Berkala **PENELITIAN HAYATI**

ISSN: 08526834 | E-ISSN:2337-389X
Volume 29| No. 2| August|2023

Original Article

Genetic diversity analysis of Wamena Arabica Coffee (*Coffea arabica* L.) based on SSR markers

Nurlita Dianingsih, Arsyam Mawardi*,
Department of Biology, Faculty of Mathematics and Natural Sciences, Cenderawasih University, Papua. Indonesia

Abstract

Wamena is notable for producing and exporting high-quality coffee. However, a molecular method based on its polymorphism and not affected by external factors is required since morphological observations alone are frequently insufficient for identifying the coffee. The objective of this study is to identify Wamena Arabica coffee in the Jayawijaya District using simple sequence repeats (SSRs) molecular markers. This research involved several stages, including DNA isolation and purification, PCR SSR amplification with 5 primers, polymorphism, and heterozygosity level analysis. The results analysis proved that 30 alleles of 32 coffee genotypes were successfully amplified, with fragment sizes ranging from 176 to 553bp. The Car M052 locus (0.29%) had the least polymorphism with only two alleles, while the CarM101 locus (0.82%) had the most with 12 alleles, that high and low polymorphism indicated a measure of PIC scores. Heterozygosity analysis showed that the coffee samples were highly heterozygous. And based on the results of the bootstrap analysis, the phenogram shows that the coffee samples are divided into six clusters, with a cophenetic correlation coefficient (r) of 0.948 (excellent fit). This study proved that all SSR loci succeeded in amplifying 30 alleles and could be identified molecularly based on the genetic variation of the Wamena Arabica coffee genotypes in Papua. It seems highly possible that there is a mix of traits through gene flow and exchange between coffee genotypes, and the heterozygosity in the Arabica population on plantations. As a result, it is very important further analysis to confirm the findings.

Keywords: Genetic diversity, Heterozygosity, SSRs, Wamena Arabica Coffee
Received: October 20, 2022 Revised: February 20, 2023 Accepted: March 28, 2023

Introduction

Coffea arabica L. is an annual plant and member of the Rubiaceae family. With 124 Coffee species (Davis, 2011), Coffee arabica and Coffee canephora are two of the most extensively planted, cultivated, and economically valuable. *Coffea arabica* L is originated in Ethiopia's southwest highlands, which is also where its genetic diversity is concentrated (Lashermes *et al*. 1996; Tesfaye *et al*. 2014), which is the sole species of coffee cultivated in the worldwide. Meanwhile in Indonesia, Wamena coffee is an important commodity cultivated in Wamena mountain plantations. And it is a source of income for thousands of people whose lives rely on Arabica coffee. Moreover, Wamena Arabica coffee is one of the plant commodities with promising economic potentials and a significant strategic function, particularly for export. In addition to its distinguishing fragrance and flavor, arabica coffee contains a range of secondary metabolic components/ In addition to its unique fragrance and flavor, arabica coffee contains a range of secondary metabolic components (AEKI, 2021; ICO, 2021). Wamena coffee has a gentle flavor, a distinct floral scent, and a natural sweet taste with low amounts of acid and caffeine. These make Wamena Arabica Coffee one of the priciest options in its category (Coffeeland Indonesia, 2021). Coffee quality is hugely impacted by the interior quality of the coffee bean (Figueiredo *et al*., 2016), which is assessed by juxtaposing interconnected environmental conditions, genotypes, and processing methods (Borém *et al*., 2016; Malta *et al*., 2020). It is the nature of Arabica coffee. When tested organoleptically, it generates a distinct scent after roasting or heating because of a combination of sour and slightly sweet flavors (Figueiredo *et al*., 2016; Borém *et al*., 2016).

The population of Wamena Arabica coffee is threatened by interspecies and intraspecies competition as a result of forest land clearing for other cultivated plants. If left uncontrolled, arabica coffee germplasm will become extinct. As environmental factors have no effect on DNA, molecular marker mapping can identify genetic variability and rare alleles within populations. Molecular marker mapping can reveal genetic heterogeneity and particular alleles within populations since DNA is unaffected by environmental variables (Salgotra & Stewart, 2020). Identifying and measuring genetic diversity in crop species is crucial for successful genetic resource conservation and enhancement of desirable characteristics in a breeding program.

The first stage in defining swiftly and accurately the elite coffee population when using the molecular approach of SSRs to analyze molecularly the coffee population. Molecular approaches will be utilized to collect reliable and useful genetic data on arabica coffee germplasm and allele variants. Molecular markers are a prominent method for recognizing genetic differences in many species (Idrees & Irshad, 2014).

SSR molecular markers have been employed in a range of genetic diversity and plant variety identification

* Corresponding Author:
Arsyam Mawardi
Department of Biology, Faculty of Mathematics and Natural Sciences, Cenderawasih University, Indonesia
Phone: +6285239558242- Fax: -
e-mail: mawardiarsyam@gmail.com

http://dx.doi.org/10.23869/bphjbr.29.2.20232
Published by © PBI East Java. Open Access www.berkalahayati.org

investigations (Idrees *et al*. 2021). This codominant marker identifies substantial allele variation and can distinguish between closely related plant accessions more effectively than other molecular markers (Mason, 2015; Hussain & Nisar, 2020). SSRs may be detected using silver-stained Polyacrilamyde Gel Electrophoresis more than sensitive, however agarose cannot be used since it cannot distinguish size variations of up to one base pair (Liu *et al*., 2017: Huang, 2018).

Ferreira *et al*., released a follow-up research, Genetic Diversity of Coffea arabica, in 2021. This research explains why *C. arabica* L. is a native coffee species that originated in Abyssinia, which is now Ethiopia. The benefits of breeding to generate new varieties for the worldwide market are directly tied to the economic implications of C. arabica genetic diversity. The genetic resources of *Coffea arabica* L. are estimated to be worth USD 420 million. the extent of trait and genetic diversity is essential for steering genotype crosses with the objective of producing new, high-value varieties. This research explains the economic value of Arabica coffee (*C. arabica* L), particularly in Brazil, one of the large producer country that distribute arabica coffee, as well as a quick overview of the world's main germplasm bank (Ferreira *et al*., 2021) demonstrate how genetic diversity studies based on morphology, agronomic features, and molecular markers help to the production of novel varieties.

