

*Full Length Research Paper*

# **Genetic relationship and the occurrence of multiple gene resistance to coffee berry disease (*Colletotrichum kahawae*, Waller & Bridge) within selected *Coffea arabica* varieties in Kenya**

**James M. Gimase<sup>1\*</sup>, Wilson M. Thagana<sup>2</sup>, Chrispine O. Omondi<sup>3</sup>, Jane J. Cheserek<sup>1</sup> and Elijah K. Gichuru<sup>1</sup>**

<sup>1</sup>Kenya Agricultural and Livestock Research Organization (KALRO) - Coffee Research Institute, P. O. Box 4-00232, Ruiru, Kenya.

<sup>2</sup>Department of Agricultural Science and Technology, School of Agriculture and Enterprise Development, Kenyatta University, P.O Box 43844 - 00100, Nairobi, Kenya.

<sup>3</sup>Kenya Agricultural and Livestock Research Organization (KALRO) Sugar Research Institute, P. O. Box 44- 40100, Kisumu, Kenya.

Received 24 September, 2020; Accepted 23 November, 2020

**The Coffee Berry Disease (CBD) epidemics destroy up to 100% of the crop on a susceptible variety. Resistance to CBD is conferred by the T-gene in Hibrido De Timor (HDT), R-gene in Rume Sudan (RS), and k-gene in K7, which were assembled in resistance varieties Ruiru 11(R11) and Batian. This study aimed to evaluate the genetic relationship between R11 and Batian with their parents' HDT, RS, SL8 and confirm the occurrence of T and R genes using DNA markers. Genome-wide single nucleotide polymorphism (SNP) markers were obtained through Diversity Arrays Technology sequencing (DArTseq). The genetic relationship was analyzed by Principal Component Analysis and hierarchical clustering. The Tgene was confirmed using Microsatellite primer, Sat 235 while the R gene was by marker sequence search within the DArTseq result files. The PC1 accounted for 42% of the total variation. Hierarchical clustering revealed less than 10% dissimilarity index apart from HDT that recorded above 20%. All the R11 and Batian genotypes carry the T gene. Eleven genotypes carry both T and R genes, therefore, with broad-based resistance to CBD. The study confirmed the narrow genetic relations within the *Coffea arabica* coffee varieties and further confirmed the occurrence of multiple gene resistance in R11 and Batian that will not break easily to new pathogen races.**

**Key words:** Coffee berry disease, diversity arrays technology sequencing, T gene, R gene, Ruiru 11, Batian, SL 28, codominant.

## **INTRODUCTION**

Coffee belongs to the family Rubiaceae and the genus *Coffea* with over 124 species that have been characterized

(Davis et al., 2011). Despite the diversity of this genus, only two species, *Coffea arabica* L. and *Coffea canephora*

\*Corresponding author. Email: [jgimase@yahoo.com](mailto:jgimase@yahoo.com).

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

P. are of economic importance contributing about 70 and 30% of the total world market share, respectively (Setotaw et al., 2020). The *C. arabica* is an allotetraploid species ( $2n = 2x = 44$ ) that exhibits a diploid-like meiotic behavior (Lashermes and Combes, 2018). *C. arabica* is the only tetraploid species of the *Coffea* genus while the rest of the species are diploid (Spiniso-Castillo et al., 2020). The *C. arabica* is believed to have been formed as a result of spontaneous hybridization between two diploid species, *C. canephora* and *Coffea eugenioides* (Lashermes et al., 1999, 2011). The species is autogamous with about 10% out-crossing (Bikila et al., 2017). On the other hand, *C. canephora* is diploid ( $2n = 2x = 22$ ), highly diverse (Bertrand et al., 2003), with resistance to common disease and thus a good source of genes for disease resistance (Ky et al., 2001).

Next-generation sequencing (NGS) technologies, such as genotyping-by-sequencing (GBS) and Diversity Arrays Technology sequencing (DArTseq), provide markers that are widely used in genome-wide analysis (Spiniso-Castillo et al., 2020). The GBS approach is more informative than predesigned single nucleotide polymorphism (SNP) arrays especially on wild germplasm as it is unbiased and provide information on rare alleles while the DArTseq method is based on the complexity reduction approach, using restriction enzymes that target the genome coding regions (Pailles et al., 2017). The restriction enzymes separate low copy sequences from the repetitive regions of the genome that are more informative for marker discovery for breeding purposes (Courtois et al., 2013; Pailles et al., 2017).

The genetic variation is controlled by the segregation of multiple genes such that the variances of individual loci are so small that they cannot be investigated individually and thus the need to analyze sets of these loci (Bikila et al., 2017). The DArTseq based SNP markers were successfully utilized in the determination of genetic relations in the *Coffea* genus by Garavito et al. (2016) and Spiniso-Castillo et al. (2020).

The coffee berry disease (CBD) caused by a specialized hemibiotrophic fungal pathogen, *Colletotrichum kahawae* (Waller & Bridge) (Waller et al., 1993), is a key constraint of Arabica coffee production in Africa (Hindorf and Omondi, 2011; Van Der Vossen et al., 2015; Diniz et al., 2017; Vieira et al., 2019). The CBD epidemics can destroy 50 to 80% (Van Der Vossen and Walyaro, 2009; Diniz et al., 2017) and at times up to 100% (Giddisa, 2016) of the developing berries, 4 to 16 weeks following anthesis on a susceptible variety when no control measure is applied (Gichuru et al., 2012). The control of CBD using intensive fungicide spray programs (8-12 rounds per year) increases the cost of production by up to 40% (Van der Vossen and Walyaro, 2009). The use of these chemicals also contributes to environmental pollution (Gichuru et al., 2008), thus the use of resistant varieties is the most cost-effective and environmentally friendly approach for CBD management. Resistance to

CBD is governed by three genes in the varieties Rume Sudan (R genes), HDT (Tgene) and K7 (k gene) where R and T are dominant while k is recessive (Van Der Vossen and Walyaro, 1980).

