RESEARCH ARTICLE

Capillary electrophoresis as a tool for genotyping S_H3 mediated coffee leaf rust resistance

Savina A. Gutiérrez-Calle[1] (iD); Rosa A. Sánchez-Díaz[2] (iD); Yolanda B. Delgado-Silva[1] (iD); Juan D. Montenegro[1] (iD); Dina L. Gutiérrez[1] (iD); Jorge L. Maicelo-Quintana[1] (iD); Juan C. Guerrero-Abad[1] (iD)

[1] Centro Experimental La Molina. Dirección de Recursos Genéticos y Biotecnología. Instituto Nacional de Innovación Agraria (INIA). Av. La Molina 1981, 15024 Lima. Peru.
[2] Departamento de Fitotecnia. Facultad de Agronomía. Universidad Nacional Agraria La Molina, La Molina. Av. La Molina s/n, 15024 Lima. Peru.

* Corresponding author: jguerreroa@inia.gob.pe, sdb@inia.gob.pe (J. C. Guerrero-Abad).

Abstract

Coffee is an important agricultural commodity in the world. However, it is susceptible to *Hemileia vastatrix* (*Hv*), an obligatory biotrophic fungus that causes coffee leaf rust (CLR). Natural resistance to rust has been identified in the wild species *Coffea canephora* and *Coffea liberica*. These species have been used in breeding programs where interspecific resistant hybrids have been generated. The S_H3 gene, derived from *C. liberica*, has been shown to confer extreme and long-lasting resistance to *Hv*. A total of 167 accessions of the INIA's Coffee Germplasm Collection of Peru (INIA-CGC) were screened with 4 markers linked to the S_H3 gene. As positive controls, EA67 (*C. liberica*) and the hybrid S.288 (*C. arabica* x *C. liberica*) were used. Separation of PCR products was done by capillary electrophoresis, which allow to discriminate the alleles of each marker. For three markers, specific alleles for either *C. arabica* or *C. liberica* species were found. In all cases, S.288 exhibited specific alleles for both species; whereas the INIA-CGC accessions had exclusively *C. arabica* alleles and EA67 had *C. liberica* alleles. The BA-48-21O-f marker did not produce PCR fragments for any of the positive controls, suggesting that this marker is not as predictive as the other three to determine the presence of S_H3. This work reports the existence of multiple alleles for the Sat244 marker; however, the collection does not have the S_H3 mediated-resistance gene. Finally, the utility of capillary electrophoresis as a tool to identify alleles linked to S_H3 was demonstrated.

Keywords: coffee; capillary; electrophoresis; *Hemileia vastatrix*; S_H3.

1. Introduction

Peru is the eighth largest coffee producer in the world (*International Coffee Organization - Historical Data on the Global Coffee Trade,* 2020) and the United States is the main export destination (*Cámara Peruana del Café y Cacao: Café - Datos,* 2020; Nolte, 2020). Peruvian coffee is produced in 10 regions between 600 and 1,800 meters above sea level (m.a.s.l.), with a total crop extension of 375,000 ha (Junta Nacional del Café, 2019).

Nowadays, Peruvian coffee production is recovering from coffee leaf rust (CLR) caused by *Hemileia vastatrix* (*Hv*), which affected 50 percent of the 2013/14 harvest (Nolte 2020; McCook and Vandermeer 2015). Faced with the Peruvian coffee crisis, National Institute of Agricultural Innovation (INIA) established a Coffee Germplasm Collection (INIA-CGC) to provide valuable information for its

genetic conservation and breeding programs. INIA-CGC located at Pichanaki Agricultural Experimental Station in Junin, it maintains 169 coffee accessions collected from the main coffee growing regions: Amazonas, Cajamarca, Huanuco, Junin, Pasco and Ucayali.

Hv is a mandatory biotrophic fungus that specifically attacks coffee plants (Talhinhas et al., 2017) and causes premature defoliation in several species of the genus *Coffea*, of which *C. arabica* is the most susceptible (Silva et al., 2018). The success in the infection of *Hv* is due to its large genetic diversity that manages to overcome the resistance of its host. The different races of *Hv* are classified according to their ability to infect different coffee materials that have different combinations of resistance genes (Talhinhas et al., 2017). Currently, the Centro de Investigação das Ferrugens do Cafeeiro (CIFC) of the Instituto Superior de Agronomia (ISA) - Universidade de

Lisboa (U. Lisboa) has around 3,600 samples of *Hv* from 40 coffee-producing countries, which are classified in more than 50 breeds of *Hv* for a selected range of 23 differential coffee plants (***Genetic resources, Instituto Superior de Agronomia, 2020***). The CIFC breed nomenclature is attributed to isolates with distinct and unique combinations of nine virulence genes as inferred in gene-by-gene theory and are described as sequential Roman numerals in order of their detection (**Talhinhas et al., 2017**). For example: race I has a virulence gene (V2,5); race II (V5); race III (V1,5); race X (V1,4,5); race XIII (V5,7); race XVII (V1,2,5); race XXII (V5,6); race XXIV (V2,4,5); race XXVIII (V2,4,5,6); race XXXVI (V2,4,5,8); race XXXVII (V2,5,6,7,9) and race XXXIX (V2,4,5,6,7,8,9). Some of the races mentioned before were reported in Brazil (**Zambolim, 2016**); Costa Rica and Guatemala (**Avelino et al., 2015**).

Up to 9 dominant resistance genes have been identified in the *Coffea* genus (S_H1 to S_H9) and correspond to the nine *Hv* virulence genes (**Mahé et al., 2008**). Of these resistance genes, S_H1, S_H2, S_H4 and S_H5 have been found in *C. arabica*, but they have not provided lasting resistance to most *Hv* races. The other resistance genes, S_H6, S_H7, S_H8 and S_H9, found in *C. canephora*, have been introgressed into *C. arabica*, to obtain resistant hybrids (**Mahé et al., 2008**). Finally, the S_H3 gene found in *C. liberica* (**Prakash et al., 2004**) seems to provide lasting resistance to the coffee plants that have it.

Following the discovery of a rust-resistant coffee plant on the island of Timor (Timor Hybrid, a spontaneous interspecific hybrid between *C. arabica* and *C. canephora*), coffee improvement programs were developed using this hybrid as the main gene source to obtain varieties resistant to coffee rust (**Talhinhas et al., 2017**; **Romero et al., 2014**). Despite of their early success which allowed the release of highly resistant and productive cultivars, most of them became susceptible due to the appearance of more aggressive strains of *Hv* (**Talhinhas et al., 2017**; **Cabral et al., 2016**).