The further study, using simple sequence repeat markers, Benti *et al*. published their findings in 2021 under the title, "Genetic diversity among commercial arabica coffee (*Coffea arabica* L.) types in Ethiopia. This study employed 42 commercial Arabica coffee varieties produced and released for production by the Ethiopian Agricultural Research Institute's Jimma Agricultural Research Center (JARC), as well as information on the level of genetic diversity among these scarce varieties. It was claimed that the molecular markers of these SSRs could reveal polymorphisms between varieties (Benti *et al*., 2021). Among varieties, there was a high mean number of polymorphic alleles (7.5) and polymorphic information content (PIC = 80%) per locus. Using the Jaccard similarity coefficient, the genetic similarity between varieties ranged from 0.14 to 0.78, with a mean of 0.38. In 92% of the possible paired combinations, the range of genetic similarity coefficient values ranged from 0.14 to 0.50, indicating the presence of genetic kinship among the varieties.

There hasn't been much studied of the genetic diversity studies of the Wamena arabica coffee and other coffee producing centers in Papua. The study aims to identify and analyze Wamena arabica coffee in the Jayawijaya district. This study is important so because results will serve as a scientific reference for molecular data on the genetic composition of Wamena arabica coffee, in the form of specific alleles and genetic variations capable of displaying polymorphisms, which can be used by researchers and the government in conservation efforts for Wamena arabica coffee, Papua Province. Hence, we used SSR markers in the present study to generate information on the level of genetic diversity and relationships among 32 widely cultivated commercial arabica coffee varieties grown in Wamena, As a result expected to provide information that will be used as a reference in differentiating Wamena coffee at international coffee market of Wamena arabica coffee as one of the world's top coffee producers.

Methods

A Total 32 Arabica coffee leaf samples were collected and observed as sample from Wamena plantations in Jayawijaya district, including the villages of Wollo, Assolokobal, and Tiom, Wamena Papua (table 1). The organ chosen for molecular study is the third leaf of a plagiotropic coffee branch, which was cut using sterile scissors. The immature leaves are then wrapped in plastic and either placed in an ice thermos or given an ice pack. The leaf samples were subsequently sent to Cenderawasih University's Biology Laboratory for DNA molecular analysis. The research was conducted between December 2021 and March 2022.

Table 1. Wamena coffee samples for genetic diversity analysis

No	Genotype Code	Location
1	WM1A2	Assolokobal
2	WM1A4	
3	WM1A3	
4	WM1A4	
5	WM1A12	
6	WM1A14	
7	WM2A3	
8	WM1A11	
9	WM1A5	
10	WM1A6	
11	WM1A3	
12	WM1A2	
13	WM2A1	
1	WM1T1	Tiom
2	WM1T2	
3	WM1T5	
4	WM1T12	
5	WM1T3	
6	WM1T4	
7	WM2T1	
8	WM2T3	
9	WM2T6	
1	WM117	Wollo
2	WM1W6	
3	WM1W3	
4	WM1W12	
5	WM1W13	
6	WM1W1	
7	WM1W23	
8	WM1W4	
9	WM1W5	
10	WM1W6	

DNA isolation and purification

Fresh leaves of Wamena Arabica coffee, liquid nitrogen, TAE buffer 1X, CTAB buffer, Etidium Bhromide, β-Mercaptoethanol, 70% alcohol, TE buffer, Agarose powder and loading dye were used to isolate DNA. Buffer PCR, Primer forward and reverse, DNA taq Polymerase enzyme, genomic DNA, nuclease free water, DNA marker/ladder, agarose powder, and polyacrylamide were utilized in the PCR analysis. Isolation DNA coffee enforced the CTAB procedure (Khan *et al*., 2004; Mawardi & Simonapendi, 2016). A

total of 250 mg of fresh leaves were crushed in liquid nitrogen and 20 mg of polyvinylpyrrolidone (PVP) was added. Transfer into a tube filled with 600 µl CTAB and 5% β-Mercaptoethanol, the solution was incubated in water for 30 min at 60°C. During incubation the solution was reversed several times and 600µl of Chloroform:Isoamyl-alcohol was added and centrifuged for 30 min at 10.000 rpm. The top layer was discarded and 600 µl of Chloroform : Isoamyl-alcohol was added. The solution was reversed as before, the supernatant was mixed with 600 µl of cold isopropanol and sodium acetate, frozen about 30 min for precipitation. DNA. The sample was centrifuged for 2 min, supernatant was discarded without damaging the pellet. Washed with 70% ethanol then centrifuged again during 30 min, the supernatant was discarded. The remaining pellet was dried and dissolved in 50µL TE buffer, the genomic DNA was stored in a freezer -20 °C. The DNA quality and quantity tests were performed using an electrophoresis technique with 1.5% agarose gel.

SSRs amplification

Five (5) pairs of SSR primers are used for DNA amplication. The fundamental core structure of the SSR used regarding on reported coffee molecular studies. The five primers were selected due to the potential for polymorphism ensuing from the amplification results, as was done by (Lashermes *et al*., 1999; Hue, 2005; and Missio *et al*., 2010) (table 2)

Polymerase Chain Reaction (PCR) was performed in 90µl reaction mixture containing *Nuclease free water 35* µL, *Green Go taq master mix 5* µL, Primer *Forward 2,5* µL, Primer *Reverse* µL, DNA template 2 µL. The total volume of PCR reaction was divided evenly into five PCR tubes, then each tube was filled with 2 µL The PCR conditions followed were: initial denaturation at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 56°C for 1 min (*varies for each primer), and extension72°C for 1 min with a final extension at 72°C for 5 min.

PCR-SSRs product detection

Approximately 4.0 µL of the PCR results were placed in the well. The power supply was set to a constant power of 60 watts, and the temperature around 50°C. After running PolyAcrylamide Gel Electrophoresis 4.5% for 55-60 min, the polyacrylamide gel on the glass plate was stained with silver staining.

Polymorphism and heterozygosity analysis

Scoring commences with labeling markers with names, dates, and other relevant information. The marker and sample fragments are labelled a, b, c, d, e, f, g, h, i, j, k, l, and m, starting from the largest fragment to the smallest fragment. Each single band under a fragment is named regarding to its letter marker followed by a number. Faint bands or bands that are too difficult to score are designated as missing data; bright bands are scored. Score each allele with a score 1 (with a band), and a score 0 (no band). If in doubt, give a score 9 for missing data (Warburton and Crossa, 2000). The formula proposed by Nei, (1973) was used to measure heterozygosity each Wamena coffee genotype:

$$\text{Percentage of heterozygosity} = \frac{\text{Number of heterozygous loci}}{\text{The total of SSR loci used}} \text{ X } 100\ \%$$

Analysis

Genetic similarity analysis (Lee, 1998) was performed by submitting SSRs data profiles into the NTSYSPc2.1 program and running the Unweighted Pair Group Method Using Arithmetic Average (UPGMA) with the Jaccard coefficient (Rohlf, 2009).

