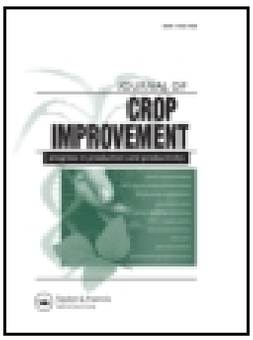

Journal of Crop Improvement

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/wcim20

Genetic diversity among commercial arabica coffee (*Coffea arabica* L.) varieties in Ethiopia using simple sequence repeat markers

Tadesse Benti , Endale Gebre , Kassahun Tesfaye , Gezahegn Berecha , Philippe Lashermes , Martina Kyallo & Nasser Kouadio Yao

To cite this article: Tadesse Benti , Endale Gebre , Kassahun Tesfaye , Gezahegn Berecha , Philippe Lashermes , Martina Kyallo & Nasser Kouadio Yao (2020): Genetic diversity among commercial arabica coffee (*Coffea arabica* L.) varieties in Ethiopia using simple sequence repeat markers, Journal of Crop Improvement, DOI: 10.1080/15427528.2020.1803169

To link to this article: https://doi.org/10.1080/15427528.2020.1803169

Published online: 13 Sep 2020.

JOURNAL OF CROP IMPROVEMENT
https://doi.org/10.1080/15427528.2020.1803169

Taylor & Francis
Taylor & Francis Group

Genetic diversity among commercial arabica coffee (*Coffea arabica* L.) varieties in Ethiopia using simple sequence repeat markers

Tadesse Benti[a], Endale Gebre[b], Kassahun Tesfaye[c], Gezahegn Berecha[d], Philippe Lashermes[e], Martina Kyallo[f], and Nasser Kouadio Yao[f]

[a]Department of Coffee Breeding and Genetics, Ethiopian Institute of Agricultural Research (EIAR, Jimma Research Center, Jima, Ethiopia; [b]Department of Agricultural Transformation and Rural Development, Policy Study Institute, Addis Ababa, Ethiopia; [c]Department of Agricultural Biotechnology, Ethiopian Biotechnology Institute, Addis Ababa, Ethiopia; [d]Department of Horticulture and Plant Science, College of Agriculture and Veterinary Medicine, Jimma University, Jimma, Ethiopia; [e]Department of Diversity, Adaptation and development of Cultivated Plants (DIADE), IRD, University of Montpellier, France; [f]Department of Biosciences Eastern and Central Africa- International Livestock Research Institute (Beca-ilri) Hub, Nairobi, Kenya

ABSTRACT

Ethiopia is the center of origin and genetic diversity of arabica coffee. Forty-two commercial arabica coffee varieties were developed by Jimma Agricultural Research Center (JARC) of Ethiopian Institute of Agricultural Research (EIAR) and released for production under diverse agro-ecologies of the country. Information on the level of genetic diversity among these varieties is scarce. Out of the 42 varieties, the genetic diversity of 40 widely cultivated commercial varieties was assessed using 14 simple sequence repeat (SSR) markers. These markers revealed polymorphism among the varieties. High average number of polymorphic alleles (7.5) and polymorphic information content (PIC = 80%) per locus were detected among the varieties. The genetic similarity among varieties using the Jaccard's similarity coefficient ranged from 0.14 to 0.78, with a mean of 0.38. The range of genetic similarity coefficient values in 92% of the possible pair-wise combinations varied from 0.14 to 0.50, indicating the presence of distant genetic relatedness among the varieties. Unweighted pair group method using arithmetic mean (UPGMA) clustering showed six major clusters and three singletons. Coffee varieties, belonging to the same geographic origin, were distributed across clusters. This study represents the first evidence of the presence of a high level of genetic diversity in Ethiopian commercial arabica coffee varieties. Divergent varieties with complementing traits could be crossed to develop productive hybrid coffee varieties.

ARTICLE HISTORY
Received 3 December 2019
Accepted 27 July 2020

KEYWORDS
Genebank; geographic origin; hybrid coffee; landraces; polymorphism; similarity coefficient

1. Introduction

Coffea arabica L. is a perennial crop, which belongs to the family *Rubiaceae*. Of the 124 species of the genus C*offea* (Davis 2011), C*offea arabica* and

CONTACT Tadesse Benti tadessebenti@gmail.com Ethiopian Institute of Agricultural Research (EIAR), Jimma Research Center, Jima, Ethiopia

Coffea canephora are economically important and most cultivated species in the world. *C. arabica* is the only tetraploid (2 n = 4x = 44) species in the genus and is self-fertile, whereas other species are diploid (2 n = 2x = 22) and genetically self-incompatible (Clarindo and Carvalho 2008). The southwest highlands of Ethiopia are the center of origin and center of genetic diversity for *Coffea arabica* (Lashermes et al. 1996; Anthony et al. 2001; Tesfaye et al. 2014) and it is the only *Coffea* species cultivated throughout the country. In Ethiopia, coffee is the most important commodity that earns foreign exchange and provides a source of income for more than 15 million people, whose livelihoods depend directly or indirectly on its production, processing and marketing.

Since the beginning of coffee research program in Ethiopia in the 1970s (Bellachew and Labouisse 2006; Fekadu et al. 2008), considerable progress has been made by the national coffee breeding program, resulting in successful development and release of 42 commercial varieties for production (Chala et al. 2012; Tadesse 2014). The presence of agro-morphological variability among the varieties has been reported by different investigators (Bellachew, Atero, and Tefera 2000; Chala et al. 2012; Tadesse 2014). Morphological variation does not reflect true genetic variation that exists at the DNA level because morphological markers are confounded by environmental, pleiotropic and epistatic effects. Besides, assessment of genetic diversity using morphological markers in perennial plants, such as coffee, often requires a lengthy and expensive evaluation during the entire vegetative growth period (Weising et al. 2005).

Detecting and quantifying genetic variation in crop species is important for successful conservation of genetic resources and improvement of desirable characteristics in a breeding program. The importance of DNA-based markers in accurately detecting genetic factors in coffee was reported by Lashermes et al. (1996). Several researchers have reported the presence of a low level of genetic diversity and a narrow genetic base of commercial varieties of arabica coffee using different molecular markers, such as amplified fragment length polymorphisms (AFLPs) (Steiger et al. 2002; Maluf et al. 2005), restriction fragment length polymorphisms (RFLPs) (Anthony et al. 2002), random amplified DNA polymorphisms (RAPDs) (Maluf et al. 2005; Silvestrini et al. 2008; Mishra et al. 2012), inter simple sequence repeats (ISSRs) and sequence-related amplified polymorphisms (SRAPs) (Mishra et al. 2012; Jingade et al. 2019) and SSR markers (Anthony et al. 2002; Maluf et al. 2005; Silvestrini et al. 2007; Teressa et al. 2010; Geleta et al. 2012; Al-Murish et al. 2013).

Comparative studies using RFLPs, RAPDs, AFLPs and SSRs have indicated that SSR markers detect a higher level of polymorphism than other markers (Bandelj et al. 2004). This has also been confirmed in previous studies on genetic diversity in wild coffee accessions and commercial varieties of arabica

coffee (Anthony et al. 2002; Maluf et al. 2005; Dessalegn, Herselman, and Labuschagne 2009). However, the level of genetic diversity among the 42 Ethiopian commercial arabica coffee varieties has not been assessed using molecular markers. Hence, we used SSR markers in the present study to generate information on the level of genetic diversity and relationship among 40 widely cultivated commercial arabica coffee varieties grown in Ethiopia and identify divergent parents for crossing to exploit hybrid vigor.

2. Materials and methods

2.1. *Genotypes and sample collection*

A total of 40 Ethiopian commercial arabica coffee varieties, widely cultivated at different altitude ranges of coffee-producing regions of the country, were used in this study (Table 1). The coffee trees of these varieties are available in the field gene bank at Jimma Agricultural Research Center (JARC), Ethiopia. Young and healthy leaf samples were collected from growing tips of a single tree of each variety using a sampling tube containing silica gel. Leaf samples were transported to a molecular laboratory at the Holetta Biotechnology Research Center (HBTRC) of Ethiopian Institute of Agricultural Research (EIAR) and dried well until DNA extraction was done.

2.2. *DNA extraction*

Approximately, 0.5 g of dried leaf of each variety was ground to a fine powder using Retsch MM400 Mixer Mill (Haan Town, Cologne City, Germany). Total DNA was extracted following a modified version of CTAB method (Borsch et al. 2003). The extracted DNA was purified using the QIAGEN PCR purification kit following the manufacturer's instructions (QIAGEN 2006). Quality and concentration of the extracted DNA were determined using a Nanodrop ND-1000 spectrophotometer (Saveen Werner Company, Malmö City, Sweden). The DNA samples were diluted to a working concentration of 10 ng/μL and used for polymerase chain reaction (PCR) amplification with SSR markers.