Another natural hybrid (S.26) between *C. arabica* and an unknown diploid coffee species was reported in India. This hybrid was used as the main source of resistance to rust in coffee improvement programs in India, from which the hybrid S.795 showed high levels of tolerance in the field for races I and II of *Hv*. More recently, it has been revealed that coffee plants from India had the S_H3 resistance gene for coffee rust from *Coffea liberica* (**Prakash et al., 2002**).

Numerous studies have used molecular markers to understand the introgression of *C. canephora* or *C. liberica* resistance genes (**Ribas et al., 2011**; **Prakash et al., 2002**; **Lashermes et al., 2000**). These markers allow the rapid and early identification of plants carrying *Hv* resistance genes through marker-assisted selection (MAS). Moreover, the use of molecular markers accompanied with a high-throughput separation method as capillary electrophoresis (CE) have allowed the develop of a rapid procedure to identify alleles (**Durney et al., 2015**).

Among the molecular markers associated with resistance to coffee leaf rust, **Mahé et al. (2008)** reported SCAR, SSR and BAC markers developed from populations that carry the S_H3 gene. These markers, used by **Alkimim et al. (2017)**,

Valencia et al. (2017) and **González-Martínez et al. (2009)**, are able to discriminate presence and absence of the S_H3 gene in different hybrid progenies of *C. liberica* and *C. arabica*.

In this work, four markers linked to the S_H3 resistance gene (**Alkimim et al., 2017**; **Valencia et al., 2017**; **González-Martínez et al., 2009**; **Mahé et al., 2008**) were used to screen the accessions from INIA-CGC, and identify potential accessions carrying the S_H3 gene in order to implement a coffee breeding program in Peru.

2. Materials and methods
Biological material
Coffee samples from INIA-CGC. The coffee germplasm collection has 169 accessions, from which 102 accessions had variety data (denomination according to producers, **Figure 1**) and 67 accessions without variety denomination (Amazonas:1; Cajamarca: 5; Huanuco: 2; Junin: 34 and Pasco: 25 accessions).
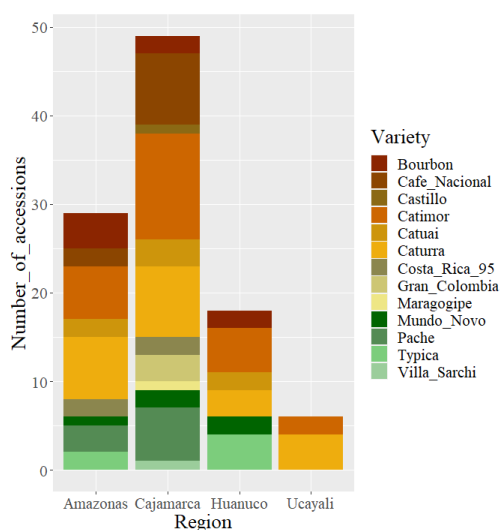

Figure 1. Geographic distribution of coffee varieties from INIA's Coffee Germplasm Collection. Variety denomination according to producers.

The method used to collect the samples consisted of taking leaves of the superior third of three plants (biological replicates) per accession; except for the accessions CA-164 and CA-165 that have only two biological replicates. A total of 505 samples were collected, stored with silica gel, and transported to INIA's Laboratory of Molecular Biology and Genomics (Lima, Peru) for further sample processing.

DNA extraction. Genomic DNA was isolated from dehydrated young leaf tissue by two methods: the CTAB method (**Doyle & Doyle 1987**) and a method provided by Dr. Lashermes from IRD-France (personal communication), for those samples that had poor quality after extraction with the CTAB method. Dr. Lashermes' protocol used 75 mg of dehydrated leaves that were snap-frozen with liquid nitrogen and ground, then 400 µL of the extraction buffer was added (1.5 M NaCl; 200 mM Tris-HCl, pH 8; 40 mM EDTA, pH 8; 3% sarkosyl; 1% sodium bisulfite; 2% MATAB, the last reagent was added before using the

buffer) and 400 µL of the lysis buffer (350 mM sorbitol; 200 mM Tris-HCl, pH 8; 40 mM EDTA, pH 8; 2% PVP, the last reagent was added just before using the buffer) and incubated at 62 °C for 30 min with regular mixing by inversion. After incubation, the mix was cooled to room temperature and 1 ml of chloroform: isoamyl alcohol solution (24:01) was added and mixed then was centrifuged for 5 min at 12,000 rpm. The supernatant was transferred to new 2 ml tubes and 20 µL of RNAse (10 mg/ml) was added, then placed on a heat block at 37 °C for 30 min. After this period, a volume of isopropyl alcohol was added to precipitate the DNA and centrifuged at 12,000 rpm for 5 min.

The supernatant was carefully discarded, and the DNA pellet was washed with 200 µL of ethanol (70%) and centrifuged at 12,000 rpm for 3 min. When finished, the ethanol was discarded, and the pellet could dry. Finally, the pellet was resuspended in 50 µL TE at 4 °C overnight and storage at -20°C. The quality of the DNA obtained from the two mentioned protocols was assessed by means of a 1% agarose gel with SYBR Safe (Invitrogen) in TAE 1X buffer, and it was visualized under UV light. The DNA obtained was quantified by spectrophotometry on an EPOCH (BiotekTM). In addition, DNA from EA67 (*C. liberica*, var. Liberica) and the hybrid S.288 (*C. arabica* x *C. liberica*), kindly provided by Dr. Lashermes, were used as positive controls for molecular markers linked to the S_H3 gene.

Marker Amplification and capillary electrophoresis.
Four molecular markers reported by **Mahé et al. (2008)** as linked to the S_H3 gene of *C. liberica* were used to screen 169 accessions of coffee plants using capillary electrophoresis. To do this, the 19 bp M13 tail (5'-CACGACGTTGTAAAACGAC-3') was added to the 5' end of every sense primer (**Table 1**). PCR conditions were similar to those described by **Mahé et al. (2008)**, with some modifications mentioned by **Schuelke (2000)**. PCR reactions were done in a 10 µL total volume which contained 40 ng of DNA, 1X Buffer Taq 10X (Thermo Fisher Scientific, USA.), 2 mM MgCl₂, 100 µM dNTP (Thermo Fisher Scientific, USA), 0.1 µM sense primer, 0.4 µM antisense primer, 0.4 µM fluorophore, 0.5 U of Taq polymerase enzyme (Thermo Fisher Scientific, USA). The fluorophores used were: FAM for the Sat244 marker, VIC for the BA124-12K-f marker, NED for the BA-48-21O-f

marker and PET for the SP-M8-SH3 marker. The PCR program included an initial denaturation stage at 95 °C for 3 min, followed by 30 cycles of denaturation at 95 °C for 45 s, hybridization at the appropriate temperature (**Table 1**) for 45 s, elongation at 72 °C for 45 s and a final extension stage at 72 °C for 10 min. In addition to the INIA-CGC accessions, positive controls (EA67 and S.288) were included in the screening for all markers.

