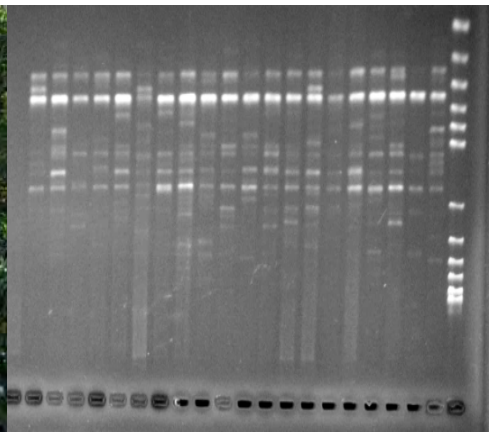
ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF BIOLOGY

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



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

Solomon Balami

ADDIS ABABA

JULY, 2007

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF BIOLOGY

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

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

BY

SOLOMON BALAMI

JULY, 2007

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF BIOLOGY

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

A Thesis Presented to the school of Graduate Studies of the Addis Ababa University in Partial Fulfillments of the Requirements for the Degree of Master of Science in Biology (Applied Genetics)

BY

SOLOMON BALEMI

Approved by Examining Board

…….. …….
Examiner

…….. . ……
Examiner

…… ……..
Chairman
Prof. Endashaw Bekele (Advisor) ……

…… …… ……... ……

Dr. Kasahun Tesfaye (Co-Advisor)

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to express my special thanks to my advisor, Prof. Endashaw Bekele, who helped me to be engaged in this research program and accepted me to work in his laboratory and made possible my dreams of working M. Sc. thesis in the area of molecular genetics. I really appreciate his unreserved support and insightful criticism in commenting on the manuscript.

I would like to extend my special thanks to my co-advisor, Dr. Kassahun Tesfaye, for his meticulous or valuable guidance, keen interest, and encouragement and constructive criticism starting from the proposal development to thesis write up.

I also extend my gratitude towards CoCE Project (Conservation and Use of the Wild Populations of *Coffea arabica* in the Montane Rain Forest of Ethiopia) for accepting me as M.Sc. students to be involved in this project and covering the whole cost of the research by supplying all necessary materials needed.

I am indebted to the School of Graduate Studies of the Addis Ababa University, Department of Biology for accepting me to join post graduate program and provide financial support for the study.
I am also indebted to Oromia Education Bureau and Jimma Teachers College for giving me the chance to join postgraduate study and paying me salary during the study time.
I am grateful for encouragement and support of my colleague Gezahegn Girma.

I am also grateful to my parents, wife, brothers, sisters, relatives, for their encouragement and support. Their all rounded and unconditional support enabled me to realize my educational goal.

Above all, I praise and glorify my only Almighty God, Waaqayyoo who offered me everything and whose help made all possible.

LIST OF TABLES **Pages**

LIST OF FIGURES **Page**

LIST OF ABBREVIATIONS.

AFLP = Amplified Fragment Polymorphisms

AMOVA = Analysis of molecular variance

CoCE = Conservation and Use of the Wild Populations of *Coffea arabica* in the Montane Rain Forest of Ethiopia

GRV = Great Rift Valley

ISSR = Inter simple sequence repeat

NJ = Neighbor Joining

PCO = Principal Coordinates

D = Dimensional

PL = Polymorphic loci

% PL = Percentage of Polymorphic loci

RAPD = Random Amplified Polymorphic DNA

RFLP = Restriction Fragment Length Polymorphism.

SSR = Simple Sequence Repeat

SW = South west

SE = South east

UPGMA = Unweighted pair- group methods using the arithmetic averages**.**

ABSTRACT

In spite of their importance, the level and distribution of wild *Coffea arabica* L. in Harrena has not been extensively examined in depth with molecular markers. The current study was conducted on the wild Arabica coffee of the Harenna Forest of Bale Mountain, southeast part of Ethiopian. The levels and distribution of within and among populations genetic diversity of one hundred wild *Coffea arabica L* individuals, representing four populations: two from semi-disturbed (Bale-I and Bale-3) and two from undisturbed (Bale-4 and Bale-6) regions of the forest, were collected and evaluated with Inter Simple Sequence Repeats (ISSR). A total of nine primers which contain different simple sequence repeat (SSR) were used and tested for PCR amplifications. A total of one hundred thirty seven bands were detected. The number of bands per ISSR markers ranges from ten to twenty one with an average of 15.2. These were then used to estimate the genetic diversity. Out of the total bands produced, 61(44.53%) were polymorphic and the number of polymorphic bands per ISSR markers ranges from one (1) to nineteen (19), with an averages of 6.77. The similarities between individual genotypes were estimated using UPGMA and NJ analysis. The populations were found to be clustered on the basis of their respective origin. The UPGMA cluster analysis showed that the four populations form two major clusters (undisturbed and semi-disturbed populations together) according to locations from which they were collected. The two major clusters further divided into two. Analysis of molecular variance (AMOVA) indicated that population level genetic diversity was relatively high (56.8%). Shannon's diversity index showed the same patterns and indicated that the within and between populations genetic diversity of *Coffea arabica* L. populations are significantly different. A considerable proportion (83.6 %) of the total genetic diversity was distributed within populations (i.e., due to differences among individuals within the populations). 16.4% of the total genetic diversity was due to differences among populations. Such distributions of the total genetic diversity could be attributed to gene flow via insect pollinators, seed flow by wild animals, birds and human. Generally, on the basis of samples of 137 bands in the four populations, ISSR was able to reveal moderate to high levels of genetic diversity within and among wild coffee populations of Harenna Forest of Bale Mountain.

Key Words: *Coffea arabica L*, ISSR marker, Harenna, Intra-regional analysis, Ethiopia

1. GENERAL INTRODUCTION

1.1. Back ground and Justification

The origin of *Coffea arabica* L. is south west and south east Ethiopia and it is here where it grows naturally as a small tree in the undergrowth of montane rain forest (Denich *et al.*, 2006). It is considered as the only wild *Coffea* species found in Ethiopia, being naturally restricted in the two isolated mountain forest on the western and eastern sides of the Great Rift Valley in the southern parts of the country (Gole, 2003).

More than 100 species are found in the genus *Coffea* out of which three are economically important. According to Bremer (1996), *Coffea* are members of the family Rubiaceae which contains some 640 genera and about 10000 species. The genus *Coffea* comprises more than 103 well identified species including the two *Coffea* species of economic importance: *Coffea arabica* and *Coffea canephora* (Bridson and Verdcourt, 1988; Charrier and Berthaud, 1985; Davis *et al.*, 2006*).* Eighty percent of the global coffee production comes from *Coffea arabica* L. because of its best cup quality and a wide choice of flavor, while *Coffea canephora* contributes nearly the remaining 20 percent. *Coffea liberica* has only minor importance in some localities of West Africa (Bridson and Verdcourt, 1988). It was also reported that *Coffea arabica*, (that occurs in the mountain forest of the western and eastern side of the Great Rift Valley) accounts for a greater proportion than the *Coffea canephora* in satisfying the world coffee demand (Gole, 2003). *Coffea arabica* and *Coffea canephora* are the two commercially most important species of *Coffea* that play a crucial role in the global market demand of coffee. It is the most important source of countries foreign currency earning for more than eighty developing countries (Vega *et al.*, 2003). For instance Ethiopia used to get up to 67 percent of its foreign currency from this commodity (Gole, 2003). This implies that agricultural-based Ethiopian economy is highly dependent on *Coffea arabica* as it contributes high percent of the country's foreign exchange earning's. *Coffea arabica* L. is considered to be allotetraploid (2n=4x=44) which is self fertile/autogamous, where as all the other *Coffea* species including *Coffea canephora* are all diploid (2n=2x=22) which are generally self incompatible. Recent

investigations indicate that *Coffea arabica* is an amphidiploid species containing two genomes that was considered to originate from two related diploid species, *C. eugenioides* and *C. canephor*a (Lashermes *et al*., 1999). However, resent work of Tesfaye (2006) showed that both *Coffea arabica* and *C. eugenioides* have common maternal parents.

Genetic diversity is crucial for the survival of wild *Coffea* species. Ethiopia is the only centers of origin and diversification of *Coffea arabica* and thus it is an important source of *Coffea* genetic resources of the world *Coffea* industry (Gole, 2003).This is mainly attributed to its diverse ecological features such as suitable altitude, ample rain forest, optimum temperature, fertile soils etc. (Yeshitila *et al.,* 2004). For instance, *Coffea arabica* naturally grows wild in the eastern part of the country, Bale Mountain (Harrena forest) which has an altitude 1513m.a.s.l (Tesfaye, 2006; Senbeta, 2006). Although Ethiopia is the centers of origin and diversification of *Coffea arabica*, the gene pools of wild *Coffea arabica* populations are severely endangered because of over unsustainable utilization coupled with rapid population growth, which is the root cause of deforestation due to demands for agricultural and settlement areas (Denich *et al.*, 2006; Senbeta, 2006; Gole, 2003). This has aggravated the erosions of genetic diversity and is increasingly likely to erode the last wild coffee (Denich *et al*., 2006). The country is the center of origin and diversification of *Coffea arabica*, where genetic resource conservation could be made and forest biodiversity could be maintained in their natural environments. However, the current situation of deforestation and land use change in SW and SE part of Ethiopia affects not only people that depends on the coffee production, but also the *Coffea* genetic resources and the biodiversity of mountain rain forest as well (Senbeta, 2006; Gole, 2003).Whether to develop effective conservation strategies or for utilization of genetic resources in the future, efficient and fast molecular markers based evaluation of the extent and distribution of *Coffea arabica* genetic diversity is required (Tesfaye, 2006; Gonzalez *et al.,* 2005). Molecular markers are the most important approaches to study the genetic diversity of plant species that specially suffers from human disturbance (Faria and Miyaki., 2006), as they provide the knowledge about the extent and distributions of C*offea arabica* genetic diversity, which is in turn very important to conserve as well as to utilize the available genetic recourse in a sustainable ways (Masumbuko *et al*., 2005).

A variety of molecular markers have been used in studying the genetic diversity of *Coffea* ranging from Isozyme systems to DNA-based systems. However, use of isozyme techniques on wild *C. arabica* species from Ethiopia failed in revealing polymorphism indicating that, isozyme as being not appropriate for diversity study in *Coffea arabica* (Louarn, 1978). In recent years, a number of DNA-based marker systems have been developed using PCR techniques (Gupta, 2003) to study genetic diversity of plants including *Coffea arabica* L., and hence to overcome constraints of isozyme based technique. The major PCR-based marker systems that are currently available for genetic study are Amplified Fragment Length Polymorphism (AFLP), Rapid Amplified Polymorphic DNA (RAPD) and Inter Simple Sequence Repeat (ISSR) which are generally described as dominant markers (Weising *et al.,* 2005). Due to the nature of the loci being found scattered throughout the genome, these dominant markers are generally very useful for assessing genetic diversity and sampling the whole genome randomly (Lowe *et al*., 2004; Weising *et al*., 2005). This is because the loci scattered through the entire genome can be rapidly genotyped using PCR techniques (Twyman, 2003). However, each of this PCR-based marker system has its own advantages and disadvantages. Hence, the choice of marker systems is important for population genetic studies depending on the objectives of the research, cost, level of diversity expected etc. (Tesfaye, 2006; Weising *et al.,* 2005). AFLP markers requires no sequence information, has medium reproducibility and generally results in a highly informative finger prints but require high degree of technical skills (Weising *et al*., 2005). RAPD marker systems is cheap, technically simple to develop, and requires no primer information, but have been severely criticized for the lack of reproducibility, product competition and product homology (Lowe *et al*., 2004). Inter simple sequence repeat (ISSR)-PCR is a technique, which involves the use of microsatellite sequences as primers in a polymerase chain reaction to generate multi-locus markers. It is a simple and quick method that combines most of the advantages of amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) (Pradeep *et al*., 2002; Zietkiewicz *et al*., 1994). Moreover, ISSR marker systems have the advantages over both AFLP and RAPD markers as it enables amplification of DNA without the need to develop species-specific primers and suitable to analyze recently diverged populations as it is derived from fast evolving sequences of SSR

(Simple Sequence Repeat)(Weising *et al.,* 2005). ISSR markers are also highly polymorphic and are useful in studies on genetic diversity, scan the whole genome, are inexpensive, requires no sequence information and easy to generate (Tesfaye, 2006, Weising *et al.,* 2005; Zietkiewicz *et al*. 1994).

Generally, the wild *Coffea arabica* from SW and SE part of Ethiopia showed a relatively high genetic diversity (Anthony *et al*., 2001; Chaparro *et al.*, 2004; Anthony *et al*., 2002). However, in most of these studies, wild *Coffea arabica* from Bale were not included. Recent study by Aga (2005) and Tesfaye (2006) indicate the presence of diverse genotypes, though they used limited samples. ISSR marker observed to differentiate closely related *Coffea* species (Gonzalez *et al.,* 2005; Masumbuko *et al*., 2005), and can show high level of genetic diversity within and among the wild C*offea arabica* (Tesfaye, 2006; Aga, 2005). The target sequence is amplified in the presence of primer complementary to a target microsatellite (Bornet and Branchard, 2001) and the amplified fragments can be separated electrophoretically and the banding patterns can be visualized by staining.

Based on the above grounds, Inter Simple Sequence Repeat (ISSR) marker systems is chosen to study the genetic diversity of wild *Coffea arabica* forest of natural populations from Harenna sites of Bale Mountain, east of the Great Rift Valley in Ethiopia which have not yet been considered for in-depth analysis with large sample size (Tesfaye, 2006; Aga, 2005). Except very few studies made by Tesfaye (2006) and Aga (2005) on the standing population of wild coffee, the genetic diversity within the forest *Coffea arabica* gene pool in Ethiopia has not been extensively studied using molecular markers. This shows that little attention is given to Harrena wild coffee which is considered as drought tolerant as compared to other wild coffee regions (Kufa, 2006). With the current climate change and increment of global heat, the plants in Harrena can play a key role in Ethiopian wild coffee production (Kufa, 2006; Tesfaye 2006). Hence, the present study is to understand the extent of genetic diversity of wild *Coffea arabica* of Harrena forest in order to generate information for conservation and sustainable use using ISSR marker systems.

2. LITRATURE REVIEWS

2.1. A brief overview of coffee

Coffee is an under story shrub or small tree that differ greatly in morphology, size and ecological adaptations, thereby leading to the description of large number of species. Particular attention has been given to the subgenus coffee (genus *Coffea* L.) which include two cultivated species that play an important role in world economy: *Coffea arabica L*. and *Coffea canephora* Pierra (Berthoud and Charrier,1998).Coffee is one of the most popularly consumed beverages in the world and the most valuable agricultural commodity, largely contributing to the economy of more than fifty countries in different part of the world as it is a very important source of foreign exchange income (Anthony *et al*., 2001; Montagnon and Bouhrnormont, 1996; OXFAM, 2002). At a present day, coffee is produced in more than eighty developing countries in the world and it was reported to be among the five most important raw goods and the second most important exported commodity on the earth next to oil (Perdergrast, 1999; Vega *et al*., 2003). According to OXFAM (2002), about eighty percent of all foreign exchange currency for many developing countries depends on coffee as the most important source, only being surpassed by oil.

Among all the economic species of coffee used as crop grown in different parts of the world, commercial coffee production depends on the two most important coffees. *Coffea arabica L*. that is more suitable for highlands contribute more than 70% of the world production. The remaining production comes from *Coffea canephora*, the lowland *coffea,* that only contributes the remaining 30% (GebreEgzihabier, 1990; Mountognon and Bouhrarmont, 1996; Herrera *et al*., 2003). Although both are not equally important in terms of economic contributions, they generally play an important role in satisfying the world‘s market coffee demand. However, *Coffea arabica* L. is the most economically important and widely cultivated *Coffea* species (Masumbuko *et al*., 2005). Like other natural products, coffee displays a great individuality with a wide range of choice of flavor. In addition, although there may be an underlying character to

a coffee from a particular origin, there will also be a lot of interesting variations carrying important characters for breeding purposes. The best quality and more flavored coffee are considered to be produced from the species of *Coffea arabica L.* (Berthaud and Charrier, 1988; Raina *et al*., 1998; Anthony *et al*., 2001). All the species of coffee used as a crop is native to tropical Africa of which *C.canephora* is native to the low lands forest from Liberia, east and south to Kenya and the Congo basin (Gebre Egzihabier, 1990) and C*offea arabica L.* is native to Ethiopia. *Coffea arabica L.* originates from south west and south east Ethiopia where wild *Coffea arabica* grows naturally as a small trees in the under story of the montane rain forests. Moreover, the entire genetic diversity is mainly believed to originate in the Afro montane rain forests of the south west areas of the country, where almost undisturbed patches of forest with coffee as an under growth can still be present (Denich *et al*., 2006; Engels and Hawkes, 1991).

According to Anthony *et al*. (2002) *Coffea* plants of the arabica type in the world are generally estimated to be 10 to 20 billion, all descendents from just a handful of original plants taken from Ethiopia, which indicates Ethiopia to be the centre of origin and diversification of the *Coffee arabica L*. where it grows naturally at altitude of 1300 to 1800 masl (Gole *et al*., 2001; Masumbuko *et al.*, 2005). Being the only centre of origin and diversification of coffee, Ethiopia possesses naturally diverse genetic resources of *Coffea arabica* for different agronomic traits such as resistant materials for coffee berry disease (CBD), leaf rust, coffee wilt, etc. and it is believed to be an important source and the home of coffee genetic resources for the world coffee production (Gole, 2003). This natural gene pool diversity of Ethiopian coffee forest, however, has been undergoing drastic erosion mainly due to deforestation and change in land use for several decades (Gole, 2003 and Senbeta, 2006).

2.2. **Taxonomy of *Coffea Arabica***

According to previous studies based on morphological data by Berthaud and Charrier (1988), Bridson and Verdcourt (1988), Bridson (1987) and Leory (1980), coffee trees, a

tropical woody plant of the Rubiaceae family, are classified into two genera. These are the genus *Coffea* L. and *Psilanthus*. The *Psilanthus* genus includes *Paracoffea and Psilanthus* (Hook f.), but the *Coffea* L. genus includes two sub-genera, *Baracoffea* and *Coffea*. The genus *Coffea* L.is differ greatly in phenotypic features like 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 *Coffea* genus 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* L. More recently, Chvalier (1947) classified the *Coffea i*nto four sections containing *Eucoffea, Argofea, Masarocoffea* and *Paracoffea*. The first three sections of the genus *Coffea* are exclusively native to Africa, Madagascar and some 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). The species of the sub-sections of *Eucoffea*, the rainforest taxa of tropical forest, are typified by *Coffea arabica L*. which are confined to Ethiopia and the rain forest belt of the Congo River drainage of central Africa and rain forest areas of West Africa.

