INT. J. BIOL. BIOTECH., 10 (2): 167-177, 2013.

AVOIDING PCR INHIBITION TO IMPLEMENT MICROSATELLITE MARKERS FOR CHARACTERIZATION OF COFFEE GENOTYPES

Kagnew Hailesellasse Sene and Belayneh Admassu

Ethiopian Institute of Agricultural Research, Holetta Agricultural Research Center, Holetta Agricultural Biotechnology Laboratory, P.O.Box 2003 code 31 Holetta, Ethiopia.
E-mail: kagnewh2004@yahoo.com

ABSTRACT

In the present study, two DNA extraction methods, namely, promega kit and CTAB method were utilized to optimize Polymerase chain reaction (PCR) protocol to amplify microsatellite of *Coffea arabica* based on a standardized PCR experiment. The experiment was conducted using explant of a newly released variety of coffee named Aba buna and tissue culture derived lines propagated from Aba buna. Primarily, the hybrid coffee, Aba buna was studied using 7 primer pairs that amplified microsatellite with genbank accession number AJ308755, AJ308779, AJ308782, CFGA91, CFGA92, AJ250251 and AJ250253. And for further experiments, three individuals of the coffee explant and the other two, tissue culture derived individuals micropropagated from Aba buna were studied at molecular level using 11 primer pairs, which amplified 11 microsatellites known to be with genbank accession number AJ308755, AJ308779, AJ308782, AJ308790, AJ308837, CFGA91, CFGA92, CFGA69, CFGA100, CFGA465 and CFGA502. The PCR amplicons were analyzed in horizontal agarose gel electrophoresis. Microsatellite allele separation was investigated using vertical polyacryl amide gel electrophoresis. PCR amplification failed using a highly colourful template DNA extracted by the promega kit. However, the PCR experiments resulted in high yield of microsatellite amplicon when a template DNA extracted by the Doyle & Doyle CTAB method was used. The present finding suggests that diverse secondary metabolites occurring in coffee leaf might have acted as PCR inhibitor though the exact mechanism of inhibition is unclear. The present research gave insight in promotion of the Doyle & Doyle CTAB method for isolating coffee DNA and recommends the need of modifying the promega kit to extract a coffee genomic DNA free of impurities. In a nutshell, the Doyle & Doyle CTAB method seems the most promising method to purify clean DNA, which is an ideal template to amplify SSR of *C.arabica* for genotyping coffee cultivars.

Key words: Agarose, Allele, *C.arabica*, CTAB, Electrophoresis, Genotyping, Microsatellite, PCR, Polyacrylamide, Promega

INTRODUCTION

Ethiopia is the origin and centre of diversity for *C.arabica*. *C.arabica* is the only tetraploid species out of all the diploid species in the genus Coffea. This cash crop is indigenous to south west Ethiopia (Sera *et al.*, 2003). *C.arabica* is allotetarploid with chromosome number 2n=4x=44 (Pinto-Maglio, 2006). *C.arabica* is a cross between *C.canephora* and *C. eugenioides* with genome size of 1,300 mega base pair (http://www.coffeegenome.org). The two cultivated species of the genus *Coffea* are *C.arabica*, Arabica coffee and *C.canephora*, Robusta coffee (Davis and Stoffelen, 2006). Harar coffee, Sidamo coffee, Wollega coffee, Limu coffee, Jimma coffee, Illubabor coffee, Yirgacheffe coffee and Tepi and Bekeka coffee are specific speciality names. These specialities are classified based on the place in which they are produced (Alemayehu *et al.*, 2010). Also breeding experiments allowed release of hybrid coffee like Ghawe, Aba buna and MCH2 (Ramos *et al.*, 2009). The genetic variation within and among arabica cultivars is limited (Steiger *et al.*, 2002). Coffee is thought to be originated from a species, which looks like *C.eugenoids*. There exists gene flow from diploid coffee species to the tetraploid *C. arabica* (Herrera *et al.*, 2004). Coffee grows under the canopy of forested land. *Citrus sinensis, Persea Americana, Cedrela odorata, Acacia mangium, Astronium graveolens, Dalbergia retusa, Inga* spp. *Erythrina* Spp. and *Senna siamea* are some of the shade trees (Albertin and Nair, 2004). In fact, there are also some coffee varieties which grow without shade. Diversity of Coffee reduces during dissemination (Anthony *et al.*, 2002).

Molecular characterization of coffee via the aid of molecular markers, characterization using biochemical markers like isozyme and morphological characterization are novel approaches in availing taxonomic variations among different genotypes or accessions of coffee. SSR are commonly used in detecting variation within accessions of *C.arabica* for they are co-dominant, polymorphic, transferable/reproducible, multiallelic and informative. Simple sequence repeats (microsatellites) of *C.arabica* could contain 1 to 13 nucleotide repeats. In earlier periods, using techniques of equilibrium buoyant-density ultracentrifugation, satellite band was distinguished from band of the principal DNA. SSR are highly concentrated around the centromere. It is also localized at the end, telomere. The

origin of microsatellite is related to backward slippage of a daughter strand when DNA replicates itself. In this case, the same short sequence is repeated twice (Lodish *et al.*, 2004). There are available several studies existed in relation to SSR based characterization of *C. arabica*. The diversity of 15 coffea species was studied via applying tools of 60 microsatellite markers. Two cultivated and 2 related wild coffees were compared (Cubry *et al.*, 2008). Genetic diversity of *C.arabica* collected from Wollega, Illubabor, Keffa, Jimma and Sidamo were studied using 32 SSR markers (Alemayehu *et al.*, 2010). SSR studies were conducted on *Coffea* species collected from north of Jimma, Kaffa, Sidamo, Dilla, Teppi-Mizan Teferi and Illubabor (Moncada and McCouch, 2004). Formerly, 367 SSR markers were developed and frequency of the SSR motif was studied (Cristancho and Gaitán, 2008). Eleven microsatellite regions of *C. arabica* were characterized via the aid of SSR primers (Combes *et al.*, 2000).