The breeding program for resistance to CBD in Kenya started in 1971 with the main breeding goal of developing cultivars that combine resistance to CBD, high production, good beverage quality, and compact growth that will be amenable to high-density planting (Van Der Vossen and Walyaro, 1980). Using conventional approaches, genes for resistance to CBD were introduced to *C. arabica* coffee varieties that are susceptible by crossing with donor varieties and backcrossing to standard varieties to restore desirable attributes (Walyaro, 1983). However, this approach takes a long time to develop a coffee variety due to the long juvenile nature of the *Coffea* genus (Moncada et al., 2016). The low genetic diversity of *C. arabica* also hinders the identification and selection of superior genotypes using traditional breeding methods (Sousa et al., 2017). To overcome this constraint, molecular markers have been used as a supporting tool to accurately discriminate genotypes and accelerate coffee breeding programs (Sousa et al., 2017). The DNA marker for the T gene was identified by Gichuru et al. (2008) and linked to Simple Sequence Repeats (SSR) primer locus Sat 235, popularly designated as *Ck-1* (Gichuru et al., 2008). This marker was validated by Alkimim et al. (2017) who confirmed that Sat 235 marker co-segregate with the T-gene. A recent study by Gimase et al. (2020a) identified the putative DNA marker for the R gene using Single Nucleotide Polymorphism (SNP) markers.

The CBD resistant cultivar R11 is an F1 hybrid derived from a cross between a specific female and male population (Omondi et al., 2001), while Batian is a pure line that was selected from the R11 pollen parents (Omondi et al., 2001), after several generations of selfing to fix the CBD resistant genes (Gichimu et al., 2014). SL 28, is a Bourbon type single-tree selection that combines high yield, high quality, and drought tolerance but highly susceptible to CBD (Walyaro, 1983). The main objective of this study was to evaluate the genetic relationship within selected *C. arabica* cultivars R11, Batian and their resistance donor parent HDT, RS and the recurrent parent SL28 using DArTseq based SNP markers and identify genotypes within R11 and Batian with multiple gene resistance to CBD conferred by the T and R genes.

## MATERIALS AND METHODS

This study was carried out at the Kenya Agricultural and Livestock Research Organization – Coffee Research Institute (KALRO-CRI) in Ruiru, Kenya. Ruiru is located within the upper midland (UM2) at 1° 06'S and 36° 45'E and an altitude of 1620 m above sea level. The rainfall pattern is bimodal with 1063 mm per annum and the annual average temperature is 19°C with a range of 12.8 to 25.2°C (Jaetzold et al., 2006).

The study materials were 91 genotypes comprising of 61 crosses

of the variety of R11 and 27 families of the variety Batian. Also included were SL 28, HDT and Rume Sudan.

### Sample collection, genomic DNA extraction and genotyping of SNP markers

Fresh leaves were randomly picked from each of the 91 coffee genotypes, kept in cool boxes and taken to the laboratory for DNA sample extraction. The LGC genomics plant sample collection kit ([www.lgcgenomics.com](http://www.lgcgenomics.com)) was used in sample collection, where 6 disks were cut and placed in each strip of the 96 deep well sample plate, and sent to the Integrated Genotyping Service and Support (IGSS) platform (<https://ordering.igssafrica.org/cgi-bin/order/login.pl>) for DNA extraction and genotyping.

Genomic DNA samples were extracted from genotypes using a standard cetyltrimethylammonium Bromide (CTAB) protocol of Doyle and Doyle (1987). The quality and quantity of the DNA samples were evaluated by running it through 0.8% agarose gel electrophoresis. The DNA concentration was adjusted to 50 ng/μl. The genomic DNA samples were sent to Diversity Arrays Technology (DArT) Pty Ltd., in Canberra-Australia (<http://www.diversityarrays.com>) for sequencing and identification of SNP markers. The GBS-SNP was performed following the standard protocol as described by Elshire et al. (2011). Next-generation sequencing was carried out using the HiSeq2500 Illumina platform.

The SNP calling was carried out by the DArT-soft14 algorithm within the KDCompute pipeline developed by Diversity Arrays Technology (<https://kdcompute.seqart.net/kdcompute/plugins>). In the primary pipeline, the FASTQ files were first processed to filter poor quality sequences to ensure that the assignments of the sequences to specific samples carried in the barcode split region were consistent and reliable (Nemli et al., 2017; Barilli et al., 2018). The identical sequences were collapsed into FASTQ call files that were used in the secondary pipeline for DArT P/L's proprietary SNPs calling algorithms (DArT-soft14) pipeline in the processing of the sequence data (Barilli et al., 2018). Since the allotetraploid *C. arabica* open-access genome assembly, with a reliable sorting of homoeologous sequences, is not yet available (Scalabrin et al., 2020), the filtered sequence reads were aligned against the finer and publicly available diploid *C. canephora* genome (<http://coffee-genome.org/coffeacanephora>) as a reference to find the SNP markers in *C. arabica* genome (Sant'Anna et al., 2018) and to determine their corresponding genomic positions.

### The SNP marker quality analysis and evaluation of the genome-wide relations

The SNP loci with >30% missing data and rare SNPs with less than 5% minor allele frequencies (MAF) and heterozygosity (Ho) above 90% were removed (Garot et al., 2018). The genetic relationships within the study genotypes were determined using the Principal Component Analysis (PCA) and hierarchical clustering, both implemented within the clustering analysis component of the KDCompute plugins system (<https://kdcompute.seqart.net/kdcompute/plugins>).

### Genomic DNA extraction, amplification, and electrophoresis using SSR primer locus Sat 235

A total of 59 genotypes comprising 27 R11 crosses, 27 Batian families, five control genotypes, HDT, Robusta, Rume Sudan, and susceptible cultivars SL28 and Caturra were analyzed. Healthy leaves were picked and genomic DNA extracted following the method of Diniz et al. (2005) with minor modifications in the

extraction buffers. About 500 mg of fresh leaves were ground and transferred to 2 mL Eppendorf tubes. After grinding, 1 mL extraction solution was added and the tubes shaken vigorously for 5 min and immediately put in a 65°C water bath for 40 min. After which the samples were centrifuged for 5 min at 13000 rpm and the supernatant transferred to a new tube, to which 1 mL CIA (chloroform: isoamyl 24:1) was added and the tubes were shaken for 10 min and centrifuged for 5 min at 12000 rpm. The supernatant was transferred to another tube and the same volume of frozen Isopropanol was added and maintained at -20°C for 1 h.

The content was centrifuged at 1300 rpm for 5 min, the supernatant discarded and the pellet washed with 70% ethanol. This step was repeated twice and after drying, the pellets were treated with 190 μL TE (Tris-EDTA buffer plus RNAse 10 mg μL<sup>-1</sup>) for 30 min at 37 and 65°C for 5 min. The DNA was then purified with the addition of 100 μL TE, 100 μL water, 100 μL NaCl 5 M and 100 μL EDTA 0.5 M. The samples were homogenized and incubated on ice for 30 min and centrifuged for 5 min at maximum speed and isopropanol added. After drying, the pellet for each genotype was diluted in an appropriate amount of TE buffer as per the amount of DNA quantified using a spectrophotometer and stored at 4°C. The extracted DNA quality was determined by running the samples in 1% agarose gel alongside a Lambda standard with a known concentration of DNA fragments for comparison and quantification of the samples.