2.3. *Polymerase chain reactions (PCRs) with SSR markers*

In the present study, 17 SSR primer-pairs, labeled with 6FAM, NED, PET or VIC fluorescent dye at 5′-end of the forward primers, were used. Annealing temperature of each primer was determined by performing gradient PCRs. All primer sets were initially screened for good amplification, polymorphism, specificity to their target loci and suitability of the allele size for multiplexing, using eight morphologically distinct coffee varieties. As a result, two primer

Table 1. Information on and categories of Ethiopian commercial arabica coffee varieties used in the study.

Variety name	Geographic origin	Genetic base (Germplasm collection)	Pedigree	Year of release	Growth habit	Altitude (m.a.s.l)
CBD[†]-resistant varieties						
E1p	Kaffa (SWE) [‡]	CBD res. Coffee coll. 1974	Pure line selection	1978	Open	1550-1970
E2	Kaffa (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1979	Open	1550-1970
E3	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1980	Open	1000-1550
E4	Kaffa (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1979	IM[§§]	1000-1970
E5	Kaffa (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1980	IM	1000-1970
E6	Kaffa (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1980	IM	1500-1970
E7	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1978	Compact	1550-1970
E8	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1978	Compact	1550-1970
E9	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1978	Compact	1550-1970
E10	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1978	Compact	1550-1970
E11	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1978	Compact	1550-1970
E12	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1978	Compact	1550-1970
E13	Illubabor (SWE)	CBD res. Coffee coll. 1974	Pure line selection	1980	Open	1550-1970
High-land varieties						
G1	Kaffa (SWE)	NCC[††] 1975	Pure line selection	2006	IM	1750-2100
G2	Jimma (SWE)	NCC 1974	Pure line selection	2006	IM	1750-2100
G3	Kaffa (SWE)	NCC 1975	Pure line selection	2006	IM	1750-2100
G4	Kaffa (SWE)	NCC 1981	Pure line selection	2006	IM	1750-2100
Mid- to low-land varieties						
M1	Mizan (SWE)	FAO collection 1965	Pure line selection	2002	MO[¶¶]	1550-1750
M2p	Kaffa (SWE)	FAO collection 1965	Pure line selection	1997	MO	1550-1750
T1	India	ICC[‡‡] 1979	Pure line selection	2002	IM	1000-1550
T2	Portugal	ICC 1979 (Catimor lines)	Pure line selection	1997	Compact	1000-1550
T3	Portugal	ICC 1979 (Catimor lines)	Pure line selection	1997	Compact	1000-1550
Hybrid varieties						
Hd1	Jimma (SWE)	NCC 1974 and FAO Coll. 1965	F1 hybrid (E7p × M2p)	2002	MC	1550-1750
Hd2	Jimma (SWE)	NCC1973 and FAO coll.1965	F1hybrid (S7395 × M2p)	1997	MO	1550-1750
Hd3	Jimma (SWE)	NCC 1974 and FAO coll.1965	F1 hybrid (E1p × M2p)	1997	MO	1550-1750
Hd4	Tepi (SWE)	FAO coll.1965 and NCC1973	F1 hybrid (M2p × SH13)	2016	MO	1550-1750
Hd5	Tepi (SWE)	FAO coll.1965 and NCC1973	F1 hybrid (M2p × SDr1)	2016	MO	1550-1750

(*Continued*)

Table 1. (Continued).

Variety name	Geographic origin	Genetic base (Germplasm collection)	Pedigree	Year of release	Growth habit	Altitude (m.a.s.l)
Hd6	Tepi (SWE)	NCC1974 and 1975	F1 Hybrid (S455 × S530)	2016	MO	1550-1750
Landrace varieties						
SR1	Gedeo (SE)[§]	NCC 1997	Pure line selection	2010	Open	1750-1850
SR2	Gedeo (SE)	NCC 1997	Pure line selection	2010	Mop.	1550-1750
SR3	Gelana (SE)	CBD res. Coffee coll. 1985	Pure line selection	2010	Open	1750-1850
SR4	Sidama (SE)	NCC 1977	Pure line selection	2006	IM	1500-1800
HR1	Hararge (EE)[¶]	NCC 1998	Pure line selection	2010	MO	1550-1750
HR2	Hararge (EE)	NCC 1998	Pure line selection	2010	IM	1650-1850
HR3	Hararge (EE)	NCC 1998	Pure line selection	2010	IM	1550-1750
HR4	Hararge (EE)	NCC 1998	Pure line selection	2010	Open	1650-1850
WR1	Wollega (WE)[#]	NCC 1998	Pure line selection	2010	IM	1550-2100
WR2	Wollega (WE)	NCC 1998	Pure line selection	2010	Open	1550-2100
WR3	Wollega (WE)	NCC 1998	Pure line selection	2010	Open	1000-1800
WR4	Wollega (WE)	NCC 1998	Pure line selection	2010	Open	1000-1800

[†]Coffee berry disease.
[‡]South-West Ethiopia.
[§]South Ethiopia.
[¶]Eastern Ethiopia.
[#]Western Ethiopia.
[††]National coffee collection.
[‡‡]International coffee collection.
[§§]Intermediate.
[¶¶]Medium open.

pairs amplified monomorphic alleles across all genotypes and one primer pair showed nonspecific amplification. Therefore, these three primer sets were excluded and detailed analyses were performed using 14 polymorphic primer sets (Table 2).

Depending on the annealing temperature of each primer set, four multiplex panels of 2–5 SSRs each were used for genotyping the coffee varieties. The PCRs for each multiplex panel were performed in a total sample volume of 10 µL, which consisted of 5 µL 2x AccuPower® Taq PCR Master Mix (BiONEER), 0.1 µL of 50 mM $MgCl_2$, 0.3 µL (3 mM) of each forward and reverse primer of one mix, 2.0 µL sample gDNA (10 ng/µL) and 0.5 µL double distilled water. The reactions were performed using the GeneAMP PCR system 9700 thermocycler. The PCR conditions followed were: 3 min initial denaturation at 94°C, followed by 30 cycles of denaturing at 94°C for 30 sec., annealing at 58°C to 62°C (depending on the multiplex group) for 1 min., and extension at 72°C for 2 min, with a final extension step at 72°C for 20 min. Amplification of each multiplex was confirmed by running 3 µL of the PCR products plus 1 µL loading dye on 1.8% agarose gel containing 1 µL GelRed®.

Then, 1 µL of the PCR product of each multiplex group was added to 8.8 µL highly deionized (Hi-di™) formamide and 0.2 µL of the Applied Biosystems GeneScan™ 500 LIZ® Size Standard. The mixed samples were then heated to 95° C for 5 min and chilled on ice for 2 min. Allele size of each variety was separated and detected by capillary electrophoresis using an ABI PRISM® 3100 Genetic Analyzer (capillary electrophoresis instrument, Applied Bioscience, Foster City, CA, USA).

Table 2. List of primer-pairs and sequences used to amplify the simple sequence repeat (SSR) loci of the arabica coffee varieties along with their annealing temperature (Ta).

		Primer sequence (5′ to 3′)		
Code	Locus name	Forward + Dye	Reverse	Ta [°C]
P1	AJ-250254[†]	GGCTCGAGATATCTGTTTAG-VIC	TTTAATGGGCATAGGGTCC	58
P2	AJ-250255[†]	CCCTCCCTGCCAGAAGAAGC-NED	AACCACCGTCCTTTTCCTCG	58
P3	AJ-250260[†]	TGATGGACAGGAGTTGATGG-6-FAM	TGCCAATCTACCTACCCCTT	58
P4	Sat-237[‡]	CAAGAGCAGACGATTCTCAATCT-6-FAM	TTGGGGTTAGGAAATCACAAT	58
P5	Sat-171[‡]	TTCCCCCATCTTTTTCTTTC-VIC	TTGTATACGGCTCGTCAGGT	58
P6	CFGA-465	ACCCTTTACTACTTATTTACTCTC-6-FAM	ACATCCCCTTGCCATTTCTTC	62
P8	AJ-250257[†]	GACCATTACATTTCACACAC-NED	GCATTTTGTTGCACACTGTA	58
P9	Sat-235[‡]	TCGTTCTGTCATTAAATCGTCAA-PET	GCAAATCATGAAAATAGTTGGTG	58
P10	MR-054[‡]	TGATGTGGAAGGCCATTG-VIC	GCCCCTATTATGACCCATGC	62
P11	Sat-180[‡]	CATGTGTAATACATTCAACAGTGA-NED	GCAATAGTGGTTGTCATCCTT	60
P12	AJ-250258[†]	AAC TCT CCA TTC CCG CAT TC-PET	CTG GGT TTT CTG TGT TCT CG	62
P13	Sat-41[‡]	AGTGTAACTTTAGTTCTTGC-PET	ATTTAATGGGCATAGGGTC	58
P15	AJ-250253[†]	CTTGTTTGAGTCTGTCGCTG-VIC	TTTCCCTCCCAATGTCTGTA	58
P16	MR-336[‡]	GAGTCGTCCACACTGCTTGA-6-FAM	CATCTGCTTTGGTCCCTGAT	60

[†]Combes et al. (2000).
[‡]Institute for Research and Development (IRD).

