4 (13.65%) was pH (0.560), which showed the lesser contribution to the variability. The third component, when included in the analysis, accumulates 72.65% of the total variance. Musoli et al. (2009) obtained 17.12% of the variance by the multivariate analysis in the first two components, and grouped the individuals in their category, using microsatellite markers of wild and cultivated robusta coffee populations from Uganda. Teressa et al. (2010) obtained 36,9% of the total variability in the first and second components of PCA, with two different groups of arabica coffee, one with Ethiopian accessions and other with commercial cultivars. Souza et al. (2013) obtained 24,9% of the total variability, in the first and

second component, using accessions that represent the *Coffea canephora* germplasm conserved and cultivated in Brazil. These results allow of the inferring on an acceptable level in the results obtained by PCA in the present study.

Total titratable acidity was more correlated with the second component (PC2) and also showed the highest eigenvector (0.699) in the sixth component (Table 1). This result contradicts that obtained by Scholz et al. (2011), in which this characteristic contributed to the formation of the first component. However, these results are the reality of the study population, which means that this characteristic could be disregarded in further studies related to genetic divergence in this conilon coffee population, based on quality parameters, which would result in economy in the qualitative analysis.

PCA were estimated with the first two PCs, and the first and third PCs, as a complementary analysis of the dissimilarity among genotypes, in order to arrange the genotypes in a bidimensional scatter plot, by means of their scores (Figure 1).

The genotypes were clustered into three groups due to the similarities observed in the behavior of the variables (Figure 2). The cluster arrangement choice was of the lower variance, and the cut point of the dendrogram was visually delimited at the point where a higher change in the level occurred. It is possible to note that cluster 1 relates more with higher values of total titratable acidity, and lower values of pH. Cluster 2,

Table 1. Estimates of eigenvalues (λj) corresponding to the percentages of variability explained by the principal components and respective eigenvectors (EV) e correlation (Cor) of variables caffeine, protein, ether extract (EE), total soluble solids (TSS), pH and total titratable acidity, evaluated in green coffee beans of 213 genotypes of conilon coffee (*Coffea canephora*), variety Robusta Tropical, under irrigation at the Cerrado.

Variables		PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
Eigen-	λj	1.745	1.669	0.945	0.819	0.464	0.358
values	λj (%)	29.08	56.9	72.65	86.3	94.03	100
Caffeine	EV	0.552	-0.257	0.223	-0.360	0.643	0.192
	Cor	0.729	-0.332	0.217	-0.326	0.438	0.115
Protein	EV	0.591	-0.230	-0.049	-0.255	-0.728	0.008
	Cor	0.781	-0.298	-0.048	-0.230	-0.496	0.005
EE	EV	0.300	0.105	-0.904	0.201	0.202	0.027
	Cor	0.396	0.136	-0.879	0.182	0.138	0.016
TSS	EV	0.159	-0.571	0.124	0.606	0.084	-0.509
	Cor	0.209	-0.738	0.120	0.549	0.057	-0.304
pH	EV	0.373	0.406	0.310	0.618	-0.057	0.463
	Cor	0.493	0.524	0.301	0.560	-0.039	0.277
Acidity	EV	-0 .302	-0 .615	-0 .140	0 .125	-0 .077	0 .699
	Cor	-0 .399	-0 .795	-0 .136	0 .113	-0 .053	0 .419