The Polymerase Chain Reaction (PCR) was carried out in a total volume of 25 μL, containing 10 ng/μL template of genomic DNA, 0.4 μM of Sat 235 SSR primer, 75 μM dNTPs (each), 2.5 μM MgCl<sub>2</sub>, PCR buffer 1x TBE [75 mM Tris-HCl; 0.5 Na<sub>2</sub> EDTA (pH 8.0)], 20 Mm Boric acid and 1 unit Taq DNA polymerase (from Gene - on company, Germany). Amplification was carried out in a Eurogene thermocycler (TECHNE, UK). The amplification program was one cycle of initial denaturation at 94°C for 5 min followed by 35 cycles of 30 s at 94°C (denaturation), 30 s at 55°C for primer annealing, and 1 min and 30 seconds at 72°C for elongations with a final extension at 72°C for 10 min.

The amplification products with SSR primer Sat 235 were electrophoresed in 2.3% (w/v) agarose gel with a 1x TBE buffer system and then visualized in a UV light trans-illuminator after staining in 60% ethidium bromide solution.

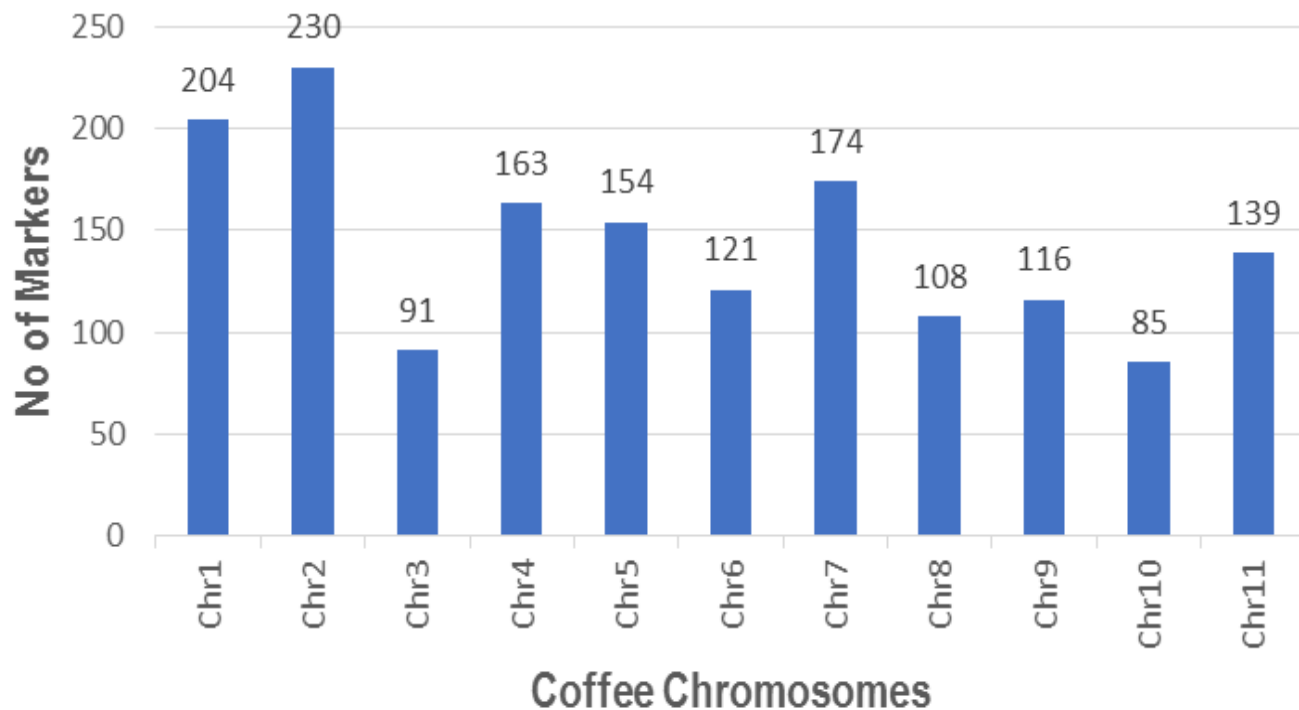
### Confirmation of the occurrence of multiple gene resistance to CBD conferred by the T and R genes

The presence/absence of the Tgene was confirmed by observation of the amplified fragment, based on the standard HDT, Robusta, and SL28 while the genotypes carrying DNA marker for R genes were identified by searching the SNP marker sequences within the GBS-based DArTseq marker result files of the study genotypes.

## RESULTS

### The analysis of the SNP marker and evaluation of the genome-wide relations

The DArTseq generated 2280 good quality SNP markers (MAF>5% and Ho<90%), out of which 1575 were aligned on the 11 chromosomes based on the *C. canephora* reference genome, markers that were well distributed with the genome (Figure 1) and therefore used in further analysis. The PCA results revealed that all the Batian genotypes from the three crosses clustered together apart from two genotypes, CR8-155 and CR30-809 that



**Figure 1.** The genome-wide distribution of DArTseq-based SNP markers.

deviated from this cluster (Figure 2). The two genotypes had a close relation with Rume Sudan. The Batian genotypes were closely related to the cultivar SL 28. About 60% of the R11 crosses (38) formed one main cluster while the rest (23) were evenly distributed within the PCA plot. HDT did not portray any close relationship with any of the genotypes in the study. PC1 was the most important and accounted for 42% of the total phenotypic variation.

Similar to the PCA, the hierarchical clustering analysis (dendrogram) revealed that all the Batian formed one cluster together with SL 28 apart from two genotypes, CR8-155 and CR30-809 that showed a close relationship with Rume Sudan (Figure 3). 38 R11 crosses formed one major cluster while the rest (23) formed four different clusters. HDT did not cluster with any of the study genotypes while Rume Sudan showed a closer relationship with Batian and three R11 genotypes (R11-6, 22 and 195). The cluster dendrogram also revealed that all the R11 crosses, Batian families from the three crosses, RS and SL 28 had low genetic diversity of about 10% or less while HDT had a relatively high diversity of about 20%.

