International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Microsatellite (SSR) Analysis on Genetic Diversity of *(Coffea canephora*) Germplasm in Kagera Region, Tanzania

N. M. Ng'homa[a*], P. M. Kusolwa[b], D. P. Mamiro[c], D. L. Kilambo[d], D. P. Massawe[e]

[a]*Tanzania Coffee Research Institute (TaCRI), P.O.Box 127, Bukoba, Tanzania*

[b]*Department of Crop Science and Horticulture, P.O.Box 3005,Sokoine University of Agriculture, Morogoro Tanzania*

[a]*Email: Nyabisi.nghoma@tacri.org*

[b]*Email: kusolowap@gmail.com*

Abstract

Coffea canephora (Robusta) is one of the two important coffee species grown in Tanzania for commercial purpose. Robusta coffee contributes 40 - 50% of the total coffee production in Tanzania. However, the production of Robusta coffee in Tanzania has been hampered by coffee wilt disease. Despites the importance of Robusta coffee in Tanzania, its genome has not thoroughly researched. A study was conducted to investigate the molecular diversity of cultivated and wild coffee found in Kagera region in Tanzania. One hundred twenty four genotypes (124) of cultivated and wild coffee were analyzed by simple sequence repeat (SSR) marker techniques using 12 microsatellite markers. Genetic diversity, similarity or dissimilarity, genetic distances between individuals and genetic differentiation between populations was analyzed. Findings indicate high genetic variations among cultivated and wild coffee genotypes ranging between 20 and 83%. Polymorphism was 80% among SSR markers with 8 loci. Two distinctive genetic groups were identified. The first genetic group comprised four distinctive genetic groups one to four.

---

* Corresponding author.

The second genetic group consisted of four genetic groups, five to eight. Wild coffee genotypes had similarities to some of cultivated *C. canephora* in groups one to four implying that some cultivated *C. canephora* originates from wild coffee. Groups' five to eight comprise genotypes from cultivated *C. canephora*. Detailed study is needed to compare the identified eight (8) genetic groups of *C. canephora* in Kagera region in Tanzania with the already know groups worldwide.

Keywords: Genetic diversity of *Coffea canephora;* (SSR) Microsatellite.

1. Introduction

Coffea canephora is one of two important coffee species grown in the world. It accounts 35-40% of world production (International Coffee Organization. www.ico.org). It is self-incompatible and diploid; while *C. arabica* which accounts 60- 65% is tetraploid and self compatible [13]. *Coffea canephora* is found in low and middle altitudes areas in Africa, far East and parts of Southern America [22]. In Africa, *C. canephora* is found in Ivory Coast, Democratic Republic of Congo, Cameroon, Uganda, Angola, Ghana, Togo, Madagascar, Republic of Central Africa and Tanzania. In Far East this coffee *spp* is found in India, Indonesia, Vietnam, Philippines and Brazil in Southern America. In most countries, C. *canephora* is conserved in ex-situ collection plots. *Coffea canephora* is self-incompatible coffee *spp* in natural, cross-pollinated perennial plant, while wind and insects being are the main pollinators. Based on the nature of *C. canephora* and its wide geographical distribution genetic characterization are very important for both effective crop improvement and conservation purposes [26, 22]. Comprehensive conversation of genetic resources for *C. canephora* is very important due to threats of extinction of the species resulting from invasion of the most devastating coffee wilt disease and rapid extinction of natural forests caused by human activities and climates changes [20]. Assessment of genetic variability between and among *C. arabica* and *C. canephora* populations using different molecular markers' techniques has be conducted by several scientists [7, 19, 15, 10,8,23, 3, 18, 17, 20, 22 and 31]. Bertrand [13] reported the genetic diversity of *C. canephora* for the first time in 1986 when studied the wild and cultivated coffee genotypes from Western and Central Africa Republic. During this study two diversity groups were identified. These groups included Congolese group, which comprised genotypes from Central African Republic and Cameroon, and a Guinean group, which comprised genotypes from Ivory Coast [20]. Montagnon *and his colleagues* [7] who further investigated the diversity of *C. canephora* and wild coffee genotypes reported three groups namely Guinea, Congolese SG1 and Congolese SG2 with groups SG1 and SG2 being the sub-groups of the Congolese group. Dussert *and his colleagues* [23] using Restriction Fragment Length Polymorphism molecular markers (RFLP) reported five diversity groups of cultivated *C. canephora* and wild coffee genotypes, adding 2 Congolese groups B and C. Cubry *and his colleagues* [18] and Cubry *and his colleagues* [17] using microsatellites on *C. canephora* confirmed the previous findings of 5 diversity groups: SG1, SG2, B, C and Guinean. Musoli *and his colleagues* [20] while studying the diversity of wild coffee and cultivated *C. canephora* from Ugandan coffee discovered 7 distinctive genetic diversity groups of which 5 were related to the previous know groups and one unique group from Uganda coffee. Thomas [4], Nyange and Marandu [9] reported that some of wild and cultivated *C. canephora* were indigenous to Uganda and the Northern part of Tanzania in Kagera. In 16th century large seed *C. canephora* was introduced from Congo, and commercial cultivation started in Uganda and the Northern part of Tanzania region in Kagera region in 1900s [2, 1].

However, due to self-incompatibility of the species and the proximity of cultivated C. *canephora* to the forests might have led to the add mixtures between introduced *C. canephora* and wild coffee [20]. The previous studies conducted to investigate the diversity of cultivated *C. canephora* and wild coffee genotypes did not include coffee genotypes from the Northern part of Tanzania. On these contexts, therefore a study was conducted to determine the diversity and relatedness and unrelatedness of cultivated *C canephora* and wild coffee genotypes found in Kagera region in Tanzania. This paper presents the results of a study of genetic diversity of Tanzanian cultivated *C. canephora* and wild coffee genotypes.

2. Materials and Methods

2.1 Samples collections

A total of one hundred and four (104) cultivated robusta coffee cultivars and twenty (20) genotypes of wild Robusta coffee were used in this study. One hundred and four cultivars of cultivated Robusta coffee collected from farmers' fields in Kagera region were selected randomly from 656 accessions in germplasm established at ARI-Maruku in Bukoba District. A total of twenty (20) trees of wild Robusta coffee sampled from Bushenyi (4 samples) and Minziro (16 samples) forests in Missenyi district were raised at Maruku (Table 1). One hundred and four (104) of cultivated *C. canephora* and twenty (20) wild coffee young leaves were harvested using Eppendorf tubes and immediately wrapped in the plastic bags. The samples were kept in the cool ice box and eventually stored at -20^{o}C before DNA extraction.