The PCR products were separated by Capillary electrophoresis (CE). The reaction mix for CE for every sample was 9.7 µl of Hi-Di Formamide, 0.3 µL of GeneScan 600 LIZ Size Standard (Applied Biosystems) and 0.8 µL of each product. The samples were denatured for 3 min at 95 °C, and then cooled to -20 °C for 3 min. The CE was performed on an ABI Prism 3130XL Genetic Analyzer® (Applied Biosystems).

Data analysis
For the representation of the coffee varieties from INIA-CGC was used "ggplot2" R package (**Wickham, 2016**). The electropherograms were analyzed with GeneMapper v4.0 (Applied Biosystems). Nonspecific products and false positive spikes were removed, as indicated in the user guide. Possible allele size parameters for each molecular marker were established. Clean data was exported as tabulated text and analyzed with "adegenet" package (**Jombart, 2008**) for the allelic frequency; the function "dist.genpop" for calculate the genetic distance; and the "dendextend" package (**Galili, 2015**) for dendrogram graphic. All packages mentioned were used in Rstudio software. The scripts used were based on **Herrera (2019)**.

3. Results and discussion
DNA isolation
The CTAB method used for DNA extraction was successful in most of the accessions, the average concentration per sample was 1,138.67 ng/µL and a spectrophotometric 260/280 ratio of 2.10. Lashermes extraction method was used only in those samples with a high concentration of polyphenols and carbohydrates, the average concentration obtained was 298 ng/µL with a 260/280 ratio of 1.6. Two accessions (CA-116 and CA-117) showed low DNA quality with both the CTAB and Lashermes methods, therefore, they were not included in further analysis.

Table 1
Primers linked to resistance to Coffee Leaf Rust through the S_H3 gene

| Marker	Primer sequence (5' → 3')	Melting temperature	Fl*
Sat244	F-M13: CACGACGTTGTAAAACGACGCAT GTGCTTTTTGATGTCGT R: GCATACTAAGGAAATTATCTGACTGCT	56 °C	FAM
BA-124-12K-f	F-M13: CACGACGTTGTAAAACGACTGAT TTCGCTTGTTGTCGAG R: TGCAGATTGATGGCACGTTA	60 °C	VIC
BA-48-21O-f	F-M13: CACGACGTTGTAAAACGACACA GTGAATTCCCCAAGCAC R: ACTTGGCAGGCGTAATTGAA	60 °C	NED
SP-M8-SH3	F-M13: CACGACGTTGTAAAACGACGAA TTCAGCGACGATTG R: GATTTGGTGGAAGGGAGC	55 °C	PET

(*) Fluorophore.
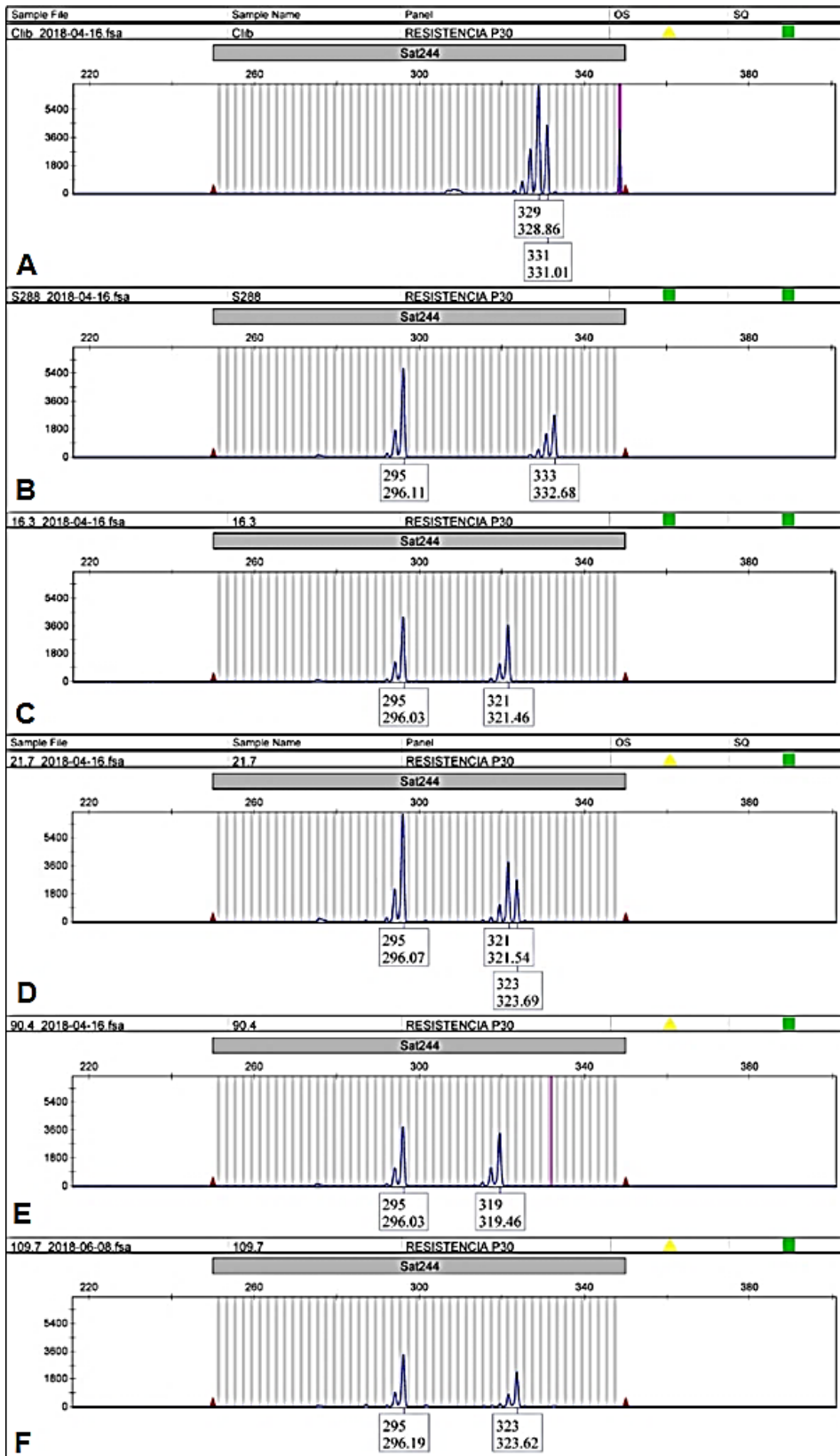
Figure 2. Sat244 electropherogram of EA67, *C. liberica* (A); S.288, *C. arabica* x *C. liberica* (B) and four accessions of INA-CGC: CA-16.3 (C); CA-21.7 (D); CA-90.4 (E) and CA-109.7 (F). The accessions CA-16.3; CA-21.7; CA-90.4 were samples with different genotype from CA-109.7, (295/323) that was the genotype most common in INIA´s Coffee Germplasm Collection.
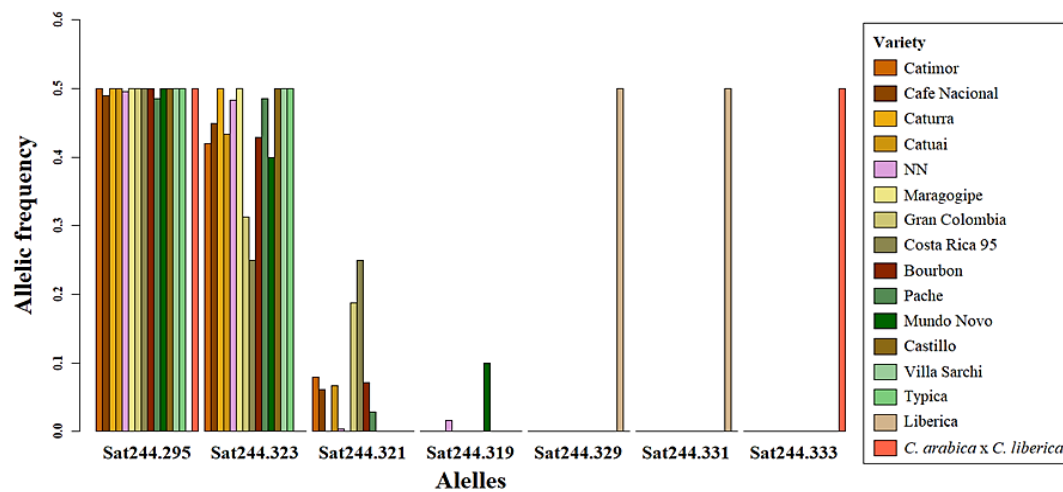
Figure 3. Allelic frequency of Sat244 marker in all samples represented by varieties. Variety denomination according to producers.