More recently, however, combination of morphological and molecular data set revealed that Rubiceae (*Coffeaeae*) is enlarged to encompass eleven genera. These include *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* is 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* L., *C. Liberica* and *Coffea canephora* Pierre which belong to the subsection *Erythrocoffea* (FAO, 1968; Wrigley, 1988) are economically important (Pearl *et al*., 2004). Out of these three, *Coffea arabica* is by far the most important commercial species and is one of the world's most important commodities (Davis *et al*., 2006; Vega *et al*., 2003). The detail classification of all these genera are summarized and found in Davis *et al*. (2007) and Davis *et al*. (2006).

2.3. **Coffee in Ethiopia**

2.3.1. History and origin of *Coffea arabica*

No body is sure of the exact location where *Coffea arabica* was originally domesticated and discovered as a beverage, except few unevidenced belief or story. The most famous story was that of the goat herd, Kaldi (who lived around 9th century) who observed his normally docile goats had suddenly behaved exceptionally lively, skipping, rearing and bleating loudly after eating the bright red berries from a shiny dark-leaved shrub nearby and that Kaldi tried a few berries him self and soon felt extraordinary, stimulated or a novel sense of elation. Moreover, according to Pursglove (1968) and Sylvia (1958), the other belief is that the Arabic 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 L. Coffea arabica* was introduced to Yemen by Arabs 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.

However, on the basis of botanical evidence, *Coffea arabica* seems to 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. According to Streinge (1956), Anthony *et a*l. (2002), Sylvian (1958) and Dench *et al*. (2006), this was confirmed by the fact that within small area, the wild coffee plants of Ethiopia have relatively high genetic variability as compared to the wild coffee

populations from Yemen that showed a characteristically low genetic diversity. 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.

2.3.2. Spreads of coffee from Ethiopia

The history of introduction of *Coffea arabica* from Ethiopian highlands to Yemen is still linked with scanty 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 from Harer as suggested by Wellman (1961). Moreover, Eskes (1989) and Haarer (1962) also reported the introduction of coffee plants from Ethiopia to Yemen three to four centuries and the fifth century ago respectively.

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 quit well documented (Wellman and Meyer, 1965). From Yemen some coffee plants traveled to wards the East to Malabar coasts of India by Boba Budan in 1760 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. In short, this plant traveled from the Arabian port of Mocha to Java across Holland to its final destination in Paris. The other *Coffea arabica L.* 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 form its place of origin of Ethiopia, ending its transcontinental journey. 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).

2.3.3. Distribution of *C.arabica*

According to Gole (2003), Gole *et al.* (2002), Gole *et al*. (2001), Senbeta (2006), Berthaud and Charrier (1988), *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 altitude ranging from 1,300 and 1,800 masl. However, Senbeta (2006) reported that an altitude ranging from 1300 to 1600 m.a.s.l contains the highest density of coffee. *C. arabica* 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. Moreover, *C. arabica* grows best in the cool, shady environment of the forest of Ethiopian highlands. The temperature requirements for *Coffea arabica* is considered 15-25^{0}C which prevails in most of the coffee growing areas of the country. Above all, these environmental conditions are suitable for coffee growth naturally in the forest of highland and make Ethiopia the 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; Dubale and Shibiru, 2000; Dubale, 1996; Willson, 1985; Teketay, 1999). The centre of origin of *C. arabica* is geographically separated from all other centre of origin of other species of coffee. Distribution of wild coffee Arabica is on the opposite sides of the Great Rift Valley, that is south west of the Rift Valley and east and south east of the Rift Valley (Monaco, 1998; Anthony *et al*., 2001). Small *C. arabica* populations from outside of Ethiopia were also reported, but are confined on the Boma Plateaus of the Sudan and Mount Marsabit in N. Kenya (Wellman, 1961; Firris, 1979).

2.3.4. The Harenna Forest and *Coffea arabica L*.

Coffea arabica has its main center of origin in the afromontane rainforests of Ethiopia. Harenna Forest of Bale Mountain is one of the afromontane rainforests occupying 50% of the Bale National Park in the southeast highland of Ethiopia, where wild *Coffea arabica* is naturally found as understory plants. The forest *Coffea arabica* occurs in the lower portion in the very wet extreme of the Harenna forest, extending between 1320m and 1700m (Taddese and Nigatu, 1996; Senbeta and Denich, 2006). However, the maximum species richness and development of the coffee plants was noted between 1550m and 1560m (Taddese and Nigatu, 1996). Although the wild arabica coffee populations of Ethiopian highland are important from economic point of view, those of the Harenna forest are significantly important for coffee breeding along the prevailing current climatic gradients as reported by Kufa (2006). Wild arabica coffee populations of Harrenna have a drought – stress avoidance mechanisms to withstand dry season in the drier lower altitudes which are promising for breeding and other purpose in rapidly fluctuating environmental condition. The Harenna wild coffees are better adjusted to drought stress through such features as high root to shoot ratio, by their extensive rooting systems, effectively exploited soil moisture, etc. The level and extent of the genetic diversity of this forest coffee, with these important traits hasn't been well investigated to demonstrate the potential diversity and its differentiation used to analyze the genetic structure for breeding purposes (Charrier and Berthaud, 1990). The genetic distribution in the Harenna forest for the *Coffea arabica* L. indicates the need for conservation and thus for sustainable use of wild coffee populations through protection of the natural forest for *in-situ* conservation of wild coffee germplasm in habitats.

2.5.5. Coffee production systems in Ethiopia

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 (Teketay, 1999; Gole *et al*., 2002). The forest coffee also referred to as wild coffee and it 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 coffee also regenerates in its natural forest. 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 referred to as semi-wild coffee, 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 part 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%of the land covered by coffee and 25% of the annual coffee production in the country.

In the garden coffee production system, coffee is planted and managed in the farmer's backyard within small area. It is mainly found in southern and eastern part of the country. Teketay (1999) and Woldetsadik and Kebede (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 about 15% of the total coffee productions. On the other hand, according to Gole *et al*. (2002), the two systems together accounts for 67% and 75% of area and production of coffee of the country, respectively.

2.3.6. Contributions of *Coffea arabica* to Ethiopian economy

The contribution of coffee to export market and foreign currency earning is very high for Ethiopia as stated earlier. According to Tsegaye (2000), coffee contributes about 10% and 5% of government revenues and gross domestic product, respectively. Besides, about 25% of Ethiopia's populations also depend on coffee for their livelihood (Gole, 2003). It can be said that, 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. Moreover, being an important export commodity, coffee plays a vital role or has tremendous impacts on both social and spiritual life of the people in different geographical location and cultural backgrounds of the country.

2.4. Genetic Diversity of Ethiopia wild *C.arabica*

According to Lowe *et al*. (2004), genetic diversity is a commonly used expression to refer to 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 centers 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.

The entire genetic diversity of *Coffee arabica* is largely found in SW and SE montane rain forest of Ethiopia at its center of origin (Gole *et al*., 2001; Senbeta, 2006; Tesfaye, 2006). The diversity in natural populations of *Coffea arabica* based on the morphological (phenotypic) characters has been clear to investigators who visited and noted the afromontane rain forest of Ethiopia. Montagnon and Bouharmont (1996) conducted multivariate analysis of phenotypic diversity of *Coffea arabica* and the result indicated clear structure within the species and identified two main groups: one west and the other east of the Ethiopian Great Rift Valley. Besides, 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, 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 as reported by Dench *et al*. (2006) and Tesfaye (2006). However, although the diversity

information comes from the molecular and agro-morphological marker indicated the high genetic diversity in the wild form of the coffee plants (Anthoney *et al*., 2001), it was confirmed 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 its predominantly self pollinating nature (Lasheremes *et al*., 1996).

Similarly, using combined analysis with AFLP and microsatellites 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 else where in the world. This can be exemplified by difficulties encountered in finding sources of resistance against disease within the cultivars of *Coffea arabica* in other parts of the world. According to Mesfin (1991), alternative to the breeding program limitations are provided by spontaneous and sub-spontaneous genotypes collected from the highlands of Ethiopia, where a high genetic diversity of *Coffee arabica L.* exists. The existence of the wide genetic diversity, in turn, is crucial for the survival of wild arabica populations and also for the cultivated crop. According to Singh (1996) and Roa (2004), the high genetic diversity provides farmers and plant breeders with options to develop new and more productive crops that withstand the changing environment and biotic stresses like diseases and virulent pests. The existence of diverse genetic resources also allows researchers to select and breed animals and plants with desired qualities. Thus, to increase agricultural productivity and to overcome stresses on crop varieties by pathogens, pests, parasites, fluctuating environment etc. that induce a reduction in productivity and extinctions, the existence of diverse genetic resources is required (Bekele, 1985).

According to Gole (2003), Richerzhagen and Virchow (2002) and Denich *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 *Coffea* industry in the future for the improvement of coffee crop. For conservation and sustainable use of the coffee germplasm, the prior knowledge 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. Such studies can provide a better understanding of the relationship between populations of the same species, and group structures that can be found (Charrier and Berthaud, 1990).

2.5. Loss and conservation of genetic diversity

2.5.1. Loss of genetic diversity

The disappearance of the genetic resources is taking place at an alarming rate more particularly in the last two to three decades (IPGRI, 1996; Whitmore, 1997). Such loses occur due to deforestations (habitat destructions), development activities, modern agriculture, accidental introduction of exotic plants and animals, the consequence of which have been aggravated in the tropical countries where the livelihood of the majority of the peoples are dependent on the existing genetic diversity (Gole, 2003; Rao, 2004; Whitmore and Sayer, 1992). Rao (2004) reported more than 15 million hectares of tropical forest are disappearing per year. Moreover, increased livestock population and change in land use pattern contribute a great damage on the wild species diversity. In Ethiopia, in general, and on afromontane rainforests in particular, the forest cover and related biodiversity are threatened due to settlement, land use pressure, habitat destruction and degradation, the consequence of which is changing the normal functioning of the ecosystems. The endangered unique wild populations of *Coffea arabica* are one of such highly vulnerable species and are in threat of extinctions in its centre of origin and diversification (Gole *et al*., 2001; Gole and Teketay, 2001; Yeshitila 2001; Richerzhagen and Virchow, 2002). The wild coffee genetic resources in threats of extinctions in the centre of origin are summarized by Richerzhagen and Virchow (2002). Besides, the gene pools of wild *Coffea arabica* populations are severely endangered because of unsustainable utilization coupled with rapid population growth, which is the root cause of deforestation due to demand for agricultural and settlement areas (Denich *et al.*, 2006; Senbeta, 2006; Gole, 2003) that have aggravated the erosions of genetic diversity (Denich *et al*., 2006). The current situation of deforestation and land use change in SW and SE part of Ethiopia affects not only the *Coffea* genetic resourses but also various people that

depend on the coffee production and the biodiversity of montane rain forest as well (Senbeta, 2006; Gole, 2003).

According to Gole *et al*. (2002), despite all these contributing factors of threats on the forest ecosystem a number of wild coffee populations are still present in the montane rain forest of SW and SE part of the great Rift Valley for conservation and sustainable utilization that prolong the economic benefits and ecological functions they provide for the survival of human beings. Since the country is the center of origin and diversification of *Coffea arabica L.,* genetic resource conservation must be made and forest biodiversity must be maintained using appropriate conservation approaches.

2.5.2. Conservation methods

Evaluations of the genetic diversity are extremely fundamental to understand the level and distribution of diversity in the species and for the design of appropriate conservation strategies. No single conservation approaches can be applied to conserve the full range of plant genetic resources of the target gene pool or species effectively. Although, there are various conservation methods, essentially two main conservation strategies can be applied for plant genetic resource conservation namely: *in-situ* and *ex-situ* conservation. The *in-situ* conservation involves maintaining and conserving the viable genetic resources in their natural habitats where the dynamic process of evolution of the species is occurring and increase the genetic diversity of species to be conserved (Bekele, 1986). Dullo *et al.* (1998) and Karp *et al*. (1997) suggested that *in-situ* conservation as the methods of choices for conserving germplasm of wild forest species and wild relatives and their cultivars in the natural environments where the target species are under natural evolutionary forces. Moreover, according to Senbeta, (2006) it is the vital mechanism of storing the world's genetic diversity or gene pool of important crop plants and their wild relatives with the potential use or vice versa. This, in turn, is crucial to meet the world's future development needs in terms of improving productivity of crops (IPGRI, 2005). *Ex-situ* conservation approach on the other hand is a mean of conservations biodiversity outside their natural habitats. It is, generally, an approach to protect population under

threats of destructions, replacement or deterioration (Bekele, 1985; Rao, 2004). *Ex-situ* conservation involves storage of seeds at low temperature and low moisture contents in the seed banks of the species, or as living in the seed gene banks or botanical gardens. Besides, pollen storage and DNA bank, tissue culture and cryopreservation also provides another means of indirect *ex-situ* conservation of plant germplasm. However, Karp *et al*. (1997) suggested the use of the combinations of both conservation methods for effective conservation of genetic resources. The detail account of both conservation approaches along with their relative strengths and limitations are provided by Maxted *et al. (*1997) and are briefly summarized by Dullo *et al*. (1998).

2.6. Genetic markers and their applications in genetic diversity analysis.

Genetic markers are measurable inherited genetic variations that can be used to understand genetic components. Many different types of genetic markers with different properties exist, each with its own advantages and disadvantages to assess the genetic variations among natural populations. Currently, the most commonly used genetic markers are agro-morphological, biochemical and molecular markers.

2.6.1. Agro-morphology

Traditionally, genetic diversity of *Coffea arabica*, like any other plant species, is analyzed by measuring visible variation of phenotypic features such as flower color, fruit size; shape, growth habit or quantitative agronomic traits like yield potential, stress tolerance, disease resistance etc. It is the oldest methods and considered as the first step in the genetic diversity assessment. However, the methods have a direct applications to the users as they are inexpensive, simple to score and easy to apply for estimating diversity (Smith and Smith, 1989; Werlemark *et al*., 1999). Moreover, Gomez *et al*. (2004) reported the complementarities of molecular and phenotypic markers and recommended the use of both markers for a complete description of the level and pattern of genetic diversity. Therefore, diversity analysis using agro-morphological characters is still relevant for evaluation of genetic diversity. For instance, Montagnon and Bouharmont (1996) used 18 agro-morphological characteristics to evaluate the genetic diversity of coffee from Ethiopia and

then classified wild and cultivated coffee genotypes on the basis of their geographic origin. Such approaches of measuring genetic diversity by selected useful traits based on phenotypic features are, however, don't adequately meet the desired properties of genetic markers and have certain limitations (Karp *et al*., 1997; Masambuko *et al*., 2005; Werlemark *et al*., 1999). They are insufficiently polymorphic to differentiate between closely related species and generally have dominance effects that make it to be poorly suited for progeny analysis. Moreover, the information provided with other traits and trait expression are strongly subjected to environmental variations as they are controlled by many genes that make it difficult to measure. Hence, the information revealed in genetic diversity study is often limited.

2.6.2. Biochemical

Biochemical markers are markers derived from study of the chemical products of gene expressions. 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 (Marshal and Brown, 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, it has been used for genetic diversity analysis in many species (Dudnikov, 2003) and the technique appears to be more informative at lower taxonomic levels, particularly for 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 in revealing polymorphism indicating that,

isozyme as being not appropriate for diversity study in *Coffea arabica* L. (Louarn, 1978) due to the small number of isozyme systems available (Berthaud and Charrier, 1988) and the low level of polymorphism detected (Louarn, 1978). Moreover, the attempt of Paillard *et al*. (1996) to construct isozyme based genetic map for coffee was unsuccessful due to the low polymorphism level. Consequently, arabica coffee researchers, like researchers in many other crop species, shifted towards using DNA-based markers.

2.6.3. 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 polymorphisms 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. However, the general characteristics that would be desirable for a good molecular marker were described in Weising *et al*. (2005) and Karp *et al*. (1997).However, it is extremely difficult to find currently 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). According to FAO (2003), different kinds of DNA based molecular markers were utilized to evaluate DNA polymorphisms in plants. The DNA based marker systems are generally classified as hybridization-based (non-PCR) markers and (PCR)-based markers (Weising *et al.*, 2005; Tesfaye, 2006; Joshi *et al*., 1999).

2.6.3.1. Non-PCR based Markers

The two main non-PCR-based techniques are restriction fragment length polymorphism (RFLP) and variable number of tandem repeats (VNTRs). 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 probe 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.

2.6.3.2. PCR-based Markers

There are a variety of PCR based marker systems currently available for the genetic variation assessment. The introduction of PCR based methods enables investigators to overcome the limitations of probe hybridization based methods and it constituted a new milestone in the field of DNA fingerprinting using molecular markers (Karp *et al*., 1997; Weising *et al*. 2005). With the development of the polymerase chain reaction (PCR), many PCR based molecular techniques have been, and still are being developed for plant genome analysis (Karp *et al*., 1997). 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). Techniques in the first category use primers which are designed arbitrarily/or semi arbitrarily, i.e., with out 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.

Anthony *et a*l. (2001) applied RAPD techniques in the genetic diversity study of wild *C. arabica* and reported the existence of low polymorphism in this species. Similarly, RAPD markers were used by Lashermes *et al.* (1993) to identify coffee accessions and confirmed that the RAPD assay can provide a highly effective and convenient means to 'fingerprint' coffee accessions. Moreover, the genetic diversity of wild and semi-wild accessions of the *Coffea arabica* L. collection from Ethiopia was evaluated with RAPD markers by Chaparro *et al*. (2004) and the result indicated the presence of a relatively large level of genetic diversity or polymorphism within the collection than previously reported.

In arabica coffee the AFLP technique has also been used to study the origin of cultivated *Coffea arabica* L. varieties (Anthony *et al*., 2002) and the result were able to show a successive reduction of genetic diversity during the dissemination of coffee from its primary centre of diversity. Similarly, AFLP were used for detection of genetic introgression by Lashermes *et a*l. (2000) and for construction of a genetic linkage map (Pearl *et al*., 2004). Generally, the genetic diversity of arabica coffee genotypes currently grown in Ethiopia was not adequately evaluated using AFLP markers, although AFLP is efficient and powerful to discriminate genetic variation among closely related genotypes.

