Journal of Agricultural Science; Vol. 12, No. 4; 2020
ISSN 1916-9752 E-ISSN 1916-9760
Published by Canadian Center of Science and Education

Genome Introgression of Híbrido de Timor and Its Potential to Develop High Cup Quality *C. arabica* Cultivars

Tesfahun Alemu Setotaw[1,2], Eveline Teixeira Caixeta[2,3], Eunize Maciel Zambolim[2], Tiago Vieira Sousa[2], Antônio Alves Pereira[4], Antonio Carlos Baião[3,4], Cosme Damião Cruz[5], Laércio Zambolim[6] & Ney Sussumu Sakiyama[7]

[1] Holetta Agricultural Research Center, Ethiopian Institute of Agricultural Research, Adiss Ababa, Ethiopia

[2] Instituto de Biotecnologia Aplicada à Agropecuária, Universidade Federal de Viçosa, Viçosa, MG, Brazil

[3] Empresa Brasileira de Pesquisa Agropecuária, Embrapa Café, Brasília, DF, Brazil

[4] Empresa de Pesquisa Agropecuária de Minas Gerais, Viçosa, MG, Brazil

[5] Departamento de Fitotecnia, Universidade Federal de Viçosa, Viçosa, MG, Brazil

[6] Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, Brazil

[7] Departamento de Biologia, Universidade Federal de Viçosa, Vicosa, MG, Brazil

Correspondence: Tesfahun Alemu Setotaw, Holetta Agricultural Research Centter, Ethiopian Institute of Agricultural Research, P.O. Box 2003, Addis Ababa, Ethiopia. Tel: 251-968-171-644. E-mail: setotaw2006@gmail.com

Received: December 17, 2019 Accepted: February 15, 2020 Online Published: March 15, 2020

doi:10.5539/jas.v12n4p64 URL: https://doi.org/10.5539/jas.v12n4p64

Abstract

Híbrido de Timor is the principal source for disease and pest resistance genes in *C. arabica* breeding program worldwide. The part of the chromosome responsible for resistance introgressed from *C. canephora* to Híbrido de Timor are claimed to affect the cup quality of the *C. arabica* cultivars derived from the crossing program of Híbrido de Timor. Therefore, this work was done to study the genome introgression of Híbrido de Timor and its impact on the cup quality on the *C. arabica* cultivars. For genome introgression and genetic relationship analysis, seventy-six accessions from *C. arabica*, *C. canephora* and Híbrido de Timor were analyzed using AFLP and SSR molecular markers. To understand the effect of genome introgressed from Híbrido de Timor on cup quality, three *C. arabica*, seven Híbrido de Timor and six cultivars derived from Híbrido de Timor × *C. arabica* were genotyped using SSR molecular markers and sensorial analysis was performed. The genetic diversity analysis among the tested genotypes showed high genetic similarity between Híbrido de Timor with *C. arabica* and clear differentiation among coffee species. The analysis of genome introgression of *C. arabica* and *C. canephora* var Robusta into Híbrido de Timor not reach 30% of *C. canephora* genome. The sensorial analysis of coffee genotypes demonstrated non-significant difference on cup quality parameters among *C. arabica* cv Bourbon and cultivars derived from Híbrido de Timor that showed the possibility of developing *C. arabica* cultivars without affecting the cup quality. Similarly, the SSR marker diversity showed high genetic similarity between the Bourbon and the *C. arabica* cultivars derived from Híbrido de Timor.

Keywords: Molecular markers, genetic diversity, multivariate analysis, Bayesian model, disease resistance, coffee quality, sensorial analysis

1. Introduction

Coffea arabica L. (2n = 2x = 44) is a true allotetraploid species (Clarindo & Carvalho, 2008), native to Africa. *C. arabica* L. and *C. canephora* Pierre are the two most cultivated and commercialized coffee species in the world. Among them *C. arabica* L. has more than 70% contribution in world coffee market. It has originated in the southwestern Ethiopia and produce high cup quality. Even if the world coffee production and consumption depend on *C. arabica*, its production was greatly affected by diseases and pests which reduce its productivity, due to lack of resistance genes for the major diseases and pests. To control the diseases (Coffee leaf rust and coffee berry disease) and pests the producer use copper-based fungicides and herbicides, which cost annually US$ 2-2.5 billion without considering its potential environmental hazards (Van der Vossen, 2009).

An economically and environmentally friendly technique to control the disease and pests is the use of resistance cultivars. Especially, the use of coffee cultivars resistant to disease gives an economic advantage for the small-scale farmers in the developing world. Considering these advantages, the coffee researchers developed *C. arabica* cultivars resistant to coffee leaf rust and coffee berry diseases. The coffee leaf rust resistant cultivars derived from the progenies of Híbrido de Timor have planted in Latin America and East Africa (Van der Vossen, 2009; Pereira et al., 2008). These cultivars contribute for ecologically sustainable coffee production and considerable socio-economic benefit due to its high yield and reduced cost of production without affecting the cup quality (Van der Vossen, 2009).

Híbrido de Timor ("HT") is the interspecific hybrid between *C. arabica* and *C. canephora,* first found in plantation of cultivar *Typica* in Timor Island in 1917 (Bettencourt, 1973). It is used as source of resistance gene for economically important diseases and pests of coffee such as coffee leaf rust (*Hemileia vastatrix*), coffee berry disease (*Colletotrichum kahawae*), root knot nematode (*Meloidgyne exigua*) and bacteriosis (*Pseudomonas syringae* pv Garçae) (Bertrand et al., 2003).

The small portion of the *C. canephora* genome introgressed into "HT" gave resistance to coffee leaf rust and other diseases. However, the concern in using this genotype in *C. arabica* breeding was the possibility of the *C. canephora* affect the cup quality. Bertrand et al. (2003) evaluated the effect of genome introgression of *C. canephora* on cup quality for lines derived from "HT" and showed the possibility of finding lines resistance to disease (coffee leaf rust) and nematodes combined with good quality as *C. arabica* cultivars. In the processes of developing resistance cultivars, breeders used the "HT" as a donor for the resistant gene and make several backcrossing with *C. arabica* that helps to maintain the desired cup quality of *C. arabica* within the new cultivars. Besides, "HT" has high natural genetic relationship with *C. arabica* (Lashermes et al., 1993, 1996, 1999). From this process, different cultivars of *C. arabica* were released in Brazil (Setotaw et al., 2013) and are being widely used in America countries.

The Brazilian germplasm has several collections of "HT", which contain important sources of gene for disease and pest resistance. These accessions have been used in large extent in the breeding program of coffee in the world. So, understanding the genome introgression from their origin (*C. arabica* and *C. canephora*), its relation to other coffee species and its impact on coffee quality of *C. arabica* cultivars released for production is an important task. The potential of "HT" accessions in the development of high cup quality *C. arabica* cultivars was reported by Sobreira et al. (2015).