Alkaloids like caffeine and trigonelline were reported to be secondary metabolites (Tesfaye *et al.*, 2007). Caffeine, 1, 3, 7-trimethyl-xanthine, is a purine alkaloid present in coffee. Caffeine is produced from the purine nucleotide AMP, GMP/IMP. It is synthesized in the leaf. There are various mechanisms in which a PCR inhibitor inhibits polymerase chain reaction. Taq polymerase utilizes Mg^{2+} as a cofactor. Inhibitors may intervene with the interaction between Mg^{2+} and Taq polymerase or the amount of Mg^{2+} available for Taq polymerase might be reduced to inhibitors. Binding of inhibitor to Taq polymerase thereby making the enzyme inactive and interaction of the inhibitor with the template DNA are mentioned to be principal causes of inhibition (Opel *et al.*, 2009). Phenol was documented to be PCR inhibitor (Roux, 2006). Secondary metabolites like polysaccharides interact with nucleic acid there by acting as oxidizing agents. This affects quality and quantity of DNA. Yellowish or Yellowish brown color of DNA is due to phenol (Bhattacharjee *et al.*, 2009). The choice of the best method of DNA isolation will have a role to isolate genomic DNA, which is free from secondary compounds. There are reports for the existence of colored DNA pellet in other crops like *Aconitum balfourii Stapf*. This colorful DNA is due to the existence of secondary metabolites in the purified DNA. The inclusion of Polyvinylpyrrolidone and β-mercaptoethanol in the CTAB extraction buffer eliminates secondary metabolites (Hatwal *et al.*, 2011). The goal of the present study is to fill the gap in knowledge about PCR performance in the presence and absence of secondary metabolites.

MATERIAL AND METHODS

Primers were screened from published articles and the nucleotide sequence of the primers was cross checked from the microsatellite sequence retrieved from Genbank, NCBI. 100 µM stock solution of primers were prepared as per instructions from the company, microsynth, Switzerland. A working solution of 10 µM was directly utilized for PCR reaction. Sequence analysis was performed using microsatellite sequence obtained from NCBI. Expected size of PCR product obtained from the original article in which the SSR is published. Also for some of the microsatellite, the expected PCR product is computed via aligning the forward and reverse primer to the retrieved microsatellite sequence. Example of computing the size of the PCR product is shown in Fig 1. Primer dimer, self complementarity and expected length of dimers were predicted using the software Vector NTI Advance 10. DNA was extracted from a fresh leaf of hybrid *C. arabica* identified with vernacular name, Aba buna, preserved in a green house. Three additional DNA samples were extracted from a single individual of explant, Aba buna and two tissue culture derived individuals, which are acclimatized in a greenhouse. The DNA was extracted by the promega kit (www.promega.com). The leaf powder was treated with nuclei lysis solution and incubated in a water bath. Immediately, the solution was incubated with RNase. The supernatant obtained after the addition of the protein precipitation solution was mixed with isopropanol to precipitate DNA. The obtained pellet was dried and dissolved in a DNA rehydration solution. The DNA solution stored at 4^{o}C. The aforementioned samples were extracted by the Doyle and Doyle CTAB method (Doyle and Doyle, 1990). Obviously, in case of the CTAB method, a leaf powder in a preheated CTAB buffer was incubated in a water bath. Chloroform: isoamyl alcohol mix was added and the most upper solution was gently taken using a micropipette. This solution was further treated with isopropanol to get the DNA pellet. The washed pellet using washing buffer was made free of RNA via a TE buffer containing RNase. The pellet was dried and dissolved in nuclease free water. Aliquot of the DNA was stored at 4^{o}C. In the next day, all of the DNA samples were stored at -20 ^{o}C. The DNA was isolated from the leaf crushed using a geno grinder machine and mortar and pestel. Some of the leaf samples were lyophilized or freeze dried in the mortar for overnight and crushed in a mortar with pestel to make leaf powder. Also the fresh leaves were directly crushed. The rest leaf samples were preserved in deep freeze for overnight and a piece of leaf in an appendof containing beads was crushed using the geno grinder. The geno grinder with the right frequency was used to grind the leaf. The amount of DNA was quantified in a nano drop, spectrophotometer. The extracted DNA was analyzed in horizontal agarose gel electrophoresis.

Table 1. List of accessions under study, PCR product and specific size.

S.No	Accession number	PCR product size in bp, Manually computed based on alignment of primers to the microsatellite sequence	PCR product size in bp from article	Type of repeat	Number of repeat	Source for type and number of repeat
1.	AJ250251	279	Ranged from 248 to 258 as per (COMBES *et al.*, 2000)	CA/CA/CA/CA/CA/CA/CA/CA	6,3,3,3,4,3,3,3	Genbank, NCBI and (COMBES *et al.*, 2000)
2.	AJ250253	291	Ranged from 240 to 270 as per (COMBES *et al.*, 2000)	GA/GT/TT/GT/TT/GT/GA/TC/CT/GT	5,8,1,4,1,7,11,2,3,1	Genbank, NCBI and (COMBES *et al.*, 2000)
3.	CFGA91	No sequence information	Ranged from 284 -303 as per (Moncada and McCouch, 2004)	AG	17	(Moncada and McCouch, 2004)
4.	CFGA92	No sequence information	Ranged from 299 – 310 as per (Moncada and McCouch, 2004)	AG-N-AG	5,8,5	(Moncada and McCouch, 2004)
5.	AJ308755	184	No information from article	CA	20	Genbank, NCBI and (Cubry *et al.*, 2008)
6.	AJ308779	116	No information from article	TG	17	Genbank, NCBI and (Cubry *et al.*, 2008)
7.	AJ308782	114	No information from article	GT	15	Genbank, NCBI and (Cubry *et al.*, 2008)
8.	AJ308790	134	No information from article	GT	21	Genbank, NCBI and (Cubry *et al.*, 2008)
9.	AJ308837	102	No information from article	TG,GA	16,11	Genbank, NCBI and (Cubry *et al.*, 2008)
10.	CFGA69	There is sequence information , but it didn't help to compute size	Ranged from 111 – 139 bp as per (Moncada and McCouch, 2004) and and other source cited by this article, 145bp	AG/GT	14,11	(Moncada and McCouch, 2004)
11.	CFGA100	There is sequence information , but it didn't help to compute size	Ranged from 177 – 185 bp as per (Moncada and McCouch, 2004) and other source cited by this article, 248bp	AG	15	(Moncada and McCouch, 2004)
12.	CFGA465	There is sequence information , but it didn't help to compute size	Ranged from 147 – 173 as per (Moncada and McCouch, 2004) and other source cited by this article, 157bp	AG	18	(Moncada and McCouch, 2004)
13.	CFGA502	There is sequence information , but it didn't help to compute size	Ranged from 140- 196 bp (Moncada and McCouch, 2004) and other source cited by this article, 180bp	AG	27	(Moncada and McCouch, 2004)