Currently, the ISSR technique has been used for genetic diversity analysis of many crop species. In arabica coffee, ISSR markers were employed by Ruas *et al*. (2003) to evaluate

genetic relationship in *Coffea* species and to determine the parentage of inter-specific hybrids. They reported that ISSR analysis could be successfully applied to study genetic relationship and to identify the species involved in the origin of inter-specific hybrids and the result provide an essential basis for a more accurate picture of genetic diversity among *Coffea* species that may assist hybrid breeding program. Masumbuko *et al.* (2005) studied the genetic diversity of *Coffea arabica L.* from Tanzania and he reported that ISSR markers have shown their ability to identify variation in coffee. Aga (2005) and Tesfaye (2006) studied diversity of forest coffee in Ethiopia which showed moderate to high diversity. Similarly, Oljira (2006) investigated the gnetic diversity status of the wild *Coffea arabica* L. populations of Yayu forest Ethiopia (SW) and reported the presence of moderate to large genetic diversity. The result indicated the appropriateness of ISSR markers in the study of intraregional genetic structure of the populations. Several contemporary reports have compared the level of polymorphism detected using RAPD, AFLP and ISSR. According to Nagoka and Ogihara (1997), the agreement is that ISSRs are very powerful to detect polymorphisms, scan the whole genome, and are inexpensive, highly sensitive, cost effective, and easy to generate. ISSR-PCR gives multi-locus and highly polymorphous patterns which are very reproducible and abundant (Zietkiewicz *et al*, 1994; Bornet and Barchard, 2001; Nagaoka and Ogihara, 1997).

Combes *et al*. (2000) made genetic diversity analysis in arabica coffee and the results indicated the presence of low genetic diversity. Anthony *et al*. (2002) studied the genetic diversity within and among Typica-, Bourbon- and sub-spontaneous derived accessions using six SSR loci and could discriminate the Typica derived accessions from the Bourbon derived accessions. Microsatellites were also used in the identification of DNA fragments introgressed from *C. canephora* in four *C. arabica* lines and assessment of polymorphism among *C. arabica* and *C. canephora* accessions. According to Huang *et al.* (2002), Aranzana *et al*. (2003), Akkaya and Buyukunal-Bal (2004), SSR as a marker can be applied to the genetic analysis of organisms with a narrow genetic base. The cultivars of arabica coffee are repeatedly reported as having a narrow genetic base, but the genetic relationship of arabica coffee genotypes currently grown in Ethiopia was not observed to show higher diversity using SSR markers (Aga, 2005).

2.7. Inter simple sequence repeats (ISSRs) Markers.

ISSR is a molecular marker technique that has been available since 1994 (Zietkiewicz *et al*., 1994). It is a glorified RAPD that has been used to study polymorphism based on the presence of the microsatillites throughout the genome (Zietkiewicz *et al*, 1994 Nagoka, and Ogihara, 1997), using longer primers (14-16 base long) that allow more stringent annealing conditions during PCR amplification (Hillis *et al*., 1996). This is because Inter simple sequence repeats (ISSR)-PCR involves the use of microsatellite sequences as primers in a polymerase chain reaction to generate multi-locus markers (Pradeep *et al*., 2002). Hence, ISSR markers are DNA sequence delimited by two inverted SSRs, that is, they are regions found between the simple sequence repeats (SSRs) composed of the same units which are amplified by a single PCR primer consisting of few units with or without anchored end. Each band corresponds to the DNA sequence delimited by the inverted microsatellites as ISSR amplify between SSR regions in the genome.

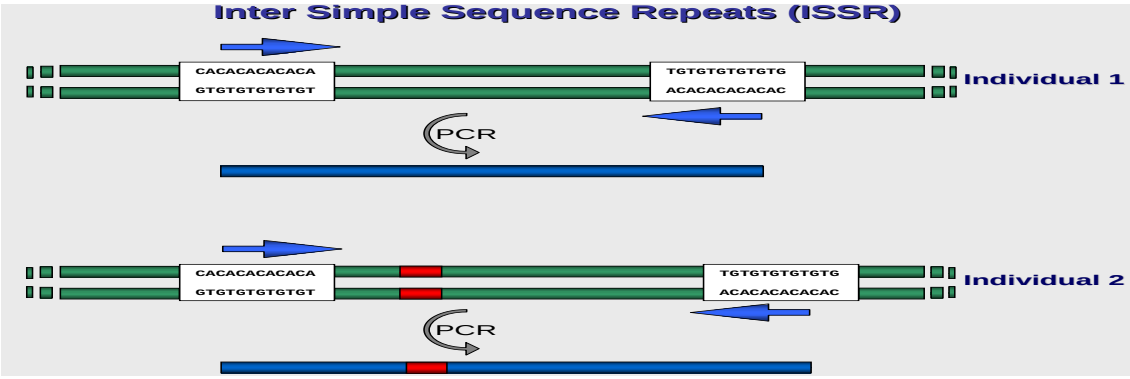


Figure-2.1. Principles of DNA amplification using PCR technique by semi-arbitrary ISSR primer. DNA segments delimited by the inverted simple sequence repeats (SSR) (Individual 1 and 2) are amplified with a single ISSR primer (green). ISSR Variation that may result from insertion or deletion (Red mark) in different individuals produces PCR fragments of different sizes (blue) of the segment (Tesfaye *et al*., 2005).

3. OBJECTIVES OF THE STUDY

3.1. General Objective

The general objective of this study was to evaluate and assess the extent and patterns of distributions of the existing genetic diversity of wild *Coffea arabica* populations in Harrena forest, Bale Mountain.

3.2. Specific Objectives

The specific objectives of this study were:

1. To study the amount of genetic diversity among wild arabica coffee populations in Harrena forest (Bale Mountain).

2. To assess genetic diversity within wild arabica coffee populations from Harrena forest.

3. To determine the relationships among wild *C. arabica* populations in Harrena forest

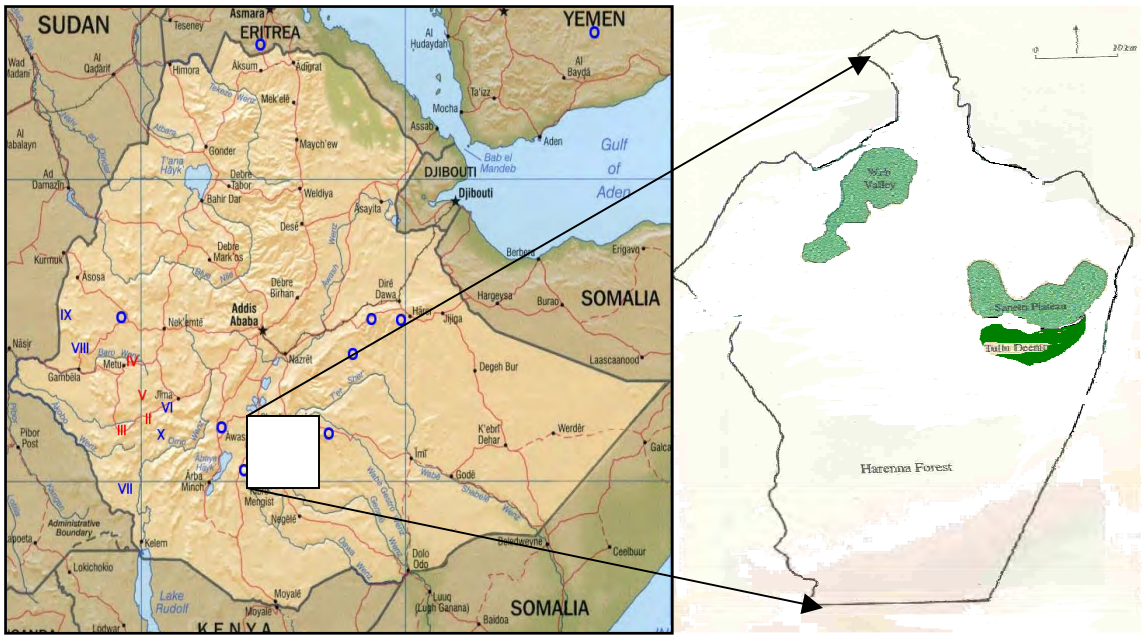
4. To assess the patterns of variation among *Coffea arabica* populations of Harrena forest and to generate information for conservation and sustainable use.

3. MATERIALS AND METHODS

3.1. Description of the study area

3.1.1. The Harenna Forest

The study was carried out on Harrena forest of Bale Mountain located in the South Eastern part of Ethiopia. Harrena forest is located in the Mena-Angetu wereda of the Bale zone of Oromiya National Regional State (Figure 3.1), at about 530 km south east of Addis Ababa. Harenna forest is one of the four research site for CoCE (Conservation and Use of Wild *Coffea arabica* L. Population in the Montane Rainforest of Ethiopia) project. It is also one of the eight important wild coffee areas in Ethiopia proposed for conservation (Dubale and Teketay, 2000). The Harrena forest lies between 6^o and 7^o N and 39^o and 40^o E. It is the most eastern Afromontane rainforest lying between 1,300 and 3,000 m.a.s.l and is one of the few remaining natural forests in Ethiopia and currently constitutes the largest sub-section of the Bale Mountain National Park (Kufa, 2006; Taddese and Nigatu, 1996). Harrena forest is within the Afromontane rain forest system of southeast Ethiopia and commonly known as the southern montane moist evergreen zones of Bale Mountain National Parks (Hillman, 1986). However, forest coffee occurs only in the lower lying area of the forest between 1300 and 1850 m.a.s.l. Although the Harrena forest is demarcated and considered as a national forest priority area partly located in the Bale Mountain National Park, there have been only few conservation efforts. The forest cover is shrinking due to the continuous human activities (Senbeta, 2006). The soil in the Harrena forest has been characterized and described by Tadesse and Nigatu (1989). Accordingly, the soil of Harrena forest is dark reddish-brown, silty-clay rich in basic exchangeable cations. The rainfall pattern in the area is the bi-modal type, i.e. March through April (short rain season) and August through October (long rain season). Annual rain is about 1000mm and the mean annual temperature is 18 ^{0}C (Senbeta, 2006).



A) B)

Figure-3. 1. A) Map of Ethiopia showing the relative sites of the study area, where samples were collected .The site is marked by small white box. B) The cross section of the study site.

3.2. Plant materials and sampling strategies

Individual sample of *Coffea arabica* trees from two years to very old was randomly sampled from four plots with a size of 50 X 50m. The fifty by fifty meter plots selected for this analysis were considered as a population. Fresh leaves were collected from individual sample and dried in silica gel. Twenty five to twenty seven individuals were collected from each population and in total 105 individuals were sampled (Table 3.1).

Table-3.1. List of *C. arabica* populations and sites included in the analysis with original sample size and Coordinates. All samples were collected from Harenna (Bale) forest and the site recognized with site code I. Pop1 & 3 and pop4 and pop6 are from semi disturbed and undisturbed site respectively.

Population Cod	Population	Sample Size	longitude/latitude	Altitude
I-1	Pop1	27	06° 28' 58.8" N/39° 45' 20.3" E/	1519m a.s.l
I-3	Pop3	27	06° 29' 54.3" N/39° 44' 41.9" E/	1576m a.s.l
I-4	Pop4	26	06° 26' 51.2" N/39° 45' 46.6" E/	1470m a.s.l
I-6	Pop6	25	06^{o}31'11.5"N/039^{o}45'091"E/	1487m a.s.l
Total		105		

3.3. DNA extraction, purification and gel test.

The genomic DNA samples used in this analysis was extracted from a total of 105 samples of silica gel dried young leaves tissues of *Coffea arabica L.* following a modified version of CTAB (2% Cetyltrimethyl ammonium Bromide, 1% polyvimylpyrolidone, 100mM Tris: PH=8, 20mM EDTA, 1.4M. NaCl, 0.2% beta-Mercapto-ethanol) procedures of Borsch *et al*.(2003).

Since only one of the three extractions is required for ISSR-PCR analysis, samples of three extractions was 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 2µl of genomic DNA sample of the stock solution and 6µ of 2x loading dye was used and applied to the 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 gel was stained and destained with ethidium bromide and distilled water for 30 minute respectively. The gels were photographed with digital camera mounted on BioDocAnalyze 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 and third extractions were found to be promising and were selected for ISSR-PCR analysis. The DNA samples of the selected extractions were further cleaned with QIAquick PCR purification kit (Qiagen GmbH, Hilden, Germany) since the secondary compounds such as alkaloids that are found in coffee plants can have inhibitory effect on ISSR- PCR. A representative example of the gel picture taken under BioDocAnalyze for some of the individuals of population Bale-4 and Bale-6 is depicted in figure below.

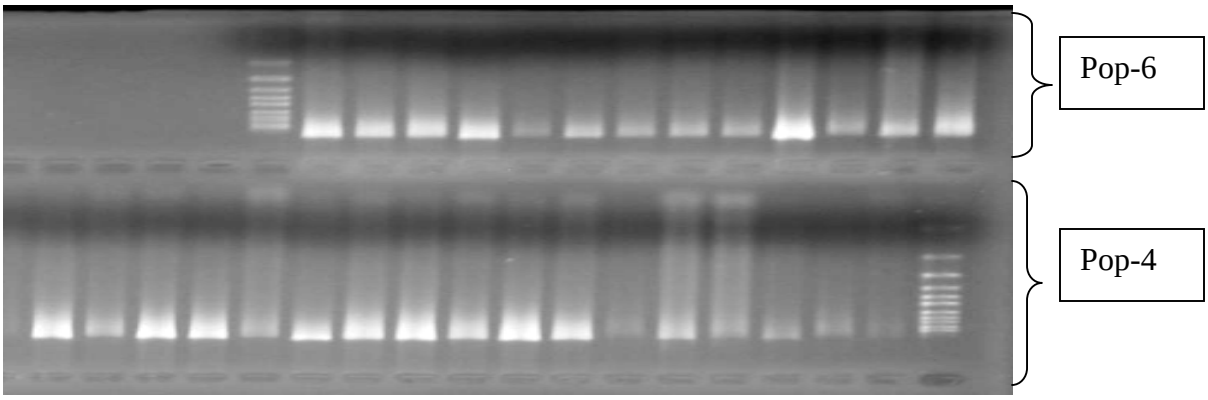


Figure-3.2. Examples of test gel for some of the individual plants of population four and six.

3. 4. Inter simple sequence repeat polymerase chain reaction (ISSR-PCR) assay

3. 4. 1. PCR optimization and primers used

Estimation of the extent and distribution of genetic diversity among and within the four populations of *Coffea arabica* from Harenna forest was analyzed using Inter Simple Sequence Repeat (ISSR) markers. As this research project execute as the follow up of CoCE-I, PCR protocol described in Tesfaye (2006) and Oljira (2006) were used for this intraregional analysis. Therefore, a total of nine primers seven di-nucleotide (810, 812, 813, 814, 818, 834, and 844) and two tetra-nucleotide (CoIS001, CoIS002) primers were chosen and used for ISSR amplification in this analysis (Table-3.2).

Table-3.2. Primers and annealing temperature of the primers used for this analysis

No.	Code of the primers	Annealing Temperature(^{o}C)	Remark
1	810	45	Dinucleotide repeats
2	812	45	Dinucleotide repeats
3	813	45	Dinucleotide repeats
4	814	45	Dinucleotide repeats
5	818	48	Dinucleotide repeats
6	834	45	Dinucleotide repeats
7	844	48	Dinucleotide repeats
8	CoIS001	45	Tetra nucleotide repeats
9	CoIS002	45	Tetra nucleotide repeats

To test reproducibility of the ISSR-PCR reaction, two separate PCR reactions were randomly run with DNA diluted at 1:5 and 1:10 for two samples randomly selected from each four populations. The PCR test was carried out using 810 primers. This primer was chosen because it was reported to have high discriminatory potential as compared to other selected primers for the analysis (Aga, 2005; Tesfaye, 2006; Girma, submitted). The PCR reaction was carried out in a total reaction volume of 24µl containing 13.2µl ddH_2O; 5.6µl dNTPs; 2.6µl Taq buffer; 2 µl MgCl2; 0.4µl 810 primers, 0.2µl Taq polymerase and 1µl DNA. The gel test was performed with PCR amplified products for the two dilution cases (diluted at 1:10 and 1:5). The test gel with sample that diluted at 1:5 and 1:10 showed good bands with a little better intensity of 1:5 dilutions. It was then decided that 1:5 dilution be used for further analysis.

3. 4.2. ISSR-PCR amplification and gel electrophoresis

Each DNA amplification reaction were performed in a final volume of 25µl containing 13.2 µl ddH_2O; 5.6 µl *Taq* buffer(10x Thermopol reaction buffer: 50mM KCl, 10mm Tris-HCl, PH 8.4), 2.6 µl of dNTPs (1.25 mM of each of dATP, dTTP, dCTP and dGTP);

2 µl $MgCl_2$ (2mM); 0.4 µl (20pmol/µl) ISSR primer; 0.2 µl (5U/µl) *Taq* DNA polymerase and 1 µl template DNA. The PCR amplification was carried out using a Biometra T3 thermal cycler PCR machine programmed with 4 min. at 94^0C for preheating and initial DNA denaturation followed by 39 cycles of denaturation at 94^0C for 5 seconds; annealing at 45^0C for primers 810-H, 812-H, 813-H, 814-H, 834-H, CoIS001 and CoIS002 and at 48^0Cfor primers 818-H and 844 for 1min.; extension at 72^0C for 1:30 min. The final cycle was followed by 7 min extensions. The samples were stored at 4^0C until electrophoresis. About 8µl of ISSR-amplification product of each sample DNA with 2 µl loading dye (X6) was loaded on and were resolved in 1.67% agarose gels in 1x TBE buffer at 100volt constant for 1:30 - 2h and stained with Ethidium Bromide which was mixed with 400 – 500ml and destained in the same amount of the distilled water for 30 min., respectively. The ISSR profiles were visualized under UV light, photographed with digital camera mounted on BioDocAnalyze and connected to PC with Biometra software, and stored for later data scoring. To estimate the molecular sizes of the resolved fragment, 100bp DNA marker was used

3.6. Scoring and Data Analysis

Out of one hundred five (105) individual coffee plants, five individual coffee plants were showed poor DNA quality and concentration on the gel picture and thus removed from the data analysis. Each of the PCR fragments generated using seven di-nucleotides (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) and two tetra-nucleotides (CoIS001 and CoIS002) from one hundred coffee plant analyzed were then scored. Data were considered as a unit character and scored for the present (1) , absent of homologous DNA fragments (0) and missing data (‘?’). Only ISSR bands showing clear polymorphisms were scored as presence or absence bands. The bands that were not clearly shown were scored as a missing data (‘?’). Intensity variations among fragments having approximately the same molecular size were not considered although in some cases intensity differences of the bands were observed. A binary matrix was constructed with the individuals in the column and the ISSR markers in the row. Each amplified fragments was named by the code of the primers across the row followed by the Arabic numbers starting from the fragment having

high molecular weight to the fragments with low molecular weight. Since ISSR markers are multi-locus markers, a locus was considered to be polymorphic if the presence and absence of bands were observed in various individuals for a given locus, and monomorphic if the bands were present in all individuals. On the basis of the recorded band profiles, different software's were employed for the analysis of the data. Genetic diversity, the numbers of polymorphic loci and the percentages of polymorphic loci at population and individual levels were calculated using POPEGENE version 1.31(Yeh and Boyle 1997).