The genetic study and relationship of different coffee species can be assessed using molecular markers such as Amplified Fragment Length Polymorphism (AFLP) and Simple Sequence Repeat (SSR) (Setotaw et al., 2010; Ferrão et al., 2013). AFLP is preferred to study the genetic diversity and genetic relationship among populations since it has a capacity of screening different regions of the genome distributed randomly throughout the chromosome (Mueller & Wolfenbarger, 1999). Another advantage of AFLP markers in relation to SSR is it does not require prior sequence information and relatively low start-up cost. In contrast SSR markers are multi-allelic, species and locus specific and relatively have high start up cost especially to develop the marker. SSR markers are used to study the genetic diversity and QTL identification in coffee by several authors (Setotaw et al., 2010; Missio et al., 2009; Ferrão et al., 2013; Silva et al., 2018).

Therefore, this work was done with the objectives: 1) to characterize accessions of "HT" using AFLP marker and assess its genetic relationship to other coffee groups; 2) to know the contribution of *C. arabica* and *C. canephora* in genome of "HT" accessions using AFLP and SSR markers; and 3) to analyze the potential of "HT" for developing *C. arabica* cultivars with disease resistance without affecting the cup quality using analysis sensorial.

2. Materials and Methods

Two separate experiments were conducted to assess the genetic relationship of "HT" with other coffee species and its impact on cup quality. The details of the experiments were described in the following section.

2.1 Genetic Materials

2.1.1 Experiment I

Seventy-six coffee accessions were used for the genome introgression analysis of "HT" and its genetic relationship study. These accessions comprise of five genotypes that belong to *C. arabica* species, 25 to *C. canephora* species (15 of the Robusta varietal group and 10 of the Conilon varietal group), 46 "HT" accessions and 1 Eugenoides (Table 1). For the genetic relationship and genome introgression analysis, AFLP and SSR markers were used.

Table 1. Coffee accessions used for genetic relationship study, genome introgression of Híbrido de Timor (HT) and sensorial analysis

Code	Genotype Name	Description	Code	Genotype Name	Description	Code	Genotype Name	Description
1	Catuaí UFV2144 [I]	*C. arabica*	31	UFV 427-55 [I]	HT	61	Conilon 3751 [I]	Conilon
2	Catuaí IAC44 [I]	*C. arabica*	32	UFV 427-56 [I]	HT	62	Conilon 3580 [I]	Conilon
3	Típica UFV2945 [I]	*C. arabica*	33	UFV 427-65 [I]	HT	63	Guarani 513 [I]	Robusta
4	Bourbon UFV2952 [I, II]	*C. arabica*	34	UFV 427-90 [I]	HT	64	Guarani 514 [I]	Robusta
5	Bourbon UFV535-1 [I, II]	*C. arabica*	35	UFV 438-52 [I]	HT	65	Robusta C2258 [I]	Robusta
6	CIFIC 832/1 [I]	HT	36	UFV 439-02 [I]	HT	66	Robusta 2257-2 [I]	Robusta
7	CIFIC 832/2 [I]	HT	37	UFV 440-22 [I]	HT	67	Robusta 2257-1 [I]	Robusta
8	CIFIC 4106 [I]	HT	38	UFV 442-108 [I]	HT	68	Robusta 640-1 [I]	Robusta
9	CIFIC 1343/269 [I]	HT	39	UFV 443-03 [I]	HT	69	Robusta 640-2 [I]	Robusta
10	UFV 376-01 [I, II]	HT	40	UFV 446-08 [I]	HT	70	Robusta 640-2 [I]	Robusta
11	UFV 376-04 [I]	HT	41	UFV 445-46 [I, II]	HT	71	Apoatã-1 [I]	Robusta
12	UFV 376-05 [I]	HT	42	UFV 428-04 [I & II]	HT	72	Apoatã-2 [I]	Robusta
13	UFV 376-35 [I]	HT	43	UFV 432-07 [I]	HT	73	Apoatã -3 [I]	Robusta
14	UFV 376-37 [I]	HT	44	UFV 437-06 [I]	HT	74	Guarini-1 [I]	Robusta
15	UFV 376-52 [I, II]	HT	45	UFV 441-03 [I]	HT	75	Guarini-2 [I]	Robusta
16	UFV 376-57 [I]	HT	46	UFV 447-48 [I]	HT	76	*C. eugenoides* [I]	Eugenoides
17	UFV 376-79 [I]	HT	47	UFV 448-69 [I]	HT	77	UFV 428-04 [II]	HT
18	UFV 377-01 [I]	HT	48	UFV 449-20 [I, II]	HT	78	Catiguá MG2 [II]	Cultivar
19	UFV 377-02 [I]	HT	49	UFV 450-61 [I, II]	HT	79	MGS Catiguá 3 [II]	Cultivar
20	UFV 377-23 [I]	HT	50	UFV 451-41 [I]	HT	80	Paraíso MG H 419-1 [II]	Cultivar
21	UFV 377-24 [I]	HT	51	Encapa 03 [I]	Conilon	81	Pau Brasil MG1 [II]	Cultivar
22	UFV 377-34 [I]	HT	52	Encapa 04 [I]	Conilon	82	Sacramento MG1 [II]	Cultivar
23	UFV 379-07 [I]	HT	53	Encapa 05 [I]	Conilon	83	UFV 971-99-313 [II]	Cultivar
24	UFV 408-18 [I]	HT	54	Encapa 06 [I]	Conilon	84	Catuaí Amarelo IAC62 [II]	Cultivar
25	UFV 408-26 [I]	HT	55	Encapa 07 [I]	Conilon			
26	UFV 408-28 [I]	HT	56	Encapa 08 [I]	Conilon			
27	UFV 427-01 [I]	HT	57	Encapa 09 [I]	Conilon			
28	UFV 427-09 [I, II]	HT	58	Conilon 66-1 [I]	Conilon			
29	UFV 427-15 [I]	HT	59	Conilon 66-2 [I]	Conilon			
30	UFV 427-22 [I]	HT	60	Conilon 66-3 [I]	Conilon			

Note. HT: Híbrido de Timor; I: genotypes used in experiment I (genome introgression and genetic relationship study), II: genotypes used in experiment II (for sensorial analysis).

2.1.2 Experiment II

To evaluate the effect of the genome introgression on the cup quality of *Coffea arabica* cultivars released in Brazil, we used pure *C. arabica* cultivars, *C. arabica* cultivars derived from the crossing program of *C. arabica* × "HT", and "HT" accessions frequently used in the breeding programs. The "HT" accessions included in this experiment were also included in the first experiment. The list of the genotypes and the coffee cultivars used in this study were presented on Table 1. To study the genetic relationship among these genotypes, SSR molecular markers were used.

