ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

COLLEGE OF NATURAL SCIENCES

MICROBIAL CELLULAR AND MOLECULAR BIOLOGY PROGRAM UNIT

Genetic Diversity of Wild Coffee (*Coffea arabica* L.) Populations from Southwestern Ethiopia as Revealed by ISSR Markers

A Thesis Submitted to the School of Graduate Studies, Addis Ababa University, in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biology (Applied Genetics)

By

BIRUK GETENET TEGEGNE

JUNE 2012

Dedication

To my beloved Grandfather and Grandmother HAILU NIGATU and WEREKA KEBEDE respectively, with gratitude

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Kassahun Tesfaye, not only for accepting me as his student, good suggestion on the manuscript, his guidance and also providing me the initial motivation and encouragement but also for facilitating laboratory materials and chemicals, that led to the realization of this work.

I also would like to thank Dr. Tileye Feyessa for kindly allowing me to use materials in Plant Propagation Laboratory. I also thank Ato Abush Zinawe and Ato Shemeket Tadele for their support during my laboratory work.

I am also grateful to the Faculty of Life Sciences of Addis Ababa University for giving me the chance for this study. My special thanks go to the Ministry of Education for allowing me to take study leave with salary.

I, certainly, must acknowledge my families for their care, support and for all what they did for me. Especially, my aunt, Aselefech Gebera who gave up her life and faces the ups and down until this date. May God Bless You All!!

Biruk Getenete Tegegne

Table of Contents

Contents……...Pages

LIST OF TABLES…….Pages

LIST OF FIGURES…….Pages

LIST OF APPENDICES…….Pages

LIST OF ACRONYMS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of molecular variance
CTAB	Cetyltrimethyl Ammonium Bromide
EDTA	Ethylene Diamine Tetra Acetic acid
FAO	Food and Agriculture Organization
GD	Gene diversity
ISSR	Inter Simple Sequence Repeats
NJ	Neighbor Joining
NPL	Number of polymorphic loci
PCO	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
PP	Percent Polymorphism
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SNNPR	Southern Nation Nationalities and People Region
SSR	Simple Sequence Repeats
UPGMA	The Un-weighted Pair Group Method with Arithmetic mean

Abstract

Coffea arabica *L., is indigenous to the highland of southwestern Ethiopia and southeastern Sudan. It is the afromontane rain forest species of south west and south east part of Ethiopia and grown naturally at altitudes ranging from 1,300 and 1,800 meter above sea level. Human population pressure and land use change in and around the afromontane rain forest highly threatened the wild coffee populations of Ethiopia. Inter simple sequence repeats (ISSRs) molecular markers were analysed. In the present study five wild populations of* ***Coffea arabica*** *from, Bench Maji-1(Maji), Bench Maji-2(Jeba), Dawero Banja, Limmu and Anfilo were included. A total of 71 individual plants were analyzed using four ISSR primers, three dinucleotides (810, 812,818) and one penta nucleotides (880) primers. A total of 30 clear and reproducible bands were amplified, and 100% were found to be polymorphic. The highest polymorphic loci were revealed by primer 812 with 9 polymorphic loci. Among the populations studied, Bench Maji-1 and Dawero showed 100% polymorphism, while Limmu and Anfilo showed moderate polymorphism of 56.67% and 63.33%, respectively, whereas Bench Maji-2showed the least polymorphism(30%). The average gene diversity relative to the overall population was 0.45. AMOVA without grouping revealed that higher percentage of variation (79.4%) is attributed to the within population variation while the remaining variation is due to the among population variation (20.6%). Forest based UPGMA analysis based on Jaccard's coefficient revealed one major group with Limu and Anfilo and two outliers (Bench Maji and Dawero Banja). NJ analysis based on individuals of wild* ***Coffea arabica*** *showed two distinct major clusters and sub-clusters within both major clusters; and individuals from Bench Maji-1 and Anfilo tend to form their own separate groups while others tends to speared all over the tree. In 3D PCO, most of Dawero Banja and Bench Maji individuals were intermixed with each other and made separate group from the other two populations but individual of Anfilo and Limmu were spread all over the plot.*

Key words: Coffee, *Coffea arabica,* genetic markers, genetic diversity and ISSR

1. INTRODUCTION

Coffee species belong to the family *Rubiaceae*, one of the largest tropical angiosperm families. Variations in chloroplast DNA (cpDNA) classified the *Coffeeae* tribe into the *Ixoroideae* monophyletic subfamily, close to *Gardenieae, Pavetteae* and *Vanguerieae* (Bremer *et al.*, 1991).Two genera, *Coffea* L. and *Psilanthus* were distinguished on the basis of flowering and flower criteria (Bridson, 1982). Each genus was divided into two subgenera based on growth habit (monopodial *vs* sympodial development) and type of inflorescence (axillary *vs* terminal flowers). More than one hundred coffee species have been identified and new taxa are still being discovered (Bridson *et al.,* 1988; Davis *et al*., 2006). Davis *et al*., (2006) recently enumerated 103, *Coffea* species naturally occurring in main land Africa, Madagascar and Mascarenes Island.

Coffee provides one of the most widely drunk beverages in the world, and is a very important source of foreign exchange for many countries. The genus *Coffea L.* contains the three species used in the production of the beverage coffee: C.*arabica* (Arabica coffee), *C. canephora* (Robusta coffee); and *C.liberica* (Liberian or *Liberica* coffee, or excels coffee). Of the three, *C.arabica* is the most important commercial species (Coste, 1992; Purseglove, 1968).

C.arabica is the only allopolyploid (2n = 4x = 44) coffee species and self-fertile at approximately above 95 % (Lashermes *et al.*, 2000; Silvarolla *et al*., 2004). The other *Coffea* species are diploid (2n = 2x = 22) and self-sterile, except for C. *heterocalyx* and C. sp. Moloundou, which are diploid but self-fertile (Coulibaly *et al*., 2002).

C. arabica could have arisen from natural hybridization between two ancestral diploid coffee species followed by unreduced gamete formation. *C. eugenioides* and *C. canephora* (*C. congenesis*) have often been assumed to be the ancestral parents of *C. arabica* (Tesfaye *et al*, 2007). Species very closely related or identical to *C. eugenioides* and to *C. canephora* (*C. congensis*) are indeed the most likely maternal and paternal progenitors of *C. arabica*, respectively (Lashermes *et al*., 1995). Recently, Tesfaye *et al*., (2007), based on more than 7kb cp DNA sequence and non-coding chloroplast genome regions strongly suggested C. *eugenioids* or its ancestor as maternal parent. Raina *et al*. (1998) also showed similar conclusions in a cytogenetic study of *C.arabica* using genomic fluorescence *in situ* hybridization techniques.

Arabica coffee plant is an evergreen shrub or small tree and it is 4-6m tall under free growth. In cultivation, arabica coffee is pruned to manageable heights of less than 2-3m with one or more stems. The growth of the coffee plant is dimorphic. The main stems (orthotropic axes) grow vertically and the branches (plagiotropic axes) grow horizontally. Horticultural propagation is relatively easy. The plant may flower once or twice a year after a rainfall of at least 10mm, which follows a period of water stress. Fruits mature in 7-9 months from date of flowering depending on the variety and environment. The coffee seeds do not behave in an orthodox manner when dehydrated or stored at low temperature. For instance, the viability of seeds of *C. arabica* decreases rapidly after 4-6 months at ambient temperature. However, short term storage (up to 3 years) is possible using adapted and controlled storage conditions (Coste, 1992; Eira *et al.*, 2006).

The coffee forest of Ethiopia lie between latitude 6° and 9° N, and longitude 34° and 40° E within Afromontane rainforest in southwest and southeast. The majority of the forest cover of Ethiopia has disappeared due to timber cutting for construction purposes and increased land requirement for cultivation of food crops which threatened the gene pool of coffee. At present a fragment of mountain rain forests is limited mostly to the southwest (Welega, Ilubabor, Jimma, Kafa and Bench Maji) and southeast (Bale) administrative zones. In altitude, these rainforest ranges from approximately 1200 to 2100 meters above sea level (Gole, 2003).

C. arabica is characterized by low genetic diversity attributed to its narrow genetic base, allopolyploidy, recent evolution and the predominantly self-pollinating nature of the species (Lashermes *et al*., 2000). However, a higher level of genetic variability with molecular markers was observed among spontaneous and subspontaneous accessions collected from Ethiopia (Lashermes *et al*., 1996a; Anthony *et al.*, 2001). The existence of two subgroups of partial genetic differentiation within germplasm of *C.arabica* into accessions collected from west (Kaffa, Ilubabor, Welega) and east (Sidamo, Harrarge) of Great Rift Valley was established by an analysis with RAPD molecular markers (Lashermes *et al*., 1996a) and also by a multivariate analysis of phenotypic characters (Montagnon *et al.,* 1996). Coffee cultivated in Yemen, from where almost all cultivated *C. arabica* derive, had its origin in Ethiopia (Van der Vossen, 2001). Also, recent analysis by Tesfaye (2006), by ISSR marker indicate the existence of high diversity within and among forest coffee populations in Ethiopia. However, not all forest coffee

populations of Ethiopia have so far been investigated for their genetic diversity. More specifically, forest coffee in Dawro, Limmu, Anfilo either not included in the previous studies or very limited samples were evaluated for genetic variability and population structure. This research is, therefore, designed to carry out diversity analysis with higher or denser sampling and also to study the genetic diversity of wild *Coffee arabica* populations in Dawro, Anfilo, Bench Maji-1, Bench Maji 2 and Limu using ISSR molecular marker.

2. LITERATURE REVIEW

2.1. Taxonomic Classification

Based on morphological data, previously Berthaud *et al.,* (1988) deduced that, coffee trees, a tropical woody plant of the *Rubiaceae* family, are classified into two genera. These are *Coffea* L. and *Psilanthus*. The genus *Psilanthus* includes two subgenera, *Paracoffea* and *Psilanthus*, and the genus *Coffea* L. also includes two sub-genera, *Baracoffea* and *Coffea*. The genus *Coffea* L. differs greatly in phenotypic features (size, adaptation habits etc) and thus its taxonomic history was very debating (Lashremes *et al*., 1997). Although the taxonomy is debating and confusing, different authors classify the genus *Coffea* in terms of genera and sections. For the first time, Linnaeus (1737) classified coffee tree as a separate genus *Coffea* with the only one known species of *Coffea arabica* collected from Arabian Peninsula. In addition, Chvalier (1947) classified the *Coffea* into four sections; *Eucoffea, Argofea, Masarocoffea* and *Paracoffea*. The first three genera of the *Coffea* are exclusively native to Africa, Madagascar and adjacent islands. On the other hand, most representatives of Paracoffea are indigenous to India, Malaysia, Ceylon and Southeast Asia. The *Eucoffea* is divided into five subsections; *Erythrocoffea, Nanocoffea Pachycoffea, Melanocoffea* and *Mozambicoffea* (FAO, 1968).

More recently, however, combination of morphological and molecular data set revealed that Rubiceae (*Coffeaeae*) is enlarged to include eleven genera. These are; *Argocoffeopsis, Belanophora, Calycosiphonia, Coffea L., Diplospora, Discospermum,Nostolachma, Psilanthus, Tricalysia, Sericanthe,* and *Xantonnea* (Davis *et al*., 2007;Davis *et al*., 2006). Among these genera, the species of the genus *Coffea* are economically the most important. Out of the 103 species of the genus *Coffea* currently enumerated by Davis *et al*. (2006), only three species namely, *Coffea arabica, C. liberica* and *C. canephora,* which belong to the subsection *Erythrocoffea* are economically important. Out of the three species, *Coffea arabica* is by far the most important commercial species and is one of the world's most important commodities (Coste 1992; Purseglove J. 1968).

2.2. Allopolyploidy origin of *C. Arabica*

The basic chromosome number for the genus *Coffea* is x = 11. Arabica coffee is the only polyploid and self-fertile (over 95%) species of the genus *Coffea,* with chromosome number 2n = 4x = 44, while others are diploid (2n = 2x = 22) and self- infertile (Lashermes *et al*., 2000; Silvarolla *et al*., 2004). Based on the analysis of a hybrid between *C. liberica* and *C. eugenioides*, Narasimhaswamy (1962) suggested these two species as the probable ancestors of *C. arabica*. Studies by Raina *et al*. (1998) based on genomic insitu hybridization (GISH) and fluorescent insitu hybridization (FISH) suggssted the allopolyploid nature of *C. arabica* and stated that *C. eugenioides* and *C. congensis* are the diploid progenitors of *C. arabica.* In particular, genomic in situ hybridization (GISH) was successfully applied for characterization of genomes and chromosomes in polyploid, hybrid plants, and recombinant breeding lines. The genome organization of *C. arabica* was confirmed by GISH using simultaneously labeled total genomic DNA from the two putative genome donor species as probes. On the other hand, Lashermes *et al*. (1999) proposed *C. eugenioides* and *C. canephora* or other ecotypes as possible ancestor of arabica coffee. However, meiotic pairing of chromosomes of *C. eugenioides* and *C. canephora* in interspecific hybrids was observed to be better than in dihaploid plants of *C. Arabica* (Raina *et al*., 1998). Species very closely related to *C. eugenioides* and to *C. canephora* (or *C. congensis*) are indeed the most likely maternal and paternal progenitors of *C. arabica*, respectively. The more recent analysis of a combined data set of non-coding chloroplast genome regions and *mat. K* gene sequencing by (Tesfaye *et al*., 2007) suggested that the cp genomes of C.*arabica* and C. *eugenioids* have diverged recently, clearly pointing to an ancestor of C. *eugenioids* as maternal parent of C. *arabica.*

2.3. History and origin of *Coffea arabica*

Origin of coffee was obtained from the fact that the knowledge of beverage and tree was described from the materials from southern Arabia (Yemen) to which Linnaeus gave a scientific name and Yemen was the only source of coffee germplasm over most of the recorded history of *Coffea arabica* (Sylvia, 1958). *Coffea arabica* was introduced to Yemen by Arab traders from Ethiopia across the Red Sea around the 6th century AD in the form of beans (Gole *et al*., 2002). Coffee was first thought to be originated from Yemen on the Arabian Peninsula when it was seen

grown there by Europeans at much later date. Ethiopian cultural ceremonies and rituals were using the beans in early periods of domestication as a stimulant and a special solid food, for instance, the ripe berries were squashed, combined with animal fats and shaped in to balls, which can be carried and eaten during the long journey since the time immemorial by Oromo people (Teketay, 1999).