2.4. *Genotyping and data analysis*

The allele peaks were visually inspected and then analyzed using GeneMapper Software 4.0 (Applied Biosystems) based on the GeneScan™ 500 LIZ® Size Standard. Each peak was considered an allele of an appropriate microsatellite locus according to size (bp) and area of the peak. Accordingly, every allele of each primer across all genotypes was scored and used for the analysis of different measures of genetic diversity.

Although microsatellites are co-dominant markers, molecular analysis of SSR loci in previous studies was mostly performed on presence/absence of amplified alleles in each genotype because of the allotetraploid nature of *C. arabica* (Moncada and McCouch 2004; Maluf et al. 2005). This data-formatting method for SSR markers might result in a loss of information when genotypes are highly heterozygous (Maguire et al. 2002). However, with predominantly self-pollinated species like *C. arabica*, this problem is likely to be minimal (Powell, Machray, and Provan 1996; Pejic et al. 1998).

Therefore, to calculate the genetic parameters, scored microsatellite alleles were formatted to present (1) or absent (0) for only polymorphic marker loci across all coffee varieties (Medini et al. 2005). Number of total (Na) and polymorphic alleles (Pa) per locus, and rate of polymorphism (Pr) were calculated, as described by Morgante et al. (1994). Observed Heterozygosity (Ho) for each SSR primer was calculated according to Hormaza (2002). Polymorphic information content (PIC) values were also determined based on allelic frequency using PowerMarker version 3.25 (Liu and Muse 2005). Rare and genotype-specific alleles were also determined, as described by Jain, Jain, and McCouch (2004).

A binary data matrix was also generated from the allelic data using NTSYS–PC 2.11 software (Rohlf 2000) and a pair-wise genetic similarity matrix was constructed using the Jaccard's similarity coefficient (Jaccard 1908). The resulting similarity matrix was used for cluster analysis, and the relationships among varieties were displayed as a dendrogram constructed based on UPGMA (Sneath and Sokal 1973).

3. Results and discussion

3.1. *SSR-marker polymorphism and genetic diversity*

The genetic diversity indices for the 14 SSR markers are summarized in Table 3. All the markers used were found to be highly polymorphic and they amplified a total of 105 alleles, of which 103 were polymorphic across the evaluated commercial arabica coffee varieties. The number of total and polymorphic alleles per primer ranged from 3 to 10, with a mean of 7.5 and 7.4, respectively. Allelic polymorphism ranged from 83.3% to a maximum of 100%, with a mean of 98.1% across all coffee varieties. Observed heterozygosity (Ho) values ranged

Table 3. Levels of genetic information generated by 14 simple sequence repeat (SSR) markers in 40 Ethiopian commercial arabica coffee varieties.

Code	Locus name	NA[†]	PA[‡]	Pr[§] (%)	Ho[¶]	PIC[#]
P1	AJ-250254	10	10	100	0.70	0.94
P2	AJ-250255	3	3	100	0.80	0.75
P3	AJ-250260	9	9	100	0.93	0.84
P4	Sat-237	6	5	83.33	1.00	0.70
P5	Sat-171	6	5	83.33	0.95	0.75
P6	CFGA-465	7	7	100	0.33	0.86
P8	AJ-250257	10	10	100	1.00	0.92
P9	Sat-235	10	10	100	0.50	0.91
P10	MR-054	10	10	100	1.00	0.76
P11	Sat-180	4	4	100	0.98	0.47
P12	AJ-250258	10	10	100	0.63	0.75
P13	Sat-41	9	9	100	0.58	0.89
P15	AJ-250253	8	8	100	1.00	0.90
P16	MR-336	3	3	100	0.50	0.72
Total		105	103			
Average		7.5	7.4	98.1	0.81	0.80

[†]Number of total alleles.
[‡]Polymorphic alleles.
[§]Polymorphism rate.
[¶]Observed heterozygosity.
[#]Polymorphic information content.

from 0.32 to 1.00, with a mean of 0.83 per primer. Similarly, the PIC values ranged from 0.47 to 0.94, with a mean of 0.80 per primer.

In the present study, the mean values for all genetic parameters calculated from the allelic data of the 14 SSR markers were much higher than previously reported values for commercial arabica coffee varieties evaluated elsewhere using SSR markers (Moncada and McCouch 2004; Maluf et al. 2005; Silvestrini et al. 2007; Missio et al. 2010; Teressa et al. 2010; Al-Murish et al. 2013), indicating the presence of a high level of genetic diversity among commercial arabica coffee varieties currently grown in Ethiopia.

In contrast to the results of our study, Anthony et al. (2002) reported the presence of a low percent polymorphism and genetic diversity among 15 commercial varieties compared to wild coffee accessions using six SSR markers. Similarly, Moncada and McCouch (2004) reported a low level of genetic diversity, with a mean of 2 alleles and 0.22 PIC per primer in 12 Colombian cultivated arabica coffee varieties using 34 SSR markers. A very low level of genetic diversity, with a mean of 2.87 alleles and 0.33 PIC per primer in 26 Brazilian commercial cultivars of arabica coffee inbred lines was also reported by Maluf et al. (2005) using 23 SSR markers. Likewise, Tornincasa et al. (2006) reported a low level of genetic diversity in 45 commercial arabica coffee varieties obtained from Brazil, Guatemala, India and Africa, but they reported wide genetic variability among 96 Ethiopian accessions using 12 SSR markers.

Using 32 SSR markers, Teressa et al. (2010) also demonstrated low genetic diversity, with a mean of 2.8 alleles and 32% polymorphism in 55 commercial arabica coffee varieties obtained from France. Similarly, Geleta et al. (2012) detected a mean of 2.0 alleles and 42% polymorphism using 12 SSR markers and reported the presence of a low level of genetic diversity among eight Nicaraguan commercial arabica coffee varieties. A low level of genetic variation, with a mean of 2.5 alleles, 0.32 PIC and 0.43 Ho per primer among 17 arabica coffee cultivars grown in Yemen, was reported by Al-Murish et al. (2013) using 58 SSR markers. The low level of genetic diversity and the narrow genetic base of the commercial cultivars of arabica coffee were not only detected by using SSR markers but also by other DNA-based marker systems, such as RAPDs and AFLPs (Anthony et al. 2002; Maluf et al. 2005), RAPDs, ISSRs and SRAPs (Mishra et al. 2012), SRAPs and TRAPs (Al-Murish et al. 2013).

In aforementioned previous studies, most researchers analyzed small number of genotypes derived from the two botanical varieties ("Typica" and "Bourbon") and Hybrido de Timor, which is a spontaneous interspecific hybrid between *C. arabica and C. canephora* (Goncalves and Rodrigues 1976) (as cited in Lashermes et al. 1993). The narrow genetic base of Typica and Bourbon and low polymorphism in these two botanical varieties has also been reported by Anthony et al. (2002), Maluf et al. (2005) and Lopez et al. (2009), which could be attributable to reduced genetic diversity favored by the predominant autogamy of *C. arabica* and several cycles of selection and backcrossing in the course of developing improved varieties (Carvalho et al. 1991).

The coffee varieties evaluated in our study, however, were developed from different batches of coffee germplasm accessions collected from fairly distant geographical regions of coffee forests and farmers' fields that harbored the total gene pool of *C. arabica* L. The presence of a high level of genetic diversity in the germplasm collections from which these varieties were identified has also been reported by different researchers, who did their investigations using morphological markers (Ermias 2005; Kebede, M., and B. Bellachew 2008). Therefore, differences in genetic bases of sources of studied materials, sample size, marker systems used, number and type of nucleotide motifs of the SSR primers selected (Vuylsteke et al. 1999), played a significant role in the variations between the results of the aforementioned investigations and the present study.