Table 2. List of primers used in this study

S.No	EMBL. Acc.No.	Primer sequence 5' → 3'	Source
1.	AJ308755	F:CCCTCCCTCTTTCTCCTCTC R:TCTGGGTTTTCTGTGTTCTCG	(Cubry *et al.*, 2008)
2.	AJ308779	F:TCCCCCATCTTTTTCTTTCC R:GGGAGTGTTTTTGTGTTGCTT	(Cubry *et al.*, 2008)
3.	AJ308782	F:AAAGGAAAATTGTTGGCTCTGA R: TCCACATACATTTCCCAGCA	(Cubry *et al.*, 2008)
4.	AJ308837	F:CTCGCTTTCACGCTCTCTCT R:CGGTATGTTCCTCGTTCCTC	(Cubry *et al.*, 2008)
5.	AJ308790	F:TTTTCTGGGTTTTCTGTGTTCTC R:TAACTCTCCATTCCCGCATT	(Cubry *et al.*, 2008)
6.	CFGA91	F: CTTCTCAGCTTTAGGTTCACTTTG R: TTTTGAATACTGGCTCGTGAACTT	(Moncada and McCouch, 2004)
7.	CFGA92	F: TGAGGGCAAAGGAGTAAGAAAG R: TCAAACTTCAACAATCAAATACCC	(Moncada and McCouch, 2004)
8.	AJ25025l	F: ATTCTCTCCCCCTCTCTGC R: TGTGTGCGCGTTTTCTTG	(COMBES *et al.*, 2000)
9.	AJ250253	F: CTTGTTTGAGTCTGTCGCTG R: TTTCCCTCCCAATGTCTGTA	(COMBES *et al.*, 2000)
10.	CFGA69	F: TGGTGGAGTGGCTTTGATTGATG R: GCAACTTATGAGCCTAATCC	(Moncada and McCouch, 2004)
11.	CFGA100	F: TTGACTCTTTTCTCTCCCAA R: ATTTAGCAGGCTTGGCATTTTT	(Moncada and McCouch, 2004)
12.	CFGA465	F: ACCCTTTACTACTTATTTACTCTC R: ACATCCCCTTGCCATTTCTTC	(Moncada and McCouch, 2004)
13.	CFGA502	F: AAGCCACCCAGAAAACAGCACAT C R: ATTTGCTTCTCATGTTCCCTTTCA	(Moncada and McCouch, 2004)

List of expected size of PCR product including other details and primers are shown in Table 1 and Table 2 respectively. The PCR reaction was conducted in a total volume of 25 µl. The PCR reaction consists of 2 µl DNA, 1 µl of 10µM forward primer, 1 µl of 10µM reverse primer, 0.5 µl Taq DNA polymerase, 0.5 µl dNTP, 2.5 µl 10x buffer, 1.5 µl $MgCl_2$ and 16 µl nuclease free water. The PCR product was loaded in agarose gel. Three different PCR programs were used. The touch down PCR program 94^{o}C 4 min, 9 cycle of 94 ^{o}C 45 s, 1 min 60 ^{o}C to 55 ^{o}C, decreasing by 0.5 ^{o}C and 1 min 30 s at 72 ^{o}C, 72 ^{o}C 1 min 30s and 72 ^{o}C and 5 min was used following (Cubry *et al.*, 2008). 94 ^{o}C 2 min, 5 cycle of 94 ^{o}C 45s, 1 min primer annealing at 60 ^{o}C with decreasing temperature of one degree at each cycle and 1 min 30 s elongation at 72 ^{o}C then 30 cycle of 90 ^{o}C 45s, 55 ^{o}C 1 min and 1 min 30 s at 72 ^{o}C and 72 ^{o}C 8 min was another touch down PCR program (COMBES *et al.*, 2000). A PCR program of 94 ^{o}C 5 min, 30 cycle of 94 ^{o}C 1 min, 55 ^{o}C 2 min and 72 ^{o}C 2 min and 72 ^{o}C 5 min was the third one (Moncada and McCouch,

2004). The PCR product was analyzed in 2% agarose, horizontal agarose gel electrophoresis running in 1xTAE buffer. DNA image was analyzed in a transilluminator. And the PCR product was analyzed in 6% polyacryl amide gel electrophoresis, vertical electrophoresis to separate alleles. The polyacryl amide gel solution, which was prepared from acrylamide:bisacrylamide, distilled water, 10X TBE buffer, 10%APS and TEMED was poured between glass plates. Following the polymerization of the gel, the PCR product was loaded and migrated vertically in a vertical electrophoresis set up for 2½ to 3 hrs in 1X TBE running buffer. The glass plates were disassembled and the arylamide gel containing the PCR product was left in the lower plate and stained in a staining box, which contained distilled water and ethidium bromide. Microsatellite alleles were visualized in a UVP transilluminator.

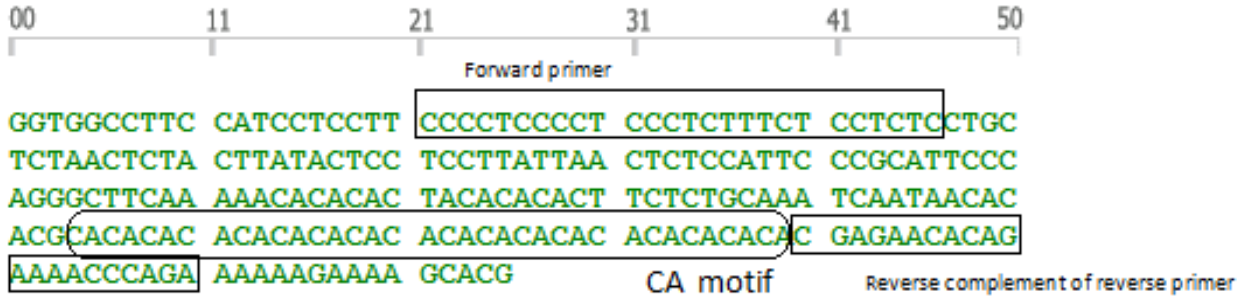

Fig 1. Example of confirmation of primer sequences for accession AJ308755. The forward primer, CCCTCCCTCTTTCTCCTCTC, is directly seen in the SSR sequence while the reverse primer, TCTGGGTTTTCTGTGTTCTCG, is reverse complemented to get it in the SSR sequence. The CA repeat is searched manually. Microsatellite sequence source: - database CoffeeDNA, An open source database on Coffee Genomics. Microsatellite sequence obtained by pasting primer sequence in the SeqMat option of the database, CoffeeDNA.

