DECLARATION

I declare that the thesis hereby submitted for the Philosophiae Doctor degree at the University of the Free State 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 favour of the University of the Free State.

Yigzaw Dessalegn Bekele
September 2005

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following institutions and persons:

- I am indebted to the Ethiopian Agricultural Research Organization (EARO) and Amhara Regional Agricultural Research Institute (ARARI) for the scholarship. I would like to express my special thanks to my promoter, Prof. Maryke Labuschagne. Her meticulous guidance, keen interest, encouragement and constructive criticism have made this study a reality. I would like to extend my special thanks to my co-promoter, Dr. Liezel Herselman for her valuable guidance, encouragement, unreserved support and insightful criticism. I am also indebted to my co-promoter Prof. Garry Osthoff for his guidance and constructive criticism adding to the success of my study. Mrs. Sadie Geldenhuys deserves special thanks for her kind and prompt response to all my administrative enquires. I would also like to thank Prof. Charl S. van Deventer and Dr. Geleta for their useful suggestions and encouragement.

- I also extend my gratitude towards Dr. P. Lashermes for supplying SSR primers sequence information, Prof. P. Mazzafera for information regarding extraction and determination of green bean biochemical compounds, Staff of the Addis Ababa Cup Quality Evaluation and Standardisation Centre for their help in the cup quality evaluation part of my work and Mr. P. Botes for his technical assistance during determining biochemical compounds using HPLC. I thank all post graduate students at the Plant Breeding section of University of the Free State for their support and encouragement.

- I am grateful for the support of my colleagues Fentahun, Dr. Getachew, Dr. Tesfaye (CTA), Dr. Yehenew, Tesfaye, Ermias, Anteneh, Yazie, Sisaye, Tazebew, Mesganaw, Yeshambele, Ahunim, Landuber, Mulugeta and Minale.

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

- Above all I thank and praise the Almighty God, who offered me everything.

DEDICATION

This piece of work is dedicated to my late mother Segede Seneshaw and my father Dessalegn Bekele who were always at my side to inspire and add courage.

QUOTATION

"The most wonderful mystery of life may well be the means by which it created so much diversity from so little physical matter, DNA". B.O. Wilson (1992)

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AA	Fraction of heavy beans retained by a no. 18/7.5 mm/ screen
AB	Fraction of heavy beans retained by a no. 15/5.95 mm/ screen
AD	Anno Domini
AEI	Assay efficiency index
AFLP	Amplified fragment length polymorphism
AP-PCR	Arbitrary primed polymerase chain reaction
ATP	Adenosine 5'-triphosphate
Av	Average
Bl	Bean length
bp	Base pair(s)
BSA	Bovine serum albumin
Bt	Bean thickness
Bw	Bean width
C	Fraction of beans retained by 2.90 mm screen
^{o}C	Degree Celsius
Cand	Canopy diameter
CBD	Coffee berry disease
CGA	Chlorogenic acid
cm	Centimetre(s)
cM	Centimorgan(s)
CTA	Coffee and Tea Authority
CTAB	Hexadecyltrimethylammonium bromide
dmb	Dry matter basis
DMRT	Duncan multiple range test
DNA	Deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
*Eco*RI	*Escherichia coli*
EDTA	Ethylenediaminetetraacetate
EMR	Effective multiplex ratio
Fa	Fair
FAO	Food and Agriculture Organization of the United Nations

Fg	Fairly good
Fmol	Femtomole
Frul	Fruit length
Fruw	Fruit width
F_1	First filial generation
g	Gram(s)
GA	Genetic advance
Gbw	Green bean weight
Gbyt	Green bean yield per tree
GCV	Genotypic coefficient of variation
GV	Genotypic variance
H	Broad sense heritability
h	Hour(s)
H’	Shannon-Weaver diversity index
ha	Hectare(s)
H_n	Expected heterozygosity
Ho	Observed heterozygosity
HPLC	High performance liquid chromatography
IBCR	Institute of Biodiversity Conservation and Research
ICO	International Coffee Organization
Ilob	Internode length of orthotropic branches
Ilpb	Internode length of primary branches
IPGRI	International Plant Genetic Resource Institute
IRAP	Inter-retrotransposon amplified polymorphism
ISSR	Inter simple sequence repeats
ITC	International Trade Centre
JSC	Jaccard’s similarity coefficient
kb	Kilobase(s)
Kg	Kilogram(s)
Km	Kilometre(s)
Leal	Leaf length
Leaw	Leaf width
LSD	Least significant difference

M	Molar
m	Metre(s)
masl	Metre above sea level
Me	Medium
Mf	Medium to Full
mg	Milligram(s)
mg/ml	Milligram per millilitre
MI	Marker index
min	Minute(s)
ml	Millilitre(s)
ml/min	Millilitre per minute
mm	Millimetre(s)
mM	Millimolar
Mp	Medium-pointed
MSe	Mean square of error
*Mse*I	*Micrococcus* species (*Meiothermus rubber*)
MSg	Mean square of genotype
µg	Microgram(s)
µM	Micromolar
µm	Micrometre(s)
µl	Microlitre(s)
NCSS	Number cruncher statistical system
ND	No data
ng	Nanogram(s)
ng/µl	Nanogram per microlitre
NGI	Number of genotypes identified
nm	Nanometre(s)
Npb	Number of primary branches
NPF	Number of polymorphic fragments
Nsb	Number of secondary branches
NTSYS	Numerical taxonomy and multivariate analysis system
NUF	Number of unique fragments
ODAP	Oxalyl Diamino Propanoic acid

ORSTOM	French institute of scientific research for development in cooperation
P	Probability
PAGE	Polyacrylamide gel electrophoresis
PB	Fraction of round beans retained by 4.43mm screen
Pbpb	Percentage of bearing primary branches
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
PCV	Phenotypic coefficient of variation
Petl	Petiol length
pg	Picogram(s)
PIC	Polymorphic information content
pmol	Picomole(s)
ppm	Parts per million
PV	Phenotypic variance
QTL	Quantitative trait loci
r	Correlation coefficient
RAPD	Random amplified polymorphic DNA
REMAP	Retrotransposon-microsatellite amplified polymorphism
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
rpm	Revolution per minute
s	Second(s)
SCAR	Sequence characterised amplified region
S-SAP	Sequence-specific amplified polymorphism
SSLP	Simple sequence length polymorphism
SSR	Simple sequence repeat
Stdv	Standard deviation
STR	Short tandem repeat
Taq	*Thermus aquaticus*
TBE	Tris-borate/EDTA
TE	Tris-Cl/EDTA
Tht	Tree height

TNF	Total number of fragments
Trud	Trunk diameter
TT	Fraction of light beans separated from AA and AB beans
U	Unit(s)
UPGMA	Unweighted pair group method using arithmetic averages
UV	Ultra-violet light
V	Volt
v	Volume
v/v	Volume/volume
W	Watt
w	Weight
w/v	Weight/volume

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

GENERAL INTRODUCTION

Coffee belongs to the family *Rubiaceae* and to the genus *Coffea* (Berthaud and Charrier, 1988; Coste, 1992). *Rubiaceae* has over 6000 species and 500 genera (ITC, 2002). Of these, the most economically important genus is *Coffea* (Wellman, 1961), comprising about 100 species (Pearl et al., 2004). The species of the genus *Coffea* are indigenous to Africa, Madagascar and the Mascarenes (Wrigley, 1988).

All species of *Coffea* are woody, ranging from small shrubs to large robust trees with heights up to 10 m. Phenotypic variation between species is enormous. Some are deciduous while others are evergreen. Leaves range in colour from yellow and dark-green to bronze and purple-green; their size vary from 1 to 40 cm in length, *Coffea liberica* L. having the largest leaves. Species differ considerably in the type of fruit they bear, ranging from being good and sweet flavoured to being distinctly inedible. Fruit size ranges from that of a small pea to a good-sized plum. Flowers range from being small, unattractive and scentless to being large and densely clustered with abundant fragrance. Some species have white flowers, some pink or almost purplish and some creamy to yellowish (Wellman, 1961).

The genus *Coffea* is not only endowed with enormous morphological variation, but also with adaptation to a wide range of environments, provided there is no frost. It grows from sea level to 2600 m altitude above sea level, habitats ranging from under shade to without shade, from sandy to humic soils and from flooded habitats (example *Coffea congensis* Frohener) to arid [example *Coffea rhamnifolia* (Chiov.) Bridson] (FAO, 1968b; Bellachew, 2001). This great diversity is a genetic reservoir to coffee breeders for achieving different breeding objectives.

Although the genus *Coffea* is diverse and reported to comprise about 100 species (Pearl et al., 2004), only two species namely arabica (*Coffea arabica* L.) and robusta (*Coffea canephora* Pierre) are under commercial cultivation (Raina et al., 1998; Lashermes et al., 1999; Anthony et al., 2002; Herrera et al., 2002; Prakash et al.,

2002; Steiger et al., 2002; Pearl et al., 2004). Arabica, the highland coffee, accounts for 60-70% of the global production while robusta coffee is more adaptable to lowlands and contributes the remaining 30-40% (Anthony et al., 2002; Herrera et al., 2002; ITC, 2002; Prakash et al., 2002). Arabica coffee produces superior quality coffee but its yield is very low, often constrained by diseases and pests (Agwanda et al., 1997; Prakash et al., 2002).

Arabica coffee is the only known tetraploid (2n = 4x = 44) and self-fertile (over 95%) species in the genus (FAO, 1968a; Charrier and Berthaud, 1985; Raina et al., 1998; Lashermes et al., 1999; Anthony et al., 2001; Herrera et al., 2002; Prakash et al., 2002; Pearl et al., 2004; Silvarolla et al., 2004). It is the most widely cultivated (Lashermes et al., 1996b) and one of the oldest known species (Coste, 1992).

Arabica coffee is monocentric (Harlan, 1992) since its centre of origin and diversity is in Ethiopia (Sylvian, 1958; Meyer, 1965; Wondimu, 1998; Bellachew et al., 2000; Anthony et al., 2002; Steiger et al., 2002). Considerable phenotypic diversity was observed in cultivated and traditionally recognised landraces of arabica coffee in Ethiopia (FAO, 1968b; Teketay, 1999; Dessalegn, 2002). Many important characteristics were observed in Ethiopian coffee, such as resistance to orange leaf rust (*Hemileia vastatrix* Berk and Br.) (Eskes, 1983; Wondimu, 1998), nematodes (*Meloidogyne incognita*) (Anzueto et al., 2001), coffee berry disease (*Colletotrichum kahawae* Waller and Bridge) (Bellachew et al., 2000), as well as variation in green bean biochemical compounds (caffeine, chlorogenic acids, sucrose and trigonelline) composition (Silvarolla et al., 2000; Ky et al., 2001), tree size and shape, bean size, shape and colour and in cup quality (Wondimu, 1998). The phenotypic variation as well as cultivation under diverse environmental conditions demonstrates the presence of *C. arabica* genetic diversity in Ethiopia. Currently, this genetic resource is under threat mainly because of deforestation of its natural habitat for timber and food crop production and replacement of the landraces by a few high yielding and disease resistant improved varieties (Woldemariam et al., 2002).

Several individuals and international organisations have been involved in efforts to collect, conserve and utilise Ethiopian coffee germplasm since 1928 (FAO, 1968a; Wondimu, 1998; Anthony et al., 2001). In addition, a number of national coffee

germplasm collection activities were carried out several times in different coffee growing areas of the country, including in some parts of northwestern Ethiopia. Consequently, the Institute of Biodiversity Conservation and Research (IBCR) of Ethiopia preserved over 4500 accessions in a coffee field genebank in Keffa, southwestern Ethiopia (Dubale and Teketay, 2000; ITC, 2002). However, the genetic diversity of these collections is not extensively analysed and documented using biochemical and molecular markers (Bellachew, 2001; Woldemariam et al., 2002; Aga et al., 2003).

Knowledge of genetic diversity within and among genotypes of any crop is fundamental to estimate the potential of genetic gain in a breeding programme and for effective conservation of available genetic resources (Sakiyama, 2000). It may also be important for selecting promising parental lines in hybrid variety development (Lashermes et al., 1996b; Ajmone-Marsan et al., 1998; Barbosa et al., 2003). Currently, different methods such as morphological, biochemical and molecular markers are available for estimation of genetic diversity within and among genotypes. Montagnon and Bouharmont (1996) classified wild and cultivated coffee accessions from Ethiopia according to their geographic origin using 18 agro-morphological characters. Arabica coffee genotypes diversity was also assessed using green bean biochemical compound composition. Ky et al. (2001) observed significant variation among arabica coffee accessions for caffeine, chlorogenic acids, sucrose and trigonelline contents of green bean. Recently, Silvarolla et al. (2004) identified three naturally decaffeinated arabica coffee trees out of 300 accessions collected from Ethiopia. Similarly, a number of investigators reported the presence of cup quality variation among different arabica coffee genotypes (Walyaro, 1983; Van der Vossen, 1985; Selvakumar and Sreenivasan, 1989; Roche, 1995).

Several DNA marker technologies have been developed and are available to study genetic diversity. The following are properties of a good DNA marker: highly polymorphic, codominant, abundant in the genome, even distribution throughout the genome, easy and fast assay, high reproducibility and easy exchange of data between laboratories (Weising et al., 1995). No DNA marker technology currently fulfils all of these criteria. However, amplified fragment length polymorphism (AFLP) and microsatellite or simple sequence repeat (SSR) fulfil most of these requirements.

Anthony et al. (2002) reported a relatively high level of polymorphism among arabica coffee accessions using AFLP. Moreover, AFLP has been used to detect genetic introgression in the arabica coffee genome (Lashermes et al., 2000a; Prakash et al., 2002; 2004) and to construct a genetic map for *C. arabica* (Pearl et al., 2004). Rovelli et al. (2000) identified polymorphic microsatellite sequences in the genome of *C. arabica*. Subsequently Anthony et al. (2002) classified cultivated and sub-spontaneous accessions of arabica coffee into their genetic origin using SSR markers. Despite the importance of genetic diversity information for efficient conservation as well as genetic improvement in any crop, the genetic diversity of coffee genotypes currently grown in northwestern Ethiopia has not been assessed and documented by any of the above methods. Therefore, this study was conducted with the following objectives:

1. To assess the presence and level of diversity among *C. arabica* genotypes from northwestern and southwestern Ethiopia using morphological, biochemical and molecular (AFLP and SSR) markers.

2. To compare the relative advantage of AFLP and SSR markers in discriminating genotypes.

3. To evaluate the cup quality of different *C. arabica* genotypes in northwestern Ethiopia and identify genotypes with superior cup quality for the area.

4. To analyse green bean caffeine, chlorogenic acids, sucrose and trigonelline contents of different genotypes and identify genotypes with low caffeine content.

5. To assess the level of association among cup quality, green bean physical characters and green bean biochemical compounds composition.

6. To assess the level of association among different agro-morphological characters.

CHAPTER 2

LITERATURE REVIEW

2.1 Economic importance and uses of coffee

Coffee has been used by Ethiopians before its migration to Arabia Felix (Yemen). However, the history of coffee as economically important crop goes back only to the 15th century when the world supply came from Yemen (FAO, 1968b). Currently, it is the world's most popular non-alcoholic beverage and the second most important commodity in global trade rated after petroleum products (Birhane, 2002; Prakash et al., 2002). It is exported in various forms to more than 165 countries (Birhane, 2002) and generates on average US $ 9.7 billion annually (ITC, 2002).

Coffee is grown in over 70 countries in the tropical and subtropical regions of the world (Birhane, 2002; ITC, 2002; Prakash et al., 2002), accounts up to 75% of the export revenue of many of these nations and provides a livelihood for about 25 million coffee farming families around the world (ICO, 2004). According to the ICO (2004) statistical report, about 6.1 million ton green coffee bean was produced in 2003 and the top 10 coffee producing nations with percentage contribution were Brazil (27.8%), Colombia (11.5%), Vietnam (10.9%), Indonesia (5.9%), India (4.5%), Mexico (4.4%), Ethiopia (4.2%), Guatemala (3.4%), Uganda (3%) and Hondurans (2.8%). Brazil and Colombia are the leading coffee producer and exporter countries and contributed 39.3% to the world coffee production in 2003 (ICO, 2004).

In Ethiopia, coffee contributes over 5% of the gross domestic product, 12% of the agricultural output, 70% of the foreign exchange earning and 10% of the government revenues. Besides, 25% of the population is employed in coffee production, processing and marketing (Tsegaye et al., 2000). About 400 000 ha land is covered by coffee and the annual national coffee production is estimated to be 200 000-250 000 ton (Yemane-Berhan, 1998; CTA, 2003). Of this, 80% is dry processed (natural) coffee (Chifra et al., 1998) and about 35% is consumed locally (Yemane-Berhan, 1998). Ethiopia ranks first in coffee consumption in Africa and the annual national

average per capita coffee consumption is 3 kg (CTA, 2003). Small-scale coffee farms contribute about 90% while large scale modern plantations account for the remaining 10% coffee production of the country (Yemane-Berhan, 1998).

Although coffee is popular as a non-alcoholic beverage, it combines valuable qualities such as medicine, food and beverage. Traditionally it served human beings since the prehistoric times to medicate different diseases. Shetty et al. (1994) demonstrated the medicinal value of coffee by testing coffee extracts for the control of *Staphylococcus aureus*, *Vibrio cholerae* and *Salmonella typhi*. All 25 tested strains of *Salmonella typhi* were sensitive to coffee extracts. Similarly, Berhanu (1998) reported the importance of coffee as a remedy against intermittent fevers and as antidote against narcotic poisoning. The most common medicinal use of caffeine is as part of headache preparations and other pain relievers. The ability of caffeine to stimulate breathing is also used in the treatment of apnea (cessation of breathing) in newborn babies and for people with lung disease who suffer from breathlessness, like asthmatic patients. There are also some controversial therapeutic uses of caffeine of coffee such as to kill skin funguses, aid fertility by improving sperm mobility, enhance the toxic effects of chemicals used in cancer therapy and facilitate the production of seizures during electroconvulsive therapy (Anonymous, 2004a).

Coffee is also used for the preparation of different food items in different parts of Africa (Wellman, 1961). Ethiopians, especially from the western part, prepare a snack known as *bunakella* from a mixture of roasted coffee bean and butter. In addition, caffeine extracted from coffee bean is currently used as additive in various cola formulations (Willson, 1999).

Coffee is mainly cultivated for its non-alcoholic stimulant beverage. The beverage can be prepared either in decoction or infusion. Decoction of the beans to extract a greenish liquid is the most ancient recipe. Currently, the most common preparation method is the infusion of roasted and ground coffee beans (Berhanu, 1998). Although the beverage that cheers and stimulates the taste buds of millions of people throughout the world is usually prepared from roasted coffee beans, other parts of the coffee tree are also used for making different types of non-alcoholic beverages. In Southeast Asia and Yemen the skin of ripe fruits is used to prepare *kishr*, a tea-like drink. In addition,

in Bolivia, Ethiopia and India sultan coffee (*kutti*) is prepared from the fruit hulls (*jenfel*) of cured coffee. Furthermore, Africans since prehistoric times used to prepare coffee leaf infusions from tender leaves and strips of green bark of twigs (Wellman, 1961). Local inhabitants of the eastern part of Ethiopia still prepare and use leaf infusions, locally known as *hojja*.

Coffee was identified as an abundant source of antioxidant phenols. The main contributors to phenolic compounds in coffee beans and pulp are chlorogenic acids (42.2%), epicatechin (21.6%), isochlorogenic acid I (5.7%) and isochlorogenic acid II (19.3%). These chemicals are considered as food antioxidants and may protect animal cells against somatic mutations associated with cancer. Currently, only caffeine is extracted and marketed around the world as a low price additive for cola drinks (Viniegra-Gonzalez, 2000) as well as in pharmaceutical companies (ITC, 2002). Therefore, assessment of green bean biochemical composition of different coffee genotypes seems essential in order to develop a variety with desirable green bean biochemical composition and to diversify the uses of coffee.

2.2 Taxonomic classification and genetics of coffee

Coffee belongs to the family *Rubiaceae* and the genus *Coffea* (Berthaud and Charrier, 1988; Coste, 1992). The family *Rubiaceae* consists of some 500 genera (ITC, 2002), of which *Coffea* is economically the most important (Wellman, 1961). Based on literature, herbarium specimens and wide field experiences, Chevailer (1947) classified the genus *Coffea* into four sections: *Eucoffea* K Schum, *Argocoffea* Pierre, *Mascarocoffea* Chev. and *Paracoffea* Miq. 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. Cultivated species of the genus *Coffea* belongs to the section *Eucoffea. Eucoffea* is divided into five subsections, *Erythrocoffea, Nanocoffea, Pachycoffea, Melanocoffea and Mozambicoffea* (FAO, 1968b). Although the genus *Coffea* comprises about 100 species, only two species namely, *Coffea arabica* L. and *Coffea canephora* Pierre which belong to the subsection *Erythrocoffea* (FAO, 1968b; Wrigley, 1988) are economically important (Pearl et al., 2004).

The basic chromosome number for the genus *Coffea* is n = 11. Arabica coffee is the only polyploid and self-fertile (over 95%) species of the genus *Coffea,* with chromosome number 2n = 4x = 44, while others are diploid (2n = 2x = 22) and self-infertile (FAO, 1968a; b; Lashermes et al., 2000a; Woldemariam et al., 2002; Silvarolla et al., 2004). Very little is known about the polyploidy origin of *C. arabica* (Pinto-Maglio and Da Cruz, 1998). Based on the analysis of a hybrid between *C. liberica* and *C. eugenioides*, Narasimhaswamy (1962) suggested these two species as the probable ancestors of *C. arabica*. Molecular characterisation of *C. arabica* also indicated its possible origin as an allotetraploid (Raina et al., 1998; Lashermes et al., 1999). According to Raina et al. (1998), the diploid wild ancestors of *C. arabica* are *C. eugenioides* and *C. congensis*. On the other hand, Lashermes et al. (1999) proposed *C. eugenioides* and *C. canephora* or other ecotypes. However, none of these suggested ancestors are found in Ethiopia, the centre of origin of the crop (FAO, 1968a; Woldemariam et al., 2002). Based on chromosome morphology analysis, Pinto-Maglio and Da Cruz (1998) suggested *C. arabica* as a segmental allopolyploid with genetically controlled chromosome pairing. Likewise, Lashermes et al. (2000a) and Rovelli et al. (2000) observed diploid-like meiotic behaviour in arabica coffee.

Cros et al. (1993) used laser flow cytometry to estimate total nuclear DNA content for 75 *Coffea* accessions, corresponding to 16 diploid species and *C. arabica*. Results indicated that three species native to east Africa (*C. sessiliflora, C. racemosa* and *C. pseudozanguebariae*) had the smallest genome size (about 1 pg per nucleus). Species native to the African evergreen forest (*C. humilis* and *C. liberica)* had the highest diploid DNA content (1.6 pg). The genome size of the tetraploid species, *C. arabica,* was 2.5 pg. The somatic chromosomes of *Coffea* are small (1 to 3.5 μm), making banding a difficult task (Pinto-Magilo and Da Cruz, 1998).

2.3 Origin, distribution and characteristics of arabica coffee

Arabica coffee is the only species found in Ethiopia (FAO, 1968a; b; Woldemariam et al., 2002). Its centre of origin is geographically isolated from the centre of origin of other species of the genus *Coffea*. It is confined to the plateau of southwestern Ethiopia and on the Boma plateau of Sudan (Wellman, 1961; FAO, 1968a; b; Lashermes et al., 1999; Anthony et al., 2002; Steiger et al. 2002). On the other hand,

the centre of origin of other coffee species overlaps elsewhere in the central and western parts of Africa (FAO, 1968b). Therefore, *C. arabica* follows the typical distribution features of polyploids, that is, peripheral expansion outside the range of distribution of the other diploid species of the genus (FAO, 1968a).

Arabica coffee is the earliest known and most widely distributed coffee species (Coste, 1992; Moncada and McCouch, 2004). It is grown throughout the tropical belt and in some areas even beyond the two tropics as in Brazil and Mozambique in the southern and in China (Taiwan) in the northern hemisphere (FAO, 1968b). Arabica coffee is found at elevations ranging from slightly above sea level to altitudes of 2350 m in Yemen and 2400 m in Ecuador (FAO, 1968b). In Ethiopia, it grows almost everywhere (FAO, 1968b; Yemane-Berhan, 1998), under diverse environmental conditions ranging in altitude from 550-2600 m above sea level and annual rainfall of 1000-2000 mm (Bellachew, 2001).

Arabica coffee is an evergreen shrub of variable size. The tree grows up to 8-10 m high and its branches are opposite, long, flexible and thin. Branches are semi-erect when young and spreading or pendulous when old (Coste, 1992). Its leaves are opposite, oval-shaped and acuminate. The leaves are between 10-15 cm long and 4-6 cm wide. It has white, Jasmine-scented flowers grouped together in the axils of the paired leaves, with two to three cymes making up whorls of 8-15 flowers. Its fruit is sub-globular, ovoid, oblong or squat-shaped. Fruits are orange-red to red on ripening, 16-18 mm long and 10-15 mm wide. Fruits have a coloured exocarp (skin), a fleshy yellow-white mesocarp (pulp) and two beans joined together along their flat sides. The calyx may or may not be pronounced as well as persistent until fruit maturity, depending on different varieties. The size and shape of the beans differ depending upon the variety, environmental conditions and management practices. On average, beans are 10 mm long, 6-7 mm wide, 3-4 mm thick and weigh between 0.15 and 0.20 g. Bean colour can be yellowish-grey to slate-grey, bluish or grey-green, depending upon the variety, method of preparation and storage condition (Coste, 1992). Bean shape may be sub-globular, ovoid, oblong, linear-oblong, either rounded at both ends or pointed at one end and rounded at the other (FAO, 1968a).

2.4 History of arabica coffee distribution and its impact on genetic diversity

Although arabica coffee has been used as a stimulant and a special kind of food in Ethiopia since time immemorial, it became known to the rest of the world after it reached Yemen (Woldemariam et al., 2002). In about 1510 the practice of coffee drinking spread to Mecca, Medina and Syria, and via Syria and Aden to Cairo and throughout the Moslem Middle East (Haarer, 1962). This habit gradually spread to the rest of the world and triggered an increased interest in some countries to smuggle viable coffee beans and produce coffee as a commodity on large scale (Haarer, 1962).

The history of introduction of arabica coffee from Ethiopia to Yemen is still vague. Steiger et al. (2002) reported that the first migration of arabica coffee from Ethiopia to Yemen was during the period of prehistoric trade. Wellman (1961) suggested that arabica coffee plants could have been introduced to Yemen as early as 575 AD and again by about 890 AD by Persian armies from Harar, Ethiopia. On the other hand, Eskes (1989) reported that *C. arabica* was introduced to Yemen three to four centuries ago. Similarly, Haarer (1962) reported that Arabs introduced coffee from Ethiopia to Yemen during the fifteenth century.

Historical data indicated that the introduction of arabica coffee to other continents first occurred from Yemen to the Malabar coast of India by Baba Budan in 1600 (Smith, 1985) and from there to Ceylon and Java in the last decade of the 17th century (Steiger et al., 2002). From Java a single coffee plant was taken to the botanical garden of Amsterdam in 1706 (Wellman, 1961; Haarer, 1962). It was from this tree that *C. arabica* var. *Typica* was propagated and spread to Asia, South America, parts of Africa and some botanical gardens in Europe. The other *C. arabica* sources were those collected by French men from Mocha (Yemen) in 1715 and 1718 and planted on Reunion (Bourbon) Island and from this source the *Bourbon* coffee of the world was propagated. Therefore, the spread of arabica coffee around the world was based on a very limited number of trees: the seven berries taken by Baba Budan to India, the small shipment to Reunion and the tree taken from Java to Amsterdam in 1706 together with its progeny in Paris. This together with its autogamous nature left the world arabica coffee production industry with a narrow genetic base (Steiger et al., 2002; Woldemariam et al., 2002; Chaparro et al., 2004).

2.5 Production systems and diversity of the Ethiopian coffee gene pool

There are four main coffee production systems in Ethiopia, namely forest, semi-forest, garden and plantation. The forest and semi-forest coffee production systems account 33% of the land covered by coffee and 25% of the annual coffee production in the country. On the other hand, the remaining two systems represent 67 and 75% of area and production of coffee of the country, respectively (Woldemariam et al., 2002). The contribution of forest coffee production system is dwindling as a result of deforestation (Yemane-Berhan, 1998). However, the forest coffee production system is still serving as a reservoir for arabica coffee genetic resource (Fig. 2.1).

Fig. 2.1 Partial view of the forest coffee production system

Considerable phenotypic diversity was observed in cultivated and traditionally recognised landraces of arabica coffee in Ethiopia (Sylvain, 1955; FAO, 1968a; Ameha, 1986; Montagnon and Bouharmont, 1996; Teketay, 1999; Dessalegn, 2002). Many important characteristics were identified in Ethiopian arabica coffee, such as resistance to orange leaf rust (*Hemileia vastatrix* Berk and Br.) (Eskes, 1983; Wondimu, 1998), nematodes (*Meloidogyne incognita*) (Anzueto et al., 2001) and coffee berry disease (*Colletotrichum kahawae* Waller and Bridge) (Bellachew et al., 2000). Substantial variation was also observed in green bean caffeine, chlorogenic acids, sucrose and trigonelline contents (Silvarolla et al., 2000; 2004; Ky et al., 2001),

tree size and shape, bean size, shape and colour and cup quality (Wondimu, 1998). Identification of several coffee berry disease (CBD) resistant and high yielding cultivars in a short period of time and cultivation of the crop under diverse environmental conditions also demonstrate the existence of diverse *C. arabica* genetic resources in Ethiopia (Ameha, 1980; Ameha and Bellachew, 1983; Bellachew et al., 2000; Bellachew, 2001). Currently, this genetic resource is under threat mainly because of deforestation of its natural habitat for timber and food crop production and replacement of landraces by a few high yielding and disease resistant improved varieties.

Since 1928 several individuals and international organisations such as the Food and Agriculture Organisation of the United Nations (FAO) and French Institute of Scientific Research for Development and Cooperation (ORSTOM) have been involved in efforts to collect, conserve and utilise coffee germplasm (FAO, 1968a; Wondimu, 1998; Anthony et al., 2001). In addition, a number of national coffee germplasm collection activities were carried out in different coffee growing areas of the country. Consequently, the Institute of Biodiversity Conservation and Research (IBCR) of Ethiopia preserved over 4500 accessions in the field coffee genebank on 115 ha land in Keffa, southwestern Ethiopia (Dubale and Teketay, 2000; ITC, 2002). However, the diversity of these genetic resources is not extensively described and documented using biochemical and molecular markers (Bellachew, 2001; Woldemariam et al., 2002; Aga et al., 2003).

2.6 Role of coffee seed for germplasm conservation

Germplasm conservation in the form of seeds is most convenient since seeds occupy a relatively small space and their transportation to various introduction centres and genebanks is economical. From their storage behaviour, seeds are characterised as “orthodox”, “recalcitrant” or “intermediate” (Razdan and Cocking, 1997). Most agricultural species are orthodox since their seeds are desiccation tolerant. Mature orthodox seeds have a moisture content of 20% or less on a weight basis and can further be dried (5% moisture content) without loss of viability. These seeds remain viable for many years and their longevity can further be increased by storing either at low temperature or via cryopreservation (Razdan and Cocking, 1997). Recalcitrant

seeds, on the contrary, are sensitive to desiccation and cannot be dried below a critical moisture level. These seeds shed from the plant in a naturally hydrated state and remain hydrated until germination or deterioration occurs. About 118 species belonging to 46 families of angiospermae are known to produce recalcitrant seeds (Pence, 1995). Sensitivity to desiccation makes recalcitrant seeds short-lived and therefore unsuitable to long-term storage by conventional drying (Razdan and Cocking, 1997). There are some species which are neither fully orthodox nor recalcitrant. Species producing seeds which is desiccation tolerant but sensitive to low temperature are designated as intermediate. Recalcitrant and intermediate seeds are short-lived and therefore unsuitable to long-term storage by conventional drying (Razdan and Cocking, 1997).

Though coffee seeds can withstand desiccation down to 6-8% water content (fresh weight basis), they can not be considered orthodox because they are cold sensitive and desiccation does not increase longevity (Dussert et al., 1997). Coffee seeds do not have a dormancy period, and are capable of germinating as soon as they are harvested (Van der Vossen, 1985; Coste, 1992). Its longevity is short (six months) when stored at ambient temperatures. A high germination capacity could be maintained for about two years if wet (40-41% moisture content) parchement seed was stored at 15-19^{o}C. Chilling to temperatures below 10 ^{o}C and reduction of seed moisture contents to 13-15% resulted in a rapid loss of viability (Van der Vossen, 1985). So far no efficient procedure is available for the long term storage of coffee seeds (Florin et al., 1995). Therefore, the longevity of arabica coffee seed is insufficient to be of use in the preservation of genetic resources. Consequently, coffee genetic resources are conventionally conserved as trees in field genebanks. Coffee trees kept in living collections allow a medium-range conservation; on average for about 50 years (Berthaud and Charrier, 1988). However, conservation of coffee in field genebanks is very expensive since it demands large areas as well as year round management. Therefore, genetic diversity analysis among accessions is vital for cost effective conservation of germplasm by removing duplicated samples. It is also essential for efficient utilisation of the available germplasm in breeding programmes.

2.7 Genetic diversity assessment methods

2.7.1 Agro-morphological characters

Currently, different methods such as morphological, biochemical and molecular markers are available to estimate the genetic diversity within and among genotypes. Of these, the morphological method is the oldest and is considered as the first step in the description and classification of germplasm (Smith and Smith, 1989). Although, morphological markers are cheap and easy to apply, they are sensitive to environmental influences and developmental stage of the plant. Moreover, morphological markers may be poorly suited for progeny analysis due to dominance effects (Werlemark et al., 1999). Newbury and Ford-Lloyd (1993) and Hardon et al. (1994) reported that although agro-morphological characters are often influenced by environmental conditions, the method is still useful and easy to apply for classification, estimating diversity and registration of cultivars. Camussi et al. (1985) also indicated that morphological data showing continuous distributions or that are polygenically controlled, may be particularly useful in inter-group classification below species level.

Natural selection affects morphological traits linked to adaptive characteristics. As a result, genetic distances from quantitative morphological character data allow inferences about adaptation and co-adaptation patterns of populations (Camussi et al., 1985). In addition, Ortiz (2001) reported that the phenotypic pattern of variation between genotypes is required for organising collections within gene banks and provides important information to plant breeders. Gomez et al. (2004) reported the complementarity of molecular and phenotypic markers and emphasised 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 resources in the context of breeding programmes.

Various agro-morphological characters have been successfully used to classify and estimate diversity in a variety of crop species (Dias et al., 1993; Zewdie and Zeven, 1997; Assefa et al., 1999; Mars and Marrakchi, 1999; Elhoumaizi et al., 2002; Lucchin et al., 2003; Tadesse and Bekele, 2003; Upadhyaya, 2003; Van de Wouw et al., 2003; Zoghlami and Zouaghi, 2003; Gomez et al., 2004; Quero-Garcia et al.,

2004). Dias et al. (1993) classified 58 Portuguese cole landraces into eight main groups using 46 morphological characters. Likewise, Zewdie and Zeven (1997) classified 67 hot pepper accessions into six clusters using 35 morphological and physiological characters. Assefa et al. (1999) classified 320 tef lines into 14 groups using 20 morphological, agronomical and phenological characters.

Upadhyaya (2003), using 12 morphological descriptors and 15 agronomic traits, analysed the phenotypic diversity of 1704 groundnut accessions. Similarly, Van de Wouw et al. (2003) characterised 454 accessions of common vetch and its relatives using 22 agro-morphological and phenological characters. On the other hand, Johns et al. (1997) failed to classify common bean landraces into their proper gene pool using 13 categorical morphological traits.

Zewdie and Zeven (1997) reported fruit weight, 1000 seed weight and fruit number per plant as the most important characters among the 35 morphological and physiological characters in clustering 67 hot pepper accessions. Upadhyaya et al. (2002) reported days to 50% flowering, plant width, apical secondary branches, tertiary branches, dots on seed testa, 100 seed weight, flowering duration, basal secondary branches, seed colour and seed testa texture as important traits in explaining multivariate polymorphism in a chickpea core collection. Tadesse and Bekele (2003) evaluated the diversity of 50 grass pea genotypes using morphological traits and observed significant differences among genotypes for days to maturity, pods per plant, seed yield per plant, harvest index and dry matter yield. These results revealed the importance of identification of important agro-morphological traits for diversity analysis in different crop species.