However, on the basis of botanical evidence, *Coffea arabica* may have originated on the plateaus of South western Ethiopia from where it spreads to Yemen and then around the world. *Coffea arabica* is endemic to the afro-montane rain forest of Ethiopia where wild coffee populations still grow in the highlands of southwest and south east parts (Gole 2003; Senbeta *et al*. 2006). Senbeta *et al*. (2006) reported that, within small area, the wild coffee plants of Ethiopia have relatively high abundance and variability as compared to the wild coffee populations from Yemen that showed a characteristically low genetic diversity in Arabian area. Gole (2003) reported that the presence of high genetic diversity of coffee in Ethiopia is attributed to the presence of indigenous traditional production system of coffee in the country. Moreover, the existence of high genetic diversity of coffee plants is due to Ethiopia's suitable altitude, ample rain fall, optimum temperature and planting materials.

2.4. The spread of coffee from Ethiopia

The history of introduction of *Coffea arabica* from Ethiopian highlands to Yemen is still linked with little evidence. However, Steiger *et al*. (2002) reported that the first migration of coffee from Ethiopia to Yemen was during the period of prehistoric trade. It has been estimated that Arabica coffee plants could have been introduced to Yemen at about 575AD and again at about 890 AD by Persian armies that, invaded the eastern part of Ethiopia (Steiger *et al*., 2002).

Although the history of introduction of *Coffea arabica* to Yemen is still not clear, its introduction from Yemen throughout the other current coffee producing countries of the world is quite well documented (Meyer *et al*., 1965). From Yemen some coffee plants traveled towards the East to Malabar coasts of India by Boba Budan in 1706 from where they were brought to Java where they flourished and established a dependable crop in 17th century (Steiger *et al*., 2002). Then coffee plant was taken from Java to the Botanical Garden of Amsterdam (Netherlands) in 1706, whose vigorous progenies (seedling from one mother tree) were sent to Paris in 1718 from which

Coffea arabica var. typica was obtained and distributed to Asia, then to Europe and South America, parts of Africa etc. The other *Coffea arabica* sources were introduced into Burboun Islands (now Reunion) by the French at about 1715 and 1718 where it was planted and produced small seed beans yielding a different variety of Arabica coffee type of the world such as Burboun, which reached the New World nearly a century later and are the progenitor of Brazils and Mexico's coffee. In 1893, the coffee from Brazil was introduced into Kenya and Tanzania, not far from its place of origin of Ethiopia. The spreads of coffee around the world was based on the limited number of trees. Originating from the limited number of plants along with its self-pollinating nature left the world coffee with narrow genetic diversity (Steger *et al*., 2002; Gole *et al*., 2002; Chaparro *et al.*, 2004).

2.5. Distribution of *Coffea arabica* in Ethiopia

According to Gole (2003) and Senbeta (2006), *Coffea arabica* is the afromontane rain forest species of south west and south east part of Ethiopian highlands being grown naturally in the forest in diverse environmental factors such as various altitudes ranging from 1,300 and 1,800 meter above sea level. In addition to this, Senbeta (2006) reported that an altitude ranging from 1300 to 1600 meter above sea level contains the highest density of coffee. It also grows within the annual rainfall of the country that varies from 1000 to 2400 mm and a wide range of soil types (from acidic to slightly acidic with low availability of phosphorous) where its fertility is maintained by organic recycling. Cool, shady environment of the forest of Ethiopian highlands were more comfortable for the growth of *C. arabica*. The temperature requirements for *Coffea arabica* is considered 15-25^{0}C which prevails in most of the coffee growing areas of the country (Willson, 1985; Dubale, 1996; Teketay, 1999; Dubale *et al.,* 2000; Gole *et al.,* 2001).

Generally, due to these suitable environmental conditions for coffee growth naturally in the highland forests, makes Ethiopia as a centre of origin and diversification of *C. arabica.* These, in turn are important for good production and are considered to be the best inherent quality that also attributed to high coffee production potential of the country (Gole *et al.,* 2001). The centre of origin of *C. arabica* is geographically separated from all other centre of origin of other species of coffee. Distribution of wild Arabica coffee is on the opposite sides of the Great Rift Valley that is west and east and east of the Great Rift Valley (Anthony *et al*., 2001).

Recently, Gole (2003) reported the southwestern forests constitute the main ecoregion of the species. Specifically, it occurs in the Illubabor, Jimma, West Wellega and Bale zones of Oromia Regional State, and in the Kaffa, Sheka, Bench and Maji zones of the Southern Nations Nationalities and People (SNNP) Regional State. Localities known for their high abundance of coffee trees, and hence important for conservation and sustainable use of wild populations of the species, are the Amora-Gedel, Berhane-Kontir, Boginda-Yeba, Dawo-Tobi, Geba-Dogi (Yayu), Harenna (Bale mountains), Maji, and Mankera forests. Except the Harenna forest, which is located east of the Great Rift Valley, all areas are located on the southwestern plateau. In its natural range, coffee is one of the dominant plant species in the shrub or small tree layer (Dubale *et al.*, 2000; Gole 2002; Senbeta 2006).

2.6. Ethiopian coffee production system

There are varieties of coffee production systems found in Ethiopia. In its area of origin such as southwest part of Ethiopia coffee is still collected in natural forests. These coffee forests form an important part of the remnant forest blocks found in Ethiopia. In several highland areas, the production of forest coffee is a major household activity, supplying between 20 and 50% of all household income; this livelihood activity is supplemented by subsistence agricultural production and collection of other forest products (Gole,2003).

The presence of wild coffee and its economic importance for the local community has contributed to the conservation of the coffee genetic resources and the forest biodiversity. Nonetheless, due to a variety of reasons, including immigration and opening up of lands for commercial estates, the conservation of these forests is not assured. In different areas, the coffee forests are considerably fragmented and damaged. In order to assure that the wild coffee genetic resources are not lost, recently three coffee gene reserves have been established; these will be managed through a community-based participatory forest management approach (Gatzweiler, 2005). Moreover, the recent establishment of Kaffa and Yayu Biosphere reserve assures the sustainable conservation and use of wild coffee populations at two coffee regions; for example in Kaffa zone the percentage of forest cover in the wild coffee producing municipalities (Kebeles) that provide coffee to the Kaffa Forest Coffee Farmers Cooperative Union ranges from 28 % to 80 %, with an average of 59%.

Due to the economic importance of coffee, the coffee forests are often consciously managed to stimulate coffee production through activities such as thinning of over storey trees, removal of ground vegetation, stimulating coppice shoots and/or transplanting of naturally regenerated or artificially raised coffee seedlings. Depending on intensity of management, level of coffee domestication, and diversity of shade trees and other plants the production systems of *Coffea arabica* in Ethiopia are mainly grouped in to four. These are; forest coffee, semi-forest coffee, garden coffee and plantation. The forest coffee also called wild coffee, is a self-sown and grown in natural forest. In this system, coffee is produced in its natural habitat by collecting coffee berries where the forest biodiversity is maintained. In forest coffee production system, there is no human intervention for coffee management to improve coffee production and regeneration in its natural forest (Senbeta *et al.*, 2006; Gole *et al*., 2002).

Forest coffee accounts for about 10% of Ethiopian total coffee lands and 5-6% of the total coffee production of the country (Gole *et al*., 2002; Teketay *et al.*, 1998). The semi-forest coffee production system also called semi-wild coffee. In this production system coffee berries are also picked and collected from naturally grown coffee trees. However, there is human intervention in the coffee management by clearing competing under story trees and shrubs regularly in order to improve production of the wild coffee. It is the main production system in south west and south east parts of the country. Semi-forest coffee production system represents about 24% of the total land covered by coffee and accounts for about 20% of total Ethiopian coffee production. Generally, the two production systems together represent about 34% and 25% of the land covered by coffee and the annual coffee production in the country, respectively (Gole *et al*., 2002).

In the garden coffee production system, coffee is planted and managed in the farmer's locality within small area, mainly found in southern and eastern part of the country. Teketay (1999) and Woldetsadik *et al.,* (2000) reported that it accounts for about 35% of the total production system of the country. Plantation coffee is grown on a large scale by private coffee farmers or the government. It is well managed, representing about15% of the total coffee productions. Recently Gole *et al*. (2002) reported that, the two systems together accounts for 67% and 75% of area and production of coffee of the country, respectively.

2.7. Contributions of *Coffea arabica* to Ethiopian economy

The contribution of coffee to export market and foreign currency earning is very high for Ethiopia. According to Tsegaye (2000), coffee contributes about 10% and5% of government revenues and gross domestic product, respectively. Besides, about 25% of Ethiopia's populations also depend on coffee for their livelihood (Gole, 2003). This means, no other product or service in Ethiopia has earned high labor intensive tree crop and provide much employment in rural areas and a means of living for many millions of people in Ethiopia. In addition to being an important export commodity, coffee plays a vital role or tremendous impacts on both social and spiritual life of the people in different geographical location and cultural backgrounds of the country (Gole, 2003). Almost 90-95% of the Ethiopian coffee is produced by smallholder coffee farmers with an average farm-size of 0.5hactar (Gole *et al.,* 2008).

2.8. Genetic Diversity of wild *Coffea. arabica* in Ethiopia

Characterization of genetic diversity within a crop plant is important, as it determines the extent to which the crop can be improved or changed by selection. Knowledge of the available variability will, among other things, lead to making decision on plant breeding options and selection strategies. Traditionally, breeders have used visible phenotypic features like fruit size, shape, color etc. Such methods of selection of useful traits based on phenotypic features are however influenced by the environment (Bar *et al.,* 2010).

Molecular markers offer a more reliable approach for genetic diversity study. Data generated from molecular markers can provide information on phylogenetic relationships, how divergent populations are and guidance on decisions concerning *in-situ* conservation strategies (Bar *et al.,* 2010).

According to Lowe *et al*. (2004), genetic diversity is a commonly used expression to refer and know heritable variation present within and among biological entities such as plants, animals and microorganisms. These genetic variations can be enumerated at three levels: species, populations and individual levels. Since Ethiopia is the only center of origin and diversifications of *Coffea arabica,* there is a high genetic diversity, which is mainly attributed to its diverse ecological

features such as suitable altitude, ample rainfall, optimum temperature, fertile soils etc. and the presence of indigenous methods of coffee production system in the country (Yeshitila *et al.*, 2004; Gole *et al*., 2001). The standing wild coffee populations are different from the other coffee types and its cultivated varieties, in terms of the levels of diversity on the basis of both morphological and molecular markers analysis (Montagnon *et al.*, 1996; Anthony *et al.*, 2001).

The entire genetic diversity of *Coffee arabica* is largely found in southwest and southwest montane rainforestes of Ethiopia at its center of origin (Gole *et al*., 2001; Senbeta, 2006). Based on multivariate analysis of phenotypic diversity of *Coffea arabica,* Montagnon *et al.,* (1996) indicated clear structure within the species and identified two main groups: one west and the other east of the Ethiopian Great Rift Valley. Additionally, Anthony *et al*. (2001) reported the presence of higher level genetic diversity with RAPD molecular markers investigation among spontaneous and sub-spontaneous coffee collected from Ethiopia. Moreover, Senbeta *et al*. (2006) and Tesfaye (2006) reported floristic and molecular genetic studies also show the high species diversity of the montane rain forest and the high genetic diversity of the wild coffee populations, respectively; although the diversity information from both molecular and agro-morphological marker indicated the high genetic diversity in the wild form of the coffee plants. In contrary to this Anthoney *et al*. (2001), also suggested that, *C. arabica* is characterized by a low genetic diversity as compared to other coffee species due to the recent evolution of the species by allopolyploidy and it predominantly self-pollinating nature.

Similarly, study done by using combined analysis with AFLP and microsatellites on four coffee cultivars growing in Yemen and 11 sub spontaneous accessions collected in the primary centre of diversity of the coffee species, Anthoney *et al*. (2002) reported the low genetic diversity in the cultivated populations of the coffee plants. Furthermore, breeding programs are thus limited due to the very narrow genetic basis of the present cultivars elsewhere in the world. This can be justified by difficulties encountered in finding sources of resistance against disease within the cultivars of *Coffea arabica* in other parts of the world (Aga et al., 2003).

According to Gole (2003), and Senbeta *et al*. (2006),wild *Coffee arabica* of the montane rain forests of Ethiopia is considered to be the important and valuable coffee genetic resources of the national and the international coffee breeding or coffee industry in the future for the

improvement of coffee crop. For conservation and sustainable use of the coffee germplasm, knowing of the extent and distribution of the genetic diversity within and among the biological entities at various levels are crucial with the help of modern molecular marker technologies.

2.9. Genetic markers for diversity analysis

Genetic markers are measurable inherited genetic variations, used to understand genetic components. There are different types of genetic markers with different properties, each having its own advantages and disadvantages to assess the genetic variations among natural populations. Currently, the most commonly used genetic markers are morphological, biochemical and molecular markers.

Due to the rapid developments in the field of molecular genetics, a variety of different techniques have emerged to analyze genetic variation during the last few decades (Whitkus *et al.*, 1994; Karp *et al*., 1996, 1997a, b; Parker *et al*., 1998; Schlötterer, 2004). These genetic markers may differ with respect to important features, such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and financial investment. No marker is superior to all others for a wide range of applications. The most appropriate and important genetic marker will depend on its specific application, the presumed level of polymorphism, the presence of sufficient technical facilities and know-how, time constraints and financial limitations.

Generally markers are divided in to three broad classes: those based on visually assessable traits (morphological and agronomic traits), those based on gene product (biochemical markers), and those relying on a DNA assay (molecular markers). Genetic markers are widely used by breeders and conservationists to study genetic diversity and assist crop improvement. Each marker system has its own advantages and disadvantages in terms of ease of use and/or the degree of information provided.