2.2 Genotyping of the Accessions Using AFLP and SSR Molecular Markers

2.2.1 Extraction of DNA

The DNA of the genotypes were extracted according to the method described by Diniz et al. (2005) from young green leaves. The DNA concentration was quantified using Spectrophotometer Smart Spec of BioRad (Hercules, California, United States). The extracted DNA was diluted in TE (Tri-HCL 10 mM, EDTA 1 mM, pH 8.0) to concentration of 50 ηg/µl for AFLP analysis and 25 ηg/µl for SSR marker analysis.

2.2.2 AFLP (Amplified Fragment Length Polymorphism) Analysis

The AFLP genotyping of coffee accessions were done according to the method described by Brito et al. (2010). The primer combinations *Mse*I-AGC/*Eco*RI-CGT and *MseI*-AGC/*Eco*RI-CTC were used to genotype the coffee accessions in this study.

2.2.3 SSR (Simple Sequence Repeat) Marker

Eighteen microsatellite primers obtained from Combes et al. (2000) and Rovelli et al. (2000) were used to genotype the coffee in this experiment. Primers AJ250254, AJ250255, AJ250258, and AJ250260 were obtained from Combes et al. (2000). Primers AJ308752, AJ308754, AJ308755, AJ308769, AJ308776, AJ308792, AJ308796, AJ308814, AJ308819, AJ308821, AJ308825, AJ308833, AJ308819, BQ448809 were obtained from Rovelli et al. (2000). The PCR reaction was realized in a total volume of 20 μL containing 50 ηg of DNA, 0.6 unit of Taq DNA polymerase, buffer 1x, 1 mM of $MgCl_2$, 150 μM of each dNTP and 0.1 μM of each primer. The amplification was done using the procedure of touchdown-PCR that consisted of denaturation at 94 °C for 2 minutes, followed by 13 cycles of denaturation at 94 °C for 30 seconds, primer annealing at 67 °C to 55 °C for 30 seconds, reducing 1 °C of each cycle and an extension of primer at 72 °C for 30 seconds. This step was followed by 30 more cycles of denaturation at 94 °C for 30 seconds, primer annealing at 55 °C for 30 seconds and primer extension at 72 °C for 30 seconds. The final extension was done at 72 °C for 8 minutes. The PCR products were separated on a 6% denaturing polyacrylamide gel and visualized by a silver staining solution.

2.3 Data Analysis

2.3.1 Genetic Relationship of "HT" With Other Coffee Accessions Using AFLP Molecular Markers

To study the genetic relationship between "HT" and other coffee species, the AFLP molecular marker were used to genotype 76 accessions (Table 1). The gels of AFLP were scored by visual inspection for presence (1) or absence (0) of specific AFLP-bands. Only distinct major bands were scored. To analyze the AFLP data using structure population genetic analysis software (Pritchard et al., 2000), the data matrix was coded according to Falush et al. (2007). The AFLP data statistical package (Ehrich, 2006) was used to manage the data conversion from txt to structure format.

The distance based and model-based clustering analysis was performed to study the genetic relationship of "HT" with other accessions. For the distance-based clustering analysis, the Jaccard similarity coefficient (Jaccard, 1908) was estimated using NTSYS-pc software (Version 2.10L; Rohlf, 2000). The clustering analysis and the dendrogram were generated from the similarity matrix by the UPGMA (*Unweighted Pair-Group Method using Arithmetic Average*) method.

The principal coordinate analysis (PCoA) was done among accessions based on genetic dissimilarity matrix (1-Jaccard similarity coefficient) using GenAlex 6.2 population genetic analysis software (Peakall & Smouse, 2006). Nei genetic diversity index (Nei, 1973), Shannon´s information and percent polymorphic bands (P%) with in populations were estimated using POPGENE statistical software version 1.3 (Yeh & Boyle, 1997). The pair-wise F_{st} analysis to understand the relationship between Híbrido de Timor with other coffee species was determined by AFLP surv (Vekemans et al., 2002). The analysis of molecular variance (AMOVA) among the group of coffee species was done using the statistical software Arlequin ver. 3.1 (Excoffier et al., 2006).

The model-based Bayesian clustering analysis was done using Structure 2.3.1 population genetic analysis software (Pritchard et al., 2000) to group the accessions of coffee species into its respective groups applying admixture model. The number of populations (k) was varied from two to twelve with twenty replicate runs per each assumed k value. It was used a burning period length of 10,000 runs and a post-burning sampling by Markov Chain Monte Carlo of 100,000 runs to estimate the number of subpopulations for each of the k values. The appropriate number of cluster was determined according to Evanno et al. (2005) using Structure Harvester program (Earl & vonHoldt, 2012).

2.3.2 Genome Introgression Analysis of "HT" From *C. arabica* and *C. canephora* var Robusta

The accessions from *C. arabica* and *C. canephora* varietal group Robusta were used to estimate the shared percentage of bands into "HT". As reported by Bettencourt (1973), the "HT" originated by the natural cross of these two species and from our first study we able to conclude the "HT" is originated from *C. arabica* and *C. canephora* var Robusta. The percentage of band shared among *C. arabica* and "HT" was calculated by counting the number of bands observed on both *C. arabica* and Híbrido de Timor. Those bands not observed in the genome of *C. arabica* were considered as obtained from Robusta varietal group (*C. canephora*). This analysis was done using the AFLP and SSR molecular marker data obtained as described in the materials and methods.

2.3.3 Study the Genetic Relationship and Cup Quality of the Different Groups of Coffee

To understand the impact of the "HT" on the cup quality of the *C. arabica* cultivars released in Brazil, a total of 16 genotypes, including three *C. arabica* (two of Bourbon and one Catuaí Vermelho), seven "HT", and six cultivars derived from the cross between "HT" and *C. arabica*, were submitted for sensorial analysis. These genotypes were laid using Randomized Complete Block Design (RCBD)—with three replications at

experimental site of EPAMIG (Empresa de Pesquisa Agropecuária de Minas Gerais), Patrocínio, Minas Gerais, Brazil. For the sensorial analysis, the coffee seed samples were collected in two growing seasons (2008 and 2009) and submitted for sensorial analysis by two qualified evaluator. Before the coffee bean submitted for sensorial analysis, the mature green cherry of each accession in each replication was harvested and prepared to be suitable for the sensorial analysis according to the coffee evaluation standard.

Cherry fruits collected from each cultivar were mechanically pulped. Seeds were prepared with and without mucilage before sun dried. After hulling, seeds were kept in a coded paper bag until three professional tasters accomplished sensorial analysis. Cup quality was analyzed according to criteria adopted for specialty coffees, with cumulative scores for aroma, uniformity, absence of defect, clean cup, flavor, acidity, body, aftertaste, balance, aspect and overall. The data obtained were subjected for analyses of variance using the following model:

$$Y_{ij} = \mu + G_i + yr_i + Pr_i + GYr + e_{ij} \quad (1)$$

Where, Y_{ij} is the response, G_i genotype effect, yr_i, year effect and GYr genotype x year interaction effect and e_{ij} residual effect.