RESULTS AND DISCUSSION

The success of the PCR experiment was confirmed by the amplification of the PCR product with the expected size and lack of amplifications in the negative control. It is always a primer dimer if any in the negative control. This was examined carefully. The failed and successful PCR experiments are shown in Fig 2 and Fig 4 respectively. The DNA extracted by the CTAB method is shown in Fig 3. This DNA extracted by the CTAB method is used as a template for the successful PCR experiments displayed in Fig 4. In Fig2, no PCR product is detected. The template DNA was extracted based on promega kit. And, this DNA is highly colourful ranging from reddish, yellowish to brownish. The pellet was also colourful. This template DNA is contaminated with secondary compounds, which inhibited PCR. The DNA extracted by the CTAB method is shown in Fig 3. This DNA extracted by the CTAB method is used as a template for the successful PCR experiments displayed in Fig 4. The DNA yield in agarose is relatively poor for the DNA extracted by the CTAB method when compared to the DNA extracted by the promega kit although the DNA concentration is pretty enough for PCR. Three individuals were selected for molecular characterization. One individual is from a parental line of *C.arabica* and two individuals are micropropagated from the parental line by techniques of somatic embryogenesis. The concentration of the genomic DNA of the parental, P_1 and TC, that is tissue culture derived individual 1 and individual 2 and the value of A_{260}/A_{280} was 217.3 ng/µl, 1.7, 47.66 ng/µl 1.71 & 53.07 ng/µl, 1.56. The DNA extracted by the promega kit is shown in Fig 2A. This template DNA shown in Fig 2A resulted in no amplification, which is displayed in Fig 2B. It has been mentioned that primer dimers could be detected in electrophoresis with size of 30 to 50 base pair with different appearance from the band of interest (http://en.wikipedia.org). It is true that cross dimer and self dimer must be considered during primer screening. In few PCR experiments, primer dimer was detected. Example of self and cros dimer is shown in Fig 7. However, dimerization is not occurring all the time. There are exceptional PCR conditions which favour dimerization. For example, primer concentration is mentioned to affect dimerization (http://en.wikipedia.org). The size of the PCR products are obtained as expected. For example, CFGA 91, Fig 4 B, lane 1 and CFGA 92, Fig 4 B, lane 2 are according to the size published in (Moncada and McCouch, 2004). AJ250251 and AJ250253 are also according to the size mentioned in (COMBES *et al.*, 2000). AJ250251, Fig 4 C, lane 1 is a bit greater than AJ250253, Fig C, lane 2. As shown in Fig 4 D, microsatellites in three individuals, from lane 1 to lane 5, lane 7 to lane 11, lane 13 to lane 17 are obtained according to the size published in (Moncada and McCouch, 2004). These microsatellites are CFGA69, CFGA100, CFGA465 and CFGA502 respectively. The PCR product size of AJ308755, AJ308779, AJ308782 and AJ308790 are displayed in Fig 4 A, from lane 1 to lane 4. AJ308755 is as expected. AJ308779 is also as expected. Regarding AJ308782 and AJ308790, the manual calculation shows AJ308790 is a bit

greater than AJ308782. That is not reflected, which could be for reasons of polymorphism. Three alleles of AJ308755 existed as displayed in Fig 4G, lane 1. Also the microsatellite sequence AJ308790 contained three alleles or fragments as displayed in Fig 4G, lane 2. Also, the PCR product of AJ308782 as displayed in Fig 4 F and CFGA69 as displayed in Fig 4 H for the three individuals were analyzed in polyacryl amide gel electrophoresis. Two alleles counted for AJ308782 in all of the three individuals. For CFGA69, three alleles were counted for three individuals. Genetic similarity was computed for locus AJ308782. For locus AJ308782, the similarity between the genotypes is computed using the coefficient of association following (Dice, 1945), that is, $r_d = 2a/2a+b+c$ where "a" stands for number of bands present on both genotype and b and c stands for number of bands unique to each genotype. Thus, r_d = 2 x number of bands in parental line and progeny / 2xnumber of bands in parent and progeny + number of bands in parent only + number of bands in progeny only. This resulted in r_d= 2x4/2x4 + 0 + 0. r_d = 8/8 = 1. The genetic similarity between explant and progeny is 100% revealing that they are identical. Microsatellite "CFGA 69" was published to contain one allele only (Moncada and McCouch, 2004). Here three alleles for CFGA 69 existed. Thus, two additional SSR/Allele CFGA69 of the tetraploid *C.arabica* are identified. This was not published before. Perhaps additional two microsatellite/ alleles of CFGA69 might have acquired in *C.arabica* through hybridization or breeding of *C.arabica* varieties to release a new hybrid coffee "Aba buna". For example, unequal crossing over may contribute to such mechanisms of acquisition of new microsatellite. It was detected that increasing the percentage of agarose to 2% may result in minute separation of SSR fragments while due to its high resolution, polyacryl amide gel electrophoresis seems to applicable to avail all existing fragments of a given microsatellite as exemplified by AJ308755, Fig A. The task of fragment analysis for locus CFGA 69 and AJ308782 showed that the number and size of alleles in the three individuals remained the same. During this protocol optimization, the fact that tissue culture derived individuals and their parent fail to exhibit polymorphism is a preliminary insight in the genetic uniformity of individuals of coffee generated via somatic embryogenesis. And this experimental result, which was not our objective, might indicate that a further experiment using significant number of tissue culture derived individuals will be a forefront task in assessing the genetic uniformity of these individuals and detecting rare mutations which may exist following events of soma clonal variations. However, testing quite many SSR markers and taking quite a number of individuals from the tissue culture derived lines is required to draw conclusions about genetic purity of the explant and the tissue culture derived lines. Infact, a polymorphic marker will be ideal to study somaclonal variation. CFGA69 couldn't be marker of choice to study variations for it is published to be monomorphic. Perhaps AJ308782 could be a helpful marker. This kind of study is no more useful for mutations like substitutions which do not show size difference can't be addressed. Automated DNA sequencing is the novel method to see difference at a level of point mutation. The extracted DNA and the PCR product are displayed in Fig 2A and Fig 2B respectively.

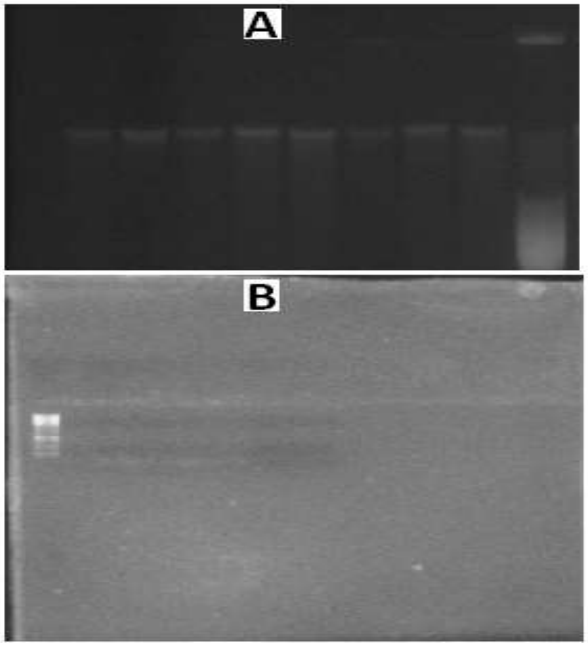

Fig 2 . A. Genomic DNA extracted by promega kit. B. PCR experiment. No PCR product except primer dimers