2.2 Coffee DNA extraction

Extraction of DNA was done at Sokoine University of Agriculture (SUA) using a modified CTAB protocol for isolating DNA from plant tissues as described by Mahuku in [24] and Dellaporta *and his colleagues* in [25]. Samples were ground using pestles in the eppendorf tubes. Fifty milligrams (50 mg) of each leaf sample were put in the 1.5 ml eppendorf tube containing DNA extraction buffer prepared from 100 mM Tris-HCL (pH 8.0), 2.0 M NaCL, 20 mM EDTA (pH 8.0) 2% (w/v) CTAB, (1% w/v) PVP (PVP K10, MW10.000) and 0.5% (w/v) activated charcoal. Charcoals and PVP were used to bind the phenolic compounds and left pure DNA. Samples were centrifuged at 16000g for 10 minutes at room temperature and the supernatants were transferred to the new microfuge tubes. One millilitre (1 ml) of chloroform / isoamyl alcohol (24:1) was added to each supernatant sample and then centrifuged at 16000 g for 10 minutes at room temperature. The aqueous (upper) phases were transferred to the fresh eppendorf tubes and 0.45 ml of isopropanol was added to each sample and mixed by inversion and thereafter samples were incubated at 25^{o}C for 1 hour and then centrifuged at 700 g for 10 minutes at room temperature to get DNA pellets. The supernatants were discarded to leave the DNA pellets. The pellets were washed by adding 1 ml of wash buffer made up of 15 mM ammonium acetate in 75% (v/v) ethanol to each tube and then centrifuged at 900 g for 10 minutes at room temperature. The supernatants were discarded and the pellets were air dried at room temperature for twenty minutes. Each pellet was dissolved in 50 µL of TE buffer prepared from (10 mM Tris-HCL (pH 8.0) and 1 mM EDTA (pH 8.0) and then centrifuged at 16000 g for 10 minutes at room temperature to remove some impurities and remain with pure supernatants. The supernatants which contained DNA were transferred to the new tubes and in forms of solutions DNA were stores at -20^{o}C in

the deep freezer. The DNAs' quantities and qualities were assessed by using 1% agarose gel electrophoresis using the standard protocol established at molecular laboratory at Sokoine University of Agriculture (SUA).

Table 1: Genotypes of 104 cultivated Robusta and 20 wild coffee characterized by using microsatellite (SSR) markers

Genotype code	Origin	Genotype code	Origin	Genotype code	Origin	Genotype code	Origin
001 MI 1	Misenyi	087 ML12	Muleba	192 ML1	Muleba	337 MI21	Misenyi
002 MI 2	Misenyi	091 KR23	Karagwe	193 ML2	Muleba	342 MI20	Misenyi
003 MI 3	Misenyi	092 KR24	Karagwe	194 ML3	Muleba	344 MI19	Misenyi
004 MI 4	Misenyi	108 BK4	Bukoba	240 BK14	Bukoba	346 MI11	Misenyi
005 MI 5	Misenyi	109 BK5	Bukoba	255 BK16	Bukoba	347 MR10	Bukoba
006 MI 6	Misenyi	112 BK6	Bukoba	257 BK18	Bukoba	348 (13/61)	Bukoba
007 MI 7	Misenyi	113 BK 7	Bukoba	259 BK19	Bukoba	349 ML2	Muleba
008 MI 8	Misenyi	114 BK4	Bukoba	263 BK20	Bukoba	MSI	Bukoba
009 MI 9	Misenyi	115 BK8	Bukoba	268 BK21	Bukoba	FB1	Busenyi forest
010 MI 10	Misenyi	117 BK	Bukoba	269 BK22	Bukoba	FB2	Busenyi forest
011 MI 11	Misenyi	118 (1/62)	Bukoba	280 KR1	Karagwe	FB3	Bushenyi forest
012 MI 12	Misenyi	120 M L13	Muleba	283 KR2	Karagwe	FB4	Bushenyi forest
020 MS5	Bukoba	123 BK 10	Bukoba	284 KR3	Karagwe	FM5	Minziro forest
023 KR20	Karagwe	125 BK11	Bukoba	287 KR4	Karagwe	FM6	Minziro forest
025 KR19	Karagwe	127 ML12	Muleba	288 KR5	Karagwe	FM7	Minziro forest
026 BK26	Bukoba	131 MS1	Bukoba	292 KR6	Karagwe	FM8	Minziro forest
030 KR18	Karagwe	139 MI11	Misenyi	293 KR7	Karagwe	FM9	Minziro forest
036 KR 12	Karagwe	142 BK13	Karagwe	294 KR8	Karagwe	FM10	Minziro forest
037 ML17	Muleba	147 KR25	Karagwe	295 BK23	Bukoba	FM11	Minziro forest
046 KR22	Karagwe	158 MI12	Misenyi	306 ML20	Muleba	FM12	Minziro forest
047 MS2	Bukoba	160 MI13	Misenyi	308 MI21	Misenyi	FM13	Minziro forest
049 KR12	Karagwe	162 MI14	Misenyi	310 MI25	Misenyi	FM14	Minziro forest
054 KR16	Karagwe	164 MI15	Misenyi	311 KR9	Karagwe	FM15	Minziro forest
055 KR15	Karagwe	165 MI16	Misenyi	312 KR10	Karagwe	FM16	Minziro forest
057 BK2	Bukoba	167 MI17	Misenyi	315 KR11	Karagwe	FM17	Minziro forest
059 BK3	Bukoba	170 ML10	Muleba	316 ML22	Muleba	FM18	Minziro forest
060 KR13	Karagwe	172 ML9	Muleba	320 KR12	Karagwe	FM19	Minziro forest
062 KR14	Karagwe	175 ML8	Muleba	323 ML24	Muleba	FM20	Minziro forest
077 ML18	Muleba	177 ML7	Muleba	324 ML15	Muleba		
079 ML17	Muleba	179 ML6	Muleba	330 MI24	Misenyi		
080 ML16	Muleba	181 ML5	Muleba	332 MI23	Misenyi		
086 ML15	Muleba	185 ML4	Muleba	333 MI22	Misenyi		

2.3 SSR genotyping

Fourteen polymorphic markers (SSRs) mapped on the *C. canephora* and *C. arabica* genomes were used to genotype the 124 genotypes collected from Bukoba, Karagwe, Missenyi and Muleba districts, and Bushenyi and Minziro forests in Kagera region in Tanzania (Table 9). Markers 394, 445, 501,364, 368, 384,355,456 DL 020 and 456 were designed from *C. canephora* clone 126 [14, 28]. Markers 774 and 782 were designed from *C. arabica* (Catura) [21]. Markers (DL013 and DL025) were designed for the study of sugar metabolism in coffee [28]. These markers were chosen because of their applications in *C. canephora* and related coffee species [29, 17, 20]. Markers were chosen based on linkage groups (n=11). In this study 10 out of 11 linkage groups which are located at a distance of at least 50 cM from each other (20) were used for the assessment of the diversity. The set of markers used were enough to allow differentiations of genetic diversity within the species [20, 17].