The CTAB method was effective for obtaining high-quality DNA from 80% of the processed samples. The remaining samples were rescued using the method developed by Dr. Lashermes. The main difficulty found on sample processing was the presence of mucilage and dark pigments in the DNA pellet that could affect its quality. Inglis et al. (2018) showed that samples with a high number of metabolites and polyphenols impair the quality of DNA extraction. To improve extraction quality, Dr. Lashermes modified the original extraction protocol and used MATAB, sarkosyl and sorbitol to remove polysaccharides (Azmat et al., 2012; Souza et al., 2012), and sodium bisulfite to prevent the action of nucleases during extraction (Byrne et al., 2001). This resulted in higher DNA purity but lower yield. However, for PCR-based genotyping methods, it is always preferable to have higher purity instead of higher yield.

Marker Amplification and capillary electrophoresis

A total of 167 accessions from the INIA-CGC were analyzed. Three biological replicates were amplified for 16 accessions (CA-1 to CA-16); 2 biological replicates for 136 accessions (CA-17 to CA-160) and 1 replicate for 15 accessions (CA-161 to CA-169) making a total of 335 samples processed.

Accessions of INIA-CGC were polymorphic with Sat 244 marker (Appendix). The collection showed 4 different alleles and 5 different genotypes for Sat244 marker (295 bp, 319 bp, 321 bp, and 323 bp; Figure 2). The most common genotype was 295/323 (303 samples, Figure 2F), followed by genotype 295/321 (21 samples, Figure 2C) and genotype 295/319 (5 samples, Figure 2E). Moreover, this marker displayed triallelic genotypes 295/321/323 (in 3 samples: CA-21.7, CA-47.3, CA-107.4, Figure 2D) and genotype 295/319/323 (in CA-97.5, figure not shown). Some samples of the biological replicates produced different genotypes among them (3.5% of the accessions evaluated), e.g., the CA-29 and CA-144 accessions showed genotypes 295/323 and 295/321. Of these alleles, only the 295 bp allele was shared with the hybrid S.288. The positive controls had exclusive alleles 329 bp, 331 bp (EA67, Figure 2A) and 333 bp (S.288, Figure 2B). Furthermore, the allelic frequency of Sat244 showed that

the accessions with variety denomination Costa Rica 95 and Gran Colombia had the allele 321 bp, and Mundo Novo the allele 319 bp (Figure 3).

The other three markers (BA-124-12K-f; SP-M8-S_H3 and BA-48-21O-f) were monomorphic for the INIA-CGC accessions. Only one allele (403 bp) was observed for BA-124-12K-f; alleles 253 bp and 259 bp for SP-M8-SH3 and the allele 332 bp for the BA-48-21O-f marker.

In all cases, the EA67 control (*C. liberica*) exhibited unique alleles different from those in the INIA-CGC (allele 349 bp for BA-124-12K-f and allele 257 bp for SP-M8-SH3, showed in Figure 4A and 5A, respectively). The hybrid S.288 control exhibited two alleles in all cases, one shared with the collection and the other shared with EA67 (349 bp and 403 bp for BA-124-12K-f; 253 bp and 257 bp for SP-M8-S_H3, showed in Figure 4B and 5B). Both positive controls did not amplify with BA-48-21O-f marker.

A variety dendrogram (Figure 6) based on the allelic frequency was obtained by the Nei's distance (Nei, 1978) and constructed through UPGMA clustering method. Varieties Costa Rica 95 and Gran Colombia (Variety denomination according to producers) were clustered in the same group. Moreover, the dendrogram showed the separation of Liberica (EA67); *C. arabica* x *C. liberica* (S.288) and all INIA-CGC accessions.