#### **The occurrence of the T (*Ck-1*) gene in variety R11 and Batian**

The amplification products revealed that all the 27 R11 crosses evaluated carry the DNA fragment for the T-gene

(*Ck-1*) that is introgressed from the *C. canephora* genome (Plate 1). The *Ck-1* DNA fragment was heterozygous for all the R11 genotypes. The *Ck-1* DNA fragment was also present in HDT and Robusta in a homozygous state but absent in Rume Sudan, SL 28, and Caturra. Similarly, all the 27 Batian families analyzed carry the *Ck-1* fragment (Plate 2), which was also present in HDT and Robusta but absent in SL28 and Rume Sudan. Out of the 27 Batian genotypes, 15 genotypes namely CR8-423, CR8-760, CR8-149, CR8-155, CR8-419, CR22-109, CR22-108, CR22-350, CR22-635, CR22-357, CR22-639, CR30-807, CR30-813, CR30-244 and CR30-809 were homozygous for the *Ck-1* gene while 12 genotypes CR8-136, CR8-154, CR8-761, CR8-420, CR22-759, CR22-114, CR22-353, CR22-111, CR30-812, CR30-242, CR30-233 and CR30-236 were heterozygous for the *Ck-1* gene (Plate 2).

#### **The occurrence of the R-gene within the variety R11 and Batian**

A total of 91 genotypes comprising 61 R11, 27 Batian and three control coffee varieties (Rume Sudan, SL 28 and HDT) were sequenced. The sequences search using the two SNP marker sequences (100025973|F|0-59:T>C-59:T>C and 100034991|F|0-44:C>T-44:C>T), that were significantly associated with CBD resistance through GWAS and QTL mapping (Gimase et al., 2020a, b), were carried out within the sample GBS result files. The



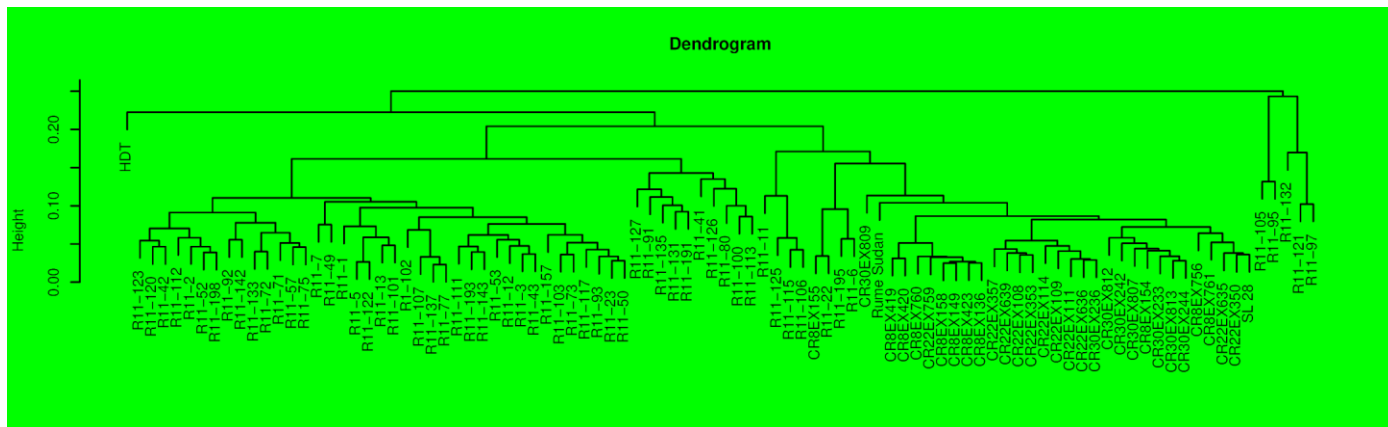


Figure 3. The clustering dendrogram of the genotypes within the selected coffee varieties and their relationships.