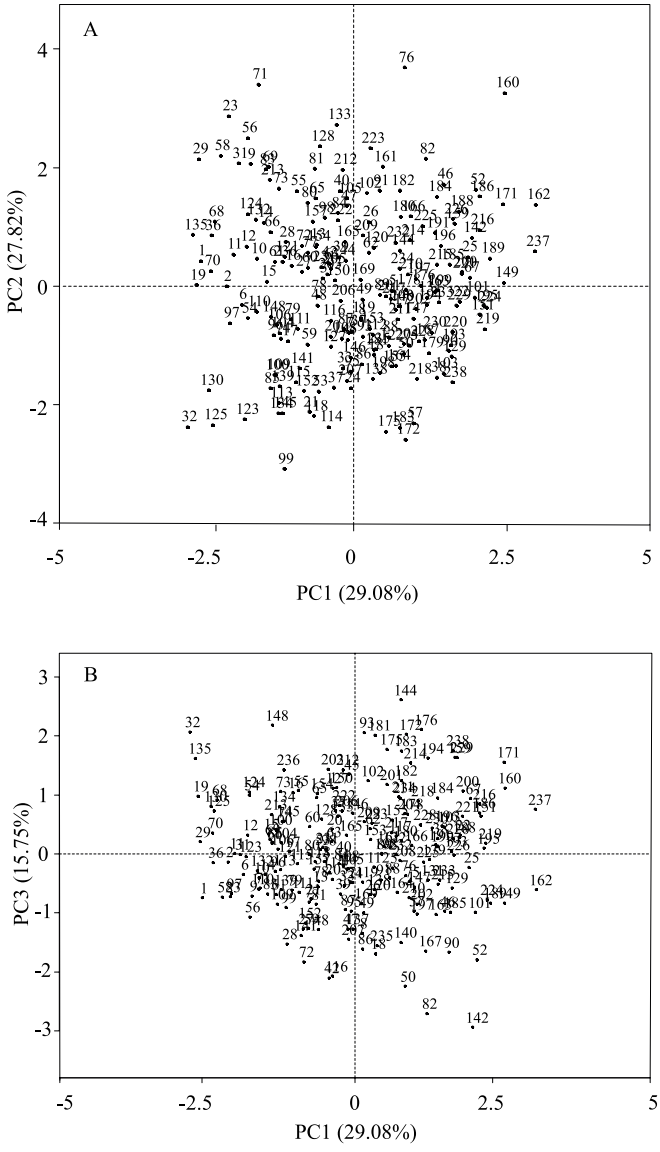

Figure 1. Dispersion of 213 genotypes of conilon coffee (*Coffea canephora*), Robusta Tropical variety, under irrigation at the Cerrado in function of the first and second principal components (A) and in relation to the first and third principal components (B), obtained based on analyzing the variables caffeine, protein, ether extract, total soluble solids, pH and total titratable acidity of green coffee beans.

however relates more with lesser values of TSS, caffeine, protein, and ether extract, and is opposite to the cluster 3, which grouped genotypes with higher values for these traits, suggesting this is a group that stands out for the studied characters.

As to the average genetic dissimilarities, the genotypes CPAC 49 showed the lowest-absolute value (1.686), whereas CPAC 114 showed the highest mean (4.375) and stands out as the most dissimilar in the group (Table 2). Based on the genetic dissimilarities obtained by the Euclidean matrix, the genotypes CPAC 160 and CPAC 32 showed the highest-dissimilarity value (7.768) for the first two principal components (PC1 and PC2). Nevertheless, these genotypes showed discrepant values for TSS (39.16 and 22.5°Brix, respectively) and caffeine (2.43 and 1.81%, respectively), which certainly contributed for the absolute value of higher distance.

Genetic dissimilarities associated with clustering and graphic results are essential for the genetic breeding program because they direct crosses involving genetically divergent parents, which may induce greater heterotrophic effects and variability in segregating generations, allowing of populations with a broader genetic base to be constructed, and providing greater genetic gains with selection cycles (Cecon, 2008; Cruz et al., 2012).

Based on the highest-genetic distances, out of the 213 genotypes, 5% were selected for further crossings considering the quality traits of coffee (Figures 1 and 2). Besides, from the selected genotypes, there were 11 ones showing a medium cycle (CPAC 76, CPAC 99, CPAC 123, CPAC 125, CPAC 130, CPAC 134, CPAC 145, CPAC 160, CPAC 162, CPAC 171, and CPAC 237), and one showing a high premature cycle (CPAC 32)