Molecular markers have been widely used to study genetic diversity of several species (Mohamed et al., 2019; Lee, 2019; Cueva-Agila et al., 2019), including coffee (Mishra, 2020). Moreover, molecular markers have been used to assist selection for agronomic characteristics (Pruvot-Woehl et al., 2020; Sousa et al., 2019; Alkimim et al., 2017; Nsabiyera et al., 2016; Mahé et al., 2008).

INIA-CGC is in the early stages of becoming a formal collection. For this reason, information about the accessions is still in process of manual curation that will include DNA fingerprinting and large-scale phenotyping. In this study, the marker Sat244 was able to discriminate between replicates (3%) of the accessions screened. Further genotyping with other microsatellites will assist in the identification of replicates that require reclassification. The use of novel genotyping techniques will facilitate curation by making it more reliable and reproducible. The capillary electrophoresis standardized in this study is a good

example of accurate and reproducible genotyping and will allow the screening of the collection and larger populations with other microsatellites already described in the literature (Sánchez et al., 2020; Andrianasolo et al., 2013).

The 4 markers used in this study were described by Mahé et al. (2008), these markers have been classified in SSR markers, SCAR markers obtained by bacterial artificial chromosome (BAC) genomic library and a SCAR from AFLP markers identified by Prakash et al. (2004). These markers were evaluated in two populations: a Matari x S.288 F2 mapping population and a backcross (BC2) [(Matari x S.288) x Matari] x Matari (Mahé et al., 2008; Prakash et al., 2004; Prakash et al., 2002). In this work, three of the four markers were useful to differentiate *C. liberica* control (EA67) from the hybrid *C. arabica* x *C. liberica* control (S.288) and the accessions from INIA-CGC.

Tornincasa (2008) used the Sat244 marker, named SAT-01-G10 (that had the same sequence), for variety analyses in *Coffea arabica*. In her work indicated that SAT-01-G10 could improve the process of variety identification of samples with a non-specific profile of the microsatellite's markers selected. Therefore, Sat244 marker clustered some *C. arabica* varieties of INIA-CGC as shown in Figure 6. The polymorphism generated by Sat244 in *C. liberica* plants was also reported by Mahé (2007); González-Martínez et al. (2009) and Alkimin et al. (2017) that observed different bands profiles in breeding population of *C. liberica*. Limited

information related to the molecular weights of the different alleles of Sat244 in *C. liberica*, makes it difficult to identify the one that co-segregates with S_H3 most frequently.

Surprisingly, the BA-48-21O-f marker could not be amplified in either positive controls, but it was amplified in the INIA-CGC. This marker was developed and evaluated in S.288 and in various mapping populations derived from S.288. Although the linkage map suggests that BA-48-21O-f is the most distant marker from S_H3 gene (Lashermes et al., 2010; Mahé et al., 2008), this does not explain the absence of an amplification product in S.288. The amplification conditions were similar for both, INIA-CGC samples and the two positive controls, which rules out the possibility of human errors. To dismiss any possible problems with capillary electrophoresis, the PCR products were also separated on 2.5% agarose gels (data not shown). As a marker belonging to the *C. arabica* genome (Ca subgenome), BA-48-21O-f could be close to the introgression / recombination point of the S_H3 locus. This would explain the absence of product in the EA67 control. For S.288, a possible determining factor was its DNA quality, which exhibited very low 260/280 and 230/260 ratios (data not shown), and agarose gel electrophoresis showed DNA degradation that could have affected the marker biding site. These factors could explain the absence of the BA-48-21O-f marker in S.288.
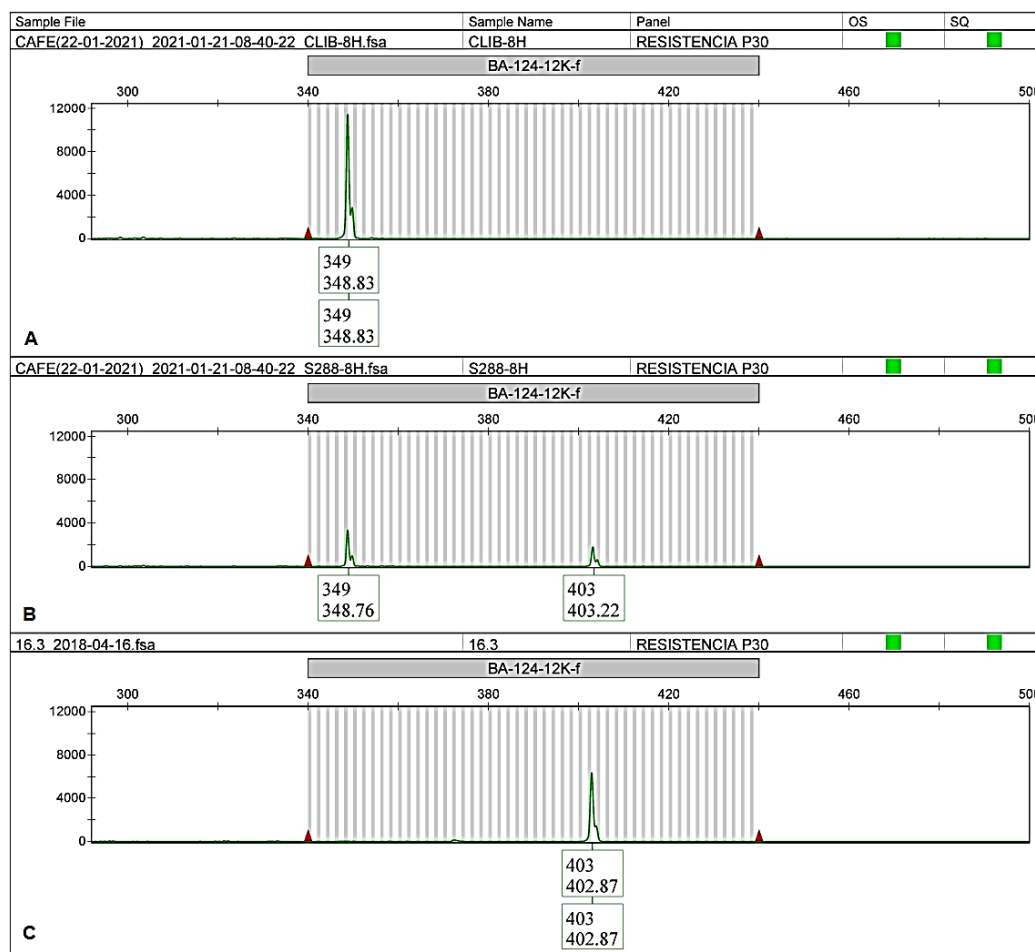