Of the 14 primer sets, 11 amplified 32 rare (with a frequency of less than 5%) and 18 variety-specific/exclusive alleles (Table 4) that were distributed among 28 and 13 varieties, respectively (Table 5). These alleles had also contributed to the observed high level of genetic diversity in the coffee varieties evaluated in this study. Six rare and four specific alleles were detected at locus AJ-250,258, followed by MR-054, with five rare and three

Table 4. Number of rare and specific alleles amplified by 10 simple sequence repeat (SSR) primer sets among studied arabica coffee varieties.

Primer name	Rare alleles	Specific alleles	Allele size and frequency
AJ-250254	2	1	P1[†]-158[‡] (0.039) [§], **P1-176**[¶] (0.026)
AJ-250260	3	1	P3-112 (0.032), **P3-120** (0.011), P3-132 (0.032)
Sat-237	1	1	**P4-126** (0.008)
Sat-171	1	1	P5-142 (0.032), **P5-146** (0.011)
CFGA-465	1	0	P6-192(0.036)
AJ-250257	4	2	P8-122 (0.027), **P8-124** (0.009), **P8-126** (0.009), P8-136 (0.018)
Sat-235	3	2	**P9-250** (0.015), **P9-254** (0.031), **P9-256** (0.015)
MR-054	5	3	**P10-134** (0.018), **P10-136** (0.009), **P10-142** (0.009), **P10-144** (0.009), **P10-156** (0.035)
AJ-250258	6	4	**P12-120** (0.033), P12-128 (0.033), **P12-130** (0.016), **P12-136** (0.016), **P12-146** (0.016), **P12-148** (0.016)
Sat-41	1	1	**P13-157** (0.016)
AJ-250253	3	2	**P15-281** (0.03), **P15-311** (0.01), **P15-319** (0.01)
Total	32	18	

[†]Primer code.
[‡]Allele size.
[§]Allele frequency.
[¶]Bold and underlined values denote specific allele size.

Table 5. List of genotypes with respective rare and specific allele (bp) of 14 simple sequence repeat (SSR) markers used in the study.

		Locus code and allele size in base pair (bp)	
Ser. No.	Variety	Rare alleles	Specific alleles
1	E1p	p1[†]-176[‡], P3-132	p1-176
2	E2	P5-142	
3	E5	p1-158, P6-192	
4	E6	P3-132, P8-136	
5	E8	P15-281	
6	E11	P8-122, P9-250, P10-136, P10-144, P12-148	P9-250, P10-136, P10-144, P12-148
7	E12	P9-254, P12-120	
8	E13	P8-122	
9	G1	P5-142	
10	G3	P9-254	
11	M1	P12-136	P12-136
12	T1	P3-112, P4-126, P10-134, P10-142, P10-156, P12-128	P4-126, P10-142
13	T2	p1-158, P10-156, P12-128	
14	T3	P12-146	P12-146
15	Hd1	P8-122	
16	Hd2	P4-136, P13-157, P15-311	P13-157, P15-311
17	Hd3	p1-176, P3-132, P5-146	p1-176, P5-146
18	Hd5	P12-130	P12-130
19	Hd4	P8-124	P8-124
20	Hd6	p1-158, P5-142, P6-192, P15-319	P15-319
21	SR3	P3-120, P10-156	P3-120
22	SR4	P10-134	
23	HR1	P8-126, P15-281	P8-126
24	HR3	P8-136, P10-156	
25	WR1	P9-256, P15-281	P9-256
26	WR2	P3-112	
27	WR3	P4-136, P12-120	
28	WR4	P3-112	

[†]Locus code.
[‡]Allele size.

specific alleles (Table 4). The number of rare alleles across the 28 coffee varieties ranged from one to six. The highest (6) number of rare alleles was detected in variety T1, followed by variety E11, with five rare and four (highest) specific alleles (Table 5). The presence of rare alleles in the studied coffee varieties reflected their rich genetic diversity, suggesting the need to include those varieties in future hybridization programs to identify segregants with desirable traits. Moreover, specific/exclusive alleles detected across the 13 varieties (Table 5) could also be used as diagnostic (fingerprint) markers for discriminating the varieties in future breeding programs using the same sets of primers. Mishra et al. (2012) also detected rare and cultivar-specific alleles in Indian arabica coffee cultivars, using RAPD, ISSR and SRAP markers. Unlike the present study, the detected rare alleles did not show the presence of a high level of genetic diversity; rather, a high degree of

genetic similarity was observed among the cultivars. This could be attributable to the distribution of the detected rare allele in a small number of individuals in the evaluated genotypes, whereas rare alleles detected in our study were distributed among 70% of the evaluated varieties.

3.2. *Genetic similarity among varieties*

The Jaccard's similarity coefficient among all possible pairs-wise combinations of the varieties ranged from 0.14 (SR1 with T1) to 0.78 (E7p with E10), with a mean of 0.38 (Table 6). Overall, the genetic similarity values among 92% of all pair-wise combinations ranged from 0.14 to 0.50, confirming the aforementioned results of genetic parameters and the presence of wide genetic distances among the varieties. Relatively higher similarity values, ranging from 0.58 to 0.78, were observed between pair-wise combinations of four CBD-resistant varieties (E7p, E8, E9 and E10) of southwest origin, indicating they were the most closely related varieties at the DNA level. These varieties were characterized by compact growth habit, short internodes on the main stem and primary branches, small and round fruits as well as narrow leaves, as observed under field conditions. The results of a separate similarity analysis among local landrace varieties developed for three regions that produce coffee with unique flavor are shown in Table 7. A low level of genetic similarity, ranging from 0.18 to 0.42, with a mean of 0.33, was observed within local landrace varieties of South Ethiopian (SE) origin. The range of similarity values within varieties from the Western origin (Wollega) was 0.38 to 0.57, with a mean of 0.47, and from the Eastern (Hararge) origin, the range was 0.40 to 0.60, with a mean of 0.49, indicating unrelatedness of commercial varieties of the same geographic origin.

The large amount of heterosis observed for desirable characteristics could be exploited by crossing distantly related varieties since high genetic recombination is expected in progenies of genetically distant parents (Sneller et al. 2005). In this regard, heterotic arabica coffee F_1 hybrids with desirable characteristics have been identified and released for commercial production from previous hybridization studies in Central America (Bertrand et al. 2011) and Ethiopia (Ameha and Bellachew 1983; Bellachew 2001; Tadesse 2014). For example, F_1 hybrid coffee varieties included in the present study resulted from the early hybridization program aimed at combining resistance to CBD and productivity.

In contrast to the results of our study, Steiger et al. (2002) reported genetic similarity values higher than 0.90 in 86% of all possible pair-wise comparisons among 58 arabica cultivars grown in Hawaii Island, U.S.A, using AFLP markers. Moncada and McCouch (2004) observed a mean similarity coefficient value of 0.59 between possible pairs of 12 Colombian cultivated varieties of arabica coffee, using 34 SSR markers. Very close genetic relationships,

Table 6. Jaccard's genetic similarity coefficient matrix among 40 Ethiopian commercial varieties of arabica coffee using 14 simple sequence repeat (SSR) markers.