The ANOVA was done using the PROC GLM procedure of SAS statistical analysis software (SAS inst. 2007).

3. Results

3.1 Genetic Relationship of "HT" With Other Coffee Accession

The accessions of coffee in this study categorized into five groups as *C. arabica*, *C. canephora* var Robusta, *C. canephora* var Conilon, "HT" and *C. eugenoides* before the clustering analysis. The individual population diversity measures were estimated for all the populations. The principal coordinate analysis (Figure 1) showed clear differentiation among different coffee groups and high similarity between "HT" and *C. arabica* accessions. In addition, the *C. eugenoides*, one of the parents for *C. arabica*, appeared in the graph near *C. arabica* and "HT" groups proved the possible origin of Arabica coffee. The high genetic diversity was observed within *C. canephora* var Robusta (Figures 1 and 2).

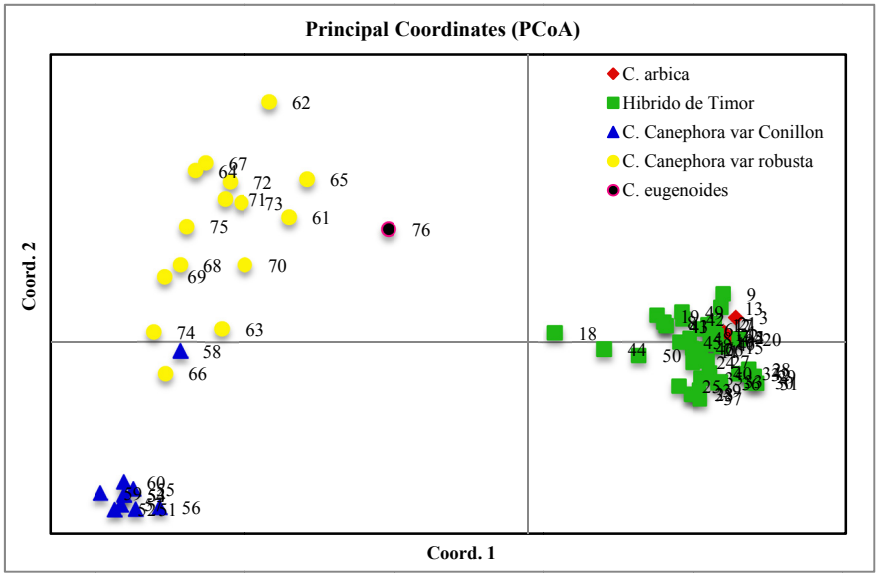

Figure 1. Principal coordinate analysis of AFLP diversity among coffee species. Where circle-red (*C. arabica*), square green (Híbrido de Timor), yellow circle (*C. canephora* var robusta), blue (*C. canephora* var conilon) and red circle with black (*C. eugenoides*)

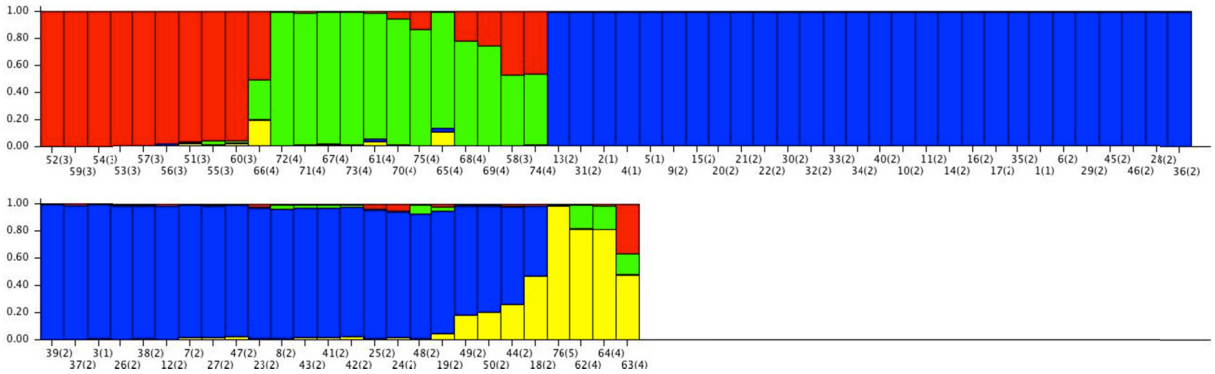

Figure 2. The horizontal graph k = 4 cluster groups produced using the Structure software, where the vertical indicate the percentage of membership coefficient and the horizontal refers the ID of the genotypes referred on material and method section. (Cluster I (red: *C. canephora* var conilon), Cluster II (green: *C. canephora* var robusta *C. arabica* & Híbrido de Timor), Cluster III (Blue: *C. arabica* & Híbrido de Timor), Cluster IV (yellow: Eugenoides). The numbers in horizontal corresponds the accession code in Table 1

The AMOVA (Analysis of Molecular Variance) confirmed high genetic differenciation among coffee groups. The total genetic variation was partitioned among population (60.5%) and within populations (39.05%) (Table 3, supplement). The overall F_{st} value (0.609) demonstrated the existence of high genetic differenciation among coffee groups. The pairwise F_{st} analysis between "HT" and other coffee species showed high genetic similarity between *C. arabica* and "HT", in contrast they showed high genetic dissmilarity with *C. canephora* (Table 4).

Table 3. AMOVA of genetic variation using AFLP markers

Source of Variation	*df*	Sum of squares	Variance component	Percent of total component variance
Among populations	4	498.62	10.4919	60.95
Within populations	72	484.10	6.7237	39.05
Total	76	982.72	17.2156	100
Fixation Index F_{st} = 0.609				

Table 4. Pairwise F_{st} (allele frequency used was estimated by square root method) between coffee species (Phylip format) investigated in this study.

	C. arabica	*C. canephora* var Robusta	Híbrido de Timor	*C. canephora* var Conilon
C. arabica	0	0.6346	0.184	0.8038
C. canephora var Robusta	0.6346	0	0.4312	0.343
Híbrido de Timor	0.184	0.4312	0	0.6148
C. canephora var Conilon	0.8038	0.343	0.6148	0

Note. F_{st} over all populations = 0.608.

The population structure analysis based on Bayesian statistics (Pritchard et al., 2000) grouped the 76 accessions into four clusters (Figure 2). The *C. arabica* and "HT" were grouped in Cluster III (blue) with shared ancestral probability greater than 0.92. Cluster IV (yellow) contains *C. eugenoides* and accessions from *C. canephora* var robusta. Cluster I (red) and Cluster II (green) contains all the accessions of Conilon and Robusta, respectively (Figure 2). The high admixture proportion was observed with the accessions of *C. canephora* var Robusta. This result also shows high level of similarity between "HT" and *C. arabica*.