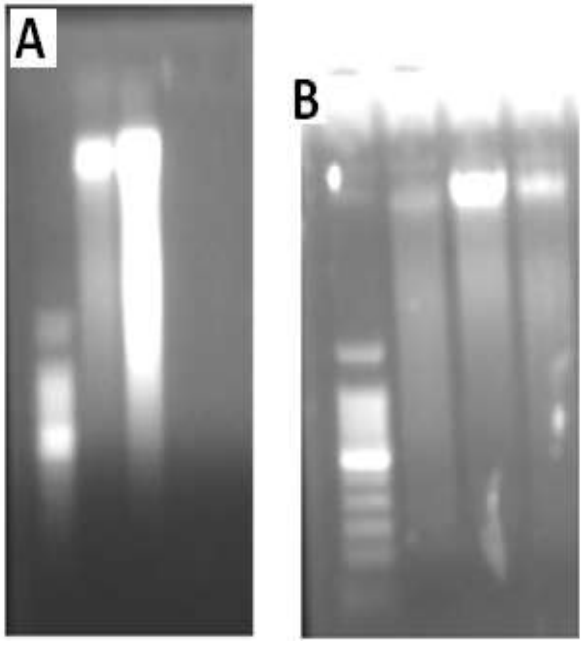

Fig 3. A & B. Genomic DNA extracted by the CTAB method

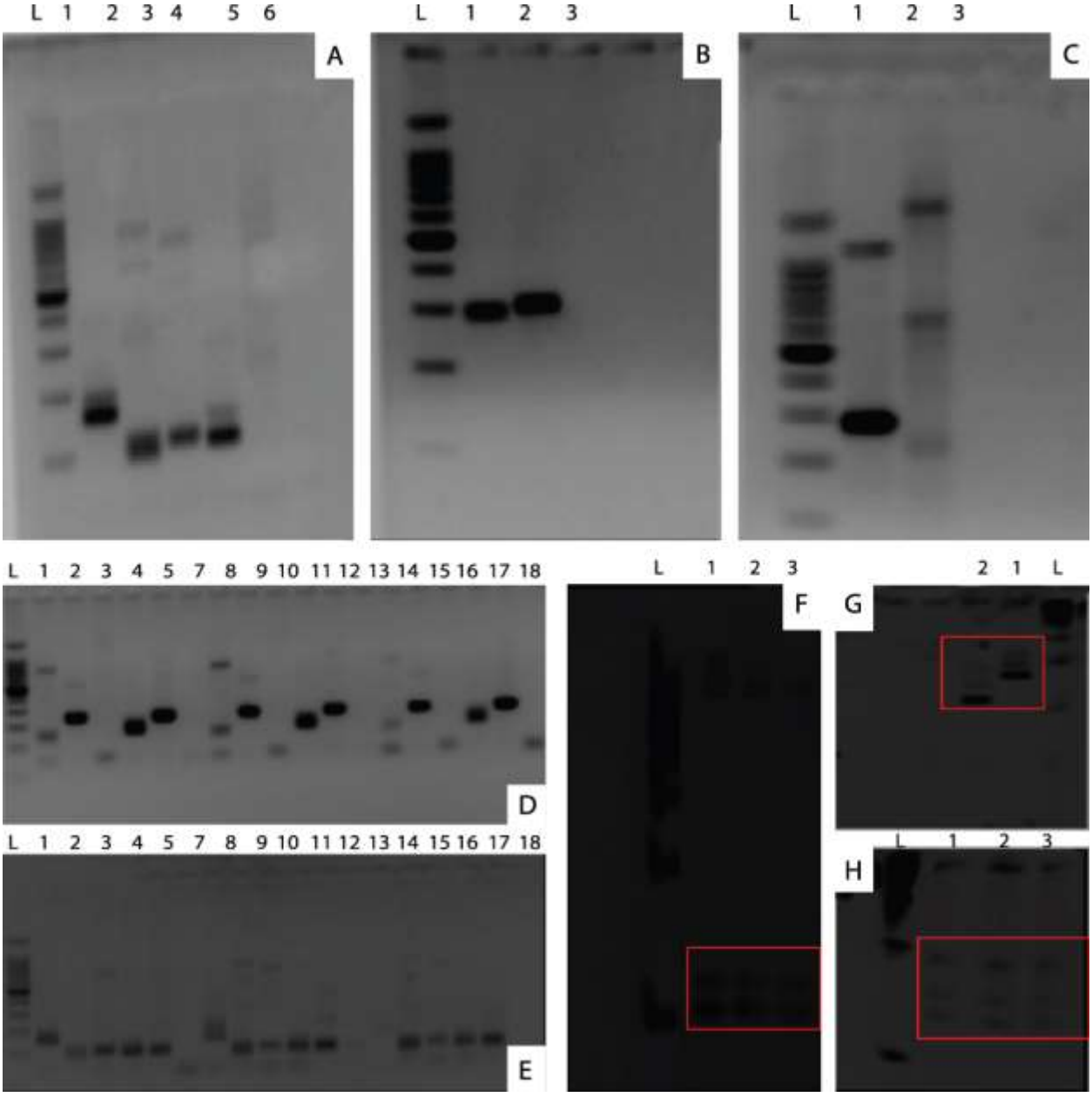

Fig 4. A,B,C D,E – PCR product in agarose gel. F,G,H – Allele separation in polyacryl amide gel. A-L- 100bp ladder,lane 1- AJ308755, Lane 2- AJ308779 , Lane 3- AJ308782, Lane 4- AJ308790,Lane 5-failed pcr reaction, Lane 6-negative control. B-Lane 1- CFGA91, Lane 2- CFGA 92, Lane 3- negative control. C, Lane 1- AJ250251, Lane 2- AJ250253, Lane-3- negative control. D-Lane 1-CFGA69explant, Lane 2-CFGA100 explant, Lane 3-failed pcr reaction, Lane 4-CFGA465explant, Lane 5-CFGA502 explant, Lane 6-negative control, Lane 7- CFGA69 tc-derived individual 1, Lane 8-CFGA100 tc-derived individual 1, Lane 9-failed PCR reaction, Lane 10-CFGA465 tc-derived individual one, Lane 11-CFGA502 tc-derived individual one, Lane 12-negative control, Lane 13-CFGA69 tc-derived individual 2, Lane 14-CFGA100 tc derived individual 2, Lane 15-failed pcr reaction, Lane 16-CFGA465 tc-derived individual 2, Lane 17-CFGA502 tc derived individual 2, Lane 18-negative control. E-Lane 1-AJ308755 explant , Lane 2-AJ308779 explant, Lane 3-AJ308782 explant, Lane 4-AJ308790 explant, Lane 5- explant AJ308837, Lane 6-negative control, Lane 7- AJ308755 tc derived individual 1, Lane 8- AJ308779 tc derived individual 1, Lane 9- AJ308782 tc derived individual 1, Lane 10- AJ308790 tc derived individual 1, Lane 11- AJ308837 tc derived individual 1, Lane 12-negative control , Lane 13- AJ308755 tc derived individual 2, Lane 14- AJ308779 tc derived individual 2, Lane 15- AJ308782 tc derived individual 2, Lane 16- AJ308790 tc derived individual 2, Lane 17-AJ308837 tc derived individual 2, Lane 18-negative control. F-lane 1-AJ308782 explant, lane 2- AJ308782 explant tc derived individual 1, lane 3- explant tc derived individual 2. G-lane 1-AJ308755, lane 2- AJ308790. H-lane 1-CFGA69 explant, lane 2-CFGA69 tc derived

individual 1, lane 3- CFGA69 tc derived individual 1. L stands for 100 bp ladder for all cases. AJ and CFGA are known accessions tags for the microsatellite. The last band of the ladder in all cases is 100 bp. TC stands for tissue culture.