Gen.	E1p	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	G1	G2	G3	G4	M1	M2p	Hd1	Hd2	Hd3
E2	0.40																					
E3	0.44	0.46																				
E4	0.40	0.46	0.54																			
E5	0.40	0.39	0.54	0.46																		
E6	0.39	0.40	0.44	0.40	0.38																	
E7	0.51	0.50	0.43	0.50	0.46	0.35																
E8	0.40	0.59	0.36	0.46	0.46	0.31	0.58															
E9	0.37	0.38	0.36	0.42	0.36	0.34	0.63	0.59														
E10	0.51	0.50	0.36	0.46	0.43	0.37	0.78	0.64	0.69													
E11	0.33	0.32	0.30	0.23	0.30	0.28	0.27	0.35	0.26	0.32												
E12	0.41	0.26	0.27	0.23	0.33	0.38	0.33	0.29	0.36	0.39	0.29											
E13	0.50	0.53	0.38	0.37	0.34	0.33	0.53	0.49	0.41	0.49	0.37	0.38										
G1	0.29	0.53	0.31	0.30	0.38	0.30	0.34	0.49	0.30	0.41	0.37	0.28	0.43									
G2	0.46	0.49	0.34	0.33	0.41	0.33	0.53	0.58	0.49	0.53	0.41	0.38	0.61	0.47								
G3	0.45	0.43	0.44	0.39	0.40	0.35	0.40	0.47	0.36	0.43	0.39	0.41	0.42	0.46	0.55							
G4	0.41	0.43	0.44	0.56	0.60	0.38	0.51	0.56	0.47	0.56	0.33	0.37	0.46	0.46	0.50	0.44						
M1	0.30	0.28	0.26	0.28	0.38	0.33	0.38	0.34	0.34	0.49	0.25	0.38	0.26	0.33	0.39	0.35	0.35					
M2p	0.40	0.36	0.45	0.45	0.36	0.35	0.42	0.39	0.39	0.42	0.36	0.36	0.34	0.37	0.40	0.36	0.46	0.48				
Hd1	0.43	0.38	0.48	0.38	0.42	0.35	0.58	0.44	0.41	0.55	0.33	0.31	0.43	0.43	0.43	0.39	0.49	0.43	0.64			
Hd2	0.41	0.37	0.49	0.43	0.43	0.38	0.43	0.43	0.43	0.40	0.34	0.35	0.41	0.33	0.38	0.38	0.50	0.28	0.57	0.53		
Hd3	0.51	0.35	0.41	0.41	0.33	0.37	0.47	0.41	0.38	0.50	0.33	0.33	0.34	0.34	0.43	0.36	0.42	0.43	0.73	0.57	0.52	
T1	0.21	0.15	0.23	0.25	0.32	0.25	0.21	0.22	0.17	0.22	0.20	0.23	0.20	0.20	0.33	0.26	0.35	0.33	0.24	0.29	0.26	0.25
T2	0.25	0.18	0.27	0.33	0.40	0.26	0.27	0.23	0.26	0.26	0.15	0.30	0.28	0.21	0.34	0.30	0.41	0.35	0.28	0.31	0.29	0.24
T3	0.35	0.30	0.40	0.51	0.40	0.36	0.40	0.30	0.33	0.37	0.27	0.34	0.38	0.35	0.46	0.41	0.49	0.36	0.43	0.43	0.44	0.37
Hd4	0.43	0.39	0.52	0.49	0.43	0.35	0.49	0.45	0.42	0.49	0.39	0.34	0.41	0.38	0.48	0.43	0.54	0.38	0.61	0.56	0.54	0.55
Hd5	0.38	0.30	0.43	0.46	0.35	0.34	0.38	0.32	0.35	0.37	0.25	0.35	0.33	0.36	0.33	0.38	0.44	0.34	0.57	0.50	0.51	0.57
Hd6	0.42	0.40	0.38	0.43	0.50	0.34	0.44	0.38	0.32	0.43	0.32	0.30	0.42	0.42	0.45	0.38	0.48	0.31	0.46	0.45	0.39	0.45
SR1	0.38	0.60	0.40	0.47	0.37	0.42	0.55	0.56	0.47	0.56	0.30	0.25	0.46	0.35	0.38	0.38	0.41	0.27	0.34	0.34	0.38	0.37
SR2	0.33	0.21	0.38	0.24	0.32	0.30	0.26	0.24	0.21	0.24	0.28	0.32	0.29	0.29	0.36	0.35	0.39	0.30	0.32	0.38	0.33	0.29
Gen.	E1p	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	G1	G2	G3	G4	M1	M2p	Hd1	Hd2	Hd3
SR3	0.34	0.36	0.36	0.43	0.30	0.29	0.33	0.36	0.24	0.33	0.27	0.22	0.31	0.41	0.31	0.37	0.40	0.32	0.42	0.45	0.37	0.36
SR4	0.37	0.29	0.30	0.32	0.33	0.28	0.39	0.32	0.32	0.42	0.26	0.39	0.37	0.33	0.41	0.39	0.39	0.45	0.39	0.38	0.34	0.35

(*Continued*)

Gen.	E1p	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	G1	G2	G3	G4	M1	M2p	Hd1	Hd2	Hd3
SR3	0.34	0.36	0.36	0.43	0.30	0.29	0.33	0.36	0.24	0.33	0.27	0.22	0.31	0.41	0.31	0.37	0.40	0.32	0.42	0.45	0.37	0.36
SR4	0.37	0.29	0.30	0.32	0.33	0.28	0.39	0.32	0.32	0.42	0.26	0.39	0.37	0.33	0.41	0.39	0.39	0.45	0.39	0.38	0.34	0.35
HR1	0.41	0.40	0.40	0.43	0.37	0.33	0.40	0.43	0.30	0.40	0.25	0.26	0.35	0.35	0.38	0.40	0.48	0.39	0.37	0.48	0.35	0.34
HR2	0.41	0.37	0.40	0.44	0.40	0.42	0.40	0.44	0.40	0.40	0.33	0.49	0.46	0.35	0.46	0.41	0.53	0.33	0.40	0.43	0.53	0.39
HR3	0.37	0.32	0.33	0.42	0.30	0.40	0.33	0.29	0.32	0.35	0.20	0.33	0.33	0.24	0.41	0.33	0.39	0.31	0.33	0.27	0.31	0.33
HR4	0.41	0.40	0.34	0.47	0.31	0.33	0.37	0.37	0.27	0.37	0.30	0.31	0.38	0.32	0.38	0.45	0.45	0.21	0.34	0.31	0.33	0.34
WR1	0.32	0.49	0.34	0.30	0.31	0.36	0.34	0.49	0.33	0.41	0.27	0.28	0.39	0.47	0.52	0.42	0.38	0.36	0.37	0.34	0.27	0.34
WR2	0.37	0.58	0.36	0.33	0.33	0.44	0.36	0.43	0.33	0.43	0.36	0.27	0.45	0.57	0.53	0.47	0.40	0.32	0.36	0.36	0.35	0.41
WR3	0.41	0.43	0.37	0.36	0.40	0.29	0.44	0.43	0.36	0.47	0.26	0.37	0.50	0.42	0.50	0.41	0.53	0.46	0.46	0.45	0.38	0.36
WR4	0.40	0.46	0.36	0.36	0.36	0.35	0.43	0.46	0.46	0.46	0.30	0.33	0.49	0.45	0.62	0.44	0.51	0.32	0.36	0.42	0.32	0.38

Gen.	T1	T2	T3	Hd4	Hd5	Hd6	SR1	SR2	SR3	SR4	SR5	HR2	HR3	HR4	WR1	WR2	WR3
T2	0.59																
T3	0.36	0.53															
Hd4	0.30	0.34	0.50														
Hd5	0.27	0.38	0.53	0.57													
Hd6	0.26	0.33	0.42	0.55	0.49												
SR1	0.14	0.22	0.35	0.38	0.36	0.45											
SR2	0.41	0.39	0.43	0.45	0.37	0.37	0.18										
SR3	0.35	0.37	0.37	0.43	0.40	0.33	0.28	0.42									
SR4	0.31	0.43	0.47	0.42	0.40	0.35	0.37	0.34	0.39								
HR1	0.42	0.44	0.41	0.52	0.36	0.36	0.29	0.43	0.54	0.40							
HR2	0.39	0.41	0.57	0.47	0.41	0.39	0.38	0.39	0.37	0.33	0.44						
HR3	0.38	0.47	0.56	0.36	0.35	0.32	0.37	0.24	0.33	0.35	0.40	0.60					
HR4	0.30	0.41	0.49	0.50	0.44	0.48	0.49	0.33	0.40	0.47	0.51	0.45	0.51				
WR1	0.20	0.28	0.38	0.35	0.31	0.36	0.42	0.29	0.28	0.33	0.38	0.35	0.41	0.38			
WR2	0.21	0.24	0.40	0.37	0.38	0.41	0.48	0.32	0.30	0.36	0.31	0.31	0.33	0.37	0.57		
WR3	0.26	0.30	0.34	0.40	0.33	0.41	0.34	0.28	0.33	0.39	0.40	0.34	0.29	0.31	0.42	0.44	
WR4	0.23	0.24	0.31	0.40	0.38	0.41	0.31	0.32	0.30	0.30	0.37	0.34	0.27	0.31	0.38	0.54	0.47

Table 7. Jaccard's genetic similarity coefficient among pair-wise combinations of local landrace varieties from different geographic origins using 14 simple sequence repeat (SSR) markers.