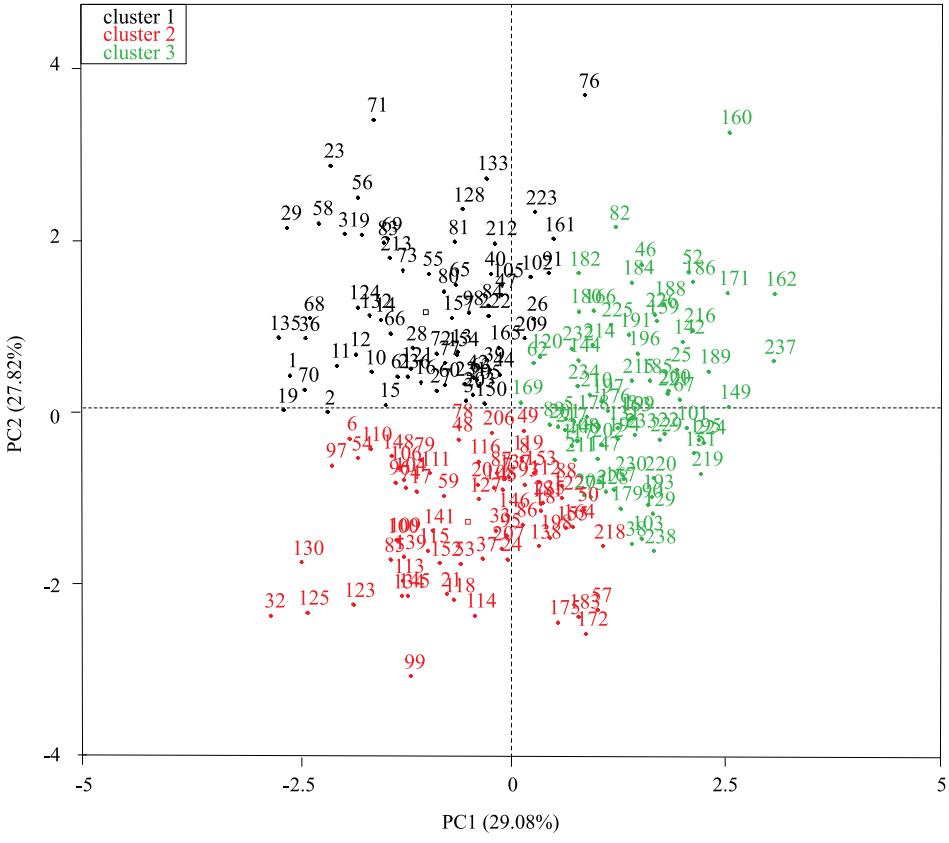

Figure 2. Cluster analysis of 213 genotypes of conilon coffee (*Coffea canephora*), Robusta Tropical variety, under irrigation at the Cerrado in function of the first and second principal components obtained based on analyzing the variables caffeine, protein, ether extract, total soluble solids, pH and total titratable acidity of green coffee beans. Each color represents a cluster.

Table 2. Average of genetic dissimilarities of 213 genotypes of conilon coffee (*Coffea canephora*), calculated based on the scores obtained from the principal component analysis for the first two components, using quality attributes.