Figure 4. BA-124-12K-f electropherogram of EA67, *C. liberica* (A); S.288, *C. arabica* x *C. liberica* (B) and CA-16.3 from INA-CGC (C).
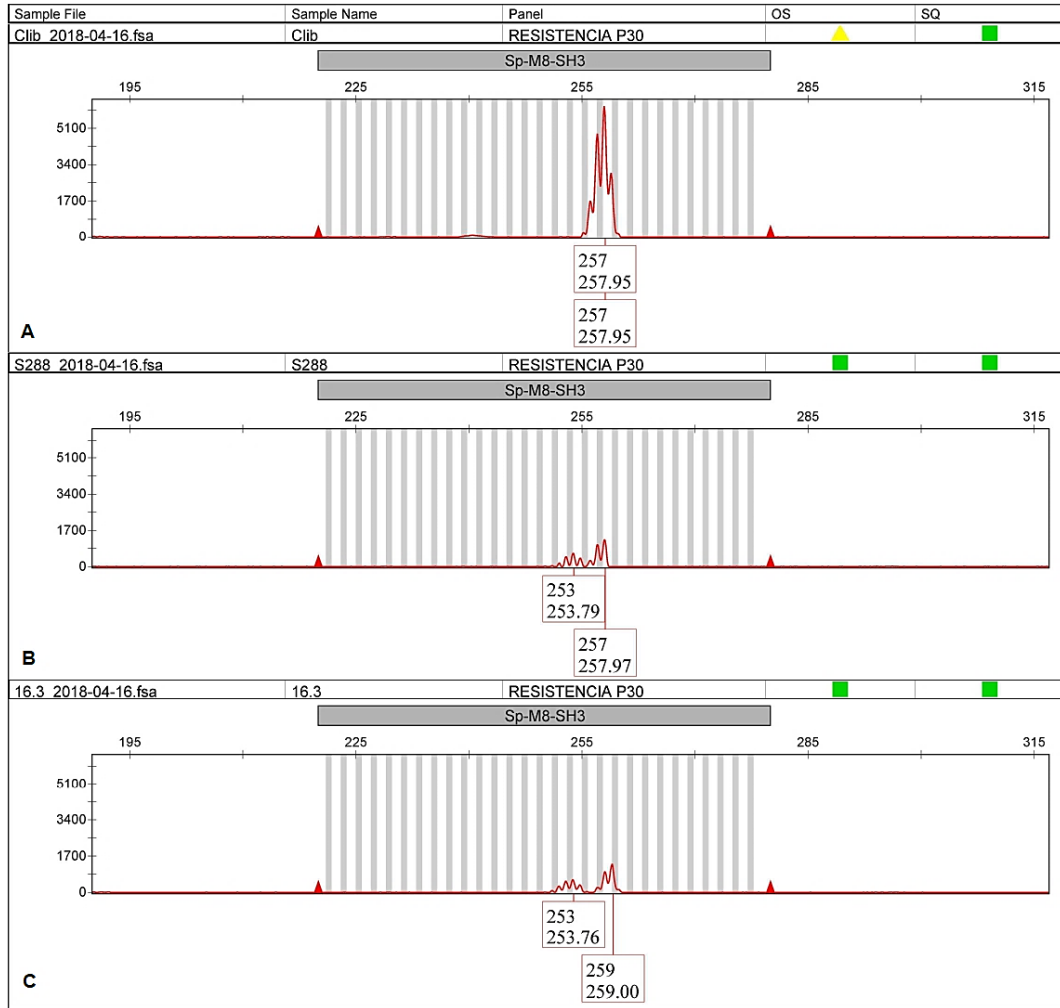
Figure 5. SP-M8-S$_H$3 electropherogram of EA67, *C. liberica* (A); S.288, *C. arabica* x *C. liberica* (B) and CA-16.3 from INA-CGC (C).
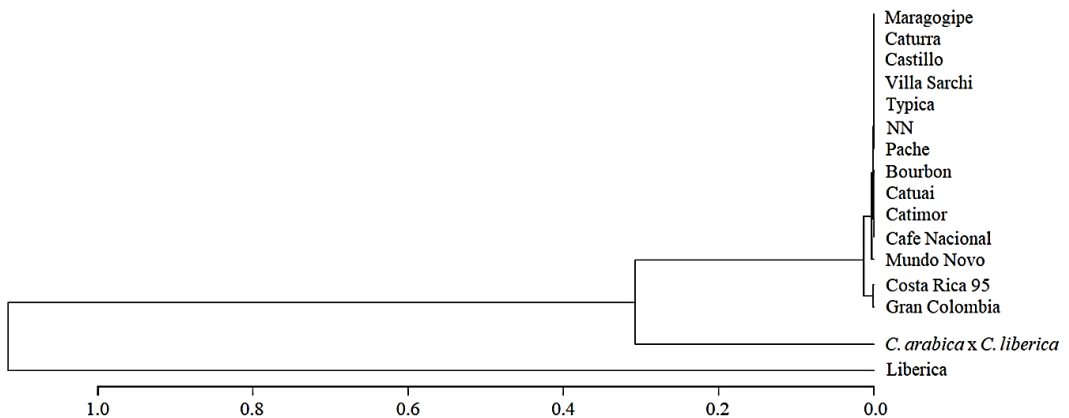

Figure 6. Variety dendrogram by Nei's distance and constructed through UPGMA clustering method. Variety denomination according to producers.

The BA-124-12K-f marker was used by **Alkimim et al. (2017)**, **Valencia et al. (2017)** and **González-Martínez et al. (2009)**. **Valencia et al. (2017)** only used BA-124-12K-f for their S$_H$3 screening in introgressed progenies and concluded that their evaluation can be useful for early selection of plants that carries S$_H$3 gene.

The SP-M8-SH3 marker was selected for screening the INIA-CGC accessions due to its proximity to the S$_H$3 gene, as shown in the linkage map done by **Mahé et al. (2008)**. The results of this study suggest that the best markers to determine the presence of the S$_H$3 gene in coffee collections are BA-124-12K-f and SP-M8-SH3 that clearly

discriminate INIA-CGC accessions from the hybrid S.288 and from *C. liberica* (EA67).

Taken together, these results demonstrate that none of the INIA-CGC accessions possess the S_H3 gene that confers resistance to *Hv*. There are internal reports of tolerance/resistance to *Hv* in accessions from the INIA-CGC. If this tolerance/resistance is confirmed that it was not introgressed from *C. liberica*.