IPGRI (1996) listed a number of morphological, agronomical and biochemical characters for characterisation of coffee. However, some of these descriptors require at least five years to be expressed. Sylvain (1955) classified Ethiopian cultivated coffee into 13 main types namely, *Agaro, Arbagugu, Cioiccie, Coulo, Dilla, Ennarea, Harar, Irgalem, Keffa, Tafari Kela, Wolkite, Wollamo* and *Zeghie* using only bean and fruit morphology. Montagnon and Bouharmont (1996) classified wild and cultivated coffee genotypes from Ethiopia according to their geographic origin using 18 agro-morphological traits.

The FAO coffee collection team has observed phenotypic variation in branching habit, young leaf colour, fruit colour, persistence of sepals, leaf and fruit size (FAO, 1968a). Ameha (1986) observed growth habit variation such as compact and spreading type genotypes from national coffee collections. Likewise, Selvakumar and Sreenivasan (1989) observed phenotypic variation among 54 coffee accessions collected from Keffa province of Ethiopia. The level of agro-morphological variation among arabica coffee genotypes currently grown in northwestern Ethiopia is yet to be evaluated.

In perennial crops like coffee, it is important to identify the most suitable age of a tree when characters can be measured easily and with utmost accuracy. Walyaro (1983) reported that characters such as tree height, girth of the stem, internode length measured on the main stem and primaries and radius of canopy can be accurately determined using a single measurement even on young trees, 18 months after field planting. The heritability of an agro-morphological character indicates the importance of a particular character for characterisation of genotypes. Walyaro (1983) reported high narrow sense heritability for the following characters: girth (0.50-0.61), tree height (0.77-0.78), number of primaries (0.58-0.62), angle of primary with the main stem (0.56-0.62), canopy radius (0.44-0.50), internode length on primaries (0.51-0.63) and bearing primaries (0.51-0.53). Van der Vossen (1985) summarised the heritability of growth, yield and quality characters estimated from diallel crosses of 11 *C. arabica* cultivars. Results indicated that most growth and bean size characters (girth of main stem, tree height, canopy radius, internode length, angle of primary with main stem, 100 bean weight, % PB, % AA, % AB, % TT and % C) had high heritability, particularly when estimated from means of four trees. The heritability of bean yield (kg/tree) for the first year was low (0.29) but high (0.72) when using total yield/tree of the first two consecutive years.

Genetic studies in arabica coffee have shown that selection efficiency for higher bean yield can be increased by taking into account various growth parameters and yield components, such as stem girth, canopy radius, percentage of bearing primaries, percentage of bearing nodes and number of berries per node (Van der Vossen, 1985).

2.7.2 Cup quality and green bean physical characters

2.7.2.1 Variation in cup quality and green bean physical characters

Cup quality, often referred to as drinking quality or liquor quality, is an important attribute of coffee (Muschler, 2001; Agwanda et al., 2003) and acts as yardstick for price determination (Walyaro, 1983; Roche, 1995; Agwanda et al., 2003). Its assessment is done organoleptically by panels of experienced coffee tasters (Van der Vossen, 1985; Agwanda, 1999) and is determined on the basis of the level of acidity, body and flavour of the brew (Raju et al., 1978; Walyaro, 1983; Moreno et al., 1995). Walyaro (1983) recommended this method as a sufficiently reliable for use as a basis of selection in quality improvement programmes. Similarly, Owuor (1988) observed close similarity among liquorers in ranking various cup quality characteristics of the cultivars, indicating that any one panel could be relied on for selection for cup quality.

Acidity indicates the bitter or acidic balance and the presence of a sweet caramelic after taste (Petracco, 2000). High acidity gives better quality and more intense aroma to the beverage (Clifford, 1985). The preferred pH range for coffee beverage is 4.9 to 5.2. Body is synonymous with mouthfeel or linked with density and viscosity of the brew (Pertracco, 2000). However, there is no simple relationship between instrumentally measured beverage viscosity and professionally judged body (Clifford, 1985). Flavour indicates fragrance of the liquor either by direct inhaling of the vapours arising from the cup or nasal perception of the volatile substance evolving in the mouth (Petracco, 2000).

Agwanda (1999) compared four traits (acidity, body, flavour and overall standard) for their suitability as selection criteria for the genetic improvement of overall liquor quality. Based on correlation, repeatability and sensitivity analysis, flavour rating was recommended as the best selection criterion for genetic improvement of cup quality in arabica coffee. The trait showed high genetic correlation with preference, was easy to determine organoleptically and had relatively high sensitivity in discriminating different coffee genotypes (Agwanda, 1999). On the other hand, Walyaro (1983) and Van der Vossen (1985) observed fairly high heritability for the overall standard of cup quality and indicated the possibility of good selection progress for this character with

the assistance of experienced coffee tasters. Carvalho (1988) reported the dominant nature of good cup quality in arabica coffee.

Walyaro (1983) reported the presence of large inherent differences among genotypes for bean and cup quality attributes. Similarly, Van der Vossen (1985) observed variation for cup quality characters among varieties and crosses of arabica coffee. SL28 had big sized beans (46% AA) and excellent cup quality, while Caturra and Rume Sudan had small sized beans and lower cup quality. On the other hand, Hybrido de Timor had fairly big sized beans but poor cup quality (Van der Vossen, 1985). Owuor (1988) compared Ruiru 11, SL28 and K7 in cup quality and bean characteristics and reported their similarity for both traits.

Selvakumar and Sreenivasan (1989) observed coffee quality variation ranging from good to excellent among 54 arabica coffee accessions collected from Keffa province of Ethiopia. The world's best quality coffees such as Harar, Limu and Yirgacheffe (ITC, 2002) are produced from the eastern and southwestern parts of Ethiopia. Likewise, farmers, consumers and agricultural development extension experts reported the presence of considerable cup quality variation among different coffee genotypes grown in northwestern Ethiopia (personal communication). However, their report was not verified by research results.

Apart from genotype, cup quality is influenced by the age of the tree. Wellman (1961) reported that samples from young trees are likely to be mild and thin, but fine in flavour. Samples from old trees produce strong taste and a harsh characteristic brew. Medium aged trees, 15 to 20 years old, bear beans with good flavour as well as acidity and body (Wellman, 1961).

Cup quality of coffee is also affected by processing methods. Wet processed arabica is aromatic with fine acidity and some astringency, while dry processed arabica is less aromatic and less acidic but with greater body (Clifford, 1985). Length and condition of bean storage also affect cup quality (Wellman, 1961; Woelore, 1995). Long time storage under high relative humidity and warm conditions increase bean moisture content and consequently reduce quality in terms of raw and roast appearance as well as liquor (Woelore, 1995). Cup quality is also affected during the roasting process

(Clifford, 1985). Uneven roast results in poor quality liquor (ITC, 2002). Dark roast enhance the body while light roast emphasizes acidity (ITC, 2002).

Agronomic treatments also affect coffee cup quality. Wellman (1961) reported that in South America, coffee grown with heavy application of nitrogen fertilizer had poorer, lighter and thinner quality than that from unfertilized fields. On the other hand, magnesium deficiency had an adverse effect on cup quality (Mitchell, 1988). Carvalho (1988) reported that shade trees did not improve cup quality. On the contratry, Muschler (2001) reported that shade improved the appearance of green and roasted coffee beans as well as the acidity and body of the brew, especially for those produced in sub-optimal (low altitude) coffee production zones, by promoting slower and balanced filling and uniform ripening of berries. Likewise, Yemane-Berhan (1998) reported that shade increased sugar concentration, which is an important factor for creating the aroma of coffee.

Agwanda et al. (2003) reported significant genotype x environment interaction effects on coffee bean and liquor quality. Walyaro (1983) reported relatively lower genotype x environment interaction effects on quality characters. Van der Vossen (1985) reported non-significant genotype x environment interaction effects on quality characters, such as bean size and cup quality. Roche (1995) evaluated the cup quality of 15 arabica coffee cultivars by coffee tasters throughout United States of America and reported general consistency in the overall ranking of cultivars between two seasons. Sivetz (1963) indicated that if other factors are kept equal, better quality coffee can be found at higher altitudes. Wellman (1961) reported that lowland coffee was somewhat bland, with considerable body, while coffee from high altitude areas was more acidic, with better aroma and flavour. Review of previous research results indicated inconsistent effects of genotype x environment interaction on cup quality.

2.7.2.2 Correlation between cup quality and green bean characters

Information on the relationships between green bean characters and cup quality is scarce (Moschetto et al., 1996; Agwanda et al., 2003). Carvalho (1988) observed inconsistent results from efforts made to determine correlations between physical and

chemical properties of green bean as well as between green bean physical characters and cup quality.

Roche (1995) assessed the association of cup quality and green bean physical characters using 15 *C. arabica* cultivars and reported that unlike the popular belief, bean size was not a good indicator of cup quality when comparing cultivars from a single production area. Similarly, Agwanda et al. (2003) reported that bean quality traits were not useful for enhancement of genetic gains on cup quality and vice-versa. Muschler (2001) reported that proper ripening and slower filling had remarkable impact on cup quality. Generally, review of previous limited research efforts carried out on the association between cup quality and green bean physical characters showed contradicting and inconclusive results (Carvalho, 1988; Agwanda et al., 2003). Therefore, further investigation seems essential.

2.7.3 Biochemical compounds

2.7.3.1 Variation in green bean biochemical compounds composition

Most biochemical compounds confer adaptive properties to plants. They participate in resistance to diseases and pests or give a characteristic odour or taste to edible plants (Perez de la Vega, 1994). Therefore, assessment of biochemical diversity is fundamental in order to develop a variety with desirable biochemical composition or quality (Perez de la Vega, 1994; Granati et al., 2003; Moufida and Marzouk, 2003). Several biochemical compounds for example phenolics, flavonoids, anthocyanins, cyanogens, glycosides, polyacetylenes and alkaloids have been used for characterisation of different plant species as well as genotypes within species. Moufida and Marzouk (2003) classified citrus varieties into two groups using juice aroma and total sugar content. Granati et al. (2003) observed significant differences among 117 accessions of *Lathyrus* for protein and oxalyl diamino propanoic acid (ODAP) contents and classified them into five groups using these traits. Magoma et al. (2000) separated 102 tea accessions into three major and five minor groups using total green leaf catechin concentrations and the ratio of dihydroxylated to trihydroxylated catechins.

In coffee, Anthony et al. (1993) used biochemical compounds to assess the diversity of the genus *Coffea*. Similarly, Martin et al. (1998) differentiated arabica and robusta coffee varieties using chlorogenic acids, caffeine, trigonelline, amino acids and polyphenols content of green beans. Ky et al. (2001) assessed the diversity of caffeine, trigonelline, chlorogenic acids and sucrose contents in wild arabica and robusta accessions and confirmed the presence of high levels of polymorphism for these biochemical compounds among accessions within a species as well as between species. Moreover, the level of polymorphism observed was high compared to polymorphism observed using molecular markers (Ky et al., 2001).

Green bean chlorogenic acids content varied greatly between species (Clifford and Jarvis, 1988; Clifford et al., 1989; Anthony et al., 1993; Ky et al., 1999; 2001). It ranged from 0.14% on dry matter basis (dmb) in *C. rhamnifolia* Bridson to 9.9% in *C. sessiliflora* Bridson (Ky et al., 1999). Robusta coffee generally had higher amounts of chlorogenic acids compared to arabica coffee (Clifford, 1985). Ky et al. (2001) recorded green bean chlorogenic acids content variation ranging from 3.4-4.8% in arabica and from 7.9-14.4% in robusta accessions on dmb. Variation observed among arabica coffee accessions was statistically significant (Ky et al., 2001).

Ky et al. (2001) assessed green bean sucrose and trigonelline content of different arabica coffee accessions and observed significant variation among accessions for both these traits. The sucrose content varied from 7.4-11.1% and the trigonelline content ranged from 0.9-1.8% on dmb.

Sylvain (1958) observed significant caffeine content variation among coffee accessions from Ethiopia. Silvarolla et al. (2000) analysed green bean caffeine content of 99 progenies of arabica coffee accessions collected from different parts of Ethiopia and observed the presence of intra and inter- progeny variability. Among 68 progenies collected from Keffa province they found caffeine content in the range of 0.5-2.8% (mean 1.2%) and among 22 progenies collected from Illubabor province values ranged from 0.4-2.9% (mean 1.1%). Some of these *C. arabica* progenies have caffeine content comparable to the genetically modified low caffeine coffee bush developed by Ogita et al. (2003). Ky et al. (2001) assessed the green bean caffeine content of different *C. arabica* accessions and recorded 1.0-1.6% on dmb of green

beans. Recently, Silvarolla et al. (2004) identified three arabica coffee trees from 300 accessions collected from Ethiopia which are almost free from caffeine.

A mutant *C. arabica* cultivar known as Laurina was identified with half the caffeine content of normal arabica coffee varieties (Baumann et al., 1998). In addition, beans of CBD resistant coffee genotypes in Ethiopia and Kenya had higher caffeine content than susceptible ones (Biratu, 1997). This variability can be exploited in breeding programmes aimed at producing beans with low or high caffeine content for different purposes.

Green bean caffeine content variability among different coffee species was reported by different researchers (Baumann et al., 1998; Montagnon et al., 1998; Ashihara and Crozier, 1999; Ky et al., 2001). Ashihara and Crozier (1999) reported that arabica coffee had higher amounts of caffeine compared to *C. salvatrix*, *C. eugenioides* and *C. bengalensis*. Arabica coffee had lower green bean caffeine content than *C. canephora* (Baumann et al., 1998; Montagnon et al., 1998; Ky et al., 2001). The pulp from robusta coffee had lower caffeine contents than the pulp from arabica cultivars (Clifford and Ramirez-Martinez, 1991).

The amount of caffeine in different parts of coffee trees has been determined by a number of chemists. Wellman (1961) reviewed different findings and reported that in *C. arabica*, beans largely contained 1.5% caffeine, flowers of a 20-year old tree 0.9%, air-dried coffee leaves 0.82%, coffee stems 0.09% and the cherry pulp 0.88% caffeine.

2.7.3.2 Correlation between cup quality and green bean biochemical composition

Several studies were conducted to assess the relationship between green bean biochemical compounds and cup quality of coffee. De Maria et al. (1994) reported sucrose, trigonelline and chlorogenic acids as important flavour precursors. Sucrose and trigonelline gave rise to appreciated flavour products, including furans, pyrazine, alkyl-pyridines and pyrroles (Clifford, 1985), whereas chlorogenic acids and caffeine increased bitterness (Leloup et al., 1995). Fat contributed to cup quality by catching aromatic compounds arising during roasting (Montagnon et al., 1998).

Chlorogenic acids (CGA) modify cup quality both directly and indirectly (Ky et al., 1999). The direct effect is due to the presence of CGA after roasting, which is known to increase astringency (Clifford, 1985). The indirect effect is due to molecular changes during roasting, which have positive or negative influences. Degradation of CGA into phenolic derivatives (Leloup et al., 1995) and inhibition of pyrazine formation by Millards reactions were negative effects. On the other hand, CGA reaction with trigonelline, sucrose and amino acids, which enriched headspace volatile profiles, was a positive effect (De Maria et al., 1994). Generally, cup quality of coffee increased when the CGA content decreased (Ky et al., 1999). This largely explains the taste differences between robusta and arabica coffees (Clifford, 1985). Bicchi et al. (1995) were unable to correlate with certainty the beverage quality, or a specific sensory attribute such as astringency, with the presence of specific chlorogenic acid(s).

In addition to its effect on cup quality, chlorogenic acids control germination and cell growth and participate in defence mechanisms against phytopathogens (Aerts and Baumann, 1994; Ky et al., 1999). This indicated the importance of CGA both for cup quality and disease resistance in coffee. Waldhauser and Baumann (1996) reported the correlated accumulation of caffeine and chlorogenic acids in coffee bean. Similarly, Baumann et al. (1998) found positive correlations between chlorogenic acids and caffeine contents of green bean in all coffee species.

Sucrose is an important component of coffee flavour. The higher the sucrose content in green beans, the more intense the cup flavour (Ky et al., 2000). Sucrose degrades into appreciable flavour product, furans (Clifford, 1985; De Maria et al., 1996). In addition, sucrose reacts with amino acids to produce pyrazines and carbonyl compounds, which are involved in flavour (Feldman et al., 1969). The sucrose/amino acid ratio in green beans widely determines the profile of volatile compounds and differences in this ratio might explain in part the organoleptic differences between arabica and robusta coffee (Tressl et al., 1982).

Trigonelline gives rise to appreciated flavour products and influences the flavour of the beverage (Clifford, 1985). In addition, trigonelline degradation during roasting results in niacin, nicotinamide and a range of aroma volatiles, which include pyridines

and pyrroles (Clifford, 1985; Ky et al., 2001). Green bean processing, dewaxing and decaffeination procedures had little effect upon the trigonelline content of coffee beans (Clifford, 1985), but roasting caused progressive destruction (Sivetz, 1963). Trigonelline showed potent mutagenic activity in roasted coffee (Wu et al., 1997) and played a significant role in excess salt stress tolerance in soybean (Tramontano and Jouve, 1997; Wood, 1999).

Coffee is consumed mainly for its stimulatory effect. The stimulating effect of coffee is associated with the biochemical compound known as caffeine. Caffeine, 1,3,7-trimethyl xanthine, is a white, bitter-tasting and crystalline substance. Caffeine was first isolated from coffee in 1820 (Anonymous, 2004b). Clifford (1985) reported the limited role of caffeine in beverage bitterness. Green bean dewaxing and wet processing after removal of the parchment reduces the caffeine content in coffee beans. Moreover, decaffeination virtually eliminates caffeine (Clifford, 1985).

Decaffeination was first practised as early as 1905 by the Hag Company in Bremen, Germany (Clarke, 1985). Decaffeination caters for those who, for whatever reason, do not want the stimulating effect of caffeine. In the European Union the absolute caffeine content of decaffeinated coffee may not exceed 0.1% in roasted and 0.3% in soluble coffee. In the United States, caffeine in decaffeinated coffee is to be reduced to less than 3% of the original content (ITC, 2002). Decaffeination is carried out using green beans and the following processing steps. First, green beans are treated with vapour and water to open up the bean surface and the cell structure to access the crystalline caffeine taken up in the cell walls. The second step is the extraction of the caffeine by extraction agents (water, carbon dioxide and organic extraction agents like methylene chloride and ethyl acetate) which possess the ability to extract only the caffeine. Wet coffee, from which the caffeine has been removed, is dried until it reaches its normal moisture content. Beans can then be roasted as usual (ITC, 2002). The process of decaffeination using carbon dioxide or organic solvents removes caffeine from beans, often along with other key flavour compounds, consequently reduce cup quality (Silvarolla et al., 2000; Ogita et al., 2003)

Caffeine present in *C. arabica* seeds and plants is assumed to protect these tissues from pathogens (Harborne, 1988). Biratu (1997) proved this assumption by studying

the effect of different concentrations of caffeine on coffee berry disease causative fungi, *C. kahawae.* Results indicated that 0.15% caffeine application inhibited mycelial growth for six isolates of the fungi. Seedlings excreted caffeine into the surrounding substrate inhibits seed germination in the vicinity of coffee seedlings (Baumann and Gabriel, 1984). Young *Coffea* foliage leaflets and cotyledons contained caffeine concentrations as high as 40-60 mM (Aerts and Baumann, 1994). This high caffeine concentration is toxic and lethal to insects and fungi (Prabhuji et al., 1983; Nathanson, 1984). This indicated the role of caffeine as chemical defence mechanism for the coffee plant.

A positive association between caffeine and chlorogenic acids content was reported for various coffee species, except for the Laurina cultivar with very low caffeine content and a normal level of chlorogenic acids (Carvalho, 1988). Caffeine content in coffee is under polygenic control (Baumann et al., 1998) and absence of caffeine is controlled by one recessive gene (Barre et al., 1998).

Review of previous results indicated the importance of green bean biochemical composition assessment to select or develop a variety with superior cup quality and/or with desirable biochemical composition for alternative uses. However, the biochemical diversity of *C. arabica* genotypes currently grown in Ethiopia is yet to be analysed.

2.7.4 Molecular markers

Traditionally like any other plant species, *C. arabica* genotypes were characterised based on phenotypic observations, but this approach is slow and subjected to environmental influences mainly due to the long juvenile period and perennial nature of the crop. Currently, molecular markers are employed in coffee breeding programmes of some countries to study genetic relationships among cultivars (Lashermes et al., 1996a; Anthony et al., 2001; Steiger et al., 2002), detect genetic introgression (Lashermes et al., 1999; 2000a; Herrera et al., 2002; Prakash et al., 2002), marker assisted selection, example for coffee berry disease, rust and root-knot nematode resistance (Agwanda et al., 1997; Noir et al., 2003; Prakash et al., 2004) and genetic map construction (Pearl et al., 2004).

The importance of molecular markers for genetic improvement in perennial cops like coffee is immense. It allows selection of desirable genotypes at an early growth stage (at seedling stage), on a large number of breeding lines, reduce the number of back-cross cycles required to restore the quality of the recurrent parent and for simultaneous improvement of different traits (Lashermes et al., 2000b). Various molecular marker techniques were developed and are available, but only those previously employed on coffee are reviewed below.

2.7.4.1 Isozymes

Isozymes were the first molecular markers used in plant breeding (Tanksley and Orton, 1983). The term isozyme was coined by Markert and Moller (1959) to describe multiple forms of enzymes that share a common substrate but differ in electric mobility. A number of studies conducted in the early 1950s provided evidence regarding the existence of multiple forms of enzymes (McMillin, 1983). Isozymes are revealed when tissue extracts are subjected to electrophoresis in various types of gels and subsequently submersed in solutions containing enzyme-specific stains. Electrophoretic separation of complex mixtures of proteins can be accomplished in several types of supporting media, including starch, polyacrylamide, and agarose gels and cellulose acetate membrane. The latter two lacks sufficient resolving power and are not generally employed for the study of enzyme polymorphisms. Despite the greater resolving power and flexibility of polyacrylamide gel electrophoresis (PAGE), starch gel electrophoresis is preferred for most studies involving large numbers of individuals and different enzymes. This could be due to the simplicity of starch gel preparation, the non-toxic nature of the material used, relative cost of equipment and the ease of loading samples onto the gel. Samples for starch gels electrophoresis are usually crude uncentrifuged homogenates, but PAGE demands clarified samples. The most compelling reason for the popularity of starch gels is the difference in amount of data generated per gel. Starch gels allow multiple assays (upto six enzyme systems per gel) whereas polyacrylamide gels are usually stained for a single enzyme (Wendel and Weeden, 1989).

Isozymes generally exhibit Mendelian inheritance, co-dominant expression, complete penetrance and are free of pleiotropic and epistatic interactions (Weeden, 1989). The

isozyme technique is fast, cheap and simple. However, isozyme markers are not as plentiful as DNA markers (Dudnikov, 2003) and sometimes interpretation of zymograms become difficult due to complex banding profiles arising from polyploidy or duplicate genes. In addition, proteins with identical electrophoretic mobility (co-migration) 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 not as plentiful as DNA markers and limited by tissue and developmental stage specificity, it has been used for genetic diversity analysis in many species (Dudnikov, 2003). The isozyme technique appears to be more informative at lower taxonomic levels, particularly for species and population level characterisation (Brown, 1990).

Using isozyme markers, Hagen and Hamrick (1998) evaluated the genetic diversity of nine red clover populations collected from southeastern and northeastern USA and reported higher levels of genetic diversity within the populations and lower levels of genetic divergence among populations. Montarroyos et al. (2003) studied the genetic variability of 28 accessions of cassava using four enzyme systems from leaf and root tissue and classified genotypes into eight clusters. Results indicated similarity between leaf tissue isozyme analysis and morphological characterisation indicating that highly inherited characters are good cassava descriptors. Dudnikov (2003) analysed the genetic control and linkage patterns of growth habit of *Aegilops tauschii* using 12 enzyme loci. Mosjidis et al. (2004) evaluated the genetic diversity of 15 wild red clover populations collected from Caucasus, Russia, using seven isozymes and reported the presence of high levels of genetic diversity within populations compared to among populations.

A number of studies were also conducted to assess the association between genetic distances estimated from isozyme loci and heterosis. Cerna et al. (1997) reported the presence of correlation between genetic distances estimated by isozyme loci and heterosis in seed yield of soybean. Qian et al. (1999) reported the association of some isozyme loci with quantitative trait loci (QTL) in maize. Similarly, Yu et al. (2005)

reported significant correlation between some isozyme loci such as acid phosphatase and quantitative traits of the hybrid in oilseed rape.

Isozymes have been applied to *C. arabica*. However, their use for arabica coffee characterisation have been limited due to the small number of isozyme systems available (Berthaud and Charrier, 1988) and the low level of polymorphism detected (Berthou and Trouslot, 1977; Louarn, 1978 as cited in Ruas et al., 2000). The attempt of Paillard et al. (1996) to construct a genetic map for coffee based on isozymes failed due to the low polymorphism level. Consequently, arabica coffee researchers like researchers in many other crop species shifted towards using DNA-based markers. DNA-based markers are abundant, have simple inheritance and consistent results, regardless of the cropping environmental condition of the plant, or of the type or age of the tissue sampled (Sakiyama, 2000). These characteristics are relevant for coffee research, since it is a perennial crop with a long juvenile period.

Recent advances in the field of plant molecular genetics have resulted in the development of a series of DNA markers. Of these, restriction fragment length polymorphism (RFLP) (Botstein et al., 1980), random amplified polymorphic DNA (RAPD) (Welsh and McClelland, 1990; Williams et al., 1990), amplified fragment length polymorphism (AFLP) (Vos et al., 1995) and simple sequence repeat (SSR) or microsatellite (Weber and May, 1989) are the most important and were employed in arabica coffee.

2.7.4.2 Restriction fragment length polymorphism (RFLP)

Restriction fragment length polymorphism was the first developed DNA-based marker technique (Fazuoli et al., 2000). Polymorphism observed in RFLP analyses is based on the length of fragments generated by digestion with restriction enzymes (Botstein et al., 1980). RFLP variability in plants can be caused by (1) base sequence changes which add or eliminate restriction sites, (2) rearrangements such as insertion or deletions, or (3) unequal crossing over or replication slippage. It appears that most RFLP variability in plants is caused by genome rearrangements rather than nucleotide sequence change. RFLPs are simply inherited Mendelian characters and proven to be abundant in most organisms (Kochert, 1994).

RFLP analysis consist of DNA isolation, digestion of DNA with a restriction enzyme, separation of DNA fragments by agarose gel electrophoresis, transfer of the separated restriction fragments to a filter by Southern blotting, detection of individual restriction fragments by nucleic acid hybridisation with a radioactively labelled cloned probe and scoring of RFLPs by direct observation of autoradiograms. RFLPs are co-dominant markers, thus distinguish both homozygotes and heterozygotes of a locus. Consequently, it is more informative than dominant markers. However, this technique requires relatively large amounts of purified and high molecular weight DNA, is time consuming and laborious. Therefore, it is inefficient for large scale analyses involving evaluation of segregating populations for genetic mapping or extensive screening of germplasm collections (Halward et al., 1992).

The RFLP technique has been used for genetic diversity analysis, mapping and marker assisted selection in different crop species (Powell et al., 1996; Mohan et al., 1997; Russsell et al., 1997; Pejic et al., 1998; Garcia-Mas et al., 2000). Ajmone-Marsan et al. (1998) employed RFLP markers to study the association between genetic distance and hybrid performance in maize. Correlation of genetic distance with hybrid performance for grain yield was positive but too small to be of predictive value.

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 hybridisation between *C. eugenioides* and *C. canephora*, or ecotypes related to these diploid species. They reported low levels of polymorphism among arabica coffee accessions using RFLP analysis. Paillard et al. (1996) failed to construct genetic maps for coffee based on RFLP markers due to the low level of polymorphism detected among genotypes. Results of previous studies were not encouraging for the use RFLP for genetic diversity analysis in *C. arabica*.

2.7.4.3 Random amplified polymorphic DNA (RAPD)

Random amplified polymorphic DNA, one of the various PCR-based marker techniques, was developed independently by Welsh and McClelland (1990) and

Williams et al. (1990). Williams et al. (1990) proposed the term random amplified polymorphic DNA for this class of DNA marker while Welsh and McClelland (1990) named it arbitrarily primed PCR (AP-PCR). The RAPD method uses arbitrary short oligomers (usually 10-mer) which anneal to random homologous target sites within the genome. Variation among individuals for RAPD profiles mostly arises from base-pair substitutions that modify or eliminate the primer site. Insertions in the genomic sequence that separate the primer sites to a distance which prohibit amplification or cause length changes of the amplified product results in RAPD profile variation among individuals (Chowdhury et al., 2002). RAPD requires small amounts of template DNA, does not require sequence information for primer construction, is easy and quick to assay, have low cost, have random distribution throughout the genome and generates multiple fragments per reaction. However, it is a dominant marker technique and has a problem of reproducibility of amplification (Rafalski, 1997). Relative to other methods, RAPD markers tend to underestimate genetic distances between more distantly related individuals, for example in inter-specific comparisons (Powell et al., 1996).

Several studies have indicated that RAPDs are useful in cultivar identification, population and pedigree analyses, phylogenetic studies and genetic mapping (Halward et al., 1992). Recently, RAPD analysis was successfully used for genetic relationship analysis among different genotypes of sorghum (Ayana et al., 2000), eucalyptus (Li, 2000), populus (Rajagopal et al., 2000), chickpea (Chowdhury et al., 2002), pea (Simioniuc et al., 2002), *Betula alnoides* (Zeng et al., 2003) and banana (Onguso et al., 2004). It was also successfully employed to study phylogenetic relationships among species of *Ipomoea* (Ardelean et al., 2004) and *Limonium* (Bruna et al., 2004). Xu et al. (2004) used RAPD markers for sex determination in *Eucommia ulmoides* Oliv.

RAPD markers were employed for marker assisted selection in arabica coffee. Coffee berry disease is the most important disease for arabica coffee production in Africa. Agwanda et al. (1997) suggested that resistance to CBD in arabica coffee is controlled by at least three genes which are present in the varieties Hibrido de Timor (*T* gene), Rume Sudan (*R* and *k* genes) and K7 (*k* gene) and identified three RAPD markers

linked to the *T* gene which could be used for CBD resistance selection in arabica coffee.

Lashermes et al. (1993) and Orozco-Castillo et al. (1994) proved the effectiveness of RAPD markers to identify coffee accessions. Lashermes et al. (1996b) successfully employed RAPD markers to analyse the genetic diversity among cultivated and subspontaneous accessions of *C. arabica* and confirmed the narrow genetic base of commercial cultivars. Anthony et al. (2001) studied the genetic relationships of 88 spontaneous and subspontaneous accessions as well as six cultivars of *C. arabica* collected from different parts of Ethiopia and two accessions derived from the genetic populations Typica and Bourbon, using RAPD markers, reported the differentiation of Ethiopian materials from Typica- and Bourbon-derived accessions. They did not observe within accession polymorphism and confirmed the low level of polymorphism in *C. arabica*. This might be due to the dominant nature of RAPD markers.

Aga et al. (2003) analysed the genetic diversity of 144 genotypes representing 16 *C. arabica* populations using 12 RAPD primers and reported genetic differentiation of 0.80 within and 0.20 between zones of sample collection sites. Chaparro et al. (2004) studied the genetic diversity of 50 wild and semi-wild accessions of *C. arabica* germplasms collected by the FAO and ORSTOM mission to Ethiopia using five polymorphic RAPD primers and observed lack of correlation between clustering of genotypes and collection region. These results indicated the presence of a high level of genetic diversity within collection sites and may be attributed to the autogamous and perennial nature of the crop as well as to the unrestricted movement of coffee germplasm to different regions of Ethiopia.

2.7.4.4 Amplified fragment length polymorphism (AFLP)

The amplified fragment length polymorphism technique, developed by Vos et al. (1995), is a powerful tool for DNA fingerprinting. In principle it is a combination of RFLP and polymerase chain reaction (PCR) techniques (Vos et al., 1995). Consequently, it combines the speed of PCR with the precision of RFLP (Powell et al., 1996). Unlike RFLP, it is PCR-based, requires small amounts of DNA and is

readily automatable. Unlike RAPD, AFLP is robust, reliable and reproducible (Jones et al., 1997) and unlike SSR it does not require sequence information of the target genome.

AFLP is carried out in four distinct steps: digestion of DNA using rare and frequent cutter restriction enzymes, ligation of double-stranded oligonucleotide adapters to the restricted sites, PCR amplification of restricted fragments with primers that bind to the adapter sequence, restriction site sequence and adjacent selective base(s) and acrylamide gel electrophoresis or capillary electrophoresis (when using ABI). The use of two restriction enzymes with different specificities enables the generation of a large number of DNA fragments suitable for PCR amplification. PCR amplification with specific primers reduces the number of DNA fragments to be amplified and ensures reliable and reproducible detection of restricted fragments. Acrylamide gel or capillary electrophoresis is sensitive and able to distinguish fragments differing in length by only one base pair (Miyashita et al., 1999). Thus AFLP is a powerful technique for detection of a large number of fragments with high reproducibility and sensitivity (Jones et al., 1997). It can be employed to DNA of any origin or complexity without any prior sequence information (Vos et al., 1995). The other advantage of the AFLP technique is observation of a large number of markers that are randomly distributed throughout the genome (Lin et al., 1996; Breyne et al., 1997).

AFLPs are usually considered dominant markers since polymorphism is detected as presence and absence of fragments (Powell et al., 1996). Since it is based on restriction digestion of DNA, it is sensitive to the quality and concentration of template DNA (Jones et al., 1997). Easyness of implementation, large numbers of polymorphisms detected per reaction, requirement of small amounts of DNA and high reproducibility of the DNA fingerprint patterns recommend AFLP as an attractive method for studying DNA polymorphisms (Innan et al., 1999). Currently, AFLP is being used to investigate genetic variation in a number of microorganisms, plants and animals (Magurie et al., 2002).

AFLP analysis has been applied to construct genetic linkage maps of barley (Waugh et al., 1997), melon (Wang et al., 1997), potato (Van der Voort et al., 1997) and *Arabidopsis* (Alonso-Blanco et al., 1998). Studies on crops such as daylily (Tomkins

et al., 2001), *Arabidopsis* (Breyne et al., 1999), rice (Aggarwal et al., 1999), papaya and its related species (Van Droogenbroeck et al., 2002), bread wheat (Barrett and Kidwell, 1998; Almanza-Pinzon et al., 2003), *Musa* (Ude et al., 2002), *Olea europaea* (Sensi et al., 2003), sorghum (Uptmoor et al., 2003), impatiens (Carr et al., 2003), barley (Turpeinen et al., 2003), groundnut (Herselman, 2003) and *Anemone coronaria* (Nissim et al., 2004) demonstrated the importance of AFLP technique to assess genetic diversity within and among species.

The AFLP technique has also been used to identify markers linked to disease resistance loci (Becker et al., 1995; Cervera et al., 1996; Herselman et al., 2004) and to assess the relationship between molecular marker polymorphism of inbred lines and hybrid performance in maize (Ajmone-Marsan et al., 1998). Barbosa et al. (2003) reported that AFLP is efficient in assigning maize inbred lines to heterotic groups and that AFLP-based genetic distance is suitable for predicting maize single cross performance for intrapopulation crosses of broad based populations. On the other hand, Geleta et al. (2004) observed lack of significant correlation between AFLP-based genetic distances and heterosis in pepper. Zeid et al. (2004) did not find significant correlation between AFLP-based genetic distances and hybrid performance within the elite European faba bean gene pool.

Among the available DNA molecular techniques, AFLP is a powerful technique for cultivar identification (Powell et al., 1996). Lombard et al. (2000) distinguished 83 rapeseed cultivars including isogenic lines using two AFLP primer combinations. Kim et al. (1998) demonstrated the potential of the method for cultivar identification by distinguishing 12 potato cultivars using any one of the AFLP primer combinations tested. Sensi et al. (2003) differentiated *Olea europaea* cultivars using six primer combinations and confirmed the importance of AFLP technology not only for studying variation between populations of the genus *Olea* but also for characterising intraspecific variation among cultivated accessions. Tyrka et al. (2002) identified 19 strawberry cultivars using 10 AFLP primer combinations.