A subordinate task of sequence analysis was performed. The amplicon AJ250251 in Fig 4C was selected for homolog scanning. The sequence of AJ250251 was retrieved from NCBI and its nucleotide sequence was used as a query sequence for BLASTN analysis. There existed four homologs of AJ250251. These homologs are AJ308804, AJ308762, AJ308884 and AJ308761. The overall information of these homologs is shown in Table 3. Multiple sequence alignment of AJ250251, AJ308804, AJ308762, AJ308884 and AJ308761 is indicated in Fig 5. There are conserved regions observed from the alignment of these microsatellites. However, it seems impossible to see a conserved SSR revealing that these microsatellites are homologues due to other sequences, not due to SSR. The depicted phylogeny as shown in Fig 6 indicates that AJ250251 clustered with AJ308761. Biological events like gene duplication might have played a leading role in divergence of these homologs. Perhaps the use of all the homologs of AJ250251 in SSR based characterization of *C.arabica* could be crucial to get consolidated information.

Table 3. BLASTN analysis. AJ250251 and its homologs.

Accession	Clone name	Max score	Total score	Query coverage	E-value	Max ident
AJ250251	ZapII.3	558	558	100%	2e-160	100%
AJ308804	75-5CTG	558	558	100%	2e-160	100%
AJ308762	E6-3CTG	547	547	99%	4e-157	99%
AJ308884	14-7CTG	449	449	100%	1e-127	93%
AJ308761	E4-3CTG	449	449	100%	1e-127	93%