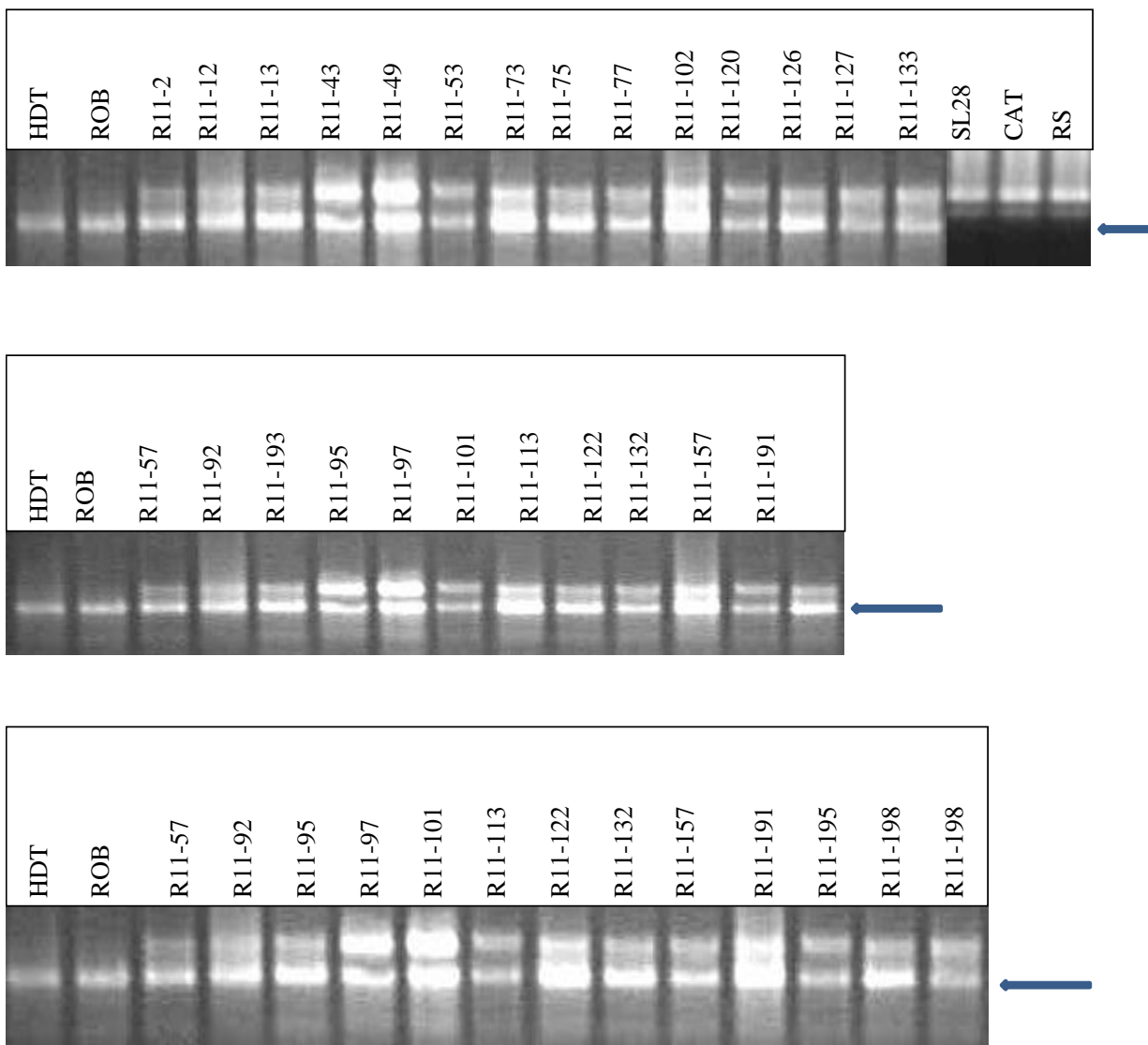
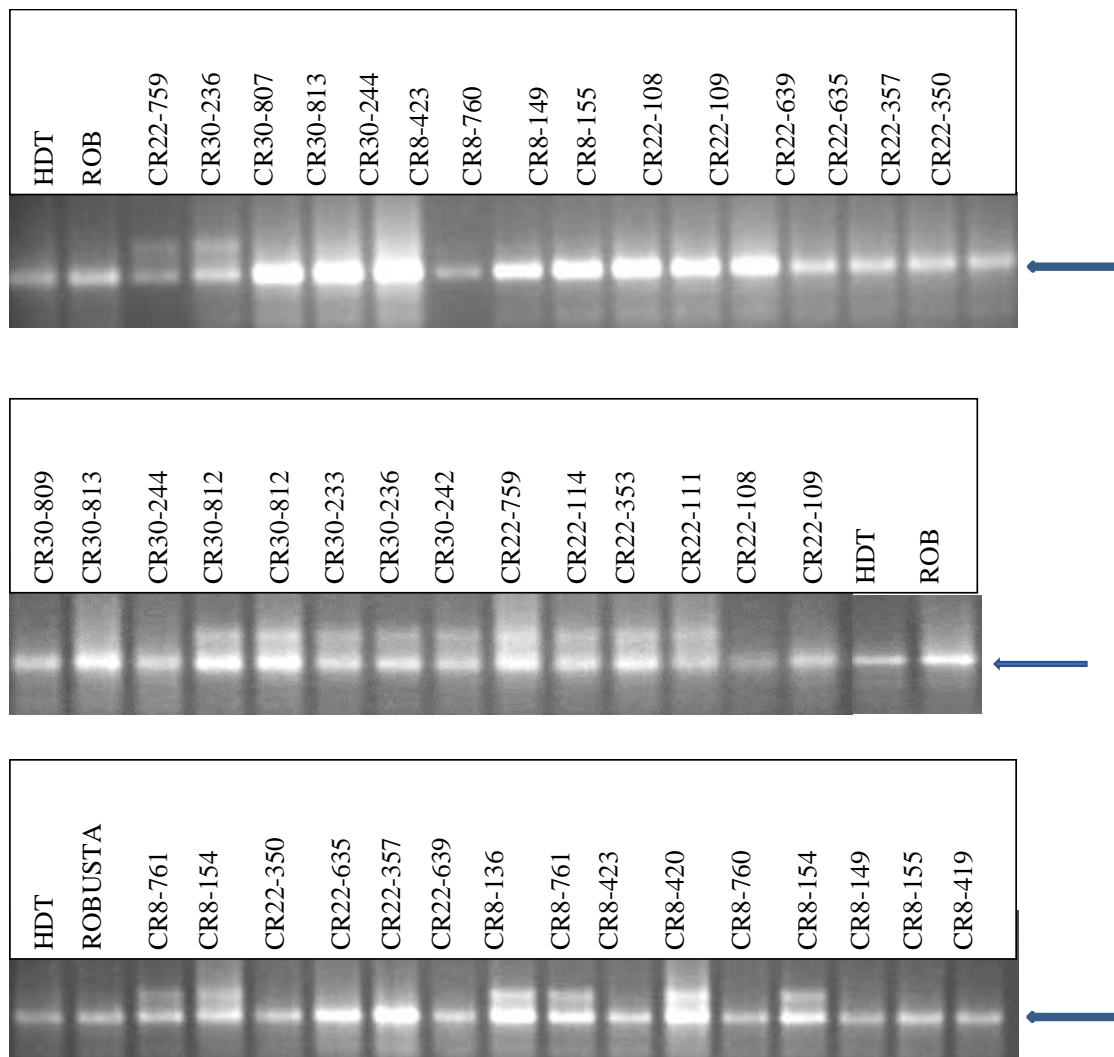


Plate 1. The occurrence of *Ck-1*-gene fragment within cultivar R11 crosses. The fragment is indicated by the arrow. Rob – Robusta, CAT – Caturra, HDT – Híbrido de Timor.



**Plate 2.** Occurrence of *Ck-1* gene fragment within the families of Batian crosses, indicated by the arrow.

**Table 1.** Occurrence of the R-gene marker within the variety R11 and Batian.

Genotype	CR30-809	R11-157	CR8-155	R11-22	R11-121	R11-195	R11-6
Marker (100034991 F 0-44:C>T-44:C>T)	1	2	1	1	2	1	1

Genotype	R11-135	R11-123	CR8-136	R11-11	Rume Sudan	SL 28	HDT
Marker (100034991 F 0-44:C>T-44:C>T)	2	1	2	2	1	0	0

1 - Homozygous, 2 - Heterozygous, 0 - Absent.

confirmed to carry the T-gene (*Ck-1*).

## DISCUSSION

### The genome-wide relationship within the study varieties

The genome-wide analysis revealed a narrow genetic

base within *C. arabica* coffee varieties R11, Batian, SL 28 and Rume Sudan, with a diversity index of less than 10%. This is attributed to the fact that these varieties are derived from a few individual collections and whose subsequent dispersal has progressively narrowed further their genetic base (Baruah et al., 2003; Setotaw et al., 2013). The variety of HDT had a higher diversity index of about 20%. The variety HDT is a natural interspecific cross between *C. arabica* and *C. canephora* and usually shows



a divergence from commercial cultivars for most of the agronomic traits (Agwanda et al., 1997). The expressed diversity in HDT is due to the introgressed genes from the *C. canephora* genome (Lashermes et al., 1999). Although HDT was utilized as a donor parent for resistance to CBD in the varieties' R11 and Batian, the high diversity was not reflected in these genotypes as in HDT. This is most likely due to filial advancement from the original crosses, such that the existing progenies contain less of the initial *C. canephora* genome (Gichuru, 2007).