2.4 PCR amplification and visualization

Polymerase chain reaction (PCR) were performed in 20 µL containing 13.4 µl PCR water, 2 µl PCR buffer, 1.2 µl of $MgCl_2$ (25 mM), 0.4 µl of dNTP (10 mM), 1 µL of primer forward (10 µM), 1 µL of primer reverse (10µM), Taq DNA polymerase 5 µ/µL and 1 µL of DNA. Polymerase chain reaction (PCR) amplifications were run in an Eppendorf PCR tubes in the PCR machine.

The amplification protocol consisted of an initial denaturation cycles of 5 minutes at 94^oC followed by 30 cycles (45s at 94^oC, 30s at TM^C in accordance to the design of each primer (Table 4.1), 2 minutes at 72^o C and the final elongation step at 72^o C for 5 minutes. The 8 µL of each PCR products together with molecular marker ladder were loaded into the wells on 2% solidified agarose gels in 1 x TBE (Tris Borate EDTA) buffer. Electrophoresis separation of DNA fragments was conducted at 120 V for 2.30 hours. Resulting DNA fragments were stained with 0.5 µg / ml ethidium bromide in the bath of water for 30 to 60 minutes.

The gel images were visualized under UV light chamber and were retrieved by using digital camera CANON and eventually modified using Picasa 3 software. The DNA fragments from different loci sizes for each primer were scored as 0 and 1 represented absence and presence of bands respectively. The pair of alleles from each locus was scored as AA, Aa and aa to represent dominant, heterozygote and recessive alleles respectively.

Data analysis

Data were entered in a computer using excel software package as binary matrices. Data were analyzed by using hierarchical structure genomic statistical package. Analysis of genotypic frequency, alleles frequency, polymorphic loci, Shannon index, HW test, allele numbers, observed heterozygosity. F-Statistic, fixed index effective allele numbers, expected heterozygote and genetic distance were performed using the POPGENE software version 1.31 using the following models:

i. The index of genetic similarity (GS): $GS= 2N_i/ (N_i+N_j)$, where N_{ij} is the number of SSR alleles common to genotypes i and j while N_i and N_j are the total numbers of SSR alleles observed for genotype i and j respectively.

ii. The mean number of alleles 'N': $N = \sum^{n}_{i=1}(N_i/n$, where N_i is the number of alleles at *i*th allele, where N is the number of alleles at *i*th allele.

iii. The effective allelic number 'N_e': $=N_e = \sum^{n}_{i=1}N_{ei}/n = \sum^{n}_{i=1} (1/\sum q^2_j)/n$, where N_e is the effective allelic number at *i*th allele, and *qj* the frequency of the *j*th allele.

iv. H_o is the observed heterozygosity: $H_o = \sum H_{oi}/n = \sum (1-\sum^{m}_{j=1}q^2_{ij})/n$, where H_{oi} represents the observed heterozygosity of the *i*th allele and q_{ij} is the frequency of *j*th homozygous allele at *i*th allele.

v. The expected heterozygosity, (index of gene diversity): $H_e = \sum H_i/n = \sum (1-\sum q^2_{ij})/n$, where H_i is the expected heterozygosity of the *i*th allele and q_{ij} is the frequency of the *j*th homozygous allele at *i*th allele.

vi. Wright fix index, defined as inbreeding coefficient 'F': $F = 1-H_o/H_e$ ranges from -1 to 1. The value of F is 1 when the population is heterozygous. 'F_{it}' 'F_{is}' and 'F_{st}' are Wright F- statistics parameters. F_{it} and F_{is} are defined as genetic deviation from Hardy- Weinberg expectation within and among genotypes respectively. Genotypes arrive at Hardy- equilibrium when F_{it} and F_{is} are 0. F_{st}, ranging from 0 to 1, is an estimate of gene differentiation between genotypes, which represents genetic variation among genotypes [32]. F_{st} is 0 if there was no genetic variation among genotypes.

vii. Genetic distances 'D' between each pair of landraces was estimated by the modified Rogers distance as follows: $D = 1/n \sum^{n}_{i=1}\sum^{m}_{j=1}1/2(p^X_{ij}-p^Y_{ij})^2$, where p^X_{ij} and q^Y_{ij} are frequencies of ith allele at $_j$th allele/locus in genotypes X and Y respectively. Molecular data were further analyzed by using the hierarchical cluster analysis method. Dendrogramme tree was drawn based on Nei's genetic distances using un-weighted pair-group method with arithmetic mean (UPGMA) generated by the nearest neighbour Jaccard similarities coefficients.

3. Results

3.1 Amplification of PCR products of 104 cultivated C. canephora and 20 wild coffee

Figure 1 summarizes amplification products of genotyped 104 cultivated Robusta coffee and 20 wild coffee genotypes using markers designed from Robusta clone 126. Among twelve (12) markers tested, 10 showed amplification and produced clear bands that were scored (Figure 1).

3.2 Genetic diversity of cultivated Robusta and wild coffee

Figure 1 and Table 2 summarize results on the dissimilarity/or similarity of assessed cultivated Robusta coffee and wild coffee in Kagera region. A total number of alleles observed for tested primers were 248. Eight out of 10 amplified markers generated 8 polymorphic loci, and 80% polymorphisms were detected from 124 genotypes investigated showing the greatest allelic diversity of the populations. For each SSR locus, the number of alleles ranged from 1 to 2 corresponding to an average of 1.8. The effective allelic number ranged from 1.0 to 1.9995 with average of 1.4779. The average of observed heterozygosity (H_o) was 0.12 ranging from 0.00 to 0.35. The average of expected heterozygosity was 0.28, ranging from 0.00 to 0.50. The overall mean for heterozygosity ranged between 0.00 and 0.18 with average of 0.05.