The method of analysis presented in this study confirmed the origin of "HT", which is the interspecific hybridization between *C. arabica* and *C. canephora* var Robusta in Timor Island. The clustering of "HT" with *C. arabica* with high membership coefficient (Figure 2) and presented in the same cluster in the PCoA (Figure 1) also confirmed this fact and shows more similarity with *C. arabica*.

3.2 Genome Introgression Analysis of Híbrido de Timor

To understand the genome introgression analysis of "HT" from their relative's *C. canephora* var Robusta and *C. arabica*, the shared bands from its parent were estimated. The maximum percentage of bands shared only with *C. canephora* var Robusta were about 30% and 20% for AFLP and SSR marker, respectively (Table 2), when all the accessions of "HT" were considered together. The genome introgression analysis using two molecular markers and a large number of "HT" accessions showed that the "HT" has more percentage of genome from *C. arabica*, making it more genetically similar to species *C. arabica* than to *C. canephora*. The slight difference in the percentage of *C. canephora* genome within "HT" in AFLP marker than SSR marker may be attributed due to the high information obtained from this marker since it produced differences in all part of the genome than SSR marker.

Table 2. The number of bands shared by the *C. arabica* and *C. canephora* var Robusta with Híbrido de Timor based on AFLP and SSR molecular markers

	AFLP		SSR	
	Nº bands	%	Nº bands	%
Present in all *C. arabica* accessions	45	62.5	25	56.82
Polymorphism in *C. arabica* and *C. canephora*	5	6.9	10	22.73
Not detected in any *C. arabica* accession	22	30.5	9	20.45
Total	72	100	44	100

3.3 Quality Analysis of C. arabica Cultivars Derived From Híbrido de Timor

To understand the genetic relationship and the cup quality attributes among the cultivars of coffee, "HT" and cultivars developed from the crossing of "HT" and *C. arabica*, some of these genotypes were analyzed using SSR molecular marker and subjected to sensorial analysis. The UPGMA clustering method was able to group 16 genotypes into three principal cluster using SSR markers (Figure 3): Cluster I (UFV 427-09, UFV 450-61, Catuai amarelo IAC62, Bourbon UFV 535-1, Bourbon UFV2952, UFV 449-20, Paraíso MG H 419-1 , UFV 445-46), Cluster II (MGS Catiguá 3, Pau Brasil MG1, UFV 376-01), and Cluster III (UFV 376-52, Catiguá MG2, UFV 428-04, Sacramento MG1, UFV 971-99-313). The grouping pattern did not follow the cultivar groups since in each cluster we found C. *arabica* cultivars, "HT" and cultivar derived from *C. arabica* × "HT". *C. arabica* cultivars var Bourbon known for its high quality were grouped with the cultivars derived from crossing of *C. arabica* × "HT" and other "HT" accessions (Figure 3). The cultivar Paraíso MG H 419-1 was grouped with UFV 449-20 ("HT") with 100 percent similarity as shown in Figure 3. Similarly, cultivar Catiguá MG2 (*C. arabica* × "HT") also grouped with "HT" (UFV 376-52, UFV428-04) and Sacramento MG1 (*C. arabica*).

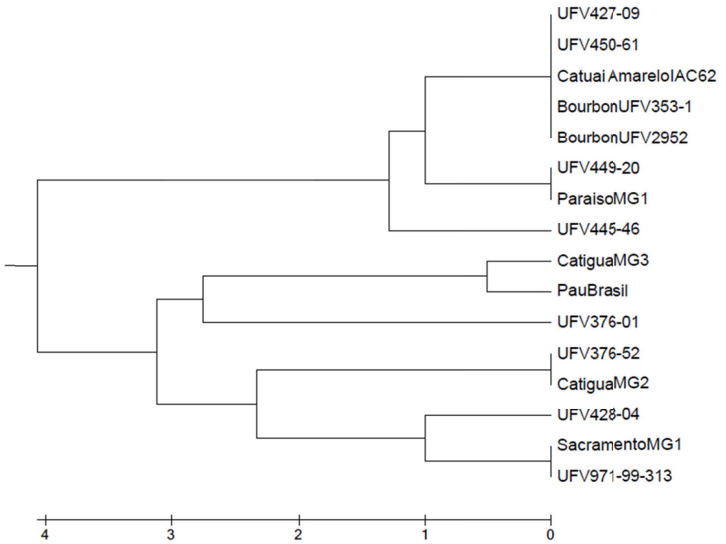

Figure 3. The dendrogram obtained by UPGMA clustering method based on the genetic distance produced from SSR molecular marker

The quality analysis among the tested genotypes showed significant difference among genotypes in over two years period (p = 0.05) on fragrance (aroma), uniformity, acidity, body, aftertaste, balance, and total quality. In general, most of the genotypes presented similar quality attributes, which revealed the little impact of the genome introgression on the quality attributes in the *C. arabica* cultivars developed from the crossing of "HT" × *C. arabica*. The bargraph on the quality attributes for the cluster groups formed using SSR marker diversity (Figure 3) showed high similarity among the genotypes (Figure 4) for all quality parameters. It proved the genome introgression from *C. canephora* via "HT" did not affect the quality even if it contributed for the resistance gene for disease and pests. The mean data of the quality parameter on different groups of coffee (*C. arabica*, "HT" and cultivars derived from the crossing of "HT" × *C. arabica*) showed non significant difference among them (Figure 4).

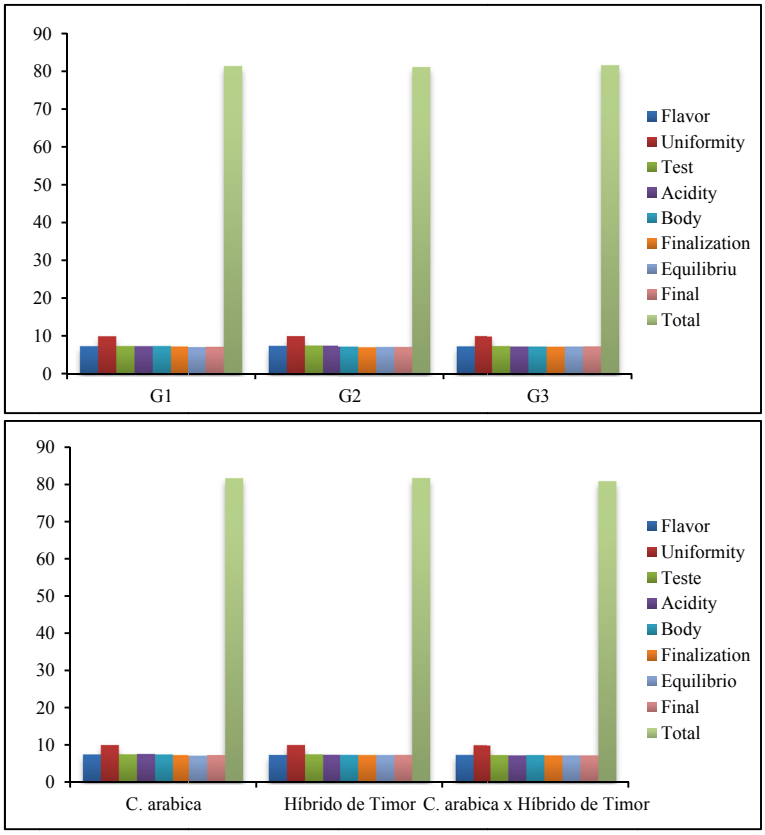

Figure 4. The bargraph of quality attributes for each groups of coffee genotypes clustered using the UPGMA clustering method based on SSR marker and different groups of coffee (*C. arabica*, *Híbrido de Timor*, *C. arabica* × Híbrido de Timor)