2.9.1. Morphological markers

The earliest studies of genetic characterization and divergence are based on morphological markers such as qualitative and quantitative traits (Arriel *et al.*, 2007; Cruz *et al.*, 2004). These markers are inexpensive and simple to score, are based on distinct phenotypes such as plant

color, plant height, seed characteristics, etc. The apparent disadvantage of such markers is that in studies of genetic diversity the expression of the phenotypes is highly influenced by environmental conditions. Lack of adequate genome coverage because of limitation of the number of markers and problems of dominance can also be mentioned as weaknesses of morphological markers (Brown, 1978). Furthermore, the expression of these characters which is influenced by the environment may require that plants be grown to a suitable stage before certain characters can be scored.

2.9.2. Biochemical markers

Biochemical markers are markers derived from study of the chemical products of gene expressions. These are also termed isozyme/allozyme markers or simply protein markers. According to Karp *et al*. (1997), the agro-morphological markers resulted in the development of biochemical markers that complement its drawbacks. According to Gottlieb (1981) the oldest biochemical techniques used to study variations is isozyme analysis. It has the power to reveal polymorphism of alleles at any particular locus on the basis of protein mobility (Brown *et al.*, 1975). Isozyme technique is fast, cheap and simple. However, isozyme markers are not as plentiful as DNA markers; it underestimates the level of genetic diversity (Dudnikov, 2003) and sometimes interpretation of bands become difficult due to complex banding profiles arising from polyploidy or duplicate genes. In addition, proteins with identical electrophoretic mobility (comigration) may not be homologous (Morell *et al*., 1995). Isozyme studies in plants have demonstrated that, pattern and band intensities differ by tissue types and developmental stages (Montarroyos *et al*., 2003). Although isozymes are with limitations, the technique has been used for genetic diversity analysis in many species (Dudnikov, 2003) and it appears to be more informative at lower taxonomic levels, particularly at species and population level characterization (Brown, 1990).

Isozymes have been applied to *C. arabica.* However, their use on wild *C. arabica* species characterizations from Ethiopia have failed to reveal polymorphism, indicating that isozyme technique is not appropriate for diversity study in *Coffea arabica* due to the small number of isozyme systems available (Berthaud *et al.*, 1988). Moreover, Paillard *et al*., (1996) tried to construct isozyme based genetic map for coffee and was unsuccessful due to the low

polymorphism level. Consequently, like many other crop species, arabica coffee researchers shifted towards using DNA-based markers.

2.9.3. Molecular markers

Molecular markers are fragments of nuclear, mitochondrial or chloroplast DNA, with specific sequences and repetitive DNA (satellite, dispersed, tandem repeats).These technologies based on polymorphism in DNA, can be considered as objective measures of variations and have catalyzed research in a variety of disciplines such as phylogeny, taxonomy, ecology, genetics, plant and animal breeding, etc. Markers are only informative if they are polymorphic in populations. Level of polymorphism is an important determinant of what a marker is useful for. Many different types of molecular markers with different properties exist, each with its own advantages and disadvantages (Weising *et al.,* 2005; Karp *et al.*, 1997). However, it is extremely difficult to find molecular markers which could adequately meet all the ideal properties of the molecular markers (Lowe *et al*., 2004). But depending up on the type of the study to be undertaken, one can identify between a varieties of marker systems that could fulfill the objectives of the study (Weising *et al*., 2005). Many authors also suggest the use of more than one type of molecular markers in a single experiment (Karp *et al.,* 1997). The DNA based marker systems are generally classified as hybridization-based (non-PCR) markers and (PCR)-based markers (Weising *et al.,* 2005; Joshi *et al*., 1999).

2.9.3.1. Non-PCR based markers

The main non-PCR-based technique is restriction fragment length polymorphism (RFLP). RFLP are co-dominant markers and are more informative than dominant markers. However, this technique requires relatively large amounts of purified and high molecular weight DNA, is time consuming and laborious, uses radioactive probes that, is difficult to handle and dispose, etc. In arabica coffee RFLP markers were employed to identify the origin of the *C. arabica* genome. By comparing the RFLP patterns of wild diploid species with those of *C. arabica*, Lashermes *et al*. (1999) concluded that the *C. arabica* genome was formed by hybridization between *C. eugenioides* and *C. canephora*, or ecotypes related to these diploid species.

2.9.3.2. PCR-based DNA markers

The development of the polymerase chain reaction (PCR), many PCR based molecular techniques have been, and still are being developed for plant genome analysis. The techniques could be categorized in to two main groups: arbitrary (or semi arbitrary) primed techniques and site targeted PCR techniques (Karp *et al*., 1997; Weising *et al*. 2005). Techniques in the first category use primers which are designed arbitrarily/or semi arbitrarily, i.e., without the knowledge of flanking sequence of the region to be amplified. These techniques includes RAPD (random amplified polymorphic DNA) which was introduced in 1990 (Williams *et al*., 1990), ISSR (inter simple sequence repeats) (Zietkiewicz *et al*., 1994) and AFLP (amplified fragment length polymorphism) (Vos *et al*., 1995). These are also called dominant molecular markers. These classes of dominantly expressed markers (RAPD, AFLP and ISSR) are multi-locus, and scattered throughout the entire genome (Lowe *et al*., 2004). Techniques in the second category depend on primers that target a single known site, such as a gene. These are alternative approaches to multi-locus profiling. Microsatellites (SSR), PCR-DNA sequencing, sequence-tagged microsatellite (STMs), thermal gradient gel electrophoresis (TGGE),denaturing gradient gel electrophoresis (DGGE), cleaved amplified polymorphic sequences (CAPS), single–strand conformational polymorphisms (SSCP), are important examples of these categories (Karp *et al*.,1997). Since they give information on a single locus, they are particularly important when information is required on gene frequency or genealogical information for genetic diversity management and when information on heterozygosity is required.

2.9.3.2.1. Inter simple sequence repeats (ISSRs) markers.

ISSRs are DNA fragments of about 100-3000 base pairs located between adjacent, oppositely oriented microsatellite regions. ISSRs are amplified by PCR using microsatellite core sequences as primers with a few selective nucleotides as anchors into the non-repeat adjacent regions (16-18 base pair). About 10-60 fragments from multiple loci are generated simultaneously, separated by gel electrophoresis and scored as the presence or absence of fragments of particular size. 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 *et al.,* 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 *et al.,* 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).

ISSR method has been used extensively to identify and determine relationships at the species and cultivar levels (Martins *et al.*, 2003). This method is widely applicable because they do not need sequence data for primer construction, rapid, inexpensive and randomly distributed throughout the genome.

The ISSR method has been reported to produce more complex marker patterns than the RAPD approach (Parsons *et al*., 1997; Chowdhury *et al*., 2002), which is advantageous when differentiating closely related cultivars. In addition, ISSR markers are more reproducible than RAPD markers (Oliveira *et al.*, 2001), because ISSR primers, designed to anneal to a microsatellite sequence, are longer than RAPD primers, allowing higher annealing temperatures to be used. It also because of the multilocus fingerprinting profiles obtained, ISSR analysis can be applied in studies involving genetic identity, parentage, clone and strain identification, and taxonomic studies of closely related species. In addition, ISSRs are considered useful in gene mapping studies (Godwin *et al*., 1997; Zietkiewicz *et al*., 1994; Gupta *et al*., 1994).

ISSR markers have proven valuable for fingerprinting studies and genetic diversity investigations in tef (Assefa *et al.*, 2003) and rice (Blair *et al.*, 1999; Joshi *et al.*, 2000; Girma *et al.*, 2003), coffee (Tesfaye, 2006; Aga, 2005;).ISSR markers have been used to study forest coffee in Ethiopia (Aga, 2005; Tesfaye, 2006; Balemi, 2007). However, due to resource limitation and inaccessibility, some forest regions were not included in the previous study. Hence, to evaluate the genetic diversity and patterns of distribution of *C. arabica* populations in previously unaddressed region of Ethiopia the ISSR marker assay has been chosen for the present study.

3. OBJECTIVES

3.1. General objective

The general objective of the study was to investigate the genetic variability among the south and south west Ethiopian wild *Coffea arabica* populations using inter simple sequence repeat markers.

3.2 Specific objectives

- To determine the level and pattern of distribution of variation among forest coffee (*C. arabica* L.) in; Dawro, Anfilo, Bench Maji and Limu areas.
- To generate information that can be used to design conservation strategies that could act as reservoir of coffee gene for sustainable use and conservation.

4. MATERIAL AND METHODS

4.1. Plant Material and sampling

The Arabica coffee plant used in this study were collected from five different forests in south and south west Ethiopia (Table-4.1). Individual of *Coffea arabica* trees were randomly sampled from five forests, with a size of 50X 50 m. Fresh leaves were collected from individual tree and dried in silica gel. Six to twenty-two individuals were collected from each population and in total 71 individuals trees were sampled. The forests are located in, Anfilo (close to Denbi-Dolo) and Dawero, Essara district (Banja forest), Limmu, Bench Maji-1 (near Maji, Gobe forest) and Bench Maji-2 (Jeba).

Table-4.1Altitude and location of sample collection used in this study, from south and south west Ethiopia.

No	Region	Specific location	Sample size	…..Longitude/Latitude	Altitude
1	Oromia	Limmu (Jimma)	6	36° 24' 00" E/9° 47' 00" N	1630
2	Oromia	Wellega (Anfilo)	15	034°35.101 E / 08 3°1.683 N	1422
3	SNNPR	Dawero Banja	20	36° 46.917E/6°57.2945N	1620
4	SNNPR	Bench Maji-1(Maji)	22	6°12′N 35°35′E/6°12′N 35°35′E	2430
5	SNNPR	Bench Maji-2(Jeba)	8	33.0461°E/7.7944°N	2088

4.2. DNA extraction

The CTAB (2% Cetyltrimethyl ammonium Bromide, 1% polyvimylpyrolidone, 100mMTris: PH=8, 20mM EDTA, 1.4M. NaCl, 0.2% beta-Mercapto-ethanol) solution was warmed in water bath up to 65°C. Then, 100-200 mg fresh leave material per sample was weighed and pulverised thoroughly using a clean mortar and pestle and grinded after sand was added. Then, the powder was transferred into an eppendorf cap and 700 µl of warm CTAB solution was added to the

powdered sample and dissolved the powder and the sample was incubeted for 30 minutes at 65°C then centrifuge for 5 minutes at 15000 rpm. The supernatant was transferred to a new eppendorf cap and 700 µl CTAB solution was added to the tissue pellet for the second extraction and stirred slightly with a new 1000 µl pipette tip and incubated for 30 min at 65°C then, 600 µl chloroform was added to the cap immediately with supernatant and shake carefully a few times upside down for approximately 5 minutes and centrifuged for 5 min at 15000 rpm, then the supernatant (only clear liquid) was transferred to a new eppendorf-cap and repeated the chloroform extraction to removed all impurities. Then, cooled isopropanol (4°C), approximately $^{2}/_{3}$ of the solution volume was added and shake carefully and allowed to freeze for more than 2 h at -20°C. After centrifuge 10 min at 15000 rpm, the liquid was aspirated using yellow tips without touched the pellet and 200 µl of 70 % ethanol was added to the pellet and centrifuge for 10 min at 15000 rpm in a cooled centrifuge and aspirated ethanol by yellow tips and dried the DNA pellet at room temperature for 15 minutes and dissolve in 100 µl TE (1x, p.a. grade) and store at 4°C. Then, 50 µl cooled 7.5 M NH_4Ac- solution was added and mixed carefully. 100 µl cool 100% ethanol was added and mixed carefully and frozen for more than 2 h at -20°C, centrifuge 30 min at 15000 rpm and aspirated fluid carefully. Then, 200 µl of 70% ethanol was added and rinsed the inner cap surface by turned the cap and centrifuge 10 min. at 15000 rpm and aspirated the liquid and dried pellet at room temperature. Then, dissolved the pellet in 100 µl TE (1x, p.a. grade) and repeated these four steps with 3 M NaAC-solution instead of 7.5 M NH_4Ac solution (4°C, half the volume) then, centrifuge 10 min. at 15000 rpm and aspirated liquid and dried the pellet at room temperature and dissolved the pellet in 100 µl TE (1x, p.a. grade).

4.3. DNA Quality checking

Mostly the second extraction is required for ISSR-PCR analysis. Samples of two extractions were run out on agarose gel at 0.98% concentration to check the presence and quality of genomic DNA and then to select on the basis of quality and concentration. Only 5µl of genomic DNA sample of the stock solution and 2µl of 1x loading dye were used and applied by adding a concentration of 0.5mg/ml ethidium bromide on the 0.98% gel in 1x TBE solution running buffer and then electrophoresed at 80 volt constant for 45 min. For comparisons, 100bp DNA ladder with known concentration was used. The gels were photographed with digital camera mounted

on Bio DocAnalyze apparatus and connected to desk top computer. The gel picture was further examined and used to make selection of good quality DNA extract. In most of the cases, genomic DNA from the second extractions was found to be promising and was selected for ISSR-PCR analysis.

4.4. PCR optimization and primers used

Estimation of the extent and distribution of genetic diversity among and within the individual sample of *Coffea arabica* from four different forests were analyzed using Inter Simple Sequence Repeat (ISSR) markers. The ISSR marker assay was conducted at Cytogenetic and Isozyme Research Laboratory of the Faculty of Life Since, Addis Ababa University, Addis Ababa. The polymerase chain reactions were conducted in Biometra T3 Thermocycler. The amplification program was 4 min at 94^{o}C; 39 x 15s at 94°C, 1min at the primer annealing optimal temperature (45^{o}C or 48^{o}C), 1.30min at 72^{o}C, and at 72^{o}C for 5 min final extension. Therefore, a total of four primers three di-nucleotide (810, 812 and 818) and one penta-nucleotide (880) primers were selected after testing and screening nine primers at initial stage.

Table 4.2:-List of primers, annealing temperature, primer sequence, remark and repeat motives used or selected for optimization.