The resolving power of AFLP in cultivar identification was demonstrated by Turpeinen et al. (2003) while assessing the genetic diversity of 94 barley genotypes using eight AFLP primer combinations. Of the eight primer combinations used, four

distinguished 94 genotypes independently and the genetic similarity coefficient varied between 0.74 and 0.98. Results of their study revealed differences among primer combinations in detecting genetic diversity among genotypes. Similar results were reported in different crop species by several investigators (Breyne et al., 1999; Yee et al., 1999; Garcia-Mas et al., 2000; Renganayaki et al., 2001; Simioniuc et al., 2002; Almanza-Pinzon et al., 2003; Sensi et al., 2003). Identification of primer combinations revealing high levels of polymorphism seems essential for cultivar identification using the AFLP technique.

In arabica coffee, AFLP has been used to study the origin of cultivated varieties (Anthony et al., 2002), to analyse genetic diversity within and among cultivars (Steiger et al., 2002), for detection of genetic introgression (Lashermes et al., 2000a; Prakash et al., 2002) and for construction of a genetic linkage map (Pearl et al., 2004). Lashermes et al. (2000a) observed 109 polymorphic fragments among 42 arabica coffee accessions using 42 AFLP primer combinations. Prakash et al. (2002) observed 35 polymorphic fragments among six accessions using 36 AFLP primer combinations. Steiger et al. (2002) reported low levels of genetic variation (an average genetic similarity of 0.933) among six arabica cultivars using six AFLP primer combinations. Anthony et al. (2002) studied the genetic diversity of 26 cultivated and subspontaneous derived accessions of *C. arabica* using 37 AFLP primer combinations and detected a total of 107 polymorphic markers. The genetic distances between cultivated groups (Typica and Bourbon) were low, on average 0.17. Similar results were reported by Lashermes et al. (1996b) and Anthony et al. (2001) using RAPD markers. Higher genetic distance values were estimated among subspontaneous derived accessions as well as between cultivated and subspontaneous derived accessions (Anthony et al., 2002). The genetic distances ranged from 0.40 to 0.71 between the Typica group and the subspontaneous derived accessions and from 0.37 to 0.70 between the Bourbon and subspontaneous derived accessions. Classification based on AFLP markers confirmed the partition of the subspontaneous coffee accessions into two groups, separated by the Great Rift Valley (Anthony et al., 2001).

The observed low level of genetic diversity especially among cultivated cultivars of arabica coffee is attributed to its origin, reproductive biology and evolution

(Lashermes et al., 2000b). In addition, most cultivars are derived from the few trees which survived various efforts to spread arabica growing world-wide (Van der Vossen, 1985; Lashermes et al., 1996b). Although AFLP is efficient to detect genetic variation among closely related genotypes, the genetic diversity of arabica coffee genotypes currently grown in Ethiopia was not estimated using AFLP markers.

2.7.4.5 Simple sequence repeats (SSR)

Simple sequence repeat is also known as microsatellite or short tandem repeat (STR) or simple sequence length polymorphism (SSLP) (Tanya et al., 2001). SSR markers are based on tandem repeats of short (2-6 base pair) DNA fragments scattered throughout the genome that lie between conserved sequences (Litt and Luty, 1989; Weber and May, 1989) and have been isolated and characterised in a large number of animal and plant species (Tautz, 1989; Wang et al., 1994; Powell et al., 1996; Lefort et al., 1999; Hormaza, 2002). It has been suggested that the three mechanisms for creating a new allele at SSR loci are replication slippage (Tachida and Iizuka, 1992), unequal crossing-over (Harding et al., 1992) and genetic recombination (Huang et al., 2002). Replication slippage is considered to be a major factor affecting the repeat number for short tandem repeat sequences, whereas unequal crossing-over is thought to result in a very large number of alleles for long tandem repeat arrays (Huang et al., 2002).

Data base searches have demonstrated that both dinucleotide and trinucleotide repeats are frequent in the plant genome, with at least one repeat greater than 20 bp in length in every 30 kilo base (kb) throughout the genome (Taramino and Tingey, 1996). In plant species, repeats containing $(AT)_n$ were found to be the most frequent dinucleotide repeat. In contrast to the human genome, $(AC)_n$ repeats were found to be much less abundant in plants (Akkaya et al., 1992). The frequency of each class of SSR also appears to be different between plant species (Wang et al., 1994).

The SSR technique uses PCR to amplify DNA fragments by repeated cycles of DNA denaturation, annealing and extension using DNA polymerase enzyme and resulting DNA fragments are separated by banding on a gel (Akkaya et al., 1992). This method has many advantages for genetic fingerprinting such as being rapid and reliable

(Diwan and Cregan, 1997), abundant (Lagercrantz et al., 1993), co-dominant (Akkaya et al., 1992), highly heterozygous (Powell et al., 1996) and highly polymorphic (Akkaya et al., 1995). However, it needs characterisation of primers in every species of interest, which is expensive and time consuming (Kubik et al., 2001). There are observations that primers developed in one species can amplify microsatellite loci of related species (Sosinski et al., 2000; Hormaza, 2002).

SSR has been used in the construction of a genetic linkage map of soybean (Akkaya et al., 1995), rice (Temnykh et al., 2000), corn (Taramino and Tingey, 1996) and sorghum (Taramino et al., 1997). It was used in genetic diversity analysis of rice (Cho et al., 2000), barley (Struss and Plieske, 1998), wheat (Prasad et al., 2000; Huang et al., 2002; Akkaya and Buyukunal–Bal, 2004), sorghum (Dean et al., 1999; Smith et al., 2000), soybean (Tanya et al., 2001), potato (McGregor et al., 2000), ryegrass (Kubik et al., 2001), apricot (Hormaza, 2002), peach (Aranzana et al., 2003), carnation (Smulders et al., 2003) and coffee (Anthony et al., 2002).

Westman and Kresovich (1999) reported that SSR polymorphism is consistent with variation patterns revealed by the species` breeding history. Results of several investigations concluded SSR as a marker of choice for genetic analysis of organisms with a narrow genetic base (Akkaya et al., 1995; Westman and Kresovich, 1999; Huang et al., 2002; Aranzana et al., 2003; Akkaya and Buyukunal-Bal, 2004).

Hormaza (2002) used 37 SSR primer pairs to assess the genetic relationship of 48 apricot genotypes and reported that 31 primer pairs resulted in correct amplification and 20 produced polymorphic and repeatable amplification patterns within the 48 genotypes. All genotypes were distinguished using 20 polymorphic SSR primer pairs. Aranzana et al. (2003) identified 87% of the 212 peach and nectarine cultivars studied using 16 polymorphic SSR primer pairs. Huang et al. (2002) estimated the genetic similarity of 998 accessions of bread wheat using 24 SSR primers and reported genetic similarity coefficients ranging from 0.15 to 0.96. Akkaya and Buyukunal-Bal (2004) assessed the genetic variation of 11 bread wheat cultivars using 19 highly polymorphic wheat SSR primers and observed 0.36 to 0.87 polymorphic information content (PIC). They attributed the observed high polymorphic information content to the use of informative markers. Huang et al. (2002) suggested that the repeat length of

SSRs and location on chromosomes should be considered for the selection of SSRs in the evaluation of genetic diversity and cultivar identification.

Kubik et al. (2001) identified seven perennial ryegrass cultivars using 22 SSR primers. Smith et al. (2000) uniquely distinguished 48 sorghum cultivars from the 50 studied using 15 SSR primers. McGregor et al. (2000) identified 24 potato cultivars using five SSR primers. Moreover, they identified one SSR primer which can discriminate all 24 potato cultivars. These results revealed efficiency difference of different SSR primers.

Dean et al. (1999) identified five redundant accessions of sorghum from 19 orange sorghum holdings using 15 SSR primers. Therefore SSR is useful to pinpoint redundant accessions from collections for efficient conservation of germplasm.

In arabica coffee, polymorphic SSR markers were characterised recently (Combes et al., 2000; Rovelli et al., 2000; Moncada and McCouch, 2004). Simple sequence repeat loci $(TG)_n$ were more polymorphic than $(ATC)_n$. Out of 20 $(TG)_n$ and 12 $(ATC)_n$ SSR loci screened 12 and one loci proved to be polymorphic, respectively. Furthermore, some of the primer pairs amplified single loci of the homologous chromosome pair, i.e. they were able to discriminate between the two chromosome-sets derived from the diploid ancestral parents (Rovelli et al., 2000). This result could offer an interesting key for studying the origin of the two sets of chromosomes, the actual contribution of the two diploid progenitors as well as the possible functional evolution of the homologous loci within this organism. Combes et al. (2000) identified five polymorphic loci out of 11 SSR loci tested for genetic diversity analysis in arabica coffee and the mean heterozygosity value was 0.04. Results illustrated the presence of low genetic diversity in *C. arabica* as a consequence of its origin, reproductive biology and evolution (Lashermes et al., 2000b). Anthony et al. (2002) studied the genetic diversity within and among Typica-, Bourbon- and subspontaneous-derived accessions using six SSR loci and identified two alleles (M24-147 and M24-149) which discriminated the Typica derived accessions from the Bourbon derived accessions. Recently, Moncada and McCouch (2004) analysed the genetic relationship of 30 genotypes of *Coffea* using 34 SSR primers and reported average genetic similarity coefficients of 0.18 among diploids, 0.34 among wild tetraploids

and 0.59 among cultivated tetraploids. Higher genetic diversity was observed among diploids than tetraploids since diploids are outcrossing and obtained from diverse geographical regions.

Results of several investigations concluded SSR as a marker of choice for genetic analysis of organisms with a narrow genetic base (Akkaya et al., 1995; Huang et al., 2002; Aranzana et al., 2003; Akkaya and Buyukunal-Bal, 2004). Arabica coffee is repeatedly reported as a narrow genetic base crop, but the genetic relationship of arabica coffee genotypes currently grown in Ethiopia was not assessed using SSR markers.

2.7.5 Comparison of different genetic diversity assessment methods

Traditionally genetic diversity of genotypes was assessed based on differences in range of expressions of morphological and agronomical characters. Currently, a variety of molecular techniques are available for measuring genetic diversity. The most common ones are RFLP, RAPD, AFLP and SSR. Recently, S-SAPs (sequence-specific amplified polymorphism) (Waugh et al., 1997, Ramsay et al., 1999), IRAPs (Inter-retrotransposon amplified polymorphisms) and REMAPs (Retrotransposon-microsatellite amplified polymorphisms) (Kalendar et al., 1999) are available and have been applied in a number of crop species (Simioniuc et al., 2002). All of them detect polymorphism by assaying subsets of the total amount of DNA sequence variation in a genome. However, they differ in principle, application, type and amount of polymorphism detected and cost and time requirements (Karp et al., 1998). Consequently, comparison of the efficiency of different types of molecular markers is vital and carried out by several investigators on a number of plant species.

Ajmone-Marsan et al. (1998) compared the efficiency of RFLP and AFLP to assess the genetic diversity of 13 maize inbred lines and reported the superiority of AFLPs. Fuentes et al. (1999) assessed the usefulness of isozyme, RAPD and AFLP markers for Cuban rice genetic diversity analysis and reported the superiority of AFLP compared to RAPD and isozymes in detecting polymorphisms. Yee et al. (1999) compared RAPD and AFLP markers in assessing diversity among selected *Vigna angularis* (Azuki) accessions and reported the superiority of AFLP for generating

polymorphisms and distinguishing accessions. Similar results were reported by Das et al. (1999) while studying genetic diversity of *Brassica campestris* cultivars using AFLP and RAPD analysis. AFLP methodology detected more polymorphism than RAPD due to a greater number of loci assayed per reaction.

Garcia-Mas et al. (2000) compared AFLP, RAPD and RFLP markers for measuring genetic diversity in melon and reported that although all three types of markers were equally informative, AFLP showed the highest efficiency in detecting polymorphism. Simioniuc et al. (2002) studied the genetic diversity and relationships among 21 pea cultivars using RAPD and AFLP and reported that although similarities between RAPD and AFLP clustering were obvious ($r = 0.79$) there were also some differences that may be due to the fact that AFLP scan different parts of the genome than RAPD.

Pejic et al. (1998) studied the genetic similarity of 33 maize inbred lines using AFLP, RAPD, RFLP and SSR markers and reported that except for RAPD, the genetic similarity trees of the three markers were highly correlated and concluded that both SSR and AFLP technologies can replace RFLP analysis for genetic similarity studies on maize. Russell et al. (1997) compared the efficiency of AFLP, RAPD, RFLP and SSR technologies to estimate the genetic relationships among 18 cultivated barley accessions and reported that all of the approaches were able to uniquely fingerprint each of the accessions but the four assays differed in the amount of polymorphism detected. Similarly, Powell et al. (1996) evaluated the utility of AFLP, RAPD, RFLP and SSR markers in both cultivated (*Glycine max*) and wild (*Glycine soja*) soybean accessions and reported high correlations indicating congruence between AFLP, RFLP and SSR assays. The correlations of RAPD marker data with those obtained using other marker systems were lower.

Milbourne et al. (1997) compared AFLP, RAPD and SSR markers in genotyping 16 potato cultivars and reported that although all the three approaches successfully discriminated 16 cultivars using a minimum of one assay, similarity matrices produced for each marker type on the basis of Nei and Li coefficients showed low correlation. AFLP had the highest effective multiplex ratio and SSR the highest diversity index.

Maguire et al. (2002) did comparative analysis of the genetic diversity in the mangrove species using AFLP and SSR and reported the congruence between AFLP and SSR data sets, which suggested that either method or a combination was applicable to assess genetic diversity of mangroves. Uptmoor et al. (2003) compared the efficiency of RAPDs, AFLPs and SSRs for analysis of genetic relatedness of sorghum accessions from South Africa and reported that RAPD and AFLP similarity indices were highly correlated (r = 0.81), while the Spearman`s rank correlation coefficient between SSRs and AFLPs (r = 0.57) and RAPDs and SSRs (r = 0.51) were relatively low.

Almanza-Pinzon et al. (2003) compared molecular markers and coefficient of parentage for the analysis of genetic diversity among spring bread wheat accessions and concluded that AFLP and SSR markers were generally in agreement with estimates of diversity measured using coefficient of parentage, especially when complete pedigree data were available. Therefore, results of several investigations on different crop species revealed the efficiency of AFLP and SSR analysis for genetic diversity assessment and identification of cultivars.

Among DNA-based markers, RFLP, RAPD, AFLP and SSR have been employed independently at different times to assess the genetic diversity among and within arabica coffee genotypes and group them according to genetic similarity. Results indicated that RFLP and RAPD showed a relatively low degree of polymorphism (Lashermes et al., 1993; 1996a; b; Orozco-Castillo et al., 1994; Paillard et al., 1996; Anthony et al., 2001). On the other hand, AFLP and SSR showed a relatively high degree of polymorphism (Lashermes et al., 2000a; Anthony et al., 2002; Steiger et al., 2002). However, the efficiency of these DNA-based markers was not compared using the same sets of genotypes.

Several investigators compared the efficiency of morphological, biochemical and molecular markers for genetic diversity assessment in different crop species. The molecular marker technique RAPD was more effective compared to morphological traits for classifying Chilean common bean landraces into the Andean and Mesoamerican gene pools (Johns et al., 1997). Bahrman et al. (1999) while assessing the genetic diversity of 26 winter barley varieties using molecular, biochemical and

morphological markers, observed higher differentiation levels between varieties using DNA markers compared to biochemical and morphological markers.

Iketani et al. (1998) observed incongruence between RFLP and morphological characters in pear classification, but suggested that these disorders could be explained as the result of hybridisations and introgression between species. Briard et al. (2002) studied the genetic diversity of wild seakale (*Crambe maritime* L.) using morphological and RAPD markers and observed lack of clear correlation between results of morphological and molecular classification. Lage et al. (2003) evaluated the genetic diversity in synthetic hexaploid wheat parents using AFLP and agronomic traits and reported non-significant correlation between genetic distances obtained from AFLP and agronomic data. Geleta et al. (2004) reported lack of similarity in clustering of pepper genotypes using AFLP and morphological markers.

Martinello et al. (2001) observed a correlation coefficient of 0.62 between molecular and morphological data while estimating genetic diversity in okra (*Abelmoschus* spp.). The correlation coefficient was increased to 0.88 when the comparison was made between molecular marker and only quantitative trait data. Results further pointed out the relative importance of quantitative traits compared to qualitative traits in genetic diversity assessment of okra. Ayele et al. (1999) reported the better resolution of quantitative traits as compared to qualitative morphological traits for characterisation of two *Eragrostis* species.

Tatineni et al. (1996) reported strong correlation (r = 0.63) between morphological characters and RAPD measurement of distances in genetic diversity analysis of elite cotton germplasm. Goulao et al. (2001) obtained a significant correlation among similarity matrices of RAPD and inter simple sequence repeats (ISSR) with binarized morphological data, r = 0.72 and 0.71 respectively, while characterising cultivated chestnut (*Castanea sativa* Mill.) cultivars.

Hamza et al. (2004) studied the genetic diversity of 26 Tunisian winter barley cultivars using SSR markers and morphological traits and reported significant correlation between the two diversity measures and correspondence of clusters constructed using morphological and SSR data. Jacoby et al. (2003) and Rotondi et al.

(2003) reported similarity in clustering of genotypes using molecular and morphological characters in *Solannum* and olive genotypes, respectively.

Although comparison among DNA markers as well as DNA and morphological markers are important for marker selection and were made on a number of crop species, to date such type of studies were not reported on arabica coffee.

CHAPTER 3

MORPHOLOGICAL VARIATION AMONG ETHIOPIAN ARABICA COFFEE GENOTYPES

3.1 Abstract

Sixteen arabica coffee genotypes of Ethiopia were evaluated for nine qualitative and 18 quantitative characters using a randomised complete block design with three replications at Finoteselam, Ethiopia. The main objectives of the study were to assess phenotypic variation among genotypes, estimate correlation coefficients among various quantitative characters, estimate phenotypic as well as genetic variances and estimate broad sense heritability and genetic advance of various characters. Cluster analysis using a combination of nine qualitative and 18 quantitative characters grouped the 16 coffee genotypes into two main clusters. The first main cluster consisted of vigorously growing and high yielding genotypes while the second main cluster had dwarf, compact growth habit and low yielding genotypes. Quantitative characters compared to qualitative characters had better resolution power in clustering genotypes. Of the 15 principal components involved in explaining the entire variation, the first four explained 82.3% of the total variation. The first principal component which accounted for 37.8% of the total variation was attributed to inter-node length, canopy diameter, seed and fruit width, trunk diameter, leaf width, tree height, petiole length, fruit length, bean thickness and length, number of primary branches and average green bean yield per tree. Average green bean yield per tree had positive correlations with all characters. Average green bean yield per tree was significantly correlated with percentage of bearing primary branches per tree, bean weight, canopy diameter, tree height, trunk diameter and seed thickness. The phenotypic and genotypic coefficients of variation ranged from 4.5-53.4% and 3.3-51.7%, respectively. Estimates of broad sense heritability varied from 0.38 for bean thickness to 0.94 for number of secondary branches per tree as well as 100 green bean weight. Values for genetic advance (expected from selection of the superior 5% of the genotype, expressed relative to the means) ranged from 4.3% for bean thickness to 103.4% for number of secondary branches per tree. This study indicated the presence of high levels of diversity among genotypes for evaluated characters.

3.2 Introduction

Knowledge of genetic diversity and relationships among elite breeding materials is important for the improvement of crop plants. It can be applied for selection of promising parents in hybrid variety and inbred line development, in assigning inbred lines to heterotic groups, variety registration and protection (Pejic et al., 1998) and to estimate the potential of genetic gain in a breeding programme (Almanza-Pinzon et al., 2003). It is essential for identification of duplicated accessions among collections and for efficient conservation and utilisation of available genetic resources (Sakiyama, 2000). It assists in maintaining genetic diversity and sustained long-term selection gain in a breeding programme (Chowdhury et al., 2002). Furthermore, evaluation and grouping of landraces of a crop of a certain region is helpful in studying the evolutionary relationships in line with the history of the crop in that region (Zeven, 1990).

Currently, different genetic diversity assessment methods such as morphological, biochemical and molecular markers are available. Of these, the morphological method is the oldest and considered as the first step in description and classification of germplasm (Smith and Smith, 1989). However, morphological markers are relatively less effective for genetic diversity analysis due to sensitivity to environmental influences and developmental stage of the plant (Werlemark et al., 1999). Newbury and Ford-Lloyed (1993) and Hardon et al. (1994) reported that although morphological characters are often influenced by environmental conditions, the method is still useful and easy to apply for classification, estimating diversity and registration of cultivars. Consequently it has been widely used to classify and estimate diversity in a variety of crop species including coffee (Dias et al., 1993; Montagnon and Bouharmont, 1996; Zewdie and Zeven, 1997; Assefa et al., 1999; Mars and Marrakchi, 1999; Elhoumaizi et al., 2002; Lucchin et al., 2003; Upadhyaya, 2003; Van de Wouw et al., 2003). Since natural selection affects morphological traits linked to adaptive characteristics, genetic distances from quantitative data allow inferences about adaptation and co-adaptation patterns of populations (Camussi et al., 1985). Ortiz et al. (2001) emphasized the importance of phenotypic patterns of variation for organising collections within gene banks. Tatineni et al. (1996), Goulao et al. (2001), Martinello et al. (2001) and Hamza et al. (2004) demonstrated the correlation of

morphological and molecular markers in discriminating genotypes in cotton, chestnut, okra and barley, respectively.

Among morphological characters, Ayele et al. (1999) and Martinello et al. (2001) observed a better discrimination power of quantitative morphological characters compared to qualitative characters in different tef as well as okra genotypes. Camussi et al. (1985) indicated the importance of morphological data showing continuous distribution for inter-group classification below species level.

In arabica coffee, Sylvain (1955) classified Ethiopian cultivated coffee into 13 main types namely *Agaro, Arbagugu, Cioiccie, Coulo, Dilla, Ennarea, Harar, Irgalem, Kaffa, Tafari kela, Wolkite, Wollamo* and *Zeghie* using fruit and bean morphology. In addition, Montagnon and Bouharmont (1996) classified wild and cultivated Ethiopian coffee genotypes according to geographic origin using 18 agro-morphological characters. Several investigators reported variation in branching habit, young leaf colour, fruit colour, calyx persistence and leaf and fruit sizes among Ethiopian coffee genotypes (FAO, 1968a; Ameha, 1986; Wondimu, 1998; Teketay, 1999; Dessalegn, 2002). Therefore, this study was conducted to (1) assess morphological variation among currently grown Ethiopian arabica coffee genotypes, (2) classify and identify similar genotypes through cluster analysis, (3) identify characters accounting for high levels of variation to total variation through principal component analysis, (4) estimate the association of different characters through correlation analysis and (5) estimate heritability (broad sense) and expected genetic advance of different traits.

3.3 Materials and Methods

3.3.1 Description of the trial site

The study was conducted at Finoteselam coffee trial site, Ethiopia, in 2003. The trial site is located at $10^0$67` North latitude, $37^0$11` East longitude and at an altitude of 1850 metre above sea level (masl). The soil at the trial site had a pH of 6.5, 3.02% organic carbon, 5.21% organic matter, 0.13% total nitrogen and 10.11 ppm available phosphorous. It is deep, well-drained and slightly acidic nitosol, thus suitable for arabica coffee production. The area receives 950 mm rainfall annually, of which 88%

is usually received from May to October. The monthly average minimum and maximum temperature of Finoteselam, ranged between 9.4-15.5 ^{o}C and 24.2-33.1 ^{o}C, respectively.

3.3.2 Genotypes

Sixteen *C. arabica* genotypes collected from the northwestern and southwestern parts of Ethiopia were used for this investigation (Table 3.1). These genotypes were obtained from Adet Agricultural Research Centre of the Amhara National Regional State Agricultural Research Institute. Coffee trees were five years old during the course of this investigation and were grown under *Sesbania sesban* shade (Fig.3.1). Each coffee tree received 20.5 g nitrogen and 19 g P_2O_5 three times per year in March, June and September as recommended (Bellachew et al., 1998). Fertilizer was applied by hand under each tree in a circular strip corresponding to the spread of the branches starting 10 cm away from the main stem. Coffee trees were irrigated during the dry season. Trial plots were cultivated and weeded by hand as required. Pesticide and fungicide were not applied since serious disease and insect pest outbreak was not observed during the study period.

The study was conducted using a randomised complete block experimental design with three replications. Each plot comprised of two rows of five coffee trees each. Spacing between trees and rows was 2 m. Shade trees were planted in rows at 5 m spacing between each other in the trial plot. The spacing between plots and replications was 2 m and 3 m, respectively.

Table 3.1 List of arabica coffee genotypes evaluated for morphological variation

No.	Genotype	Collection Site Altitude (masl)	Collection Region
1	AD0191	1800	Southwestern Ethiopia
2	AD0291	1800	Southwestern Ethiopia
3	AD0391	1900	Southwestern Ethiopia
4	AD0491	1920	Southwestern Ethiopia
5	AD0591	1700	Southwestern Ethiopia
6	AD0691	1700	Southwestern Ethiopia
7	AD0791	1700	Southwestern Ethiopia
8	AD0891	1710	Southwestern Ethiopia
9	AD0991	1710	Southwestern Ethiopia
10	AD1091	1710	Southwestern Ethiopia
11	AD1191	1710	Southwestern Ethiopia
12	AD1291	1820	Northwestern Ethiopia
13	AD1391	1720	Southwestern Ethiopia
14	AD1491	1575	Southwestern Ethiopia
15	AD1591	1900	Southwestern Ethiopia
16	AD1691	1950	Northwestern Ethiopia

masl = metre above sea level

Fig. 3.1 Partial view of the coffee experimental field at Finoteselam, Ethiopia

3.3.3 Data collection and statistical analysis

During the course of this study, data of 27 different agro-morphological characters were collected from each genotype (Table 3.2). Data for 18 quantitative characters such as average green bean yield per tree (g), 100 green bean weight (g), bean length (mm), bean width (mm), bean thickness (mm), fruit length (mm), fruit width (mm), tree height (cm), canopy diameter (cm), trunk diameter (mm), inter-node length of orthotropic branch (cm), inter-node length of primary branches (cm), leaf length (cm), leaf width (cm), petiole length (mm), number of primary branches per tree, number of secondary branches per tree and percentage of bearing primary branches per tree were collected as an average value of the central six trees of each plot. Method of quantitative characters data collection is described in Table 3.2. Data for nine qualitative characters such as young leaf colour, leaf shape, leaf apex shape, stipule shape, plant habit, fruit shape, calyx limb persistence, fruit colour and overall appearance of the tree were collected according to the International Plant Genetic Resources Institute (IPGRI, 1996) coffee descriptor on a plot basis.

Data of quantitative characters were subjected to analysis of variance using MSTAT-C statistical package (Anonymous, 1991) to examine the presence of statistically significant differences among genotypes for these characters. Duncan Multiple Range Test (DMRT) was performed to identify genotypes that were significantly different from each other. Correlations among quantitative characters and phenotypic similarity of genotypes were estimated using correlation and cluster analyses, respectively. The importance of different traits in explaining multivariate polymorphism was assessed using principal component analysis (PCA) (Upadhyaya et al., 2002).

Table 3.2 List of recorded morphological characters and brief descriptions

Character no.	Character and descriptive value
	Quantitative characters
1	Green bean yield per tree (Gbyt) (g): Weight of green beans harvested per tree
2	100 green bean weight (Gbw) (g): Average of four samples of 100 beans weight of each tree
3	Bean length (Bl) (mm): Average of ten normal beans of each tree measured at the longest part
4	Bean width (Bw) (mm): Average of ten normal beans of each tree measured at the widest part
5	Bean thickness (Bt) (mm): Average of ten normal beans of each tree measured at the thickest part
6	Fruit length (Frul) (mm): Average of ten normal and mature green fruits of each tree measured at the longest part
7	Fruit width (Fruw) (mm): Average of ten normal and mature green fruits of each tree measured at the widest part
8	Tree height (Tht) (cm): The length from the ground level to the tip of the tree
9	Canopy diameter (Cand) (cm): Average length of tree canopy measured twice, east-west and north-south, from the widest portion of the tree canopy
10	Main stem diameter (Trud) (mm): Measured as a diameter of the main stem at five cm above the ground
11	Inter-node length on orthotropic branch (Ilob) (cm): Computed per tree as (TH–HFPB)/TNN-1, where TH = total plant height, HFPB =height up to first primary branch, TNN = total number of main stem nodes
12	Inter-node length on primary branches (Ilpb) (cm): Estimated from four primaries per tree. The first four undamaged primaries from the bottom were selected. For each primary the length divided by the number of nodes and inter-node length for each tree was taken as the average value of the four primaries
13	Leaf length (Leal) (cm): Average of five normal (> node 3 from the terminal bud) leaves, measured from petiole end to apex
14	Leaf width (Leaw) (cm): Average of five normal (> node 3 from the terminal bud) leaves, measured at the widest part
15	Leaf petiole length (Petl) (mm): Average of five normal (> node 3 from the terminal bud) petioles, measured from the base to the insertion with the blade
16	Number of primary branches (Npb) (no): total number of primary branches counted per tree
17	Number of secondary branches (Nsb) (no): total number of secondary branches counted per tree
18	Percentage of bearing primary branches (Pbpb) (%): computed per tree as (NBPB/Npb) * 100, where NBPB = number of bearing primary branches per tree, Npb= total number of primary branches per tree
	Qualitative characters
19	Young leaf colour (Ylc): 0 (bronze), 1 (green)
20	Leaf shape (Leas): 0 (lanceolate), 1 (ovate)
21	Leaf apex shape (Leaas): 0 (apiculate), 1 (acuminate)
22	Stipule shape (Sts): 0 (ovate), 1 (deltate)
23	Plant habit (Ph): 0 (tree), 1 (shrub)
24	Fruit shape (Frus): 0 (round), 1 (obovate)
25	Fruit colour (Fruc): 0 (yellow), 1 (light red), 2 (dark red)
26	Calyx limb persistence (Clp): 0 (not persistent), 1(persistent)
27	Overall appearance (Oaa): 0 (elongated), 1 (pyramid), (2) bushy

In order to reduce the influence of outliers and scale differences during cluster and principal component analyses, data of quantitative characters were standardised as follows. The mean observation for each genotype was standardised by subtracting the mean value of the character and by subsequently dividing with its respective standard deviation (Zewdie and Zeven, 1997; Upadhyaya et al., 2002; Hamza et al., 2004).

This resulted in standardised values for each trait with an average of zero and standard deviation of one or less. These standardised values were used to perform principal component and cluster analyses. During principal component and cluster analyses, 16 genotypes were considered as operational taxonomic units and were represented by rows and 18 quantitative and nine qualitative characters taken as variables by columns. Principal component, cluster and correlation analyses were performed using the Number Cruncher Statistical System, NCSS 2000 (Hintze, 1998). The unweighted pair group method with arithmetic averages (UPGMA) of cluster analysis was used to construct the dendrogram of coffee genotypes. The similarity of dendrograms constructed using nine qualitative, 18 quantitative and a combination of nine qualitative and 18 quantitative characters were tested by calculating correlation coefficients using pair-wise Euclidian distances estimated during cluster analysis (Ayele et al., 1999; Hamza et al., 2004). The diversity of qualitative characters was estimated using Shannon-Weaver (1949) diversity index, $H` = -\Sigma_j^i PilogPi$, where Pi is the relative frequency in the i^{th} category of j^{th} trait.

Genetic parameters were estimated to identify genetic variability among genotypes and determine genetic and environmental effects on various characters. These genetic parameters were estimated by adapting the formulae suggested by Allard (1960) and Singh and Chaudhary (1985). Accordingly, genotypic and phenotypic components of variance and coefficients of variability were estimated using the following formulae:

i. Genotypic variance, GV= (MSg –MSe)/r, where MSg = mean square of genotypes, MSe = mean square of error, and r = number of replications

ii. Phenotypic variance, PV= GV + MSe, where GV = genotypic variance and MSe = mean square of error

iii. Phenotypic coefficient of variation, PCV = (√PV/X) * 100, where PV = phenotypic variance and X = mean of the character

iv. Genotypic coefficient of variation, GCV = (√GV/X) * 100, where GV = genotypic variance and X = mean of the character

v. Heritability (Broad sense heritability), H = GV/PV, where GV = genotypic variance and PV = phenotypic variance

vi. Genetic advance (GA) expected and GA as percent of the mean assuming selection of the superior 5% of the genotypes were estimated in accordance with the methods illustrated by Assefa et al. (1999) as:
GA = K * (√PV/X)*H and

GA (as % of the mean) = (GA/X) * 100, where K is a constant (which varies depending upon the selection intensity and, if the latter is 5%, stands at 2.06), √PV/X is phenotypic standard deviation, H is heritability and X refers to mean of the character being evaluated

3.4 Results

Analysis of variance showed highly significant ($p \leq 0.01$) differences among genotypes for all 18 quantitative characters, indicating the presence of substantial variability among tested genotypes (Table 3.3). Average green bean yield per tree ranged from 144.6-566.7 g among genotypes. The lowest and highest average green bean yields per tree were recorded from genotype AD1091 and AD1291, respectively. AD1291 had significantly higher average green bean yield per tree compared to all other genotypes except for AD0291, AD0391 and AD1691. Dwarf genotypes with compact growth habits had relatively lower average green bean yield per tree compared to genotypes with vigorous growth habits. Although dwarf genotypes were characterised by low average green bean yield per tree, productivity can be increased by planting large numbers of trees per unit area of land. Similarly, genotypes were significantly different in 100 green bean weight. AD1691, a vigorous genotype had significantly heavy beans compared to genotypes AD0891 and AD1091.

Most genotypes were not significantly different in terms of trunk diameter. Genotype AD1091 had the smallest trunk diameter and was statistically different from all other genotypes except AD0891, AD0991 and AD1191. Genotypes with small trunk diameter were also characterised with relatively low average green bean yield per tree and dwarf growth habit.

Highly significant differences among genotypes were observed for tree height. Among all genotypes tested, AD0491 and AD1691 were the shortest and tallest, respectively. AD0491 was significantly shorter than all genotypes except AD0891, AD0991, AD1091 and AD1191. Similarly, genotypes were significantly different for canopy diameter. Of the tested genotypes, AD0991 had the smallest and AD1391 the largest canopy diameter. AD1391 had significantly higher canopy diameter compared to all other genotypes. AD0991 had significantly lower canopy diameter compared to all other genotypes except AD0891, AD1091 and AD1191.

Genotypes were significantly different in inter-node length of both orthotropic and primary plagiotropic branches. AD0891 and AD0391 had the shortest and the longest orthotropic branch inter-node length, respectively. On the other hand, AD1191 and AD1391 had the shortest and longest inter-node length of primary branches, respectively. Genotypes with short inter-node length are important for the development of varieties with compact growth habit.

Genotypes were significantly different for number of primary and secondary branches per tree. Among all genotypes tested, AD0491 and AD1691 had the lowest and highest number of primary branches per tree, respectively. AD1091 had the lowest and AD0791 the highest number of secondary branches per tree. Genotypes were significantly different for percentage of bearing primary branches per tree. Of all genotypes tested, AD0891 had the lowest and AD1291 the highest percentage of bearing primary branches per tree. Generally, coffee genotypes with large numbers of primary branches, high percentage of bearing primary branches and few numbers of secondary branches are desirable.