Table 2. Oligonucleotide Primer PCR-SSRs

No	Primer	Sequence	SSRS primer product range (bp)
1	Car M101	F: TAT GTC TCT AAC TTT CTA TTT T R: AGA GAC TAC ATT TAC ACA CAG AAG A	178 – 321
2	Car M096	F: TAC TGG GGA AGA ATT TAT CAT C R: TTA GGC CAT CCA AGA GTA TTC	176 – 343
3	Car M092	F: TAC TGG GGA AGA ATT TAT ACT C R: TTA GGC CAT CCA AGA GTA TTC	383 – 425
4	Car M051	F: GAT GTG GAG GAG GCT GCT GCT GAA R: TAG GGC GCC ATC TGG TAG GGT TGT	275 – 288
5	Car M052	F: AGC AGC TGC AGC CACAAC A R: GAG TAA AG CCC CAG AGC GTA ACC T	296 – 309

Results

The isolation of coffee DNA resulted in a wide range of quality and quantity of genomic DNA isolates. Calculations with a spectrophotometer yielded genomic DNA concentrations ranging from 1213 ng/l to 2998 ng/l. The ability to obtain high concentrations is related to the precision with which the method of isolation and extraction of genomic DNA is chosen. Figure 1 clearly shows the level of luminescence produced and the thickness of the DNA when run in conjunction with i.e. λ-DNA, which is similar to the quality of the DNA. The use of a DNA concentration of 10 ng/l and a reaction volume of 2.5 l/reaction is the best choice for ensuring that the DNA amplification process at the SSRs stage runs smoothly and accurately tells the profile of a bright band.

Amplification of five SSR loci using 32 copies of genomic DNA resulted in a total of 30 alleles. The lowest allele is the primer CarM052 with a PIC value of 0.29. CarM101 was the highest with a total of 12 alleles, and a PIC value of 0.82. The average PIC value of the 5 primers was 0.57, indicating that the genetic diversity of the samples analysed have high genetic diversity. The sizes of the amplification products range from 176.37 to 553 bp. The results showed that the five SSR primers were used, all primers produced polymorphic bands and demonstrated the existence of genetic diversity in Arabica coffee populations in Wamena Papua.

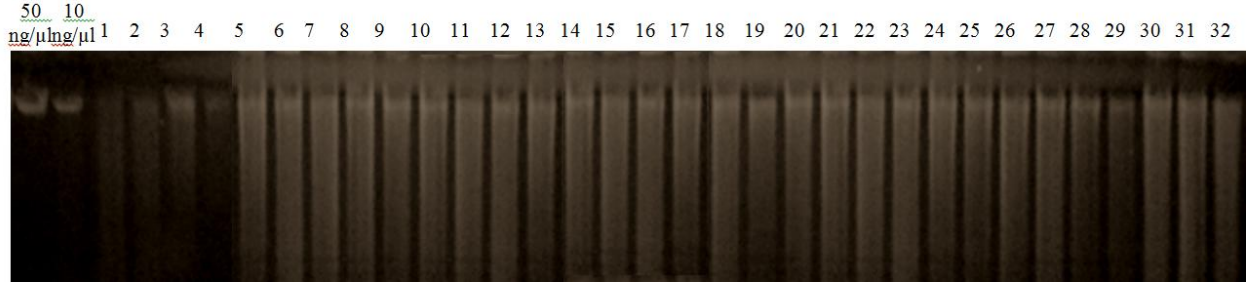

Figure 1. Electroferogram of Wamena coffee genomic DNA isolate. Notes: DNA lambda (λ –DNA) 50 ng/µl, and 10 ng/µl. 1-32 : Genomics of 32 coffee genotypes

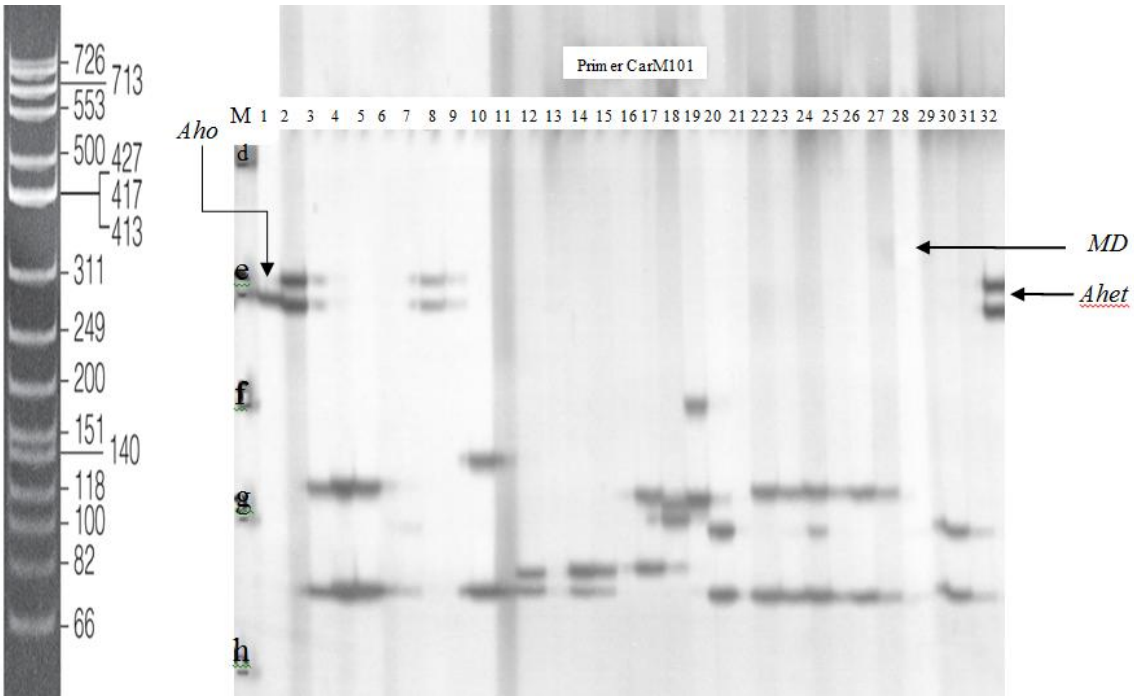

Figure 2. Amplicon electroferogram with SSRs Car M101 primer. Notes: M: DNA ladder marker, Aho: Homozygous allele, Ahet: Heterozygous allele, MD: Missing data, 1-32 (sample serial number code)

Figure 2 showed that the marker used is polymorphic. The electroferogram below shows the attachment and elongation of the CarM101 primer inside the genomic DNA sequence of 32 genotypes. The presence of specific alleles in only one genotype indicates that genotype has the potential for specific traits. However, because few primers and coffee genotypes have been analyzed, this finding cannot be confirmed that the allele is a specific character. In this study, the specific allele of Assolokobal Wamena was detected using Car096 loci with sizes ranging 180 and 200 bp, CarM101 loci specific to arabica Typica and robusta genotypes with four size combinations, such as 311(n:1), 250(n:1), and 200bp (n:2). At 300 bp, the CarM051 locus was specific for the Assolokobal robusta genotype, whereas the CarM92 locus was specific for the arabica genotype at 550 bp.