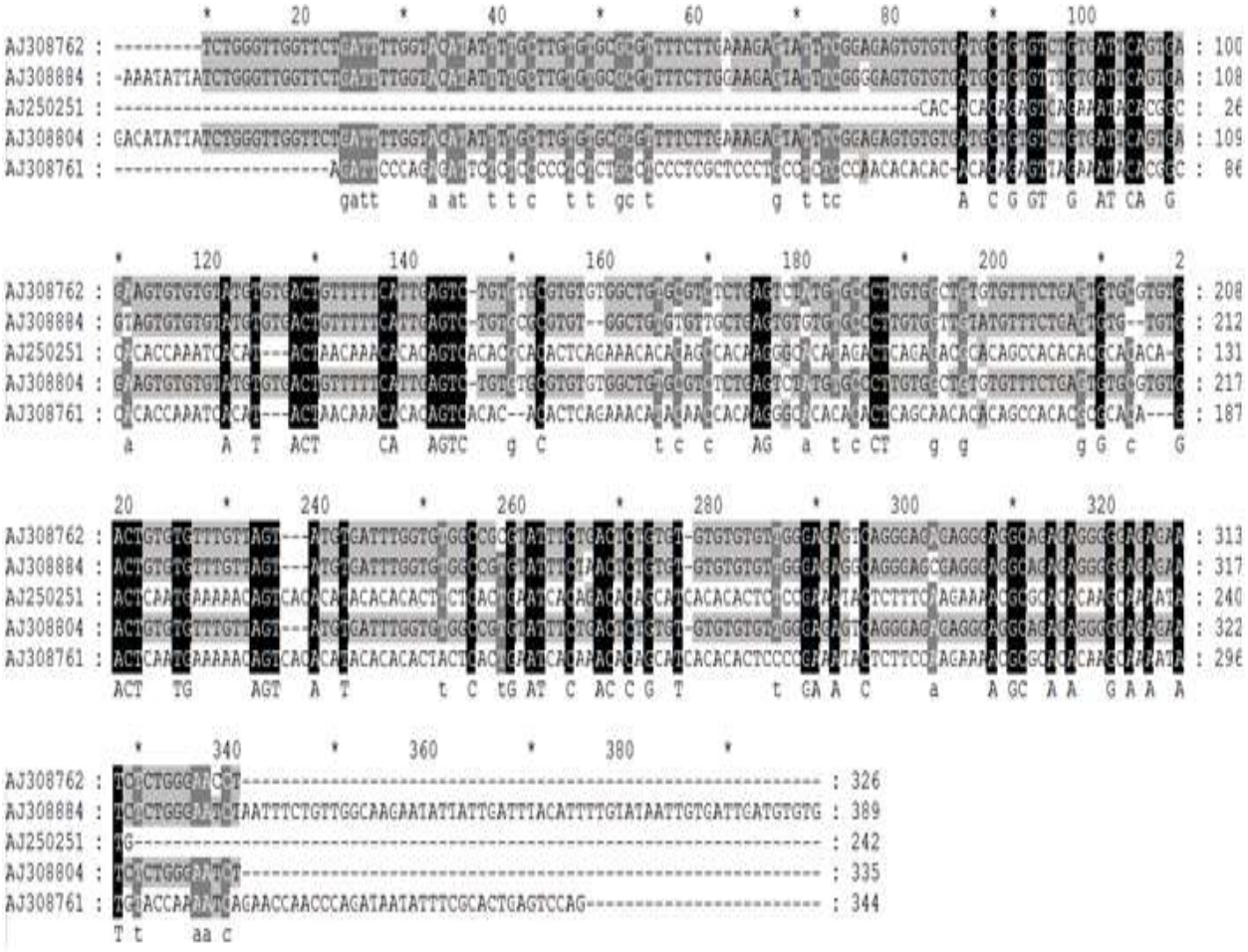

Fig 5. Multiple sequence alignment of homologs of AJ250251

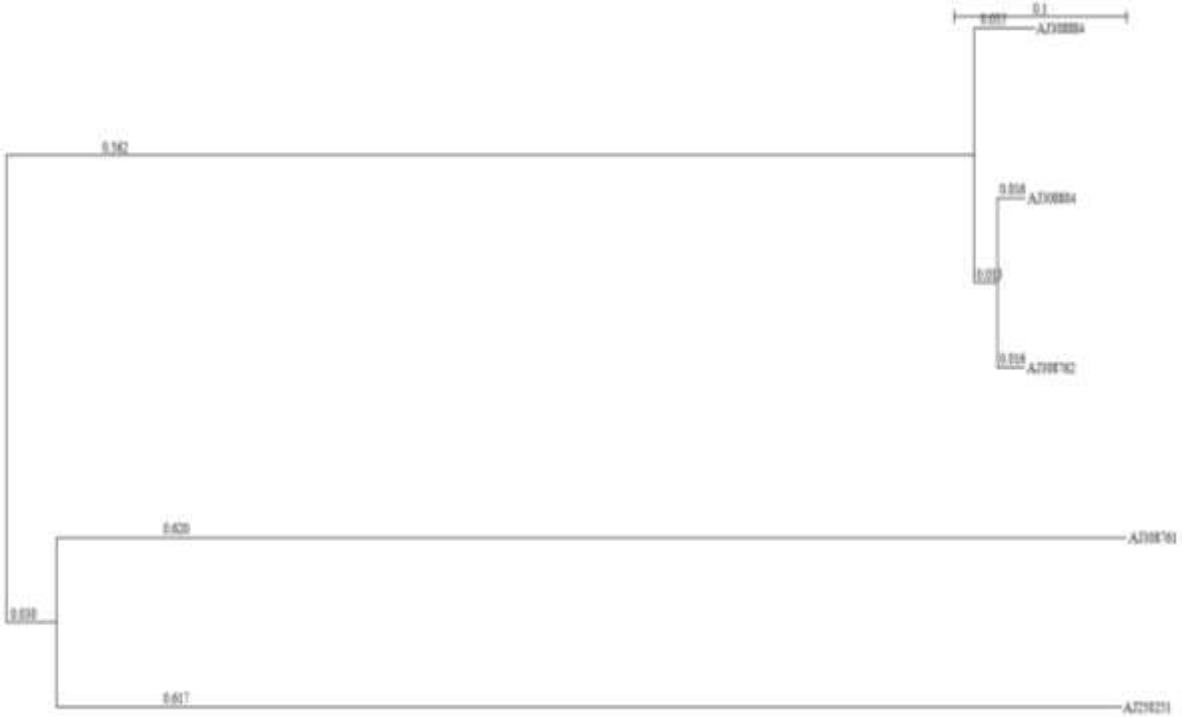

Fig 6. Phylogeny of homologs of AJ250251

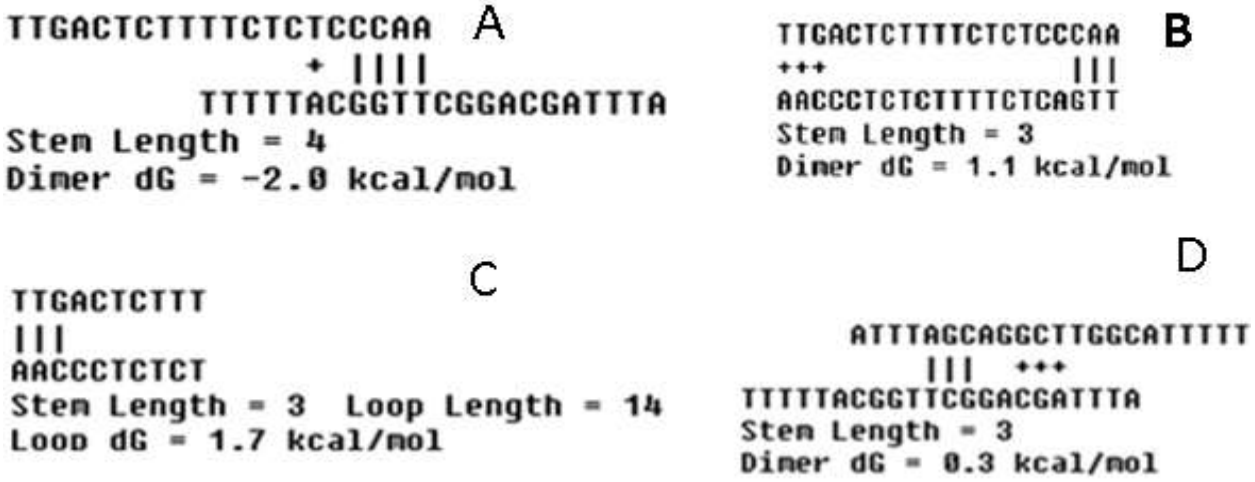

Fig 7. Oligo analysis using one primer pair; CFGA100 F,TTGACTCTTTTCTCTCCCAA and CFGA100R, ATTTAGCAGGCTTGGCATTTTT. A. cross dimer between CFGA100F and CFGA100R. B. Self dimer 100F. C. CFGA100F hairpin loop. D. Self dimer CFGA100R.

CONCLUSION AND RECOMMENDATIONS

SSR of *C. arabica* can be amplified with maximum efficiency in the presence of a pure DNA free of contaminants like secondary metabolites. It is highly recommendable to extract the genomic DNA of *C.arabica* under the consideration of reduction of the secondary compounds that add impurities to the genomic DNA. Indeed the mechanism of PCR inhibition is unclear. Such hypotheses as secondary compounds might inhibit PCR, secondary compounds might have modified the template DNA, secondary compounds might have modified the PCR enzyme and secondary compounds might compete with Taq polymerase for Mg^{++} needs to be tested further. Nevertheless, other unknown reasons might have caused the PCR inhibition. This needs in depth analysis. There is a need to zoom in to the mode of action of the secondary metabolites in the PCR reaction. The investigation of the specific metabolite which acted as a target inhibitor could help to investigate the mechanism of inhibition and identify the candidate inhibitor out of bunch of the compounds. The knowledge of the chemistry of the coffee leaf is a key issue to screen the diverse secondary compounds. Indeed tracing the particular inhibitor and determining its concentration was suggested to be a relevant step in investigating PCR inhibition (Opel *et al.*, 2009). Thus, the inhibitor and its quantity should be determined. A detailed study is required to find out the principal cause of the inhibition. Generally, minimization of secondary metabolites in the DNA extract of *C.arabica* and avoiding other contaminants is crucial for achieving a good PCR yield. It is desirable to extract clean genomic DNA of Coffee with care for successful PCR experiment. Added to this, non specific annealing is detected in some of the PCR reactions. For example, in the PCR reaction, which amplified AJ250251, a PCR product of greater than 1000 bp is detected. And in the PCR reaction which amplified AJ250253, two PCR, one greater than 500 bp and the other greater than 1000 bp is detected. In AJ308779 and AJ308782, a PCR product greater than 500 bp is detected. This calls for sequence analysis of the primers against the coffee genome, which is unavailable. In case the primers have homology regions in other part of the Coffee genome, these kind of PCR experiments may be detected. Alternatively, it is highly recommendable to undertake primer blast against the *C.arabica* genome sequence, which is not yet released to the public. Sometimes, the non target amplicon could be a microsatellite in case the marker occurs in more copies. Annealing temperature of the primers may also cause such problems. Under these complications, since the large fragment is not an expected PCR product of microsatellite, automated DNA sequencing will be the sole option to annotate the unexpected fragment. In case of special requirements, target fragments can be purified from gel and alleles can be sequenced, and the non target fragment can be neglected for it has no value. Repeating the PCR experiments by manipulating the annealing temperature several times may also assist in avoiding non target region amplification. Thus, a detailed work on the sequence analysis part is demanded and for this the complete or partial genome sequence of *C.arabica* is required. Sequencing of the coffee genome is in progress. This might help to resolve the task of sequence analysis in the near future. Polyacryl amide gel