Table 3.3 Means of 18 quantitative morphological characters measured for 16 arabica coffee genotypes at Finoteselam, Ethiopia

Genotype	Gbyt	Gbw	Bl	Bw	Bt	Frul	Fruw	Tht	Cand	Trud	Ilob	Ilpb	Leal	Leaw	Petl	Npb	Nsb	Pbpb
AD0191	228.1	11.0	8.6	6.2	3.8	14.0	12.7	150.7	192.2	39.6	5.5	4.3	12.5	5.5	6.8	47.7	96.0	54.7
AD0291	480.2	13.0	9.3	6.2	3.7	14.5	12.6	154.2	208.3	38.3	5.7	5.0	12.2	5.2	7.4	42.3	43.0	54.7
AD0391	425.1	11.3	8.9	6.0	3.6	13.7	12.5	167.7	219.1	36.0	6.8	5.1	11.6	5.0	6.8	43.7	60.3	61.3
AD0491	276.1	11.0	8.1	5.8	3.6	13.1	12.3	107.5	196.8	32.2	4.6	3.6	10.4	4.2	6.0	35.7	72.7	64.7
AD0591	310.3	12.0	8.9	6.6	3.8	13.1	12.8	151.3	208.3	38.3	6.0	4.8	11.3	5.1	6.4	38.7	42.0	55.7
AD0691	308.4	14.0	9.8	6.0	3.7	15.1	13.2	141.9	183.9	36.4	5.1	4.6	12.0	5.3	7.0	45.0	63.7	52.3
AD0791	354.9	14.0	9.9	6.2	3.8	14.4	12.5	161.2	201.2	35.2	4.9	4.4	12.2	5.4	6.5	51.7	117.7	55.0
AD0891	180.6	10.3	8.3	5.7	3.5	12.7	12.2	121.2	138.3	27.0	3.8	3.1	10.2	4.0	5.0	52.7	38.0	45.7
AD0991	161.1	11.0	8.0	5.9	3.4	12.4	12.0	124.0	137.1	26.2	4.0	3.4	10.8	3.9	5.1	48.7	23.7	47.7
AD1091	144.6	10.3	8.2	5.9	3.5	12.7	12.2	119.1	138.7	24.6	4.1	3.6	12.5	4.5	5.7	51.0	21.3	51.3
AD1191	152.5	9.3	8.1	5.8	3.5	12.7	12.3	125.3	143.6	30.2	4.3	3.0	11.9	4.3	5.2	54.0	40.7	52.3
AD1291	566.7	13.3	9.3	6.1	3.8	13.5	12.5	179.3	205.9	39.0	5.1	4.6	11.4	4.6	5.5	60.3	116.7	66.7
AD1391	320.9	13.0	8.9	6.4	3.9	14.4	14.1	179.6	246.5	38.7	6.4	5.4	11.0	5.1	6.2	43.7	53.3	59.7
AD1491	266.0	12.3	9.8	6.1	3.6	14.1	12.1	179.6	193.6	32.4	5.8	4.5	10.4	4.9	6.3	53.3	40.3	54.7
AD1591	417.8	12.0	9.3	6.3	3.7	14.3	13.0	179.0	212.5	39.1	6.0	4.9	12.8	5.4	6.6	47.3	29.7	59.3
AD1691	532.6	16.0	8.8	6.2	3.9	13.9	12.5	182.8	213.2	35.6	5.0	4.3	12.0	5.0	6.7	62.0	60.7	64.0
F test	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**	**
LSD	133.8	1.0	0.8	0.4	0.3	0.9	0.7	25.3	20.6	6.9	1.3	1.2	1.2	0.5	1.4	7.9	17.4	13.6

** indicates significant at $p \leq 0.01$; Abbreviations as explained in Table 3.2; LSD = least significant difference

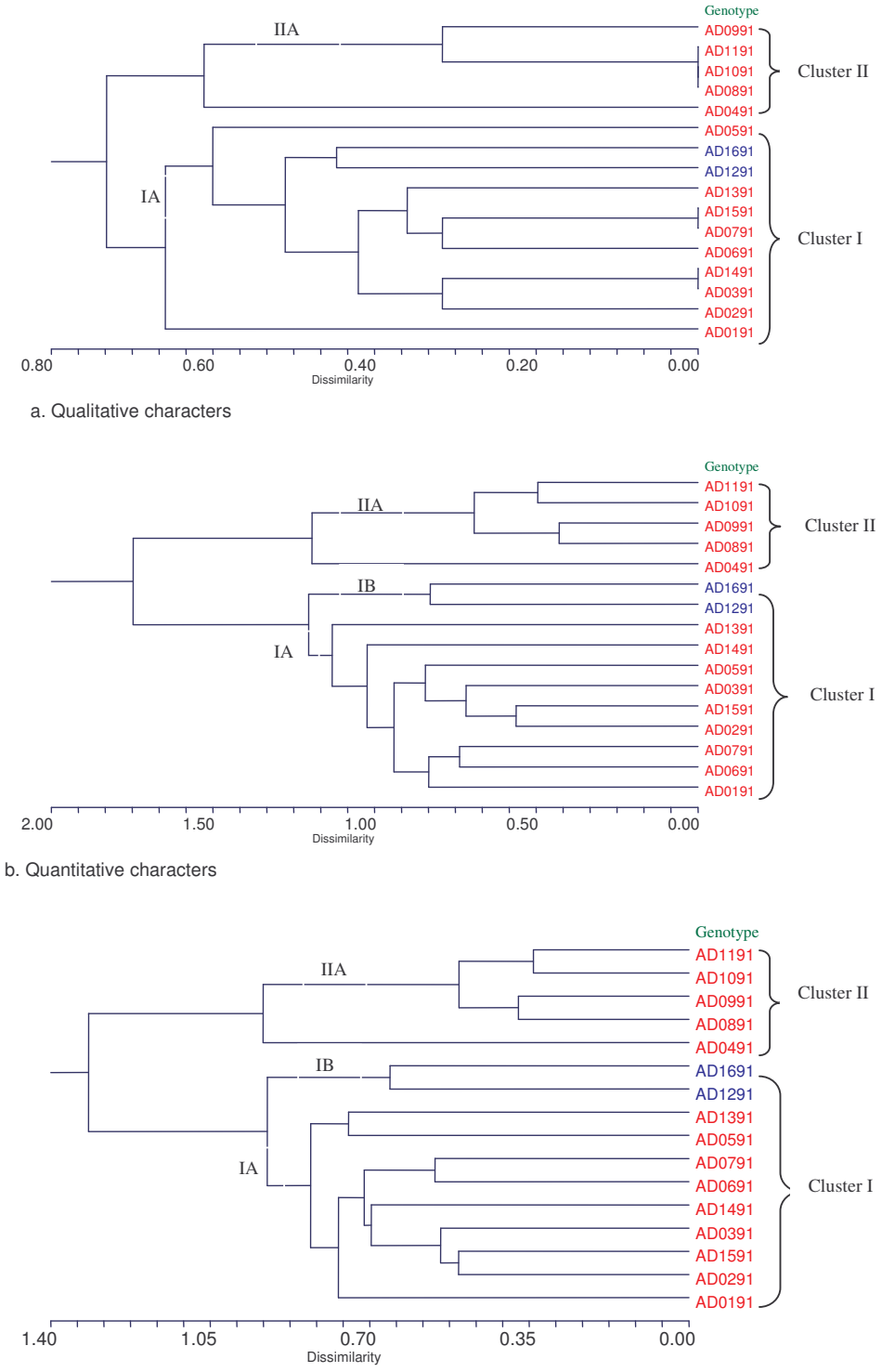
Substantial variability among genotypes was also revealed by big difference between minimum and maximum values as well as high standard deviations for 18 quantitative characters (Table 3.4). For example, average green bean yield per tree varied from 144.6-566.7 g and 100 green bean weight ranged from 9.3-16.0 g. Tree height varied from 107.5-182.8 cm, canopy diameter from 137.1-246.5 cm, trunk diameter from 24.6-39.6 mm, number of primary branches per tree from 35.7-62.0 and number of secondary branches per tree from 21.3-117.7. Differences between minimum and maximum values for other characters were also big. The minimum and maximum values of different characters were recorded on different genotypes. This entails conservation of all these genotypes for the improvement of different characters of arabica coffee.

Table 3.4 List of arabica coffee genotypes with extreme values for 18 quantitative morphological characters

Character	Minimum		Maximum		Mean±Stdv
	Value	Genotype	Value	Genotype	
Green bean yield per tree (g)	144.6	AD1091	566.7	AD1291	320.3±140.13
100 green bean weight (g)	9.3	AD1191	16.0	AD1691	12.1±1.71
Bean length (mm)	8.1	AD1191, AD0491	9.9	AD0791	8.9±0.68
Bean width (mm)	5.7	AD0891	6.6	AD0591	6.1±0.28
Bean thickness (mm)	3.4	AD0991	3.9	AD1391	3.7±0.19
Fruit length (mm)	13.1	AD0591	15.1	AD0691	13.7±0.83
Fruit width (mm)	12.0	AD0991	14.1	AD1391	12.6±0.56
Tree height (cm)	107.5	AD0491	182.8	AD1691	151.5±26.73
Canopy diameter (cm)	137.1	AD0791	246.5	AD1391	189.9±33.37
Trunk diameter (mm)	24.6	AD1091	39.6	AD0191	34.3±5.49
Orthotropic branch inter-node length (cm)	4.0	AD0991	6.8	AD0391	5.2±1.02
Primary branches inter-node length (cm)	3.0	AD1191	5.4	AD1391	4.3±0.85
Leaf length (cm)	10.2	AD0891	12.8	AD1591	11.6±0.91
Leaf width (cm)	3.9	AD0991	5.5	AD0191	4.8±0.56
Petiole length (mm)	5.0	AD0891	7.4	AD0291	6.2±0.89
No. of primary branches per tree	35.7	AD0491	62.0	AD1691	48.6±7.60
No. of secondary branches per tree	21.3	AD1091	117.7	AD0791	57.5±30.13
% of bearing primary branches	45.7	AD0891	66.7	AD1291	56.2±7.66

The two genotypes collected from the northwestern part of Ethiopia, AD1291 and AD1691, had relatively higher average green bean yields per tree, high 100 green bean weight and vigorous growth habits compared to all other genotypes. On the other hand, genotypes collected from southwestern Ethiopia were more diverse in growth habit, average green bean yield per tree, bean size and weight. For example, some had open canopies while others had compact growth habits. Generally, results of this study indicated the presence of enormous variability for the assessed characters among tested coffee genotypes.

The phenotypic similarity of 16 coffee genotypes was assessed by cluster analysis using nine qualitative, 18 quantitative and a combination of nine qualitative and 18 quantitative characters. According to nine qualitative characters of coffee the 16 coffee genotypes were classified into two main clusters (Fig. 3.2a). In the first main cluster 11 genotypes were assigned and further classified into one sub-cluster (IA, Fig.3.2a) and a singleton (AD0191). Sub-cluster IA of the first main cluster contained 10 genotypes, which were characterised by tree growth habits. Five coffee genotypes namely AD0991, AD1191, AD1091, AD0891 and AD0491 were assigned in the second main cluster. These genotypes were characterised by green young leaves, lanceolate leaf shape, shrub growth habit, dark-red fruit colour and persistent calyx limb. The second main cluster bifurcated into one sub-cluster (IIA, Fig.3.2a) with four genotypes and a singleton (AD0491). The singleton from cluster I, AD0191, had an elongated conical shape and persistent calyx limb. AD0491, the singleton from cluster II, was bushy in appearance, had yellow fruits and a lanceolate leaf shape. It was not possible to differentiate genotypes AD1191, AD1091 and AD0891 using qualitative characters. Similarly, genotypes AD1591 and AD0791 as well as AD0391 and AD1491 were identical based on the above mentioned nine qualitative characters.



Blue for northwestern Ethiopia and red for southwestern Ethiopia genotypes

Fig. 3.2 Dendrograms of 16 *C. arabica* genotypes constructed by cluster analysis using qualitative (a), quantitative (b) and combination of qualitative and quantitative (c) characters

According to 18 quantitative characters, the 16 coffee genotypes were grouped into two main clusters (Fig. 3.2b). The first main cluster consisted of 11 genotypes, which were vigorous in growth habit and high yielding compared to genotypes assigned in the second main cluster. The first main cluster bifurcated into two sub-clusters, IA and IB. The first sub-cluster, IA, comprised nine coffee genotypes, which were vigorous in growth habit, high yielding and collected from the southwestern part of Ethiopia. The second sub-cluster, IB, consisted of two genotypes, AD1291 and AD1691, collected from the northwestern part of the country. They were vigorous in growth habit, high yielding and adaptable to marginal areas. The second main cluster consisted of five genotypes, which were characterised by dwarf and compact growth habits and small sized leaf, bean and fruit. The second main cluster in turn bifurcated into one sub-cluster (IIA) with four genotypes and a singleton. The singleton, AD0491, was characterised by dwarf growth habit and weak apical dominance.

Cluster analysis performed using 27 agro-morphological (18 quantitative and nine qualitative) characters also grouped the 16 coffee genotypes into two main clusters (Fig. 3.2c). Eleven coffee genotypes appeared together in the first main cluster. The first main cluster further bifurcated into two sub-clusters (IA and IB, Fig. 3.2c). In the first sub-cluster (IA) nine genotypes were assigned together. These genotypes were vigorous in growth habits, had green young shoots and collected from the southwestern part of the country. The second sub-cluster (IB) of the first main cluster had two cultivars which were vigorous in growth habit, high yielding, adaptable to marginal environmental conditions and were collected from the northwestern part of the country. The second main cluster had five genotypes, which were compact in growth habit with small tree, bean, fruit and leaf sizes. The second main cluster bifurcated into one sub-cluster (IIA) with four genotypes and a singleton. The singleton, AD0491, was unique in its overall appearance and growth habit.

The ability of qualitative and quantitative characters for grouping genotypes similarly was assessed with correlation analysis using pair-wise Euclidian distances calculated during cluster analysis. Results indicated that the correlation coefficient between qualitative and quantitative characters for grouping genotypes similarly was 0.72. Correlation coefficients between qualitative characters alone and qualitative and

quantitative characters together and between quantitative characters alone and qualitative and quantitative characters together in classifying genotypes were 0.78 and 0.99, respectively. All three approaches grouped genotypes similarly, but quantitative characters were better than qualitative characters for grouping coffee genotypes according to geographic origin.

Principal component analysis grouped the 18 quantitative characters into 15 components, which accounted for 100% of the variability evident among tested genotypes. The first six principal components explained 91.5% of the total variation. The first four principal components (the only ones with eigenvalues over one) accounted for a cumulative of 82.3% of the entire variability apparent among tested genotypes (Table 3.5). The first principal component which explained 37.8% of the gross variability among genotypes had been attributed to variation in inter-node length on orthotropic and primary branches, canopy diameter, bean and fruit width, trunk diameter, leaf width, tree height, petiole and fruit length, number of primary branches, bean length and average green bean yield per tree. Likewise, 13.2% of the total variability among the tested genotypes accounted for the second principal component originated from variation in number of primary branches per tree, tree height, 100 green bean weight, bean length and average green bean yield per tree. The third principal component, which explained 14.8% of the total variation among genotypes were due to variation in leaf, petiole, fruit and bean length and leaf width. The fourth principal component which explained 16.5% of the total variation was associated with variation due to percentage of bearing primary branches, number of secondary branches, average green bean yield per tree, bean thickness and trunk and canopy diameters. Of all the characters evaluated, average green bean yield per tree and bean length contributed to variation in three out of the first four principal components.

Table 3.5 Eigenvectors and eigenvalues of the first four principal components for 18 quantitative morphological characters of 16 arabica coffee genotypes

Character	Eigenvectors			
	PC1	PC2	PC3	PC4
Green bean yield per tree (g)	-0.24	0.30	0.14	0.01
100 green bean weight (g)	-0.23	0.30	-0.10	0.15
Bean length (mm)	-0.23	0.10	-0.28	0.23
Bean width (mm)	-0.25	-0.15	0.04	0.16
Bean thickness (mm)	-0.27	0.12	0.07	-0.13
Fruit length (mm)	-0.26	-0.05	-0.27	0.01
Fruit width (mm)	-0.21	-0.26	0.09	0.02
Tree height (cm)	-0.26	0.19	0.02	0.40
Canopy diameter (cm)	-0.29	-0.04	0.28	-0.00
Trunk diameter (mm)	-0.29	-0.04	0.09	-0.18
Inter-node length on orthotropic branch (cm)	-0.26	-0.25	0.18	0.16
Inter-node length on primary branches (cm)	-0.29	-0.15	0.08	0.17
Leaf length (cm)	-0.12	-0.03	-0.52	-0.37
Leaf width (cm)	-0.27	-0.13	-0.31	-0.11
Petiole length (mm)	-0.25	-0.19	-0.23	-0.16
Number of primary branches	0.03	0.58	-0.23	0.29
Number of secondary branches	-0.14	0.36	0.06	-0.55
Percentage of bearing primary branches	-0.19	0.24	0.45	-0.26
Eigenvalue	6.80	2.37	2.66	2.97
Percent variation explained	37.80	13.19	14.80	16.50
Cumulative percent variation explained	37.80	50.99	65.79	82.29

PC = principal component

Shannon-Weaver diversity indices (H`) were calculated to compare phenotypic diversity among qualitative characters. A low H` indicates unbalanced frequency classes for an individual trait and lack of diversity for the trait. In this study Shannon-Weaver diversity values were variable among traits and ranged from 0.410 to 0.989 (Table 3.6). Traits such as leaf apex shape, plant habit, fruit shape and fruit colour were more diverse compared to overall appearance of the tree and young leaf colour. About 88% of the tested genotypes had green young leaves. Generally the diversity indices of all evaluated traits were above 0.400, indicating the presence of adequate variability for these traits among evaluated genotypes.

Table 3.6 Shannon-Weaver diversity indices for nine qualitative morphological characters of 16 evaluated *C. arabica* genotypes

Character	H` value
Young leaf colour	0.557
Leaf shape	0.893
Leaf apex shape	0.989
Stipule shape	0.701
Plant habit	0.951
Calyx limb persistence	0.811
Fruit shape	0.951
Fruit colour	0.933
Overall appearance of the tree	0.410

Correlation coefficients between 18 quantitative characters are given in Table 3.7. Average green bean yield per tree had positive associations with all characters and the association was statistically significant with percentage of bearing primary branches, 100 green bean weight, canopy diameter, tree height, trunk diameter, inter-node length on primary branches, bean thickness, bean length, petiole length and inter-node length on orthotropic branch. Similarly, 100 green bean weight showed positive correlations with all characters. Its association was statistically significant with bean thickness and length, fruit length, tree height, canopy diameter, inter-node length on primary branches, leaf width, petiole length, trunk diameter and bean width. On the other hand, number of primary branches per tree showed negative but statistically non-significant association with most of the characters considered in this study.

Table 3.7 Correlation coefficients among 18 quantitative morphological characters of arabica coffee

	Gbw	Bl	Bw	Bt	Frul	Fruw	Tht	Cand	Trud	Ilob	Ilpb	Leal	Leaw	Petl	Npb	Nsb	pbpb
Gbyt	0.774**	0.545*	0.463	0.664**	0.532*	0.264	0.733**	0.735**	0.707**	0.523*	0.673**	0.251	0.474	0.533*	0.195	0.441	0.760**
Gbw		0.667**	0.508*	0.747**	0.684**	0.368	0.666**	0.614*	0.537*	0.315	0.574*	0.197	0.544*	0.561*	0.269	0.410	0.460
Bl			0.491	0.522*	0.836**	0.330	0.684**	0.550*	0.590*	0.520*	0.682**	0.235	0.725**	0.616*	0.094	0.395	0.212
Bw				0.776**	0.494*	0.601*	0.681**	0.733**	0.750**	0.710**	0.785**	0.314	0.738**	0.565*	-0.187	0.144	0.334
Bt					0.650**	0.660**	0.707**	0.822**	0.833**	0.562*	0.704**	0.302	0.734**	0.582*	0.061	0.560*	0.599*
Frul						0.618*	0.630**	0.686**	0.723**	0.608*	0.756**	0.383	0.849**	0.821**	-0.099	0.358	0.289
Fruw							0.441	0.661**	0.643**	0.585*	0.679**	0.208	0.574*	0.417	-0.311	0.113	0.267
Tht								0.764**	0.706**	0.725**	0.787**	0.229	0.656**	0.488	0.326	0.289	0.530*
Cand									0.863**	0.857**	0.894**	0.130	0.694**	0.680**	-0.240	0.393	0.744**
Trud										0.779**	0.825**	0.351	0.801**	0.704**	-0.180	0.500*	0.598*
Ilob											0.914**	0.177	0.714**	0.691**	-0.348	0.118	0.490
Ilpb												0.267	0.786**	0.751**	-0.259	0.228	0.498*
Leal													0.629**	0.501*	0.125	0.127	0.079
Leaw														0.867**	-0.135	0.348	0.281
Petl															-0.331	0.206	0.325
Npb																0.191	0.022
Nsb																	0.491

* and ** indicates significant correlations at $p \leq 0.05$ and 0.01, respectively.

Abbreviations as explained in Table 3.2

Phenotypic and genotypic variances for 18 quantitative characters of coffee were estimated and are given in Table 3.8. The level of variation was different among different quantitative characters. The highest phenotypic coefficient of variation (PVC) was observed on number of secondary branches per tree (53.4%) while the lowest value was recorded on fruit width (4.5%). The phenotypic coefficient of variation for average green bean yield per tree was relatively high (44.6%) compared to 100 green bean weight (14.4%), bean length (7.8%), bean thickness (5.3%) and bean width (4.6%). In addition, relatively low phenotypic coefficient of variation was calculated for fruit and bean size characters compared to other quantitative characters.

The genetic coefficient of variation (GCV) for 18 quantitative characters ranged from 3.3-51.7% (Table 3.8). The maximum value was recorded for number of secondary branches per tree and the minimum for bean thickness. The genotypic coefficient of variation for fruit and bean size was relatively low compared to average green bean yield per tree, 100 green bean weight and other vegetative characters. Generally characters with higher phenotypic coefficients of variation showed higher genetic coefficients of variation, indicating the minimum effect of environment on the phenotypic expression of these characters.

Broad sense heritability (H) for 18 quantitative characters of arabica coffee was calculated using 16 coffee genotypes and values ranged from 0.38 to 0.94 (Table 3.8). Results indicated that most characters were highly heritable. The maximum value was for 100 green bean weight and number of secondary branches per tree while the minimum value was for seed thickness. Characters such as bean weight, number of secondary branches per tree and canopy diameter showed high broad sense heritability (>0.90). However, characters such as seed thickness, percentage of bearing primary branches, petiole length and inter-node length on primary branches had relatively low broad sense heritability (<0.60). Selection for traits with lower heritability values will be difficult due to environmental effects. Therefore in these genotypes, selection for number of secondary branches per tree, 100 green bean weight and canopy diameter will be much easier than for bean thickness and percentage of bearing primary branches per tree.

Genetic gain (GA) that could be expected from selecting the top 5% of the genotypes as percentage of the mean, varied from 4.3% for seed thickness to 103.4% for number of secondary branches per tree (Table 3.8). Relatively higher selection advance was expected for average green bean yield per tree, tree height and canopy diameter. Larger genotypic coefficients of variation along with high heritability and high genetic advance provide better information than each parameter alone. Therefore, characters that exhibited a high genotypic coefficient of variation, heritability and genetic advance would be useful as a base for selection. In this investigation, characters such as average green bean yield per tree, number of secondary branches per tree, canopy diameter, tree height and 100 green bean weight showed larger values for these three parameters. Moreover, four of them were positively and significantly correlated among each other. Therefore, selection for the improvement of one of the traits will also help for the simultaneous improvement of the other trait.

Table 3.8 Estimates of phenotypic (PCV) and genotypic (GCV) coefficients of variation, broad sense heritability (H) and genetic advance (GA) for 18 quantitative morphological characters in 16 *C. arabica* genotypes at Finoteselam, Ethiopia

Character	PCV (%)	GCV (%)	H	GA	GA (% of mean)
Green bean yield per tree (g)	44.6	40.5	0.83	244.20	76.2
100 green bean weight (g)	14.4	14.0	0.94	3.40	27.8
Bean length (mm)	7.8	6.7	0.73	1.00	11.7
Bean width (mm)	4.6	3.7	0.63	0.36	5.9
Bean thickness (mm)	5.3	3.3	0.38	0.16	4.3
Fruit length (mm)	6.2	5.5	0.78	1.40	10.0
Fruit width (mm)	4.5	3.8	0.72	0.85	6.8
Tree height (cm)	18.0	16.4	0.83	46.60	30.8
Canopy diameter (cm)	17.9	17.2	0.93	65.10	34.3
Trunk diameter (mm)	16.3	13.6	0.69	7.90	23.1
Orthotropic branch internode length (cm)	19.6	15.8	0.65	1.40	26.4
Primary branches internode length (cm)	19.9	15.2	0.58	1.00	24.0
Leaf length (cm)	8.0	6.5	0.65	1.30	10.8
Leaf width (cm)	11.7	10.7	0.84	0.99	20.5
Petiole length (mm)	14.3	10.4	0.53	0.97	15.6
No. primary branches/tree	15.9	14.1	0.79	12.60	25.8
No. secondary branches/tree	53.4	51.7	0.94	59.50	103.4
% bearing primary branches	13.8	8.6	0.39	6.20	11.1

3.5 Discussion

Analysis of variance showed highly significant variation ($p \leq 0.01$) among tested genotypes for all characters recorded (Table 3.3). Phenotypic variation among Ethiopian arabica coffee germplasm has been reported previously by several investigators for different characters (Sylvain, 1955; FAO, 1968a; Montagnon and Bouharmont, 1996). FAO (1968a), Ameha (1986) and Selvakumar and Sreenivasan (1989) observed phenotypic variation in branching habit, young leaf colour, fruit colour, calyx limb persistence, leaf and fruit size and cup quality among Ethiopian coffee genotypes. Therefore, results of this study were in agreement with previous findings. Broad trait variation apparent among the tested coffee genotypes entails the presence of ample opportunities and prospects for genetic improvement of these characters through selection either directly or following recombination through intra-specific hybridisation of desirable genotypes.

The range between minimum and maximum values of different characters closely corresponded to those established in previous coffee genotype evaluation results (FAO, 1968a; Coste, 1992). For example, according to the FAO (1968a) report, leaf length (cm) varied between 9.8-13.7, leaf width (cm) between 3.8-5.9, fruit length (mm) between 15.5-18.2 and fruit width (mm) between 13.1 and 15.3. Similarly, Coste (1992) reported values ranging between 10-15 cm for leaf length, 4-6 cm for leaf width, 10-15 mm for fruit width, 16-18 mm for fruit length, 10 mm for bean length, 6-7 mm for bean width, 3-4 mm for bean thickness and 0.15-0.20 g for bean weight among arabica coffee genotypes. Although characters vary depending upon variety, age of the tree, environmental conditions and cropping practices (Coste, 1992), most of the findings of the present study were in agreement with reports of FAO (1968a) and Coste (1992). The slightest discrepancies in extreme values for some characters could be ascribed to differences in number and type of tested genotypes, age of the tree and the location in which these studies were conducted.

Cluster analysis confirmed the presence of variation among genotypes. According to qualitative characters, genotypes were grouped into two main clusters, each consisting of one sub-cluster and a singleton (Fig. 3.2a). Based on quantitative morphological characters, genotypes were grouped into two main clusters subdivided into three sub-

clusters and one singleton (Fig. 3.2b). Similarly, based on both quantitative and qualitative characters, genotypes were grouped into two main clusters subdivided into three sub-clusters and one singleton (Fig. 3.2c). Results of this study revealed the better resolution power of quantitative characters for grouping coffee genotypes. Similar results were reported in tef by Ayele et al. (1999) and in okra by Martinello et al. (2001). All three approaches separated genotypes with compact growth habits, which are suitable for high population density planting, from those with open branching and vigorous growth habits. Cluster analysis using only quantitative characters as well as qualitative and quantitative characters together discriminated genotypes according to their geographic origin. Montagnon and Bouharmont (1996) classified wild and cultivated coffee genotypes from Ethiopia according to geographic origin (east and west of the Great Rift Valley) using 18 agro-morphological characters. Therefore, classification in this study based on agro-morphological characters was in agreement with previous reports.

The existence of broad morphological variation among genotypes was further substantiated by principal component analysis (Table 3.5), which indicated that the overall diversity observed could not be explained by a few eigenvectors. Examination of the first four principal components with eigenvalues between 2.97 and 6.80 accounted for 82.3% of the total variability of genotypes (Table 3.5). The major contributing characters were fairly well distributed among inter-node lengths, tree height, canopy diameter, number of branches and of leaf, bean and fruit characters. Therefore, almost all characters evaluated contributed to the discrimination of tested genotypes. Characters such as average green bean yield per tree and bean length were important to variations in three out of the first four principal components. Among qualitative characters, traits such as leaf apex shape, plant habit, fruit shape and fruit colour were more diverse compared to overall appearance of the tree and young leaf colour. About 88% of tested genotypes had green young leaves, indicating genotypes with green young leaf colour were more common than those with bronze young leaf colour. Leroy et al. (1993) reported that 72% of the total variation was explained by the first two principal components; the first axis separated drought susceptible genotypes from genotypes with high number of internodes and the second axis was related to height of trees and average inter-node length on the stem of *C. canephora*.

Knowledge of correlations among characters is useful in designing an effective breeding programme for any crop. There are several reasons for using indirect selection. Sometimes the main character is expressed late or measurement of the indirect character is much easier than for the direct character. Moreover, complex plant characters such as yield are quantitatively inherited and influenced by genetic effects, as well as by genotype x environment interaction. Due to these reasons, selections to improve yield directly may be difficult and time consuming especially for perennial crops with a long juvenile period such as coffee. Therefore, identification and use of highly correlated characters are appropriate. In this investigation, average green bean yield per tree had positive and significant association with percentage of bearing primary branches per tree, bean weight, canopy diameter, tree height, trunk diameter, inter-node length on primary branches, bean thickness and length, fruit length, inter-node length on orthotropic branch and petiole length. Selection for any of these characters, which are highly heritable and easy to measure, will help to improve coffee bean yield per tree in this population.

Walyaro and Van der Vossen (1979) reported significant and positive genotypic correlations between bean yield and girth at the base of the main stem (0.64), percentage of bearing primary branches (0.86), radius of the canopy (0.62) and tree height (0.53). Ameha (1982) reported positive and significant correlations between growth and girth diameter (0.84), growth and number of fruits (0.69) and growth and number of nodes on primary branches (0.54). Ameha and Bellachew (1983) reported positive correlations with mean F_1 yield and the following characters; girth (0.88), number of flowers and fruits (0.74), length of the first primary branch (0.49), number of nodes on primary branches (0.52), number of bearing nodes on primary branches (0.59) and number of secondary branches (0.46). Walyaro (1983) reported positive correlations among growth characters, yield characters and between yield and growth characters.

Results of this study showed that except for number of primary branches, almost all characters had positive phenotypic correlations amongest each other. Results of this study were in agreement with findings of previous investigations. Selection for higher bean yield per tree will probably be possible indirectly through selection for higher percentage of bearing primary branches, bean weight, canopy diameter, tree height,

trunk diameter, inter-node length on primary branches, bean thickness, bean length, fruit length, inter-node length on orthotropic branch and petiole length. In agreement to this finding, Van der Vossen (1985) reported that selection efficiency for higher yield is considerably increased by taking into account various growth parameters and components of yield, such as stem girth, canopy radius, percentage of bearing primaries, percentage of bearing nodes and number of berries per node.

Estimates of broad sense heritability for various characters of coffee ranged from 0.38 for bean thickness to 0.94 for bean weight (Table 3.8). Walyaro and Van der Vossen (1979) reported 0.28, 0.65, 0.31, 0.44, 0.48, and 0.81 broad sense heritability for tree height, stem girth, number of primary branches, percentage of bearing primaries, canopy radius and cherry yield (kg/tree) per annum, respectively in 16 arabica coffee genotypes. Ameha (1980) reported 0.55 and 0.44 broad sense heritability in arabica coffee for bean yield during the 1971-1974 and 1975-1978 cropping seasons, respectively. Hulupi et al. (2004) reported a high heritability value (69.9 %) for elephant bean and Van der Vossen and Walyaro (1983) reported quite high heritability (0.97) for coffee berry disease resistance (CBD) in arabica coffee.

Van der Vossen (1985) reported 0.64, 0.70, 0.65, 0.74, 0.60, 0.39, 0.23 and 0.74 broad sense heritability for girth of main stem, tree height, canopy radius, internode length, angle of primaries with main stem, % of bearing primaries, % of bearing nodes and 100 bean weight, respectively. In addition, Montagnon et al. (1998) reported high levels of narrow sense heritability for caffeine content (0.80), fat content (0.74) and bean weight (0.73), intermediate levels for trigonelline (0.38) and chlorogenic acids (0.36) and low levels for sucrose content (0.11) in *C. canephora*. Cilas et al. (2000) reported relatively low broad sense heritability (0.34) but with sufficient genetic variability for wood breakage in *C. canephora*. Results of this study were in agreement with most of the previous findings. Some discrepancies could be attributed to differences in number and type of genotypes tested and number of seasons and locations from which the data for estimation of these parameters were collected.

Assefa et al. (1999) suggested that genetic coefficients of variation together with heritability estimates would give the best picture of genetic advance to be expected from selection. In addition, Johnson et al. (1955) reported that heritability estimates

show only the effectiveness with which selection of genotypes could be based on phenotype but its utility increased when used along with estimates of genetic advance. The combined use of genetic coefficient of variation, heritability and genetic advance information seems vital for effective improvement of a particular trait in a population. Of all the characters evaluated in this study, characters such as number of secondary branches per tree, canopy diameter, tree height, average green bean yield per tree and bean weight had relatively high values of genotypic coefficient of variation, broad sense heritability and genetic advance. Improvement of these characters in the tested genotypes by selection seems relatively easy.

3.6 Conclusions

The present study revealed the presence of high levels of diversity for several agro-morphological characters among evaluated arabica coffee genotypes. These genotypes should be properly conserved and could serve as raw material for the genetic improvement of different characters of the crop through recurrent selection after hybridisation. In addition, most of the characters of coffee had positive correlations amongest each other. This will assist in the combined improvement of these characters by selecting only for highly heritable and easily measurable characters. Some of the examined characters were highly heritable in this investigation as well as in other studies carried out by different researchers from different countries at different times using different genotypes. Therefore, this is also a good opportunity for coffee breeders to select important characters with less environmentally induced errors.

CHAPTER 4

VARATION IN CUP QUALITY AND GREEN BEAN CHARACTERS AMONG ETHIOPIAN ARABICA COFFEE GENOTYPES

4.1 Abstract

Quality, especially liquor or cup quality, determines the relative price as well as the usefulness of a given quantity of coffee. Therefore, the quality of 42 arabica coffee genotypes collected from northwestern and southwestern parts of Ethiopia was assessed in this study. Objectives of the study were to determine the level of variation among these genotypes based on quality characteristics and assess correlations among different coffee quality attributes. Genotypes were evaluated using four cup quality traits (acidity, body, flavour and overall standard of the liquor) and four green bean physical characteristics (bean shape, size, uniformity and weight) by professional coffee tasters at the Addis Ababa Coffee Cup Quality Evaluation and Standardisation Centre, Ethiopia. Considerable variation was observed among genotypes both for cup quality and green bean physical characteristics. Genotypes did not cluster according to collection region. This indicated the presence of coffee genetic diversity for quality characteristics in each region. Among the tested genotypes, AD0691 and AD1691 had both desirable cup quality and green bean physical characteristics. These genotypes could serve as sources of desirable genes both for cup quality and green bean physical characters improvement in arabica coffee. All cup quality traits were positively and significantly correlated and correlation coefficients ranged from 0.69 to 0.93. Flavour showed relatively high correlation with all other cup quality traits. Selection for better flavour would lead to overall cup quality improvement in arabica coffee. Green bean physical characteristics and cup quality traits were positively correlated. However, most of the associations were statistically non-significant. Therefore, green bean physical characteristics would not be good indicators for cup quality improvement or vice-versa. Of the green bean physical characteristics, only bean size and weight showed statistically significant associations. Generally this study showed the presence of variation for coffee quality attributes and statistically significant correlations among cup quality traits.

4.2 Introduction

Coffee has only one value: to give the consumer pleasure and satisfaction through flavour, aroma and desirable physiological and psychological effects (Sivetz, 1963). Therefore coffee quality, especially liquor or cup quality, determines both the relative price and usefulness of a given quantity of coffee (Walyaro, 1983; Roche, 1995; Agwanda et al., 2003). Currently the price of coffee is declining at the international market mainly due to over production (ITC, 2004). Production and supply of coffee with excellent quality seems more crucial than ever before for coffee exporting countries. Consequently, some countries consider assessment of coffee quality as important as disease resistance and productivity in their coffee variety development programmes.

Cup quality is a complex characteristic, which depends on various factors such as genetic composition, environment, agronomic practices, ripeness of the fruit and post-harvest processing (ITC, 1992; Moreno et al., 1995; Chifra et al., 1998; ITC, 2002). Its assessment method is subjective since it is done organoleptically by panels of experienced coffee tasters on the basis of level of acidity, body, flavour and overall standard of the brew (Walyaro, 1983; Van der Vossen, 1985; Owuor, 1988; Moreno et al., 1995; Petracco, 2000; Agwanda et al., 2003). Although cup quality is a complex trait and its assessment method is subjective, Owuor (1988) and Moreno et al. (1995) improved the cup quality of different coffee genotypes with the assistance of professional coffee tasters. Both authors observed close similarity among different coffee tasters in ranking coffee genotypes based on cup quality characteristics. In addition, Roche (1995) reported a consistent performance of cultivars for cup quality attributes between two seasons. Walyaro (1983) observed lack of variation for bean and berry characteristics measured for different year old coffee trees. These research results indicated the possibility of estimating coffee quality characteristics of different genotypes with the assistance of professional coffee tasters and even based on results of one season at any age of the tree.