No	primers	Annealing Temperature	Primer Sequence	Repeat motives	Remark
1	810	45	$(GA)_8T$	Di nucleotide	Good
2	812	45	$(GA)_8A$	Di nucleotide	Good
3	813	45	$(CT)_{8T}$	Di nucleotide	Reproducible but not polymorphic
4	818	45	$(CA)_8G$	Di nucleotide	Good
5	834	48	$(AG)_8GY$	Dinucleotide	Not reproducible
6	873	45	$(GACA)_4$	Dinucleotide	Reproducible but not polymorphic
7	880	48	$(GGAGA)_3$	Pentanucleotide	Good
8	CoIS001	45	$(CCTA)_4$	Tetranucleotide	Reproducible but not polymorphic
9	CoIS002	45	$(GGTA)_4$	Tetranucleotide	Reproducible but not polymorphic

Source: Primer kit 900(UBC 900), Y=(C,T)

To determine the optimum concentration and dilution proportion, two separate PCR reactions were randomly run with DNA diluted at 1:5, 1:10 and 1:20 for two samples randomly selected from each four populations using primer 810 and 813. The PCR reaction was carried out in a

total reaction volume of 24µl containing a concentration of 10 ng template DNA, 11.05µl H_20, 1.25mM dNTPs, Taq buffer (1xThermopol reaction buffer), 2mM $Mgcl_2$, 2.7mM beitein, 5u/µl Taq Polymerase and 20pmol/µl, 810 and 813 primers. The test gel was performed with PCR amplified products the test gel with sample that was diluted at 1:5, 1:10 and 1:20 showed good bands with a better intensity for 1:20 dilutions. Hence, 1:20 dilution was used for all working samples for further analysis.

4.5. PCR and gel electrophoresis

The polymerase chain reaction was conducted in Biometra 2000 T3 Thermo cycler. PCR amplification was carried out in a 25 µl reaction mixture containing a concentration of 10 ng template DNA, 11.05µl H_20, 1.25mM dNTPs, Taq buffer (1xThermopol reaction buffer), 2mM $Mgcl_2$, 2.7mM beitein, 20pmol/µl primer and 5u/µl Taq Polymerase. The amplification program was 4 minutes preheating and initial denaturation at 94^0C, then 40 x 15 seconds at 94^0C, 1 minute annealing at (45^0C/ 48^0C) based on primers used, 1.30 minutes extension at 72^0C and the final extension for 7 minutes at 72^0C. The PCR reactions were stored at 4^0C until loading on to gel for electrophoresis. The amplification products were differentiated by electrophoresis using 1.67 % agarose gel (with 100 ml 1xTBE) and 8µl amplicence product of each sample with 2µl loading dye (1 times concentrated) was loaded on gel. DNA marker of 1k bp was used to estimate molecular weight and size of the fragments. Electrophoreses were done for 3 hours at voltage of 100V. The DNA was stained with (0.5mg/ml) ethidum bromide (EtBr) which were mixed with 450 ml distilled water for 30 minutes and washed with distilled water for 30 minutes.

4.6. Data analysis

ISSR profiles/bands were scored manually for each individual accession from the gel photograph. The bands were recorded as discrete characters, presence ‘1’ or absence ‘0’ and ‘?’ for missing and ambiguous data. Based on recorded bands, different softwares were used for analysis. POPGENE version1.32 software (Yeh *et al.,* 1999) was used to calculate genetic diversity for each population as number of polymorphic loci and percent polymorphism. Analysis of molecular variance (AMOVA) was used to calculate variation among and within population using Areliquin version 3.01 (Excoffier*et al.,* 2006). Shannon–Weaver diversity index (H) was calculated as:

H= -∑pilog2pi,

Where,

pi is the frequency of a given band for each population (Lewontin, 1972).

NTSYS- pc version 2.02 (Rohlf, 2000) and Free Tree 0.9.1.50 (Pavlicek *et al.,* 1999) software's were used to calculate Jaccard's similarity coefficient which is calculated with the formula:-

$$S_{ij} = \frac{a}{a+b+c}$$

Where,

'a ' is the total number of bands shared between individuals i and j,

'b' is the total number of bands present in individual i but not in individual j and

'c' is the total number of bands present in individual j but not in individual i.

The Unweighted Pair Group Method with Arithmetic mean (UPGMA) (Sneath *et al.,* 1973) was used to analyze and compare the population and generate phenogram using NTSYS- pc version 2.02 (Rohlf, 2000). The Neighbor Joining (NJ) method (Saitou *et al.,* 1987; Studier *et al.,* 1988) was used to compare individual genotypes and evaluate patterns of genotype clustering using Free Tree 0.9.1.50 Software (Pavlicek*et al.,* 1999).

The patterns of variation among individual samples on 3D, a principal coordinated analysis (PCO) was performed based on Jaccard's coefficient (Jaccard, 1908). The calculation of Jaccard's coefficient was made with PAST software version 1.18 (Hammer *et al.,* 2001). The first three axes were used to plot the three dimensional PCO with STATISTICA version 6.0 software (Hammer *et al.,* 2001; Statistica Soft, Inc.2001).

5. RESULTS

5.1. Banding patterns of the ISSR primers

From the nine primers tested, four (three di-nucleotide and one penta-nucleotide) gave relatively clear banding pattern were selected and used in this study (Table 5.1). The molecular weight of the bands amplified using these four primers were in the range of 400 bp to 1000 bp. A total of 30 clear and scoreable bands were scored, from 71 wild *Coffea arabic* populations collected from different localities. All of the 30 scored fragments were found to be polymorphic among coffee individuals. The highest polymorphic bands (9) were scored from primer 812 and the remaining three primers; 810, 818 and 880 shows 6, 7 and 8 polymorphic loci, respectively. All the latter four primers showed 100 per cent polymorphism. Figure 5.1 shows the fingerprint pattern of primer 812. Primer 812 showed the highest values of gene diversity (0.47) and Shannon value (0.66) (Table 5.3).

Table-5.1 Banding patterns generated using the four primers, their repeat motifs, remark and number of polymorphic loci.

Primers	**Repeat Motives**	**Number of Polymorphic loci**	**Remark**
810	$(GA)_8T$	6	Good
812	$(GA)_8A$	9	Good
818	$(CA)_8G$	7	Good
880	$(GGAGA)_3$	8	Good
	Total	**30**	

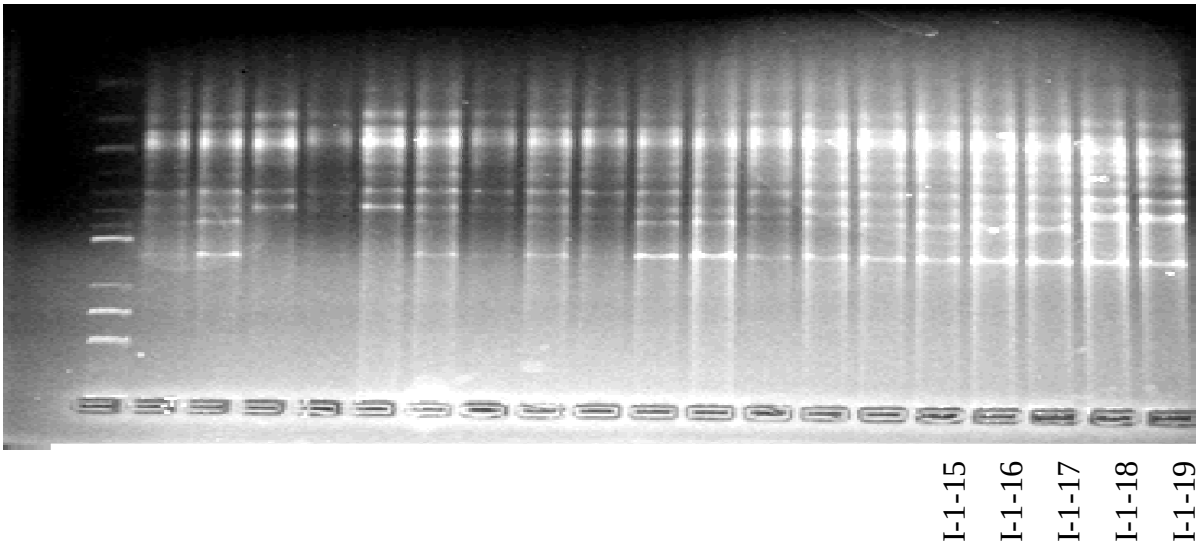


Figure 5.1: ISSR fingerprint generated from some individuals of wild *Coffea arabica* using primer 812.

5.2. Polymorphism based on ISSR analysis

In all populations, the number of polymorphic loci were; 6, 7, 8 and 9 for primer; 810, 812, 818 and 880 respectively. Of the total 30 loci scored, all (100%) were observed to be polymorphic. From all the populations studied; Bench Maji-1 and Dawero Banja showed 100% polymorphism and followed by Anfilo (63.33%), Limmu (56.67%) and Bench Maji-2 (33.3%).

5.3. Genetic diversity

Samples from Dawero Banja were the most diverse in terms of gene diversity value (*GD*=0.46) followed by population from Bench Maji-1 (*GD*= 0.40) and both Anfilo and Limmu showed equal value (GD= 0.23). Population from Bench Maji-2 was the least diverse with gene diversity value of 0.12. The average gene diversity relative to the total population (GD_T) was 0.45. The same patterns of Shannon diversity index were observed as gene diversity with the highest value for Dawro Banja populations with 0.65 and least for Bench Maji with 0.17 (Table 5.2).

Table 5.2:-Number of scorable bands (NSB), number of polymorphic loci (NPL), percent of polymorphism (PP), gene diversity (GD) and ShanonIndex (I), with 5 populations of *Coffea arabica* for all primers

Population	NPL	PP	GD±SD	I±SD
Anfilo	19	63.33	0.23±0.21	0.32± 0.29
Bench				
Maji-1	30	100	0.4± 0.1	0.6 ±0.1
Bench				
Maji-2	9	30	0.12± 0.2	0.17± 0.28
Dawero				
Banja	30	100	0.46± 0.05	0.65± 0.06
Limmu	17	56.67	0.23± 0.22	0.34± 0.31
Average	**21**	**70**	**0.288**	**0.416**
Total	**30**	**100**	**0.45± 0.06**	**0.64± 0.07**

Table 5.3:-Number of scorable bands (NSB), number of polymorphic loci (NPL), percent polymorphism (PP) genetic diversity (GD) and ShanonIndex (I) for each primer.

Primers	**NSB**	**NPL**	**PP**	**GD**	**I**
810	6	6	100	0.4± 0.1	0.57± 0.11
812	9	9	100	0.47± 0.02	0.66± 0.02
818	7	7	100	0.46± 0.06	0.65± 0.06
880	8	8	100	0.46 ±0.04	0.06 ±0.04
Total	**30**	**30**	**100**	**0.45±0.055**	**0.5±0.057**

5.4. Analysis of molecular variance (AMOVA)

Analysis of molecular variance was carried out at two levels; I) populations grouped in to Oromia and SNNPR, (II) for the entire populations The analysis was carried out by computation of the distance between 'haplotypes', each individuals data pattern as one 'haplotype' and computing variance components for each level (i.e. presence 1, absence 0) (Exocoffer *et al*., 1992).

Partitioning of genetic diversity by analysis of molecular variance using grouped populations revealed that out of the total genetic diversity, most of the diversity was attributed to individual plants within the populations (73.16%) with the remaining diversity being distributed among populations within groups (10.77%), (Populations within Oromia and SNNPR groups) and among groups (16.07%). Similarly partitioning of genetic diversity by analysis of molecular variance without grouping populations revealed that out of the total genetic diversity, most of the ISSR diversity is due to differences between individual plants within the populations (79.4%), while the remaining is due to differences among populations (20.6%). In both cases, the results of analysis of molecular variance revealed the same patterns of genetic diversity and supports the

larger genetic diversity was found within the populations rather than among populations. The AMOVA result is summarized in table (5.4, 5.5).

Table5.4:- Analysis of molecular variance (AMOVA) of wild *Coffea arabica* in south and south west Ethiopia based on 30 ISSR without grouping.

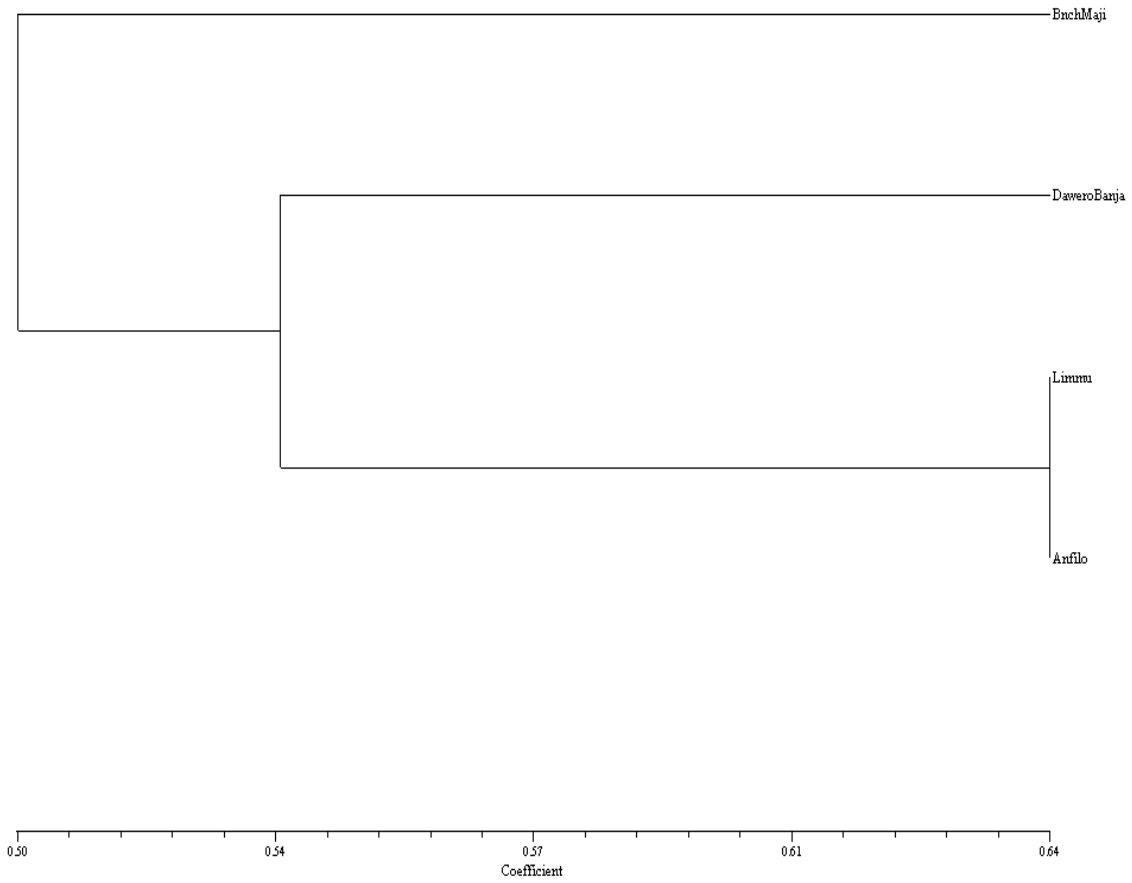
Source of Variation	**Sum of squares**	**Variance components**	**Percentage variation**	**fixation**	**p**
Between Populations	85.95	1.3	20.6	0.2	0.00
Within Populations	309.18	5.2	79.4		0.00
Total	**395.13**	**6.5**			

Table-5.5:- Analysis of molecular variance (AMOVA) of wild *Coffea* arabica in south and south west Ethiopia based on 30 ISSR with grouping.

