Inter-simple sequence repeat (ISSR) variation in forest coffee trees (*Coffea arabica* L.) populations from Ethiopia

Esayas Aga¹,² Endashaw Bekele² & Tomas Bryngelsson¹

¹Department of Crop Science, Swedish University of Agricultural Sciences, Box 44, 230 53, Alnarp, Sweden (Fax: +40-415519; E-mail: esayas.agaa@vv.slu.se); ²Department of Biology, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia

Received 1 November 2004 Accepted 28 January 2005

**Key words:** *Coffea arabica* L., Ethiopia, forest coffee trees, gene diversity, ISSR, Shannon’s index

**Abstract**

Genetic variation of forest coffee trees (*Coffea arabica* L.) from four regions of Ethiopia was investigated using inter-simple sequence repeat (ISSR) markers. A total of 160 individuals representing 16 populations were sampled. Eleven ISSR primers amplified a total of 123 fragments of which 31 fragments (25%) were polymorphic. Estimate of total gene diversity (*H*<sub>T</sub>), and the coefficient of genetic differentiation (*G*<sub>ST</sub>) were 0.37 and 0.81, respectively. This indicates that most of the variability is between populations than within populations. The partitioning of genetic variation into within and between populations based on Shannon’s information index also revealed more differentiation between populations (0.80) than within populations (0.20). In the phenogram most of the coffee tree samples were clustered on the basis of their regions of origin but failed to cluster according to their respective populations, which could be attributed to the presence of substantial gene flow between adjacent populations in each region assisted by man in the process of transplantation or by wild animals such as monkeys, which eat the berries and defecate the seeds elsewhere. On the other hand, the inter-regional clustering of some coffee tree samples from Bale and Jimma regions could be due to the transport of coffee seeds across regions and their subsequent planting. Although ISSR markers detected lower polymorphic loci than previously reported results with random amplified polymorphic DNA (RAPD) markers on the same materials, it can be used as an alternative method for molecular characterization of *C. arabica* populations. The results may provide information to select sites for *in situ* conservation.

**Introduction**

Coffee trees (family Rubiaceae) are classified in two genera, *Coffea* and *Psilanthus*. Particular attention has been given to the genus *Coffea*, which includes two cultivated species of economic importance, *C. arabica* L. and *C. canephora* Pierre. *Coffea arabica* (*2n = 4x = 44*) is an amphidiploid and self-fertile while other *Coffea* species are diploid (*2n = 2x = 22*), and self-sterile except for *C. heterocalyx* and C. sp. Moloundou, which are identified as being self-fertile (Lashermes et al., 1999; Coulibaly et al., 2002). *Coffea arabica* has its origin and primary center of diversity in the southwestern highlands of Ethiopia (Sylvain, 1955).

Assessment of the genetic variability within and between forest *C. arabica* populations in Ethiopia has important consequences in coffee trees breeding and the conservation of coffee trees genetic resources. Forest or wild coffee tree refers to coffee trees growing naturally in the forest or young wild (forest) coffee plants, which have been irregularly transplanted in the forest (Van der Graaff, 1981).
Molecular markers have been replacing or complementing traditional morphological and agronomic characterization, since they are virtually unlimited, cover the whole genome, are not influenced by the environment, and less time consuming.

Markers such as inter-simple sequence repeat (ISSR) (Zietkiewicz, Rafalski & Labuda, 1994) are widely used in genetic diversity studies because they need no prior DNA sequence information, development costs are low, and laboratory procedures can easily be transferred to any plant species (Barth, Melchinger & Lubberstedt, 2002). The commonly used polymerase chain reaction (PCR) based DNA marker systems are random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and more recently simple sequence repeats (SSRs) or microsatellites (Gupta & Varshney, 2000). The major limitations of these methods are low reproducibility of RAPD, high cost of AFLP and the need to know the flanking sequences to develop species specific primers for SSR polymorphism (Reddy, Sarla & Siddiq, 2002). ISSR is a technique that overcomes most of these limitations (Gupta et al., 1994; Wu et al., 1994; Zietkiewicz, Rafalski & Labuda, 1994).

The ISSR primer is composed of microsatellite sequences either unanchored (Gupta et al., 1994; Wu et al., 1994) or anchored at the 5' or 3' end by two or four arbitrary, often degenerate nucleotides (Zietkiewicz, Rafalski & Labuda, 1994; Fang et al., 1997). The addition of a different base at the 5' or 3' end renders their binding sites more specific and reproducible (Barth, Melchinger & Lubberstedt, 2002). The sequence between the two binding sites in opposite orientation within suitable distance is amplified, and indels within this region and loss or gain of binding sites are detected as band polymorphism (Yang et al., 1996). For 5' anchored primers, polymorphisms also occur due to the length of the microsatellite (Goulão & Oliveira, 2001). ISSR markers have proven valuable for fingerprinting studies and genetic diversity investigations in Tef (Assefa, Merker & Tefera, 2003), rice (Blair, Panaud & McCouch, 1999; Joshi et al., 2000; Virk et al., 2003), citrus (Fang & Roose, 1997), wheat (Nagaoka & Ogihara, 1997), oilseed (Charter et al., 1996), finger millet (Salimath et al., 1995) and maize (Kantety et al., 1995), but so far not in coffee trees. The objective of the present study was to investigate the possibility of using ISSR markers for genetic diversity analysis of (*C. arabica* L.).

**Materials and methods**

**Plant material**

One hundred and sixty samples representing 16 populations of forest coffee trees (*C. arabica* L.) collected from major coffee growing regions of Ethiopia (Welega, Ilubabor, Jimma and Bale) were used in this study (Table 1 and Figure 1). DNA was extracted from lyophilized leaves following the hexadecyl trimethyl ammonium bromide (CTAB) method (Wang et al., 1996).

**PCR amplification**

A total of 15 primers were tested for ISSR amplification in the *Coffea arabica* genome. The primers were from the Biotechnology Laboratory, University of British Colombia (UBC). Eleven primers, which showed clear and reproducible banding pattern were selected. The 11 ISSR primers include eight di-nucleotide repeats (six were anchored at 3' and two were anchored at 5' ends), two tri-nucleotide, and one penta-nucleotide repeats (Table 2). A single primer concentration of 0.2 μM was used in each PCR reaction, which was carried out in a total volume of 25 μl containing 5–10 ng of genomic DNA, 1 × PCR reaction buffer without MgCl₂ (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 2.5 mM of MgCl₂, 0.2 mM of each dNTPs, 2% formamide, 1 unit of Sigma D-4545 Taq DNA polymerase.

Amplification was performed in 96-well plates on a GeneAMP® PCR System 9700 version 3.01 thermocycler (PE Biosystems, California, USA) under the following conditions: a hot start of 94°C for 1 min; followed by 29 cycles of 1 min at 94°C, with 22% ramp rate to 2 min at 55°C, with 57% ramp rate to 2 min at 72°C, and with 57% ramp rate to the next cycle. The last cycle was followed by an additional 5 min of product extension at 72°C. Equal volume of amplification