Limited research results are available on the association and heritability of different cup quality characteristics of coffee. Walyaro (1983) reported positive correlations among all tested cup quality characteristics. The level of association for most of the

characteristics was statistically significant. Walyaro (1983) and Van der Vossen (1985) reported high heritability for the overall standard of the brew. Good cup quality in arabica coffee is a dominant character (Carvalho, 1988). Results of previous studies demonstrated the possibility of selection for cup quality.

A number of investigators reported the presence of cup quality variation among different coffee genotypes. Walyaro (1983), Van der Vossen (1985) and Roche (1995) observed significant differences among different arabica coffee cultivars and crosses for various cup quality attributes. Selvakumar and Sreenivasan (1989) observed coffee cup quality variation ranging from good to excellent among 54 arabica coffee accessions collected from Keffa, Ethiopia. Coffee produced in some parts of Ethiopia, especially from Harrar, Limu and Yirgachefe, is always sold at a premium price both at domestic and international coffee markets because of its distinctive fine quality (Chifra et al., 1998; ITC, 2002). Consumers, farmers and agricultural development extension experts in northwestern Ethiopia reported the presence of cup quality variation among coffee genotypes growing in the area. However, this was not verified by professional coffee tasters. The association of different cup quality and green bean physical characteristics under northwestern Ethiopia condition is unknown. Therefore, this study was conducted in order to (1) determine the cup quality of different coffee genotypes under northwestern Ethiopia condition, (2) assess the level of cup quality and green bean physical characteristics variability among currently grown coffee genotypes, (3) identify and classify genotypes on the basis of quality parameters and (4) estimate the type and magnitude of correlations between different coffee quality characteristics.

4.3 Materials and Methods

4.3.1 Genotypes and description of the trial site

Forty-two *C. arabica* genotypes collected from two coffee producing regions of Ethiopia (Fig. 4.1 and Table 4.1) were used in this investigation. These genotypes were obtained from Adet Agricultural Research Centre of the Amhara Regional Agricultural Research Institute, Ethiopia. Genotypes were phenotypically different and collected from different agro-ecologies. Each genotype was represented by five

trees that were five years old during the course of this study. Genotypes were planted at the germplasm maintenance block of Finoteselam coffee trial site, Ethiopia. Management practices were applied as recommended (Bellachew et al., 1998). The description of the trial site is given in section 3.3.1.

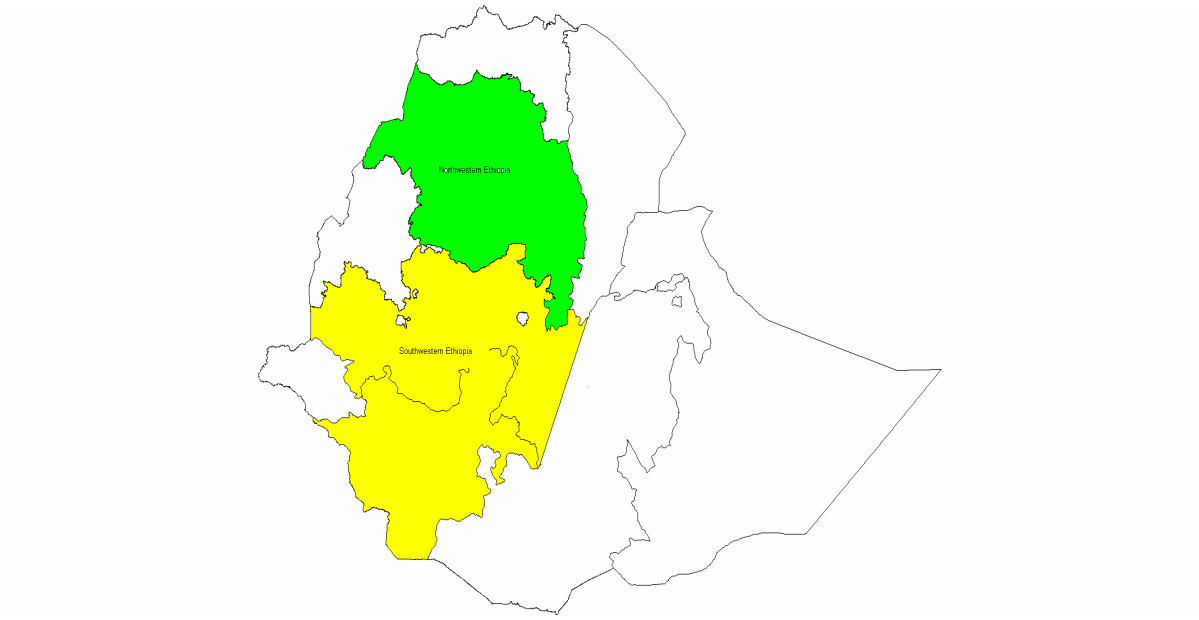


Fig 4.1 Map of Ethiopia depicting collection regions of evaluated coffee genotypes

Table 4.1 List of *C. arabica* genotypes evaluated for cup quality and green bean physical characters variability

Serial no.	Accession no.	Collection region
1	AD0191	Southwestern Ethiopia
2	AD0291	Southwestern Ethiopia
3	AD0391	Southwestern Ethiopia
4	AD0491	Southwestern Ethiopia
5	AD0591	Southwestern Ethiopia
6	AD0691	Southwestern Ethiopia
7	AD0791	Southwestern Ethiopia
8	AD0891	Southwestern Ethiopia
9	AD0991	Southwestern Ethiopia
10	AD1091	Southwestern Ethiopia
11	AD1191	Southwestern Ethiopia
12	AD1291	Northeastern Ethiopia
13	AD1391	Southwestern Ethiopia
14	AD1491	Southwestern Ethiopia
15	AD1591	Southwestern Ethiopia
16	AD1691	Northwestern Ethiopia
17	AD1791	Northwestern Ethiopia
18	AD1891	Northwestern Ethiopia
19	AD1991	Northwestern Ethiopia
20	AD2091	Northwestern Ethiopia
21	AD2191	Northwestern Ethiopia
22	AD2291	Northwestern Ethiopia
23	AD2391	Southwestern Ethiopia
24	AD2491	Northwestern Ethiopia
25	AD2591	Northwestern Ethiopia
26	AD2691	Northwestern Ethiopia
27	AD2791	Northwestern Ethiopia
28	AD2891	Southwestern Ethiopia
29	AD2991	Northwestern Ethiopia
30	AD3091	Southwestern Ethiopia
31	AD3191	Southwestern Ethiopia
32	AD3291	Northwestern Ethiopia
33	AD3391	Southwestern Ethiopia
34	AD3491	Northwestern Ethiopia
35	AD3591	Northwestern Ethiopia
36	AD3691	Northwestern Ethiopia
37	AD3791	Northwestern Ethiopia
38	AD3891	Northwestern Ethiopia
39	AD3991	Northwestern Ethiopia
40	AD4091	Northwestern Ethiopia
41	AD4191	Northwestern Ethiopia
42	AD4291	Northwestern Ethiopia

4.3.2 Sample preparation

During peak harvesting time, only healthy and red-ripe berries were harvested in bulk by hand from five trees of each genotype and processed according to the dry processing method (berries were sun dried on a cemented floor for about three weeks and when fully dried, dehulled by pounding with a pestle and mortar). Mechanically undamaged beans were used both for cup quality and green bean physical characteristics analyses.

4.3.3 Cup quality and green bean physical characteristics evaluation

For each genotype, 600 g dry processed green beans were submitted by code to the Addis Ababa Coffee Cup Quality Evaluation and Standardisation Centre, Ethiopia. Cup quality and green bean physical characteristics of each genotype were evaluated from three samples by three professional coffee tasters (Fig. 4.2). Acidity, body, flavour and overall standard of the brew among cup quality attributes and shape, size, uniformity and weight among green bean physical characteristics were used in this evaluation. Descriptions are given in Table 4.2.

Fig. 4.2 Professional coffee tasters of the Centre for cup quality evaluation

Table 4.2 Evaluated cup quality and green bean physical characters

No.	Character and descriptive value
	Cup quality parameters
1	Acidity: 1 (fair), 2 (medium), 3 (medium pointed)
2	Body: 1 (fair), 2 (medium), 3 (medium to full)
3	Flavour: 1 (fair), 2 (average), 3 (fairly good)
4	Overall standard: 1 (fair), 2 (average), 3 (fairly good)
	Green bean physical characteristics
1	Bean shape: 1 (round), 2 (long)
2	Bean size (screen size): small or 1(<14 mm), medium or 2 (14-16 mm), bold or 3 (> 17 mm)
3	Bean uniformity: 1 (mixed), 2 (uniform)
4	100 Bean weight (g)

4.3.4 Statistical analysis

Relationships among coffee genotypes were assessed using the unweighted pair group method (UPGMA) of cluster analysis. The diversity of seven coffee quality characteristics were estimated using the Shannon-Weaver (1949) diversity index, H` = $-\Sigma_j^i$PilogPi, where Pi is the relative frequency in the $i^{th}$ category of $j^{th}$ trait. Correlation coefficients among different coffee quality attributes were estimated using the Spearman`s method of correlation analysis. One way analysis of variance was performed to observe the presence of variation in average green bean weight between genotypes collected from the southwestern and northwestern parts of Ethiopia. Analysis of variances were not performed for acidity, body, flavour, overall standard, bean shape, bean uniformity and bean size since they are qualitative characters. All data analyses were performed using the Number Cruncher Statistical System, NCSS 2000 (Hintze, 1998).

4.4 Results

Forty-two *C. arabica* genotypes were evaluated for cup quality variation and results are given in Table 4.3 and summarised in Table 4.4. According to overall liquor quality standard, four genotypes namely AD0191, AD0691, AD0791 and AD1691

had fairly good overall liquor quality standards. Most of the genotypes (27) had an average and 11 genotypes were fair in overall liquor quality standard.

Table 4.3 Cup quality and green bean physical characteristics for 42 Ethiopian *C. arabica* genotypes

Variety	Cup quality				Green bean physical characteristics			
	Acidity	Body	Flavour	Overall standard	Bean size	Bean uniformity	Bean shape	100 bean weight (g)
AD0191	Mp	Mf	Fg	Fg	Medium	Uniform	Round	14.2
AD0291	Me	Me	Av	Av	Medium	Uniform	Round	11.7
AD0391	Mp	Mf	Fg	Av	Medium	Uniform	Round	10.6
AD0491	Mp	Me	Av	Av	Small	Uniform	Round	11.5
AD0591	Mp	Mf	Fg	Av	Small	Uniform	Round	11.6
AD0691	Mp	Mf	Fg	Fg	Bold	Uniform	Long	13.9
AD0791	Mp	Mf	Fg	Fg	Medium	Uniform	Round	13.7
AD0891	Me	Fa	Fa	Av	Small	Uniform	Round	9.6
AD0991	Fa	Fa	Fa	Fa	Small	Mixed	Round	9.5
AD1091	Me	Fa	Fa	Fa	Small	Uniform	Round	9.8
AD1191	Fa	Fa	Fa	Fa	Small	Uniform	Round	10.1
AD1291	Fa	Fa	Fa	Fa	Bold	Uniform	Round	17.4
AD1391	Fa	Fa	Fa	Fa	Medium	Mixed	Round	11.9
AD1491	Me	Me	Av	Av	Medium	Mixed	Long	12.5
AD1591	Me	Me	Av	Av	Medium	Uniform	Round	12.6
AD1691	Mp	Mf	Fg	Fg	Bold	Uniform	Round	17.0
AD1791	Fa	Fa	Fa	Fa	Medium	Mixed	Round	12.2
AD1891	Me	Me	Av	Av	Medium	Mixed	Long	14.8
AD1991	Me	Me	Av	Fa	Medium	Mixed	Round	11.3
AD2091	Fa	Fa	Fa	Fa	Medium	Mixed	Long	11.2
AD2191	Me	Mf	Fa	Fa	Bold	Mixed	Round	14.4
AD2291	Fa	Me	Av	Av	Medium	Uniform	Round	13.9
AD2391	Mp	Mf	Fg	Av	Bold	Uniform	Round	14.7
AD2491	Me	Me	Av	Av	Medium	Uniform	Round	12.5
AD2591	Mp	Me	Av	Av	Medium	Mixed	Round	14.2
AD2691	Me	Me	Av	Av	Medium	Uniform	Round	13.9
AD2791	Mp	Me	Av	Av	Bold	Mixed	Round	15.1
AD2891	Mp	Mf	Fg	Av	Medium	Mixed	Round	12.3
AD2991	Mp	Mf	Fg	Av	Medium	Mixed	Round	12.2
AD3091	Mp	Mf	Fg	Av	Medium	Uniform	Round	14.4
AD3191	Me	Me	Av	Av	Medium	Mixed	Round	14.5
AD3291	Mp	Mf	Fg	Av	Medium	Uniform	Round	12.0
AD3391	Mp	Mf	Fg	Av	Small	Mixed	Round	12.3
AD3491	Mp	Mf	Fg	Av	Bold	Uniform	Round	15.6
AD3591	Fa	Fa	Fa	Fa	Medium	Uniform	Round	14.0
AD3691	Mp	Mf	Fg	Av	Bold	Mixed	Round	13.7
AD3791	Mp	Mf	Fg	Av	Bold	Uniform	Round	15.0
AD3891	Mp	Mf	Fg	Av	Bold	Uniform	Round	18.2
AD3991	Me	Fa	Fa	Fa	Medium	Mixed	Long	14.3
AD4091	Mp	Mf	Fg	Fg	Medium	Uniform	Round	12.6
AD4191	Me	Me	Av	Av	Bold	Mixed	Round	15.1
AD4291	Me	Me	Av	Av	Bold	Uniform	Long	16.2

Av = average; Fa = fair; Fg = fairly good; Me = medium; Mf = medium to full; Mp = medium pointed.

Among the 19 genotypes collected from southwestern Ethiopia, three were fairly good, 12 were average and four were fair in overall liquor quality standard. On the other hand, out of 23 genotypes collected from northwestern Ethiopia, one was fairly good, 15 were average and seven were fair in overall liquor quality standard. Relatively more genotypes from southwestern Ethiopia were better in overall liquor quality compared to those from northwestern Ethiopia.

Table 4.4 Number and percentage of coffee genotypes classified in different classes of cup quality traits

Traits	Classes	Northwestern Ethiopia (23)		Southwestern Ethiopia (19)		Both regions (42)	
		no.	%	no.	%	no.	%
Overall standard							
	Fairly good	1	4.4	3	15.8	4	9.5
	Average	15	65.2	12	63.2	27	64.3
	Fair	7	30.4	4	21.0	11	26.2
Acidity							
	Medium pointed	7	30.4	6	31.6	13	30.9
	Medium	10	43.5	8	42.1	18	42.9
	Fair	6	26.1	5	26.3	11	26.2
Body							
	Medium to full	2	8.7	3	15.8	5	11.9
	Medium	15	65.2	11	57.9	26	61.9
	Fair	6	26.1	5	26.3	11	26.2
Flavour							
	Fairly good	2	8.7	3	15.8	5	11.9
	Average	14	60.9	11	57.9	25	59.5
	Fair	7	30.4	5	26.3	12	28.6

no = number of genotypes

Acidity indicates the bitter or acidic balance and the presence of a sweet caramelic after taste. Traditionally coffee graders consider acidity as a desirable attribute. In

terms of acidity of liquor, 13 genotypes were scored as medium pointed while 18 and 11 genotypes were as medium and fair in liquor acidity, respectively (Table 4.4). Of the 19 genotypes collected from southwestern Ethiopia, six, were scored as medium pointed, eight as medium and five as fair in liquor acidity. Similarly, out of 23 genotypes collected from northwestern Ethiopia, seven were scored as medium pointed, 10 as medium and six as fair in liquor acidity. Percentages of coffee genotypes of the two regions assigned in different liquor acidity classes were comparable.

Body assesses the beverage mouth feel or the property linked with density and viscosity of the brew. Of the total genotypes evaluated in this study, five were scored as medium to full, 26 as medium and 11 as fair in liquor body (Table 4.4). Among 19 genotypes collected from southwestern Ethiopia, three were scored as medium to full, 11 as medium and five as fair in their beverage body. Out of 23 genotypes from northwestern Ethiopia, two genotypes had medium to full, 15 medium and six fair liquor body.

Flavour indicates fragrance of the liquor either by direct inhaling of the vapours arising from the cup or nasal perception of the volatile substances evolving in the mouth. Out of the 42 genotypes evaluated in this study, five were fairly good, 25 were average and 12 were fair in their beverage flavour (Table 4.4). From 19 genotypes collected in southwestern Ethiopia, three genotypes were fairly good, 11 were average and five were fair in beverage flavour. Out of 23 genotypes collected from northwestern Ethiopia, two genotypes were fairly good, 14 were average and seven were fair in their beverage flavour.

In addition to cup quality, these genotypes were evaluated for different green bean physical characteristics such as bean shape, size, uniformity and weight and results are given in Table 4.5. Bean size is one of the most important green bean physical characteristics, which plays a significant role in coffee marketing. In this study, genotypes with bold, medium and small sized beans were identified. Of all the genotypes evaluated, 12 genotypes had bold, 23 medium and seven small sized beans. Among 19 coffee genotypes collected from southwestern Ethiopia two genotypes had

bold, 10 medium and seven small sized beans. Coffee genotypes collected from northwestern Ethiopia were characterised by bold and medium size beans.

Table 4.5 Average 100 bean weight and percentage *C. arabica* genotypes classified into different classes of green bean physical characters

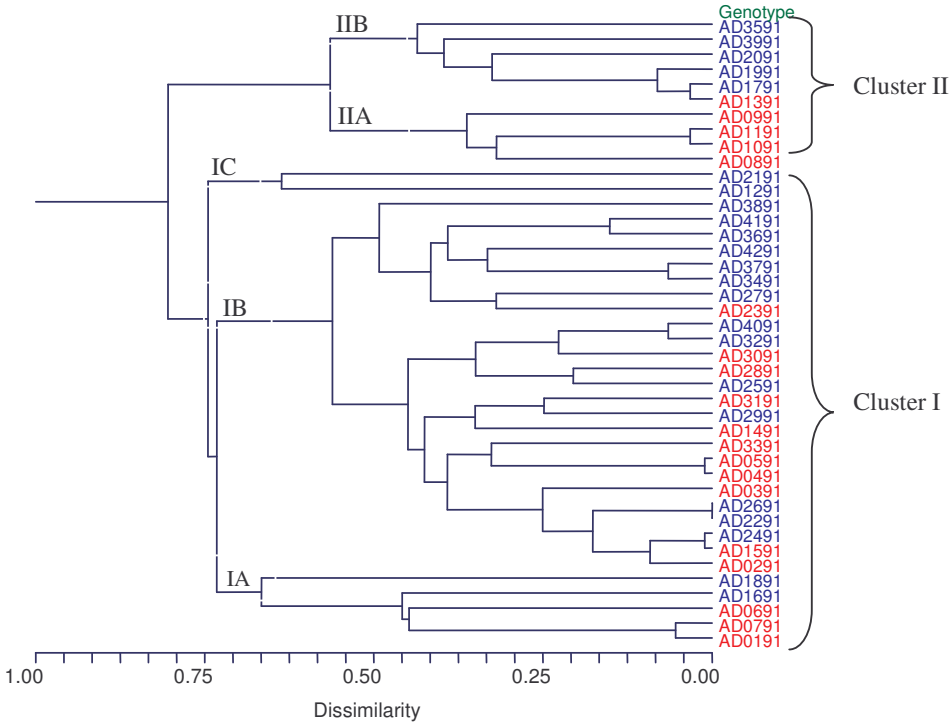
Regions	100 bean weight (g)	Bean size			Bean uniformity		Bean shape	
		Bold	Medium	Small	Uniform	Mixed	Long	Round
Northwestern	14.2	43.5	56.5	0.0	52.2	47.8	17.4	82.6
Southwestern	12.2	10.5	52.6	36.9	68.4	31.6	10.5	89.5
Both	13.3	28.6	54.8	16.6	59.5	40.5	14.3	85.7

Bean uniformity is another important green bean physical characteristic of coffee. It affects the roasting process and subsequently reduces cup quality. Of all the genotypes evaluated, 25 had uniform while 17 had mixed beans. The number of genotypes with uniform beans is relatively high. Genotypes collected from the southwestern part of the country had more uniform beans compared to those from the northwestern part. Evaluation for bean shape revealed that six genotypes had long beans while 36 had round beans. The proportion of genotypes with long beans was almost similar and low in both regions.

Another important green bean physical character considered for coffee quality evaluation was bean weight. Genotypes were diverse in average 100 bean weight, ranging from 9.5 to 18.2 g. Genotypes such as AD0891, AD0991, AD1091 and AD1191 had light beans while AD1291, AD1691, AD3891 and AD4291 had relatively heavy beans. T-test performed using data of this trait indicated that coffee genotypes collected from northwestern Ethiopia had significantly heavier beans compared to those from the southwestern part of the country. Relatively more genotypes from the northwestern Ethiopia had bold size and heavy beans compared to those from southwestern Ethiopia. More genotypes from northwestern Ethiopia had mixed beans, which is undesirable.

The genetic relationship of 42 coffee genotypes based on cup quality and green bean physical characteristics was assessed using the UPGMA method of cluster analysis.

Genotypes grouped into two main clusters (Fig. 4.3). The first main cluster comprised of 32 genotypes and bifurcated into three sub-clusters. The first sub-cluster (IA) contained five genotypes namely AD0191, AD0691, AD0791, AD1691 and AD1891. These genotypes were characterised by good cup quality and desirable green bean physical characteristics. The second sub-cluster (IB) consisted of 25 coffee genotypes, which were characterised by average cup quality. The third sub-cluster (IC) consisted of two coffee genotypes namely, AD1291 and AD2191 which were characterised by poor cup quality but with desirable green bean physical characteristics. The second main cluster consisted of 10 genotypes characterised by poor cup quality and small to medium size beans. The first sub-cluster of the second main cluster (IIA) consisted of four genotypes namely, AD0891, AD1091, AD1191 and AD0991. These genotypes were characterised by poor cup quality as well as small sized beans. The second sub-cluster of the second main cluster (IIB) had six genotypes which were characterised by poor cup quality and medium sized beans. Genotypes were not clustered according to geographic origin (collection region) or similarity of morphological characteristics.



Blue for northwestern Ethiopia and red for southwestern Ethiopia genotypes

Fig. 4.3 Dendrogram of 42 *C. arabica* genotypes constructed by cluster analysis using cup quality and green bean physical characters

Shannon-Weaver diversity indices (H`) were calculated to compare the diversity of seven coffee quality characteristics. A higher H` value indicates presence of diversity for the trait among the tested genotypes. Results of this study indicated the presence of diversity for all of the characteristics assessed. Shannon-Weaver diversity indexes ranged from 0.592 to 0980, bean shape having the lowest and liquor acidity the highest value (Table 4.6). Genotypes from the southwestern part of the country were more diverse in overall liquor quality, body and flavour compared to those genotypes collected from the northwestern part of the country. On the other hand, genotypes collected from the northwestern part were more diverse in bean uniformity and shape than those from the southwestern part. This result showed the presence of greater diversity among genotypes from the southwestern part compared to those collected from the northwestern part of the country.

Table 4.6 Shannon-Weaver diversity indices for seven coffee quality characters

Characters	Regions where genotypes collected		
	Northwestern Ethiopia	Southwestern Ethiopia	Both
Overall standard	0.708	0.828	0.782
Acidity	0.978	0.983	0.980
Body	0.766	0.873	0.820
Flavour	0.798	0.873	0.838
Bean size	0.988	0.858	0.897
Bean uniformity	0.999	0.900	0.974
Bean shape	0.667	0.485	0.592

Correlations of four cup quality attributes and four green bean physical characteristics were assessed. All cup quality characteristics were positively and significantly correlated among themselves (Table 4.7). All green bean physical characteristics had positive associations with all cup quality attributes. Correlations were significant between bean size and body as well as between bean uniformity and overall standard of the liquor. Of all the green bean physical characteristics, only bean size and bean weight were significantly correlated. Round beans were generally small in size and light in weight.

Table 4.7 Correlation coefficients among different coffee quality characters

	Body	Flavour	Overall standard	Bean size	Bean uniformity	Bean shape	Bean weight
Acidity	0.86**	0.88**	0.71**	0.22	0.19	0.18	0.22
Body		0.93**	0.69**	0.36*	0.21	0.19	0.29
Flavour			0.78**	0.28	0.27	0.16	0.25
Overall standard				0.22	0.36*	0.02	0.29
Bean size					0.01	-0.12	0.79**
Bean uniformity						0.22	0.07
Bean shape							-0.12

* and ** indicate significant correlations at p≤0.05 and 0.01, respectively.

4.5 Discussion

Results of this study indicated the presence of variation among evaluated coffee genotypes both for cup quality and green bean physical characteristics. Selvakumar and Sreenivasan (1989) reported cup quality variation ranging from good to excellent among 54 arabica coffee accessions collected from southwestern (Keffa province) parts of Ethiopia. Silvarolla et al. (2000), Ky et al. (2001) and Silvarolla et al. (2004) reported significant variation among *C. arabica* accessions collected from different parts of Ethiopia for bean caffeine, chlorogenic acids, sucrose and trigonelline contents. Montagnon and Bouharmont (1996) and Anzueto et al. (2001) reported diversity among Ethiopian coffee genotypes for different agro-morphological characteristics. Therefore, results of the present study further confirmed the presence of diversity for different characteristics of arabica coffee in the Ethiopian arabica coffee gene pool. Coffee genotypes collected from the southwestern part were relatively more diverse in coffee quality attributes compared to genotypes from the northwestern part of the country. This is in agreement with previous findings that indicated southwestern Ethiopia as a centre of origin and diversity for this crop (Sylvain, 1955; Steiger et al., 2002). Therefore, coffee germplasm collection and evaluation efforts should primarily focus on the southwestern part of the country. However, coffee landraces in northwestern Ethiopia are currently threatened by environmental degradation and replacement by high yielding as well as coffee berry disease resistant coffee cultivars. Therefore, emphasis should also be given for the collection and conservation of coffee genotypes currently grown in northwestern Ethiopia.

Cluster analysis based on coffee quality traits grouped 42 coffee genotypes into two main clusters. Interestingly, genotypes were not clustered according to area of collection. This could be ascribed to the unrestricted movement of coffee seed from region to region by man as well as wild animals. Moreover, in each region the agro-ecology is diverse and might contribute to the occurrence of different coffee genotypes through natural selection. Out of 42 coffee genotypes evaluated, only two genotypes, AD2291 and AD2691, were identical both in cup quality and green bean physical characteristics. This indicated the presence of adequate diversity within the Ethiopian arabica coffee gene pool for cup quality as well as green bean physical

characteristics. These genetic resources should be properly conserved in order to utilise them for genetic improvement of coffee quality in the future.

Most genotypes were average in cup quality. However, five genotypes, AD0191, AD0691, AD0791, AD1691 and AD1891 showed fairly good cup quality. The first three were collected from the southwestern and the last two from the northwestern part of the country. This indicated the availability of genetic resources for cup quality improvement in both regions. The possibility of genetic improvement for berry size and cup quality in arabica coffee was demonstrated and reported by different researchers (Walyaro, 1983; Owuor, 1988; Moreno et al., 1995; Agwanda et al., 2003). These genotypes could be used as sources of desirable genes for cup quality improvement in arabica coffee in the future. The number of genotypes with good cup quality was relatively low. This could be since trees were young (five years old) and beans used for the evaluation were dry processed. According to Wellman (1961) coffee from young trees produce mild and thin liquor, while beans from 15 to 20 year old trees provide liquor characterised with good flavour, acidity and body. In addition, cup quality is affected by the processing method. According to Clifford (1985) wet processed arabica was aromatic with a fine acidity and some astringency, while dry processed arabica was less aromatic and less acidic but with greater body.

Bean physical characteristics such as bean size, density and shape are unified criteria for conducting coffee business within the international market (ITC, 1992; Agwanda et al., 2003). Uniform, bold and heavy coffee beans are the most preferred and are usually priced high. Identification of gene sources for desirable green bean physical characteristics is crucial for the development of coffee varieties with desirable green bean physical characteristics. Of the 42 genotypes evaluated, 25 genotypes had uniform beans while the remaining 17 genotypes had mixed beans. The number of genotypes with uniform beans was high and could be attributed to the high self-pollination (over 95%) habit of the crop (Silvarolla et al., 2004). In terms of bean shape, most genotypes had round beans. However, six genotypes namely AD0691, AD1491, AD1891, AD2091, AD3991 and AD4291 had long beans. Genotypes were diverse for average bean weight. The average 100 bean weight ranged from 9.5 to 18.2 g. Coste (1992) reported average 100 bean weight ranging from 15 to 20 g among arabica coffee cultivars. The differences between results of these two studies

could be attributed to differences in genotypes as well as locations where these two studies were conducted. Generally the wide variation among genotypes for different green bean physical characteristics indicated the presence of an opportunity for genetic improvement of green bean physical characteristics using these genotypes. Of all the genotypes evaluated, AD0691, AD1291, AD1691, AD2391, AD3491, AD3791, AD3891 and AD4291 had bold, uniform and relatively heavy beans. On top of this AD0691 and AD1691 had desirable cup qualities. Therefore, these two genotypes could serve as sources of desirable genes both for cup quality and green bean physical characteristics improvement in arabica coffee.

Knowledge of correlations among different characteristics is fundamental to design an effective breeding programme for any crop, especially for perennial crops like coffee. Some characteristics, for example cup quality, are economically important but its evaluation method is subjective and difficult. Hence identification of characteristics, which are highly correlated with cup quality and are easily measurable seem important for possible indirect selection. Results of this study showed positive and significant correlations among all cup quality attributes such as acidity, body, flavour and overall standard of the liquor. This is in agreement with previous findings (Walyaro, 1983). Selection using one of these traits will enable the improvement of other cup quality traits. Van der Vossen (1985) recommended overall standard as the best cup quality selection trait due to its high heritability. On the other hand, based on correlation, repeatability and sensitivity analyses, Agwanda (1999) recommended flavour rating as the best selection criterion for genetic improvement of cup quality in arabica coffee. In this study, flavour showed positive, statistically highly significant and relatively high correlation coefficients with all other cup quality characteristics. Therefore, selection using liquor flavour alone would enable all rounded cup quality improvement in arabica coffee.

The association between green bean physical characteristics and cup quality attributes were assessed during this study. All green bean physical characteristics (shape, size, uniformity and weight) showed positive correlations with all cup quality traits (acidity, body, flavour and overall standard). Correlations were statistically significant between bean size and body as well as between bean uniformity and overall standard of the brew. Therefore, green bean physical characteristics are not good indicators for

cup quality improvement and vice-versa. Similar results were reported by Roche (1995) and Agwanda et al. (2003). Correlations among most green bean physical characteristics were statistically non-significant. Simultaneous selection for green bean physical characteristics seems possible since most of them were positively correlated.

4. 6 Conclusions

Results indicated that genotypes were diverse both for cup quality and green bean physical characteristics. There is therefore an opportunity to select genotypes with desirable green bean physical characteristics and cup quality. Of all the genotypes, AD0691 and AD1691 had desirable cup quality and green bean physical characteristics. Genotypes collected from southwestern Ethiopia were more diverse for coffee quality attributes compared to those collected from northwestern Ethiopia. Coffee germplasm collection, conservation and evaluation attempts should focus more on the southwestern than the northwestern part of the country. The correlation coefficients of all cup quality traits were positive and statistically significant. Flavour had relatively high correlation coefficients with all other cup quality traits. Selection for all rounded cup quality improvement can be achieved using flavour rating alone. Associations between most cup quality and green bean physical characteristics were statistically non-significant. Green bean physical characteristics are therefore not useful for cup quality indirect selection. However, simultaneous selection for both traits seems possible since they were positively correlated. Correlations among most green bean physical characteristics were statistically non-significant. Hence, selection made for one green bean physical characteristic will not cause a significant effect on the other trait.

CHAPTER 5

GREEN BEAN BIOCHEMICAL COMPOSITION VARIATION AMONG ETHIOPIAN ARABICA COFFEE GENOTYPES

5.1 Abstract

The presence of green bean biochemical composition variability among coffee genotypes currently grown in Ethiopia is unknown. Therefore 42 arabica coffee genotypes collected from northwestern and southwestern parts of Ethiopia were evaluated for green bean biochemical composition variation. Objectives of the study were to (1) determine green bean caffeine, chlorogenic acids, sucrose and trigonelline contents of these genotypes, (2) evaluate the level of green bean biochemical composition variability among genotypes within and between regions and (3) assess the presence of correlations among biochemical compounds, cup quality and green bean physical characteristics. Genotypes were significantly different for green bean caffeine, chlorogenic acids, sucrose and trigonelline contents. Genotypes from northwestern Ethiopia were more diverse in green bean caffeine, chlorogenic acids and sucrose contents compared to those from southwestern Ethiopia. Coffee genotypes collected from southwestern Ethiopia had relatively higher green bean caffeine, chlorogenic acids, sucrose and trigonelline contents compared to those from northwestern Ethiopia. Genotypes did not cluster according to collection region. This indicated the presence of biochemical diversity among genotypes within a region as well as the presence of close similarity between some coffee genotypes collected from the two regions. Among green bean biochemical compounds, caffeine showed statistically significant and negative correlations with desirable cup quality attributes. This indicated the possibility of simultaneous selection for low caffeine content and better cup quality. Sucrose had positive correlations with better cup quality attributes and most desirable green bean physical characteristics. This also indicated the possibility of simultaneous selection for high green bean sucrose content, better cup quality and desirable green bean physical characteristics. Generally, genotypes were diverse in green bean biochemical composition and could serve as sources of genes to develop varieties with desirable biochemical composition in the future.

5.2 Introduction

Biochemical compounds are metabolic products and confer adaptive properties to plants. They participate in resistance to diseases as well as pests and give a characteristic odour or taste to edible plants (Perez de la Vega, 1994). Therefore, biochemical variability assessment seems essential in all crop species in order to develop varieties with desirable quality, disease and pest resistance and adaptation to different environmental conditions.

Caffeine, chlorogenic acids, sucrose and trigonelline were used for characterisation of coffee species as well as varieties within a species (Clifford et al., 1989; Rakotomalala et al., 1992; Bicchi et al., 1995; Ky et al., 2001). These biochemical compounds are important in beverage quality since they are aroma precursors. For example, sucrose and trigonelline give rise to appreciated flavour products, including furans, pyrazine, alkyl-pyridines and pyrroles (Clifford, 1985; Ky et al., 2001). On the other hand, chlorogenic acids and caffeine increase bitterness, the former after degradation into phenol derivatives and the latter without any degradation (Ky et al., 2001). These biochemical compounds also have some other uses. For instance, chlorogenic acids inhibit indole acetic acid oxidase that subsequently controls germination and cell growth. It was also shown to have allelopathic properties (Aerts and Baumann, 1994) and participates in defence mechanisms against phytopathogens (Ky et al., 1999). Knowledge of green bean biochemical composition variability seems vital to develop coffee varieties with desirable cup quality, disease resistance and adaptation to different environmental conditions.

Caffeine was responsible for the discovery of coffee as well as the spread of the coffee drinking habit throughout the world. The practice of decaffeination was started as early as 1905 by the Hag Company in Bremen, Germany (Clarke, 1985). Currently the demand for decaffeinated coffee is increasing and constitutes 10% of the world coffee consumption (Silvarolla et al., 2004). Manufacturers of instant coffee have devised processes for artificial removal of caffeine from coffee. But the process is expensive and removes some aroma precursors such as sugar, oil and phenolic compounds (Clarke, 1985; Silvarolla et al., 2000). Previous studies indicated the presence of caffeine content variability among arabica coffee accessions (Silvarolla et

al., 2000; 2004; Ky et al., 2001). Assessment of caffeine content variability seems vital to develop naturally decaffeinated coffee varieties for the growing decaffeinated coffee market.