Shannon's Weaver genetic diversity index (H) was used to quantify the degree of within population diversity, estimated as $-\sum p_3 \log_2$ pi, where *pi* is the frequency of the presence and absence of an ISSR bands (Lewontin, 1972). The average or mean of intra-population diversity over the different populations (H_{pop}) and the total diversity calculated from the phenotypic frequencies *p* in all populations considered together (H_{sp}) was calculated from $-\sum p \log_2$ pi since Shannon's diversity index. The overall diversity of the species /whole population was then partitioned in to the proportion of the diversity within (H_{pop})/H_{sp}) and between populations [(H_{sp}-H_{pop})/H_{sp}] for each locus, respectively. Analysis of molecular variance (AMOVA) were carried out to estimate variance components of the ISSR data to partition the genetic diversity within and between populations and groups based on the Arlequin soft ware version 3.01 (Excoffer *et al*., 2006). Genetic similarity matrix among hundred individual samples of *Coffea arabica L*. was calculated in all pair-wise comparisons following Jaccard's similarity coefficients (Sneath and Sokal, 1973) using binary character matrix synthesized from band profiles with the formula:

J xy = a/a+b+c Where: 'a' is the number of bands shared between 'X' and 'Y' populations

'b' is the number of bands present in Population 'Y' but absent in population 'X'

'c' is the number of bands present in 'X' but absent in 'Y'

The Jaccard's similarity coefficients were used to construct dendrogram using UPGMA (Unweighted pair group methods on arithmetic average) and NJ (Neighbor Joining) algorithms. This similarity coefficients was calculated using NTSYS-pc-version 2.1 (Rohlf, 2000) and Free Tree 0.9.1.50 (Pavlicek *et al*., 1995) software. The UPGMA algorithms of NTSYS-pc version of 2.1 software were employed to perform a sequential, agglomerative, hierarchical, and nested (SAHN) cluster analysis using the similarity matrix. A dendrogram generated was used to analyze and compare individual coffee plants, i.e., to show the genetic relationship and differences of each individual plants. The NJ method based dendrogram (Saitou and Nei, 1987; Studier and Keppl, 1988) was used to compare individual plants and evaluate patterns of individual clustering using tree. Patterns of genetic variation among individual samples were also further examined with three dimensions with the help of principal coordinate analysis (PCO) on the basis of Jaccard's coefficients of similarities (Jaccard's, 1908), which was calculated using PAST software version1.18 (Hammer *et al*., 2001). The first three axis were later used to construct the scatter plot with STATISTICA version 6.0 software's (Hammer *et al*., 2001; Statistica soft, Inc., 2000)

4. RESULTS

4.1. ISSR marker banding patterns

Fragment patterns generated by the seven di-nucleotide and two tetra-nucleotide ISSR primers were analyzed among one hundred individuals of *Coffea arabica* L. representing four populations. The pattern of DNA amplification obtained was moderately better, clear and reproducible banding patterns based on the results from gels pictures taken for each primer. The size of the band generated ranged from 400 to 3000bp (Table 4-1). The number of bands produced by each primers varied from ten bands (the lowest numbers of bands) for 814 to twenty one (the highest number of bands) for CoIS001 primers. According to the previously reported studies, the ISSR technique amplifies microsatellite regions that are potentially polymorphic and thus expected to reveal very high levels of genetic variation. In this study, one hundred thirty seven bands were scored in which only

61 fragments showed polymorphisms. Each individuals sample did not indicated abundant number of unique ISSR phenotype, indicating the low level of genetic variation in the individual plant analyzed. In addition, in both di and tetra nucleotide ISSR primer analysis, very different banding patterns was observed. While the seven di-nucleotide primers showed almost the same banding patterns, almost all individuals analyzed were observed to have a very unique banding pattern in assessment with tetra- nucleotide primer. Generally, individual specific markers were clearly observed in the assessment with tetra than in di-nucleotide primers. However, population-specific markers for each primer were not observed in the whole analysis. A representative electrophoresis pattern obtained with ISSR markers for di and tetra-nucleotide is depicted in figure-4.1 below.

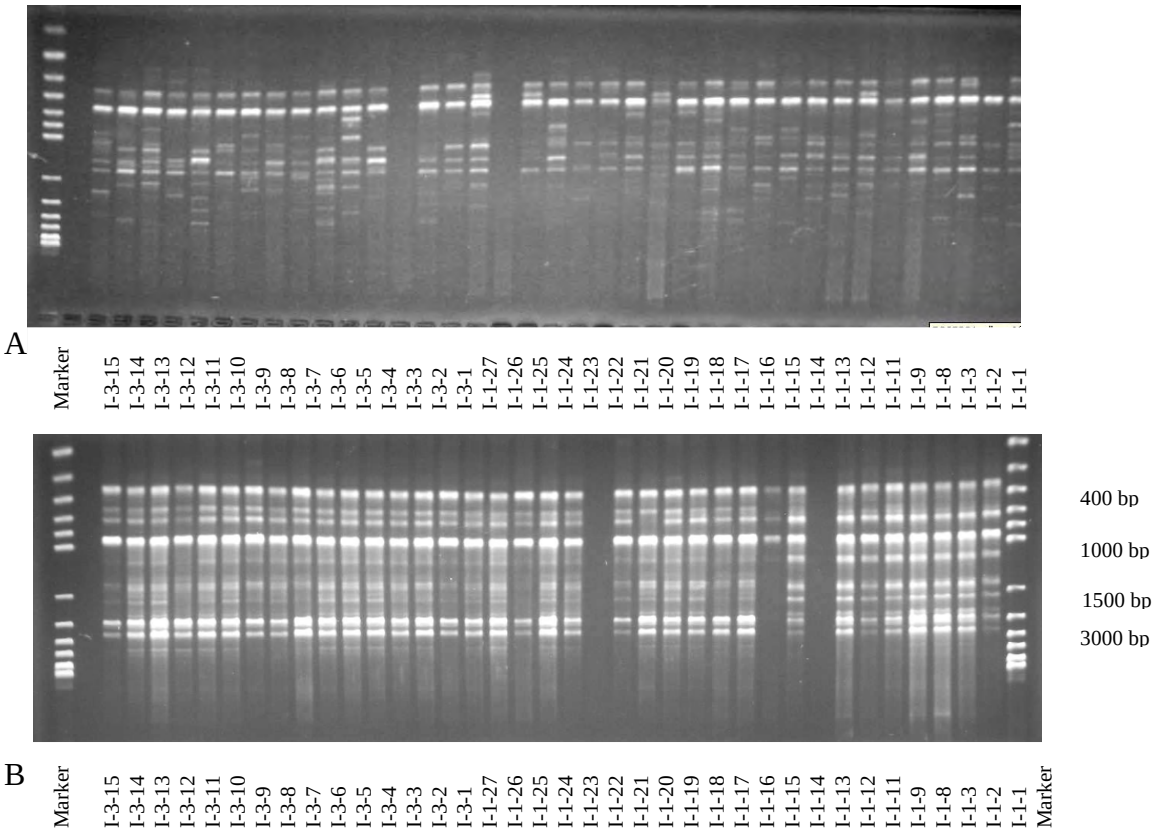


Figure 4.1 Inter simple sequence (ISSR) repeat electrophoresis assessment of genetic diversity in 37 wild *Coffea arabica* L. individuals. A) is based on tetra nucleotide primer CoIS001 and B) is based on di-nucleotide primer (844-H).

4.2 Genetic Diversity

4.2.1. Diversity of the locus

In this intraregional analysis, the total locus diversity across the four populations was determined using the data from an over all nine primers: seven di- and tetra nucleotide ISSR primers that were tested for their capacity to differentiate the populations under consideration. Generally, all primers together were able to produce one hundred thirty seven (137) scorable and reproducible bands, of which 61 were polymorphic. The seven di-nucleotides primers were able to generate a total of ninty nine (99) loci of which twenty six (26) bands were polymorphic across the four populations of *Coffea arabica* L. samples analyzed, which is comparable with the locus diversity obtained by Oljira (2006). According to the result, analysis with five di-nucleotide primers generates a total of 55 bands of which 21 were polymorphic. The tetra nucleotide produced a total of thirty eight (38) fragments of which thirty five (35) were found to be polymorphic. The bands obtained were well separated for tetra nucleotide primers as compared to the di-nucleotide primer. The number of amplified fragment and polymorphic loci produced per individual varied depending on the primer used. Ten (for 814 primers) to twenty one bands (for CoIS001 primers) and one to nineteen polymorphic fragments for these primers were obtained with a mean of 15.2 and 6.78 respectively.

4.2.2. Levels of polymorphism

Among *Coffea arabica* L. population analyzed to evaluate the level of polymorphisms between and within populations, PCR amplification with ISSR primers produced a total of one hundred thirty seven (137) bands per individual samples. The levels of polymorphism both in terms of the number and percentages were not equally contributed by both di- and tetra-nucleotide primers employed in the analysis (Table 4.1). Out of the total bands produced, sixty one (61) were polymorphic. Among these, twenty six (26) and thirty five (35) polymorphisms were contributed by the seven di-nucleotides (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) and two tetra-nucleotides (CoIS001 and CoIS002-

nucleotide) primers, respectively. The numbers and percentages of polymorphic bands also varied per primer from one (10%) for 814_H di-nucleotide primer to nineteen (94.2%) for tetra-nucleotide CoIS001 with an average of 6.78 (5.59%). Moreover, tetra-nucleotide CoIS002 also yielded high number polymorphism loci as compared to the di-nucleotides, indicating that the two tetra-nucleotides surprisingly contributed more polymorphisms. In addition, the overall analysis with di-nucleotide and tetra-nucleotide primers, the number and percentage of polymorphic bands within populations varied from 40 (29.20%) for population 1 to 50 (36.50%) for population 3. However, the analysis with di-nucleotides only, ranged from 9 (9.09%) to 19 (19.19%) and with tetra nucleotides 25 (65.79%) for population 4 to 31 (81.58%) for population Pop1 and Pop3. This also supports the high polymorphisms resolution power of the tetra-nucleotide primers.

Moreover, polymorphic loci and percentages of polymorphic loci for the four populations varied (Table 4.2). The number and percentages of polymorphism with all primers combined ranged from 40 to 50 and 29.2% to 36.5% respectively for population Pop1 and Pop3. Similar patterns were obtained with di-nucleotide primer and the number and percentages of polymorphic loci ranged from 9 to 19 and 9.1% to 19.2% respectively for population Pop1 and Pop3. However, different patterns was observed with tetra-nucleotide primers in which number and percentages of polymorphic loci ranged from 25 (65.79%) for population Pop4 to 31 (81.56%) for population Pop1 and Pop3, again indicating that tetra-nucleotides reveal more genetic diversity than di-nucleotide primer.

Hence, among the four populations analyzed in this study, Pop1 is the least diverse while Pop3 is more variable for all primers. In the same manner, Pop1 is the least diverse while Pop3 is the most diverse with an over all and di-nucleotide primers analysis. In the tetra-nucleotides, Pop4 (which showed a slightly high variability with the two cases) is the least divers while Pop1 (which showed least variability previously) and Pop3 are the most variable. In all cases Pop6 showed similar levels of polymorphism and occupy approximately an average values. This again clearly indicates tetra-nucleotides repeats are more polymorphic than di-nucleotides. Generally, out of the total primers employed in the analysis, tetra-nucleotide showed a significantly maximum polymorphism while the

smallest amount of polymorphism resulted from di-nucleotide primers both in terms of the number and percent polymorphisms. In terms of the mean percentage polymorphisms, the mean diversity levels revealed by tetra-nucleotides primers (M=75) was higher than the mean diversity of di-nucleotide primer (M=15.15 on the basis of t-test, t (6) = 11.81, P = 0.001, at significance level of $\alpha = 0.05$. Since the probability value (0.001) is less than the significance level (0.05), the variability of the tetra-nucleotide primer is highly significant. Moreover, these patterns of diversity analysis identified the high diverse and least diverse populations and more variable primers to investigate *Coffea arabica* L. populations (Table 4.2).

Table- 4.1. Lists of all primers used in the analysis, primer sequence, number of scorable bands, no of PL, percentages of PL and estimated molecular size range in comparisons between the *Coffea arabica L.* individuals from wild populations.

Primer Code	Primer Sequence	No of Scorable bands	No of PL loci	%PL	Molecular size range in bp
810	GAGAGAGAGAGAGAGAT	16	5	31.25	2500-500
812	GAGAGAGAGAGAGAGAA	18	4	22.22	3000-500
813	CTCTCTCTCTCTCTCTT	16	3	18.75	2500-600
814	CTCTCTCTCTCTCTCTA	10	1	10	2500-500
818	CACACACACACACACAG	11	3	27.27	2500-500
834	AGAGAGAGAGAGAGAGYT	12	4	33.33	2000-400
844	CTCTCTCTCTCTCTCTRC	16	6	37.5	2500-600
CoIS001	CCTACCTACCTACCTA	21	19	90.48	2500-600
CoIS002	GGTAGGTAGGTAGGTA	17	16	94.2	2000-400
Total		137	61	43.53	

Table-4.2. Genetic variation of *Coffea arabica* L. populations' in terms of number of polymorphic locus (PL) and percent polymorphism (%PL) based on 137 ISSR bands with all primers combined, di-nucleotide and tetra-nucleotide.

Populations	No of plant Examined	Total No PL			% PL		
		Di-+ tetra	Di-nucl.	Tetra-nucl.	Di-+ tetra	Di-nucl	Tetra-nucl.
Pop1	22	40	9	31	29.20	9.09	81.58
Pop3	27	50	19	31	36.50	19.19	81.58
Pop4	26	42	17	25	30.66	17.17	65.79
Pop6	25	42	15	27	30.66	15.15	71.05
Total	100	61	26	35	44.53	26.26	92.11
Mean	25	43.5	15	28.5	31.73	15.15	75

4.3. Partitions of genetic variation

Levels variations can be attributed to within and between population components. Shannon's diversity index (Table 4.3) and analysis of molecular variance (AMOVA) (Table 4.4 and 4.5) were used to partition the existing genetic variation in to different components.

4.3.1. **Shannon Diversity Index**

Shannon's diversity indexes for each sub-divided populations (Hpop) and for whole data set (Hsp) were calculated for di and tetra-nucleotides and for the combination of both primer (over all) and is summarized in table-4.3. Considering estimates Shannon's genetic

diversity index (H_o) of each population with all primer (over all), it ranged from 0.24 (for populations P1) to 0.30 (for population P3), and is consistent with the number and percentages of polymorphic bands (Table 4.2). For di-nucleotides and tetra nucleotide primers, the same patterns were observed. The mean Shannon's genetic diversity for populations (Hpop) and mean Shannon's genetic variations for entire data (Hsp) was observed to be higher for tetra-nucleotides as compared to the di-nucleotides (Table 4.3). This was also being supported with t-test conducted at a significance level of 0.05 % and df of 6. Accordingly, the mean Shannon's diversity level revealed by tetra-nucleotide primer (M =0.64) was significantly higher than the Shannon's diversity levels revealed by di-nucleotide primers (M = 0.122), t (6) = 12.29, p= 0.001, at significance level of α = 0.05. Since the probability value (0.001) is less than the significance level (0.05), the Shannon's diversity levels revealed by of the tetra-nucleotide primer is highly significant. The Shannon's diversity index is summarized in table 4.3 below.

Table-4.3. Summary of Shannon's genetic diversity index for each population and all individuals of wild *Coffae arabica* L. from Harenna forest and partitioning of the genetic variation in to within and among populations.

Parameters	Di-nucleotide primers	Tetra-nucleotide primers	Over all
Pop1	0.237	0.078	0.556
Pop3	0.300	0.136	0.730
Pop4	0.264	0.152	0.652
Pop6	0.264	0.122	0.636
Hpopn	0.122	0.644	0.266
Hspecies	0.208	0.725	0.351
Hpopn/Hspp	0.827	0.858	0.836
1-Hpopn/Hspp	0.173	0.142	0.164

Hpop=mean genetic variation for the populations; Hsp = mean genetic variation for the entire data set; Hpop/Hsp = proportion of the genetic variation within the population; [(Hsp-Hpop)/Hsp] = proportion of the variation among the populations.

4.3.2. Analysis of Molecular Variance (AMOVA)

Analysis of molecular variance was carried out in two phases; one was done using the populations grouped in to semi-disturbed and undisturbed and the other was done for the entire populations (i.e., using the four populations as it is with out grouping) over all loci by considering them as one geographic region. 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 (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 ISSR diversity was distributed between individual plants within the populations (56.8%), with the remaining diversity being distributed among populations within groups (11.11%) (i.e. populations within semi-disturbed and undisturbed groups) and among groups (33.2%). Similarly partitioning of genetic diversity by analysis of molecular variance with out 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 (63.7 %), while the remaining is due to differences among populations (36.4%). In both cases, the results of analysis of molecular variance revealed the same patterns of genetic diversity and supports the larger genetic diversity found within the populations rather than among populations and is similar to Shannon's diversity index. The AMOVA result is summarized in table 4.4 and 4.5 below.

Table-4.4. Partitioning of the genetic variation in to among groups, among populations within groups and within populations by analysis of molecular variance based on n 137 ISSR bands generated from *Coffea arabica L*.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation Indices	P
Among Groups	1	149.966	2.761 Va	32.2	0.162	0.00
Among Populations Within Groups	2	51.865	0.939 Vb	11.0	0.432	0.00
Within Populations	86	418.336	4.864Vc	56.8	0.322	
Total	89	620.167	8.563			

Table-4.5. Partitioning of genetic variations in to within and among populations by analysis of variance based on n 137 ISSR bands generated from *Coffea arabica L*.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation Indices	P
Among populations	3	201.831	2.778 Va	36.4	0.36350	0.00
Within						

populations	86	418.336	4.864 Vb	63.7
Total	89	620.167	7.64	

4.4. Cluster Analysis.

Intraregional cluster analysis of UPGMA (Unweighted Pair-Group Methods Using arithmetic Averages) and NJ (Neighbor Joining) were computed for all individuals of *Coffea arabica* L populations that were analyzed with ISSR primers and had a complete data set. The UPGMA dendrogram resulting from a SAHN clustering analysis and NJ analysis on the basis of Jaccard's coefficients of similarity was constructed. The Jaccard's coefficient of similarity was obtained after pair-wise comparisons performed using binary character matrices (of the presence and absence) that were produced from amplified fragments. In UPGMA clustering of an over all analysis, all the individuals clustered according to their respective origin of the populations in the region in to two major groups representing semi-disturbed (Pop1 and Pop3) and undisturbed (Pop4 and Pop6) groups (Figure-4.2). Analysis with only di-nucleotide primer showed a similar patterns as those with the same site grouped together in an over all assessment (Figure-4.3). Tetra-nucleotide revealed no clear patterns of associations (Figure 4.4). In the former two cases, clear patterns of grouping were observed except in few intermixing of individuals from other population. In terms of grouping intensity, Pop1 and Pop3 were observed to form strong grouping as compared to Pop4 and Pop6 in the case of UPGMA. Moreover, similar patterns of grouping in the case of NJ were also observed but with more intermixing than with UPGMA clustering (Figure- 4.5, 4.6 and 4.7). In terms of grouping intensity, however, Pop4 and Pop6 tend to form strong grouping than Pop1 and Pop3 in NJ analysis. The UPGMA and NJ clustering methods of hundred individuals of coffee for the overall nine primers, seven di- and two tetra nucleotides, did not produce exactly the same tree topology

Generally, on the basis of UPGMA, each individuals of the respective population were observed to form moderately clustered groups with few intermixing from the other population, except for the tetra- nucleotide primers. Similar patterns of grouping were also observed in the case of NJ again except with tetra-nucleotide primers analysis. However,

few individuals were observed to have a long extended branch from their respective group in the case of NJ analysis. The UPGMA and NJ based dendrograms are depicted in indicated figure below.