Table 3. Polymorphism analysis of 5 primer SSRs in 32 coffee genotypes

No	Primer	Total Allele	Size range (bp)	(PIC)
1	Car M101	12	178-321	0,82
2	Car M096	8	176-343	0,73
3	Car M092	5	449-553	0,70
4	Car M051	3	269-303	0,33
5	Car M052	2	296-309	0,29
	Total	30		2.87
	Average	6	176-553	0,57

Notes: PIC: polymorphism Information Content; bp: base pairs

Table 4. The genetic conditions of Wamena typica coffee types.

No	Code	# of het	% het	Type	Location
1	WM1_W1	3	60	*Typica*	Wollo01
2	WM1_A1	3	60	*Typica*	Assolokobal1
3	WM1_W12	2	40	*Typica*	Wollo 12
4	WM1_W23	2	40	*Typica*	Wollo 23
5	WM1_A5	4	80	*Typica*	Assolokobal 5

Notes: #of het: number of heterozygosities; %het: heterozygozygosity percentage; WM1 W1: the first sample code for an arabica coffee of the Wamena typica variety from Wollo. A1 and A5: Assolokobal 1 and 5; W12 and W23 are Wollo 12 and 23, respectively.

Table 5. The specifications of detected alleles

No	SSR loci	Size range (bp)	Type
1.	Car M101	311	Arabika *Typica*
		250	Robusta
		200	Arabika *Typica*
		200	Arabika *Typica*
2	Car M096	200	Arabika *Typica*
		180	Arabika *Typica*
3	Car M092	550	Arabika *Typica*
4	Car M051	300	Robusta

Notes: Loci: similar to primer; bp: base pairs

Referring to the phenogram results, the level of grouping confidence derived from the bootstrapping analysis based on the UPGMA. There is undeniable consistency. Confidence levels for the three coffee groups were 95.8%, 91.7%, and 97.3%, respectively. The bootsrapping analysis produces a high percentage of consistency and confidence above 90%, thus the results will not change significantly even when using far more primers. The results of the analysis showed that there were genetic variations between the Wamena Arabica coffee genotypes, *Typica* and non-*Typica*, so that six clusters were formed. One of the clusters revealed genetic similarity between members of the *Typica* Arabica group. There is one cluster (cluster IV) which contains only typhus from the Assolokobal Wollo and Tiom areas. While the other clusters describe high genetic kinship between *Typica*, USDA, Catimor and Linies. The mixing of different coffee types in cluster is due to trait mixing through gene flow and exchange between coffee genotypes, as well as heterozygosity in the arabica population on plantations, therefore further study is needed to confirm these findings.

The phenogram obtained from the bootstrapping analysis showed that the coffee samples were grouped into two large clades and six clusters, which is typica arabica clustering with a confidence level of 89.0. Meanwhile, robusta has a confidence level of 91.8. USDA Assolokobal arabica clustered in a group with the typica arabica genotype with a confidence 25.8, indicating that USDA is likely to separate from the group if additional primers are added. Similarly, typica arabica from Assolokobal and Tiom clustered in one cluster with a confidence level of 23.5 of experiencing the same situation. Furthermore, the confidence value for the Wollo and Asssolokobal types in one cluster is 79.3, indicating that the addition of primers will remain stable, maintaining them in the same position in one cluster.