Table 2: SSR primers used for PCR analysis of 124 coffee genotypes

SN	EMBL.acc.no.	Marker Name	Primer Sequences (5´- 3´)	Allele size (bp)	TM[C] Designer	TM[C] PCR products	Primer origin	Species origin
1	AM231563	394	Forward. GCCGTCTCGTATCCCTCA	124	52.9	54.0	Poncet and his colleagues (29)	*Coffea canephora*
			Reverse. GAAGCCAGAAAGTCAGTCACATAG			53.8		
2.	AM231567	445	Forward. CCACAGCTTGAATGACCAGA	275	52.1	53.3	Poncet *and his colleagues* (29)	*Coffea canephora*
			Reverse. AATTGACCAAGTAATCACCGACT			53.3		
3.	AM231576	501	Forward. CACCACCATCTAATGCACCT	343	51.9	52.4	Poncet and his colleagues (29)	*Coffea canephora*
			Reverse. CTGCACCAGCTAATTCAAGC			52.4		
4.	AJ871899	DL020	Forward.TGCTCAAACTTCTTGCT	250	42.5	42.5	Leroy and his colleagues (28)	*Coffea canephora*
			Reverse. CGCCAACTCTAATGTGT			42.5		
5.	AJ871892	DLO13	Forward. AGAGGGATGTCAGCATAA	267	44.1	44.2	Leroy *and his colleagues* (28)	*Coffea canephora*
			Reverse. ATTTGTGTTTGGTAGATGTG			44.3		
6.	AJ871904	DL025	Forward.TTGTTGAGAGTGGAGGA	197	42.0	44.0	Leroy *and his colleagues* (28)	*Coffea canephora*
			Reverse. CCAAAGACAGTGCAGTAA			43.0		
7.	AM231556	364	Forward. AGAAGAATGAAGACCAAACACA	90	50.5	50.4	Poncet *and his colleagues* (29)	*Coffea canephora*
			Reverse. TAACGCCTGCCATCG			48.3		
8.	AM231558	368	Forward. CACATCTCCATCCATAACCATTT	160	54.2	54.5	Poncet *and his colleagues* (29)	*Coffea canephora*
			Reverse. TCCTACCTACTTGCCTGTGCT			53.0		
9.	AM231560	384	Forward. ACGCTATGACAAGGCAATGA	255	52.9	54.5	Poncet *and his colleagues* (29)	*Coffea canephora*
			Reverse. TGCAGTAGTTTCACCCTTTATCC			54.0		
10.	AM231552	355	Forward. CTATGATGTCTTCCAACCTTCTAAC	177	52.2	52.5	Poncet *and his colleagues* (29)	*Coffea canephora*
			Reverse. GGTCCAATTCTGTTTCAATTTC			51.8		
11.	AJ308774	774	Forward. GCCACAAGTTTCGTGCTTTT	228	54.2	55.0	Poncet *and his colleagues* (29)	*Coffea arabica catura*
12.			Reverse. GGGTGTCGGTGTAGGTGTATG			53.8		
13.	AJ308782	782	Forward. AAAGGAAAATTGTTGGCTCTGA	114	54.4	53.0	Poncet *and his colleagues* (30)	*Coffea arabica catura*
			Reverse. TCCACATACATTTCCCAGCA			53.4		
14.	AM231568	456	Forward.TGGTTGTTTTCTTCCATCAATC	297	53.0	53.0	Poncet *and his colleagues* (29)	*Coffea canephora*
			Reverse. TCCAGTTTCCCACCCTCT			52.5		

Note: Source Cubry *And his colleagues* [17] and Musoli *And his colleagues* [20]

The results showed two distinctive main genetic groups among 104 cultivated robusta and 20 wild coffee genotypes (Fig. 1). The first genetic group consisted groups I, II, III and IV. The second genetic group composed groups V, VI, VII and VIII. The distances of each group from the two main groups varied significantly ($P = 0 \leq 0.05$) (Fig. 1).

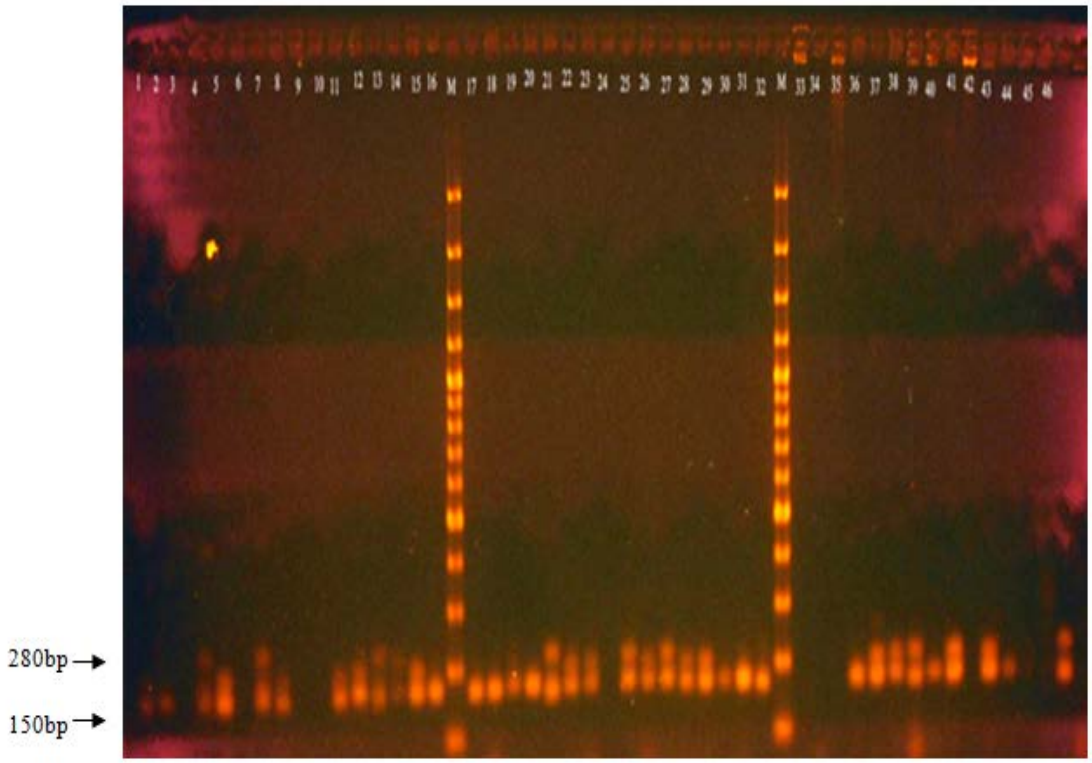

Figure 2: Gel electrophoresis of DNA of *C. canephora* genotypes, M is a 100bp Ladder, amplification product from SSR marker 355 with alleles ranging between 150bp - 280bp) using accessions 1- 46 (Table 1).