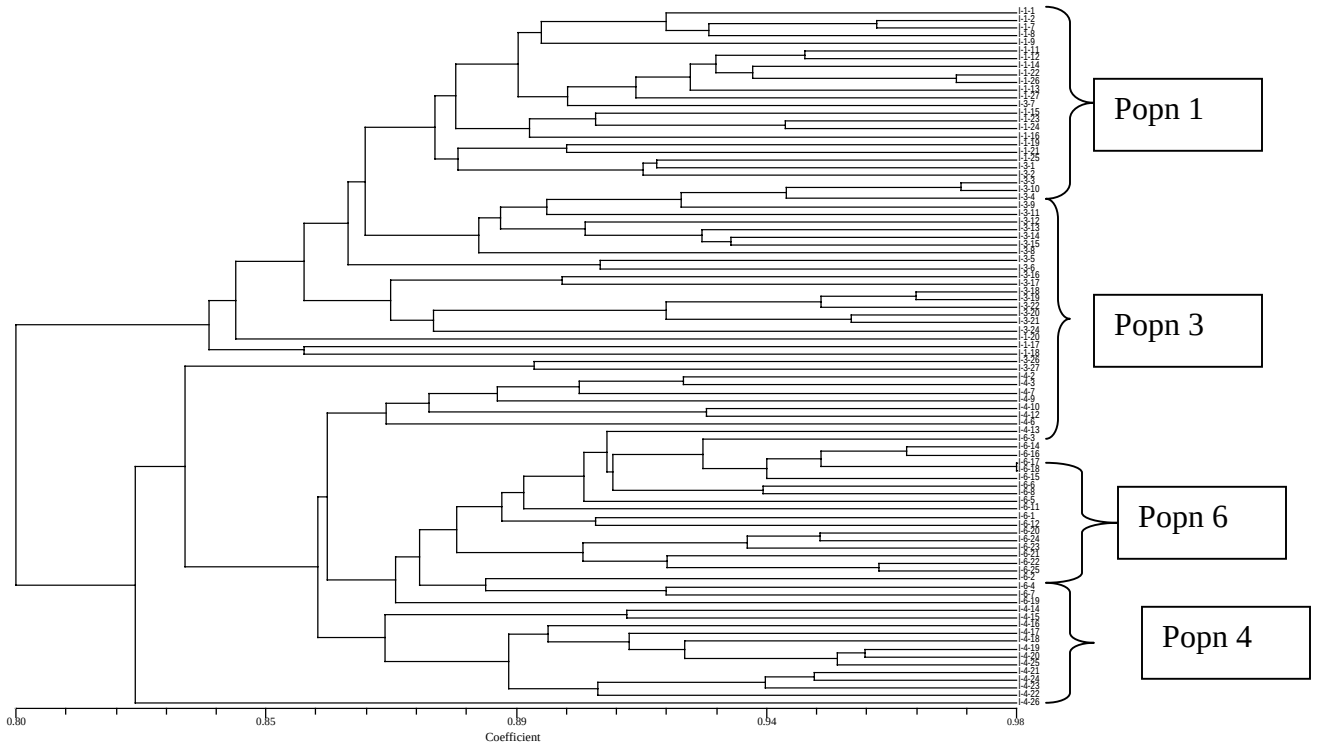


Figure-4.2. Dendrogram generated based on UPGMA analysis demonstrating the genetic similarity between one hundred individuals of forest *Coffea arabica* L population using seven di (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) and two tetra-nucleotides (CoISoo1 and CoISOO2) data. The diagram was base on the Jaccard's coefficients of similarity from 137 ISSR fragments.

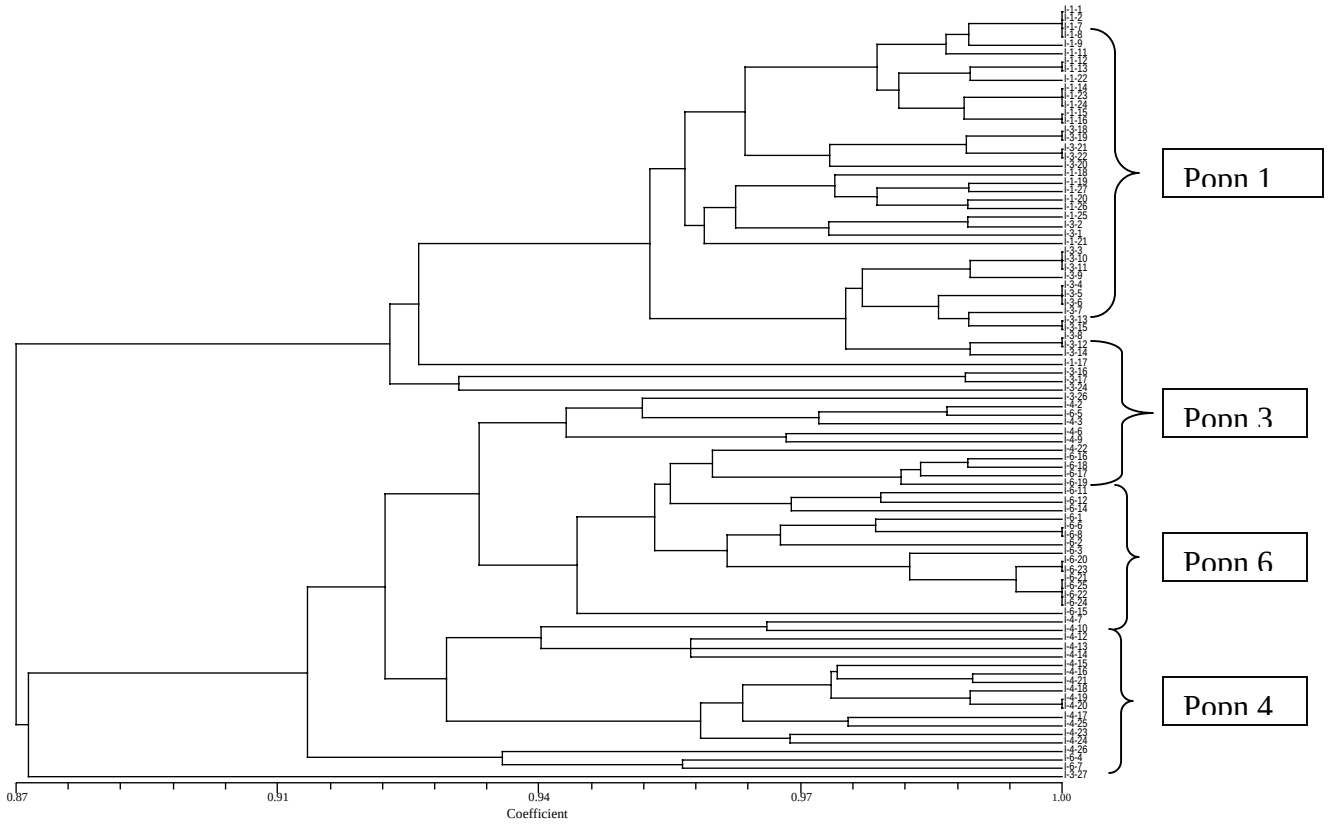


Figure-4.3. Dendrogram generated based on UPGMA analysis demonstrating the genetic similarity between one hundred individuals of forest *Coffea arabica* L population using seven di- nucleotide primers only (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H). The diagram was base on the Jaccard's coefficients of similarity from 99 ISSR fragments.

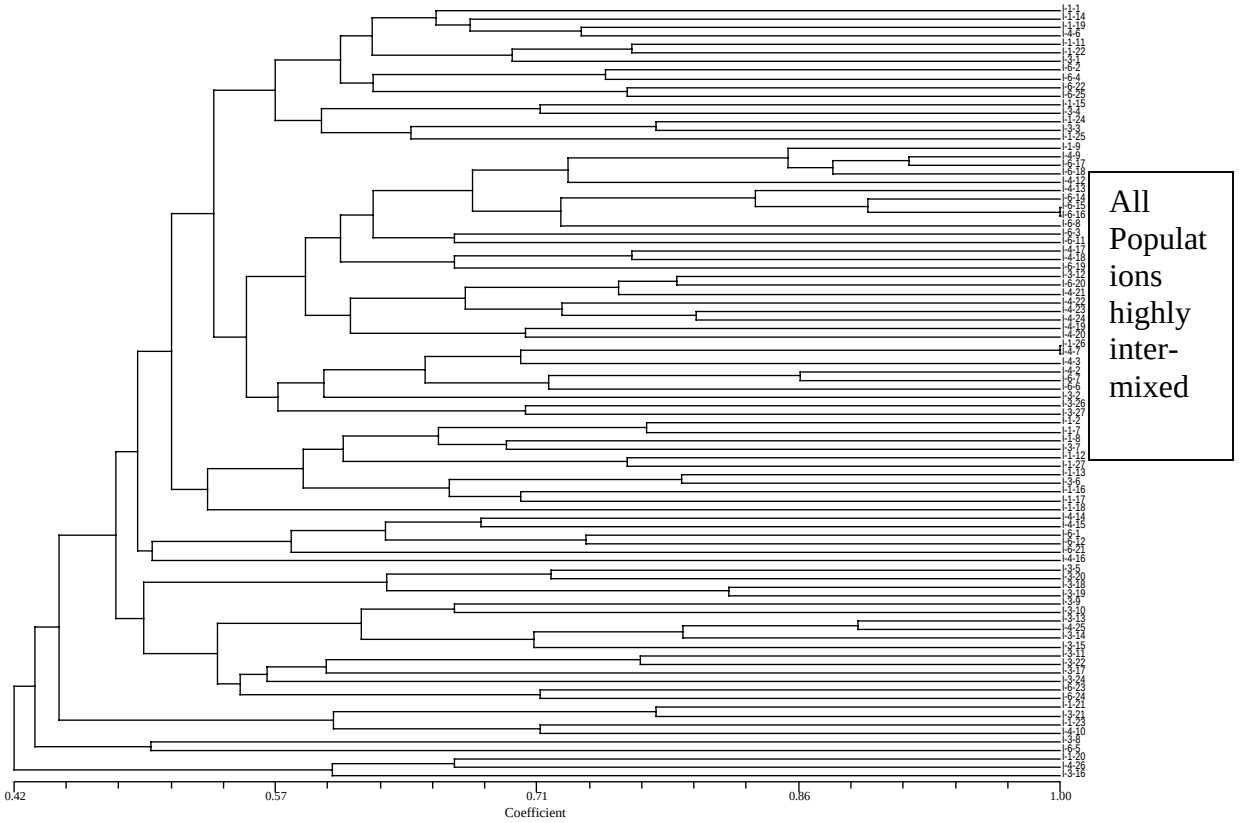


Figure-4.4. Dendrogram generated based on UPGMA analysis demonstrating the genetic similarity between one hundred individuals of forest *Coffea arabica* L population using two tetra-nucleotides (CoIS001 and CoIS002) data. The diagram was based on the Jaccard's coefficients of similarity from 38 ISSR fragments.

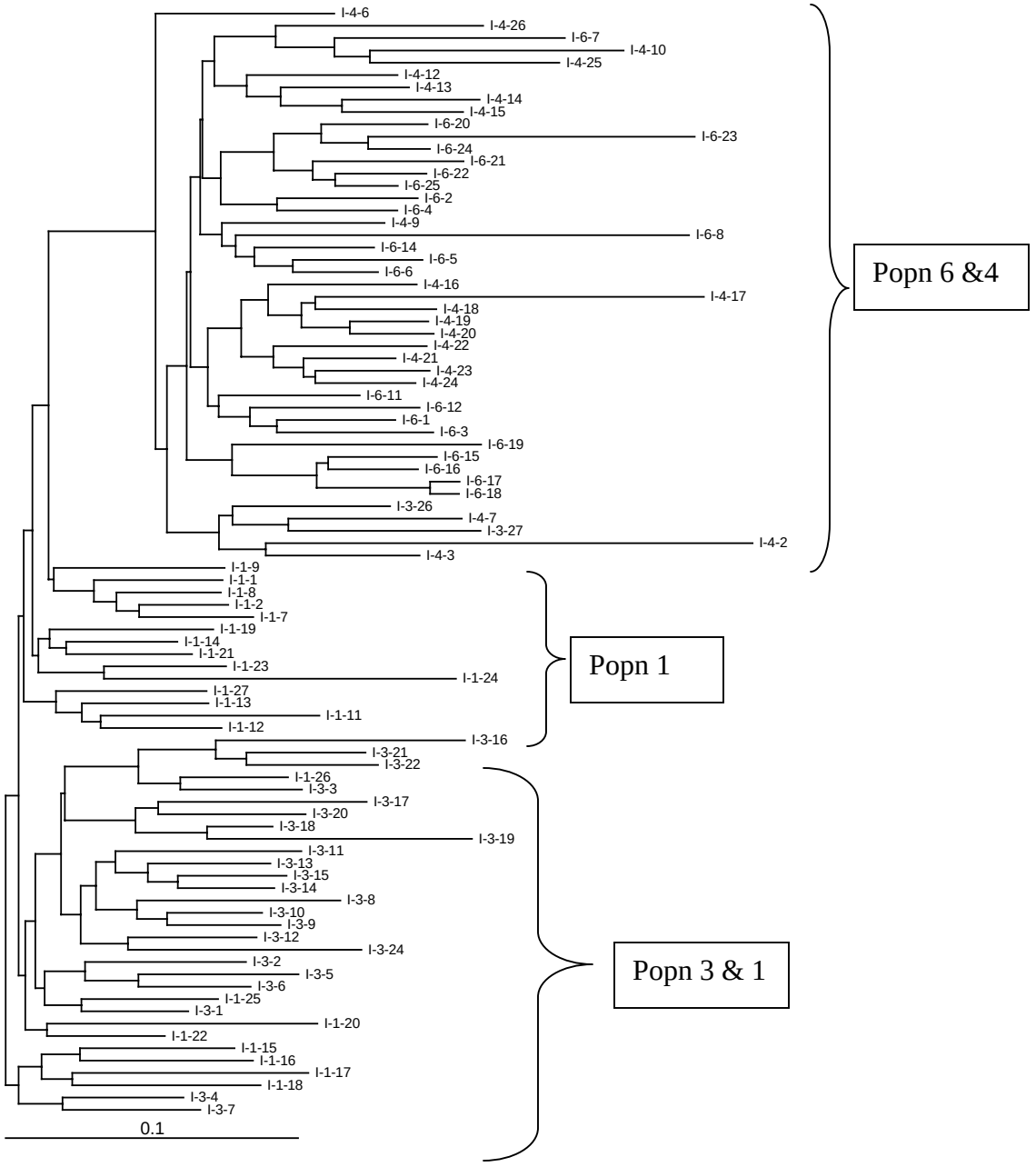


Figure-4.5. Dendrogram generated based on NJ analysis of complete intraregional data set of one hundred wild coffee based on seven di (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) and two tetra-nucleotides (CoIS001 and CoIS002) data. The NJ tree was based on the Jaccard's coefficients of similarity from 137 ISSR fragments of hundred individuals of forest *Coffea arabica* L. The Roman number indicate the site (Bale),the first and the second Arabic number indicate the population and individual plants of the population respectively.

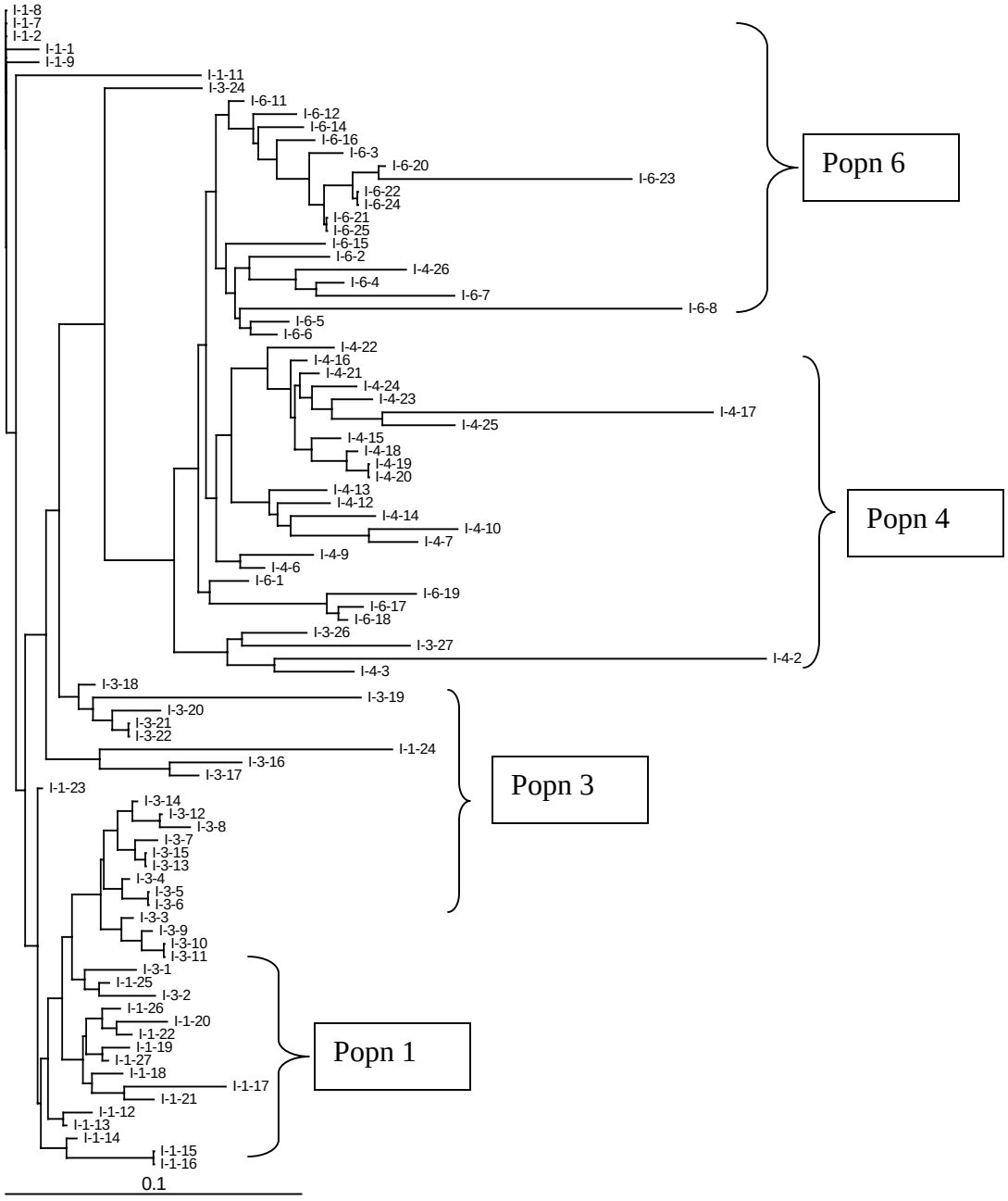


Figure-4.6. Dendrogram generated based on NJ analysis of complete intraregional data set of one hundred wild coffees based on seven di (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H). The NJ tree was based on the Jaccard's coefficients of similarity from 99 ISSR data of hundred individuals of forest *Coffea arabica* L. The Roman number indicate the site (Bale),the first and the second Arabic number indicate the population and individual plants of the population respectively.