Source of Variation	**Sum of squares**	**Variance components**	**Percentage variation**	**fixation**	**p**
Among Groups	42.90	1.14	16.1	0.13	0.00
Among Populations Within Groups	43.04	0.76	10.77	0.27	0.00
Within Populations	309.18	5.16	73.16	0.16	0.00
Total	**395.13**	**7.06**			

5.5. Clustering analysis

Interregional cluster analysis of UPGMA (Unweighted Pair-Group Methods Using Arithmetic Averages) and NJ (Neighbor Joining) were computed for the 4 populations having a total of 71 individuals of *Coffea arabica* based on 30 ISSR- PCR fragments amplified by three di-nucleotides (810,812 and 818) and one penta nucleotides (880) primers. UPGMA analysis based on wild *Coffee arabica* populations (Fig 5.2) revealed one major group and two outliers (Bench Maji and Dawero Banja). This major cluster again forked into two subgroups, the first containing Anfilo, while the second contained Limmu population. But individual based UPGMA clustering of an overall analysis (Fig 5.3), most individuals clustered according to their respective populations. In terms of grouping intensity on individual based UPGMA; Bench Maji-1, 2 tend to intermix with Dawero Banja, while Anfilo tends to intermix with Limmu. The dendrogram derived from neighbor-joining analysis of the whole ISSR data which includes 71 individual of wild *Coffea arabica* (Fig 5.4) showed two distinct major clusters and sub-clusters within both major clusters. With the exception of some; Bench Maji-1 and Anfilo individual which are clustered in cluster II sub cluster I and II, respectively in the denderogramm, other individual did not show a clear grouping



Figure, 5.2: UPGMA based dendrogram for four south and southwest Ethiopian *Coffea arabica* populations using 4 ISSR (3 di and1 tetra nucleotide) primers.

Key:

- I-1-1 up to I-1-22-Bench Maji-1
- I-2-1 up to I-2-8- Bench Maji-2
- II-1-1 up to II-1-20-Dawero Banja
- III-1-1 up to III-1-6-Limmu
- IV-1-1 up to IV-1-15-Anfilo

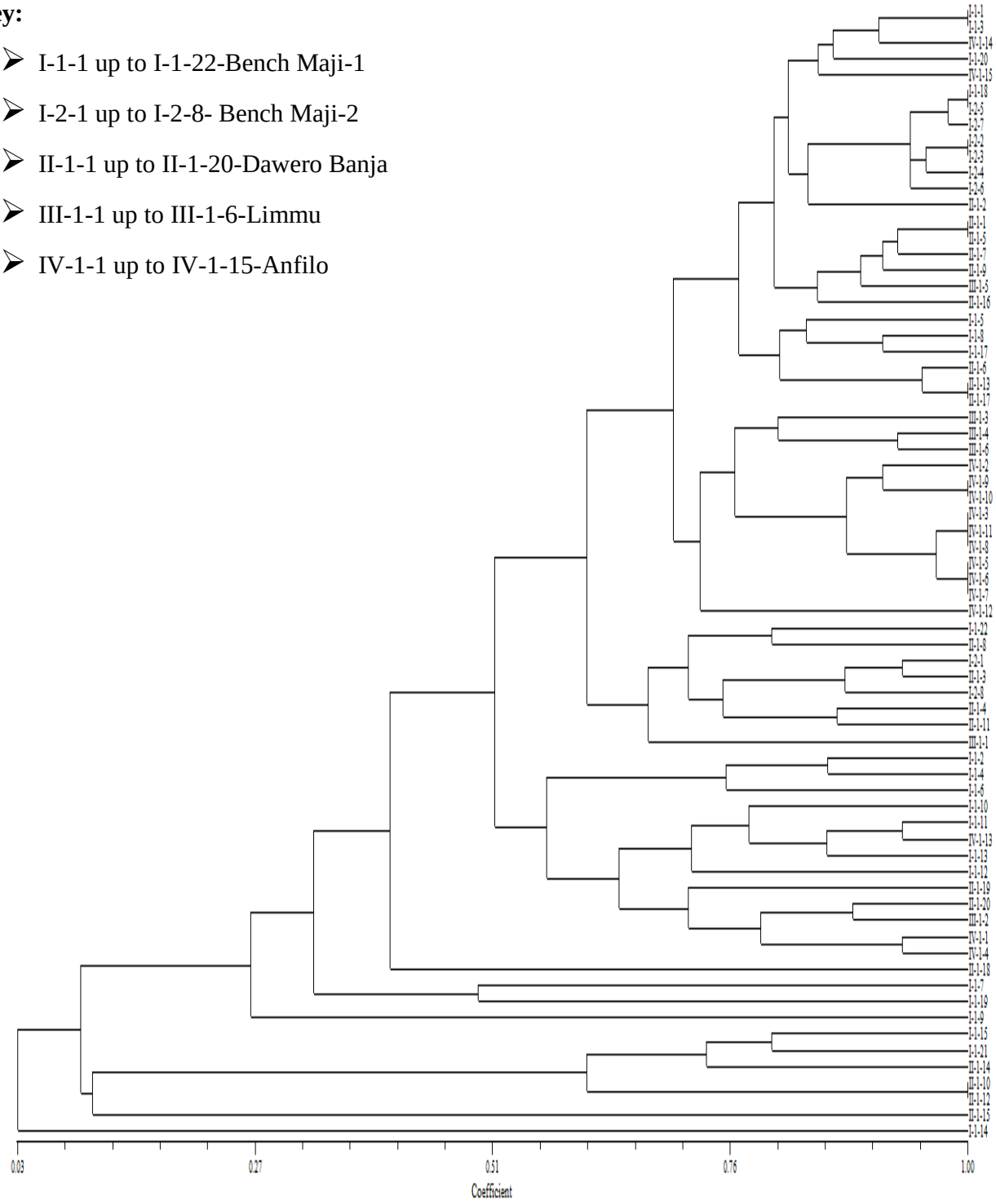


Figure 5.3, UPGMA based dendrogram for 71 individuals of *Coffea arabica* using 4 ISSR (3 di and1 tetra nucleotide) primers.

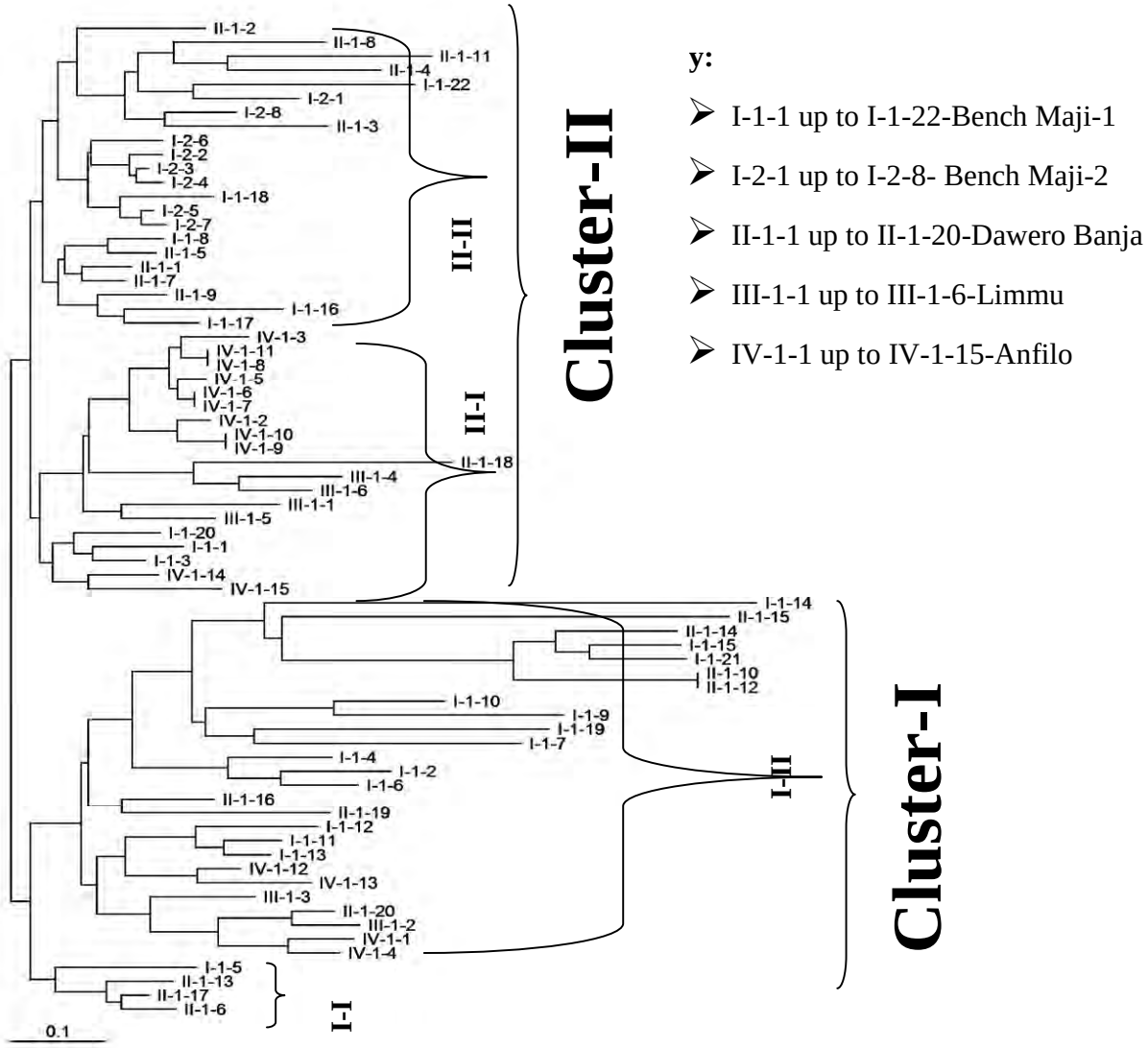


Figure5.4: Neighbor-joining analysis of 71 individuals of *Coffea arabica* based on 30 PCR bands amplified by: three di-nucleotide (810,812 and 818) and one penta nucleotide (880) primers.The neighbor joining algorithm is based on Jaccard's coefficients obtained after pair wise comparison of the presence-absence fingerprint

5.6.Principal coordinate (PCO) analysis

All the data obtained using 4 ISSR primers were used in PCO analysis using Jaccard's coefficients of similarity. The first two coordinates of the PCO having Eigen values of 1.79 and 0.76 with variance of 31.77 % and 13.48 %, respectively used to show the grouping of individuals using two and three coordinates (Figure 5.5 and Figure 5.6).In 3D, most of Dawero Banja and Bench Maji individuals were intermix with each other and made a separate group from the other two population but individual of Anfilo and Limmu were spread all over the plot. Using two coordinates, almost all individuals of each population were spread all over the plot

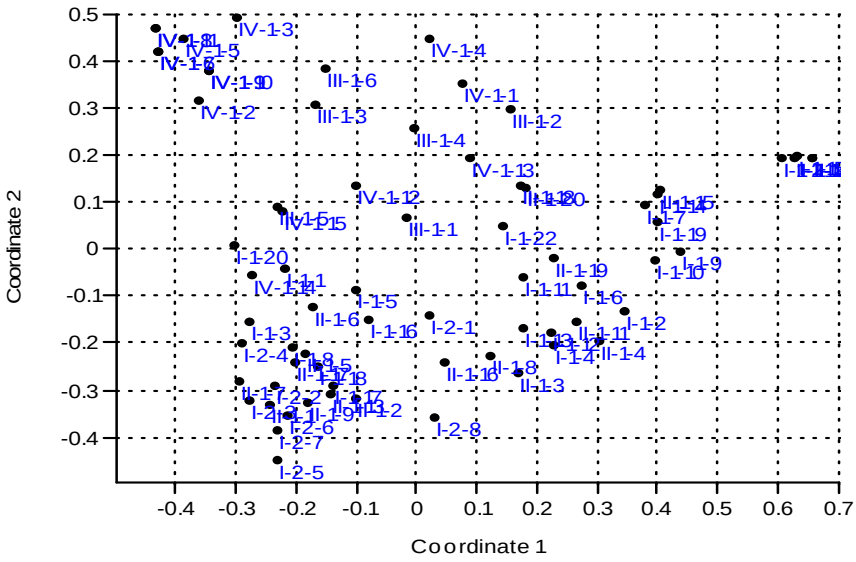


Figure 5.5: Two dimensional representation of principal coordinate analysis of genetic relationships among 71 individuals of wild *Coffea arabica.*