Genetic structure of cultivated and wild coffee populations of Tanzanian germplasm

The results for the genetic structure of cultivated and wild coffee population varied from F_{st}= 0 for loci DL 025 and 774 to F_{st} = 0.97 for locus DL013. The overall mean of F_{st} for 104 cultivated *C. canephora* and 20 wild coffees was 0.81 indicating that 81% of genetic variations were observed among investigated genotypes and only 19% of genetic variations were noted within cultivated *C. canephora* and wild coffees populations.

The structured analysis revealed ten groups of cultivated *C. canephora* and wild coffee coffees in Kagera region of Tanzania with four distinctive populations of cultivated C. *canephora* (Fig. 1).

The wild coffee from Bushenyi forest was clearly identified as different group with admixture of very few individuals from cultivated *C. canephora* (Table 4.3). However, the results showed that some wild coffees from Bushenyi and Minziro forests were genetically closely related with some cultivated *C. canephora*.

Genetic distances between cultivated *C. canephora* genotypes from Tanzanian germplasm

Tables 4 and 5 present genetic distances of 14 selected cultivated *C. canephora* and coffee populations from Missenyi, Muleba, Bukoba, Karagwe and wild coffees from Bushenyi and Minziro forests. The genetic distances among 14 selected *C. canephora* ranged from the lowest value (0.00) between MI-4, MI-6 and MI-5 to the highest (0.64) between MI-13 and MI-2 (Table 4). Furthermore, the genetic distances between the genotypes of cultivated *C. canephora* and wild coffee varied greatly among the six origins.

The lowest genetic distance was 0.0034 between the genotypes for cultivated C. *canephora* from Karagwe and Bukoba districts. The overall results of genetic distances of cultivated *C. canephora* from Missenyi, Bukoba, Karagwe and Muleba were genetically closely related (Table 5).

The highest genetic distance was observed between wild coffee from Bushenyi forest and cultivated *C. canephora* from Muleba (0.13), Karagwe (0.11) and Missenyi (0.10) districts. The results also showed that wild coffee from Minziro forest were closely related to cultivated *C. canephora* from Bukoba and Karagwe districts.

Table 3: Genetic variations of 104 cultivated *C. canephora* and 20 wild coffees based on 10 SSR loci in Tanzania

Locus	Simple Size	N_a	N_e	I	H_o	H_e	Ne**	H_{av}	F_{st}	$Null_m$
456	248	2	1.3402	0.4213	0.1532	0.2549	0.2539	0.0766	0.6982	0.1081
364	248	2	1.4018	0.4612	0.0242	0.2878	0.2866	0.0121	0.9578	0.0110
368	248	2	1.7763	0.6288	0.2581	0.4388	0.4370	0.1290	0.7048	0.1047
394	248	2	1.0411	0.0987	0.2442	0.0397	0.0395	0.0121	0.6938	0.1103
501	248	2	1.6103	0.5667	0.1210	0.3805	0.3790	0.0605	0.8404	0.0475
355	248	2	1.9995	0.6930	0.3548	0.5019	0.4999	0.0121	0.6451	0.1376
384	248	2	1.6205	0.5710	0.0645	0.3845	0.3829	0.0605	0.9158	0.0230
DL013	248	2	1.9895	0.6905	0.0242	0.4994	0.4974	0.1774	0.9757	0.0062
DL025	248	2	1.000	0.000	0.000	0.000	0.000	0.0324	0.000	*****
774	248	2	1.000	0.000	0.000	0.000	0.000	0.0121	0.000	*****
Mean	248	1.8	1.4779	0.4131	0.1024	0.2787	0.2776	0.0512	0.8155	0.0565
St.Dev		0.42	0.3849	0.2775	0.1209	0.1997	0.1997	0.0605		

N_a = Observed number of alleles per locus, N_e = Effective number of alleles {(Kimura and Crow (1964)}, I = Shannon's Information index {Lewontin (1972)}, H_o = Observed heterozygosity computed using Levene (1949), H_e = Expected heterozygosity computed using Levene (1949), Nei's = (1973) Expected heterozygosity, H_{av} = Average heterozygosity and $Null_m$ = Estimated from $F_{st} = 0.25(1-F_{st})/F_{st}$. At $P = \leq 0.05$)