Coffee is an important commodity in the world economy, accounting for trade worth on average US$ 9.7 billion annually (ITC, 2002). For the last few years the price of coffee was steadily declining mainly due to overproduction and undiversified outputs (Viniegra-Gonzalez, 2000). Diversification of coffee outputs is required to buffer these crises. Assessment of biochemical variability will be of paramount importance for coffee output diversification. Generally, green bean biochemical composition analysis seems essential for characterisation of varieties, development of coffee varieties with desirable quality, disease and pest resistance, adaptability to different environmental conditions and for diversification of coffee outputs. The presence of biochemical variability among coffee genotypes currently grown in Ethiopia is unknown. Therefore, this study was conducted to (1) determine the green bean caffeine, chlorogenic acids, sucrose and trigonelline contents of 42 arabica coffee genotypes collected from northwestern and southwestern parts of Ethiopia, (2) evaluate the level of biochemical variability among genotypes within and between regions and (3) assess the presence of correlations among green bean caffeine, chlorogenic acids, sucrose and trigonelline contents and with cup quality as well as green bean physical characteristics.

5.3 Materials and Methods

5.3.1 Genotypes and sample preparation

Forty two arabica coffee genotypes collected from the northwestern and southwestern parts of Ethiopia were used for this study. Descriptions of the trial site and genotypes are given in Section 4.3.1 and Table 4.1. Samples were prepared as follows: healthy and red-ripe berries were harvested in bulk by hand from five trees of each genotype. Berries were processed according to the dry processing method (berries were sun dried on a cemented floor for about three weeks and when fully dried, were dehulled by pounding with a pestle and mortar). From these dry processed beans, 300 mechanically undamaged beans per genotype were selected, frozen in liquid nitrogen

and ground to a fine powder using a coffee grinder. The powder was divided into six samples, three to estimate dry matter content and three for extraction and determination of caffeine, chlorogenic acids, sucrose and trigonelline contents. Three samples were extracted and analysed to determine the value of each biochemical compound of each genotype. Dry matter was determined by oven drying three samples per genotype overnight at 105 ^{o}C (Clifford, 1985; Baumann et al., 1998).

5.3.2 Extraction and purification

Caffeine and trigonelline were extracted and purified according to the method of Ky et al. (2001) with some modifications. Fifty mg coffee powder, 500 mg magnesium oxide and 25 ml double distilled water were mixed, followed by heating at 121 ^{o}C in an autoclave for 20 min. Extracts were filtered (0.45 μm) and used for caffeine and trigonelline determination using High Performance Liquid Chromatography (HPLC).

Sucrose was extracted and purified according to the method of Ky et al. (2000). Fifty mg coffee powder and 25 ml double distilled water were mixed and boiled at 60 ^{o}C for 15 min. Tubes were hand shaken at 5 min intervals. Colloidal materials present in an aqueous extract were precipitated by adding 0.5 ml Carrez solution I (Carrez solution I was prepared by dissolving 21.9 g of crystallised zinc acetate and 3 ml of glacial acetic acid in 100 ml double distilled water) and 0.5 ml Carrez solution II (Carrez solution II was prepared by dissolving 10.6 g potassium hexacyanoferrate (Fe^{+2}) in 100 ml double distilled water). The volume of the solution was increased to 50 ml by adding double distilled water, filtered (0.45 μm) and used for sucrose determination using HPLC. Chlorogenic acids were extracted and purified following the method of Ky et al. (1997). Two hundred and fifty mg green bean powder and 100 ml of a methanol: water (70/30 v/v) mixture and 0.5% Na_2SO_3 were mixed and shaken overnight in darkness at 125 rpm at 4 ^{o}C. Then 0.5 ml Carrez solution I and 0.5 ml Carrez solution II were added to precipitate colloidal materials present in the solution. Extracts were filtered (0.45 μm) and used for chlorogenic acids analyses using HPLC.

5.3.3 HPLC analysis

Caffeine, chlorogenic acids and trigonelline chromatography were carried out on a Hewlett Packard system consisting of a Quaternary pump, auto-sampler, a Shimadzu SPD 10A UV-vis detector, a C18 pre-column and a 250 x 4.6 mm Phenomenex Luna 18(2) column with a 5 μm pore size. Sucrose was measured using a Waters Breeze system equipped with Differential Refractive Index Detector (Waters Corp. Milford, Massachusetts, USA) and a Waters SUGARPACK1 300 x 7.8 mm column.

Caffeine and trigonelline contents were measured according to the elution programme described by Barre et al. (1998) with some modifications. Two solvents were used for co-elution of caffeine and trigonelline: solvent A consisted of 2 mM phosphoric acid, pH 2.7 containing 5% methanol. Solvent B consisted of methanol containing 5% of 2 mM phosphoric acid, pH 3.9. These two solvents were filtered (0.45 μm) and degassed before use. Samples and standards (10 μl) were analysed at room temperature using the following elution program: A-B mixture (65/35) for 8 min, 5 min to reach 100% of solvent B by linear gradient, 5 min with 100% of solvent B and 5 min to return to initial condition (65/35). The flow rate was 1 ml/min and samples were analysed randomly. Detection was at 266 nm and quantifications according to caffeine and trigonelline standards from Sigma Chemical Co.

Chlorogenic acids were measured according to the elution programme described by Ky et al. (1999). The same solvents A and B as described for caffeine and trigonelline were used according to the following elution program: A-B mixture (75/25) to pure solvent B in 45 min of linear gradient at a flow rate of 1 ml/min. Samples and standards (10 μl) were analysed at room temperature. Detection was at 325 nm and quantification was according to a chlorogenic acid standard from Sigma Chemical Co. The mobile phase for sucrose measurement was de-ionised water at 0.5 ml/min flow rate at 84 ^{o}C. Twenty μl of each sample or standard were automatically injected into the system. Quantification was according to a known sucrose standard from Sigma Chemical Co.

5.3.4 Statistical analysis

Three samples were analysed per genotype to determine the quantity of each biochemical compound. All statistical analyses were performed using the mean of these three samples. The biochemical variability among genotypes within and between regions was assessed using descriptive statistics such as mean and standard deviation. Critical differences (CD) at 0.05 and 0.01 probability levels were computed to test the presence of statistically significant differences between any two genotypes. The same mean value was used for correlation analysis and to classify genotypes into similar groups using the UPGMA method of cluster analysis. Principal component analysis was performed to assess the contribution of each biochemical compound in grouping genotypes. Correlation, cluster and principal component analyses were performed using the Number Cruncher Statistical System, NCSS 2000 (Hintze, 1998).

5.4 Results

Genotypes were significantly different for green bean caffeine content. Green bean caffeine content ranged from 0.91 to 1.32% with an average of 1.10% in dry matter basis (dmb) (Table 5.1). Among all genotypes, AD2691 and AD2291 had the lowest and highest green bean caffeine contents, respectively. The level of caffeine content variability was relatively low compared to that of the other biochemical compounds (Table 5.1).

Chlorogenic acids were the second highest biochemical compound in coffee green beans and the second most variable biochemical compound among genotypes. Values varied from 2.34 to 4.67% dmb and genotypes were significantly different (Table 5.1). Among all genotypes, AD3991 and AD1691 had the lowest and highest green bean chlorogenic acids contents, respectively.

Table 5.1 Green bean caffeine, chlorogenic acids, sucrose and trigonelline contents (% dmb) for 42 Ethiopian *C. arabica* genotypes

No.	Genotype	Caffeine	Trigonelline	Chlorogenic acids	Sucrose
1	AD0191	1.09	1.41	3.62	6.78
2	AD0291	0.93	1.35	3.18	7.66
3	AD0391	1.06	1.24	3.58	6.78
4	AD0491	1.11	1.38	3.21	7.23
5	AD0591	0.98	1.39	3.68	6.63
6	AD0691	1.01	1.17	3.78	6.23
7	AD0791	1.05	1.23	3.05	7.79
8	AD0891	1.25	1.63	3.44	7.13
9	AD0991	1.13	1.60	4.09	7.74
10	AD1091	1.24	1.28	4.41	6.89
11	AD1191	1.23	1.52	4.16	7.77
12	AD1391	1.18	1.41	3.55	7.24
13	AD1491	1.14	1.27	3.85	7.45
14	AD1591	1.27	1.35	3.90	7.59
15	AD2391	1.17	1.23	3.49	8.61
16	AD2891	0.97	1.16	3.10	8.04
17	AD3091	1.12	1.38	2.68	8.30
18	AD3191	1.16	1.28	2.77	5.48
19	AD3391	1.10	1.71	3.13	8.67
20	AD1291	1.23	1.48	3.28	7.72
21	AD1691	1.23	1.22	4.67	8.98
22	AD1791	1.10	1.25	2.85	6.96
23	AD1891	1.19	1.38	3.14	6.59
24	AD1991	1.23	1.34	4.03	7.40
25	AD2091	1.08	1.39	2.78	6.28
26	AD2191	1.12	1.41	2.73	5.69
27	AD2291	1.32	1.11	2.72	7.32
28	AD2491	0.94	1.40	2.99	7.16
29	AD2591	1.01	1.31	2.82	6.75
30	AD2691	0.91	1.31	3.68	7.22
31	AD2791	0.96	1.36	3.09	5.30
32	AD2991	1.13	1.43	3.52	6.73
33	AD3291	1.04	1.34	2.70	7.81
34	AD3491	1.08	1.04	3.77	7.30
35	AD3591	1.02	1.20	2.62	6.63
36	AD3691	1.07	1.28	3.50	6.48
37	AD3791	1.08	1.22	2.36	8.09
38	AD3891	1.02	1.31	2.78	8.36
39	AD3991	1.11	1.28	2.34	6.21
40	AD4091	1.06	1.39	3.34	7.68
41	AD4191	1.08	1.29	3.04	7.29
42	AD4291	1.11	1.20	3.61	6.93
Mean		1.10	1.33	3.31	7.21
Standard deviation		0.098	0.132	0.545	0.824
Minimum		0.91	1.04	2.34	5.30
Maximum		1.32	1.71	4.67	8.98
CD at 0.05		0.20	0.27	1.10	1.66
CD at 0.01		0.26	0.36	1.47	2.23

CD = critical difference; dmb = dry mass basis

Sucrose was the highest biochemical compound in coffee green bean with an average of 7.21% dmb and the most variable biochemical compound among evaluated genotypes (Table 5.1). Green bean sucrose content ranged from 5.30 to 8.98% dmb and genotypes were significantly different. Of all the genotypes, AD2791 had the lowest and AD1691 the highest green bean sucrose contents. Green bean trigonelline content ranged from 1.04 to 1.71% dmb with an average of 1.33% dmb. The lowest and highest trigonelline values were recorded from genotypes AD3491 and AD3391, respectively (Table 5.1). Statistically significant differences were observed among genotypes for green bean trigonelline content.

The levels of variability for green bean caffeine, chlorogenic acids, sucrose and trigonelline contents among coffee genotypes within as well as between regions were assessed and results are given in Table 5.2. Among 23 coffee genotypes collected from northwestern Ethiopia, green bean caffeine content ranged from 0.91 to 1.32% dmb with an average of 1.09% dmb. Green bean caffeine content of 19 coffee genotypes collected from southwestern Ethiopia ranged from 0.93 to 1.27% dmb with an average of 1.12% dmb (Table 5.2). Caffeine content variability was relatively high among coffee genotypes in northwestern Ethiopia compared to those genotypes from southwestern Ethiopia. On average, genotypes from southwestern Ethiopia had relatively higher green bean caffeine content compared to those from northwestern Ethiopia.

The green bean chlorogenic acids content of coffee genotypes from northwestern Ethiopia ranged from 2.34 to 4.67% dmb with an average of 3.15% dmb. Nineteen coffee genotypes collected from southwestern Ethiopia had green bean chlorogenic acids contents ranging from 2.68 to 4.41% dmb with an average of 3.51% dmb (Table 5.2). The level of chlorogenic acids variability was relatively higher among coffee genotypes collected from northwestern Ethiopia compared to those from southwestern Ethiopia. On average, coffee genotypes from southwestern Ethiopia had higher levels of chlorogenic acids than those from northwestern Ethiopia.

The sucrose content of coffee genotypes collected from northwestern Ethiopia ranged from 5.30 to 8.98% dmb with an average of 7.08% dmb, while that of the southwestern Ethiopia genotypes ranged from 6.23 to 8.67% dmb with an average of

7.37% dmb (Table 5.2). Genotypes from northwestern Ethiopia were more variable in green bean sucrose content compared to genotypes from southwestern Ethiopia. On average, genotypes from southwestern Ethiopia had relatively higher green bean sucrose contents.

Table 5.2 Green bean caffeine, chlorogenic acids, sucrose and trigonelline contents variability among *C. arabica* genotypes within and between regions

Traits	Minimum		Maximum		Mean	Standard deviation
	Value	Genotype	Value	Genotype		
Northwestern Ethiopia						
Caffeine	0.91	AD2691	1.32	AD2291	1.09	0.099
Trigonelline	1.04	AD3491	1.48	AD1291	1.30	0.104
Sucrose	5.30	AD2791	8.98	AD1691	7.08	0.837
Chlorogenic acids	2.34	AD3991	4.67	AD1691	3.15	0.561
Southwestern Ethiopia						
Caffeine	0.93	AD0291	1.27	AD1591	1.12	0.098
Trigonelline	1.16	AD2891	1.71	AD3391	1.37	0.155
Sucrose	6.23	AD0691	8.67	AD3391	7.37	0.803
Chlorogenic acids	2.68	AD3091	4.41	AD1091	3.51	0.466
Both regions						
Caffeine	0.91	AD2691	1.32	AD2291	1.10	0.098
Trigonelline	1.04	AD3491	1.71	AD3391	1.33	0.132
Sucrose	5.30	AD2791	8.98	AD1691	7.21	0.824
Chlorogenic acids	2.34	AD3991	4.67	AD1691	3.31	0.545

The trigonelline content of coffee genotypes collected from northwestern Ethiopia ranged from 1.04 to 1.48% dmb with an average of 1.30% dmb (Table 5.2), while it ranged from 1.16 to 1.71% dmb with an average of 1.37% dmb among genotypes from southwestern Ethiopia. Unlike other biochemical compounds, green bean trigonelline content variability was higher among coffee genotypes collected from southwestern Ethiopia compared to those from northwestern Ethiopia. Moreover, genotypes from southwestern Ethiopia had relatively higher trigonelline content.

Correlations among green bean caffeine, chlorogenic acids, sucrose and trigonelline contents as well as with cup quality and green bean physical characteristics were assessed and results are given in Table 5.3. Green bean caffeine content showed negative and statistically significant correlations with all cup quality attributes. Its correlations with green bean physical characteristics were also negative but statistically non-significant. Green bean caffeine content showed positive correlations with green bean sucrose, trigonelline and chlorogenic acids contents. The correlation between green bean caffeine content and chlorogenic acids content was statistically significant.

Green bean chlorogenic acids content showed statistically non-significant correlations both with cup quality and green bean physical characteristics (Table 5.3). Among cup quality attributes, body of the liquor showed negative correlation with green bean chlorogenic acids content. Green bean chlorogenic acids content had positive and statistically non-significant correlations with green bean sucrose and trigonelline contents.

Table 5.3 Correlation coefficients among cup quality, green bean physical characters and green bean caffeine, chlorogenic acids, sucrose and trigonelline contents of 42 *C. arabica* genotypes at Finoteselam, Ethiopia

Traits	Acidity	Body	Flavour	Overall standard	Bean size	Bean uniformity	Bean shape	Bean weight	Caffeine	Sucrose	Trigonelline	Chlorogenic acids
Acidity		0.851**	0.863**	0.711**	0.215	0.181	0.152	0.230	-0.378**	0.166	-0.175	0.005
Body	0.856**		0.927**	0.706**	0.365*	0.198	0.184	0.324*	-0.312*	0.246	-0.258	-0.029
Flavour	0.879**	0.928**		0.782**	0.281	0.267	0.157	0.288	-0.319*	0.354*	-0.288	0.035
Overall std	0.714**	0.694**	0.780**		0.222	0.369*	0.016	0.292	-0.251	0.205	-0.222	0.108
Bean size	0.220	0.357*	0.275	0.218		0.002	-0.132	0.800**	-0.145	-0.051	-0.540**	-0.172
Bean uifor	0.191	0.206	0.273	0.364*	0.011		0.218	0.105	-0.008	0.375*	-0.187	0.174
Bean shape	0.183	0.192	0.165	0.023	-0.125	0.218		-0.105	-0.017	0.299	0.154	0.047
Bean weigh	0.223	0.293	0.251	0.297	0.794**	0.070	-0.115		-0.087	0.073	-0.413**	0.271
Caffeine	-0.407**	-0.336*	-0.359*	-0.305*	-0.166	-0.064	-0.045	-0.075		0.167	0.148	0.317*
Sucrose	0.169	0.253	0.338*	0.181	-0.018	0.350*	0.354*	0.032	0.146		0.089	0.206
Trigonelline	-0.175	-0.215	-0.276	-0.226	-0.452**	-0.174	0.191	-0.367*	0.172	-0.016		0.101
Chlorogenic acids	0.002	-0.024	0.046	0.096	-0.164	0.154	-0.011	-0.308*	0.288	0.103	0.111	

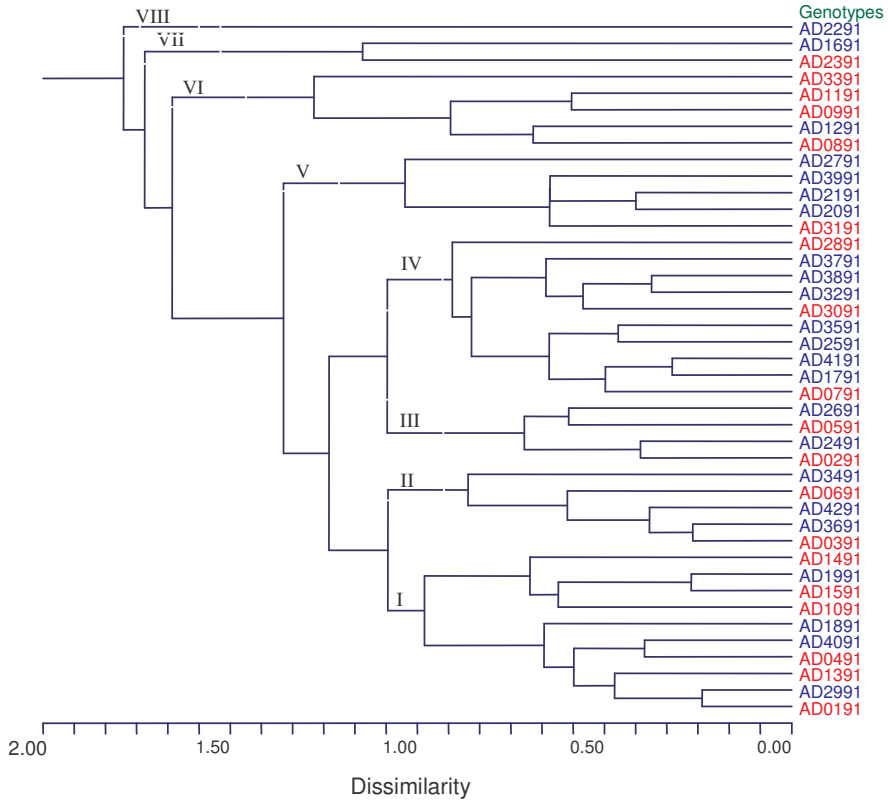
* and ** indicate significant correlations at p≤ 0.05 and 0.01, respectively.

Above the diagonal are Pearson while below the diagonal are Spearman rank correlation coefficients

Green bean sucrose content had positive correlations with all cup quality attributes. The correlation between green bean sucrose content and flavour of the liquor was statistically significant (Table 5.3). It also had positive correlations with all green bean physical characteristics except bean size. Among all green bean physical characteristics, bean uniformity had statistically significant correlation with green bean sucrose content. Green bean trigonelline content showed negative and statistically non-significant correlations with all cup quality attributes. Its associations with green bean physical characteristics were also negative, except for bean shape. Green bean trigonelline content had statistically significant correlations with bean size and bean weight (Table 5.3).

The relationship of 42 arabica coffee genotypes based on four green bean biochemical compounds was investigated using the UPGMA method of cluster analysis. Genotypes were grouped into eight clusters (Fig. 5.1). Cluster I comprised of 10 coffee genotypes, of which six were collections from southwestern Ethiopia. This group of genotypes were characterised by above average green bean caffeine, chlorogenic acids and trigonelline contents (Table 5.4). Five genotypes were assigned in cluster II, of which two were collections from southwestern Ethiopia. This group of genotypes were characterised by relatively low green bean caffeine, trigonelline and sucrose contents and higher chlorogenic acids contents.

Cluster III consisted of four genotypes of which two were collections from southwestern Ethiopia. This group of genotypes were characterised by low green bean caffeine content. Cluster IV had 10 genotypes of which three were collections from southwestern Ethiopia. This group of genotypes were characterised by low chlorogenic acids content. Cluster V consisted of five genotypes of which one was collected from southwestern Ethiopia. This group of genotypes were characterised by average green bean caffeine and trigonelline contents and low chlorogenic acids and sucrose contents. Cluster VI consisted of five genotypes of which four were collections from southwestern Ethiopia. This group of genotypes were characterised by high green bean trigonelline content. Two genotypes were assigned to cluster VII and they were characterised by high green bean sucrose and chlorogenic acids contents. Genotype AD2291 was assigned to cluster VIII and was characterised by high green bean caffeine and low trigonelline and chlorogenic acids contents.



Blue for northwestern Ethiopia and red for southwestern Ethiopia genotypes

Fig. 5.1 Dendrogram of 42 *C. arabica* genotypes constructed by cluster analysis using green bean caffeine, chlorogenic acids, sucrose and trigonelline contents

Table 5.4 Mean green bean caffeine, chlorogenic acids, sucrose and trigonelline contents for eight clusters of 42 *C. arabica* genotypes

Cluster	Number of genotypes	Caffeine	Trigonelline	Chlorogenic acids	Sucrose
I	10	1.16	1.36	3.66	7.16
II	5	1.07	1.19	3.65	6.74
III	4	0.94	1.36	3.38	7.17
IV	10	1.05	1.27	2.80	7.60
V	5	1.09	1.34	2.74	5.79
VI	5	1.19	1.59	3.62	7.81
VII	2	1.20	1.22	4.08	8.79
VIII	1	1.32	1.11	2.72	7.32
Average		1.10	1.33	3.31	7.21

Principal component analysis was performed to assess the relative importance of each biochemical compound for characterisation of genotypes and results are given in Table 5.5. Hundred percent of the variation present among genotypes was explained by four principal components. The first principal component explained 24.99% of the variation and caffeine and chlorogenic acids were more important. The second principal component explained 25.01% of the variation and this variation was mainly attributed to green bean trigonelline content variation. The third principal component explained 25.02% of the total variation and green bean sucrose and caffeine contents were more important for this variation. The fourth principal component explained 24.98% of the total variation and contributed to chlorogenic acids and caffeine contents variability.

Table 5.5 Eigenvectors and eigenvalues of four principal components for four green bean biochemical compounds

Character	Eigenvectors			
	PC1	PC2	PC3	PC4
Caffeine	-0.578	-0.049	-0.450	0.679
Trigonelline	-0.344	0.918	0.152	-0.125
Chlorogenic acids	-0.581	-0.265	-0.297	-0.710
Sucrose	-0.459	-0.290	0.828	0.137
Eigenvalue	0.9996	1.0006	1.0007	0.9991
Percent variation explained	24.99	25.01	25.02	24.98
Cumulative percent variation explained	24.99	50.01	75.02	100.00

PC = principal component

5.5 Discussion

Statistically significant green bean caffeine content variability was observed among evaluated arabica coffee genotypes. Values ranged from 0.91 to 1.32% with an average of 1.10% dmb. Similar results were reported by Clifford et al. (1989), Mazzafera and Carvalho (1992), Silvarolla et al. (2000) and Ky et al. (2001) among 21, nine, 99 and 38 *C. arabica* genotypes, respectively. Recently, Silvarolla et al. (2004) identified three naturally decaffeinated arabica coffee trees from 300 accessions collected from Ethiopia. Among all genotypes, AD0291, AD0591, AD2891, AD2491, AD2691 and AD2791 had green bean caffeine contents of less than 1.00% dmb. Therefore, results of the present study were in agreement with previous reports and some coffee genotypes with a relatively low green bean caffeine contents were identified.

Chlorogenic acids are phenolic compounds commonly found in green coffee beans. Genotypes were significantly different in green bean chlorogenic acids content and values ranged from 2.34 to 4.67% dmb with an average of 3.31% dmb. Ky et al. (2001) reported green bean chlorogenic acids content ranging from 3.4 to 4.8% dmb among 38 *C. arabica* accessions. Clifford et al. (1989) reported green bean

chlorogenic acids contents of 5.17 to 7.49% dmb among 21 *C. arabica* genotypes. The current results are lower compared to these previous reports. This could be since dry processed green beans were used for the analysis. Most of coffee berry disease resistant genotypes had relatively higher chlorogenic acids than susceptible genotypes. The association between coffee berry disease resistance and green bean chlorogenic acids content could be a future area of investigation. Genotypes were also significantly different in green bean trigonelline content and values ranged from 1.04 to 1.71% dmb. This result was in agreement with the report of Ky et al. (2001). Genotypes from southwestern Ethiopia had relatively higher trigonelline contents.

The sucrose content of coffee bean is an important parameter for coffee flavour. The higher the sucrose content in green beans, the more intense the coffee cup flavour (Ky et al., 2000). Genotypes were significantly different in green bean sucrose content and ranged from 5.30 to 8.98% dmb. Result of the present study is in agreement with the report of Ky et al. (2001). On average, genotypes from southwestern Ethiopia had higher green bean sucrose contents compared to genotypes from northwestern Ethiopia. This could be since most southwestern Ethiopia genotypes are improved varieties.

Genotypes from northwestern Ethiopia were more diverse for green bean caffeine, chlorogenic acids and sucrose contents compared to those from southwestern Ethiopia. On the other hand, genotypes from southwestern Ethiopia were more diverse in green bean trigonelline content. In addition, genotypes from southwestern Ethiopia had relatively higher caffeine, chlorogenic acids, trigonelline and sucrose contents compared to those from northwestern Ethiopia. The variation for these biochemical compounds was considerable among genotypes within each region. This indicates that varieties with desirable biochemical composition may be developed for each region using the available coffee genetic resource in the region.

Green bean caffeine content showed negative and statistically significant correlations with all cup quality attributes. Therefore, selection for both better cup quality and relatively low caffeine content seems possible. Correlations of green bean caffeine content with green bean physical characteristics were negative and statistically non-significant. Hence, the importance of green bean physical characteristics for indirect

selection of green bean caffeine content in arabica coffee seems minor. Green bean caffeine content showed positive correlations with chlorogenic acids, trigonelline and sucrose contents. Therefore, simultaneous selection for low caffeine content and high chlorogenic acids or sucrose or trigonelline content seems impossible. The correlation between green bean chlorogenic acids and caffeine contents was statistically significant. Waldhauser and Baumann (1996) reported a correlated accumulation of caffeine and chlorogenic acids in *Coffea* seeds. Therefore, one of the traits could be used for the indirect selection of the other trait. Arabica coffee has better cup quality since its beans have low caffeine and chlorogenic acids and high trigonelline and sucrose contents compared to robusta (Montagnon et al., 1998; Ky et al., 2001).

Chlorogenic acids had statistically non-significant correlations with all cup quality attributes. Its correlations were positive with most of the cup quality attributes. Therefore, both characteristics can be selected simultaneously in these arabica coffee genotypes. Green bean chlorogenic acids content showed negative correlations with most of the desirable green bean physical characteristics. Moreover, its correlation with bean weight was statistically significant. Light beans had higher chlorogenic acids compared to heavy beans. This could be since plants were triggered to synthesise more chlorogenic acids during biotic and abiotic stresses. Therefore, simultaneous selection both for high green bean chlorogenic acids and desirable green bean physical characteristics seems impossible. Green bean chlorogenic acids content showed positive but statistically non-significant correlations with bean sucrose and trigonelline contents. Hence, selection both for high chlorogenic acids, sucrose and trigonelline contents can be done simultaneously.

Green bean sucrose content showed positive correlations with all cup quality attributes. The correlation between green bean sucrose content and liquor flavour was statistically significant. Therefore, selection for higher green bean sucrose content will result in better cup quality. This is in agreement with the report of Ky et al. (2000). Correlations of green bean sucrose content with most of the desirable green bean physical characteristics were positive. Its correlations with bean uniformity and shape were statistically significant. Selection for higher green bean sucrose content and desirable green bean physical characteristics can be done simultaneously. Green bean sucrose and trigonelline contents showed negative but statistically non-significant

correlation. Therefore, simultaneous selection both for high green bean sucrose and trigonelline contents seems impossible.

Green bean trigonelline content showed negative and statistically non-significant correlations with all cup quality attributes. Therefore, selection for low trigonelline content and better cup quality seems possible in these genotypes. Its correlations with most of the desirable green bean physical characteristics were negative and associations with bean size and weight were statistically significant.

The UPGMA method of cluster analysis using green bean caffeine, chlorogenic acids, sucrose and trigonelline contents classified coffee genotypes into eight clusters. All clusters, except cluster VIII, consisted of coffee genotypes collected from both southwestern and northwestern Ethiopia. This indicated the presence of diverse genetic resources within each region as well as the close similarity of coffee genotypes collected from the two regions. All 42 genotypes were independently distinguished using data of these four biochemical compounds. This indicated the presence of adequate variability among genotypes for these biochemical compounds. Ky et al. (2001) reported the presence of high biochemical diversity in arabica coffee. Therefore, these genotypes could be used as resources to develop varieties with desirable biochemical composition in the future.

5.6 Conclusions

Assessment of green bean biochemical composition in coffee is vital since it gives information on cup quality, biotic and abiotic resistances and for diversification of the uses of coffee. Results of this study showed the presence of significant variation among genotypes for green bean caffeine, chlorogenic acids, sucrose and trigonelline contents. These genotypes could be used to develop coffee varieties with desirable green bean biochemical compositions, biotic and abiotic resistances as well as with diverse uses. Genotypes were not clustered according to collection region, demonstrating variability among genotypes within a region as well as close similarity among genotypes collected from the two regions. Evaluated biochemical compounds showed positive correlations among themselves. Therefore, it is possible to select for all of them simultaneously. Some of these biochemical compounds showed

statistically significant correlations with desirable cup quality and green bean physical characteristics. This indicated the importance of these traits for better cup quality and desirable green bean physical characteristics indirect selection.

CHAPTER 6

AFLP-BASED GENETIC RELATIONSHIPS OF ETHIOPIAN ARABICA COFFEE GENOTYPES

6.1 Abstract

Knowledge of genetic relationships among genotypes expedites crop improvement. This study was conducted to determine the genetic relationships of 28 *C. arabica* genotypes collected from the northwestern and southwestern parts of Ethiopia and evaluate the efficiency of 10 AFLP primer combinations in detecting genetic variation among arabica coffee genotypes. Jaccard similarity coefficients were calculated and a dendrogram was constructed following the UPGMA method of cluster analysis using NTSYS-pc software version 2.02i. A total of 712 fragments were amplified of which 220 (30.9%) were polymorphic among genotypes. Eighteen markers were uniquely linked to 10 genotypes and could therefore be utilised for cultivar fingerprinting. The number of polymorphic fragments amplified per primer combination varied from 10 to 39. Of the 10 primer combinations, *Eco*RI-ACA/*Mse*I-CAA, *Eco*RI-ACC/*Mse*I-CAG, *Eco*RI-ACT/*Mse*I-CAG and *Eco*RI-AAC/*Mse*I-CAA were more efficient in detecting genetic variation among genotypes. These four primer combinations should be useful for genetic diversity analysis in arabica coffee. Pair-wise genetic similarity coefficients ranged from 0.851 to 0.982**,** with an average of 0.915. AD1491 was the most divergent genotype with an average genetic similarity coefficient of 0.870. On the other hand, AD5091 was the most closely related genotype to all genotypes with an average genetic similarity coefficient of 0.930. Genotypes from southwestern Ethiopia were relatively more dissimilar than those from the northwestern Ethiopia. Genotypes were not clustered according to geographic origin. This unveiled the presence of considerable coffee genetic resource diversity within each region. Results of this study demonstrated the presence of genetic variation among evaluated coffee genotypes and the potential of AFLP analysis for genetic diversity analysis and cultivar fingerprinting in arabica coffee.

6.2 Introduction

Arabica coffee is indigenous to Ethiopia (Sylvain, 1955; Steiger et al., 2002) and is grown throughout the country except in some extremely low and high altitude areas (Yemane-Berhan, 1998). Several researchers reported the presence of phenotypic diversity among *C. arabica* genotypes grown in Ethiopia (Sylvain, 1955; FAO, 1968a; Ameha, 1986; Selvakumar and Sreenivasan, 1989; Montagnon and Bouharmont, 1996; Wondimu, 1998; Teketay, 1999; Bellachew et al., 2000; Silvarolla et al., 2000; 2004; Anzueto et al., 2001; Ky et al., 2001; Dessalegn, 2002). The genetic relationship of coffee genotypes currently grown in different parts of the country has not been extensively investigated using DNA-based markers (Aga et al., 2003). Consequently, coffee genetic resource utilisation and conservation attempts made in the country are not as efficient as required.

Recent advances in the field of molecular genetics resulted in the development of a number of DNA-based marker techniques that are widely used for cultivar fingerprinting as well as for genetic diversity analysis in different plant species. Each technique has advantages and disadvantages (Shan et al., 2004). The AFLP technique (Vos et al., 1995) has several advantages over other DNA marker systems. The most important of these are, it does not require sequence information, produces a large number of informative polymorphic markers per primer, requires a small amount of DNA, inspects an entire genome and is highly reproducible (Blears et al., 1998; Portis et al., 2004; Shan et al., 2004). It is extensively used for genetic diversity analysis in different plant species including coffee (Vos et al., 1995; Maughan et al., 1996; Ellis et al., 1997; Barrett and Kidwell, 1998; Aggarwal et al., 1999; Breyne et al., 1999; Erschadi et al., 2000; Tomkins et al., 2001; Anthony et al., 2002; Steiger et al., 2002; Ude et al., 2002; Carr et al., 2003; Herselman, 2003; Sensi et al., 2003; Turpeinen et al., 2003; Uptmoor et al., 2003; Burton et al., 2004; Nissim et al., 2004; Portis et al., 2004; Saini et al., 2004; Shan et al., 2004; Guzman et al., 2005).

Comparative studies using Restriction Fragment Length Polymorphic (RFLP), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and Simple Sequence Repeat (SSR) indicated AFLP as an efficient method to estimate genetic diversity because of its high reproducibility and

multiplex ratio (Powell et al., 1996; Russell et al., 1997; Pejic et al., 1998; Maguire et al., 2002; Uptmoor et al., 2003). These authors also reported correlations among genetic estimates using RFLP, AFLP and SSR.

RFLP, RAPD, AFLP and SSR were employed previously by a number of investigators to assess genetic diversity among and within arabica coffee genotypes and to group them according to genetic similarity. RFLP and RAPD showed relatively low degrees of polymorphism (Lashermes et al., 1996b; Paillard et al., 1996). The AFLP technique was efficient to identify the origin of cultivated cultivars (Anthony et al., 2002), analyse genetic diversity within and among cultivars (Steiger et al., 2002), detect genetic introgression (Lashermes et al., 2000a; Prakash et al., 2002; 2004) and construct genetic maps in arabica coffee (Pearl et al., 2004). Therefore, in this study, the AFLP technique was selected and employed to (1) investigate the presence of genetic variation among arabica coffee genotypes currently grown in Ethiopia, (2) explore the possibility of using AFLP markers for cultivar identification, (3) estimate genetic distances among arabica coffee genotypes and (4) compare the efficiency of different AFLP primer combinations in detecting genetic variation.

6.3 Materials and Methods

6.3.1 Genotypes

A total of 28 arabica coffee genotypes collected from the northwestern and southwestern parts of Ethiopia were evaluated for genetic variation (Table 6.1). These genotypes were obtained from Adet Agricultural Research Centre of the Amhara Regional Agricultural Research Institute, Ethiopia. Of these genotypes, five had bronze young leaves while 23 had green young leaves. These genotypes had different levels of coffee berry disease and coffee leaf rust resistances. In addition, some had compact growth habits, suitable for management practices and high density planting while others had open branching and vigorous growth habits. These genotypes were also different in cup quality and green bean biochemical composition. Generally genotypes were phenotypically diverse and collected from different areas with a range of ecological and altitudinal variations (Table 6.1).