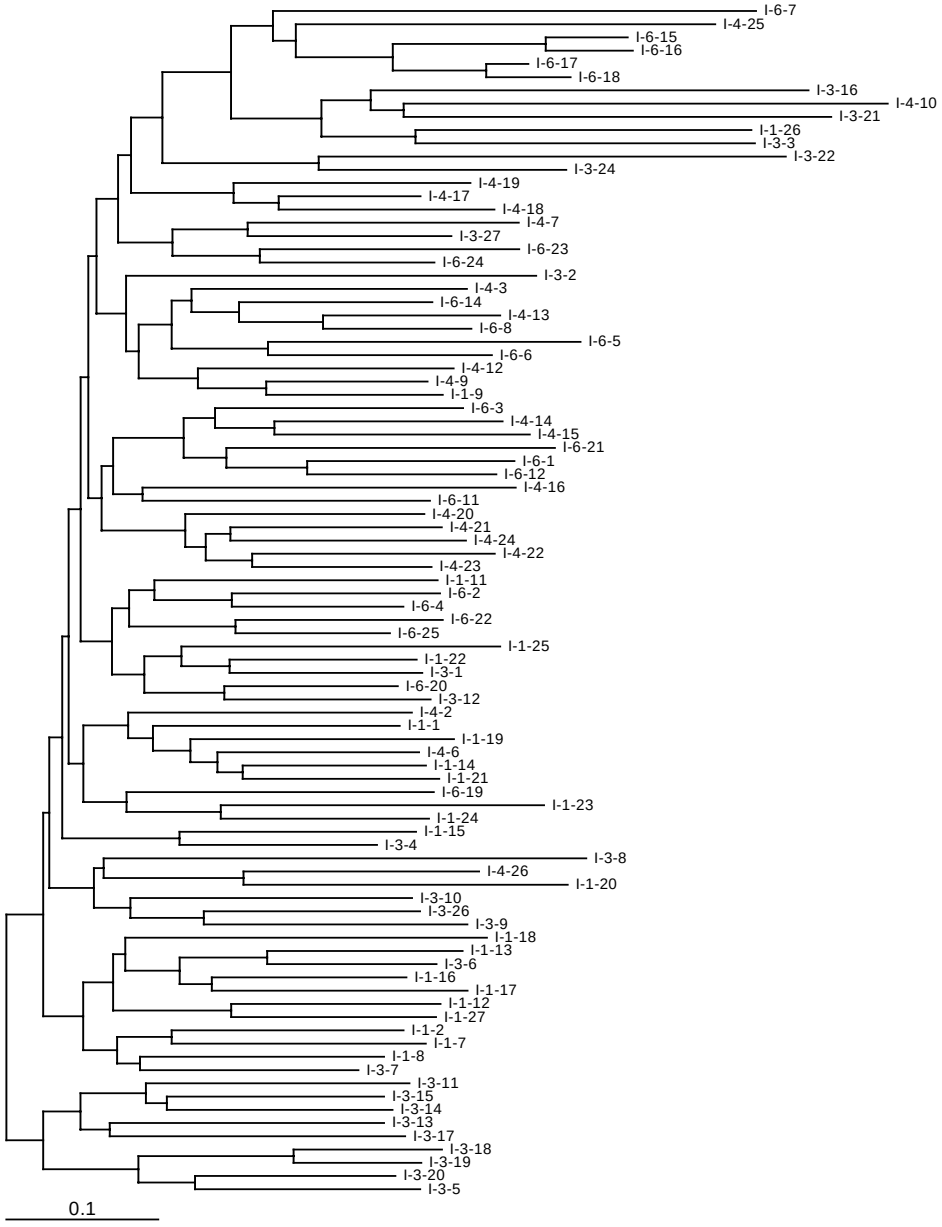
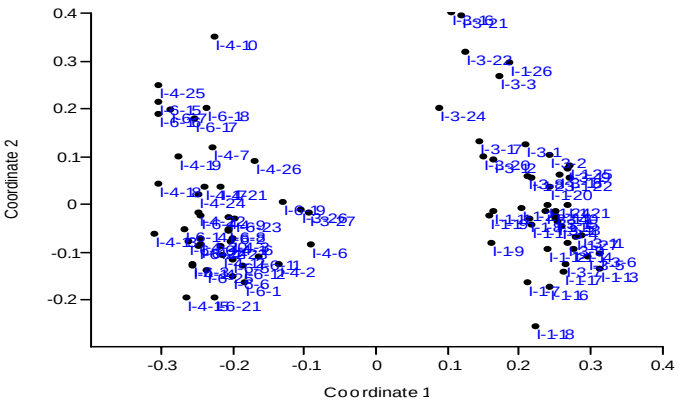


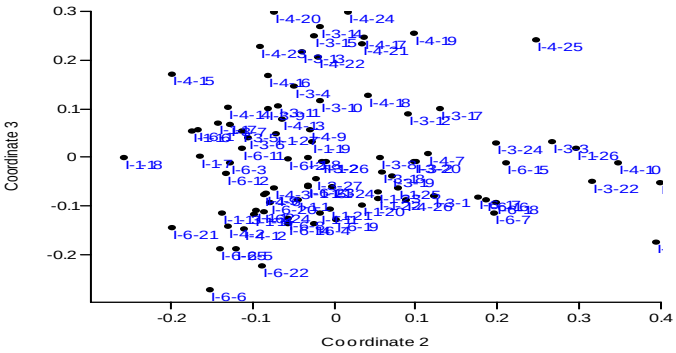
Figure-4.7. Dendrogram generated based on NJ analysis of complete intraregional data set of one hundred wild coffee based on two tetra-nucleotides (CoIS001 and CoIS002) data. The NJ tree was based on the Jaccard's coefficients of similarity from 38 ISSR data of hundred individuals of forest *Coffea arabica* L. The Roman number indicate the site (Bale),the first and the second Arabic number indicate the population and individual plants of the population respectively.

4.5. Principal Coordinate (PCO) Analysis

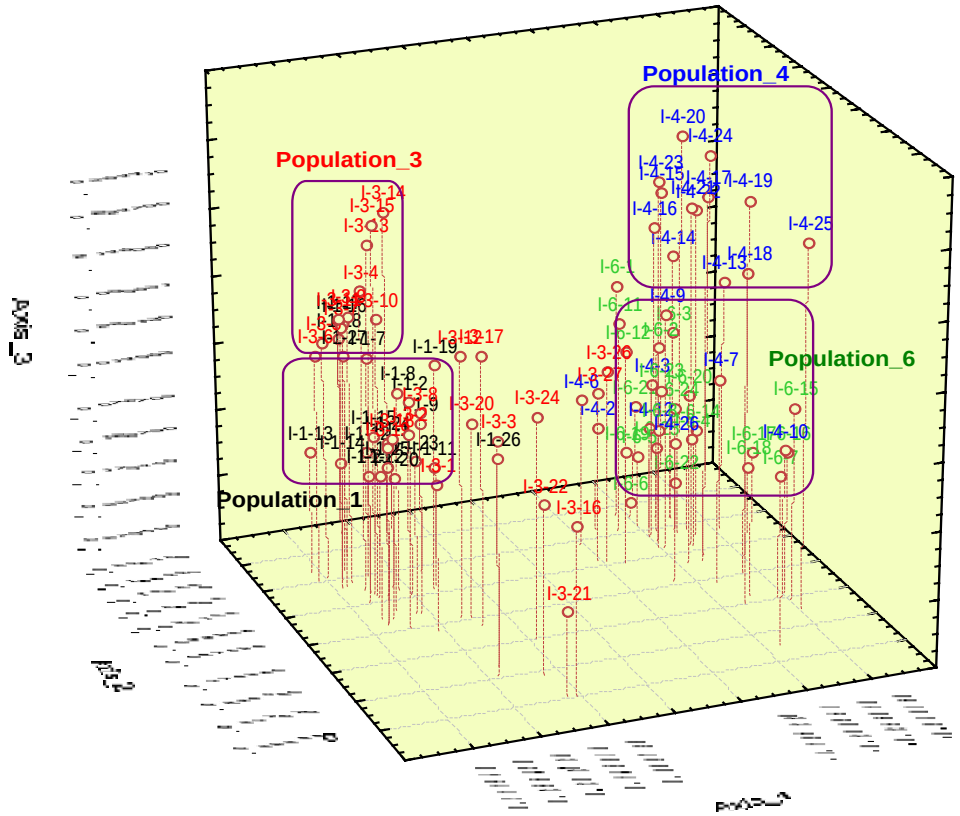
Because of the intra-specific nature of the *Coffea arabica L* individuals analyzed, the UPGMA and NJ algorithms extensively used migh not be appropriate for their efficient clustering. Due to this reason, it was decided to examine the clustering of the individual samples from all primers together; di and tetra-nucleotide were subjected to PCO analysis or metric multi-dimensional scaling to resolve the patterns of clustering among the one hundred *Coffea arabica L* individuals. The matrix of genetic similarity was also used in PCO analysis. Populations are arranged relative to their principal axis and are sorted by descending order of their eigen values which are equal to the total variation among the populations in all cases. While the plot of the first three PCO for all primers produced 4.76, 1.69 and 1.46 eigen values which in turn accounts for 14.8%, 5.3% & 4.5%, the three axis of PCO for di-nucleotides primer yielded 5.5, 1.5 and 1.1 which also accounts for 27.9%, 7.8% & 5.5% respectively. Similarly, the tetra-nucleotides generate eigen values of 4.0, 3.2 and 2.8 which accounts for 5.7%, 4.5 and and 4.1% respectively. These values were used to display the grouping levels of individuals of the populations using three co-ordinates. The two and three dimensional representation of PCO analysis per classes of primers is indicated in figure 4.8, 4.9 and 4.10 below. With an over all analysis, the two dimensional representation (2D) of PCO analysis revealed poor patterns of grouping on the basis of population and the individuals are more scattered and intermixed with each other. However, the three dimensional representation (3D) showed better patterns of groupings. Similarly, with di-nucleotide primers analysis, a clear patterns of grouping based on populations was observed in the case of 3D representation than 2D representation. Due to high polymorphisms in the tetra-nucleotides primes, no clear patterns of grouping were observed both with 3D and 2D representation of PCO analysis. A scatter plot of the ISSR depicts the levels of variation and similarity among four populations indicating how populations cluster in groups for each dimension as it can be seen from the figure below (Figure-4.8, 4.9 and 4.10).



I)

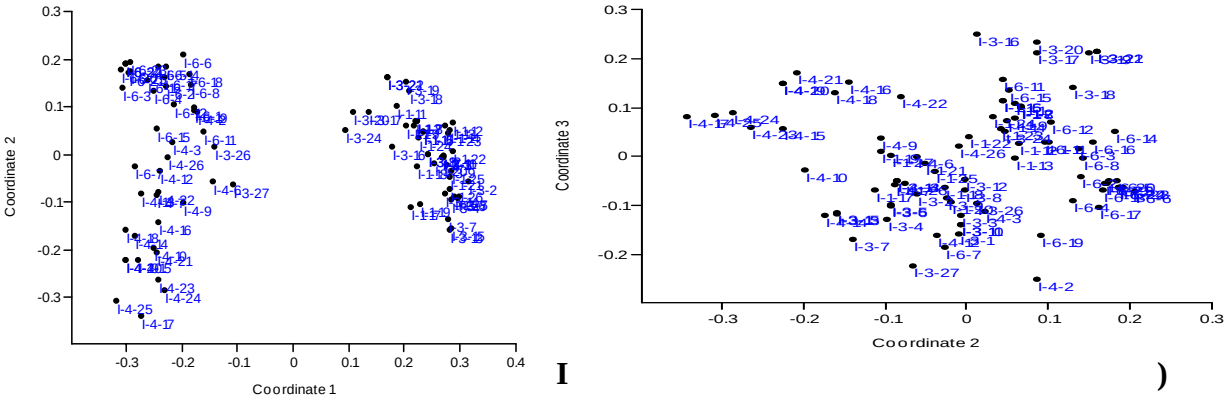


II)

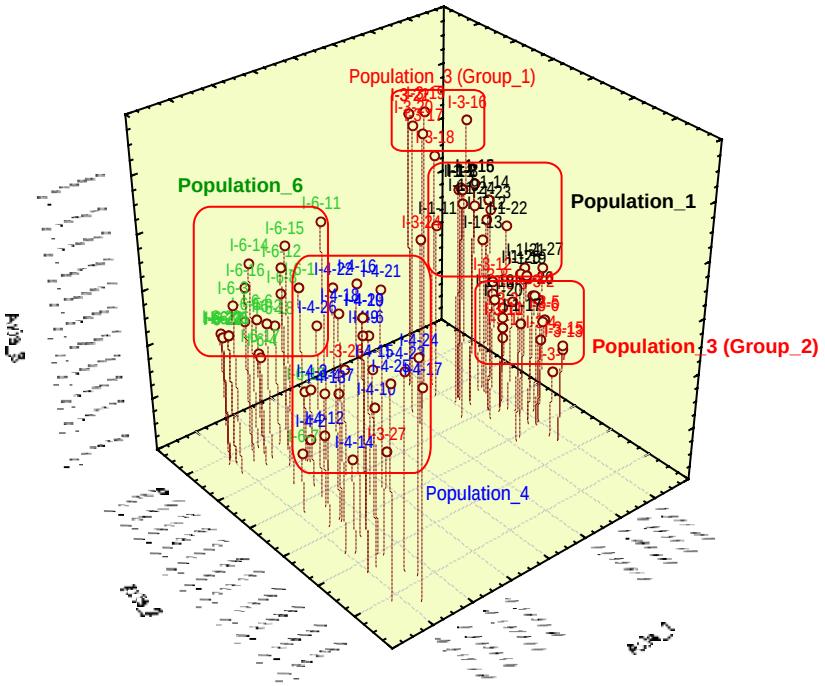


III)

Figure-4.8. The two (I and II) and three dimensional (III) representation of PCO analysis of the genetic relation ship among 100 individuals of wild *Coffae arabica* L. obtained from Harenna on the basis of Jaccard's similarity matrix of 137 ISSR markers using nine primers: seven di (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) and two tetra-nucleotides (CoIS001 and CoIS002).

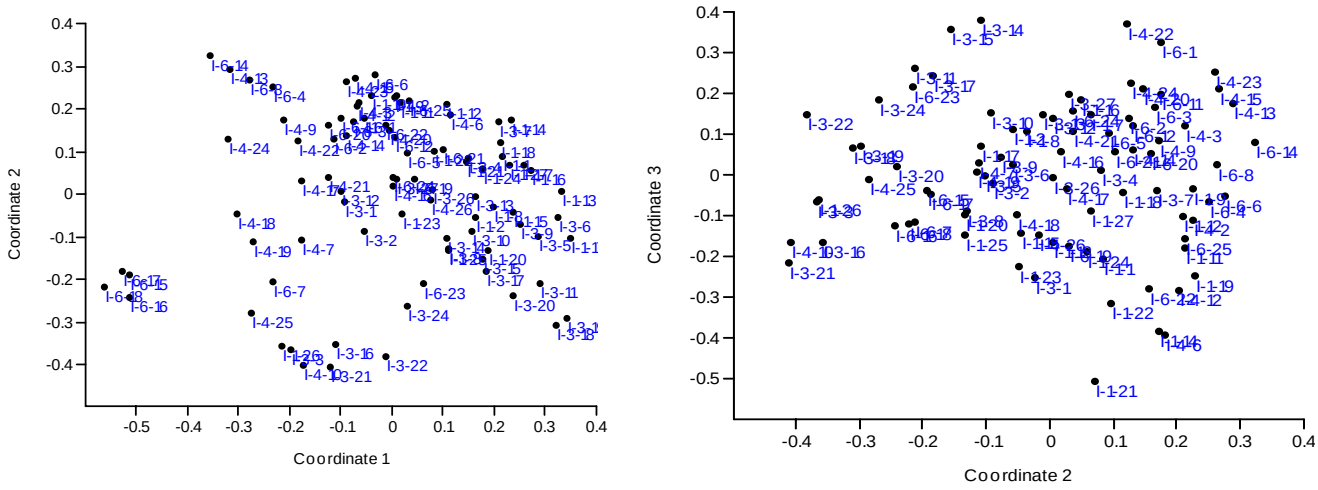


II)



III)

Figure-4.9. The two (I and II) and three dimensional (III) representation of PCO analysis of the genetic relationship among 100 individuals of wild *Coffea arabica* L. obtained from Harenna on the basis of Jaccard's similarity matrix of 99 ISSR markers using seven di-nucleotide (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) primers



I) **II)**

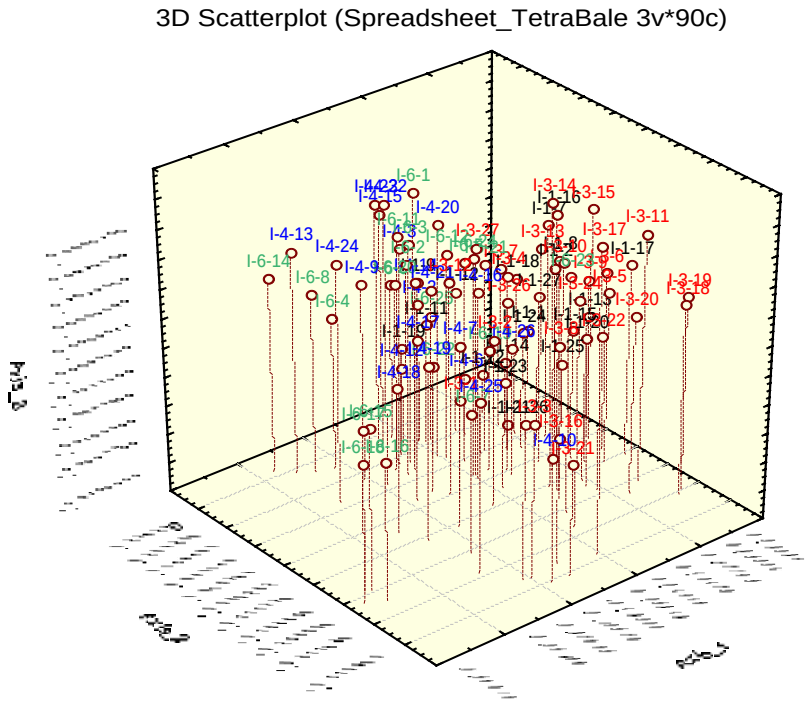


Figure-4.10. The two (I &II) and three dimensional (III) representation of PCO analysis of the genetic relation ship among 100 individuals of wild *Coffae arabica* L. obtained from Harenna on the basis of Jaccard's similarity matrix of 38 ISSR markers using two tetra-nucleotides (CoIS001 and CoIS002).