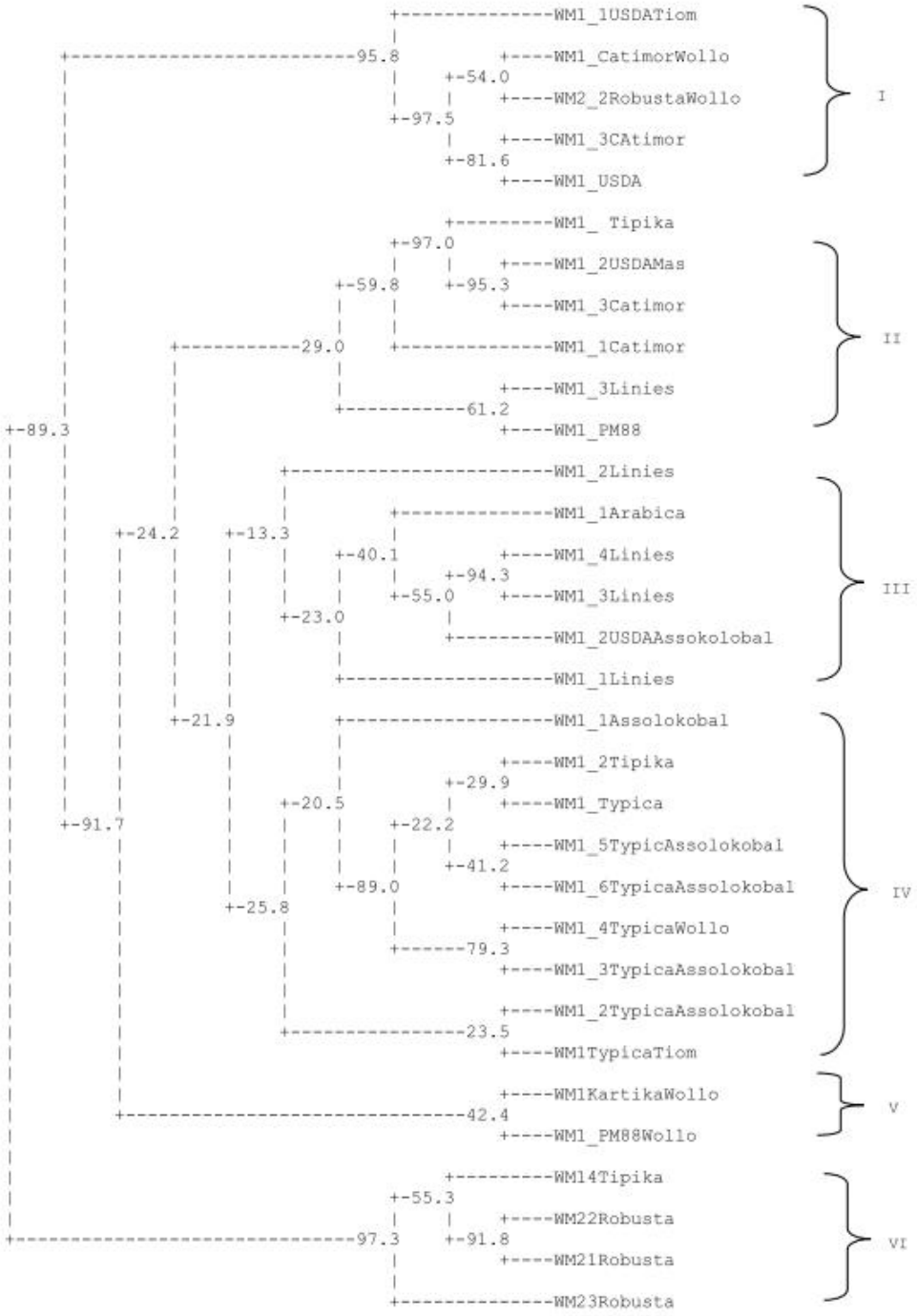

Figure 3. Phenogram of 32 Wamena coffee genotypes.

Discussion

There are two forms of genomic DNA: DNA with poor quality appears as smeared white lines, which are isolated products, and DNA with superior quality appears as a thick white line that is intact. Smears are caused by an inefficient sample grinding method and a high temperature. The quality of DNA revealed that the isolated DNA had a high concentration. This demonstrates that the method used to isolate DNA samples is correct and precise.

The visual appearance of DNA can be used to determine DNA quality by indicating the thickness and luminosity of the sample DNA to standard λ-DNA DNA as a comparison. The concentration of DNA in a research sample is heavily influenced by the isolation technique used, as well as the level of experience and skill with which the isolation or extraction process was handled (Thieman & Palladino, 2014). Cell hemolysis was caused

by grinding the cells and adding a hypotonic buffer solution. As a DNAse cofactor, the EDTA content forms metal ion complexes such as Mg2+. Meanwhile, the CTAB solution binds to the DNA complex, and the Tris NaCl solution stabilizes it.

SSR molecular markers do not always require a high concentration of DNA to function properly, this is one of the benefits that this method with a low concentration can still operate effectively. The PCR-SSRs amplification process on the genomic DNA template was successful. The electropherogram produced by 4.5% polyacrylamide gel electrophoresis revealed the banding pattern profile in each genotype.

Figure 2 shows three conditions: a band pattern for the heterozygous allele category, a band pattern for the homozygous allele and no band (missing data). The electropherogram of the CarM101 primer consists of these three conditions. The absence of DNA bands was caused by samples of genetic material that were not complementary to the type of primer. Errors in PCR composition mixing, inhomogeneity, and contamination also were triggers. The heterozygosity, homozygosity, and missing allelic data have impact on the data profile, which has implications for the clustering of existing genotypes as well as genotype instability in the grouping. The SSRs data profile was obtained using binary data by evaluating the presence of bands, specifically by assigning a value of 1 if a band appeared and a value of 0 if no band appeared. Table 3 presents the findings of polymorphism analysis using five SSR primers on thirty two genotypes of Wamena coffee for further study, as well as an illustration of its species diversity.

An investigated the genetic diversity of 84 Arabica coffee genotypes using six SSR loci. The results revealed that one of the six primers used failed to amplify the coffee genome (Hue, 205). On the other hand, Vieira *et al*., (2010) used 127 SSR primers to identify coffee in Brazil, but only twenty two were effective in amplifying genome and resulting in polymorphism. Which are CarM051, CarM052, CarM101, CarM092, and CarM096 used. The using CarM101 generated the highest number of twelve alleles, the number of alleles per locus varied in our study included two to twelve alleles. Furthermore, despite the fact that twenty two primers produced 55 alleles of the Brazilian coffee genome, five primers in our study was able to amplify thirty alleles of Wamena coffee, indicating that the genotype of Wamena coffee is quite polymorphic.

SSR molecular markers are more frequently used to detect high and low levels of polymorphism. The relationship between the length of the SSR and the frequency of polymorphisms demonstrates that a decrease in polymorphism frequency is associated with an increase in repetitive sequences (Zhao *et al*., 2012; Jingade *et al*. 2019). Meanwhile, the SSRS marker has a co-dominant character that can detect the presence of homozygous and heterozygous alleles. This is consistent with the findings proposed by (Mason, 2015), SSRs have been the most widely used markers for genotyping plants over the last 20 years because they are codominant, extremely informative, multi-allele genetic markers that are experimentally reproducible and transferable between related species (Berecha *et al.,*2014). Because each genotype had a large number of heterozygous alleles, detecting the heterozygosity degree at the SSR locus resulted in a fairly high proportion of heterozygosity. The percentage of heterozygosity in all samples ranged from 20% to 80%, depending on the number of primers used. The Robusta Wollo genotype had the lowest percentage of heterozygozity, with a value of 20%. This implies that Wamena Coffee requires additional primers to analyze its polymorphism. While the Assokolobal genotype has the highest percentage of heterozygosity (80%). This indicates that the total number of primers used is adequate to describe the genetic diversity. Specifically, regarding heterozygosity and varieties of coffee genotypes that correlate with the highest genetic diversity are shown in Table 4 which shows that the level of heterozygosity for the genotypes of Wamena coffee of typica varieties has been identified 40 to 80%.

Heterozygosity data analysis revealed that the coffee sample had a high heterozygous state of genes, with heterozygozygosity values of 40/2 (WM1 W12, WM1 W23), 60/3 (WM1 W1, WM1 A1), and 80/4 (WM1 A5), respectively. The missing data from the DNA band profile is also significant. The WM2 sample contained a DNA segment that did not correspond to any of the five primers used. This is indicated by the fact that 20% of the 30 alleles identified have missing data. The higher the missing data value, the greater the likelihood that no match exists between the sample DNA segment and the primary segment.