4. Conclusions

This is the first study in Peru to demonstrate the usefulness of capillary electrophoresis to evaluate markers linked to the S_H3 gene which confers resistance to coffee leaf rust (*Hemileia vastatrix)* in 335 samples representing 167 accessions of the INIA-CGC. This work will facilitate the use of marker-assisted selection (MAS) in a massive way in large coffee populations and will accelerate the molecular characterization of coffee germplasm to identify samples carrying the S_H3 gene that could be used in breeding programs as a source of potential resistance.

ORCID

S.A. Gutiérrez-Calle [iD] https://orcid.org/0000-0003-1077-9153
R.A. Sánchez-Díaz [iD] https://orcid.org/0000-0003-2348-3968
Y.B. Delgado-Silva [iD] https://orcid.org/0000-0002-2593-3873
J.D. Montenegro [iD] https://orcid.org/0000-0003-4898-5761
D.L. Gutiérrez [iD] https://orcid.org/0000-0002-2649-0525
J.L. Maicelo-Quintana [iD] https://orcid.org/0000-0001-9109-0504
J.C. Guerrero-Abad [iD] https://orcid.org/0000-0002-7285-9506

References

Alkimim, E. R., Caixeta, E. T., Sousa, T. V., Pereira, A. A., de Oliveira, A. C. B., Zambolim, L., & Sakiyama, N. S. 2017. Marker-assisted selection provides arabica coffee with genes from other *Coffea* species targeting on multiple resistance to rust and coffee berry disease. *Molecular Breeding*, 37, 6.

Andrianasolo, D. N., Davis, A. P., Razafinarivo, N. J., Hamon, S., Rakotomalala, J. J., Sabatier, S. A., & Hamon, P. 2013. High genetic diversity of in situ and ex situ populations of Madagascan coffee species: Further implications for the management of coffee genetic resources. *Tree Genetics & Genomes*, 9, 1295-1312.

Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A. J., & Morales, C. 2015. The coffee rust crises in Colombia and Central America (2008–2013): Impacts, plausible causes and proposed solutions. *Food Security*, 7, 303-321.

Azmat, M. A., Khan, I. A., Cheema, H. M. N., Rajwana, I. A., Khan, A. S., & Khan, A. A. 2012. Extraction of DNA suitable for PCR applications from mature leaves of *Mangifera indica* L. *Journal of Zhejiang University. Science B*, 13, 239-243.

Byrne, M., Macdonald, B., & Francki, M. 2001. Incorporation of Sodium Sulfite into Extraction Protocol Minimizes Degradation of Acacia DNA. *BioTechniques*, 30, 742-748.

Cabral, P. G. C., Maciel-Zambolim, E., Oliveira, S. A. S., Caixeta, E. T., & Zambolim, L. 2016. Genetic diversity and structure of *Hemileia vastatrix* populations on *Coffea spp*. *Plant Pathology*, 65, 196-204.

Cámara Peruana del Café y Cacao: Café - Datos. 2020. Available in: https://camcafeperu.com.pe/ES/cafe-datos.php

Cueva-Agila, A., Vélez-Mora, D., Arias, D., Curto, M., Meimberg, H., & Brinegar, C. 2019. Genetic characterization of fragmented populations of *Cinchona officinalis* L. (Rubiaceae), a threatened tree of the northern Andean cloud forests. *Tree Genetics & Genomes*, 15(6), 81.

Doyle, J. J., & Doyle, J. L. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, 11-15.

Durney, B. C., Crihfield, C. L., & Holland, L. A. 2015. Capillary electrophoresis applied to DNA: Determining sequence and structure to advance bioanalyses (2009–2014). *Analytical and Bioanalytical Chemistry*, 407, 6923-6938.

Galili, T. 2015. Dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, 31, 3718-3720.

Genetic resources, Instituto Superior de Agronomia. 2020. Available in: https://www.isa.ulisboa.pt/en/cifc/genetic-resources

González-Martínez, L. F., Cortina-Guerrero, H. A., & Herrera-Pinilla, J. C. 2009. Validación de Marcadores Moleculares Ligados al Gen SH3 de Resistencia contra la Roya en Introducciones de la Colección Colombiana de Café. *Cenicafé*, 60, 375-389.

Herrera, S. 2019. Caracterización de la diversidad genética de variedades de albaricoquero mediante marcadores microsatélites (SSR). Thesis Máster Bioinformática y Bioestadística, Universitat Oberta de Catalunya. España. 42 pp.

Inglis, P. W., Pappas, M. de C. R., Resende, L. V., & Grattapaglia, D. 2018. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLOS ONE*, 13(10), e0206085.

International Coffee Organization-Historical Data on the Global Coffee Trade. 2020. [Data Base]. Available in: http://www.ico.org/new_historical.asp

Jombart, T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403-1405.

Junta Nacional del Café. 2019. La caficultura peruana está en riesgo por bajos precios y altos costos de producción. *El Cafetalero*, 62, 10-12.

Lashermes, P., Combes, M.C., Ribas, A., Cenci, A., Mahé, L., & Etienne, H. 2010. Genetic and physical mapping of the SH3 region that confers resistance to leaf rust in coffee tree (*Coffea arabica* L.). *Tree Genetics & Genomes*, 6, 973-980.

Lashermes, P., Andrzejewski, S., Bertrand, B., Combes, M. C., Dussert, S., Graziosi, G., Trouslot, P., & Anthony, F. 2000. Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theoretical and Applied Genetics*, 100, 139-146.

Lee, J. 2019. Development and Evolution of Molecular Markers and Genetic Maps in Capsicum Species. N. Ramchiary & C. Kole (Eds.), The Capsicum Genome. Springer International Publishing. Pp. 85-103.

Mahé, L., Combes, M. C., Várzea, V. M., Guilhaumon, C., & Lashermes, P. 2008. Development of sequence characterized DNA markers linked to leaf rust (*Hemileia vastatrix*) resistance in coffee (*Coffea arabica* L.). *Molecular Breeding*, 21(1), 105-113.

Mahé, L. 2007. Contribution à l'amélioration génétique de la résistance des caféiers (*Coffea arabica* L.) à la rouille (Hemileia vastatrix): De l'étude des hybrides interspécifiques naturels de Nouvelle-Calédonie à la cartographie d'un locus de résistance. Tesis Doctoral. École nationale supérieure agronomique (Montpellier). Francia. 135 pp.