The study revealed uniformity within individuals of the three crosses that make up the variety Batian and revealed further that they are closely related to SL 28. The variety Batian was obtained from complex crosses between CBD donor parents (HDT, Rume Sudan and K7) and susceptible varieties (SL 28, SL 34, Bourbon and Tanganyika drought-resistant selections) and then backcrossed to SL28 and selfed to restore and fix genes for desirable attributes of superior beverage quality (Gichimu et al., 2014). From this result, it is most likely that this attribute was successfully restored as SL 28 was closely related to the individual genotypes of Batian.

The cultivar R11 is also relatively uniform as 60% of the crosses clustered together while the rest were in four different small clusters. R11 is an F1 hybrid between complex crosses (that were selected as Batian) as pollen parents and Catimor (a cross between HDT and Caturra) isolines as seed parents (Omondi et al., 2001). Variation within R11 crosses on various traits have been reported in previous studies; resistance to CBD (Omondi et al., 2001; Gichimu et al., 2014); beverage quality (Gichimu et al., 2012); and yield (Gichimu et al., 2013).

#### **The occurrence of the T-gene (*Ck-1*) within *C. arabica* variety R11 and Batian**

All the 27 R11 crosses were confirmed for the occurrence of the T-gene. R11 is a composite F1 hybrid made up of 66 different crosses (Omondi et al., 2001). The previous study by Gichimu et al. (2014), confirmed the occurrence of T-gene in 34 R11 crosses and therefore this study brings the total number of R11 crosses confirmed to carry the T-gene through the marker-assisted selection to 61. The cultivar R11 inherited the T-gene from two different sources, the seed parent Catimor and Pollen parents. The seed parent Catimor comprises several lines derived from a cross between HDT (a spontaneous cross between *C. canephora* and Caturra, a *C. arabica* cultivar, highly susceptible to CBD) (Gichuru et al., 2008). The pollen parents are complex crosses between CBD resistance var HDT, Rume Sudan and susceptible SL28, SL34, SL4, N39 and Bourbon (Van Der Vossen et al., 1981; Agwanda et al., 1997; Omondi et al., 2001).

Similarly, all the 27 Batian genotypes analyzed were confirmed to carry the *Ck-1* gene where 15 genotypes were homozygous and hence stable for the gene. Batian is a selection from the pollen parents of R11 hence

inherited the T-gene from HDT, one of the resistance sources in the complex crosses (Gichimu et al., 2014).

A study by Alkimim et al. (2017) using three CBD resistance genotypes in Brazil revealed the occurrence of *Ck-1*, within the genotypes, where two genotypes were homozygous while one was heterozygous and confirmed that Sat 235 marker co-segregates with the gene. Similarly, a study by Mtenga (2016) reported the occurrence of *Ck-1* in CBD resistant genotypes from Tanzania and Ethiopia accessions. In the study, Sat 235 could not amplify the T-gene fragment in Rume Sudan since the gene for resistance to CBD in this variety is in a different locus. This therefore confirmed further, the findings by Mtenga (2016) as the *Ck-1* gene was not amplified in Rume Sudan.

#### **The occurrence of the R-gene within *C. arabica* variety R11 and Batian**

Eleven genotypes were confirmed for the occurrence of the SNP 100034991|F|0-44:C>T-44:C>T while none of the genotypes carry the SNP marker 100025973|F|0-59:T>C-59:T>C. The SNP marker 100025973|F|0-59:T>C-59:T>C was comparatively a rare variant where minor allele frequency (MAF) was lower (0.28) as opposed to 100034991|F|0-44:C>T-44:C>T (0.42) out of the possible maximum of 0.5 (Zhang et al., 2015; Gimase et al., 2020a). The SNP marker 100025973|F|0-59:T>C-59:T>C could have been lost in both R11 and Batian during backcrossing of the complex crosses to SL28 to restore good cup quality and high yield (Agwanda et al., 1997) since the selection process was not guided by the use of DNA markers (Omondi, 1998). This, therefore, confirmed that the SNP marker 100034991|F|0-44:C>T-44:C>T as a reproducible marker within *C. arabica* genotypes carrying resistance gene inherited from Rume Sudan. The polymorphic occurrence of this locus in Rume Sudan and SL 28 signifies its ability to discriminate variants in terms of resistance to CBD and its suitability for MAS (Rouet et al., 2019). The polymorphic genomic loci are used as genetic markers in the determination of the co-segregation of genetic alleles with qualitative traits emanating from populations of crosses or naturally occurring populations (Motazedhi et al., 2019). The SNP marker 100034991|F|0-44:C>T-44:C>T was found to be linked to two genes by Gimase et al. (2020b), a finding that is further supported by the inheritance study by Van Der Vossen et al. (1980) that reported that the occurrence of two alleles for R gene in Rume Sudan as R1R1.

#### **Conclusion**

The study confirmed the close relationship within the *C. arabica* coffee varieties as exhibited by a narrow genetic base, with Batian as a very uniform cultivar and genetically very close to SL 28 while R11 was fairly



uniform. As per the previous studies, this work affirmed further the uniformity within the *C. arabica* genome. The study also revealed that all the genotypes within the CBD resistant varieties R11 and Batian carry the T gene, while 11 carries the R gene and therefore with multiple gene resistance to CBD. The study further confirmed the codominant nature of the DNA marker for T gene (*Ck-1*) and R gene due to their ability to discriminate between homozygous and heterozygous variants within the resistant genotypes. From this study, it is evident that the selection of arabica coffee varieties with multiple gene resistance to CBD is a reality. The genotypes that were confirmed to carry the two genes for resistance to CBD are recommended for further distribution to growers since resistance will not break easily to new disease races.

## CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

## ACKNOWLEDGEMENTS

This work was supported by KALRO-CRI, International Livestock Research Institute (ILRI), and the National Research Fund (NRF). The authors thank the Director-General, KALRO (through the Institute Director, CRI), the Director-General ILRI, and the Chief Executive officer, NRF. This work is published with the permission of the Institute Director, CRI.