The use of primers that have been tested repeatedly on all samples, if it produces a lot of missing data, indicates that the primer is of poor quality. Five primers were successfully used in this study to describe polymorphisms in coffee genotypes. This is related to primers' ability to produce high levels of polymorphism, which will allow many genotypes to be distinguished, whether they are close or far apart. SSRs, on the other hand, was more accurate in identifying variations in cultivars (Garland *et al*., 1999), where 14 cultivars were tested, 21.5% heterogeneous based on phenotype and RLFP data, and 64% heterogeneous based on SSR marker data. The results of the heterozygosity analysis also revealed four SSR loci capable of amplifying eight specific alleles, i.e. alleles that were present in only one genotype of all coffee genotypes.

Individual identification, particularly for specific alleles, is accomplished through DNA extraction per individual plant per genotype. Specific alleles will be useful as individual identities or genotype fingerprints. Furthermore, if the specific character is an interesting character, the genotype must be preserved. The DNA used in this study was the result of DNA extraction per individual plant per genotype.

The NTSYSPc2.1 program's (Rohlf, 2009) analysis of SSRs genetic characteristic data profiles resulted an

output of genetic links across 32 coffee genotypes using 5 primers. The analysis constructed heterotic groups, which were visualized using a phenogram. The emergence of opportunities for the shape of this phenogram is due to inconsistent genotypes and several factors such as: 1. the number of markers (primer) used, 2. the impact of missing data, 3 the impact of heterozygous alleles, and 4. the scoring method' accuracy. The analysis results also show that there are inconsistent genotypes, as previously described. As a result, group formation will differ.

A phenogram obtained from the matrix of genetic similarity reveals group formation and kinship patterns. The phenogram was constructed using UPGMA at 5 primer SSR-identified genotypic loci. A value of 0.948 was determined for the cophenetic correlation coefficient (r). This value is categorized as an excellent fit (Rohlf, 2009). This value shows that there were adequate primers used in this investigation to create the phenogram. The value of the cophenetic correlation coefficient, according to (Peijic *et al*., 1998), describes how much accuracy the genotypic grouping can produce based on the estimated genetic similarity between the identified genotypes and the markers used. The cophenetic coefficient value should be greater than 0.9 for a more accurate visualization of the phenogram, which is classified as excellent fit.

The degree of genetic similarity obtained by pairing or complementation of each coffee genotype with other genotypes. The calculation results show that the genotypes of Wamena coffee are divided into two major clusters with a relatively large genetic distance, consisting of arabica coffee and robusta coffee. Arabica arabica is classified into Typica and non-Typica groups, which include catimor, USDA, PM88, kartika, and Linies. Referring to the calculation of genetic similarity, the liberica group has very low similarity with the Robusta group, Robusta Tiom (WM2 T3) vs Robusta Wollo (WM2 T6) and Linies Assolokobal (WM1 A14) vs Robusta Wollo (WM2 W3), with the lowest similarity or indicating indicating the high the genetic distance about 0.91. This indicates how far the kinship between Arabica and Robusta. There is a robusta coffee genotype that has a high genetic similarity to the arabica coffee group, with genetic similarity ranging from 0.31 between Tiom robusta (WM2 T1) and Wollo catimor arabica (WM1 W17) to 0.54 between Assolokobal robusta (WM2 A1) and Arabica typica Tiom). This suggests that the genetic gap between the robusta and arabica groupings is just 0.66 to 0.89, showing that robusta and arabica coffees are descended from the same parents. Robusta coffee is one of arabica coffee's two seniors, Robusta coffee is one of arabica coffees 2n elders.

The results of the analysis showed that there were genetic variations between the Wamena Arabica coffee genotypes, *Typica* and non-*Typica*, so that six clusters were formed. One of the clusters showed genetic closeness between several individuals in the *Typica* Arabica population. There is one cluster (cluster IV) which contains only typica from the Assolokobal Wollo and Tiom areas. While the other clusters describe high genetic kinship between *Typica*, USDA, Catimor and Linies.

The mixing of different coffee types in clusters is assumed to be due to trait mixing through gene flow and exchange across coffee genotypes, as well as heterozygosity in the arabica population on plantations, therefore further research is needed to confirm these findings.

The Arabica coffee group has a generally low genetic diversity, which causes all coffee varieties in the same group to be classified together into a single cluster. Nevertheless, due to the heterozygosity factor that the coffee's parents possess, it is feasible to classify some coffee varieties individually (Steiger *et al*., 2002; Tounekti *et al*. 2017) and the occurrence of natural cross-pollination leading in gene flow across cultivars (Geleta *et al* 2012). Conversely, it is possible that Arabica coffee varieties grown in different growing areas have the same genetic identity. Both species are the result of tissue culture propagation.

Further phenogram analysis revealed significant results on the hypothesis that Wamena's coffee population was a premium variety of typica arabica that survived in the long term. The Wamena arabica coffee population has a relatively high genetic diversity. There are coffee varietals that cluster together but are not linear with one another. Yet, the study's findings reveal that varieties of typica arabica coffee may still be found in Jayawijaya Regency. Genetic polymorphisms in the DNA of Wamena Arabica coffee cause this variation (Budi & Mawardi, 2021). Sousa *et al*., (2017) support this by saying that while the number of polymorphic loci was the same for certain cultivars, diversity was also identified in their populations.

Variations and polymorphisms are also associated with the likelihood of phenogram formation, with the bootstrapping study producing confidence in the grouping of the coffee genotypes studied. Bootstrapping analysis was performed to determine the number of loci employed in the categorization procedure, which was constant across all coffee genotypes examined. The phenogram results that we obtained are consistent with reports from the international coffee agency that the arabica coffee samples in Wamena, Papua, belong to elite coffee varieties. According to the International Coffee Organization, the arabica variety vanished at the end of the 18th century, leaving only the typica variety. According to the phenogram, the USDA, Catimor, Linies, and PM 88 varieties were more separated from the typica clusters. This demonstrates that the typica and the other four varieties have a distant relationship. As previously stated, typica arabica is an old arabica that differs greatly from young arabicas such as catimor arabica, USDA arabica, PM88 arabica, and Linies arabica. All of these samples are not pure arabica typica derivatives because they have undergone genetic changes as a result of cross-pollination with several other types of coffee. Arabica catimor was a cross between typica and Bourbon, USDA was indeed a cross between typica and a Liberika, Arabica PM 88 was a cross between typica and a USDA, and Linies was a cross between typica and a canephor

(robusta). This strongly suggests that rabica coffee varieties, particularly Arabica Typica, can still be found in Wamena, Jayawijaya Regency, Papua Province. Genetic variation is caused by genetic polymorphism in the DNA composition of Wamena arabica coffee.