Table 6.1 List of *C. arabica* genotypes evaluated for molecular level genetic diversity

No.	Accession number	Collection region	Altitude (masl)	Characters
1	AD0391	Southwestern Ethiopia	1900	CBD resistant, average in cup
2	AD0491	Southwestern Ethiopia	1920	CBD resistant, weak apical dominance
3	AD0591	Southwestern Ethiopia	1700	Vigorous growth habit, average in cup
4	AD0791	Southwestern Ethiopia	1700	Vigorous growth habit, good in cup
5	AD0991	Southwestern Ethiopia	1710	Dwarf growth habit, poor in cup
6	AD1191	Southwestern Ethiopia	1710	Small size bean, poor in cup
7	AD1491	Southwestern Ethiopia	1575	Branchy and had long beans
8	AD1291	Northwestern Ethiopia	1820	Has bronze young leaf and bold bean
9	AD2191	Northwestern Ethiopia	1940	Poor in cup quality
10	AD4391	Northwestern Ethiopia	1960	Green young leaf
11	AD4491	Northwestern Ethiopia	2290	Bronze young leaf
12	AD2291	Northwestern Ethiopia	2210	Green young leaf
13	AD2691	Northwestern Ethiopia	ND	Green young leaf
14	AD4591	Northwestern Ethiopia	ND	Bronze young leaf
15	AD4691	Northwestern Ethiopia	ND	Bronze young leaf
16	AD4791	Northwestern Ethiopia	1850	Green young leaf
17	AD3891	Northwestern Ethiopia	2060	Bold, heavy and uniform bean
18	AD4091	Northwestern Ethiopia	ND	Green young leaf
19	AD4891	Northwestern Ethiopia	ND	Green young leaf
20	AD2991	Northwestern Ethiopia	2040	Rust susceptible, bear on orthotropic branch, long bean, poor in cup
21	AD3991	Northwestern Ethiopia	1900	Green young leaf
22	AD3591	Northwestern Ethiopia	2000	CBD susceptible
23	AD3791	Northwestern Ethiopia	1880	Bold, heavy and uniform bean
24	AD4191	Northwestern Ethiopia	2000	Green young leaf, round and bold bean
25	AD4991	Northwestern Ethiopia	2020	Green young leaf
26	AD5091	Northwestern Ethiopia	1940	Green young leaf
27	AD1891	Northwestern Ethiopia	1930	Green young leaf, long bean
28	AD3491	Northwestern Ethiopia	1500	Adaptable to dry areas, bronze young leaf, average in cup, round bean

CBD = coffee berry disease; ND = no data; masl = metre above sea level

6.3.2 DNA extraction protocol

Total genomic DNA was isolated using the CTAB (hexadecyltrimethylammonium bromide) method (Saghai-Maroof et al., 1984). Young and healthy leaves were collected from each genotype and lyophilised using a freeze drier over a period of two days at -60 ^{o}C. Lyophilised leaves were ground to a fine powder with sterilised mortars and pestles. DNA was extracted in a 1.5 ml microfuge tube by adding ±50 mg ground leaf tissue and 750 µl extraction buffer (100 mM Tris pH 8.0, 20 mM EDTA (ethylenediaminetetraacetate) pH 8.0, 1.4 mM NaCl, 2% (w/v) CTAB, 0.2% (v/v) ß-mercaptoethanol and 200 mM urea). The leaf sample and extraction buffer were thoroughly mixed and incubated at 65 ^{o}C for 1 h with agitation by hand at every 15 min interval. To each incubated sample, 500 µl chloroform:isoamyl alcohol (24:1 (v/v)) was added, mixed well by gentle shaking and centrifuged at 12000 rpm for 3 min. The resultant supernatant was carefully transferred to a new microfuge tube. Isolated DNA was precipitated from the aqueous phase by adding 500 µl isopropanol, mixed well by gentle shaking, incubated for 20 min at room temperature and centrifuged at 12000 rpm for 5 min. The resultant supernatant was discarded and tubes kept upside down for some time to be drained. Precipitated DNA was washed by adding 500 µl ice-cold 70% (v/v) ethanol, incubated for 20 min at room temperature and centrifuged at 12000 rpm for 5 min. The resultant supernatant was discarded and the pellet was air-dried for 1 h at room temperature. The air-dried pellet was re-suspended in 200 µl TE buffer (10 mM Tris-Cl pH 8.0, 1 mM EDTA pH 8.0), pH 8.0. Twenty µl 7.5 M ammonium acetate and 200 µl chloroform:isoamyl alcohol were added, mixed well and centrifuged at 12000 rpm for 3 min. The resultant supernatant was carefully transferred to a new microfuge tube and DNA was precipitated overnight at -20 ^{o}C from the aqueous phase by adding 500 µl ice-cold 100% ethanol. DNA was recovered by centrifugation for 15 min at 12000 rpm and the resultant DNA pellet was washed twice with 500 µl ice-cold 70% ethanol by centrifuging for 10 min at 12000 rpm in each cycle. The pellet was air-dried and re-suspended in 50 µl TE buffer pH 8.0 and treated with 0.4 mg/ml DNase free Rnase A and incubated at 37 ^{o}C for 2 h to remove RNA from the isolated genomic DNA.

The concentration and purity of isolated DNA was determined using a spectrophotometer. The quality of isolated genomic DNA was assessed through

electrophoresis in 0.8% (w/v) agarose gel in 1x UNTAN (40 mM Tris-Cl, 2 mM EDTA, pH adjusted to 7.4 with acetic acid) buffer at 60 V for 45 min. DNA was diluted to a working concentration of 200 ng/µl and stored at 4 ^{o}C.

6.3.3 AFLP analysis

6.3.3.1 DNA digestion and adapter ligation

AFLP reactions were performed according to the protocol of Vos et al. (1995) with minor modifications. Genomic DNA (500 ng) was digested with 5 U of *Mse*I at 37 ^{o}C for 5 h followed by overnight digestion with *Eco*RI. The restricted DNA fragments were ligated to double-stranded *Eco*RI- and *Mse*I-adapters using T4 DNA Ligase by overnight incubation at 16 ^{o}C. The adapter ligation solution consisted of 50 pmol *Mse*I-adapter, 5 pmol *Eco*RI-adapter, 1 U T4 DNA Ligase, 0.4 mM ATP and 1x T4 DNA Ligase buffer. Primers and adapters were synthesised by Life Technologies Inc. (Glasgow, UK) and oligonucleotides used for adapters were HPLC (high performance liquid chromatography) purified. Adapters were prepared by adding equimolar amounts of both strands, heating for 10 min at 65 ^{o}C in a water bath and then leaving the mixture to cool down to room temperature.

6.3.3.2 Pre-selective and selective amplifications

Pre-selective amplification was carried out using digested and adapter ligated DNA fragments as templates and primers complementary to the adapters with one added selective nucleotide: E+A and M+C. The letter following E or M corresponds to the selective nucleotide at the 3` end of *Eco*RI and *Mse*I primers, respectively. Each sample for pre-selective amplification reaction contained 50 µl reaction mixture composed of undiluted template DNA, 1x Promega polymerase buffer, 2 mM $MgCl_2$, 200 µM of each dNTP, 30 ng *Mse*I-C, 30 ng *Eco*RI-A and 1 U Promega *Taq* polymerase. The PCR programme for pre-selective amplification consisted of 30 cycles of DNA denaturing at 94 ^{o}C for 30 s, annealing at 56 ^{o}C for 1 min and elongation at 72 ^{o}C for 1 min. The pre-selective amplified DNA was diluted 1:10 with TE buffer. The diluted pre-selective amplification product were used for selective amplification in a reaction tube containing 1x Promega polymerase buffer, 2 mM $MgCl_2$, 200 µM of each dNTP, 100 µg/µl bovine serum albumin (BSA), 30 ng *Mse*I

primer, 30 ng *Eco*RI primer and 0.75 U Promega *Taq* DNA polymerase. *Eco*RI and *Mse*I primers for selective amplification contained three selective nucleotides and the *Eco*RI primers were fluorescently labelled as Fam (blue) or Ned (yellow). The list of 10 primer combinations used in this study is given in Table 6.2. These primers were selected based on previous reports (Anthony et al., 2002; Steiger et al., 2002).

Table 6.2 List of adaptors, pre-selective and selective AFLP primers

*Eco*RI- adapter	*Mse*I-adapter
5`-CTCGTAGACTGCGTACC-3`	5`GACGATGAGTCCTGAG-3`
3`-CATCTGACGCATGGTTAA-5`	3`-TACTCAGGACTCAT-5`
*Eco*RI pre-selective primer	*Mse*I pre-selective primer
5`-GACTGCGTACCAATTCA-3`	5`-GATGAGTCCTGAGTAAC-3`
*Eco*RI primer sequence	*Mse*I primer sequence
5`-GACTGCGTACCAATTCNNN-3`	5`-GATGAGTCCTGAGTAANNN-3`

No	Primer combination code	*Eco*RI primer	*Mse*I primer
1	E1M1Fam	*Eco*RI + ACA	*Mse*I + CAA
2	E1M1Ned	*Eco*RI + AAC	*Mse*I + CAA
3	E2M1Fam	*Eco*RI + ACT	*Mse*I + CAA
4	E2M1Ned	*Eco*RI + ACC	*Mse*I + CAA
5	E2M6Fam	*Eco*RI + ACT	*Mse*I + CAG
6	E2M6Ned	*Eco*RI + ACC	*Mse*I + CAG
7	E1M7Fam	*Eco*RI + ACA	*Mse*I + CTG
8	E1M7Ned	*Eco*RI + AAC	*Mse*I + CTG
9	E2M7Fam	*Eco*RI + ACT	*Mse*I + CTG
10	E2M7Ned	*Eco*RI + ACC	*Mse*I + CTG

The PCR programme for selective amplification consisted of an initial 10 cycles of denaturing at 94 ^{o}C for 30 s, primer annealing at 65 ^{o}C for 30 s and elongation at 72 ^{o}C for 1 min, with a 1 ^{o}C decrement in annealing temperature at each cycle. This was followed by 25 cycles of 30 s at 94 ^{o}C, 30 s at 56 ^{o}C, and 1 min at 72 ^{o}C and a final elongation cycle of 5 min at 72 ^{o}C. PCR products were prepared for capillary

electrophoresis by mixing equal amounts of Fam- and Ned-reactions, 99.5% formamide and 6 fmol GENESCAN-1000 Rox^{TM} size-standard. Samples were denatured at 94^{o}C for 10 min and cooled in ice slurry for 5 min. Capillary electrophoresis was performed using a Perkin Elmer Prism 310 Automated capillary sequencer (PE Bio-systems).

6.3.4 Data analysis

DNA fragments amplified by 10 different AFLP primer combinations were scored as present (1) or absent (0) in all genotypes. Genetic similarity among accessions was calculated according to Jaccard`s Similarity Coefficient (JSC) (Jaccard, 1908) in all pair-wise combinations: JSC_{xy} = a/(n-d) where a = number of fragments shared by genotype x and y, n is total number of fragments amplified and d is the number of missed fragments both in x and y. The similarity coefficients were used to construct a dendrogram using the UPGMA method of cluster analysis. The cophenetic correlation coefficient was calculated in order to test the association between the cluster in the dendrogram and the JSC matrix. The NTSYS-pc software package version 2.02i (Rohlf, 1998) was used to calculate JSC, the cophenetic correlation coefficient and construct a dendrogram.

The ability of each primer combination to generate genetic information was compared based on number of polymorphic fragments, number of genotypes uniquely identified, Shannon-Weaver diversity index, marker index and polymorphic information content. Shannon-Weaver diversity index (H`) was used to compare levels of diversity detected by different primer combinations (Lanteri et al., 2004) and calculated over all loci amplified by each primer as follows: H` = -Σpilogpi, where pi is the frequency of the ith fragment in the sample. Marker index (MI) indicates the utility of a given marker system and was calculated according to Powell et al. (1996) as the product of expected heterozygosity (H_n) and effective multiplex ratio (EMR). H_n of a locus is defined as: 1-Σpi^2, where pi is the frequency of the ith allele (DNA fragment). EMR of a primer is defined as: ßn, where ß is the percentage of polymorphic loci (DNA fragments) and n is the number of polymorphic loci detected per primer combination (Milbourne et al., 1997). In addition, polymorphic information content (PIC) for each primer combination was calculated to estimate its discriminatory power (Lanteri et al.,

2004). Polymorphic information content was calculated according to Smith et al. (2000) as follows: PIC = 1-Σfi^2, where fi is the percentage of genotypes in which the fragment is present. Spearman's rank correlation coefficients were calculated between genetic similarity coefficient matrices resulted from 10 different primer combinations.

6.4 Results

A total of 712 fragments were amplified among 28 *C. arabica* genotypes using 10 different AFLP primer combinations, of which 220 fragments (30.9% of the total amplified fragments) were polymorphic among tested genotypes (Table 6.3). The number of polymorphic fragments per primer combination ranged from 10 to 39 with an average of 22. Among the 10 primer combinations tested, E2M7Ned and E1M1Ned showed the lowest and highest number of both total and polymorphic fragments, respectively.

Percentage of polymorphic fragments (P %), H`, PIC and MI were calculated for each primer combination and are given in Table 6.3. Primer combination E2M7Fam showed the highest value for H`, while primer combination E2M6Fam showed the highest values for P% and MI and distinguished all (28) coffee genotypes. The highest PIC value was recorded for primer combination E2M6Ned. On the other hand, primer combinations E1M7Fam, E1M1Ned and E2M7Ned showed the lowest values for H`, PIC and MI, respectively. In addition, E1M7Fam distinguished only 28.6% of the genotypes tested (Table 6.3). Therefore, primer combination E2M6Fam was highly discriminative compared to other primer combinations.

The importance of each primer combination was assessed based on the number of unique or distinctive fragments (i.e. fragments present only in one genotype) generated. A total of 18 unique fragments were recorded from eight primer combinations. As a result, 10 arabica coffee genotypes were uniquely characterised using these distinctive fragments (Table 6.4). E2M1Fam and E2M1Ned did not reveal unique fragments among genotypes. E2M6Ned amplified the highest number (six) of unique fragments on a large number (four) of genotypes. Of the 10 genotypes with distinctive fragments, genotype AD1491 had the maximum number (six) of unique fragments.

Table 6.3 Levels of genetic information generated by 10 AFLP primer combinations on 28 *C arabica* genotypes

No.	Primer combination code	Selective nucleotide sequence		TNF	NPF	P (%)	H`	PIC	MI	NGI	NUF	Genetic similarity			Fragment size (bp)	
		*Eco*RI (5`→3`)	*Mse*I (5`→3`)									Min.	Max.	Mean	Min.	Max
1	E1M1Fam	ACA	CAA	90	25	27.8	0.620	0.52	3.61	26	1	0.839	1.000	0.915	40	512
2	E1M1Ned	AAC	CAA	108	39	36.1	0.473	0.35	4.93	26	2	0.788	1.000	0.921	46	522
3	E2M1Fam	ACT	CAA	71	21	29.6	0.530	0.39	2.42	20	0	0.843	1.000	0.929	44	466
4	E2M1Ned	ACC	CAA	67	19	28.4	0.610	0.42	2.27	24	0	0.821	1.000	0.909	43	487
5	E2M6Fam	ACT	CAG	70	31	44.3	0.653	0.63	8.65	28	1	0.714	0.981	0.841	39	452
6	E2M6Ned	ACC	CAG	69	24	34.8	0.517	0.73	6.10	25	6	0.754	1.000	0.906	45	536
7	E1M7Fam	ACA	CTG	50	16	32.0	0.376	0.42	2.15	8	2	0.809	1.000	0.948	39	365
8	E1M7Ned	AAC	CTG	69	17	24.6	0.529	0.45	1.88	17	3	0.864	1.000	0.939	42	437
9	E2M7Fam	ACT	CTG	72	18	25.0	0.716	0.65	2.93	22	2	0.794	1.000	0.905	38	518
10	E2M7Ned	ACC	CTG	46	10	21.7	0.494	0.37	0.80	9	1	0.844	1.000	0.951	39	432
Total				712	220						18					
Mean				71	22	30.4	0.552	0.49	3.57	20.5	1.8	0.851	0.982	0.915		

TNF = total number of fragments; NPF = number of polymorphic fragments; P (%) = percentage of polymorphic fragments; H` = Shannon-Weaver diversity index; PIC = polymorphic information content; MI = marker index; NGI = number of genotypes identified; NUF = number of unique fragments; bp = base pair

Table 6.4 List of AFLP primer combinations with unique fragment(s), fragment size and genotypes on which unique fragment(s) were identified

No.	Primer code	Selective nucleotide sequence		Unique fragments (bp) (identified genotype)
		*Eco*RI (5`→3`)	*Mse*I (5`→3`)	
1	E1M1Fam	ACA	CAA	168bp (AD1291)
2	E1M1Ned	AAC	CAA	156bp (AD4791) & 325bp (AD1491)
3	E2M6Fam	ACT	CAG	105bp (AD4591)
4	E2M6Ned	ACC	CAG	122bp, 211bp & 241bp (AD1491), 59bp (AD2691), 279bp (AD3991), & 272bp (AD4991)
5	E1M7Fam	ACA	CTG	198bp (AD1891) & 226bp (AD3491)
6	E1M7Ned	AAC	CTG	137bp (AD1491), 332bp (AD2291), & 77bp (AD3491)
7	E2M7Fam	ACT	CTG	103bp (AD1291) & 115bp (AD3491)
8	E2M7Ned	ACC	CTG	137bp (AD1491)

Correlation coefficients between genetic similarity matrices estimated by different AFLP primer combinations were calculated and results are presented in Table 6.5. The lowest correlation coefficient (-0.001) was recorded between genetic similarity matrices estimated by primer combinations E1M1Ned and E2M7Fam, while the highest value (0.447) was recorded between genetic similarity matrices estimated by primer combinations E1M1Fam and E1M1Ned (Table 6.5). Except for E1M1Ned and E2M7Fam and E1M7Fam and E2M7Fam, all other primer combinations showed positive correlations in estimating genetic similarity among tested coffee genotypes. This illustrated the similarity of most primer combinations in estimating genetic similarity among tested genotypes. The correlation of genetic similarity matrices estimated by E1M1Fam and E1M7Fam, E1M7Fam and E1M7Ned, E1M1Ned and E2M7Fam, and E1M7Fam and E2M7Fam were statistically non-significant. Levels of correlation were generally low, less than 0.500. This indicated low levels of duplication or redundancy of information while using different primer combinations.

Table 6.5 Correlation coefficients among 10 AFLP primer combinations in estimating genetic similarity among genotypes

Primer code	E1M1Ned	E2M1Fam	E2M1Ned	E2M6Fam	E2M6Ned	E1M7Fam	E1M7Ned	E2M7Fam	E2M7Ned	All primers
E1M1Fam	0.447**	0.378**	0.376**	0.269**	0.279**	0.096	0.330**	0.401**	0.236**	0.725**
E1M1Ned		0.348**	0.186**	0.153**	0.372**	0.299**	0.321**	-0.001	0.193**	0.608**
E2M1Fam			0.293**	0.315**	0.262**	0.129*	0.222**	0.353**	0.350**	0.623**
E2M1Ned				0.129*	0.248**	0.206**	0.207**	0.314**	0.202**	0.525**
E2M6Fam					0.387**	0.171**	0.109*	0.230**	0.207**	0.558**
E2M6Ned						0.333**	0.162**	0.125*	0.425**	0.622**
E1M7Fam							0.073	-0.095	0.225**	0.348**
E1M7Ned								0.118*	0.110*	0.456**
E2M7Fam									0.224**	0.501**
E2M7Ned										0.449**

* and ** indicate significant correlations at p≤ 0.05 and 0.01, respectively

The correlation of genetic similarity estimated by each primer combination and from combined data of all primer combinations was analysed. Of all the primer combinations, E1M7Fam and E1M1Fam had the lowest and highest correlation coefficients, respectively, with the genetic similarity estimated using combined data of all primer combinations. However, no primer combination estimated genetic similarity identical to that of combined data of all primer combinations. Each primer combination probably scanned different parts of a genome, providing different but complementary information.

Pair-wise genetic similarity coefficients were calculated using 712 fragments generated by all primer combinations. Pair-wise genetic similarity coefficients varied from 0.851 to 0.982 with an average of 0.915 (Table 6.6). Genotype AD1291 and AD1491 were the most dissimilar, whereas AD3591 and AD3991 were the most similar genotypes. AD1291 and AD1491 were phenotypically distinct and collected from different agro-ecologies and regions. AD3591 and AD3991 were collections from northwestern Ethiopia within a 25 km distance. Of the 378 pair-wise genotype combinations, 82 combinations had genetic similarity coefficients less than 0.900 and genotype AD1491 was one of the genotypes occurring in most of these combinations.

High genetic similarity coefficients were recorded between genotypes collected from different regions. For example, genotype AD0591, collected from southwestern Ethiopia, showed relatively high genetic similarity with genotypes AD3591 and AD3991, which were collections from northwestern Ethiopia. Relatively low genetic similarity was observed among genotypes collected from the same region. Genotypes AD4491 and AD4591 were collections from northwestern Ethiopia but were genetically dissimilar.

Table 6.6 Pair-wise genetic similarity coefficient matrix for 28 *C. arabica* genotypes calculated from AFLP analysis data

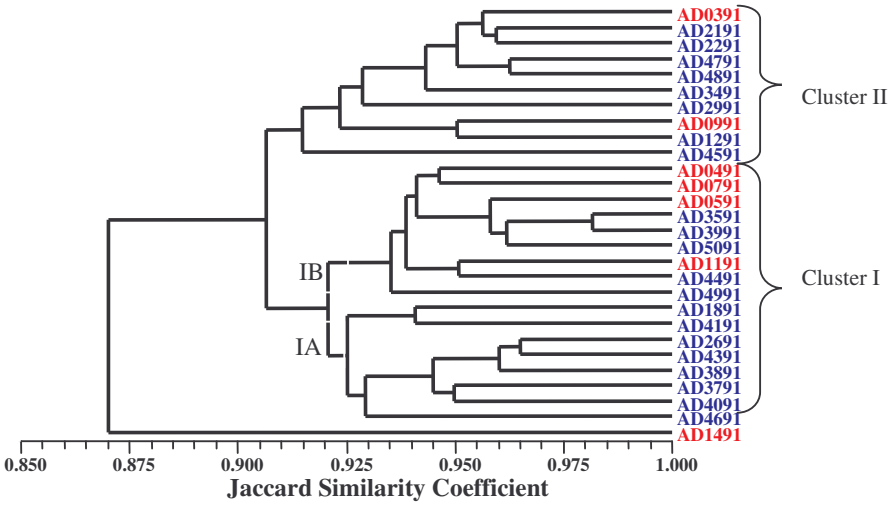
Genotype	AD0391	AD0491	AD0591	AD0791	AD0991	AD1191	AD1491	AD1291	AD1891	AD2191	AD2291	AD2691	AD2991	AD3491
AD0491	0.911													
AD0591	0.913	0.956												
AD0791	0.909	0.946	0.936											
AD0991	0.955	0.921	0.929	0.922										
AD1191	0.918	0.937	0.930	0.938	0.925									
AD1491	0.869	0.871	0.875	0.871	0.861	0.863								
AD1291	0.925	0.910	0.920	0.908	0.951	0.899	0.851							
AD1891	0.906	0.902	0.918	0.920	0.907	0.920	0.862	0.899						
AD2191	0.957	0.891	0.901	0.895	0.934	0.903	0.865	0.919	0.908					
AD2291	0.957	0.880	0.892	0.881	0.922	0.886	0.860	0.922	0.883	0.959				
AD2691	0.907	0.912	0.922	0.907	0.900	0.904	0.883	0.889	0.912	0.910	0.893			
AD2991	0.925	0.895	0.911	0.904	0.917	0.892	0.875	0.915	0.898	0.930	0.928	0.920		
AD3491	0.936	0.895	0.902	0.893	0.920	0.895	0.858	0.903	0.887	0.942	0.942	0.899	0.934	
AD3591	0.909	0.939	0.956	0.938	0.922	0.940	0.868	0.919	0.928	0.908	0.891	0.921	0.901	0.890
AD3791	0.902	0.910	0.920	0.911	0.898	0.902	0.879	0.890	0.919	0.905	0.891	0.951	0.918	0.889
AD3891	0.924	0.914	0.925	0.915	0.920	0.915	0.883	0.911	0.921	0.924	0.906	0.956	0.922	0.899
AD3991	0.911	0.942	0.959	0.941	0.922	0.940	0.860	0.913	0.928	0.908	0.889	0.927	0.901	0.890
AD4091	0.923	0.914	0.936	0.923	0.916	0.920	0.891	0.916	0.923	0.914	0.909	0.942	0.930	0.907
AD4191	0.918	0.920	0.930	0.926	0.919	0.926	0.868	0.916	0.941	0.914	0.894	0.933	0.916	0.890
AD4391	0.916	0.915	0.931	0.919	0.911	0.915	0.898	0.900	0.921	0.915	0.898	0.965	0.922	0.905
AD4491	0.907	0.941	0.939	0.933	0.923	0.951	0.864	0.909	0.901	0.890	0.887	0.896	0.894	0.899
AD4591	0.913	0.884	0.897	0.896	0.915	0.887	0.861	0.904	0.896	0.913	0.916	0.911	0.926	0.905
AD4691	0.911	0.915	0.917	0.908	0.912	0.913	0.872	0.909	0.901	0.896	0.893	0.929	0.926	0.891
AD4791	0.957	0.911	0.913	0.912	0.943	0.915	0.863	0.914	0.900	0.947	0.948	0.904	0.925	0.945
AD4891	0.948	0.900	0.910	0.901	0.925	0.903	0.860	0.914	0.912	0.960	0.944	0.916	0.931	0.951
AD4991	0.921	0.929	0.930	0.936	0.920	0.936	0.883	0.902	0.926	0.909	0.886	0.928	0.913	0.899
AD5091	0.925	0.941	0.960	0.936	0.932	0.930	0.881	0.929	0.918	0.919	0.910	0.931	0.917	0.914

Table 6.6 (Contd.)

Genotype	AD3591	AD3791	AD3891	AD3991	AD4091	AD4191	AD4391	AD4491	AD4591	AD4691	AD4791	AD4891	AD4991
AD3791	0.928												
AD3891	0.923	0.935											
AD3991	0.982	0.928	0.926										
AD4091	0.934	0.950	0.945	0.931									
AD4191	0.946	0.938	0.933	0.946	0.935								
AD4391	0.929	0.948	0.965	0.929	0.948	0.936							
AD4491	0.944	0.900	0.899	0.941	0.915	0.912	0.902						
AD4591	0.898	0.904	0.908	0.896	0.913	0.896	0.905	0.877					
AD4691	0.907	0.918	0.941	0.907	0.925	0.928	0.932	0.891	0.897				
AD4791	0.914	0.897	0.916	0.914	0.915	0.903	0.913	0.910	0.919	0.899			
AD4891	0.911	0.902	0.924	0.914	0.915	0.912	0.921	0.899	0.922	0.893	0.963		
AD4991	0.947	0.932	0.930	0.941	0.944	0.932	0.933	0.919	0.916	0.917	0.915	0.915	
AD5091	0.962	0.930	0.934	0.962	0.942	0.942	0.940	0.949	0.903	0.917	0.922	0.925	0.945

The genetic relationship of coffee genotypes collected in each region was investigated. A total of 186 polymorphic AFLP markers were identified among 21 coffee genotypes collected from northwestern Ethiopia. For seven genotypes from southwestern Ethiopia, a total of 143 markers were identified. Pair-wise genetic similarity coefficients ranged from 0.877 to 0.982 and 0.861 to 0.956 among coffee genotypes collected from northwestern and southwestern Ethiopia, respectively. This indicated the presence of relatively higher levels of genetic dissimilarity among coffee genotypes collected from southwestern Ethiopia compared to those from northwestern Ethiopia. Of the 220 polymorphic AFLP markers detected among 28 coffee genotypes, 24 were missing and 53 were monomorphic across all seven coffee genotypes collected from southwestern Ethiopia. On the other hand, six were missing and 28 were monomorphic among all 21 coffee genotypes collected from northwestern Ethiopia. Genotypes from northwestern Ethiopia contained 97.3% of the detected AFLP markers.

Cluster analysis was employed to explore the pattern of genetic relationships among Ethiopian coffee genotypes. A dendrogram was constructed using the UPGMA method of cluster analysis from pair-wise genetic similarity coefficients calculated using combined data of all primer combinations. The first bifurcation separated genotype AD1491 from the other 27 coffee genotypes tested in this study (Fig. 6.1). Subsequently, the 27 coffee genotypes were classified into two clusters. Cluster I contained 17 coffee genotypes, including four collections from southwestern Ethiopia. Cluster II comprised of 10 coffee genotypes, of which two were collections from southwestern Ethiopia. The first main cluster further divided into two sub-clusters. Sub-cluster IA contained eight genotypes collected from northwestern Ethiopia. Sub-cluster IB comprised of nine coffee genotypes collected both from northwestern and southwestern Ethiopia. Cluster analysis performed using combined data of 10 AFLP primer combinations did not assign genotypes according to collection regions. The cophenetic correlation coefficient was 0.87, suggesting a good fit between the dendrogram clusters and the genetic similarity matrices.



Blue for northwestern Ethiopia and red for southwestern Ethiopia genotypes

Fig. 6.1 Dendrogram of 28 *C. arabica* genotypes constructed by cluster analysis using AFLP analysis data

6.5 Discussion

Knowledge of genetic relationships among genotypes is useful in any plant breeding programme. It provides useful information for the efficient utilisation and conservation of available genetic resources. This study assessed the genetic relationship of 28 arabica coffee genotypes currently grown in northwestern and southwestern parts of Ethiopia using 10 different AFLP primer combinations.

AFLP fingerprinting in any crop species requires selection of primer combinations that provide sufficient and reliable information. All primer combinations tested in this study detected polymorphisms among 28 *C. arabica* genotypes. The number of polymorphic fragments detected per primer combination was variable and ranged from 10 to 39, with an average of 22. Previous studies using AFLP markers in different plant species confirmed the presence of variability among different primer combinations in detecting polymorphisms among genotypes (Barrett and Kidwell, 1998; Baker et al., 1999; Breyne et al., 1999; Yee et al., 1999; Garcia-Mas et al., 2000; Renganayaki et al., 2001; Turpeinen et al., 2003). Steiger et al. (2002)

evaluated the polymorphism rate of 24 AFLP primer combinations, using 58 arabica coffee accessions and three accessions from two diploid *Coffea* species and reported the presence of high levels of variability among primer combinations. The result of the present study was in agreement with previous findings.

The efficiency of 10 primer combinations were compared based on number of polymorphic fragments, Shannon-Weaver diversity index, polymorphic information content, marker index, number of genotypes identified, number of distinctive fragments and correlation of pair-wise genetic similarity coefficients calculated from data of each primer combination and combined data of all primer combinations. In terms of number of polymorphic fragments detected per primer combination, E1M1Ned, E2M6Fam, E1M1Fam and E2M6Ned were superior. E2M7Fam, E2M6Fam, E1M1Fam and E2M1Ned had relatively higher Shannon-Weaver diversity indices. Primer combinations E2M6Ned, E2M7Fam, E2M6Fam and E1M1Fam generated high polymorphic information contents. E2M6Fam, E2M6Ned, E1M1Ned and E1M1Fam had higher marker indices. E2M6Fam distinguished 28, E1M1Fam and E1M1Ned 26 and E2M6Ned 25 genotypes of the 28 genotypes tested. E2M6Ned and E1M7Ned had more than two unique or distinctive fragments that uniquely distinguished genotypes. Positive and highly significant correlations were observed between pair-wise genetic similarity coefficients calculated from combined data of all primer combinations and data of primer combinations E1M1Fam, E2M1Fam, E2M6Ned and E1M1Ned. Although this study did not identify a superior primer combination for all parameters considered, primer combinations E1M1Fam, E2M6Ned, E2M6Fam and E1M1Ned were superior for most of the parameters considered. The last three primer combinations were also used by Anthony et al. (2002) to evaluate genetic diversity among wild and cultivated *C. arabica* genotypes. Two of these primer combinations (E1M1Ned and E2M6Ned) were recommended as being more informative out of 24 primer combinations tested by Steiger et al. (2002). These primer combinations seem promising for arabica coffee genetic diversity analysis.

Using all 10 primer combinations, a total of 712 fragments were amplified, of which 220 fragments were polymorphic among the 28 arabica coffee genotypes. With the help of these 220 AFLP markers, all 28 genotypes were distinctly identified. All 28

coffee genotypes were distinguished using 31 polymorphic fragments generated by a single primer combination, E2M6Fam. Anthony et al. (2002) identified 26 arabica coffee accessions using 107 AFLP markers. Steiger et al. (2002) identified 58 arabica coffee accessions using 274 informative AFLP markers. AFLP analyses were employed by Noir et al. (2003) and Prakash et al. (2004) to identify introgressed root-knot nematode and leaf rust resistance genes, respectively, from the *C. arabica* genome. Pearl et al. (2004) constructed the genetic map of arabica coffee with a total length of 1,802.8 cM and an average distance of 10.2 cM between adjacent markers using AFLP analysis. These research results demonstrated the importance and efficiency of the AFLP technique for different genetic studies in arabica coffee.

The genetic similarity coefficient computed for 28 coffee genotypes using 712 AFLP markers ranged from 0.851 to 0.982 with an average of 0.915. Despite the presence of wide morphological variations among tested genotypes, variation at DNA level was relatively low. Similar results have been observed in tomato (Miller and Tanksley, 1990), melons (Shattuck-Eidens et al., 1990), soybean (Keim et al., 1990) and common bean (Gepts, 1991). Steiger et al. (2002) reported an average genetic similarity coefficient of 0.933 among 58 arabica coffee accessions. Results of the present study were in agreement with previous research results. This result confirmed the low genetic diversity characteristic of *C. arabica,* which is attributable to its reproductive biology and evolution (Lashermes et al., 2000b). Observed wide agro-morphological variation among genotypes may have resulted from spontaneous mutations of few major genes conditioning plant, fruit and seed characteristics (Carvalho, 1988). Major morphological and physiological traits are often controlled by a few major genes and were subjected to intense selection pressure during the process of domestication. As a result, morphological variation is likely to increase while molecular markers, which are not subject to direct selection often decreases (Gepts, 1991).

The most similar genotypes were AD3591 and AD3991. These genotypes were collected in northwestern Ethiopia, in locations with similar agro-ecologies. The most dissimilar genotypes were AD1291 and AD1491. These two genotypes were morphologically different, prefer different agro-ecologies and were collected from different regions. Despite this result, high genetic similarities were also observed

among genotypes collected from different regions. This could be attributed to the unrestricted movement of coffee planting materials from region to region by different agents. The perennial nature as well as the highly self-pollination (over 95%) (Silvarolla et al., 2004) habit of this crop may also contribute to this observation. On the other hand, some genotypes collected from the same region had low genetic similarities. For example, genotypes AD1191 and AD1491 were both collected from southwestern Ethiopia, but were genetically dissimilar. The low genetic similarity between genotypes from the same region could be attributed to the vast agro-ecological diversity present within each region that can contribute to the development of genetically diverse genotypes through natural selection.

Duplicated accessions were not detected among studied genotypes and reasonable genetic distances were observed between some of the genotypes. Ajmone-Marsan et al. (1998) reported positive correlations between pair-wise genetic distances and specific combining ability effects for yield in maize. Bellachew et al. (1993) reported a large heterosis effect in F_1 crosses between coffee genotypes from southwestern and southeastern Ethiopia. A number of researchers identified these genotypes as genetically distinct using DNA markers (Lashermes et al., 1996b; Anthony et al., 2001; 2002). Hence results of the present study may help to select parental genotypes for hybrid variety development.

Results of the present study and the previous study by Anthony et al. (2002) revealed the presence of genetic variation among arabica coffee genotypes from Ethiopia. In addition, in the past many coffee breeders identified some highly desirable characteristics including CBD resistance (Bellachew et al., 2000), coffee leaf rust resistance (Wondimu, 1998), nematode resistance (Anzueto et al., 2001), naturally decaffeinated coffee (Silvarolla et al., 2004) and excellent cup quality coffee (Selvakumar and Sreenivasan, 1989) genotypes from Ethiopian arabica coffee germplasm collections. These research indicated the importance of the Ethiopian arabica coffee gene pool for the world coffee industry.