electrophoresis seems to be an absolute method for mcrosatellite allele separation unlike agarose gel electrophoresis, which has less resolution. It is noted that 2 to 3 alleles were counted in the vertical gel. This was not possible in agarose. Such computational work as homolog scanning, motif analysis and SSR pattern study, phylogenetic studies and sequence alignment will have great role towards supporting the experimental work. It appears that the method of DNA extraction is the governing factor to avoid PCR inhibition. The Doyle and Doyle CTAB method seems the right novel method for extracting coffee DNA. It is also good not to neglect the promega kit. But, after extracting the coffee DNA using promega kit, a further step of purifying the secondary compounds from the extracted DNA is required. The eventual goal is to get a pure clean DNA free of secondary compounds.

ACKNOWLEDGMENTS

The authors are grateful for the tissue culture laboratory of the Ethiopian Institute of Agricultural Research (EIAR) for providing coffee leaf samples. The laboratory work is actualized in the agricultural biotechnology laboratory of EIAR. We extend our thanks to EIAR for provision of molecular laboratory and computer facility.

REFERENCES

Albertin, A. and P.K.R. Nair (2004). Farmers' Perspectives on the Role of Shade Trees in Coffee Production Systems: An Assessment from the Nicoya Peninsula, Costa Rica. *Human Ecology* 32 (4):443-463.

Alemayehu, T., C. Dominique, P. Vincent and B. Pier (2010). Genetic diversity of Arabica coffee (*Coffea arabica* L.) collections. *EJAST.,* 1:63-79.

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. *Theor Appl Genet.,* 104:894-900.

Bhattacharjee, R., M. Ferguson, M. Gedil, D. Dumet and I. Ingelbrecht1 (2009). Field collection, preservation and large scale DNA extraction procedures for cassava (*Manihot esculenta* Crantz.). *African Journal of Biotechnology,* 8:3424-3430.

COMBES, M.C., S. Andrzejewski, F. Anthony, B., Bertrand, P. Rov-elli, G. Graziosis and P. Lashermes (2000). Characterization of microsatellite loci in Coffea arabica and related coffee species. *Molecular Ecology,* 9:1171-1193.

Cristancho, M. A. and Á. L. Gaitán (2008). Isolation, characterization and amplification of simple sequence repeat loci in coffee. *Crop Breeding and Applied Biotechnology,* 8:321-329.

Cubry, P., P. Musali, H. Legnate, D. Pot, F.D. Bellis, V. Poncet, F. Anthony, 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.

Davis, A. and P. Stoffelen (2006). An annotated taxonomic conspectus of the genus Coffea (Rubiaceae) *Bot. J. Linn. Soc.,* 152:465 – 512.

Dice, L.R. (1945). Measures of the amount of ecologic association between species. *Ecology,* 26 (3):297-302.

Doyle, J.J. and J. L. Doyle (1990). Isolation of plant DNA from fresh tissue. *Focus,* 12:13-15.

Hatwal, D., R. Bist, K. Pathak, P. Chaturvedi, J. P. Bhatt A.K. and Gaur1 (2011). A simple method for genomic DNA isolation for RAPD analysis from dry leaves of Aconitum balfourii Stapf. (Ranunculaceae) *J. Chem. Pharm. Res.,* 3(3):505-510.

Herrera, C.J., C. Marie.Combes, H. Cortina and P. Lashermes (2004). Factors influencing gene introgression into the allotetraploid Coffea arabica L. from its diploid relatives. *Genome,* 47: 1053 – 1060.

http://en.wikipedia.org. Primer dimer.

http://www.coffeegenome.org. INTERNATIONAL COFFEE GENOME NETWORK (ICGN).

Lodish, H., A. Berk, P. A. Matsudaira, C. Kaiser, M.P. Krieger, M. Scott, S.L. Zipursky and J. Darnell (2004). *Molecular Cell Biology*. 5th Edition. Printed in the United states of America. W.H.Freeman and Company.

Moncada, P. and S. McCouch (2004). Simple sequence repeat diversity in diploid and tetraploid Coffea species. *Genome,* 47: 501 - 509.

Opel, K.L., D. Chung and B.R. McCord (2009). A Study of PCR Inhibition Mechanisms Using Real Time PCR. *J Forensic Sci*. 55(1):25-33.

Pinto-Maglio, C.A.F. (2006). Cytogenetics of coffee. *Braz. J. Plant Physiol.,* 18 (1):37-44.

Ramos, R., T. Wondyifraw, F. Martínez, M. Gonzalez, G. Endale, T. Alemayehu and A. Zerihun (2009). Plant regeneration through somatic embryogenesis in three ethiopian Coffea arabica Lin. hybrids. *Biotecnología Vegetal* 9 (1):19-26.

Roux, K.H. (2006). Optimization and Troubleshooting in PCR. *Cold Spring Harb Protoc*.

Sera, T., P.M. Ruas, C.D.F. Ruas, L.E.C. Diniz, V.D.P. Carvalho, L. Rampim, E.A. Ruas and S.R.D.Silveira (2003). Genetic polymorphism among 14 elite Coffea arabica L. cultivars using RAPD markers associated with restriction digestion. *Genetics and Molecular Biology*, 26: 59-64.

P.H Moore C.Nagai, C.W Morden, R.V Osgood and R.Ming (2002) AFLP analysis of genetic diversity within and among Coffea arabica cultivar coffee beans. *Theor Appl Genet* 105: 209–215.

Tesfaye, K., T. Borsch, K. Govers and E. Bekele (2007). Characterization of Coffea chloroplast microsatellites and evidence for the recent divergence of C.arabica and C.eugenioides chloroplast genome. *Genome* 50:1112-1129.

www.promega.com. Technical manual. Wizard Genomic DNA purification kit. Isolating genomic DNA from plant tissue.

(*Accepted for publication March 2013*)