## REFERENCES

- Agwanda CO, Lashermes P, Trouslot P, Combes MC, Charrier A (1997). Identification of RAPD markers for resistance to coffee berry disease, *Colletotrichum kahawae*, in arabica coffee. *Euphytica* 97(2):241-248.
- Alkimim ER, Caixeta ET, Sousa, TV, Pereira AV, de Oliveira ACB, Zambolim L, Sakiyama NS (2017). Marker-assisted selection provides Arabica coffee with genes from other *Coffea* species targeting multiple resistance to rust and coffee berry disease. *Molecular Breeding* 37(6):1-10.
- Barilli E, Cobos MJ, Carrillo E, Kilian A, Carling J, Rubiales D (2018). A High-Density Integrated DArTseq SNP-Based Genetic Map of *Pisum fulvum* and Identification of QTLs Controlling Rust Resistance. *Frontiers in Plant Science* 9:167.
- Baruah A, Naik V, Hendre P, Rajkumar P, Aggarwal RK (2003). Isolation and characterization of nine microsatellite markers from *Coffea arabica* L. showing wide cross-species amplification. *Molecular Ecology Notes* 3:647-650.
- Bikila BA, Sakiyama NS, Caixeta ET (2017). SNPs Based Molecular Diversity of *Coffea canephora*. *Journal of Microbiology and Experimentation* 5(1):00136.
- Courtois B, Audebert A, Dardou A, Roques S, Ghneim-Herrera T, Droc G, Rouin J, Rouan L, Gozé E, Kilian A, Ahmadi N (2013). Genome-wide association mapping of root traits in a japonica rice panel. *PLoS ONE* 8(11):e78037.
- Diniz LE, Sakiyama NS, Lashermes P, Caixeta ET, Oliveira AC, Zambolim EM, Loureiro ME, Pereira AA, Zambolim L (2005). Analysis of AFLP markers associated with the *Mex-1* resistance locus in Icatu progenies. *Crop Breeding and Applied Biotechnology* 5(4):387-393.
- Diniz I, Figueiredo A, Loureiro A, Batista D, Azinheira H, Várzea V, Pereira AP, Gichuru E, Moncada P, Guerra-Guimarães L, Oliveira H (2017). A first insight into the involvement of phytohormones pathways in coffee resistance and susceptibility to *Colletotrichum kahawae*. *PLoS ONE* 12(5):e0178159.
- Davis AP, Tosh J, Ruch N, Fay M (2011). Growing coffee: *Psilanthus* (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of *Coffea*. *Botanical Journal of the Linnean Society* 167(4):357-377.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11-15. Available at: <https://worldveg.tind.io/record/33886/>
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler E, Mitchell S (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS one* 6(5):e19379.
- Garavito A, Montagnon C, Guyot R, Bertrand B (2016). Identification by the DArTseq method of the genetic origin of the *Coffea canephora* cultivated in Vietnam and Mexico. *BMC Plant Biology* 16:242.
- Garot E, Joet T, Combes MC, Lashermes P (2018). Genetic diversity and population divergences of an indigenous tree (*Coffea mauritiana*) in Reunion Island: the role of climatic and geographical factors. *Heredity* 122:833-847.
- Gichimu BM, Gichuru EK, Mamati GE, Nyende AB (2012). Selection within *Coffea Arabica* cv. Ruiru 11 for high cup quality. *African Journal of Food Science* 6(18):456-464.
- Gichimu BM, Gichuru EK, Mamati GE, Nyende AB (2013). Yield Selection within *Coffea arabica* cv. Ruiru 11. *American Journal of Experimental Agriculture* 3(1):76-88.
- Gichimu BM, Gichuru EK, Mamati GE, Nyende AB (2014). Occurrence of *Ck-1* gene conferring resistance to Coffee Berry Disease in *Coffea Arabica* cv. Ruiru 11 and its parental genotypes. *Journal of Agricultural and Crop Research* 2(3):51-61.
- Gichuru EK (2007). Characterization of genetic resistance of Coffee berry disease (*Colletotrichum kahawae* Waller and Bridge) in Arabica coffee (*Coffea arabica* L.) that is introgressed from *Coffea canephora* (PhD thesis, University of Nairobi, Kenya).
- Gichuru EK, Agwanda CO, Combe MC, Mutitu EW, Ngugi ECK, Bertrand B, Lashermes P (2008). Identification of molecular markers linked to a gene conferring resistance to Coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica*. *Plant Pathology* 57:1117-1124.
- Gichuru EK, Ithiru JM, Silva MC, Pereira AP, Varzea VMP (2012). Restructured sampling plan enables the characterization of more virulence genes of *Hemileia vastatrix* in Kenya. 24<sup>th</sup> International Conference on Coffee Science, San José, Costa Rica.
- Giddisa G (2016). A Review on the Status of Coffee Berry Disease (*Colletotrichum kahawae*) in Ethiopia. *Journal of Biology, Agriculture and Healthcare* 6(19):140-151.
- Gimase JM, Thagana WM, Omondi CO, Cheserek JJ, Gichimu BM, Gichuru EK, Ziyomo C, Sneller CH (2020a). Genome-Wide Association Study Identify the genetic loci conferring resistance to Coffee Berry Disease (*Colletotrichum kahawae*) in *Coffea arabica* var. Rume Sudan. *Euphytica* 216:86.
- Gimase JM, Thagana WM, Omondi CO, Cheserek JJ, Gichimu BM, Gichuru EK (2020b). Quantitative trait Loci (QTL) mapping of Resistance to Coffee Berry Disease (*Colletotrichum kahawae* Waller & Bridge) in *Coffea arabica* variety Rume Sudan. *African Journal of Agricultural Research* 16(8):1184-1194.
- Hindorf H, Omondi CO (2011). A review of three major fungal diseases of *Coffea arabica* L. in the rainforests of Ethiopia and progress in breeding for resistance in Kenya. *Journal of Advanced Research* 2(2):109-120.
- Jaetzold R, Schimidt H, Hornez H, Shisanya C (2006). Farm Management Handbook of Kenya, Vol. II/C: Natural Conditions and Farm Management Information, Central Kenya. Ministry of Agriculture, Nairobi; Kenya. P. 573.
- Ky CL, Guyot B, Louarn J, Harmon S, Noirot M (2001). Trigonelline inheritance in the interspecific *Coffea pseudozanguebaria* x *C. liberica* var. dewevrei cross. *Theoretical and Applied Genetics* 102(4):630-634.
- Lashermes P, Combes MC (2018). Diversity and genome evolution in