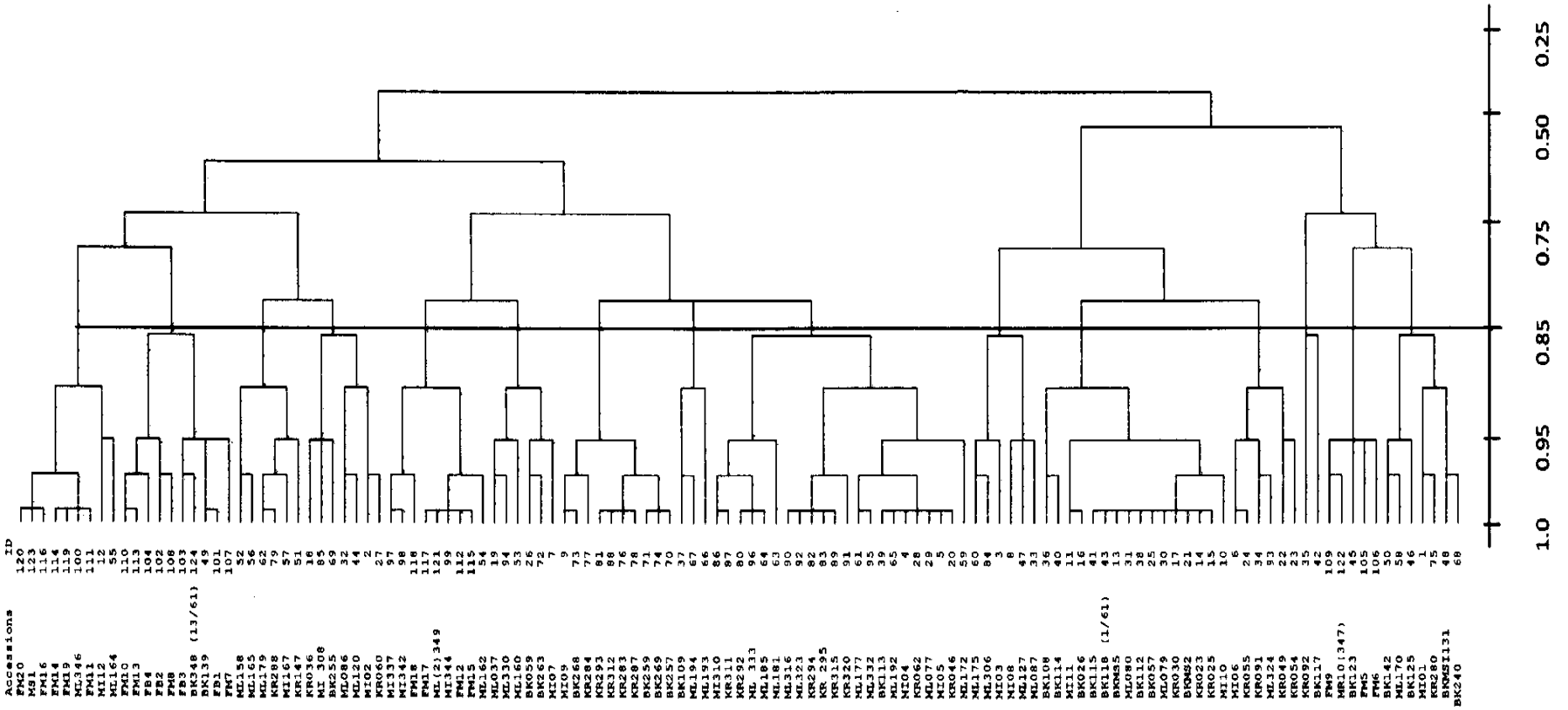

Figure 1: The Cl**u**ster Dendogramme representing the dissimilarity among cultivated *C. canephora* and wild coffee genotypes from Kagera region, Tanzania obtained by UPGMA method based on the weighted Jaccard index estimated from the polymorphism of 12 microsatellite

Table 4: Genetic distances between selected 14 cultivated *C. canephora* from Missenyi district

	MI-1	MI-2	MI-3	MI-4	MI-5	MI-6	MI-7	MI-8	MI-9	MI-10	MI-11	MI-12	MI-13	MI-14
MI-1	******													
MI-2	0.3567	******												
MI-3	0.0256	0.2620	******											
MI-4	0.3495	0.2064	0.3239	******										
MI-5	0.3495	0.2064	0.3239	0.0000	******									
MI-6	0.3495	0.2064	0.3239	0.0000	0.0000	******								
MI-7	0.5816	0.2451	0.5559	0.0949	0.0949	0.0949	*******							
MI-8	0.3567	0.5108	0.4051	0.0813	0.0813	0.0813	0.2451	******						
MI-9	0.2520	0.5722	0.3054	0.1808	0.1808	0.1808	0.2936	0.1369	******					
MI-10	0.3567	0.2231	0.2620	0.3495	0.3495	0.3495	0.3993	0.5108	0.2620	******				
MI-11	0.5166	0.2064	0.4039	0.1942	0.1942	0.1942	0.1638	0.3495	0.2498	0.0813	*******			
MI-12	0.0256	0.2620	0.0541	0.3239	0.3239	0.3239	0.4606	0.4051	0.2364	0.2620	0.4039	******		
MI-13	0.3040	0.6405	0.2783	0.5592	0.5592	0.5592	0.5289	0.4581	0.3525	0.3040	0.2968	0.3247	*****	
MI-14	0.1054	0.5106	0.1369	0.3495	0.3495	0.3495	0.5816	0.2231	0.1369	0.2231	0.3495	0.1280	0.1516	******

Genetic distance among 14 selected Robusta accessions (P = 0.05)

Table 5: Genetic distance of cultivated *C. canephora* and wild coffees based on geographical locations

Population(ID)	MI	ML	BK	KR	FB	FM
MI	*****					
ML	0.0345	*****				
BK	0.0185	0.0055	*****			
KR	0.0215	0.0063	0.0034	*****		
FB	0.1006	0.1294	0.0948	0.1141	*****	
FM	0.0785	0.1076	0.0783	0.0902	0.0076	*****

Genetic distances among population of cultivated robusta and wild coffee species (P = 0.05).

Abbreviations represent geographical locations where the materials were collected: MI-

Missenyi, ML= Muleba, BK = Bukoba, KR = Karagwe, FB = Bushenyi forest and FM = Minziro forest.

4. Discussion

The analysis of genetic diversity of cultivated *C. canephora* and wild coffee revealed unrelated and relatedness of coffee species in Kagera region of Tanzania. The high polymorphism (80%), variations on F- statistics, expected heterozygosity; coefficient of similarity and genetic distances found in this study reflects a highest genetic diversity among 124 genotypes of both wild and cultivated C. *canephora*. The highest observed genetic diversity on this study provides evidence that SSR markers are adequate for assessing intra-specific and inter specific variations, and informative for detecting genetic diversity and relationships among cultivated *C. canephora* and their related wild coffee genomes. These findings are in agreement with those reported in the previous studies on the global diversity of *C. canephora* and wild coffee using SSR markers [18, 17, and 20]. In this study, two main groups of cultivated *C. canephora* with eight [8] subgroups had been identified of which four sub groups composed of mixtures of cultivated and wild coffee genotypes and the other four sub groups composed of only cultivated *C. canephora*. The first and second groups comprised 75 and 25 % of the total genotypes investigated, respectively indicating that high proportions of robusta coffee cultivars growing in Kagera region are mixtures of cultivated and wild coffees. Furthermore, two distinctive groups of wild coffee genotypes were observed from Bushenyi and Minziro forests. The SSR techniques showed that coffee genotypes found in Minziro and Bushenyi forests were closely related to cultivate *C. canephora* indicating that all investigated coffee genotypes could have a common genome. Moreover, observed genetic dissimilarity between cultivated C. *canephora* and some wild coffee in Minziro and Bushenyi imply that some coffee genotypes lack common genome. However, cultivated *C. canephora* and wild coffee species from Minziro are not easily distinguishable morphologically and they were either erect or bending types.

The wild genotypes from Bushenyi forest despite of being genetically related to cultivated *C. canephora*, but they have distinctive morphological characteristics. Individual coffee trees found in Bushenyi forest were short, with small branches, few branches, few berry clusters with small seeds, resistant to coffee leaf rust and coffee wilt disease. Similar relationships between wild coffee materials and Nganda - erect populations had been observed in Uganda [20] and Ivory Coast [16].