Based on the findings, it is possible to conclude that 30 alleles were successfully amplified from 32 coffee genotypes, with fragment sizes ranging from 176 to 553bp. The CarM052 locus had the lowest polymorphisms with only two alleles, while the CarM101 locus had the most with twelve alleles, it means that high and low polymorphisms indicate the size of the PIC score.

The average PIC score is 0.57 (Table 3), placing it in the top category (Zhang *et al*. 2014). The research results obtained were slightly different the findings from Missio *et al*. (2010) that demonstrated a relative low PIC value (0.22) in six Arabica coffee types using 33 pairs of primers. In this study CarM101 primer had the highest PIC (0.82) compared to the others. This finding proves that the CarM 101 primer can be applied to analyze Wamena Arabica coffee and allows SSRs marker primers to have the potential to be used in plant breeding.

According to the findings of this study, the average value of genetic parameters calculated from the acquisition of 5 primer SSR allele data is much higher than the values previously reported for Arabica coffee varieties elsewhere using SSR markers, indicating a high genetic diversity among coffee varieties Wamena arabica. Teressa *et al*. (2010) reported low genetic diversity and polymorphism in 55 coffee varieties in France using 32 SSR markers, which differed from our findings. The allele average is 2.8, and the polymorphism is 32%. Geleta *et al*. (2012) reported low genetic variation among 8 Nicaraguan arabica coffee varieties, 12 SSR markers yielded an average of 2.0 alleles, and 42% polymorphism. Whereas Al-Murish *et al*. (2013) showed low amounts of genetic diversity in 17 Arabica coffees in Yemen, involving 58 SSRs bases with an average of 2.5 alleles, and 0.32 PIC

Heterozygosity data analysis revealed that the coffee samples were highly heterozygous, with heterozygosity samples ranged from 40/2 (WM1 W12, WM1 W23), 60/3 (WM1 W1, WM1 A1), and 80/4 (WM1 A5). The phenogram confirmed that the coffee samples were divided into six clusters based on the results of the bootstrapping analysis, with a cophenetic correlation coefficient (r) of 0.948 (excellent fit). This study proved that all SSR loci were successfully amplified and identified molecularly based on the degree of genetic variation of the Wamena Papua Arabica coffee genotypes. The results also showed that the markers have the potential to be used in genotyping, quantifying genetic diversity, and distinguishing between elite breeding lines using fingerprinting.

This is in line with earlier discoveries made by a number of scientists who documented the effective use of SSR markers to the diversity study of several genetic subgroups of *Coffea Arabica* L (Sousa *et al*., 2017; Pruvot-Woehl *et al*., 2020; Montagnon *et al*., 2021).

Acknowledgements

Awards and thanks are extended to the Cenderawasih University Research and Community Service Institute (LPPM), funding the research activities through the PNBP financing scheme.

References

AEKI, 2021. Asosiasi Eksportir & Industri Kopi Indonesia. https://www.aeki-aice.org/

Al-Murish T. M., Elshafei A. A., Al-Doss,A. A., Barakat M. N., 2013. Genetic diversity of coffee (*Coffea arabica L.*) in Yemen via SRAP, TRAP and SSR markers. *J Food Agric Environ*, 11(2): 411-416.

Benti T., Gebre E., Tesfaye K., Berecha G., Lashermes P., Kyallo M., Kouadio Yao N., 2021. Genetic diversity among commercial arabica coffee (*Coffea arabica L.*) varieties in Ethiopia using simple sequence repeat markers. *Journal of Crop Improvement*, 35(2): 147–168.

Berecha G., Aerts R., Vandepitte K., Van Glabeke, S., Muys B., Roldán-Ruiz I., Honnay O. 2014. Effects of forest management on mating patterns, pollen flow and intergenerational transfer of genetic diversity in wild Arabica coffee (*Coffea arabica L.*) from Afromontane rainforests. *Biological Journal of the Linnean Society*, 112(1): 76-88.

Budi I. M., & Mawardi A., 2021. Identifikasi Molekular Kekerabatan Genetik Kopi Wamena Berbasis Marka Random Amplified of Polymorphic DNA (RAPD). *Jurnal Biologi Papua*, 13(1): 8–18.

Coffeeland Indonesia, 2021. Arabika Wamena Mengenal Kopi daerah Ujung Timur Indonesia – Coffeeland. https://coffeeland.co.id/arabika-wamena-mengenal-kopi-daerah-ujung-timur-indonesia/

Davis A. P. , 2011. Psilanthus Mannii, the Type Species of Psilanthus, Transferred to Coffea. *Nordic Journal of Botany,* 29: 471–472.

Ferreira J., Caixeta E., Caniato F. F., Setotaw T., Sant'Ana G. C., Ferreira L. M., 2021. Genetic Diversity of Coffea arabica. *Genetic Variation*. https://doi.org/10.5772/intechopen.94744

Figueiredo L. P., Ribeiro F. C., Giomo G. S., Salva T. J., 2016. The relationship between organic acids, sucrose and the quality of specialty coffees. *African Journal of Agricultural Research*, 11(8): 709-717.

Garland S. H., Lewin L., Abedinia M., Henry R., &Blakeney A., 1999. The use of Microsatellite Polymorphisms for the Identification of Australian Breeding Lines of Rice (*Oryza sativa L.*). *Euphytica*, 108(1): 53–63.

Geleta M., Herrera I., Monzón A., Bryngelsson T., 2012. Genetic diversity of arabica coffee (*Coffea arabica L.*) in Nicaragua as estimated by simple sequence repeat markers. *The Scientific World Journal*, *2012*, 939820.

Gimase J. M., Thagana W. M., Kirubi D. T., Gichuru E. K., Kathurima C. W., 2014. Beverage quality and biochemical attributes of arabusta coffee (*C. arabica L. x C. canephora Pierre*) and their parental genotypes. *African Journal of Food Science*, 8(9): 456-464.

Hue T. M. T., 2005. Genetic variation in cultivated coffee (*Coffea arabica L* .) accessions in northern New South Genetic Variation In Cultivated Coffee (*Coffea arabica L* .) Accessions in by Tran Thi Minh Hue BSc Centre for Plant Conservation Genetics.

Hussain H., & Nisar M., 2020. Assessment of plant genetic variations using molecular markers : A review. *Journal of Applied Biology & Biotechnology*, 8(05): 99-109.

ICO, 2021. Coffee Market ReportSeptember 2021. September, 1–11. https://www.ico.org/Market-Report-20-21-e.asp

Idrees M., & Irshad M., 2014. Molecular Markers in plants for analysis of genetic diversity: A review. *European Academic Research*, 2(1): 1513–1540.