CPAC	1	2	3	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	2.997	2.625	1.758	1.748	2.453	1.791	1.729	3.054	2.300	2.589	2.462	1.898	2.410	2.160	1.965	2.124	1.949	3.037	1.857
CPAC	21	22	23	24	25	26	27	28	29	31	32	33	36	37	38	39	40	42	44
	2.665	2.299	3.825	2.268	2.451	1.983	1.880	2.105	3.657	3.171	3.925	2.065	2.934	2.287	2.512	1.766	2.274	1.773	1.745
CPAC	45	46	47	48	49	50	52	53	54	55	56	57	58	59	60	61	62	63	65
	1.832	2.733	2.115	1.797	1.686	2.057	2.994	2.380	2.419	2.424	3.382	2.876	3.438	2.014	1.850	2.115	1.785	1.762	2.264
CPAC	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84
	2.289	2.423	2.961	2.876	2.837	4.018	1.967	2.568	2.168	2.023	4.063	1.904	1.783	2.004	2.252	2.589	2.894	2.865	2.058
CPAC	85	86	87	88	89	90	91	93	95	96	97	98	99	100	101	102	103	104	105
	2.652	2.016	1.779	1.868	1.711	2.390	2.306	1.807	2.087	2.217	2.651	2.053	3.572	2.492	2.479	2.251	2.537	2.149	2.184
CPAC	106	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125
	2.144	1.835	2.476	2.298	1.991	1.809	2.735	2.799	2.393	1.781	1.783	2.696	1.722	1.818	2.050	1.927	3.224	2.628	3.618
CPAC	127	128	129	130	131	132	133	134	135	137	138	139	140	141	142	144	145	146	147
	1.919	2.857	2.468	3.327	1.972	2.505	3.124	2.867	3.185	1.778	2.182	2.555	1.797	2.245	2.583	1.901	2.834	1.961	1.916
CPAC	148	149	150	151	152	153	154	155	157	159	160	161	162	163	164	165	166	167	169
	2.166	2.872	1.715	2.583	2.441	1.772	1.891	2.110	2.070	2.477	4.375	2.593	3.624	2.037	2.120	1.837	2.188	2.103	1.692
CPAC	171	172	175	176	178	179	180	181	182	183	184	185	186	188	189	191	193	194	195
	3.190	3.063	2.884	1.921	1.822	2.228	2.132	1.913	2.380	2.876	2.550	2.222	2.961	2.617	2.722	2.237	2.379	1.981	2.605
CPAC	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214
	2.218	1.898	2.134	2.055	2.339	1.728	1.905	1.747	1.985	1.746	1.705	2.189	1.793	1.875	1.855	1.824	2.513	2.722	2.012
CPAC	215	216	217	218	219	220	221	222	223	224	225	226	228	229	230	231	232	233	234
	2.102	2.713	1.747	2.378	2.704	2.323	2.331	1.999	2.802	2.649	2.191	2.498	2.070	2.268	2.189	1.784	1.921	2.086	1.822
CPAC	235	236	237	238															
	1.906	2.051	3.387	2.707															

that is also among five out of the eleven suggested crosses, which enhances variability for the choice of cycle in the breeding program of conilon coffee in the Cerrado.

Conclusions

1. The conilon coffee (*Coffea canephora*) accessions show genetic variability for the evaluated quality characteristics of green coffee beans; and CPAC 160 and CPAC 32 are the most promising genotypes to the breeding program of conilon coffee for cultivation under irrigation in the Cerrado.

2. Protein and caffeine content, total soluble solids, total lipids, and titratable acidity are the characteristics that contribute for the genetic variability among 213 genotypes of conilon coffee.

3. The most dissimilar genotypes are CPAC 160, that is a promising selection for high-level soluble solids and caffeine, and CPAC 32, that is favorable in the selection for its low-level caffeine.

Acknowledgments

To Consórcio Pesquisa Café, for financial support.

References

BARBIN, D.F.; FELICIO, A.L. de S.M.; SUN, D.-W.; NIXDORF, S.L.; HIROOKA, E.Y. Application of infrared spectral techniques on quality and compositional attributes of coffee: an overview. **Food Research International**, v.61, p.23-32, 2014. DOI: https://doi.org/10.1016/j.foodres.2014.01.005.

BONOMO, R.; OLIVEIRA, L.F.C. de; SILVEIRA NETO, A.N.; BONOMO, P. Produtividade de cafeeiros arábica irrigados no Cerrado goiano. **Pesquisa Agropecuária Tropical**, v.38, p.233-240, 2008.

CECON, P.R.; SILVA, F.F. e.; FERREIRA, A.; FERRÃO, R.G.; CARNEIRO, A.P.S.; DETMANN, E.; FARIA, P.N.; MORAIS, T.S. da S. Análise de medidas repetidas na avaliação de clones de café 'Conilon'. **Pesquisa Agropecuária Brasileira**, v.43, p.1171-1176, 2008. DOI: https://doi.org/10.1590/S0100-204X2008000900011.