5. DISCUSSION

5.1. ISSR Marker Polymorphism in the entire Data set.

Currently, a number of molecular markers have been widely used to study diversity in many plants (Karp *et al*., 1996). Given the proliferation of molecular markers, a comparison between the markers seems highly inevitable on the basis of study objectives and the nature of the markers. Of many desired quality of molecular markers, automation (PCR-based), polymorphisms and reproducibility are the highly demanded features of the molecular techniques to be used in the intraregional diversity analysis. ISSR markers are thus one of the molecular markers that have these characteristics (Zietkiewicz *et al*., 1994, Wolf and Liston, 1998). ISSR markers are observed to be highly variable within the species and reveal many more polymorphisms since they use longer primers that allow more stringent annealing temperatures (Hillis *et al*., 1996). In this study, also the ISSR markers observed to be an appropriate molecular marker for generating the detailed intra-specific genetic diversity data to evaluate extent and distribution of genetic diversity within and among *Coffea arabica L*. Out of the total 137 scorable bands produced with the total of nine; seven di- and two tetra-nucleotides, 61 bands were polymorphic. In terms of number of polymorphic fragment detected and percentage of polymorphic loci, per classe of primer, tetra nucleotides were found to be superior. While the seven di-nucleotide primers generated 99 bands, the two tetra-nucleotide primers generated 38 bands, of which 26 and 35 were polymorphic loci respectively. Out of all primers employed in this analysis of genetic diversity, the two tetra-nucleotide primers showed unique patterns where almost all individuals were observed to have different banding patterns. Almost none of the plants were genetically identical; instead every individual plant were almost unique according to ISSR analysis using tetra-nucleotide primers, indicating that the levels of polymorphism or resolution of tetra nucleotide primer was extremely higher and sufficient to distinguish all genotype of the *arabica coffea L.* analyzed in the present study. In general the detection of high levels of polymorphisms make ISSR analysis with tetra nucleotides primers a powerful technique for measuring the intraregional genetic diversity in wild *Coffea arabica L.* The same patterns were reported by Tesfaye (2006) and Oljira (2006).

Because the number of bands generated and levels of polymorphisms observed were very high and made scoring very difficult, they, however, removed the tetra-nucleotide primers from data analysis. The result of the present study showed similar pattern and are comparable to their results obtained from analysis with di-nucleotide primers. Moreover, Raus *et al.* (2003) obtained 100% polymorphism with five tetra primers including CoIS001 and CoIS002 in the study of genetic relation ship in *Coffea* species and parentage determination of inter-specific hybrid crosses. Moreover, Joshi *et al*. (2000) reported the high resolution power of ISSR primers in the determinations of diversity and phylogenetic relationship of rice. Forty two individuals including wild and cultivated rice were analyzed using thirty ISSR primers representing di, tri, tetra and penta nucleotide repeats. The result indicated the usefulness of ISSR primers in studying genetic components of plant species. These all clearly indicate that the extent and distribution of tetra-nucleotide repeat polymorphism in *C. arabica* genome studied here is very high and every individual coffee plant is almost unique for these primers.

5.2. Intraregional Genetic Variation

Using the over all nine ISSR primers: seven di (810-H, 812-H, 813-H, 814-H, 818-H, 834-H and 844-H) and two tetra-nucleotides (CoIS001 and CoIS002) we found evidence of intraregional genetic diversity between sub-divided populations of the same region. The results of the genetic diversity study indicated the presence of moderate to high levels of genetic diversity in wild coffee within and among these sub-divided populations in a region on the basis of diversity measurement parameter such as percentage of polymorphic loci and Shannon's diversity index.

The genetic diversity within population ranges from P=29.20% for population P1 (Bale-1) to P=56.56% for population P3 (Bale-3) both of which are from semi-disturbed region. However, the genetic diversity for Population P4 and P6, both of which are from undisturbed regions was found to be equal with P=30.66% slightly larger than population P1 with least genetic diversity (Table-4.2). This means that the genetic diversity was high for Bale-3 and lower for Bale-1, both of which from semi-disturbed region, indicating that

higher and lower variability were obtained from semi-disturbed regions of the forest. Both higher and lower variability resulted from semi disturbed region. This could be attributed to the levels of management implemented in the two populations. In this case Pop3 has been well managed than population Pop1 since population Pop3 is very close to Magnette village. The closer the plot to the village, the more chance to replace the standing original populations with more uniform high yielding one. However, Tesfaye(2006) observed general trends of a slightly higher diversity of semi-disturbed plots than undisturbed one in south western Ethiopia. Similarly, Oljira (2006) reported the general patterns of high genetic diversity in semi-disturbed region than undisturbed region in the Yayu forest of Ethiopia (SW). On the basis of percent polymorphisms, the total genetic diversity for the whole populations was 44.53%. A relatively larger genetic variability was observed in this study which is comparable with what Tesfaye (2006) found. Shannon's diversity index ranged from 23.73% for population 1 (Bale-1) to 30.08 % for population 3 (Bale-3) with di-nucleotide primer analysis and followed the same patterns of genetic diversity as with percentages of PL above in the case of assessment with an overall and tetra-nucleotide primer (Table 4.3). The same pattern of intraregional genetic variation was obtained by Tesfaye (2006) in materials collected from south west forest (Brhane Kontir and Yayu) of Ethiopia using six di-nucleotide primers. However, the observed diversity of Bale (Harenna) is slightly lower than that of Yayu and Brahane Kontir This could be because of the absence of forest coffee other than Harrena. The other possible reason could be fire incident in Bale which is common as compared to South West. Moreover, extreme environment in both sides of the Harrena forest (Chilling temperature on the top of Bale Mountain and extremely dry and higher temperature from the lowland of Bale) could also be other reasons.

5.3. Genetic Similarity and Relationship

Cluster and PCO analysis were employed to observe the genetic relationship in one hundred individual *Coffea arabica* L. The UPGMA and NJ dendrogram using 137 ISSR bands resulted in slightly different tree topologies.

The UPGMA dendrogram has two main cluster or group in which individuals from semi disturbed and undisturbed regions are separated. The grouping pattern observed was similar to those of Oljira (2006). These could be because of human selection process in the forest for particular traits. Both main clusters further separated into some sub-clusters forming four groups that were being clearly distinguished from each other except for few inter-mixing. Generally, individual from the same populations were largely grouped together. Further more, it was observed that the samples from semi-disturbed and undisturbed populations tend to show a stronger grouping according to their population of origin. The semi-disturbed populations tend to be more scattered and inter mixed with individuals from the other populations as compared to the undisturbed populations. These patterns of relationship were observed in an overall analysis and with di-nucleotide primers. However, weak grouping was obtained with tetra-nucleotides and thus individuals from one population were found to be highly inter-mixed with individuals from other populations and there were no clear distinction between populations. This is also observed in the analysis of Yayu and Brhane Kontir in SW Ethiopia whereby 100% polymorphism of fragments wasobserved in the tetra primers. Similarly, in an overall analysis and with di-nucleotide primers, the NJ dendrogram constructed grouped individuals into three main clusters. Cluster one consists of individuals from population four and six which form sub-clusters partially on the basis of their vulnerability to external pressure (i.e., they are from undisturbed wild forest). These are the group on the top of the tree (cluster one) that is more dominated by population four (Figure-4.5). Similarly, cluster two includes individuals from population one with out any inter-mixing from the other populations. In addition, cluster three consists of individuals from population one and three that were grouped together on the basis of the sites from which the samples were collected (i.e., they are from semi-disturbed forest and close to each other). This group is

those found at the bottom of the tree which is more dominated by population-3. For tetra-nucleotide the same patterns of relationship was obtained for NJ as with UPGMA analysis (That means, NJ and UPGMA have similar patterns of grouping for the tetra-nucleotide primers). In the case of NJ analysis, some of the individuals from each population appear to have longer branches which indicate that they are more variable. That means, individuals escaped from the population might have accumulated many diverse adaptive gene complexes adapted to environmental changes than the other individuals of the same population. Generally, the dendrogram analysis using *Coffea arabica L*. individual plant form cluster on the basis of their respective populations since low levels genetic variation is detected in all populations investigated. However, the tree or dendrogram based on the tetra repeat seems to be not well resolved due to the higher polymorphism of these primers as compared to trees with the di-and all primers together.

PCO analysis was performed to graphically display the genetic associations between the four populations considered. The orientation of the axis was defined by coordinates, whose lengths are given by eigen values, which provide information's about the dimensionality of the data and how the individuals of the populations are related to each others and the main axes. The three dimensional representation of the PCO analysis was observed to have better grouping of individuals in to their respective population except a few intermixing than the two dimensional representation. Using two dimensional scaling of PCO, ISSR data revealed no clear distinction between the four populations and the scatter arrangement observed (Figure 4.8, 4.9 and 4.10). However, in the three dimensional scaling there are obvious grouping linking populations together based on their respective origin as it can be seen from the scatter plot except for the tetra-nucleotides cases. Similar populations are placed close together with strong affinity (e.g. Population 1 and 3) and dissimilar population far apart with low affinity (Population P4 and P6) suggesting that the 3D indicate great resolution as compared to the 2D. Moreover, 3D scaling of PCO analysis showed better resolutions compared to NJ and UPGMA which can only revealed linear relationships of the individuals in the populations like 2D scaling of PCO analysis. Those are, strong grouping of the individual plant in to their respective populations as well as among semi-disturbed and undisturbed groups in a region were

observed. However, in the case of tetra-nucleotide, 3D was not revealing clear grouping like 2D and this could also again is linked to the higher variability of the tetra primers. PCO also revealed that population Pop3 and Pop4 were clearly separated from the rest with few intermixing. Further more population Pop1 and Pop6 tends to form separate group but with more intermixing from the other populations.

5.4. Distribution of Genetic Diversity

5.4.1 Estimation of within and among population genetic diversity

Among population and within population, genetic diversity was estimated by AMOVA and Shannon's diversity index. In *Coffea arabica L.* population, higher among population diversity was expected than within population genetic diversity as the plant is predominantly self-pollinated species (Raus *et al*., 2003). Moreover, Mayer (1965) also observed higher variation among wild and semi-wild *Coffea arabica* populations' from montane rain forests of Ethiopia as compared to the cultivated materials. However, contrary situations were observed in this study. The result generally indicated more genetic diversity within populations (83.60%) rather than among populations (16.4) in an over all analysis by Shannon's diversity index analysis. Oljira (2006) observed the general trends of the larger within population (0.573) than among population (0.427) Shannon's diversity from genetic diversity study conducted at Yayu forest of Ethiopia. Similar patterns of Shannon's diversity index were observed with di- and tetra-nucleotides primers independently.

Analysis of molecular variance also indicated that most of the ISSR diversity was distributed between individual plants within the populations (56.8 %), with the remaining diversity distributed among populations within groups (11.0%) in the case of grouped populations (i.e. semi-disturbed and undisturbed groups) and among populations (33.2%)(Table-4.4). Moreover, AMOVA with out grouping populations revealed that out of the total genetic diversity, 63.7%, is found between individual plants within the populations and the rest 36.4% among populations (Table 4.5). Although the levels of

variations are different, the patterns of this result were similar to the result of Tesfaye (2006) obtained from studies on samples from south west Ethiopia (Brhane Kontir and Yayu). The result of this study indicated high genetic diversity within populations (63%) than between population (37%) in Brhane Kontir wild coffee populations. Similar patterns of high genetic diversity within populations (53%) than between population (47%) were obtained in Yayu. In addition, AMOVA result of Oljira (2006) supports the larger within population variation (54.06) as compared to among population variation (45.94). In both cases i.e., with or without grouping, the results of analysis of molecular analysis showed the same patterns. The higher within populations' genetic diversity, rather than among population diversity, might be accounted to two contrary reasons. Like any other plant species, the population genetic diversity of a *Coffea arabica* L. species is affected by multiple evolutionary forces which operate within historical and biological context of the plant species. This includes the mating types, gene flow, mode of reproduction and natural selection etc. (Hamrik and Godt, 1989). For this reasons, it could be speculated from the result that *Coffea arabica L.* might have mixed mating system (partial out-crossing by pollen and seed and partial selfing) for which some extent of gene flow is expected as reported by Loveless and Hamrik, (1984) which could result in high within genetic diversity. The outcrossing nature of the *Coffea arabica* L was reported by some investigators, and rates of outcrossing reported by different author are different. For instance, Meyer (1965) showed that the natural cross-pollination rate was 40% to 60% in wild population of *Coffea arabica* L. Carvalho (1974) on the other hand reported the low out crossing rates of *Coffea arabica* L. cultivars with proportion of 7% to 15% which would probably be affected by insect and birds. Although these report support the current result the reported rates was not confirmed in the study areas and need to be investigated.

In contrary to the above, it could also be speculated that the high genetic diversity observed within populations in our investigation of the *Coffea arabica L.* population studied might be due to preferential adaptive gene complexes adapted to environmental changes being evolved during long evolutionary period in the region. This might be due to the location of the sample collection from Harenna Forest of Bale Mountain, where it is believed to be one of the primary diversification centre and origin for the genus *Coffea*

arabica. In this case, *Coffea arabica* L population uses selfing as mechanisms to prevent influx of the gene from another portion of the populations that might reduce diversity through disrupting adaptive genes (Lowe *et al*., 2004). In both context, the fact that the *Coffea arabica* genetic resources have decreased markedly recently due to over exploitation for different purposes and shrinking of their natural habitat, indicates that the threats to the survival of the species mainly come from human activity.

6. CONCLUSION.

In the present study, nine ISSR primers; seven di and two tetra-nucleotides were employed. They were able to reveal that genetic diversity ranged from moderate to high levels and identified the high diverse and least diverse populations in Harenna forest. The tetra-nucleotide primers surprisingly showed high levels of genetic diversity as compared to the di- nucleotides. This clearly implied that tetra-nucleotide ISSR primers are appropriate for the intraregional diversity analysis of the sub-divided populations. While high genetic diversity was observed in Pop3 from semi-disturbed region, moderate levels of variation was shown in Pop4 and Pop6, both from undisturbed regions. These assessments of the levels and extents of genetic diversity indicated that the Harenna forest is a valuable site for *Coffea arabica* conservation. Although the abundance and genetic diversity of coffee in the Harenna forest has been significantly vanished by deforestation, the existing genetic diversity of wild *Coffea arabica* population in the region is still found to be moderate to large but require urgent conservation for future use and breeding program. Until the present day, information available on the reproductive biology of *Coffea arabica* suggested that it is a predominantly self-pollinating plant. However, the result of this study might be attributed to two reasons one against and the other in favor of the self-pollinating nature of the *Coffea arabica* plant. In the former case, the result obtained could be accounted to mixed type of mating, typical of plant species, in which there is a gene flow, and thus there may be moderate gene flow among the local populations by effectors such as wind, insect, human (seedling movement) and birds. The other is they might have preferential or diverse adaptive genes that are not fixed through self-pollination until the present day.

7. RECOMMENDATIONS: Implication for Conservation and Breeding

7.1. Implication for Conservation.

Gonzalez *et al*. (2005) reported that accurate and fast molecular marker based assessment of the levels of genetic diversity and degree of genetic relatedness are in demand for the efficient management of the existing genetic resources either for conservation purposes or for breeding program utilization. For instance, molecular markers are needed to generate information about the extent and distribution of genetic diversity which is essential for the conservation of genetic resources. In this context, one of the objectives of the present analysis was to generate genetic diversity information to develop appropriate conservation strategies for *Coffea arabica* in Harrena forest. This study provided information on the genetic variation at the intra-population levels of a region in Harenna *Coffea arabica* plant and helps to identify candidate populations for conservation. This study suggested that the programs of *in-situ* conservation should be implemented to create strategies for maintaining wild coffee populations' genetic diversity by appropriate forest management since diverse populations were observed in the managed area of the forest. This *in-situ* conservation strategy is considered to be most important that allows the dynamic evolutionary events to continue in the original sites. However, care should be taken during forest management not to disrupt other biodiversity components in the forest to avoid loss of natural forest to the point of no return.

Alternatively, Berthud and Charrier (1998) suggested that genetic diversity can also be increased by establishing coffee field gene bank where collections of the coffee genetic resources are conventionally maintained as living trees or shrubs. Since the coffee field gene bank should be established at a nearby area where management could be easily implemented, it would provide environment nearly similar to the forest. In the case of Ethiopia, currently, IBC (Institute of Biodiversity Conservation) has two *ex situ* field gene bank site in SW (Chocie) and Easter (Badessa) part of the country. This could act as part of the maintenance of diverse coffee populations.