Reports by Musoli *and his colleagues* [20] and Montagnon *and his colleagues* [6] showed that cultivated *C. canephora* in most of African, *C. canephora* growing countries resulted from natural crossing of wild coffee materials and introduced genotypes from other regions or countries leading to a mixture of all genotypes. Genetic relationships between cultivated *C canephora* observed in Bukoba, Karagwe, Missenyi and Muleba districts with wild coffee found in Bushenyi and Minziro forests support that hypothesis and indicate that the early established coffee bushes in Kagera region in Tanzania originated in Kagera region and Uganda [9,5,11]. From this study therefore, it was believed that cultivated coffee in Kagera region in Tanzania originated from wild coffee genotypes found in the region and Uganda [5, 4]. The presence of both Nganda and Erect coffee types in Uganda and Tanzania implies that both countries share the genetic pool of *C. canephora*. The variations of identified genetic groups between coffee species found in Tanzania and those of Uganda probably attributed to the origin sources of tested materials. Coffee genotypes used in this study represented the actual natural origin where there was no any kind of improvement while those used in Uganda probably could have been undergone through improvements. The genotypes of wild coffee probably have been cultivated in admixtures with introduced genotypes from Congolese sub groups SG2, SG1 B, C and Guinean leading to existence of spreading (Nganda, Erect and Semi erect cultivated types). The high diversity of cultivated C. *canephora* genotypes could be due to multiples origins of coffee species, resulting from successive natural hybridization of wild species and introduced Congolese genotypes [20]. According to TCB [27] cultivated *C. canephora* was introduced in Kagera region during the 16^{th} century from Congo. Musoli *and his colleagues* [20] reported the presence of the genetic diversity of the Congolese C. *canephora* groups B, C, SG1, SG2 and Guinean in Ugandan *C. canephora* genotypes. Since Tanzania and Uganda share borders there is a high possibility of having Guineans and Congolese groups in Tanzanian cultivated *C. canephora*. The genomic molecular study on cultivated *C. canephora* and wild coffee genotypes revealed the diversity and the complex mixtures between wild and cultivated *C. canephora* in Kagera region in Tanzania. The value of expected dissimilarity coefficient within Tanzanian genotypes ranged from 0.08 to 0.83 with overall mean of 0.51. The mean value of gene diversity observed in this study was close to the previous values (0.55) observed within *C. canephora* [17]. Furthermore, the values of expected variations on this study were close to those reported by Musoli *and his colleagues* [20] who reported the diversity of wild and *C. canephora* in Uganda using 18 SSRs microsatellite markers ranged 0.48 to 0.59, and 0.47 to 0.68 for out crossing perennial plants [12].

The molecular analysis of wild coffee genotypes found in Bushenyi and Minziro forests revealed the genetic relatedness and distinctiveness to minority and majority of cultivated *C. canephora* found in Kagera region in Tanzania respectively. The existence of the similarity between cultivated *C. canephora* and wild genotypes mixture observed in this study is further supported by Cubry *and his colleagues* [16] who reported the highest diversity of the Guinean group which comprised a large number of natural populations from the forests of Ivory Coast and smallholder plantations. The overall findings with regards to wild coffee support the previous reports

that *C. canephora* is indigenous to Uganda and Tanzania [9, 20, 11].

5. Conclusion and Recommendations

In conclusion, this study revealed the richness of genetic diversity of cultivated *C. canephora* and wild coffee in Kagera region in Tanzania. The results obtained from this study will be useful in planning strong breeding programme for genetic conservation and improvement of *C. canephora* the second important coffee species in Tanzania. The observed genetic diversity in this study will be utilized in hybridization of coffee species to develop coffee varieties which are high yielding, resistant to diseases, with good qualities and which can grow well in various agro ecological environments. In addition, these findings will be the basis for exploring more information on existence of coffee species within the country, collect and conserve them for further uses in the breeding programme.

The identified eight (8) genetically diverse groups of cultivated *C. canephora* and two (2) genetic diversity groups of wild coffee should be thoroughly studied to compare their relationships with the genotypes found in the other countries which grow robusta coffee worldwide.

Acknowledgement

We express our grateful thanks to European Union (EU) and Coffee Stakeholders for the financial support of this study. We also thank the management teams of Tanzania Coffee Research Institute (TaCRI) and Sokoine University of Agriculture SUA for the support they give me during the implementation of this study.

We also extent our sincere appreciation to all technicians at SUA, Molecular Biology Laboratories, TaCRI and ARI- Maruku for their valued technical assistance in data collection and processing.

References

[1]. A.A.Thomas (1947). The cultivation and selection of Robusta coffee in Uganda. The Emp, J. Exp. Agric. 15: 66 – 81.

[2]. A. D. Kaiza., W. Heemskert and F. E (1999). Coffee Production in Kagera Region: Past, Present and Future In: 50th Anniversary of ARI-Maruku, Planning the Future: Past, Present and Future Perspectives of Land Use In the Kagera Region. (Edited by Former, E. C. R., Schouten, C. and Baijukya F.P.) Ministry of Agriculture and Co-operatives Lake Zone Agricultural Research and Training Institute, Kagera. pp. 91 – 97.

[3]. A. P. Chaparro., M.A. Cristancho., H.A.Cortina and Gaitan, A. L. (2004). Genetic variability of Coffea arabica L. accessions from Ethiopia evaluated with RAPDs. Genetic Resource Crop Evolution 51: 291 – 297.

[4]. A.S. Thomas. (1935). Types of Robusta coffee and their selection in Uganda. East African Agricultural Journal I: 193 – 198.

[5]. B. A. Eskes (1997). Report of Coffee Mission. CIRAD-CP, Montpelier. 24pp.

[6]. C. Montagnon, T. Leroy. and B.A. Eskes (1998). Varietal improvement of Coffea canephora II.

Breeding programmes and their results. Plantation Recherché Development 5: 18 – 33.

[7]. C. Montagnon., T. Leroy and B.A. Yapo, (1992). Diversite genotypique et phenotypique de quelques groupes de cafeiers (Coffea canephora Pierre) en collection Consequences sur leur utilisation en selection. Café Cacao The 36: 187 – 198.

[8]. E. Aga., T. Bryngelsson., E. Bekele. and B. Salomon. (2003). Genetic diversity of forest Arabica Coffee (Coffea arabica L.) in Ethiopia as revealed by random amplified polymorphic DNA analysis. Heriditas 138: 36 – 46.