CHENG, B.; FURTADO, A.; SMYTH, H.E.; HENRY, R.J. Influence of genotype and environment on coffee quality. **Trends in Food Science & Technology**, v.57, p.20-30, 2016. DOI: https://doi.org/10.1016/j.tifs.2016.09.003.

CRUZ, C.D.; REGAZZI, A.J.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: Ed. da UFV, 2012. 514p.

EMBRAPA CERRADOS. **Monitoramento da irrigação no Cerrado**. Available at: <http://hidro.cpac.embrapa.br/index.php>. Accessed on: Apr. 19 2009.

FERNANDES, A.L.T.; PARTELLI, F.L.; BONOMO, R.; GOLYNSKI, A. A moderna cafeicultura dos cerrados brasileiros. **Pesquisa Agropecuária Tropical**, v.42, p.231-240, 2012. DOI: https://doi.org/10.1590/S1983-40632012000200015.

FONSECA, A.F.A. da; FERRÃO, M.A.G.; FERRÃO, R.G. Vantagens e riscos no uso de mudas clonais de *Coffea canephora*. **Visão Agrícola**, v.12, p.17-18, 2013.

GUERRA, A.F.; ROCHA, O.C.; RODRIGUES, G.C.; SANZONOWICZ, C.; SAMPAIO, J.B.R.; SILVA, H.C.; ARAÚJO, M.C. de. **Irrigação do cafeeiro no Cerrado**: estratégia de manejo de água para uniformização de florada. Planaltina: Embrapa Cerrados, 2005. 4p. (Embrapa Cerrados. Comunicado técnico, 122).

HELRICH, K. (Ed.). **Official methods of analysis of the Association of Official Analytical Chemists**. 15th ed. Arlington: AOAC, 1990. 684p. Official Method 920.39, Official Method 942.15.

IEZZONI, A.F.; PRITTS, M.P. Applications of principal component analysis to horticultural research. **HortScience**, v.26, p.334-338, 1991. DOI: https://doi.org/10.21273/HORTSCI.26.4.334.

JOËT, T.; SALMONA, J.; LAFFARGUE, A.; DESCROIX, F.; DUSSERT, S. Use of the growing environment as a source of variation to identify the quantitative trait transcripts and modules of co-expressed genes that determine chlorogenic acid accumulation. **Plant, Cell & Environment**, v.33, p.1220-1233, 2010. DOI: https://doi.org/10.1111/j.1365-3040.2010.02141.x.

JONES, R.W.; REINOT, T.; FREI, U.K.; TSENG, Y.; LÜBBERSTEDT, T.; MCCLELLAND, J.F. Selection of haploid maize kernels from hybrid kernels for plant breeding using near-infrared spectroscopy and SIMCA analysis. **Applied spectroscopy**, v.66, p.447-450, 2012. DOI: https://doi.org/10.1366/11-06426.

MARDIA, K.V.; KENT, J.T.; BIBBY, J.M. **Multivariate analysis**. London: Academic Press, 1979. 521p.

MUSOLI, P.; CUBRY, P.; ALUKA, P.; BILLOT, C.; DUFOUR, M.; DE BELLIS, F.; POT, D.; BIEYSSE, D.; CHERRIER, A.; LEROY, T. Genetic differentiation of wild and cultivated populations: diversity of *Coffea canephora* Pierre in Uganda. **Genome**, v.52, p.634-646, 2009. DOI: https://doi.org/10.1139/G09-037.

PEZZOPANE, J.R.M.; PEDRO JÚNIOR, M.J.; THOMAZIELLO, R.A.; CAMARGO, M.B.P. de. Escala para avaliação de estádios fenológicos do cafeeiro arábica. **Bragantia**, v.62, p.499-505, 2003. DOI: https://doi.org/10.1590/S0006-87052003000300015.

PINHEIRO, L.C. de M.; GOD, P.I.V.G.; FARIA, V.R.; OLIVEIRA, A.G.; HASUI, A.A.; PINTO, E.H.G.; ARRUDA, K.M.A.; PIOVESAN, N.D.; MOREIRA, M.A. Parentesco na seleção para produtividade e teores de óleo e proteína de soja via modelos mistos. **Pesquisa Agropecuária Brasileira**, v.48, p.12461253, 2013. DOI: https://doi.org/10.1590/S0100-204X2013000900008.