4. Discussion

4.1 Genetic Relationship of "HT" With Other Coffee Accessions

The AMOVA analysis using AFLP molecular marker revealed that more variation was partitioned between coffee groups (Gst = 0.608) than within groups of coffee. Similar results were also reported using RAPD molecular marker in coffee species (Silvestrini et al., 2008). The pairwise F_{st} estimated between coffee groups also demonstrated that "HT" is more related to *C. arabica*, which is in accordance with result reported by Lashermes et al. (1993, 1996, 2000). This is the principal characteristics of autogamous plants where they maintain homozygosity at all loci that result low heterozygosity within the population. The multivariate statistical analysis (Principal coordinate analysis), clustering analysis and population genetic analysis clearly showed similar results and proved the existence of well-defined population structure among coffee species. In addition, it showed high genetic similarity between *C. arabica* and "HT". The result also confirmed that the "HT" is the result of interspecific hybridization of *C. arabica* and *C. canephora* as reported by Lashermes et al. (2000).

The model based Bayesian cluster analysis using Structure program (Pritchard et al., 2000) was frequently used to study the population structure and genetic diversity in different crops (Holsinger & Wallac, 2004; Kwak et al,. 2009; López-Gartner et al., 2009). In this study, the Bayesian model was also used and showed that *C. arabica* and "HT" accessions have similar genetic architecture and grouped in Cluster III (blue) with shared ancestral probability > 0.92 (Figures 1 and 2). This result proved the existence of high genetic similarity between them. The neighbor joining tree produced by Structure software (Table 3, supplement) also proved the proximity of *C. arabica* and *C. canephora* var Robusta to "HT" that affirmed that "HT" is the interspecific hybrid of *C. arabica* and *C. canephora* var Robusta.

Cluster I include the clones of Conilon with shared ancestral probability > 0.95, indicating high uniformity among these clones. The highest admixture probability was observed within accessions of *C. canephora* var Robusta that demonstrated the existence of high genetic diversity within this group. The clustering analysis based on Bayesian statistics proved to be an efficient method in assigning genotypes in its respective group. López-Gartner et al. (2009) used the same method and also grouped *C. arabica* in its respective geographic locations. Our study also showed the efficiency of the Bayesian model-based clustering method in grouping the population used in this study.

The individual population diversity measures estimated (Shanon Inoformation index, Nei genetic diversity and percent polymorphism) showed that *C. canephora* var Robusta has the highest genetic diversity followed by "HT" and *C. canephora* var Conilon. Similar results also reported by Lashermes et al. (1993, 2000), Orozco-Castillo (1994) and Silvestrini et al. (2008).

The high genetic similarity between *C. arabica* and "HT" demonstrated by different statistical analysis methods proved that the selection of "HT" accessions with similar architecture and agronomic characters with *C. arabica* in the breeding programs influenced the genetic composition of "HT". In addition, high genetic similarity of "HT" and *C. arabica* can be interpreted from the origin of "HT", where 2/3 of its genome came from *C. arabica*. The directional selection and backcrossing with *C. arabica* reduced the introgressed alien genetic material from *C. canephora* but maintained the genes responsible for resistance to disease and pests, which are important for the breeding programs. The low genetic differentiation between *C. arabica* and "HT" (Figure 2) indicated the high gene flow between these two groups of coffee.

In addition, our result showed the existence of considerable genetic diversity among "HT" accessions. The existence of this genetic diversity within "HT" has great importance in the improvement program of *C. arabica*. Since this coffee group is the most important coffee group extensively used in the breeding programs as a source of resistance gene for coffee leaf rust (*Hemileia vastatrix*), coffee berry disease (*Colletotrichum kahawae)*, root knot nematode (*Meloidgyne exigua*) and bacteriosis (*Pseudomonas syringae* pv garçae) (Bettencourt, 1973; Charrier and Eskes, 1997; Bertrand et al., 2003; Pereira et al., 2005; Sera et al., 2005). Besides, this group of coffee has been used to develop high quality cup cultivars resistant to pest and disease in Brazil (Setotaw et al., 2013).

4.2 Genome Introgression Analysis of "HT"

The genome introgression analysis demonstrated that the maximum percentage of shared band between *C. canephora* var Robusta with "HT" was 30% and 20% for AFLP and SSR marker, respectively, suggesting high similarity with *C. arabica*. The genome introgression analysis of individual "HT" accessions show very low genome introgression from *C. canephora* that showed the possibility of identifying "HT" accession to be used in the crossing programs of coffee Arabica in the breeding program from the Brazilian collections.

CIFC 4106 shared more bands with *C. arabica* and "HT" than *C. canephora* var Robusta, supported by high genetic similarity between these coffee groups. The "HT" accession showed only 18.9% of alien genetic material introgressed from *C. canephora,* this percentage was lower than the expected percentage for F_1 triploid plant. According to Pereira et al. (2005), CIFC 4106 is considered the first plant, probably F_1, obtained in Timor Island and has some features that suggest that it is an interspecific hybrid. CIFC 4106 shows high flowering and low fruiting capacity, produces fruit type Moca and shows high self-incompatibility under the Viçosa soil and climatic condition. In addition, it did not produce fruit when used as female parent (Personal Field observation).

These results showed the importance of "HT" in the future *C. arabica* breeding in Brazil and the world. The results obtained widen the information on "HT" and their predecessors, which have great importance in the breeding program of coffee. The low percentage of *C. canephora* introgression reported here suggest that the Híbrido de Timor accessions maintained their Arabica quality, which is good to develop cultivars with good cup quality. Bertrand et al. (2003) reported the possibility of developing new cultivars with good cup quality and disease resistance from *C. canephora* introgressed lines. The high genetic variability of "HT" with its resistance

to disease and pests demonstrated the future potential of this coffee group in developing *C. arabica* cultivars (Pereira et al., 2005; Bertrand et al., 2003). Sobreira et al. (2013) showed the potential of "HT" in developing high quality *C. arabica* cultivars in Brazilian condition. In addition, "HT" will continue to be used to transfer resistance gene of coffee leaf rust and other diseases to *C. arabica* since this coffee group is highly compatible with *C. arabica* during crossing. Van der Vossen (2009) also showed the success of *C. arabica* breeding in developing resistant cultivar with quality using accessions of "HT" as source of resistant gene without affecting the coffee quality.