[9]. E. N. Nyange and F. E. Marandu. (1996). Report on the Improvement of Coffea Canephora Germplasm, Exploration and collection of new robusta materials from farmers' plots. Maruku Agricultural Research Institute, Kagera, Tanzania. 22pp.

[10]. F. Antony., M.C. Combes., C. Astorga., B. Bertrand., G. Graziosi and P. Lashermes. (2002). The origin of cultivated Coffea arabica L. varieties revealed by AFLP and SSR markers. Theoretical Applied Genetics 104: 894 – 900.

[11]. G. Wrigley (1988). Coffee Longman Scientific and Technical. John Wiley and Sons, New York. 639pp.

[12]. H. Nybom .(2004). Comparison of different nuclear DNA markers for estimating intraspecifi genetic diversity in plants. Molecular Ecology (13): 1143 – 1155.

[13]. J. Berthaud (1986). Les ressources genetiques pour I amelioration des caferers africans diplides: evaluation de la richesse genetique des populations sylvestres et de ses mecanismes organisateurs, consequences pour I' application. ORSTOM, Paris.

[14]. M. Dufour., P. Hamon., M. Noirot., A.M. Risterucci and T. Lerory (2002). Potential use of SSR markers for Coffea spp. genetic mapping. In: Proceedings of the 19th International Scientific Colloquim on Coffee . (Edited by ASIC), Paris, France.

[15]. O. C. Agwanda., P. Lashermes., P. Trousiot., C. M. Combes. and A. Charrier (1997). Identification of RAPD markers for resistance to coffee berry disease, Colletrichum kahawae, in Arabica coffee. Euphytica 97: 241 – 248.

[16]. P. Cubry., F. de Bellis., K. Avia., S. Bouchet, D. Pot., M. Dufour., H. Legnate and T. Leroy. (2013). An initial assessment of linkage disequilibrium (LD) in coffee tree: LD patterns in groups of Coffea canephora Pierre using microsatellite analysis. BMC Genomics, 14 (10):1471-2164

[17]. P. Cubry., P. Musoli., H. Legnate., D. Pot., F. de Bellis., V. Poncet., F. Antony., M. Dufour and T. Leroy. (2008). Diversity in coffee assessed with SSR markers; structure of the genus Coffea and perspectives for breeding. Genome 51: 50 – 63.

[18]. P. Cubry., F. De Bellis., D. Pot., P. Musoli., H. Legnate., T. Leroy. and M. Dufour. (2005). Genetic diversity analyses and linkage disequilibrium evaluation in some natural and cultivated populations of Coffea canephora. In: Proceedings of the 4th Plant Genomics European Meeting, Amsterdam. Centre for Bio systems genomics, 20-23 September 2005, Wageningen, The Netherlands.

[19]. P. Lashermes., J. Cros., P. Marmey and A. Charrier. (1993). Use of random amplified DNA markers to analysis genetic variability and relationships of Coffea species. Genetic Resource Crop Evolution 40: 91 – 99.

[20]. P. Musoli,, P. Cubry., P. Aluka., C. Billot., M. Dufour.,F. De Bellis., D. Pot., D. Bieysse., A. Charrier.

and T. Leroy. (2009). Genetic differentiation of wild and cultivated populations: diversity of Coffea canephora Pierre in Uganda. Genome 52: 634 – 646.

[21]. P. Rovelli., R. Mettulio., F. Anthony., F. Anzueto and P. Lashermes (2000). Microsatellites in Coffea arabica L. In: (Edited by Sera,T., Soccol, C.R., Pandey, A. and Roussos, S) Coffee biotechnology and quality. Kluwer Academic Publishers, the Netherlands, pp 123-133.

[22]. P. Tshilenge., K.K Nkongolo., M. Mehes and A. Kalonji (2009). Genetic variation in Coffea canephora L. (Var. Robusta) accessions from the founder gene pool evaluated with ISSR and RAPD. African Journal of Biotechnology 8(3):380 – 390.

[23]. S. Dussert., P. Lashermes., F. Anthony., C. Montagnon., P. Trouslot and M.C. Combes (2003). Coffee (Coffea canephora). In Genetic diversity of cultivated tropical plants. (Edited by P. Hamon., M. Seguin., X. Perrier. and J. C. Glaszmann) Science Publishers, Inc., Enfield. pp. 239 – 258.

[24]. S. G. Mahuku. (2004). A simple extraction method suitable for PCR- based analysis of plant, fungal and bacterial DNA. Journal of Plant Molecular Biology Reporter 22(1): 71 – 81.

[25]. S. L. Dellaporta., J. Wood.and J.B.Hicks. (1983). A plant DNA mini preparation version II. Plant Molecular Biology Reporter 1: 19 – 21.

[26]. S. N. Prakash., C.M. Combes., S. Dursset, S. Naveen and P. Lashermes (2005). Analysis of genetic diversity in Indian robusta coffee gene pool (Coffea canephora) in comparison with representative core collection using SSRs and AFLPs. Genetic Resource Crop Evolution 52: 333 – 343.

[27]. TCB (2011). Miaka 50 ya Uhuru wa Tanzania na Mapinduzi ya Kilimo cha Kahawa. Tanzania Coffee Board, Moshi, Tanzania. 46pp

[28]. T. Leroy., Marraccini, M. Dufour, C. Montagnon., P. Lashermes and X. Sabau .(2005) Construction and characterization of Coffea canephora BAC libarary to study the organization of sucrose biosynthesis genes. Theoretical Applied Genetics 111(6): 1032-1041.

[29]. V. Pocent., M. Dufour., P. Hamon., S. Hamon., A. De Kochko and T. Leroy (2007). Development of genomic microsatellite markers in Coffea canephora and their transferability to other coffee species. Genome 50(12): 1156 – 1161.

[30]. V. Pocent., P. Hamon., J. Minie., C. Carasco., S. Hamon.and M. Noirot (2004). SSR cross-amplification and variation within coffee trees.(Coffea spp.). Genome 47: 1071- 1081.

[31]. W. C. Kathurima., M. G. Kenji., M. S. Muhoho., R. Boulanger., M. B. Gichimu. and K. E. Gichuru (2012). Genetic diversity among commercial coffee varieties, advanced selections and museum collection in Kenya using molecular markers. International Journal of Biodiversity and Conservation 4(2): 39 – 46.

[32]. Y. Qilun., Y. Kecheng., P. Guangtang and R. Tingzhao (2007). Genetic Diversity of Maize (Zea may L.) Landraces from Southwest China Based on SSR Data. Journal Genetics and Genomics 34(9): 851 – 860.