- coffee. In: Lashermes P (ed) Achieving sustainable cultivation of coffee. BDS Publ., Cambridge. pp. 3-20.
- Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A (1999). Molecular characterisation and origin of the *Coffea arabica* L. genome. *Molecular and General Genetics* 261(2):259-266.
- Lashermes P, Combes MC, Ansaldo C, Gichuru E, Noir S (2011). Analysis of alien introgression in the coffee tree (*Coffea Arabica* L.). *Molecular Breeding* 27(2):223-232.
- Mtenga JD (2016). Diversity, combining ability and Coffee Berry Disease (*Colletotrichum kahawae*) resistance among Ethiopian and Tanzanian Arabica coffee genotypes, (PhD thesis, Sokoine University of Agriculture, Tanzania).
- Moncada MP, Tovar E, Montoya JC, González A, Spindel J, McCouch S (2016). A genetic linkage map of coffee (*Coffea Arabica* L.) and QTL for yield, plant height, and bean size. *Tree Genetics and Genomes* 12(1):5.
- Motazed E, Maliepaard C, Finkers R, Visser R, Ridder D (2019). Family-Based Haplotype Estimation and Allele Dosage Correction for Polyploids Using Short Sequence Reads. *Frontiers in Genetics* 10:335.
- Nemli S, Ascogul TK, Ates D, Esiyok D, Tanyola MB (2017). Diversity and genetic analysis through DArTseq in common bean (*Phaseolus vulgaris* L.) germplasm from Turkey. *Turkish Journal of Agriculture and Forestry* 41:389-404.
- Omondi CO (1998). Genetic diversity among isolates of *C. kahawae* causing Coffee Berry Disease and their interactions with varieties and breeding populations of Arabica coffee (PhD thesis, University of Nairobi, Kenya).
- Omondi CO, Ayiecho PO, Mwang'ombe AW, Hindorf H (2001). Resistance of *Coffea Arabica* cv. Ruiru 11 tested with different isolates of *Colletotrichum kahawae*, the causal agent of coffee berry disease. *Euphytica* 121:19-24.
- Pailles Y, Ho S, Piresl S, Tester M, Negrão S, Schmöckel SM (2017). Genetic Diversity and Population Structure of Two Tomato Species from the Galapagos Islands. *Frontiers in Plant Science* 8:138.
- Rouet C, Lee EA, Banks T, O'Neill J, LeBlanc R, Somers DJ (2019). Identification of polymorphism within the *Rosa multiflora muRdr1A* gene linked to resistance to multiple races of *Diplocarpon rosae* W. in tetraploid garden roses (*Rosa x hybrida*). *Theoretical and Applied Genetics* 133(1):103-117.
- Sant'Ana GC, Pereira LF, Pot D, Ivamoto ST, Domingues DS, Ferreira RV, Leroy T (2018). Genome-wide association study reveals candidate genes influencing lipids and diterpenes contents in *Coffea arabica* L. *Scientific Reports* 8(1):1-12.
- Scalabrin S, Toniutti L, Gaspero G, Scaglione D, Magris G, Michele Vidotto M (2020). A single polyploidization event at the origin of the tetraploid genome of *Coffea Arabica* is responsible for the extremely low genetic variation in wild and cultivated germplasm. *Scientific Reports* 10:4642.
- Setotaw TA, Caixeta ET, Pereira AA, Oliveira ACB, Cruz CD, Zambolim EM, Zambolim L, Sakiyama NS (2013). Coefficient of Parentage in *Coffea arabica* L. Cultivars Grown in Brazil. *Crop Science* 53:1237-1247.
- Setotaw TA, Caixeta ET, Zambolim EM, Sousa TV, Pereira AA, Baião AC, Cruz CD, Zambolim L, Sakiyama NS (2020). Genome Introgression of Híbrido de Timor and Its Potential to Develop High Cup Quality *C. Arabica* Cultivars. *Journal of Agricultural Science* 12(4):64-76.
- Sousa TV, Caixeta ET, Alkimim ER, de Oliveira AC, Pereira AA, Sakiyama NS, de Resende Júnior MF, Zambolim L (2017). Population structure and genetic diversity of coffee progenies derived from Catuai and Híbrido de Timor revealed by genome-wide SNP marker. *Tree Genetics and Genomes* 13(6):124.
- Spiniso-Castillo JL, Escamilla-Prado E, Aguilar-Rincón VH, Ramos VM, Santos GG, Pe' rez-Rodríguez P, Corona-Torres T (2020). Genetic diversity of coffee (*Coffea spp.*) in Mexico evaluated by using DArTseq and SNP markers. *Genetic Resources and Crop Evolution* 67(7):1795-1806.
- Van der Vossen HAM, Walyaro DJA (1980). Breeding for resistance to coffee berry disease in *Coffea arabica* L. II. Inheritance of the resistance. *Euphytica* 29:777-791.
- Van der Vossen HAM, Walyaro DJA (1981). Coffee breeding program in Kenya. A review of progress made since 1971 and a plan of action for the coming years. *Kenya Coffee* 46:113-130.
- Van der Vossen HAM, Walyaro DJ (2009). Additional evidence for oligogenic inheritance of durable host resistance to coffee berry disease (*Colletotrichum kahawae*) in Arabica coffee (*Coffea arabica* L.). *Euphytica* 165:105-111.
- Van der Vossen H, Bertrand B, Charrier A (2015). Next-generation variety development for sustainable production of arabica coffee (*Coffea arabica* L.): a review. *Euphytica* 204:243.
- Vieira A, Diniz I, Loureiro A, Pereira AP, Silva MC, Várzea V, Batista D (2019). Aggressiveness profiling of the coffee pathogen *Colletotrichum kahawae*. *Plant Pathology* 68(2):358-368.
- Waller JM, Bridge PD, Black RL, Hakiza G (1993). Differentiation of the coffee berry disease pathogen. *Mycological Research* 97:989-994.
- Walyaro DJA (1983). Considerations in breeding for improved yield and quality in Arabica Coffee (*Coffea arabica* L.) (PhD thesis, Wageningen, the Netherlands).
- Zhang J, Liu T, Feng R, Liu C, Chi S (2015). Genetic Map Construction and Quantitative Trait Locus (QTL) Detection of Six Economic Traits Using an F<sub>2</sub> Population of the Hybrid from *Saccharinalongissima* and *Saccharina japonica*. *PLoS ONE* 10(5):e0128588.