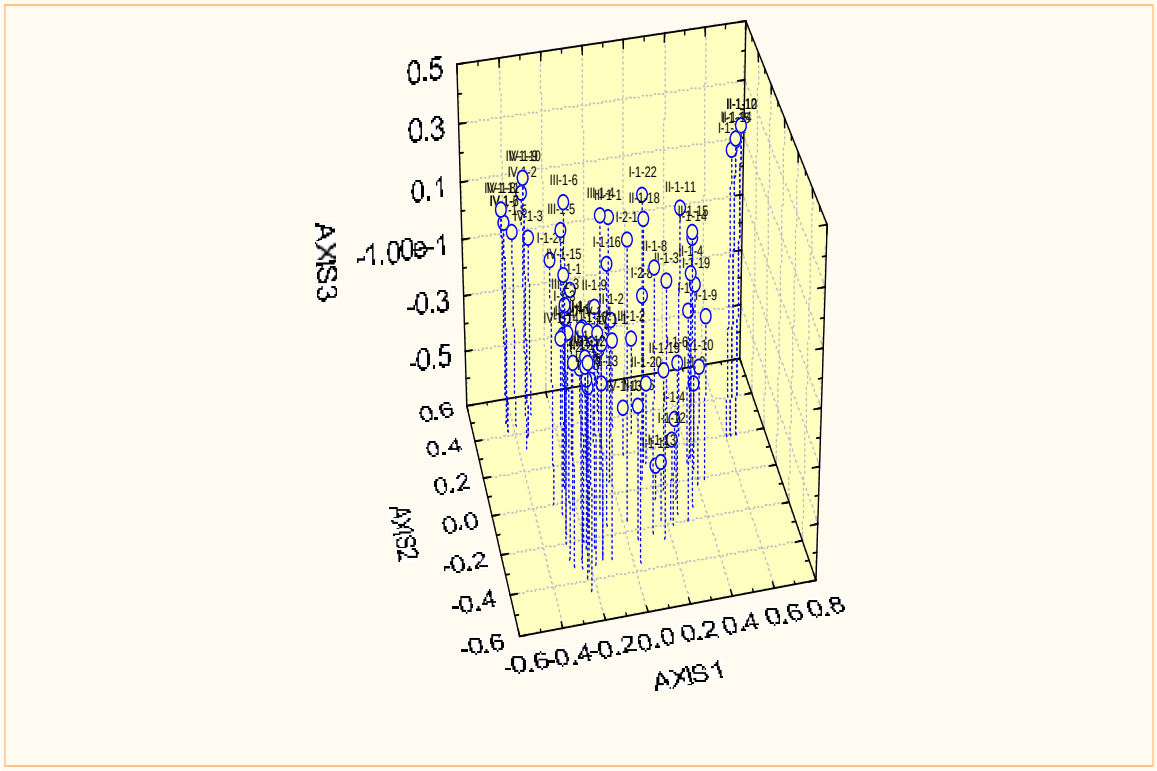


Figure 5.6: Three dimensional representation of principal coordinate analysis of genetic relationships among 71 individuals of wild *Coffea arabica.*

6. DISCUSSION

ISSR amplification technique has been found to be useful and potentially powerful technique for genotypic studies in many cultivars or crops. It was used to detect genetic diversity of coffee (Aga, 2005; Balemi; 2007 Tesfaye, 2006), Sesame (Admas, 2011; Teshome, 2011), lentils (Fikru, 2006) and rice (Gezahegne, 2007; Mitiku, 2011) collected from different parts of Ethiopia. In addition to revealing high level of polymorphism, ISSR is highly reproducible than RAPD and easy to use, less time demanding and cost effective markers as compared to SSR, RFLP and AFLP markers.

Analysis of molecular variance indicated that most of the ISSR diversity was distributed between individual plants within the populations, 79.4% and the rest 20.6% among populations. Tesfaye (2006) also obtained similar pattern with populations collected from nine different regions of Ethiopia whereby, 22.6% of the variance attributed to among groups and 77.4% was due to differences within groups. In addition, Balemi (2007) carried out similar investigation on Harenna forest coffee and observed similar pattern with higher proportion variation (63.7%) detected within population. Aga *et al.,* (2003) also observed similar pattern of diversity among Ethiopian coffee populations with Shannon-Weaver diversity index, whereby high genetic differentiations (80%) within zones of sample collection sites were observed.

The genetic structure of predominantly self-pollinating populations of a species, which are characterized by a relatively high value of total gene diversity, a low value of gene diversity within populations, high value of gene diversity among populations and a high value of the coefficient of gene differentiation . Self-pollinating species maintain high genetic diversity at their polymorphic loci, and most of this variation is found among populations (Aga *et al.,* 2003). The genetic diversity of a *Coffea arabica* could be affected by the mating system, gene flow, mode of reproduction and natural selection (Hamrik *et al.,* 1989). The genetic diversity of *C. arabica* in Ethiopia is strongly influenced by long distance gene flow (Tesfaye, 2006). Human activity in the forest such as, collecting seeds, seedling from productive forest and exchange with other region could cause long distance gene flow between populations and regions. Wild animals such as monkeys, baboons, insects and birds could also be other factors. On the other hand Balemi (2007) reported mixed mating system (partial out-crossing by pollen and seed and partial selfing) of *Coffea arabica* may probably affect extent of gene flow which could result in high

within genetic diversity. Tesfaye, (2007) reported single chloroplast haplotype was found in *C. arabica* indicates that all current populations have originated from a common ancestor in such short time, that no mutations could accumulate. This could either be due to a very recent allopolyploidization event or a strong bottleneck situation in the evolutionary history of *C. arabica*, perhaps caused by habitat changes. In contrary to the above, it could also be believe that the high genetic diversity observed within Arabica coffee populations studied, might be due to preferential adaptive gene complexes adapted to environmental changes being evolved during long evolutionary period in the regions.

Cluster and PCO analysis were employed to observe the genetic relationship in 71 individual *Coffea Arabica* collected from five populations. On the basis of UPGMA most individuals of the respective population were observed to form moderate grouping with few intermixing with other population. The intermixing goes with geographic proximity and administrative relationship which could facilitate seed and seedling flow and population migration. For instance, the intermixing of Bench Maji-1, 2 with Dawero Banja could be because of their geographic location and interaction among ethnic groups in Dawro and Maji; the same explanation could also work for Anfilo with Limmu which is located in western side of Oromia regional state. In the case of NJ, each individual of the respective population were observed to form small clustered groups with highly inter mixing with other group. Three-dimensional representation of a principal coordinate analysis of the genetic relationships among 81 individuals of *C. arabica* from Yayu, inferred from a distance matrix using the Jaccards index, by Tesfaye, 2006 shows tendency that populations found in spatial proximity group together. However, population-9 of Berhane Kontir formed a group with population-1 and -2 of Yayu, this implies the relative degree of similarity or difference in terms of floristic composition or other complex environmental variables; and indicates the history of these plots might be similar, except the area or region difference, this study also revealed similar patterns in 3D, most of Dawero Banja, Bench Maji-1 and Bench Maji-2 individuals were intermixe each other and made separate group from the rest. While the individual from Anfilo and Limmu were spread all over the plot, which could show the levels of variation among these individuals. Using two coordinates almost all individual of each population were spread all over the plot. Generally, on the basis of UPGMA, most individuals of the respective population were observed to form moderately clustered groups with few intermixing from the other population. In the case of NJ, each individual of the respective

population were observed to form small clustered groups with highly inter mixing with other group. This lack of association of the individual plant in to their respective populations in case of NJ and 2D may indicate the vulnerability or exposure of the forest to human intensive management and exchange of coffee seeds via extension package within and among SNNP and Oromia regions. Moreover, 3D and UPGMA analysis showed better resolutions as compared to 2D and NJ, respectively.

Both crop-improvement and conservation programmes depend on the availability and knowledge of the genetic resources in a germplasm collection. This study revealed that, there is higher genetic diversity in Bench Maji-1(Maji) and Dawero Banja than other three population area. Hence, higher conservation priority should be given for areas with higher diversity. In this case Bench Maji-1(Maji) and Dawero Banja forest should be targeted for conservation and further collection effort for coffee improvement program at national level. In the case of Limmu, the expansion of state frames and strong coffee agricultural extension program could lead to gradual replacement of the local populations with that of introduced. In this study a limited number of sample size was represented from Bench Maji-2, Anfilo and Limmu populations. This could undermine the levels of genetic diversity observed in this population and could be the main reason for the low genetic diversity in Bench Maji-2,Anfilo and Limmu areas.

7. CONCLUSIONS

The present study was conducted with the main objective of assessing the extent and distribution of genetic diversity of wild coffee (*C. arabica*) in Ethiopia using Inter Simple Sequence Repeat (ISSR) marker. In the present study, four ISSR primers (three di and one tetra-nucleotide) were employed. They were able to reveal the existence of high genetic diversity and identify highly diverse and least diverse populations in south and south west Ethiopia.

High genetic diversity was observed in both Bench Maji-1 and Dawero Banja populations whereas low genetic diversity was observed in Anfilo, Bench-Maji-2(Jeba) and Limmu populations. These assessments of the levels and extents of genetic diversity indicated that both Bench Maji-1 and Dawero Banja forests are valuable sites for *Coffea arabica* conservation.

Analysis of molecular variance for the population studied showed that the highest proportion of genetic variation was attributed to within population than among populations. This shows presence of gene flow that caused by human interference and other factors. The UPGMA and NJ clustering methods of seventy-one individuals of coffee for the overall four primers, three di- and one tetra nucleotides, did not produce exactly the same tree topology. In two coordinates almost all individuals of each population were spread all over the plot, but in 3D, individuals from Dawro Banja and Bench Maji population showed some tendency of grouping while others spread all over the space. Over all, 3D PCO plot and UPGMA analysis showed better resolutions than 2D and NJ.

8. RECOMMENDATIONS

- The present study was based only on ISSR markers. Hence, further analysis using co-dominant markers system, like microsatellites and SSR should be conducted for better understanding of the genetic diversity of wild coffee and estimating the gene flow, population size and levels of inbreeding.

- There is a need for further collection of additional accessions to better represent populations or regions. In line with the finding of this study, prior consideration should be given to Bench Maji-1 and Dawero Banja populations that show higher genetic diversity. In addition to targeting areas of high genetic diversity, the limited collections from areas of low genetic diversity (Anfilo, Bench-Maji-2(Jeba) and Limmu) should be supplemented with additional collections.

- In spite of its significant social and economic values for local, national, and global contribution, the coffee forest ecosystem is threatened by deforestation. Therefore, the Institute of Biodiversity Conservation and regional governments should take action to safeguard the last remaining coffee forest patches of the country, before they become completely lost.

9. REFERENCES

Admas, A. (2011). Genetic diversity of sesame (*Sesamum indicum*) from north western ethiopia using inter simple sequence repeate markers. MSc Thesis Presented to the School of Graduate Studies of the Addis Ababa University.

Aga, E., Bryngelsson, T., Bekele E., and Salomon, B. (2003). Genetic diversity of forest arabica coffee (*Coffea arabica*) in Ethiopia as revealed by random amplified polymorphic DNA (RAPD) analysis. *Hereditas* **138:** 36–46.

Aga, E. (2005). Molecular genetic diversity study of forest coffee tree (*Coffea Arabica* L.) populations in Ethiopia: Implications for conservation and breeding, Faculty of Landscape Planning, Horticulture and Agricultural Science, Swedish University of Agricultural Sciences (SLU).

Anthony, F., Bertrand, B., Quiros, O., Wilches, A., Lashermes, P., Berthaud, J. and Charrier, A. (2001). Genetic diversity of wild coffee trees (*Coffea arabica* L.) using molecular markers. *Euphytica.* **118:** 53-65.

Anthony, F., Combes, M.C., Astorga, C., Bertrand, B., Graziosi, G., and Lashermes, P. (2002). The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theor. Appl. Genet.***104:**894-900.

Arriel N. H.C. , Mauro A.O.D., Arriel E.F., Unêda-Treviso S. H., Costa M.M., Bárbaro I. M., and Muniz, F. R. S. (2007). Genetic divergence in sesame based on morphological and agronomic traits. *Crop Breeding and Applied Biotechnology*, **7:** 253-261

Assefa, K. (2003). Phenotypic and molecular diversity in the Ethiopian cereal, Tef [*Eragrostis tef* (Zucc.) Trotter]. Doctoral Dissertation, Department of Crop Science, SLU. Acta Universitatis Agriculturae Sueciae.Agraria vol. 426.

Balami, S. (2007). Genetic Diversity Analysis of the Wild *Coffea arabic*a L. Populations from Harenna Forest, Bale Mountains of Ethiopia, Using Inter Simple Sequence Repeats (ISSR) Marker. MSc Thesis Presented to the School of Graduate Studies of the Addis Ababa University, Addis Ababa.

Berthaud, J and Charrier, A. (1988). Genetic Resources of *Coffea*. In: R.J. Clarke & R. Macrae (eds) *Coffee Vol*. 4 Agronomy, Elsevier Applied Science, London & New York, pp. 1-42.

Blair, M.W., Panaud, O., and McCouch, S. R. (1999). Inter-simple sequence repeats (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa L.*). *Theor. Appl. Genet*. **98**:780-792.

Borsch, T., Hilu, K.W., Quandt, D., Wilde, V., Neinhuis, C. and Barthlott, W. (2003). Non coding plastid trnT-trnF sequences reveal a well resolved phylogeny of basal angiosperms. *J. Evol. Biol.* **16**: 558-576.

Brar, D.S., and Mohan, J.S. (2010). *Molecular Techniques in Crop Improvement.* University of Helsinki, Finland.

Bremer B, and Jansen RK. (1991). Comparative restriction site mapping of chloroplast DNA implies new phylogenetic relationships within *Rubiaceae*. *Am J Bot* **78:**198–213

Bridson, D.M., (1987). Nomenclature on *Psilanthus*, including coffee sec. *Paracoffea* (*Rubiaceae* tribe *Coffeeae*). *Kew Bulletin*.**42**: 453 – 460.