Origin	Variety name	1	2	3	4	5	6	7	8	9	10	11	12
South Ethiopia	SR1 (1)	1.00											
	SR2(2)	0.18	1.00										
	SR3(3)	0.28	0.42	1.00									
	SR4(4)	0.37	0.34	0.39	1.00								
East Ethiopia	HR1(5)	0.29	0.43	0.54	0.40	1.00							
	HR2(6)	0.38	0.39	0.37	0.33	0.44	1.00						
	HR3(7)	0.37	0.24	0.33	0.35	0.40	0.60	1.00					
	HR4(8)	0.49	0.33	0.40	0.47	0.51	0.45	0.51	1.00				
West Ethiopia	WR1(9)	0.42	0.29	0.28	0.33	0.38	0.35	0.41	0.38	1.00			
	WR2(10)	0.48	0.32	0.30	0.36	0.31	0.31	0.33	0.37	0.57	1.00		
	WR3(11)	0.34	0.28	0.33	0.39	0.40	0.34	0.29	0.31	0.42	0.44	1.00	
	WR4(12)	0.31	0.32	0.30	0.30	0.37	0.34	0.27	0.31	0.38	0.54	0.47	1.00

with average genetic similarity values of 0.96, 0.90 and 0.87, were observed among all possible pairs of 24 Indian arabica coffee cultivars, using RAPD, ISSR, and SRAP markers, respectively (Mishra et al. 2012). Genetic similarity values ranging from 0.69 to 0.90 among 17 arabica coffee cultivars were reported by Al-Murish et al. (2013) using 58 SSR markers. Majority of the genotypes evaluated in these studies were descendants of the botanical varieties Typica and Bourbon, and selections from segregating progenies of Hybrido de Timor. Thus, the observed high level of genetic similarity reported by the researchers could be attributed to sharing of a large proportion of common alleles (Cidade et al. 2013) because of the narrow genetic base of the gene pool, from where the studied genotypes were derived.

3.3. *Clustering pattern among varieties*

Based on the Jaccard's similarity coefficient, the UPGMA cluster analysis identified six major clusters and three singletons at a genetic distance of >0.54 (Figure 1). Cluster I consisted of six varieties that were assigned to two sub-clusters. The sub-cluster Ia contained one CBD-resistant variety from southwest and one local landrace variety from the south; whereas four compact CBD-resistant varieties from the southwest were grouped into sub-cluster Ib. Cluster II comprised four varieties from three categories: two local landrace varieties from the west, one CBD-resistant and one highland variety from the southwest. Cluster III consisted of two local landrace varieties from the western region (recommended for mid- to high-altitude areas) and two highland varieties from the southwest. Cluster IV, the largest cluster, consisted of 11 varieties from the southwest, which formed two sub-clusters. Sub-cluster IVa contained one highland and three CBD-resistant varieties, whereas sub-cluster IVb comprised of all hybrid varieties and one mid-altitude variety from the southwest that was involved as a male parent in

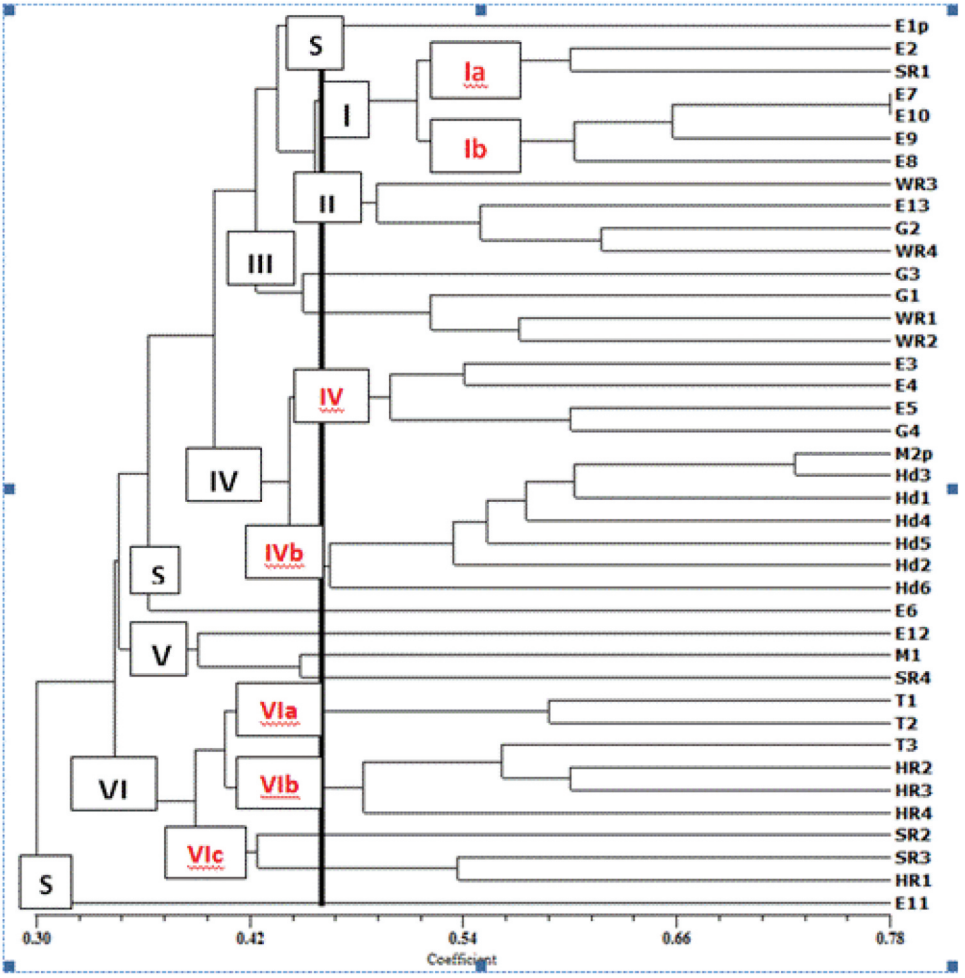

Figure 1. UPGMA dendrogram (based on Jaccard's similarity coefficient) showing genetic relationship among Ethiopian commercial varieties of arabica coffee based on 14 SSRs markers.

four of the six hybrid varieties. Cluster V consisted of a single CBD-resistant variety and a single mid-altitude variety from the southwest, and one local landrace from the south.

The second largest cluster was Cluster VI that contained nine varieties, which formed three sub-clusters. Sub cluster VIa comprised two lowland varieties, whereas three eastern varieties and one lowland variety were grouped in sub-cluster VIb. Two local landrace varieties from the south and one from the east were grouped into sub-cluster VIc. Three Varieties, E1p, E6 and E11, were singletons (S) in the dendrogram (mean genetic similarity of 0.46).

The results of cluster analyses also corroborate the results of genetic parameters and the estimates of genetic distances. In the present study, the number of clusters and the genetic distance value, at which the clusters were

determined (0.46 similarity coefficient), were higher than previously reported values for commercial varieties of arabica coffee. Three clusters at a genetic distance, ranging from 0.10 to 0.33, were reported in previous studies by several authors using RFLPs and SSRs (Anthony et al. 2002), RAPDs and SSRs (Maluf et al. 2005), EST-SSRs and gSSRs (Missio et al. 2011), RAPDs, ISSRs and SRAPs (Mishra et al. 2012) and SRAPs, TRAPs and SSRs (Al-Murish et al. 2013).

The diversity representation in the dendrogram also indicates distribution of varieties into different clusters regardless of their geographic origin, which could be attributed to the presence of a high level of genetic diversity within varieties of the same geographic origins. This could facilitate easy identification of divergent parents for hybridization to exploit heterosis to develop productive hybrid varieties for each region. The grouping of the southwest and eastern local landrace varieties into different clusters clearly indicated the presence of high genetic variation between them at the molecular level, which might be associated with genes responsible for adaption to various environmental conditions. Poor performance of majority of the southwest origin CBD-resistant varieties in major coffee-producing areas of Eastern part (Hararge) of the country was reported by Ameha and Bellachew (1987). The grouping of hybrid varieties into a cluster also indicates sharing of more common alleles among themselves than with pure line varieties. In fact, majority of them had Mp2 as the male parent; therefore, a close relationship among them was naturally expected.

4. Conclusion

This study has demonstrated the highly polymorphic nature of the SSRs used and has clearly revealed the presence of a high level of genetic diversity among commercial arabica coffee varieties grown in Ethiopia. The selection of well-known divergent genotypes as parents for crossing could increase the level of variation present in a segregating population, which should be useful in a coffee improvement program. The studied materials could serve as a potential source of genes responsible for broad adaptation in the current climate-change scenario that threatens coffee production. It may also imply the need for exploiting the observed diversity by designing a strong crossing program to develop hybrid varieties that combine high yield, quality and resilience to climate change. The present fingerprint data could be used to construct a DNA reference database for the molecular identification of the varieties in future breeding program.