The incorporation of four CIFC materials (CIFC 4106, CIFC 832/1, CIFC 832/2 and CIFC 1343/269) and 42 segregating accessions of "HT" in this study showed the better representation of "HT" derivatives in relation to past works. Therefore, the experiment was more informative in the genome introgression analysis of "HT" in relation to *C. arabica* and *C. canephora* and its relationship with other coffee accessions. The obtained results are useful to choose the best accessions, the ones with lower *C. canephora* introgression, to be introduced in the *C. arabica* breeding programs.

The high genetic similarity observed within accessions of "HT" in relation to *C. arabica* and high genetic diversity within accessions of "HT" showed the importance of "HT" in the development of *C. arabica* cultivars with resistance to disease and high coffee quality.

4.3 The Genetic Relationship and Sensorial Analysis of C. arabica Cultivars Derived From "HT"

The quality analysis of three groups of coffee (*C. arabica*, "HT" and Cultivars derived from cross of *C. arabica* and "HT") showed the absence of negative effect of *C. canephora* genome introgression on the cup quality of coffee cultivars developed in Brazil. The result showed no significant difference among the coffee groups on sensorial analysis. The quality data obtained in this work (Figure 4) supported the conclusion made by the genome introgression analysis using AFLP marker and SSR markers. In addition, this result confirmed the possibility of developing *C. arabica* cultivars using "HT" as a source of resistant gene, without affecting the quality in contrast to most people's believe the genome introgressed from *C. canephora* affect the quality of coffee cultivars. Sobreira et al. (2015) also showed the importance of "HT" in developing quality *C. arabica* cultivars without influencing the cup quality.

Our result also showed the absence of significant difference in quality parameters presented (Figure 4) among the cultivars derived from the crossing program of "HT" × *C. arabica* with Bourbon one of the principal arabica cultivar known for its quality. The diversity analysis using UPGMA clustering method grouped the Arabica cultivars and cultivar derived from the crossing of *C. arabica* × "HT" in the same group confirmed these groups are highly similar genetically and also in quality (Figure 4). The absence significant difference among the different coffee groups (Figure 4) affirmed the possibility of developing *C. arabica* cultivars with good cup quality as Bourbon using "HT" as a source of resistance gene. In addition, it also revert the conclusion made by some groups the use of "HT" affect the quality of *C. arabica* cultivars developed using "HT" in the crossing programs.

The review about developing resistant coffee cultivars using "HT" as source of resistant gene that combine the good cup quality presented by Van der Vossen (2009) disclosed its impact on environment protection by reducing the quantity of fungicide applied and make the farmers more competitive. Especially the use of resistant varieties will help the farmers of the developing nations those did not have financial capacity to use fungicide where coffee is the only source of income. Therefore, our work showed the importance of "HT" in developing cultivars of *C. arabica* resistant to diseases and pests without affecting the cup quality which is an important parameter for coffee. It also proved the importance of "HT" and the effort of the breeding programs using "HT" as source genes resistance to disease and pest by releasing more productive and quality coffee cultivars for the commercial production. The study also draw conclusion using molecular marker analysis (AFLP and SSR) and sensorial analysis in relation to the impact of "HT" in cup quality of *C. arabica* cultivars released using "HT" as source of resistant gene for major diseases and pests.

5. Conclusion

The hypothesis that the varieties developed through introgression of "HT" genome as a source of resistance with the Arabic coffee was accepted. The recently released varieties from such crosses proved have similar cup quality as that of the well-known *C. arabica* coffee variety Burbon. This shows the importance of future investigations including more genotypes and marker types.

References

Bertrand, B., Guyot, B., Anthony, F., & Lashermes, P. (2003). Impact of the *Coffea canephora* gene introgression on the beverage quality of *C. arabica. Theorotical and Applied Genetics, 107*, 387-394. https://doi.org/10.1007/s00122-003-1203-6

Bettencourt, A. (1973). Considerações gerais sobre o 'Hibrido de Timor'. *Instituto Agronomico de Campinas, Circular No. 23* (p. 20). Campinas, Brasil.

Brito, G. G., Caixeta, E. T., Gallina, A. P., Maciel-Zambolim, E., Zambolim, L., Diola, V., & Loureiro, M. E. (2010). Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. *Euphytica, 173*, 255-264. https://doi.org/10.1007/s11295-011-0406-2

Charrier, A., & Eskes, A. (1997). Coffee. In A. Charrier, M. Jacquot, S. Hamon, & D. Nicolas (Eds.), *Tropical Plant Breeding* (pp. 128-152). Coll. Repères, Cirad, France.

Clarindo, W. R., & Carvalho, C. R. (2008). First Coffea arabica karyogram showing that this species is a true allotetraploid. *Plant Systematics and Evolution, 274*, 237-241. https://doi.org/10.1007/s00606-008-0050-y

Combes, M. C., Andrzejewski, S., Anthony, F., Bertrand, B., Rovelli, P., Graziosi, G., & Lashermes, P. (2000). Characterization of microsatellite loci in Coffea arabica and related coffee species. *Molecular Ecology, 9*(8), 1178-1180. https://doi.org/10.1046/j.1365-294x.2000.00954-5.x

Diniz, L. E. C., Sakiyama, N. S., Lashermes, P., Caixeta, E. T., Oliveira, A. C. B., Zambolim, E. M., … Zambolim, L. (2005). Analysis of AFLP marker associated to the *Mex-1* resistance locus in Icatu progenies. *Crop Breeding and Applied Biotechnology, 5*, 387-393. https://doi.org/10.12702/1984-7033.v05n04a03

Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359-361 https://doi.org/10.1007/s12686-011-9548-7

Ehrich, D. (2006). AFLPdat: A collection of R functions for convenient handling of AFLPdata. *Mol. Ecol. Notes, 6*, 603-604. https://doi.org/10.1111/j.1471-8286.2006.01380.x

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol., 14*, 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Laval, G., & Schneider, S. (2006). *Arlequin ver. 3.1: An integrated software package for population genetics data analysis*. https://doi.org/10.1177/117693430500100003

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Mol. Ecol. Notes, 7*, 574-578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Ferrão, L. F. V., Caixeta, E. T., Souza, F. F., Zambolim, E. M., Cruz, C. D., Zambolim, L., & Sakiyama, N. S. (2013). Comparative study of different molecular markers for classifying and establishing genetic relationships in *Coffea canephora*. *Plant Systematics and Evolution, 299*(1), 225-238.