7.2. Implication for Breeding.

The present investigation provided information's not only to design appropriate conservation strategies, but also genetic diversity information needed in breeding programs. Breeders are in demand of genetic diversity information in order to focus on the populations with high genetic variability to select parental coffee for crossing. By crossing genetically more diverse parent, the level of variation present in the segregating populations can be maximized. The populations with moderate variability from unmanaged areas are observed to have very unique genotypes hence, special attention or priority should be given to these populations. The increment of genetic diversity through breeding mechanisms in the subsequent hybrid populations is evidenced from some previously reported studies. For instance, Amha (1990) reported breeding as an appropriate mechanism to significantly and rapidly increase productivity in hybrid varieties which is attributed to presence of diverse genotype. Based on his investigation, he observed 30 to 69% heterosis in hybrid varieties over the better parent in Ethiopia *Coffea arabica L.*

7.3. Recommendations

- The patterns of genetic diversity obtained in this study suggested that:
  - Further intraregional analyses using SSR sequence of tetra-nucleotide primers are required.
  - Rates of self- pollination and cross-pollination require in-depth investigation to account for the speculated gene flow.
  - Co-dominant, especially SSR based intraregional genetic diversity analysis should be carried out to provide heterozygosity based genetic diversity information and confirm the high within and low among population diversity observed in this study
  - The levels of management should be investigated and standardized so as to optimize the management of the forest and the coffee as well.

8. REFERENCES

Aga, E. (2005). Molecular Genetic Diversity Study of Forest Coffee-Tree (*Coffea arabica* L.)Populations in Ethiopia: Implications for Conservation and Breeding. Doctorial Thesis, Swedish University of Agricultural Sciences.

Akkaya, M.S. and Buyukunal-Bal, E.B. (2004). Assessment of genetic variation of bread wheat varieties using microsatellite markers. *Euphytica,* **135**: 179-185.

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.

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

Aranzana, M.J. J. Carbo and P. Arus. (2003). Microsatellite variability in peach-*Prunus persica* (L) Batsch- cultivar identification, marker mutation, pedigree inferences and population structure. *Theor. and Appl. Genet.* **106**: 1341-1352.

Bekele, E. (1985). The biology of cereal landrace populations. *Hereditas* **103**: 119 – 134.

Bekele, E. (1986). The interfaces of ecology, genetics, and biosystematics in conservation. *Symb. Bot. Ups. XXVI.* **2**: 101 – 113.

Berthud and Charrier.(1988). Genetic resources of *Coffea*. **In** : R.J. Clark and R, Macrae (eds) *Coffee Vol. 4*. Agronomy Elsevier Applied Science, London and New York. Pp1– 42.

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

Bremer, B. (1996).Combined and separate analysis of morphological and molecular data in the plant family Rubiaceae. *Cladistics,* **12:**21-40.

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

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

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.

Carvalho, A., Ferwerda, F.p., Frham-Leiveld, J.A., Medina, D.M., Mendes, A.J.T. and Monaco, L.C. (1969). Coffee (*Coffea arabica* L. and *C. canephora* Pierre ex Froehner. **In**: *Outlines of Perennial Crop Breeding in the Tropics* (Ferwerda, F.P.and Wit, F., eds.) Veenman and Zoneh, Wageningen, Pp. 189 -241. Chang, R.Y., O' Donoughue, L.S. and Bureau, T.E. 2001. Inter. MITE.

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.

Charrier, A.and Berthaud, J. (1985). Botanical classification of coffee. **In**: Clifford, M. N. and Wilson,K.C.(eds), *Coffee Botany,Biochemistry and Production of Beans and Beverage*, Croom Helm, London, Pp.13-47.

Charrier, A.and Berthaud, J. (1990). Use and values of genetic resources of *Coffea* for breeding and their long term conservation. *Mitt Inst. Allg. Bot. Hamburg* 23a, 65 – 72.

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.

Combes, M.C., S. Anderzejewski, F. Anthony, B. Bertrand, P. Rovelli, G. Graziosi and P. Lashermes (2000). Characterisation of microsatellite loci in *Coffea arabica* and related coffee species. *Molecular Ecology.* **9**: 1178-1180.

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. *Ame. J. of 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 the Linnean Society*.**15:** 465 – 512.

Denich, M., Gatzweiler, F. and Gole, T.W.M. (2006).”Prosperity and poverty in a globalized world Challenges For agricultural Research”: *Conservation and use of Wild Coffea Arabica Populations in Montane Rainforests of Ethiopia Forest Ecology and Management* **232**: 68- 74

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**:*The Proceedings of the Workshop on Control of Coffee Berry Disease in Ethiopia*, pp. 125-135. Ethiopian Agricultural Research Organization (EARO), Addis Ababa.

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.

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

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

Dullo, M.E.., Guarino, L., Engelman, F., maxted, N., Newberry, J.H., Attena, F. and Ford-LIoyd, B.U. (1998). Complementary conservation strategies for the genus *Coffea*: a case study of mascarena *Coffea* species. Genetic Resour. *Crop. Evol*. **45**: 565 – 579.

Engles, J.M.M. and Hawkes. (1991). *The Ethiopian Gene Centre and its Genetic Diversity*. Cambridge University Press. Cambrige, UK. Pp 354 – 359.

Eskes, A.B. (1989). Identification, description and collection of coffee types in P.D.R. Yemen. IPGRI, Rome.

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.

Food and Agriculturs1 Organization (1968). FAO Coffee Mission to Ethiopia.1964-1965. Report, FAO. Ed., Rome (Italy), 200p.

FAO (2003). Molecular marker assisted selection as a potential tool for genetic improvement of crops, forest trees, livestock and fish in developing countries. *Biotechnology in Food and Agriculture*. Conference 10: 17 November – 14 December

Faria, P.J. and Miyaki, C.Y. (2006).Molecular markers for population genetic analysis in the family Psittacidae (Psittaciformes, Aves). *Genet. Mol.Biol*..**29 (2)**. Pp213- 227.

Firris, I.B. (1979). The wild populations of *Coffea arabica* L. and cultivated coffee. **In**: *Taxonomic Aspects of African Economic Botany* (ed., Kunkel, G.), Proceeding of the IX Plenary of AETFAT< Las paqlmasde Gran Canaria, 18 -23 mark, 1978.

Gebre-Egzihabher, T.B. (1990). The importance of Ethiopian forests in the conservation of the arabica coffee gene pools. **In:** Ihlenfeldt, H.D., proceedings of the 12th plenary Meeting of AETFAT 1988, Humberg. Pp 65 – 72.

Girma, G. (2007). Relationship between wild rice species of Ethiopia with cultivated rice based on ISSR marker. M. Sc Thesis (Submitted). Addis Ababa University, Addis Ababa, Ethiopia.

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

Gole, T.W., Denich, M., Teketay, D. and Borsch, T. (2001). Diversity of tradditional coffee production systems in Ethiopia and their contributions to the conservation of coffee genetic diversity. Conference on International Agricultural Research for Developpment. *Deutscher 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**: R. Rao, A. Brown, M. Jackson (eds.) *Managing Plant Genetic Diversity*. CABI and IPGRI, 237-247.

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

Gomez, O.J., M.W. Blair, B.E. Frankow-Lindberg and Gullberg, U. (2004). Molecular and phenotypic diversity of common bean landraces from Nicaragua. *Crop Science*. **44**: 1412-1418.

Gonzalez, A., Wong, A., Delgado-salinas, A. Papa, R.and Gepts, P. (2005).Assessment of Inter Simple Sequence Repeat (ISSR) Markers to differentiate sympatric wild and domesticated populations of common bean. *Crop Sci*. **45**:606-615.

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

Gupta, P.K. (2003). *Cell and Molecular Biology* 3 rd ed. RASTOGI Publications SHIVAJI ROAD,India

Haarer, A.E. (1962). *Modern Coffee Production*, 2nd ed. London.

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.

Herrera, J.C., Combes, M.c., Cortina, H.and Lashermes (2003).Factors Influencing gene introgression in to the alloteraploid *Coffea arabica* L.From its diploid relatives. *Genome* **47**(6):1053-1060.

Hillis, D. M., C. Moritz and B .K. Mable (1996). *Molecular Systematics*. 2nd ed. Sunderland, MA: Sinauer Associates, Inc.

Hillman, J.C. (1986). Bale Mountain National Park Management Plan. Ethiopian. Wild Life Conservation Organization. Addis ababa, Ethiopia

Huang, X.Q., A. Borner, M.S. Roder and M.W. Ganal. (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theor. and Appl. Gene.*. **105**: 699-707.

IPGRI (1965). Crop Genetic Resources: An Economic Appraisals/EIB-2, Economic Research Servicee/ USDA, IPGRI, ROME, Italy.

IPGRI (1996). Description for Coffee (*Coffea* sp. and *Psilanthus* sp.). International Plant Genetic Resource Institute, Rome.

IPGRI (2005). Crop Genetic Resources: An Economic Appraisal/EIB-2, Economic Research Service/USDA, IPGRI, Rome, Italy.

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

Joshi, S.P., Prabharkar, K. RanjekAR AND Vidya, s. Gupta (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., S. Kresovich, K.V. Bhat, W.G. Ayad and T. Hodgkin. (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.

Kufa, T. (2006).Ecophysiological diversity of wild arabica coffee population in Ethiopia: Common water relation hydraulic characteristis along a climate gradient. Doctoral Thesis.Center for Development Research (ZEF) University of Bonn, Germany.

Lashermes, P., Cros, J. Marmey, P. and Charrier, A. (1993). Use of random amplified DNA markers to analyze genetic variability and relationships of *Coffea* species. *Genetic Resources and Crop Evolution*. **40**: 91-99.

Lashermes, P., Trouslot, P., Anthony, F., Combes., 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., Combes, M., C., Robert, J.,Trouslot, P., D'Hont,.P. A., Anthony, F. and Carries, A. (1999).Molecular characterization and origin of the *Coffea arabica* L. Geneome. *Mol. Gen. Genet*. **26**:259-266).

Lashermes, P.J., Andrzejewiski., Bertrand, B., Combes, M.C., Dussert, S., Graziosi, G., Trouslot, P. and Anthony, F. (2000). Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theor. Appl. Genet.* **100**: 139-146.

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

Leory, J.F. (1980). Evolution et taxogenese Ches les cafeiers: hypoehese sur l'origine C R A cad Sci (paris). **291** : 593 – 596.

Louarn, J. (1998).Detection of genetic diversity and selection gene introgression in coffee using RAPD markers *Theor.Appl.Genet*.**87**:934-940.

Lowe. A., Harris, S. and Ashton, P.(2005). *Ecological Genetics*: *Design, Analysis and Applications*. Blackwell Publishing. Australia.

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.

Masumbuko, l., I., Bryngelsoson, T., Mneney, E.E.and Salomon, B. (2003). Genetic diversity in Tanzania arabica coffeeUsing random Amplified Polymorphic DNA (RAPD) markers. *Hereditas*.**139**:56-63.

Mesfin, A. (1991). Siginificance of Ethiopian coffee genetic resources to coffee improvements. **In:** Plant Genetic Resources of Ethiopia, (Engles, J.M.M., Hawkes, J.G. and Melaku Worede, eds). Cambridge University Press Cambridge, UK, Pp 354 – 359.

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

Monaco, L.C. (1968). Considerations on the genetic variability of the *Coffea arabica* populations in Ethiopia. **In**: FAO (Eds), *FAO Coffee Mission to Ethiopia* 1964-65, FAO, Rome, Italy. Pp 49 – 69.

Montagnon, C. and Bouharmont, P. (1996). Multivariate analysis of phenotypic diversity of *Coffea arabica* . *Genetic. Resour. 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., R. Peakall, R. Appels, L.R. Preston and H.L. Lloyd, (1995). DNA profiling techniques for plant variety identification. *Australian Journal of Experimental Agriculture* **35:** 807-819.

Nagaoka T. & Ogihara Y. (1997) Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theor. and Appl. Genet.* **94**: 597–602.

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..

OXFAM (2002). *Bitter Coffee: How the Poor are Paying for the Slump in Coffee Prices.* Oxfam International, London.

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

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

Pearl, H.M., C. Nagai, P.H. Moore, D.L. Steiger, R.V. Osgood and R. Ming. (2004). Construction of a genetic map for arabica coffee. *Theor. and Appl. Genet.* **108**: 829-835.

Pendegrast, M. (1999). *Uncommon Grounds: The History of Coffee and how it Transformed our World.* Basic books, New York, USA.

Pradeep Reddy M., sarla,N. and Siddif, E.A.(2002). Inter Simple Sequence Repeat (ISSR) polymorphism and its application in plant breeding. *Ephytica*, **128**. Pp : 9-17.

Purseglove,J.W.(1968).*Tropical Crops :Monocotyledons*. Longman group limited, London, 605pp

Raina, S.N., Mukai, Y. and Yananoto, M. (1998). In situ hybridization identifies the Diploid progenitor's species of *Coffea arabica* (Rubiaceae). *Theor. Appl. Genet*. **97**: 1204 – 1209.

Rao, N.K. (2004). Palnt genetic resources: advancing conservation and use through biotechnology. *African J . Biotech*. **3**: 136 – 145.

Richenzhagen, C. and Virchow, D. (2002). Sustainable Utilization of crop Genetic Diversity through Property rights mechanisms: The case of Coffee Genetic Resources in Ethiopia. Bio Econ. Workshop, may30-may 31, 2002.

Rolf, F.J. (2000). NTCYS-pc v 2.11T Numerical Taxonomy and Multivariate Analysis system.Exter software.New York.

Ruas, P.M., Ruas, C.F., Rampim, L., Carvalho, V.P., Ruas, E.A. and Sera, T. (2003). Genetic relationship in *Coffea* species and parentage determination of inter-specific hybrids using ISSR (Inter Simple Sequence Repeat) markers. *Genet. and Mol. Biol*. **26**: 319 – 327.

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

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 and Management*. **232**:68–74.

Singh, P. (1996). *Essentials of Plant Breeding*, 1st ed. Kalayani pubilishers, Ludhiana. New Delhi.Noida (U.P). Hyderbad. Madras. Culcutta. Cuttack.

Smith, J.S.C. and Smith, O.S. (1989). The description and assessment of distances between inbred lines of maize: The utility of morphological, biochemical and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. *Maydica* **34**: 151-161.

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

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

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

Strenge, H. (1956). Wild coffee in Kaffa Province of Ethiopia. *Trop. Agr*. **33**: 297-301.

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

Sylvain, P.G., 1958. Ethiopian Coffee: Its significance to world coffee problems. *Botany* **12**: 111-139.

Taddese,M. and Nigatu, L. (1996). An ecological and ethinobotanical study of wild or spontaneous coffee, *Coffea arabica* in Ethiopia. In: Vander Maesen,J.G.., et al.(eds). *The Biodiversity of African Plants*. Pp 277 -294.

Tadesse, M. and Nigatu, L. (1989). An ecological study of the vegetation of Harenna forest. Bale, Ethiopia. *SINET: Ethiop. J. sci.* **12**(1):63-93 .

Teketay, D. (1999). *History, Botany and Ecological Requirements of Coffee*. Ethiopian Agricultural Research Organaization. Addis Ababa, Ethiopia.

Tekatay, D., Paulos Dubale and Ababu Anage (1998). Report on the first field trip. Report of the team of experts organized to prepare a FCCP proposal. Momeagraphed, Addis Ababa.

Tesfaye, K. (2006).Genetic Diversity of wild *Cofea arabica* populations in Ethiopia as a contribution to conservation and use planning. *Ecology and Development Series*. University of Bonn,Germany. No.44.

Tesfaye, K., Govers, K., Bekele, E. and Borsch, T. (2005). ISSR fingerprinting of *wild coffea* arabica in Ethiopia reveals high levels of genetic diversity within regions. BioTeam Status seminar, Bonn, Germany, March 14th – 16th .2005.

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.

Twyman, R.M. (2003). *Advanced Molecular Biology*: A *Concise Reference*.3rded. Viva Books Private Limited,New Dehi .

Vega, Fernando E., Rosenquist, E., and Collins, W. (2003). Global project needed to Tackle coffee crisis. ‘ A sharp drop in coffee prices has caused widespread suffering and hindered research’. *Nature Publishing Group. Vol*. **425**. pp 343.

Vos, P., R. Hogers, M. Bleeker, M. Rijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper andM. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. **23**:4407-4414.

Wellman, F.L. (1961). *Coffee: Botany, Cultivation and Utilization*. World Crop Series. Inter Science, New York.

Weising, K., Nybom, H.,Wollf, K. and Kahl, G. (2003). DNA Fingerprinting in Plants: *Principle, Methods, and Applications*.2nd ed.Taylor and Francis, London.

Werlemark, G., M. Uggla and H. Nybom, 1999. Morphological and RAPD markers show a highly skewed distribution in a pair of reciprocal crosses between hemisexual dogrose species, Rosa Sect. Caninae. *Theor. and Appl. Genet.* **98**: 557-563.

Whitemore, T.C.(1997). Tropical Forests disturbances and species loss. **In**: Laurance, W.F. and Bierregaard, R.O.(Eds), Tropical Forests Ruminants: Ecology Management and Conservation of Fragmented Communities, University of Chicago Press, Chicago. Pp 3 – 12.

Whitemore, T.C. and Sayer, J.A. (1992). Deforestation and species extinction in tropical moist forests. **In**: White more, T.C. and Sayer,J.A(eds),*Tropical Deforestation and Species Extinction*, Chapman and Hall, London, pp 1-14.

Williams J.G.K, Kubelik A.R, Livak K.J, Rafalski J.A, 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**:EARO(ed), *Proceeding of Coffee Berry, Disease Workshop*, Ethiopia agricultural Research Organization, Addis Ababa, Ethiopia. Pp 99 – 106.

Wolf, A.D. and Liston,A. (1998). Contributions of PCR-based methods to plant systematic and evolutionary biology. **In**: Soltis, D.E. Soltis, P.S. and Doyle, J.J. (Eds). *Plant Molecular Systematics* II. Kluwer Academic Publisher, Boston, MA, Pp 43-86.

Wrigley G. 1988. Coffee. Tropical Agriculture Series, London, John Wiley and Sons, Inc., Nw York.

Yeh, F.C., Yang,R-Cai.and Boyle,T.(1999). Population genetic analysis of codominant 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.

Yeshitila, K. (2004). Prospects of forest genetic resources conservation, **In**: *Proceeding of the public meeting on integrated forest policy Dervelopment in Ethiopia, Forum for Environment (FfE),* Institute of biodiversity, Addis Ababa.

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeats (SSR)-anchored PCR amplification. *Genomics*, **20**, 176–183.

DECLARATION

I declare that the thesis hereby submitted for the M.Sc. degree at the University of Addis Ababa is my own work and has not been previously submitted by me at another University for any degree. I cede copyright of the thesis in favor of the University of the Addis Ababa.

Name: Solomon Balami.

Signature: _____

Date of submission: _____

Approved by: Prof. Endashaw Bekele.

Signature: _____

.