Acknowledgments

We are grateful to the financial and material support provided by the Ethiopian Institute of Agricultural Research (EIAR); Agricultural Growth Program (AGP) II, Ethiopia; Mocca Plantation Coffee (MOPLACO) Plc, Ethiopia; Collaborative Coffee Sources (CCS), Norway; the African Biosciences Challenge Fund (ABCF) program at the BecA-ILRI hub, Nairobi, Kenya; (The ABCF Program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill & Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID) and the Swedish International Development Cooperation Agency (SIDA)) and Jimma University College of Agriculture and Veterinary Medicine (JUCAVM). We are also greatly thankful to the Jimma Agricultural Research Center (JARC) for providing the genetic materials for this study as well as Ethiopian Biodiversity Institute (EBI), for permitting transfer of the gDNA samples to Nairobi, Kenya for genotyping at BecA ILRI-Hub molecular laboratory. We thank the editor, Professor Manjit S. Kang, and two anonymous reviewers for detailed comments and suggestions that greatly improved the manuscript.

Disclosure statement

The authors declare that they have no conflict of interest

Funding

This work was supported by the Ethiopian Institute of Agriculture (EIAR); Collaborative Coffee Sources (CCS), Norway; Mocca Plantation Coffee (MOPLACO) plc, Ethiopia; and the African Biosciences Challenge Fund (ABCF) program through the BecA-ILRI hub, Nairobi, Kenya.

References

Al-Murish, T. M., A. A. Elshafei, A. A. Al-Doss, and M. N. Barakat. 2013. "Genetic Diversity of Coffee (*Coffea Arabica* L.) In Yemen via SRAP, TRAP and SSR Markers." *Journal of Food, Agriculture and Environment* 11 (2): 411–416.

Ameha, M., and B. Bellachew. 1983. "Heterosis in Cross of Indigenous Coffee Selected for Yield and Resistance to Coffee Berry Disease II-first Three Years." *Ethiopian Journal of Agricultural Science* 1: 13–21.

Ameha, M., and B. Bellachew. 1987. "Genotype × Environment Interaction in Coffee (*Coffea Arabica* L.)" Pp476–482. In: 14th International Scientific Colloquium on coffee. 4 June - 3 July 1987, Montreux, France.

Anthony, F., B. Bertrand, O. Quiros, A. Wilches, P. Lashermes, J. Berthaud, and A. Charrier. 2001. "Genetic Diversity of Wild Coffee (*Coffea Arabica* L.) Using Molecular Markers." *Euphytica* 118: 53–65. doi:10.1023/A:1004013815166.

Anthony, F., 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 and Applied Genetics* 104 (5): 894–900. doi:10.1007/s00122-001-0798-8.

Bandelj, D., J. Jakse, and B. Javornik. 2004. "Assessment ofgenetic variability of olive varieties by microsatellite and AFLP markers." Euphytica 136: 93–102

Bellachew, B. 2001. "Arabica Coffee Breeding for Yield and Resistance to Coffee Berry Disease (*Colletotrichum Kahawaesp.nov.*)." Dissertation, Imperial College at Wye University of London, 272.

Bellachew, B., B. Atero, and F. Tefera. 2000. "Breeding for Resistance to Coffee Berry Disease in Arabica Coffee: Progress since 1973." In: *Proceedings of the workshop on control of Coffee Berry Disease (CBD) in Ethiopia held in Addis Ababa*. 13–15 August 1999, Addis Ababa: Ethiopian Agricultural Research Institute, 85–98.

Bellachew, B., and J. P. Labouisse. 2006. "Arabica Coffee (*Coffea Arabica* L.) Local Landrace Development Strategy in Its Center of Origin and Diversity." In *21st International Scientific Colloquium on Coffee*, 11–15, Montpelier, France.

Bertrand, B., E. Alpizar, L. Lara, R. SantaCreo, M. Hidalgo, J. M. Quijano, C. Montagnon, F. Georget, and H. Etienn. 2011. "Performance of *Coffea Arabica* F1 Hybrids in Agro Forestry and Full-sun Cropping Systems in Comparison with American Pure Line Cultivars." *Euphytica* 181: 147–158. doi:10.1007/s10681-011-0372-7.

Borsch, T., W. K. Hilu, D. Quandt, V. Wilde, C. Neinhuis, and W. Barthlott. 2003. "Non-coding Plastid 42 *trn*T-*trn*F Sequences Reveal a Well Resolved Phylogeny of Basal Angiosperms." *Journal of Evolutionary Biology* 16: 558–576. doi:10.1046/j.1420-9101.2003.00577.x.

Carvalho, A., H.P. Medina Filho, L.C. Fazuoli, O. Guerreiro Filho, M.M. Lima. 1991. "Aspectos genéticos do cafeeiro." Revista Brasileira de Genética l4:135–183

Chala, J., A. Girma, T. Demelash, and Z. Arega. 2012. "Development and Release of Coffee Berry Disease Resistant Varieties to Specialty Coffee Producing Regions in Ethiopia." In: *Proceedings 24th ASIC, International Scientific Conference on Coffee Science*, 409, San José, Costa Rica, November 11–16.

Cidade, F.W., B.B. Vigna, F.H. de Souza, J.F.M. Valls, M. Dall'Agnol, M.I. Zucchi, T.T. de Souza-Chies, and A.P. Souza. 2013. "Genetic variation in polyploid forage grass: Assessing the molecular genetic variability in the Paspalum genus." BMC Genetics, 14(1): 50. doi: doi:10.1186/1471-2156-14-50

Clarindo, W. R., and C. R. Carvalho. 2008. "First *Coffea Arabica* Karyogram Showing that This Species Is a True Allotetraploid." *Plant Systematics and Evolution* 274: 237–241. doi:10.1007/s00606-008-0050-y.

Combes, M.C., S. Andrzejewski, F. Anthony, B. Bertrand, P. Rovelli, G. Graziosi, and P. Lashermes. 2000. "Characterization of microsatellite loci in *Coffea arabica* and related coffee species." *Mol. Ecol.* 9:1178–1180.

Davis, A. P. 2011. "*Psilanthus Mannii*, the Type Species of *Psilanthus*, Transferred to *Coffea*." *Nordic Journal of Botany* 29: 471–472. doi:10.1111/j.1756-1051.2011.01113.x.

Dessalegn, Y., L. Herselman, and M. T. Labuschagne. 2009. "Comparison of SSR and AFLP Analysis for Genetic Diversity Assessment of Ethiopian Arabica Coffee Genotypes." *South African Journal of Plant and Soil* 62 (2): 119–125. doi:10.1080/02571862.2009.10639943.

Ermias, H. 2005. "Evaluation of Wellega coffee germplasm for yield, yield components and esistants to coffee berry disease at early bearing stage." An MSc thesis submitted to school of graduate studies of Alemaya University, Ethiopia. 69p

Fekadu, T., B. Bellachew, B. Atero, A. Ayano, and T. Benti. 2008. "Germplasm Collection and Maintenance of Coffee (*Coffea Arabica* L)." In *Coffee Diversity and Knowledge, Proceedings of a National Workshop on Four Decades of Coffee Research and Development in Ethiopia*, edited by G. Adugna, B. Belachew, T. Shimber, E. Taye, and T. Kufa, 45–49. Addis Ababa, Ethiopia. 14–17 August 2007.

Geleta, M., I. Herrera, A. Monzón, and T. Bryngelsson. 2012. "Genetic Diversity of Arabica Coffee (*Coffea Arabica* L.) In Nicaragua as Estimated by Simple Sequence Repeat Markers."

The Scientific World Journal, Article ID 9398202012 (2012): 1–11. doi:10.1100/2012/939820.

Hormaza, J. I. 2002. "Molecular Characterization and Similarity Relationships among Apricot (*Prunus Armeniaca* L.) Genotypes Using Simple Sequence Repeats." *Theoretical and Applied Genetics* 104: 321–328. doi:10.1007/s001220100684.

Jaccard, P. 1908. "Nouvelles recherchés sur la distribution florale." *Bulletin de la Société Vaudoise des Sciences Naturelles* 44: 223–270. doi:10.5169/seals-268384.

Jain, S., R. K. Jain, and S. R. McCouch. 2004. "Genetic Analysis of Indian Aromatic and Quality Rice (*Oryza Sativa* L.) Germplasm Using Panels of Fluorescently Labeled Microsatellite Markers." *Theoretical and Applied Genetics* 109: 965–977. doi:10.1007/s00122-004-1700-2.