Holsinger, K. E., & Wallac, L. (2004). Bayesian approaches for the analysis of population genetic structure: An example from *Platanthera leucophaea* (Orchidaceae). *Molecular Ecology, 13*(4), 887-894. https://doi.org/10.1046/j.1365-294X.2002.01512.x

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Societé Vanddoise des Sciences Natureles, 44*, 223-270.

Kwak, M., Kami, J. A., & Gepts, P. (2009). The putative Mesoamerican domestication center of Phaseolus vulgaris is located in the Lerma-Santiago basin of Mexico. *Crop Sci, 49*, 554-563. https://doi.org/10.2135/cropsci2008.07.0421

Lashermes, P., Andrzejewski, S., Bertrand, B., Combes, M. C., Dussert, S., Graziosi, G., … Anthony, F. (2000). Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theor Appl Genet, 100*, 139-146. https://doi.org/10.1007/s001220050019

Lashermes, P., Combes, M. C., Robert, J., Trouslot, P., D'Hont, A., Anthony, F., & Charrier, A. (1999). Molecular characterization and origin of the *Coffea arabica* L. genome. *Mol Gen Genet, 261*, 259-266. https://doi.org/10.1007/s004380050965

Lashermes, P., Cros, J., Combes, M. C., Trouslot, P., Anthony, F., Hamon, S., & Charrier, A. (1996). Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus Coffea L. *Theor. Appl. Genet., 93*, 626-632. https://doi.org/10.1007/BF00417958

Lashermes, P., Cros, J., Marmey, P., & Charrier, A. (1993). Use of random amplified DNA marker to analyse genetic variability and relationships of *Coffea* species. *Gen Res Crop Evolution, 40*, 91-99. https://doi.org/10.1007/BF00052639

López-Gartner, G., Cortina, H., McCouch, S. R., & Moncada, M. D. P. (2009). Analysis of genetic structure in a sample of coffee (*Coffea arabica* L.) using fluorescent SSR markers. Tree Genetics & Genomes 5:435–446. https://doi.org/10.1007/s11295-008-0197-2

Missio, R. F., Caixeta, E. T., Zambolim, E. M., Pena, G. F., Ribeiro, A. P., Zambolim, L., … Sakiyama, N. S. (2009). Assessment of EST-SSR markers for genetic analysis on coffee. *Bragantia, Campinas, 68*(3), 573-581. https://doi.org/10.1590/S0006-87052009000300003

Mueller, U. G., & Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. *Trend in Ecology and Evolution, 14*(10), 389-394. https://doi.org/10.1016/S0169-5347(99)01659-6

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA, 70*, 3321-3. https://doi.org/10.1073/pnas.70.12.3321

Orozco-Castillo, C., Chalmers, K. J., Waugh, R., & Powell, W. (1994). Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. *Theor Appl Genet, 87*, 934-940. https://doi.org/10.1007/BF00225787

Peakall, R., & Smouse, P. (2006). Genalex 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes, 6*(1), 288-295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Pereira, A. A., Oliveira, A. C. B., & Sakiyama, N. S. (2008). Híbrido de Timor como fonte de resistência a doenças e de qualidade da bebida do cafeeiro. *Núcleo de Estudos em Fitopatologia. Manejo fitossanitário da cultura do cafeeiro* (pp. 13-24). Brasília: Sociedade Brasileira de Fitopatologia.

Pereira, A. A., Sakiyama, N. S., Zambolim, L., Moura, W. M., Zambolim, E. M., & Caixeta, E. T. (2005). Identification and use of sources of durable resistance to coffee leaf rust in the UFV/EPAMIG breeeding program. In L. Zambolim, E. M. Zambolim, & V. M. P. Várzea (Eds.), *Durable resistance to coffee leaf rust* (pp. 215-232). Viçosa, UFV.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945-959.

Rohlf, F. J. (2000). *NTSYS-pc Numerical taxonomy and multivariate analysis system* (Version 2.1). Setauket, New York: Exeter Software.

Rovelli, P., Mettulio, R., Anthony, F., Anzueto, F., Lashermes, P., & Graziosi, G. (2000). Microsatellites in *Coffea arabica* L. *Coffee biotechnology and quality* (pp. 123-133). Springer, Netherlands. https://doi.org/10.1007/978-94-017-1068-8_9

Sera, T., Sera, G. H., Ito, D. S., & Doi, D. S. (2005). Coffee breeding for durable resistance to leaf rust disease at Instituto Agronômico do Paraná. In L. Zambolim, E. M. Zambolim, & V. M. P. Várzea (Eds.), *Durable resistance to coffee leaf rust* (pp. 187-214). Viçosa, UFV.

Setotaw, T. A., Caixeta, E. T., Pereira, A. A., Oliveira, A. C. B., Cruz, C. D., Zambolim, E. M., … Sakiyama, N. S. (2013). Coefficient of Parentage in Coffea arabica L. Cultivars Grown in Brazil. *Crop Science, 53*(4), 1237-1247. https://doi.org/10.2135/cropsci2012.09.0541

Silva, R.A., Zambolim, L., Castro, I.S.L. et al. The Híbrido de Timor germplasm: identification of molecular diversity and resistance sources to coffee berry disease and leaf rust. Euphytica 214, 153 (2018). https://doi.org/10.1007/s10681-018-2231-2

Silvestrini, M., Maluf, M., Silvarolla, M. B., Guerreiro-Filho, O., Medina-Filho, H. P., Vanini, M. M. T., … Fazuol, L. C. (2008). Genetic diversity of a Coffea Germplasm Collection assessed by RAPD markers. *Genet Resour Crop Evol, 55*, 901-910. https://doi.org/10.1007/s10722-007-9295-5

Sobreira, F. M., Oliveira, A. C. B., Pereira, A. A., & Sakyiama, N. S. (2015). Potential of Híbrido de Timor germplasm and its derived progenies for coffee quality improvement. *Australian Journal of Crop Science, 9*(4), 289-295.

Van der Vossen, H. A. M. (2009). The cup quality of disease-resistant cultivars of Arabica coffee (*Coffea arabica*). *Expl .Agric., 45*, 323-332. https://doi.org/10.1017/S0014479709007595

Vekemans, X., Beauwens, T., Lemaire, M., & Roldan-Ruiz, I. (2002). Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol. Ecol., 11*, 139-151. https://doi.org/10.1046/j.0962-1083.2001.01415.x

Yeh, F. C., & Boyle, T. J. B. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian J. Bot., 129*, 157.

Abbreviations

AFLP: Amplified Fragment length polymorphism; RAPD: Randomly Amplified polymorphism DNA; DNA: Deoxyribonucleic acid; QTL: Quantitative trait loci; SSR: Simple Sequence Repeat; Tri-HCL: Tris Hydrochloride acid; EDTA: Ethylenediaminetetraacetic acid; PCR: Polymorphic Chain reaction; dNTps: deoxyribonucleotide triphosphate; AMOVA: Analysis of molecular variance; PCoA: Principal coordinate analysis.