PLANS, M.; SIMÓ, J.; CASAÑAS, F.; SABATÉ, J. Near-infrared spectroscopy analysis of seed coats of common beans (*Phaseolus vulgaris* L.): a potential tool for breeding and quality evaluation. **Journal of Agricultural and Food Chemistry**, v.60, p.706-712, 2012. DOI: https://doi.org/10.1021/jf204110k.

POJIĆ, M.M.; MASTILOVIĆ, J.S. Near infrared spectroscopy — advanced analytical tool in wheat breeding, trade, and processing. **Food and Bioprocess Technology**, v.6, p.330-352, 2013. DOI: https://doi.org/10.1007/s11947-012-0917-3.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>. Accessed on: May 6 2015.

SANTOS, J.R.; SARRAGUÇA, M.C.; RANGEL, A.O.S.S.; LOPES, J.A. Evaluation of green coffee beans quality using near infrared spectroscopy: a quantitative approach. **Food Chemistry**, v.135, p.1828-1835, 2012. DOI: https://doi.org/10.1016/j.foodchem.2012.06.059.

SCHOLZ, M.B. dos S.; FIGUEIREDO, V.R.G. de; SILVA, J.V.N. da; KITZBERGER, C.S.G. Características físico-químicas de grãos verdes e torrados de cultivares de café (*Coffea arabica* L.) do IAPAR. **Coffee Science**, v.6, p.245-255, 2011.

SCHOLZ, M.B. dos S.; KITZBERGER, C.S.G.; PEREIRA, L.F.P.; DAVRIEUX, F.; POT, D.; CHARMETANT, P.; LEROY, T. Application of near infrared spectroscopy for green coffee biochemical phenotyping. **Journal of Near Infrared Spectroscopy**, v.22, p.411-421, 2014. DOI: https://doi.org/10.1255/jnirs.1134.

SIESLER, H.W.; OZAKI, Y.; KAWATA, S.; HEISE, H.M. (Ed.). **Near-infrared spectroscopy**: principles, instruments, applications. Weinheim: J. Wiley & Sons, 2008. 361p.

SOUZA, F. de F.; CAIXETA, E.T.; FERRÃO, L.F.V.; PENA, G.F.; SAKIYAMA, N.S.; ZAMBOLIM, E.M.; ZAMBOLIM, L.; CRUZ, C.D. Molecular diversity in *Coffea canephora* germplasm conserved and cultivated in Brazil. **Crop Breeding and Applied Biotechnology**, v.13, p.221-227, 2013. DOI: https://doi.org/10.1590/S1984-70332013000400001.

TERESSA, A.; CROUZILLAT, D.; PETIARD, V.; BROUHAN, P. Genetic diversity of Arabica coffee (*Coffea arabica* L.) collections. **Ethiopian Journal of Applied Sciences and Technology**, v.1, p.63-79, 2010.

TEUBER, R. Geographical indications of origin as a tool of product differentiation: the case of coffee. **Journal of International Food & Agribusiness Marketing**, v.22, p.277-298, 2010. DOI: https://doi.org/10.1080/08974431003641612.

WILES, P.G.; GRAY, I.K.; KISSLING, R.C. Routine analysis of proteins by Kjeldahl and Dumas methods: review and interlaboratory study using dairy products. **Journal of AOAC International**, v.81, p.620-632, 1998.

ZAIDAN, Ú.R.; CORRÊA, P.C.; FERREIRA, W.P.M.; CECON, P.R. Ambiente e variedades influenciam a qualidade de cafés das matas de Minas. **Coffee Science**, v.12, p.240-247, 2017. DOI: https://doi.org/10.25186/cs.v12i2.1256.

ZENEBON, O.; PASCUET, N.S.; TIGLEA, P. (Coord.). **Métodos físico-químicos para análise de alimentos**. 4.ed. São Paulo: Instituto Adolfo Lutz, 2008. 1020p.