McCook, S., & Vandermeer, J. 2015. The Big Rust and the Red Queen: Long-Term Perspectives on Coffee Rust Research. *Phytopathology*, 105, 1164-1173.

Mishra, M. K. 2020. Genetic Resources and Breeding of Coffee (*Coffea spp*.) | SpringerLink. En Advances in Plant Breeding Strategies: Nut and Beverage Crops (Al-Khayri J.; Jain S.; Johnson D. (Eds). Springer. Pp. 475-515.

Mohamed, A., García-Martínez, S., Loumerem, M., Carbonell, P., Ruiz, J. J., & Boubaker, M. 2019. Assessment of genetic diversity among local pea (*Pisum sativum* L.) accessions cultivated in the arid regions of Southern Tunisia using agro-morphological and SSR molecular markers. *Genetic Resources and Crop Evolution*, 66(6), 1189-1203.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 23, 341-369.

Nolte, G. E. 2020. Peru: Coffee Annual [USDA Agency]. The Foreign Agricultural Service (FAS) - United States Department of Agriculture. Available in: https://www.fas.usda.gov/data/peru-coffee-annual-4

Nsabiyera, V., Qureshi, N., Bariana, H. S., Wong, D., Forrest, K. L., Hayden, M. J., & Bansal, U. K. 2016. Molecular markers for adult plant leaf rust resistance gene Lr48 in wheat. *Molecular Breeding*, 36(6), 65.

Prakash, N. S., Marques, D. V., Varzea, V. M. P., Silva, M. C., Combes, M. C., & Lashermes, P. 2004. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. *Theoretical and Applied Genetics*, 109, 1311-1317.

Prakash, N.S., Combes, M. C., Somanna, N., & Lashermes, P. 2002. AFLP analysis of introgression in coffee cultivars (*Coffea arabica* L.) derived from a natural interspecific hybrid. *Euphytica*, 124, 265-271.

Pruvot-Woehl, S., Krishnan, S., Solano, W., Schilling, T., Toniutti, L., Bertrand, B., & Montagnon, C. 2020. Authentication of *Coffea arabica* Varieties

through DNA Fingerprinting and its Significance for the Coffee Sector. *Journal of AOAC INTERNATIONAL, 103*(2), 325-334.

Ribas, A. F., Cenci, A., Combes, M. C., Etienne, H., & Lashermes, P. 2011. Organization and molecular evolution of a disease-resistance gene cluster in coffee trees. *BMC Genomics, 12*, 240.

Romero, G., Vásquez, L. M., Lashermes, P., & Herrera, J. C. 2014. Identification of a major QTL for adult plant resistance to coffee leaf rust (*Hemileia vastatrix*) in the natural Timor hybrid *(Coffea arabica* x *C. canephora)*. *Plant Breeding, 133*, 121-129.

Sánchez, E., Solano, W., Gatica-Arias, A., Chavaría, M., & Araya-Valverde, E. 2020. Microsatellite DNA fingerprinting of *Coffea sp.* Germplasm conserved in Costa Rica through singleplex and multiplex PCR. *Crop Breeding and Applied Biotechnology, 20*(1), e27812013.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology, 18*, 233-234.

Silva, D. N., Várzea, V., Paulo, O. S., & Batista, D. 2018. Population genomic footprints of host adaptation, introgression and recombination in coffee leaf rust. *Molecular Plant Pathology, 19*, 1742-1753.

Sousa, T. V., Caixeta, E. T., Alkimim, E. R., Oliveira, A. C. B., Pereira, A. A., Sakiyama, N. S., Zambolim, L., & Resende, M. D. V. 2019. Early Selection Enabled by the Implementation of Genomic Selection in *Coffea arabica* Breeding. *Frontiers in Plant Science, 9*, 1934.

Souza, H. A. V., Muller, L. A. C., Brandão, R. L., & Lovato, M. B. 2012. Isolation of high quality and polysaccharide-free DNA from leaves of Dimorphandra mollis (*Leguminosae*), a tree from the Brazilian Cerrado. *Genetics and Molecular Research, 11*, 756-764.

Talhinhas, P.; Batista, D.; Diniz, I.; Vieira, A.; Silva, D. N.; Loureiro, A.; Tavares, S.; Pereira, A. P.; Azinheira, H. G.; Guerra-Guimarães, L.; Várzea, V. & Silva, M. D. C. 2017. The coffee leaf rust pathogen *Hemileia vastatrix*: One and a half centuries around the tropics. Molecular Plant Pathology 18: 1039-1051.

Tornincasa, P. 2008. Marcatori genetici per l'analisi, la caratterizzazione e la tracciabilità del caffè e delle due specie vegetali *Coffea arabica* L. e Coffea canephora. PhD Thesis Genetica, Università degli Studi di Trieste. Italia. 110 pp.

Valencia, A., Morales, A. Y., del Pilar Moncada, M. I., Cortina, H. A., & Herrera, J. C. 2017. Introgresion of the SH3 gene resistant to rust (*Hemileia vastatrix*) in improved lines of CASTILLO variety (*Coffea arabica* L.). *Journal of Plant Breeding and Crop Science, 9*, 130-138.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. Springer.

Zambolim, L. 2016. Current status and management of coffee leaf rust in Brazil. *Tropical Plant Pathology, 41*, 1-8.

Appendix

List of genotypes in INIA's Coffee Germplasm Collection samples obtained by Sat244 marker (Detailed)

Samples	Alleles for Sat244 (bp)						
	295	319	321	323	329	331	333
CA 1-169*	1			1			
CA-16.3	1		1				
CA-16.5	1		1				
CA-16.8	1		1				
CA-20.4	1		1				
CA-20.7	1		1				
CA-21.7	1		1	1			
CA-29.4	1		1				
CA- 45.2	1		1				
CA- 45.3	1		1				
CA-47.3	1		1	1			
CA-55.4	1		1				
CA-55.5	1		1				
CA- 56.4	1		1				
CA- 56.8	1		1				
CA- 58.3	1		1				
CA- 58.6	1		1				
CA-61.5	1		1				
CA-61.7	1		1				
CA-64.4	1		1				
CA-64.6	1		1				
CA-69.3	1		1				
CA-69.5	1		1				
CA-90.4	1	1					
CA-90.5	1	1					
CA-94.4	1	1					
CA- 97.5	1	1		1			
CA-107.4	1		1	1			
CA-140.3	1	1					
CA-140.4	1	1					
CA-144.4	1		1				
S.288	1						1
EA67					1	1	

(*) CA 1-169*: The group of 303 samples with the same genotype 295/323.