Bridson, D.M. and Verdcourt, B., (1988). *Rubiaceae* part 2. **In:** *Flora of Tropical East Africa*. The Netherlands, pp. 703 – 727.

Brown, A.H.D. (1978). Isozymes plant population genetic structure and genetic conservation. *Theor Appl. Genet.* **52:**145-157.

Brown, A.H.D. (1990). The role of isozyme studies in molecular systematics. **In**: P.Y.Ladiges and L.W. Martinelli (Eds.), *Plant Systematics in the Age of Molecular Biology*, pp. 39-46. CSIRO, Melbourne.

Chaparro, A. P., Christancho, M.A., Cortina, H.A., and Gaitan, A.L. (2004). Genetic variability of *Coffea arabica* L. accessions from Ethiopian evaluated with RAPDs. *Genet. Resour. Crop. Evol*. **51**:291-297.

Chevalier, A. (1947). Les cafeiers du globe. Systematique des cafeiers et faux cafeiers.Malades et insectes nuisibles. Encycl. Biol. XXVII, Fas III, P. Lechevalier, Paris, 356 p.

Chowdhury, M.A., Vandenberg, B. and Warkentin, T. (2002). Cultivar identification and genetic relationship among selected breeding lines and cultivars in chickpea (*Cicer arietinum* L.). *Euphytica.* **127:** 317-325.

Coste, R. (1992). Coffee; *The Plant and the Product.* CTA Waginingen, The Netherland.

Coulibaly, I., Noirot, M., Lorieux, M. &Charrier,A. 2002.Introgression of self compatibility from *Coffea heterocalyx* to cultivated species *Coffea canephora*. *Theor. Appl. Genet.***105:** 994-999.

Cruz ARR, Paixao ALD, Machado, FR, Barbosa, MFF, Junqueira, CS. (2004). Metodologia para obtenção de plantas transformadas de *Coffea canephora* por co-cultivo e calosembriogênicoscom *A.tumefaciens*.Boletim de Pesquisa e Desenvolvimento. Embrapa, Brasilia, Brasil,vol.58:15 p

Davis, A.P., Chester, M., Maurin, O. and Fay, M.F. (2007). Searching for the relatives of *Coffea* (*Rubiaceae, Ixoroideae*). The circumscriptions and phylogeny of *Coffea* based on plasmid sequence data and morphology. *Am. J. Bot.* **94 (3**): 313 – 329.

Davis, A.P., Govaerts, R., Bridson, D.M. and Stoffelen, P.(2006). An annotated taxonomic conspectus of the genus *Coffea (Rubiaceae*). *Botanical Journal of theLinnean Society*.**15:** 465 – 512.

Dubale, P. (1996). Availability of phosphorous in the coffee soils of southwest Ethiopia. **In**: Teshome, Y., Eyasu, M., Mintesinot, B., (Eds).Proceeding of the Third Conference of Ethiopian Society of Soil Science, Ethiopia. Science and Technology Commission, Pp 119-127.

Dubale, P. and D. Teketay (2000). The need for forest coffee germplasm conservation in Ethiopia and its significance in the control of coffee diseases. **In**: *TheProceedings of the Workshop on Control of Coffee Berry Disease in Ethiopia,* Ethiopian Agricultural Research Organization (EARO).Addis Ababa. Pp 125-135.

Dubale, P. and Shimber, T. (2000). Some ecological parameters occurring in the major coffee growing areas of southwestern and southern Ethiopia. **In**: EARO (Ed), *Proceeding of Coffee Berry Disease Workshop*, Ethiopian Agricultural Research Organization, Addis Ababa, Ethiopia, Pp 107 – 124.

Dudnikov, A.J. (2003). Allozymes and growth habit of *Aegilops tauschii*: genetic control and linkage patterns. *Euphytica*, **129**: 89-97.

Excoffier, L.; Smouse, P.E. and Quattro, J.M. (1992) Analysis of Molecular Variance inferred from metric distances among DNA haplotypes **In:***Application to human mitochondrial DNA restriction data*. ***Genetics,*****131:**479-491.

Excoffier, L., (2006). An Integrated Soft Ware Package for Population Genetics Data Analysis. Computational and Molecular population Genetics Lab.(MPG), Institute of Zoology, University of Bern, Switzerland.

Fang, D.Q. and Roose, M.L. (1997). Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theor. Appl. Genet.***95:**408–417

Fikru, E. (2006). Morphological and molecular diversity in the Ethiopian lentil (*Lensculinaris medikus*) land race accessions and their comparison with some 50 exotic genotypes. MSc Thesis Presented to the School of Graduate Studies of the Addis Ababa University.

Food and Agricultural Organization., (1968). FAO Coffee Mission to Ethiopia, 1964-1965. Report FAO, Ed., Rome (Italy).

Gatzweiler, F.W. (2005). Institutionalising biodiversity conservation –The case of Ethiopian coffee forests. *Conservation and Society* .**3(1):** 201-223.

Girma, G. (2007). Relationship between wild rice species of Ethiopian rice with cultivated rice based on issr marker. MSc Thesis Presented to the School of Graduate Studies of the Addis Ababa University.

Godwin, I.D., Aitken, E. A.B. and Smith, L.W., (1997). Application of inter simple sequence repeat (ISSR) markers to plant genetics. *Electrophoresis* **18:**1524-1528.

Gole, T.W. (2003). Conservation and Use of *Coffea* Genetic Resources in Ethiopia: *Challenges and Opportunities in the Context Current Global Situations.* Ecology and Developmental Series.

Gole, T.W., Denich, M., Teketay, D. and Borsch, T. (2001). Diversity of traditional coffee production systems in Ethiopia and their contributions to the conservation of coffee genetic diversity. Conference on International Agricultural Research for Developpment; *Deu, Scher Tropentag,* Bonn, 9-11 October.

Gole, T.W., Denich, M., Teketay, D. Vlek, P.L.G. (2002).Human impacts on *Coffea arabica* gene pools in Ethiopia and the need for its *in situ* conservation. **In**: *Managing Plant Genetic Diversity*. CABI andIPGRI, 237-247.

Gole, T. W.M.and Teketay, D. (2001). The forest coffee ecosystems: ongoing crises, problems and opportunities for gene conservation and utilization. **In**: *Imperative Problems Associated with Forestry in Ethiopia,* Biological Society of Ethiopia. Pp 131–142.

Gole T.W.(2003). Vegetation of the Yayu forest in SW Ethiopia: *Impacts of Human use and Implications for in situ Conservation of Wild Coffea arabica L. populations. Ecoloyg and Development Series*, No.10, Zentrum fur Entwicklungsforschung, Center for Development Research, University of Bonn.

Gole, T.W., Borsch, T., Denich, M. and Teketay, D. (2008). Floristic composition and environmental factors characterizing coffee forests in southwest Ethiopia. *Forest Ecol. Management.* **255:**2138-2150.

Gottlieb, L.D. (1981). Electrophoretic evidence and plant populations. *Prog. Phytoch.***7:** 1 – 46.

Gupta, M., Chyi, Y.S., Romero-Sevenson, J. and Owen, J.L. (1994). Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theor. Appl. Genet.* **8:** 998-1006.

Hammer, O., Harper, D.A.T. and Ryan, P.D. (2001). PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4:9, http://palaeo-electronica.org/2001_1/past/issue1_01.htm**.**

Jaccard.(1908).Nouvelles researchers Sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat*. **44:** 223 – 270.

Joshi, S.P., Prabharkar, K., Ranjek, A., Vidya, R and Gupta, S. (1999). Molecular markers in plant genome analysis. Plant Molecular Biology Group, Division of Biochemical Science, Natural Chemical Laboratory, Pune 411008, India.

Joshi, S.P., Gupta, V.S., Aggarwal, R.K., Ranjekar, P.K. and Brar, D.S. (2000). Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theor Appl Genet*. **100:**1311-1320.

Karp, A., Seberg, O. and Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. *Annals of Botany* **78:**143–149.

Karp, A., S. Kresovich, K.V., Bhat, W.G. and Ayad, T.H., (1997). *Molecular Tools* **In***: Plant Genetic Resources Conservation: A Guide to the Technologies. IPGRI Technical Bulletin No. 2*. International Plant Genetic Resources Institute, Rome,Italy.

Lashermes, P., Comes, M.C., Cros, J., Trouslot, P., Anthony, F. and Charrier, A. (1995). Origin and genetic diversity of *Coffea arabica*L. Based on DNA molecular markers: **In**: *Proceedings of the 16 ASIC Colloquium (Koyoto),*pp. 528-535.ASIC, Paris, France.

Lashermes, P., Trouslot, P., Anthony, F., Comes, M.C. and Charrier, A. (1996). Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. *Euphytica.* **87:** 59-64.

Lashermes, P., Combes, M.C., Robert, J., Trouslot, P. and Carries, A. (1997). Phylogenetic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. *Theor. Appl. Genet*. **94:** 947 – 955).

Lashermes, P., Comes, M.C., Robert, J., Trouslot, P., D'Hont, A., Anthony, F. and Charrier, A. (1999). Molecular characterization and origin of the *Coffea arabica* L. genome. *Mol. Gen. Genet.* **261:** 259-266.

Lewontin, R.C. (1972). The apportionment of human diversity. *Evol. Biol*. **6:** 381–398

Lowe, A.J., Moule, C., Trick, M., and Edwards, K.J. (2004). Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species, *Theor. Appl. Genet*. **108**: 1103–1112.

Marshal, D.R. and Bron, A.H.D. (1975). Optimum sampling strategies in genetic conservation, **In:** *Crop Genetic Resources for Todays and Tomorrow*. Pp 53 – 80\ (Frankel, O.H. and Hawkes, J.G. (eds)). IBP2, Cambridge University Press, Cambridge.

Martins, M., Tenreiro, R. and Oliveira, M. (2003). Genetic relatedness of Portuguese almond cultivars assessed by RAPD and ISSR markers. *Plant Cell Reports*. **22:** 71–78.

Meyer, G.F. (1965). Notes on wild *Coffee arabica* from South Western Ethiopia: with some historical considerations. *Econ. Bot.* **19:** 136-151.

Mitiku, A. (2011). Inter Simple Sequence Repeats (ISSRs) Fingerprinting, Phenotypic Variability and Trait Associations in Released and Elite Rice (*Oryza sativa* L.) Genotypes of Ethiopia. MSc Thesis Presented to the School of Graduate Studies of the Addis Ababa University.

Montagnon, C. and Bouharmont, P. (1996). Multivariate analysis of phenotypic diversity of *Coffea arabica* genetic resource. *Crop Evol.* **43:** 221-227.

Montarroyos, A.V.V., Lima, M.A.G. de Santos, E.O. dos and Franca, J.G.E. de (2003). Isozyme analysis of an active cassava germplasm bank collection. *Euphytica.* **130**:101-106.

Morell, M.K., Peakall, R., Appels, L.R. and Preston Lloyd, H.L. (1995). DNA profiling techniques for plant variety identification. *Australian Journal of Experimental Agriculture* **35:** 807-819.

Narasimhaswamy, R.L. (1962). Some thoughts on the origin of *Coffea arabica* L. *Coffee*: **4 (12):** 1-5.Turrialba, Costa Rica.

Oljira, T. (2006). Genetic Diversity of The Wild Populations of *Coffea arabica* L. in Yayu Forest of Ethiopia Using Inter Simple Sequence repeat (ISSR) marker, M.Sc thesis, Addis Ababa University, Addis Ababa, Ethiopia.

Paillard, M., Lashermes, P. and Petiard, V. (1996). Construction of a molecular linkage map in coffee. *Theor. Appl. Genet.* **93**: 41-47.

Parker, P.G., Snow, A.A., Schug, M.D., Booton, G.C. and Fuerst, P.A. (1998). What molecules can tell us about populations: choosing and using molecular markers. *Ecology*. **79:**361–382.

Parsons, B.J., Newbury, H.J., Jackson, M.T. and Ford-Lloyd, B.V. (1997). Contrasting genetic diversity relationships are revealed in rice (*Oryza sativa* L.) using different marker types. *Molecular Breeding.* **3:**115–125.

Pavlicek, A., Hrda, S. and Flegr, J. (1999). Free tree free ware program for construction of phylogenetic trees on the basis of distance data and bootstrap/Jack Knife analysis of the tree robustness. Application in the RAPD analysis of genus.

Purseglove, J. (1968). *Tropical Crops: Dicotyledons*. Harlow, UK: Longman Group Ltd., pp 458-492.

Raina, S.N., Mukai, Y. and Yamamato, M. (1998). In situ hybridization identifies the diploid progenitor species of *Coffea arabica* (Rubiaceae). *Theor. Appl. Genet.* **97:** 1204-1209.

Rohlf, F.J. (2000). NTSYS-PC, Numerical Taxonomy and Multivariate Analysis System, Version 2.11.Exeter Software New York.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*. **4:** 406–425.

Schlötterer, C. (2004). The evolution of molecular markers—just a matter of fashion? *Nature Reviews Genetics*, **5:**63–69.

Senbeta, F. (2006). Biodiversity and Ecology of Afromontane Rainforests with Wild *Coffea arabica* L. Populations in Ethiopia: PhD thesis, *Ecology and Development Sreries*, No 38, Cuvillier Verlag Göttingen, University of Bonn, Germany.No.38.

Senbeta, F. and Denich, M. (2006). Effects of wild coffee management on species diversity in the afromontane rain forest of Ethiopia. *Forest Ecology andManagement*, **232**:68–74.

Silvarolla, M.B., Mazzafera, P. and Fazuoli, L.C. (2004). A naturally decaffeinated Arabica coffee. *Nature*, **429:** 826.

Sneath, P.H.A. and Sokal, R.R. (1973). *Numerical Taxonomy* .Freeman. San Francisco. Pp573.

Steiger, D.L., Nagai, C., Moore, P.H., Morden, C.W., Osgood, R.V. and Ming, R. (2002). AFLP analysis of genetic diversity within and among *Coffea arabica* cultivars. *Theor. Appl. Genet.* **105**: 209-215.