Jingade, P., A. K. Huded, B. Kosaraju, and M. K. Mishra. 2019. "Diversity Genotyping of Indian Coffee (*Coffea Arabica* L.) Germplasm Accessions by Using SRAP Markers." *Journal of Crop Improvement* 33: 327–345. doi:doi.10.1080/15427528.2019.1592050.

Kebede, M., and B. Bellachew. 2008. "Phenotypic Diversity in the Hararge Coffee (*Coffea arabica L.*) Germplasm for Quantitative Traits." *East African Journal of Science* 2(1): 13–18. doi:10.4314/eajsci.v2i1.40358.

Lashermes, P., J. Cros, P. Marmey, and A. Charrier. 1993. "Use of Random Amplified DNA Markers to Analyze Genetic Variability and Relationships of *Goflea* Species." *Genetic Resources and Crop Evolution* 40: 91–99. doi:10.1007/BF00052639.

Lashermes, P., P. Trouslot, F. Anthony, M. C. Combes, and A. Charrie. 1996. "Genetic Diversity for RAPD Markers between Cultivated and Wild Accessions of *Coffea Arabica*." *Euphytica* 87: 59–64. doi:10.1007/BF0002296.

Liu, K., and S. V. Muse. 2005. "Power Marker Integrated Analysis Environment for Genetic Marker Data." *Bioinformatics* 21 (9): 2128–2129. doi:10.1093/bioinformatics/bti282.

López-Gartner, G., H. Cortina, S.R. Maria, P. Moncada. 2009. "Analysis of genetic structure in a sample of coffee (Coffea arabica L.) using fluorescent SSR markers." Tree Genetics and Genomes 5: 435–446.

Maguire T.L., R. Peakall, and P. Saenger. 2002. "Comparative Analyzes of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (*Avicenniaceae*) detected by AFLPs and SSRs." *Theoretical and Applied Genetics* 104:338–398.

Maluf, M. P., M. Silvestrini, L. M. Ruggiero, C. O. de Guerreiro-filho, and C. A. Colombo. 2005. "Genetic Diversity of Cultivated Coffea Arabica Inbred Lines Assessed by RAPD, AFLP and SSR Marker Systems." *Scientia Agricola* 62: 366–373. doi:10.1590/S0103-90162005000400010.

Medini, M., S. Hamza, A. Rebai, and M. Baum. 2005. "Analysis of Genetic Diversity in Tunisian Durum Wheat Cultivars and Related Wild Species by SSR and AFLP Markers." *Genetic Resources and Crop Evolution* 52: 21–31. doi:10.1007/s10722-005-0225-0.

Mishra, M. K., N. Sandhyarani, N. Suresh, S. Satheesh Kumar, P. R. Soumya, M. H. Yashodha, A. Bhat, and Jayarama. 2012. "Genetic Diversity among Indian Coffee Cultivars Determined via Molecular Markers." *Journal of Crop Improvement* 26: 727–750. doi:10.1080/15427528.2012.696085.

Missio, R. F., E. T. Caixeta, E. M. Zambolim, G. F. Pena, L. Zambolim, L. A. S. Dias, and N. S. Sakiyama. 2011. "Genetic Characterization of an Elite Coffee Germplasm Assessed by gSSR and EST-SSR Markers." *Genetics and Molecular Research* 10: 2366–2381. doi:10.4238/2011.October.6.2.

Missio, R. F., E. T. Caixeta, E. M. Zambolim, L. Zambolim, C. D. Cruz, and N. S. Sakiyama. 2010. "Polymorphic Information Content of SSR Markers for *Coffea Spp*." *Crop Breeding and Applied Biotechnology* 10: 89–94. doi:10.12702/1984-7033.v10n01a12.

Morgante, M., M. Hanafey, and W. Powell. 1994. "Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes." *Nature Genetics* 30(2):194–200. doi:10.1038/ng822.

Moncada, P., and S. McCouch. 2004. "Simple Sequence Repeat Diversity in Diploid and Tetraploid *Coffea* Species." *Genome* 47: 501–509. doi:10.1139/g03-129.

Pejic, I., P. Ajmone-Marsan, M. Morgante, V. Kozumplick, P. Castiglioni, G. Taramino, and M. Motto. 1998. "Comparative Analysis of Genetic Similarity among Maize Inbred Lines Detected by RFLPs, RAPDs, SSRs, and AFLPs." *Theoretical and Applied Genetics* 97: 1248–1255. doi:10.1007/s001220051017.

Powell, W., G. C. Machray, and J. Provan. 1996. "Polymorphism Revealed by Simple Sequence Repeats." *Trends in Plant Science* 1: 215–222. doi:10.1016/S1360-1385(96)86898-0.

QIAGEN. 2006. "DNeasy® Plant Handbook." Qiagen, July 07. http://www.qiagen.com

Rohlf, F. J. 2000. *NTSYS-pc Numerical Taxonomy and Multivariate Analysis System Version 2.11. Exeter* Publishing Ltd. Setauket, New York.

Silvestrini, M., M. G. Junqueira, A. C. Favarin, O. Guerreiro-Filho, M. P. Maluf, M. B. Silvarolla, and C. A. Colombo. 2007. "Genetic Diversity and Structure of Ethiopian, Yemen and Brazilian *Coffea Arabica* L. Accessions Using Microsatellites Markers." *Genetic Resources and Crop Evolution* 54: 1367–1379. doi:10.1007/s10722-006-9122-4.

Silvestrini, M., M. P. Maluf, M. B. Silvarolla, O. Guerreiro-Filho, H. P. Medina-Filho, M. M. T. Vanini, A. S. Oliveira, C. de Gaspari-pezzopane, and L. C. Fazuoli. 2008. "Genetic Diversity of a *Coffea* Germplasm Collection Assessed by RAPD Markers." *Genetic Resources and Crop Evolution* 55: 901–910. doi:c10.1007/s10722-007-9295-5.

Sneath, P. H. A., and R. R. Sokal. 1973. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. San Francisco, CA: W. H. Freeman.

Sneller, C. H., R. L. Nelson, T. E. Carter, and Z. Cui. 2005. "Genetic Diversity in Crop Improvement: The Soybean Experience." In *Genetic and Production Innovations in Field Crop Technology: New Developments in Theory and Practice*, edited by M. S. Kang, 103–144. Delhi: International Book Distributing .

Steiger, D. L., C. Nagai, P. H. Moore, C. W. Morden, R. V. Osgood, and R. Ming. 2002. "AFLP Analysis of Genetic Diversity within and among *Coffea Arabica* Cultivars." *Theoretical and Applied Genetics* 105: 209–215. doi:10.1007/s00122-002-0939-8.

Tadesse, B. 2014. "Progress in Arabica Coffee Breeding in Ethiopia: Achievements, Challenges and Prospects." *International Journals of Sciences: Basic and Applied Research (IJSBAR)* 33 (2): 15–25.

Teressa, A., D. Crouzillat, V. Petiard, and P. Brouhan. 2010. "Genetic Diversity of Arabica Coffee (*Coffea Arabica* L.) Collections." *EJAST* 1 (1): 63–79.

Tesfaye, K., K. Govers, E. Bekele, and T. Borsch. 2014. "ISSR Fingerprinting of *Coffea Arabica* Throughout Ethiopia Reveals High Variability in Wild Populations and Distinguishes Them from Landraces." *Plant Systematics and Evolution* 300: 881–897. doi:10.1007/s00606-013-0927-2.

Tornincasa, P., R. Dreos, B. De Nardi, E. Asquini, J. Devasia, M. K. Mishra, L. Del Terra, P. Crisafulli, A. Pallavicini, and G. Graziosi. 2006. "Genetic Diversity of Commercial Coffee (*C. Arabica* L.) From America, India and Africa Assessed by Simple Sequence Repeats (Ssrs)." In *Proceedings of the 21st International Association for Coffee Science (ASIC)*, 778. Montpellier, France.

Vuylsteke, M., R. Mank, R. Antonise, E. Bastiaans, M. L. Senior, C. W. Stuber, A. E. Melchinger, et al. 1999. "Two High-density AFLP Linkage Maps of *Zea Mays* L.:

Analysis of Distribution of AFLP Markers." *Theoretical and Applied Genetics* 99: 921–935. doi:10.1007/s001220051399.

Weising, K., H. Nybom, K. Wolff, and G. Kahl. 2005. *DNA Fingerprinting in Plants: Principles, Methods, and Applications*. 2nd ed ed. CRC Press: Boca Raton, FL.