Idrees, M., Irshad, M., Pathak, L.M., Tariq, A., & Naeem R. (2021). Assessment of Microsatellite Markers (SSRs) for genetic diversity in *Asparagus Officinalis* L. and allied species. *Bangladesh Journal of Botany* 50(3): 595–602.

Jan C., Dawson D. A., Altringham J. D., Burke T., Butlin R. K., 2012. Development of conserved microsatellite markers of high cross-species utility in bat species (Vespertilionidae, Chiroptera, Mammalia). *Molecular Ecology Resources*, 12(3): 532–548.

Jingade P., Huded A. K., Kosaraju B., Mishra M. K., 2019. Diversity genotyping of Indian coffee (*Coffea arabica L.*) germplasm accessions by using SRAP markers. *Journal of Crop Improvement*, 33(3): 327-345

Huang L., Deng X., Li R., Xia Y., Bai G., Siddique K. H. M., Guo P., 2018. A fast silver staining protocol enabling simple and efficient detection of SSR markers using a non-denaturing polyacrylamide gel. *Journal of visualized experiments*, 134: 1-7.

Khan I. A., Awan F. S., Ahmad A., Khan A. A., 2004. A modified mini-prep method for economical and rapid extraction of genomic DNA in plants. *Plant Molecular Biology Reporter*, 22(1): 624892.

Lashermes P., Trouslot P., Anthony F., Combes M. C., Charrie A., 1996. Genetic Diversity for RAPD Markers between Cultivated and Wild Accessions of Coffea Arabica. *Euphytica*, 87: 59–64.

Lashermes P., Combes M. C., 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.

Liu W., Li R., Ayalew H., Xia Y., Bai G., Yan, G., Siddique K. H. M., Guo P., 2017. Development of a simple and effective silver staining protocol for detection of DNA fragments. *Electrophoresis*, 38(8): 1175–1178.

Lee M., 1998. DNA Markers for Detecting Genetic Realtionship among Germplasm Revealed for Establishing Heterotic Groups. Presented at The Maize Training Course.

Malta M. R., de Oliveira Fassio L., Liska G. R., Carvalho G. R., Pereira A. A., Botelho C. E., Pereira R. G. F. A., 2020. Discrimination of genotypes coffee by chemical composition of the beans: Potential markers in natural coffees. *Food Research International*, *134*, 109219.

Mason A. S., 2015. SSR Genotyping. Methods in molecular biology (Clifton, N.J.). 1245, 77–89. https://doi.org/10.1007/978-1-4939-1966-6

Mawardi A., & Simonapendi M. L., 2016. Uji Efektivitas Metode Isolasi DNA Genom Kopi Arabika (*Coffea arabica* L.) Asal Kabupaten Jayawijaya. *Jurnal Biologi Papua*, 8(1): 7–12.

Missio R. F., Caixeta E. T., Zambolim E. M., Zambolim L., Cruz C. D., Sakiyama N. S., 2010. Polymorphic information content of SSR markers for Coffea spp. *Crop Breeding and Applied Biotechnology*, 10(1): 89–94.

Montagnon C., Mahyoub A., Solano W., &Sheibani F., 2021. Unveiling a unique genetic diversity of cultivated Coffea arabica L. in its main domestication center: Yemen. Genetic Resources *and Crop Evolution,* 68 (6): 2411-2422.

Nadeem M. A., Nawaz M. A., Shahid M. Q., Doğan Y., Comertpay G., Yıldız M., Hatipoğlu R., Ahmad F., Alsaleh A., Labhane N., Özkan H., Chung G., Baloch F. S., 2018. DNA Molecular Markers in Plant Breeding: Current Status and Recent Advancements in Genomic Selection and Genome Editing. *Biotechnology and Biotechnological Equipment*, 32(2): 261–285.

Nei M., 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70(12 (I)), 3321–3323.

Peijic I., Ajmone-Marsan P., Morgant M., Kozumplick V., Castiglioni P., Taramino G., Motto M., 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, andAFLP. *Theor Appl Genet*, 97: 1248–1255.

Pruvot-Woehl S., Krishnan S., Solano W., Schilling T., Toniutti L., Bertrand B., Montagnon C., 2020. Authentication *Coffea arabica* Varieties through DNA Fingerprinting and its Significance for the Coffee Sector. Journal of AOAC International, 103(2): 325–334Rohlf F. J., 2009. *NTSYS-pc - Numerical Taxonomy and Multivariate Analysis System Version 2.2*. Applied Biostatistics Inc.

Salgotra R. K., & Stewart C. N., Jr., 2020. Functional Markers for Precision Plant Breeding. International journal of molecular sciences, 21(13), 4792. Sousa T. V., Caixeta E. T., Alkimim E. R., de Oliveira A. C. B., Pereira A. A., Zambolim L., Sakiyama N. S. 2017. Molecular markers useful to discriminate *Coffea arabica* cultivars with high genetic similarity. *Euphytica*, 213(75): 1–15.

Teressa A., Crouzillat D., Petiard V., Brouhan P., 2010. Genetic diversity of Arabica coffee (*Coffea arabica L.*) Collections. *EJAST,* 1(1): 63-79.

Tesfaye K., Govers K., Bekele E., Borsch T., 2014. ISSR Fingerprinting of *Coffea arabica* Throughout Ethiopia Reveals High Variability in Wild Populations and Distinguishes Them from Landraces*. Plant Systematics and Evolution*, 300(5): 881–897.

Tounekti T., Mahdhi M., Al-Turki T. A., Khemira H., 2017. Genetic diversity analysis of coffee (*Coffee arabica L.*) Germplasm accessions growing in the Southwestern Saudi Arabia using quantitative traits. *Natural Resources*, 8(5): 321-336.

Thieman W. J., & Palladino M. A., 2014. *Introduction to Biotechnology. In British Library Cataloguing-in-Publication Data*. Pearson New International Edition.

Vieira E. S. N., von Pinho É. V. d. R., Carvalho M. G. G., Esselink D. G., & Vosman B., 2010. Development of microsatellite markers for identifying Brazilian *Coffea arabica* varieties. *Genetics and Molecular Biology*, 33(3): 507–514.

Warburton, M. and J. Crossa. 2000. Data Analysis in the CIMMYT Applied Biotechnology Center for Finger printing and Genetic Diversity Studies. 1 st Ed. Pp.21.

Zhao Y., Prakash C. S., & He, G., 2012. Characterization and compilation of polymorphic simple sequence repeat (SSR) markers of peanut from public database. *BMC Research Notes*, 5(362): 1–7.