Statistica Stat Soft, Inc.2001. STATISTICA (data analysis software system) Version 6.0.www.statsoft.com.

Studier, J.A. and Keppler, K.J. (1988). A note on the neighbor-joining algorithm of Saitou and Nei. *Mol. Bol. Evol*, **5:** 729-731.

Sylvain, P.G. (1958). Some observations on *Coffea arabica* L. in Ethiopia. *Turrialba,* **5:** 37-53.

Teketay, D., Anaga, A., Mulat, G. and Enyew, M. (1998). Study on forest coffee conservation. CIP/CTA, Addis Ababa.

Teketay, D. (1999). History, botany and ecological requirements of Coffee. *Walia,* **20:** 28-50.

Templeton, A.R. (1993). Translocation as conservation tool. **In:** Szaro R. (Ed.), *Biodiversity in Mangrove Landscapes, Theory and Practice.* Oxford University Press.

Tesfaye, K.G. (2006). Genetic diversity of wild *Coffea arabica* populations in Ethiopia as a contribution to conservation and use planning. Ecology and development series no. 44: Doctoral thesis, University of Bonn, Germany.

Tesfaye, K.G., Thomas, B., Kimm.G. and Bekele.E. (2007). Characterization of *coffea* chloroplast microsatellite and evidence for the recent divergence of C *arabica* and C. *eugenioides* chloroplast genome .Nees-Inistitute for Biodiversity of Plants, University of Bonn.

Teshome, D.W. (2011). Genetic diversity of sesame germplasm collection by ISSR marker implication for conservation and improvement, Master thesis, Addis Ababa University, Addis Ababa.

Tsegaye, Y., Getachew, O. and Tesfaye, Z. (2000). Some socio-economic issues related to fungicide use against CBD in Ethiopia. **In:** *The Proceedings of Coffee Berry Disease (CBD) Workshop in Ethiopia*, pp. 72-84. Ethiopian Agricultural Research Organization (EARO), Addis Ababa

Vander Vossen, H.A.H. (2001). Agronomy I: Coffee breeding Practices. **In:** *Coffee Recent Development*, Blackwell Science Ltd, London, pp.184-195.

Vos, P., Hogers, R., Bleeker, M., Rijans, M., Vande Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. **23**:4407-441

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). *DNA Fingerprinting in Plants: Principles, Methods and Applications*, second edition. CRC Press, Taylor and Francis Group.

Whitkus, R., Doebley, J. and Wendel, J.F. (1994). Nuclear DNA markers in systematics and evolution. **In:** *DNA-based Markers in Plants (Advances in Cellular and Molecular Biology of Plants, vol. 1.)* (RL Phillips and IK Vasil, eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 116–141.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, **18:** 6531–6535

Wilson, K.C. (1985). Climate and soil, **In:** Clifford, M.N. and Willson, K.C.(Eds), *Coffee: Botany, Biochemistry and Production of Beans and Beverages.* Crown Helm, London and Sydeny. Pp 97 – 107.

Woldetsadik, W. and Kebede, K. (2000). Coffee production systems in Ethiopia, **In:** *Proceedings of the Workshop on Control of Coffee Berry Disease in Ethiopia Addis Ababa* (Ghion Hotel) 13-15 August 1999. Addis Ababa, Ethiopia.

Wu, K., Jones, R., Dannaeberger, L. and Scolnik, P.A. (1994). Detection of microsatellite polymorphisms without cloning: *Nucleic Acids Res.***22:** 3257-3258.

Yang, R.C., Boyle, T.J.B., Ye Z. and Mao, J.X. (1999). POPGENE, the user friendly shareware for population genetics analysis, version 1.31, Molecular Biotechnology center, University of Aleberta, Canada.

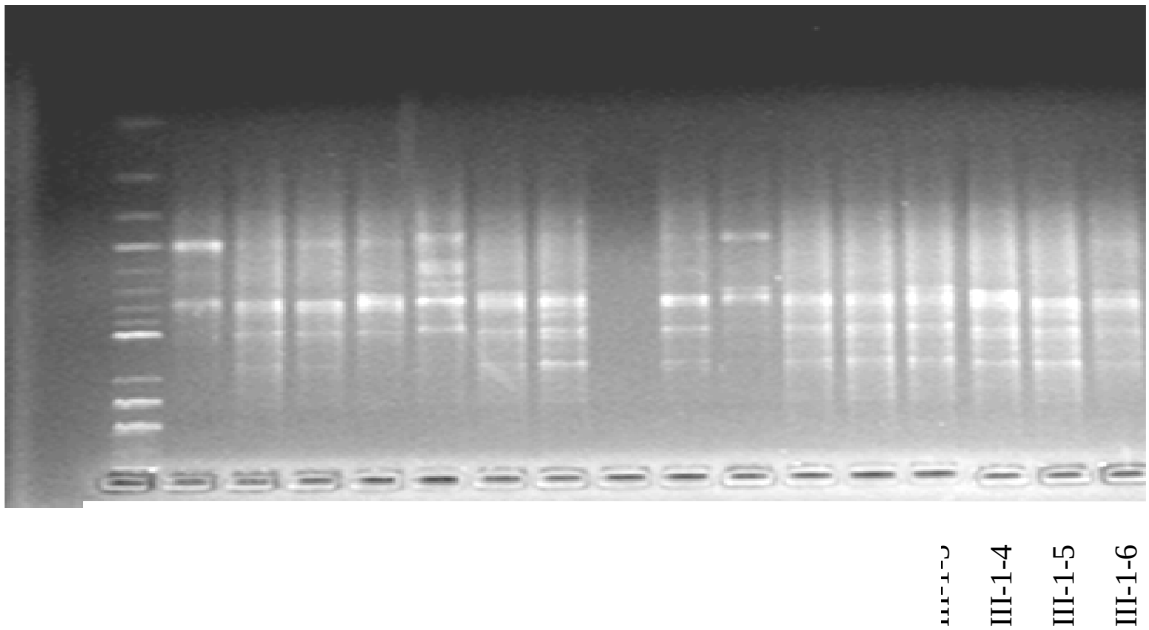
Yeh, F.C., Yang, R-Cai. and Boyle, T. (1999). Population genetic analysis of co-dominant markers and qualitative traits. *Belgian J. Bot*. **129:**157.

Yeshitila, K. (2001). Loss of forest biodiversity associated with changes in land use: The case of Chewaka-Utto Tea Plantation. **In**: *Imperative Problems Associated with Forest in Ethiopia*, Biological Society of Ethiopia, pp 115 – 122.

Zietkiewicz, E., Rafalski, A. and Labuda, D. (1994). Genomic fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, **20**:176-183.

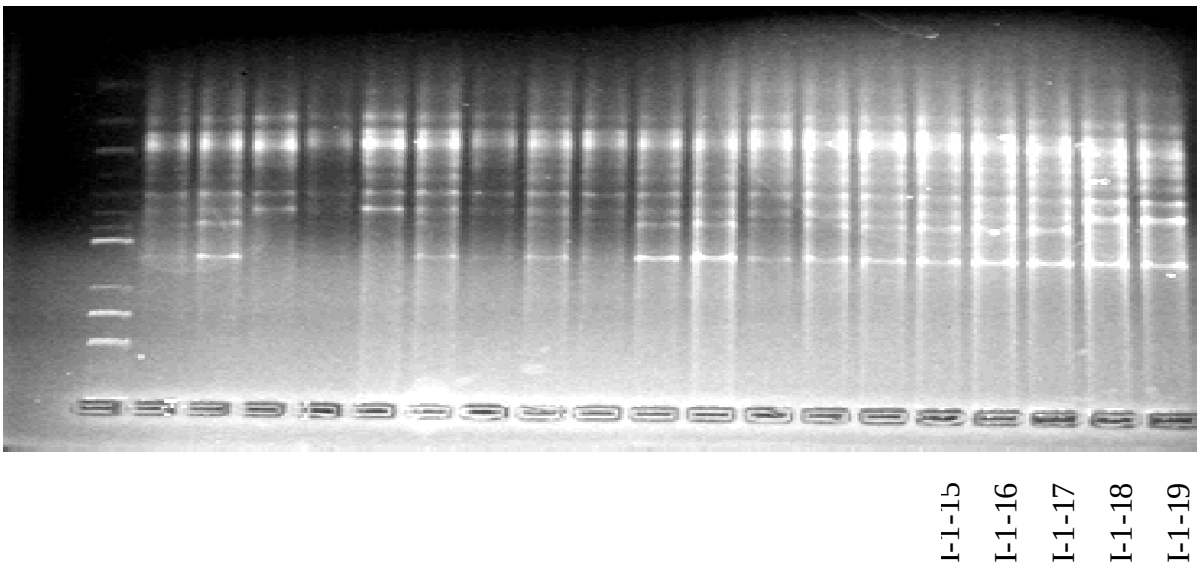
10. APPENDICES

Appendix 1



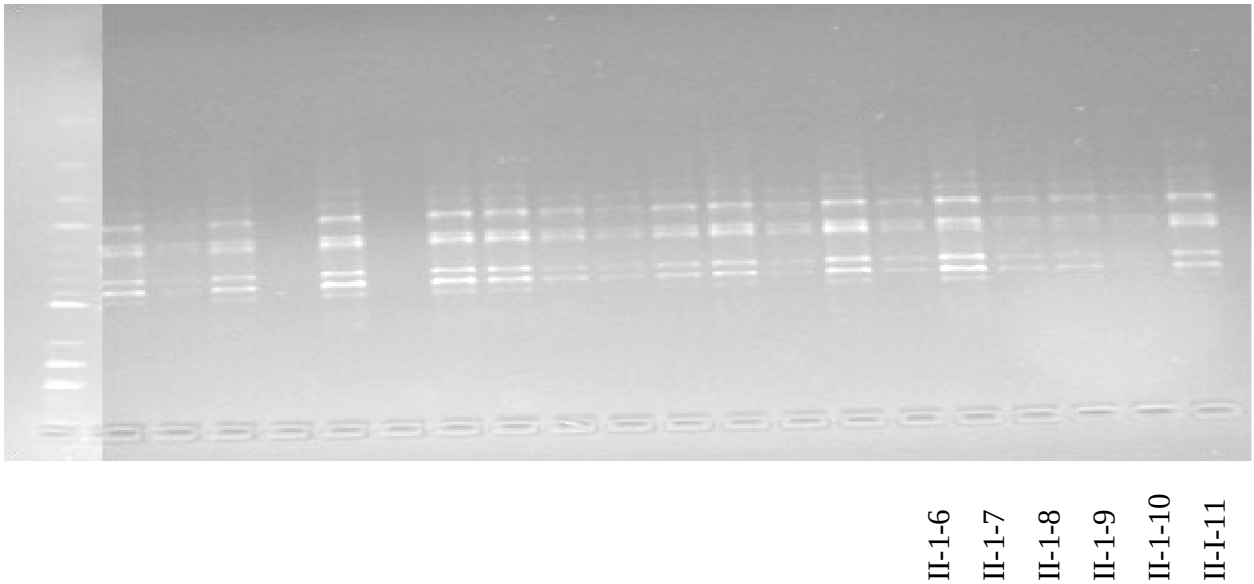
ISSR fingerprint generated from some individuals of wild *Coffea arabica* using primer 810.

Appendix 2



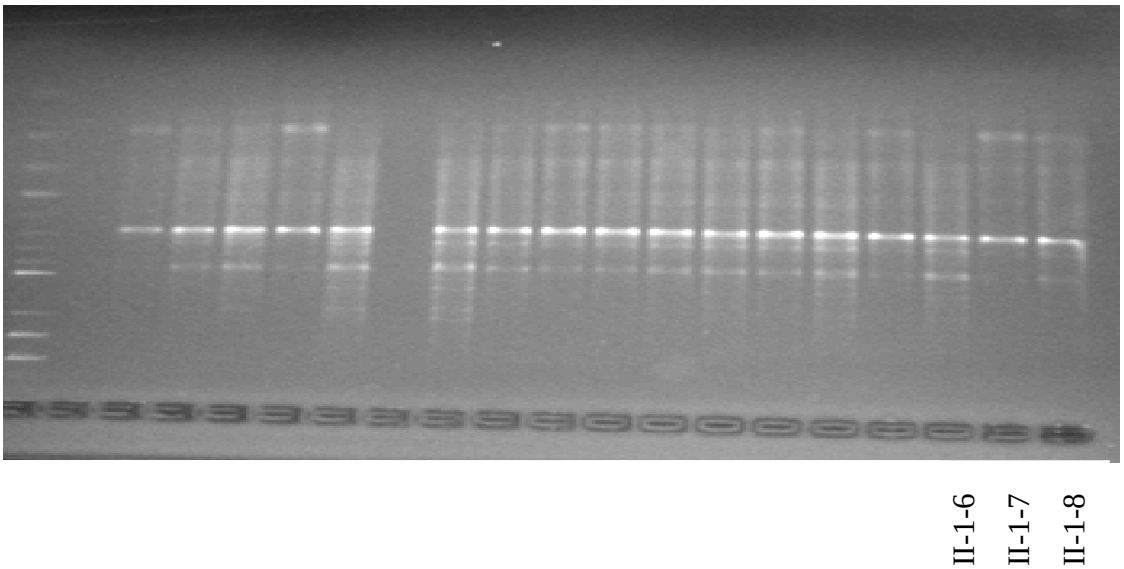
ISSR fingerprint generated from some individuals of wild *Coffea arabica* using primer 812.

Appendix 3



ISSR fingerprint generated from some individuals of wild *Coffea arabica* using primer 818

Appendix 4



ISSR fingerprint generated from some individuals of wild *Coffea arabica* using primer 880.

Declaration

I, the undersigned declare that this thesis is my original work and that all sources of material used for the thesis have been correctly acknowledge.

Name: Biruk Getenete

Signature: ----------

Date of submission: 06/06/12

Advisor: Dr. Kassahun Tesfaye

Sign: ----------

Date: ----------