The average genetic distance among coffee genotypes from southwestern Ethiopia was higher than from genotypes collected in northwestern Ethiopia. This illustrated the presence of higher genetic diversity among coffee genotypes in southwestern than

northwestern Ethiopia. This was expected since arabica coffee originated in southwestern Ethiopia and spread latter to the northwest. In agreement with this result, Mengesha (1975) reported the presence of low levels of arabica coffee genetic resource diversity in northwestern Ethiopia. Anthony et al. (2001) observed narrow genetic diversity among coffee accessions collected from northwestern Ethiopia.

Cluster analysis was employed to observe the genetic relationship among 28 coffee genotypes. According to the dendrogram constructed using 712 AFLP markers, genotype AD1491 clustered separately from other genotypes. This indicated the genetic divergence of AD1491 compared to all other genotypes. This observation was in agreement with its ecological adaptation differences from other genotypes. The remaining 27 genotypes were grouped into two clusters. The first cluster comprised 17 genotypes and the second cluster contained 10 genotypes. Both clusters contained genotypes collected from different regions. Therefore, cluster analysis did not group genotypes according to collection regions. Lashermes et al. (1996b) reported similar results while studying the genetic diversity of arabica coffee genotypes collected from different locations of southwestern Ethiopia using RAPD markers. During a study by Anthony et al. (2001) coffee genotypes from northwestern and southwestern Ethiopia were clustered together. Therefore, results of the present study were in agreement with previous reports.

It was not possible to consistently correlate the genetic clustering with the morphological characteristics or growth habits usually adopted for cultivar identification. For example, AD0991 and AD1191 are morphologically similar and characterised by dwarf and compact growth habits but were assigned in different clusters. Similarly, AD4591 and AD4691 were collected from the same coffee orchard and both had bronze young leaves but were assigned in different clusters. Similar results were reported by Portis et al. (2004) on oleander and Guzman et al. (2005) on capsicum. Ruiz et al. (2000) attributed such observations to randomness of identified AFLP markers on a genome compared to the possibility that these morphological characteristics are controlled by a small portion of the genome. The chance that any of the identified markers targeted a genomic region involved in the control of the differentiating characteristics will be very low. Charcosset and Moreau (2004) attributed the above observation to the neutrality of molecular markers which

means that their polymorphism did not contribute directly to the variation in traits of interest.

Montagnon and Bouharmont (1996) using agro-morphological characters, Anthony et al. (2001) using 29 polymorphic RAPD markers and Anthony et al. (2002) using 107 AFLP markers classified wild and cultivated coffee genotypes from different parts of Ethiopia only into two groups namely southwestern group and south/southeastern group. Results of the present study together with previous research reports indicated that coffee genotypes from northwestern Ethiopia were obtained from the southwestern coffee genetic resource base because no clustering according to collection region. In agreement with the above suggestion, farmers from major coffee growing areas of northwestern Ethiopia reported that coffee was brought to their area from the southwestern part of the country by a Christian monk, Aba Beteremariam in the 13th century.

6.6 Conclusions

The genetic relationship of 28 *C. arabica* genotypes collected from the northwestern and southwestern parts of Ethiopia was analysed using 712 fragments generated by 10 AFLP primer combinations. Of the 10 AFLP primer combinations, E1M1Fam, E2M6Ned, E2M6Fam and E1M1Ned were more efficient for detection genetic variations among genotypes and seem promising for genetic diversity analysis in arabica coffee. Ten genotypes namely AD1291, AD1491, AD1891, AD2291, AD2691, AD3491, AD3991, AD4591, AD4791 and AD4991 were uniquely distinguished using 18 specific markers detected by eight primer combinations. These specific fragments should be converted into sequence characterised amplified region (SCAR) markers for more cost effective fingerprinting of these genotypes. Genotypes did not cluster according to geographic origin due to unrestricted movement of coffee planting materials from region to region and the existence of vast agro-ecological diversity within each region. Results indicated that coffee genotypes from northwestern Ethiopia were derived from the southwestern Ethiopian coffee gene pool. This genetic resource is currently under threat due to deforestation of its natural habitat, disease outbreak like CBD and replacement of landraces with few high yielding and disease resistant varieties. The current low market price for coffee also

contributed to coffee genetic resource erosion since it is enforcing farmers to replace coffee by other crops. Therefore, concerted efforts from all sides are required to save this internationally valuable genetic resource.

CHAPTER 7

SSR-BASED GENETIC RELATIONSHIPS OF ETHIOPIAN ARABICA COFFEE GENOTYPES

7.1 Abstract

The genetic relationships of 21 arabica coffee genotypes collected from northwestern Ethiopia and seven genotypes from southwestern Ethiopia were investigated using six simple sequence repeat (SSR) primers. The aims of the study were to assess the genetic relationship of coffee genotypes collected from the same as well as different regions using SSR marker technology and evaluate the efficiency of six SSR primers in detecting genetic variation in arabica coffee. Jaccard`s genetic similarity coefficients were calculated and a dendrogram was constructed following the UPGMA method of cluster analysis using NTSYS-pc software version 2.02i. A total of 22 alleles were amplified, of which 20 were polymorphic and 64.3% of the tested genotypes were uniquely identified. The number of amplified alleles per primer ranged from two to six with an average of 3.7. Polymorphic information content (PIC) and observed heterozygosity (Ho) per primer ranged from 0.004 to 0.750 and 0.00 to 1.00, respectively. M24 and M32 were superior in detecting genetic variation compared to other primers. Two alleles, M32-145 and M32-136, were specific to genotypes AD1491 and AD2991, respectively. Jaccard`s pair-wise genetic similarity coefficients ranged from 0.286 to 1.000 with an average of 0.560. Higher genetic diversity was detected among coffee genotypes collected from southwestern Ethiopia compared to those from northwestern Ethiopia. Genotypes grouped into five main clusters and there was a tendency to cluster according to collection locality. Results from this study demonstrated the presence of considerable genetic variation among some of the tested coffee genotypes.

7.2 Introduction

Genetic diversity is the basis for genetic improvement. Information regarding the genetic diversity of available germplasm is vital to devise efficient plant breeding programmes as well as to maintain genetic diversity in a given gene pool. Genetic diversity can be estimated using morphological, biochemical and DNA-based markers. Morphological markers are often influenced by prevailing environmental conditions. In addition, morphological markers are not suitable for perennial crops as it takes a long time to generate data. DNA-based markers offer consistent results regardless cropping conditions and type and age of sampled tissue (Sakiyama, 2000). These characteristics of DNA-based markers make it suitable for coffee research, since coffee is a perennial crop with a long juvenile period.

Simple sequence repeat (SSR) is a relatively new class of plant DNA marker. It has many advantages including being rapid, reliable (Diwan and Cregan, 1997), abundant, co-dominant (Sanchez-Perez et al., 2005), highly heterozygous (Powell et al., 1996), highly polymorphic, evenly dispersed along the genome, highly reproducible, somatically stable (Rovelli et al., 2000; Hormaza, 2002) and easy to assay using polymerase chain reaction (PCR) (Kuleung et al., 2004). It has been used, for example, for construction of genetic linkage maps of maize (Taramino and Tingey, 1996), sorghum (Taramino et al., 1997), soybean (Cregan et al., 1999) and rice (Temnykh et al., 2000). It has been used for cultivar identification of potato (McGregor et al., 2000), peach (Aranzana et al., 2003) and carnation (Smulders et al., 2003) and in genetic diversity analysis of different plant species including coffee (Dean et al., 1999; Combes et al., 2000; Rovelli et al., 2000; Smith et al., 2000; Kubik et al., 2001; Tanya et al., 2001; Anthony et al., 2002; Hormaza, 2002; Huang et al., 2002; Pinto et al., 2003; Akkaya and Buyukunal-Bal, 2004; Bandelj et al., 2004; Bond et al., 2004; Hamza et al., 2004; Tonguc and Griffiths, 2004; Vaz-Patto et al., 2004).

The primary disadvantage of SSR as a molecular marker is the cost and research effort required for development, since the technique demands specific primers for every species of interest (Kuleung et al., 2004). However, transferability, the application of SSR markers developed for one species on another related species, has been proposed and was successfully demonstrated in many species of *Glycine*

(Peakall et al., 1998), *Coffea* (Combes et al., 2000), *Lycopersicon* (Alvarez et al., 2001), *Prunus* (Dirlewanger et al., 2002; Hormaza, 2002), *Convolvulaceae* (Hu et al., 2004) and among wheat, rye and triticale (Kuleung et al., 2004). Comparative studies using RFLP, RAPD, AFLP and SSR indicated that SSR markers detected higher levels of polymorphism compared to other markers (Powell et al., 1996; Pejic et al., 1998; Maguire et al., 2002; Bandelj et al., 2004). The highly polymorphic nature of SSR markers make them particularly useful for genetic diversity analysis in species with a narrow genetic basis (Akkaya et al., 1992; Huang et al., 2002).

In arabica coffee, polymorphic SSRs were identified by Combes et al. (2000) and Rovelli et al. (2000). Rovelli et al. (2000) constructed two genomic libraries, one enriched for $(ATC)_n$ and the other for $(TG)_n$. These libraries were tested on arabica coffee germplasm for SSR polymorphism. SSR loci with $(TG)_n$ repeats proved to be more polymorphic than $(ATC)_n$ repeats. Combes et al. (2000) identified five polymorphic SSR markers out of 11 tested on arabica coffee and the mean heterozygosity value was 0.04. Anthony et al. (2002) tesed six polymorphic SSR markers and identified two alleles, namely M24-147 and M24-149, which distinguished the typica from the bourbon derived accessions. These authors observed seven alleles common in all typica- and bourbon-derived accessions as well as Yemen cultivars. However, the genetic relationship of coffee genotypes currently grown in different parts of Ethiopia was not evaluated using SSR markers. This study was conducted to evaluate the efficiency of six SSR primers in detecting genetic variation among *C. arabica* genotypes collected from the northwestern and southwestern parts of Ethiopia, estimate genetic similarity coefficients among these genotypes and classify them according to genetic relationships.

7.3 Materials and Methods

7.3.1 Genotypes and DNA extraction protocol

Twenty-eight arabica coffee genotypes collected from the northwestern and southwestern parts of Ethiopia were used for this study. Their description, collection region and source are given in section 6.3.1 and Table 6.1. Genomic DNA was extracted from lyophilized leaves as described in section 6.3.2.

7.3.2 SSR primers and analysis protocol

Six SSR primers which previously showed clear polymorphisms in arabica coffee (Anthony et al., 2002) were used in this study (Table 7.1). SSR assays were performed according to Combes et al. (2000) with minor modifications. Reaction mixtures for PCR amplification of SSR loci contained 100 ng genomic DNA, 1x Promega *Taq* buffer, 2 mM $MgCl_2$, 200 μM of each dNTP, 50 ng forward primer, 50 ng reverse primer and 1 U *Taq* DNA polymerase (Promega) in a final volume of 20 μl. Reactions were performed using a Touchdown Hybaid PCR machine. The amplification cycle consisted of an initial 2 min denaturing at 94 ^{o}C, followed by 5 cycles of denaturation at 94 ^{o}C for 45 s, a 1 min primer annealing at 60 ^{o}C, with a 1 ^{o}C decrease in temperature at each cycle and 1 min and 30 s elongation at 72 ^{o}C. This was followed by 30 cycles of 45 s at 94 ^{o}C, 1 min at 55 ^{o}C and 1 min and 30 s at 72 ^{o}C and a final 8 min elongation at 72 ^{o}C. Amplification products were separated on 5% denaturing polyacrylamide gels (19:1 acrylamide:bis-acrylamide) containing 7 M urea and 1x TBE (89 mM Tris-borate; 2.0 mM EDTA) buffer**.** A 25 bp DNA ladder was used as size standard. Electrophoresis was performed at constant power of 80 W for 1 hour.

Table 7.1 Repeat motifs and sequences for tested SSR primers

Primer code	Accession number of genomic sequence	Repeat motif	Primer sequence (forward (F) and reverse (R)) 5`→3`
M20	AJ250253	$(GA)_5(GT)_8TT(GT)_4TT(GT)_7$ $(GA)_{11}(TC)_2(CT)_3GT$	F: CTTGTTTGAGTCTGTCGCTG R: TTTCCCTCCCAATGTCTGTA
M24	AJ250254	$(CA)_{15}(CG)_4CA$	F: GGCTCGAGATATCTGTTTAG R: TTTAATGGGCATAGGGTCC
M25	AJ250255	$(GT)_5CT(GT)_2/(GT)_{12}$	F:CCCTCCCTGCCAGAAGAAGC R:AACCACCGTCCTTTTCCTCG
M29	AJ250257	$(CTCACA)_4/(CA)_9$	F: GACCATTACATTTCACACAC R: GCATTTTGTTGCACACTGTA
M32	AJ250258	$(CA)_3/(CA)_3/(CA)_{18}$	F: AACTCTCCATTCCCGCATTC R: CTGGGTTTTCTGTGTTCTCG
M47	AJ250260	$(CT)_9(CA)_8/(CT)_4/(CA)_5$	F: TGATGGACAGGAGTTGATGG R: TGCCAATCTACCTACCCCTT

7.3.3 Silver nitrate staining

The silver staining (Silver Sequence^{TM} DNA Sequencing System of Promega) process of the acrylamide gels included fixing the gel in 10% (v/v) acetic acid for 30 min, rinsing three times in de-ionised water (5 min per rinse), staining for 30 min in a solution containing 0.1% (w/v) sliver nitrate and 0.056% (v/v) formaldehyde. All the above steps were performed with slow agitation on a shaker. The stained gel was rinsed with de-ionised water for 5 s and immersed in a cold (4-10 ^{o}C) developing solution containing 3% (w/v) sodium carbonate, 0.056% (v/v) formaldehyde and 0.002 mg/ml sodium thiosulphate and agitated manually until DNA fragments became visible. To stop the development process, 10% acetic acid was added directly to the developing solution and shaking continued for two to three min. The gel was briefly rinsed with de-ionised water and left upright overnight to air dry. It was photographed by exposing photographic paper (Kodak Polymax II RC) directly under the gel to dim light for approximately 20 s.

7.3.4 Data analysis

DNA fragments were scored visually both from the gel as well as from the photograph. A fragment was considered polymorphic if it was present in at least one genotype and absent in the others. A binary data matrix was generated in which each fragment was scored as present (1) or absent (0). The NTSYS-pc software package version 2.02i (Rohlf, 1998) was used to calculate Jaccard`s similarity coefficients (Jaccard, 1908) among the genotypes as follows: JSC = a/(n-d) where a = the number of shared fragments between the two genotypes, n = the total number of amplified fragments and d = the number of missed fragments in both genotypes. A dendrogram was constructed from the similarity coefficient matrix using the UPGMA method of cluster analysis. Cophenetic correlation coefficients were calculated to test the goodness of fit of the clustering to the genetic similarity data. To investigate the discriminatory power of each SSR primer, the polymorphic information content (PIC) was calculated according to Smith et al. (2000) as follows: PIC = $1-\Sigma^{n}fi^{2}$ where fi is the frequency of the i^{th} allele in the set of 28 coffee genotypes. Observed heterozygosity (Ho) for each primer was calculated by dividing the number of heterozygous genotypes at a given locus with the total number of genotypes analysed (Hormaza, 2002).

7.4 Results

A total of 22 alleles were amplified among 28 arabica coffee genotypes using six SSR primers. Of these amplified alleles, 20 were polymorphic and two, M32-145 and M32-136, were specific to genotypes AD1491 and AD2991, respectively. The number of amplified alleles per primer varied from two for primers M25 and M29 to six for M24 and M32, with an average value of 3.7 alleles (Table 7.2). Sizes of amplified alleles ranged from 100 to 310 bp. M20 amplified larger size alleles compared to other primers (Fig. 7.1).

Table 7.2 Levels of genetic information generated by six SSR primers on 28 Ethiopian *C. arabica* genotypes

Primer	Allele size range (bp)	Number of alleles	PIC	Ho
M20	250-310	3	0.157	1.00
M24	162-182	6	0.750	0.07
M25	170-182	2	0.025	0.21
M29	103-122	2	0.132	0.00
M32	100-155	6	0.683	1.00
M47	106-123	3	0.004	0.46
Total		22		
Mean		3.7	0.292	0.32

PIC = polymorphic information content; Ho = observed heterozygosity

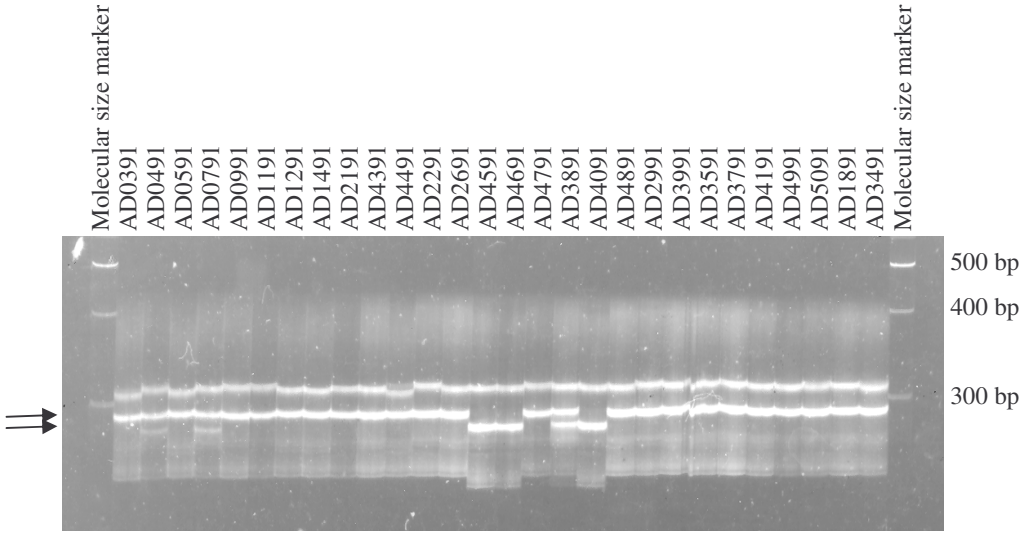


Fig. 7.1 Microsatellite alleles of M20 on 28 arabica coffee genotypes

The discriminative power of each SSR primer was assessed by calculating polymorphic information contents. The PIC per primer ranged from 0.004 to 0.750 with an average of 0.292 (Table 7.2). The lowest value was recorded for primer M47 and the highest for M24. Except for M24 and M32, other primers polymorphic information contents were less than 0.500. PIC values were positively correlated (r = 0.72) with number of amplified alleles per primer.

Observed heterozygosity (Ho) were calculated for each primer and ranged from 0.00 to 1.00 with an average of 0.32 (Table 7.2). The lowest Ho value was recorded for primer M29 while the highest value recorded for primers M20 and M32. Except for M20 and M32, the observed heterozygosity values of other primers were less than 0.50. Allele frequencies were low, particularly for primers with higher number of amplified alleles (Table 7.3). Only 20% of the polymorphic alleles had a frequency higher than 50%.

Table 7.3 Allele frequencies in 28 *C. arabica* genotypes using six SSR primers

Primer	Allele					
	1	2	3	4	5	6
M20	1.000	0.893	0.214			
M24	0.179	0.107	0.393	0.143	0.143	0.107
M25	0.214	0.964				
M29	0.929	0.071				
M32	0.357	0.036	0.393	0.036	0.179	1.000
M47	0.107	0.500	0.857			

Pair-wise genetic similarity coefficients among genotypes were calculated according to Jaccard (1908) and are presented in Table 7.4. Similarity coefficients ranged from 0.286 to 1.000 with an average of 0.560. The lowest pair-wise genetic similarity coefficient was recorded between genotypes AD1491 and AD0591 as well as between AD1491 and AD4591. The highest pair-wise genetic similarity coefficient (1.000) was recorded among 10 genotypes. Based on average genetic similarity coefficients AD1491 (0.434) and AD4491 (0.444) were the most dissimilar while AD3491 (0.654) was the most similar genotype to all genotypes.

Table 7.4 Pair-wise genetic similarity coefficient matrix for 28 *C. arabica* genotypes calculated from SSR analysis data

Genotype	AD0391	AD0491	AD0591	AD0791	AD0991	AD1191	AD1491	AD1291	AD1891	AD2191	AD2291	AD2691	AD2991	AD3491
AD0491	0.429													
AD0591	0.700	0.357												
AD0791	0.583	0.750	0.385											
AD0991	0.500	0.539	0.546	0.462										
AD1191	0.546	0.462	0.600	0.385	0.700									
AD1491	0.462	0.500	0.286	0.539	0.462	0.385								
AD1291	0.546	0.462	0.778	0.500	0.700	0.600	0.636							
AD1891	0.385	0.539	0.417	0.462	0.636	0.546	0.462	0.546						
AD2191	0.417	0.462	0.455	0.385	0.546	0.600	0.385	0.455	0.889					
AD2291	0.500	0.546	0.500	0.455	0.455	0.667	0.333	0.364	0.455	0.500				
AD2691	0.417	0.462	0.455	0.385	0.546	0.600	0.385	0.455	0.889	1.000	0.500			
AD2991	0.417	0.462	0.455	0.385	0.700	0.600	0.385	0.455	0.546	0.600	0.500	0.600		
AD3491	0.546	0.500	0.600	0.417	0.778	0.778	0.417	0.667	0.778	0.778	0.500	0.778	0.600	
AD3591	0.417	0.462	0.455	0.385	0.546	0.600	0.385	0.455	0.889	1.000	0.500	1.000	0.600	0.778
AD3791	0.385	0.539	0.417	0.462	0.636	0.546	0.462	0.546	1.000	0.889	0.455	0.889	0.546	0.778
AD3891	0.357	0.615	0.385	0.539	0.583	0.500	0.429	0.500	0.900	0.800	0.417	0.800	0.500	0.700
AD3991	0.385	0.539	0.417	0.462	0.636	0.546	0.462	0.546	1.000	0.889	0.455	0.889	0.546	0.778
AD4091	0.385	0.539	0.417	0.462	0.800	0.546	0.357	0.546	0.500	0.417	0.333	0.417	0.546	0.600
AD4191	0.417	0.583	0.455	0.385	0.546	0.600	0.385	0.455	0.700	0.778	0.500	0.778	0.600	0.600
AD4391	0.417	0.462	0.455	0.385	0.546	0.600	0.385	0.455	0.889	1.000	0.500	1.000	0.600	0.778
AD4491	0.539	0.467	0.462	0.500	0.539	0.462	0.750	0.583	0.429	0.357	0.308	0.357	0.357	0.500
AD4591	0.417	0.462	0.455	0.385	0.700	0.600	0.286	0.455	0.417	0.455	0.364	0.455	0.600	0.600
AD4691	0.500	0.539	0.417	0.583	0.636	0.546	0.462	0.546	0.500	0.417	0.333	0.417	0.417	0.600
AD4791	0.546	0.462	0.600	0.385	0.700	1.000	0.385	0.600	0.546	0.600	0.667	0.600	0.600	0.778
AD4891	0.385	0.539	0.417	0.462	0.636	0.546	0.462	0.546	1.000	0.889	0.455	0.889	0.546	0.778
AD4991	0.500	0.818	0.417	0.583	0.500	0.546	0.462	0.417	0.500	0.546	0.667	0.546	0.546	0.546
AD5091	0.357	0.615	0.385	0.539	0.583	0.636	0.429	0.500	0.727	0.636	0.700	0.636	0.500	0.700

Table 7.4 (Contd.)

Genotype	AD3591	AD3791	AD3891	AD3991	AD4091	AD4191	AD4391	AD4491	AD4591	AD4691	AD4791	AD4891	AD4991
AD3791	0.889												
AD3891	0.800	0.900											
AD3991	0.889	1.000	0.900										
AD4091	0.417	0.500	0.583	0.500									
AD4191	0.778	0.700	0.636	0.700	0.417								
AD4391	1.000	0.889	0.800	0.889	0.417	0.778							
AD4491	0.357	0.429	0.400	0.429	0.429	0.357	0.357						
AD4591	0.455	0.417	0.500	0.417	0.889	0.455	0.455	0.357					
AD4691	0.417	0.500	0.583	0.500	0.800	0.417	0.417	0.539	0.700				
AD4791	0.600	0.546	0.500	0.546	0.546	0.600	0.600	0.462	0.600	0.546			
AD4891	0.889	1.000	0.900	1.000	0.500	0.700	0.889	0.429	0.417	0.500	0.546		
AD4991	0.546	0.500	0.462	0.500	0.385	0.700	0.546	0.429	0.417	0.385	0.546	0.500	
AD5091	0.636	0.727	0.667	0.727	0.462	0.500	0.636	0.400	0.385	0.462	0.636	0.727	0.583

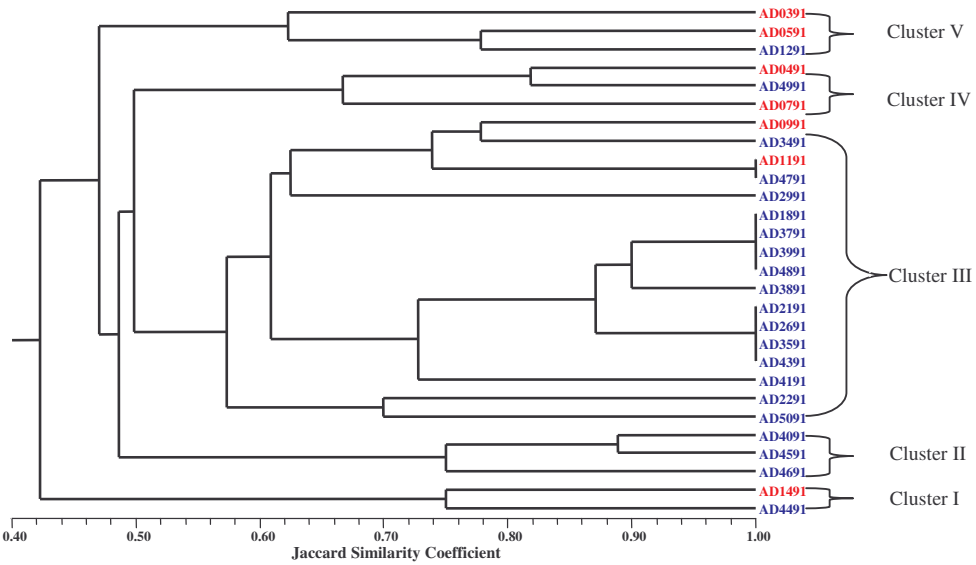
The genetic similarity of coffee genotypes collected in each region was assessed independently. Coffee genotypes collected from northwestern Ethiopia had pair-wise genetic similarity coefficients ranging from 0.308 to 1.000 with an average of 0.603 (Table 7.5). Pair-wise genetic similarity coefficients among coffee genotypes collected from southwestern Ethiopia ranged from 0.286 to 0.750 with an average of 0.503.

Of the 20 polymorphic alleles amplified in all 28 genotypes, two were monomorphic and three were missing across the seven coffee genotypes collected from southwestern Ethiopia. Only one allele was missing in the 21 coffee genotypes collected from northwestern Ethiopia. Genotypes from northwestern Ethiopia contained 95.5% of the total amplified alleles.

Table 7.5 Genetic relationships among *C. arabica* genotypes of each region

Regions	Number of genotypes	Total number of alleles	Number of polymorphic alleles	Genetic similarity coefficient		
				Minimum	Maximum	Mean
Northwest	21	21	19	0.308	1.000	0.603
Southwest	7	19	15	0.286	0.750	0.503
Both	28	22	20	0.286	1.000	0.560

The pattern of genetic relationships among genotypes was assessed using UPGMA method of cluster analysis. Genotypes were classified into five main clusters (Fig.7.2). Except for the second main cluster (Cluster II, Fig.7.2), all other main clusters contained coffee genotypes collected from both northwestern and southwestern Ethiopia. Although genotypes were not clustered according to collection region, some genotypes clustered according to collection locality. For example genotypes AD0991 and AD1191 were collected from the same locality in southwestern Ethiopia and were assigned in the same cluster (Cluster III, Fig.7.2). Similarly, genotypes AD4091, AD4591 and AD4691 were collected from Finoteselam area, in northwestern Ethiopia and were assigned to cluster II. The cophenetic correlation coefficient was 0.89, suggesting the presence of a good fit between the dendrogram and the similarity matrices.



Blue for northwestern Ethiopia and red for southwestern Ethiopia genotypes

Fig. 7.2 Dendrogram of 28 *C. arabica* genotypes constructed by cluster analysis using 22 SSR markers

7.5 Discussion

The total number of amplified alleles (22) as well as the average number of amplified alleles per primer (3.7) was relatively lower compared to previous results. Anthony et al. (2002) reported a total of 28 and an average of 4.7 amplified alleles per primer among 26 coffee accessions using the same (six) SSR primers. Recently, Moncada and McCouch (2004) reported an average of 2.5 amplified alleles per primer among 11 wild arabica coffee genotypes using 34 SSR primers. These differences could be attributed to differences in genotypes as well as SSR primers. Anthony et al. (2002) used more diverse genotypes from different sources compared to the genotypes included in the present study. The average Ho value in this study was 0.32, higher compared to the 0.04 reported by Combes et al. (2000) using 11 SSR primers. Only previously reported polymorphic SSR primers were employed in the present study. The low average PIC value of 0.292 indicated the presence of high genetic similarity among genotypes and entails development of additional polymorphic SSR primers for

effective characterisation of *C. arabica* genotypes. Moncada and McCouch (2004) reported an average PIC value of 0.35 among 11 arabica coffee genotypes.

Ten of the 28 coffee genotypes (35.7%) could not be independently distinguished using 20 polymorphic alleles. Similarly, Anthony et al. (2002) distinguished 52.6% of the genotypes analysed using the same six SSR primers. This illustrated the inadequacy of the available SSR primers to scan different parts of the genome and to pinpoint genetic differences between coffee genotypes. Hence, a wider range of informative SSR primers need to be developed for successful fingerprinting of *C. arabica* genotypes.

Two genotypes, AD1491 and AD2991, had unique alleles which could be exploited for genotype fingerprinting. AD1491 differed from all other genotypes since it is adaptable to hot, humid and low altitude areas. Genotype AD2991 was distinguished from all other genotypes since it regularly bears berries on the nodes of its orthotropic branch. Assessment of the association between these specific alleles and the above mentioned peculiar morphological characteristics could be considered as an area of investigation for the future.

The lowest genetic similarity coefficient (0.286) was recorded between genotypes, AD1491 and AD0591 as well as AD1491 and AD4591. Genotype AD1491 is rust resistant, has long beans and green young leaves. AD0591 was collected from a mid altitude area in southwestern Ethiopia and characterised by a vigorous growth habit, average cup quality and green young leaves. AD4591 was collected from a mid altitude area of Finoteselam, in northwestern Ethiopia and characterised by bronze young leaves. These three genotypes could be important for studying the association between heterosis and genetic distance in arabica coffee in the future. Similarly, Lashermes et al. (1996b) suggested the possibility of employing RAPD- based genetic distance measures for predicting hybrid performance in arabica coffee.

The genetic relationship of coffee genotypes collected in each region was assessed independently. Among coffee genotypes collected from southwestern Ethiopia, the lowest and highest pair-wise genetic similarity coefficients were recorded between genotypes AD0491 and AD0791 and between AD0591 and AD1491, respectively.

Among northwestern Ethiopia coffee genotypes, AD2191, AD3591, AD4191 and AD4391 as well as, AD1891, AD3791, AD3991 and AD4891 were genetically identical, with a pair-wise genetic similarity coefficient of 1.000. Genotypes AD2291 and AD4491 as well as AD2291 and AD4691 were the most dissimilar, with a pair-wise genetic similarity coefficients of 0.308. In both regions the lowest genetic similarity coefficients were recorded between genotypes collected from different localities. The highest genetic similarity coefficients were not recorded between coffee genotypes collected from the same locality. This indicated the availability of genetic variability among coffee genotypes collected from the same locality. Moreover, it indicated the inability of collection locality information to indicate the genetic relationship of coffee genotypes.

Results from this study indicated the presence of relatively higher genetic diversity among coffee genotypes from southwestern Ethiopia compared to those from northwestern Ethiopia. This could be since southwestern Ethiopia was identified as the centre of origin and genetic diversity for *C. arabica* (Sylvain, 1955; Steiger et al., 2002).

The UPGMA method of cluster analysis classified genotypes into five main clusters. The first main cluster contained genotypes AD4491 and AD1491. AD1491 and AD4491 were genetically the most divergent genotypes from all other genotypes with average genetic similarity coefficients of 0.434 and 0.444, respectively. The three genotypes assigned to the second main cluster (AD4091, AD4591 and AD4691) were collected from the same locality in northwestern Ethiopia. Genotypes AD4591 and AD4691 had bronze young leaves and were collected from the same orchard.

The third main cluster contained 17 coffee genotypes and was further divided into three sub-clusters. The first sub-cluster contained two coffee genotypes, AD2291 and AD5091, collected in relatively high altitude areas and are characterised by green young leaves. Ten genotypes were assigned to the second sub-cluster and it was impossible to independently distinguish most of these genotypes. These genotypes were characterised with green young leaves and low green bean chlorogenic acids content. Genotypes AD0991, AD1191, AD2991, AD3491 and AD4791 were assigned in the third sub-cluster of the third main cluster. These genotypes were characterised

by dwarf and relatively compact growth habits and high green bean chlorogenic acids content. Most of them were collected from low to mid altitude areas and have green young leaves. AD0991 and AD1191 were collected from the same locality in southwestern Ethiopia, are morphologically similar, CBD resistant and have small size beans and leaves.

The fourth main cluster contained three genotypes namely AD0491, AD0791 and AD4991, characterised by green young leaves, average green bean chlorogenic acids content and adaptation to mid and high altitude areas. Genotypes AD0491 and AD0791 were collected from southwestern Ethiopia and were CBD resistant. Three genotypes, AD0391, AD0591 and AD1291 were assigned to the fifth main cluster. These genotypes were characterised by vigorous growth habits and adaptation to mid altitude areas. Most of them had high green bean chlorogenic acids content. AD1291 was collected from northwestern Ethiopia and was characterised by bronze young leaves and better adaptation to relatively marginal areas for coffee production. Except for the second main cluster, all clusters contained coffee genotypes collected from both regions. This indicated the presence of genetic similarity between some coffee genotypes from the two regions. Moreover, this study demonstrated the presence of genetic variation among coffee genotypes both within and between regions.

7.6 Conclusions

Results of the present study demonstrated the presence of genetic variation among coffee genotypes within and between regions. Genotypes collected from southwestern Ethiopia were more diverse than those collected from the northwest. Three alleles detected among coffee genotypes from northwestern Ethiopia were not found among coffee genotypes collected in southwestern Ethiopia. Coffee germplasm collection and conservation attempts should address both regions. Currently, farmers from northwestern Ethiopia are replacing their coffee genotypes with the above listed few, high yielding and CBD resistant coffee genotypes of southwestern Ethiopia. This practice will result in considerable genetic erosion unless it is supported by adequate germplasm collection and conservation attempts. The six SSR primers independently distinguished 18 out of the 28 genotypes evaluated in this study. Hence, further analysis using large numbers of SSR primers is suggested. M24 and M32 were the

most efficient SSR primers in detecting genetic variation among genotypes. The lowest pair-wise genetic similarity coefficient was recorded between genotypes AD0591 and AD1491 as well as AD1491 and AD4591. Thus, these genotypes will be used to study the association between genetic distance and heterosis in arabica coffee. Generally, results of this study will be helpful to design crosses among these genotypes as well as for the conservation of these genetic resources.

CHAPTER 8

COMPARISON OF AFLP AND SSR MARKERS FOR COFFEE GENETIC DIVERSITY ANALYSIS

8.1 Abstract

The genetic relationship of 28 arabica coffee genotypes was evaluated using 10 AFLP primer combinations and six SSR primer pairs. The objectives of the study were to compare the efficiency of AFLP and SSR markers in detecting genetic variation among coffee genotypes and estimate the genetic similarity of coffee genotypes using combined data of AFLP and SSR analyses. Jaccard`s similarity coefficients were calculated and dendrograms were constructed using UPGMA method of cluster analysis with NTSYS-pc software package. AFLP and SSR markers were positively and significantly correlated (0.217) in estimating genetic similarity among coffee genotypes. The average genetic similarity coefficient calculated using SSR markers was much lower (0.560 with a range of 0.286-1.000) compared to AFLP markers (0.915 with a range of 0.860-0.982). This indicated the higher information content generated by SSR markers. AFLP markers distinguished all coffee genotypes analysed, while SSR markers distinguished 64.3% of the genotypes. The two dendrograms constructed using only AFLP and only SSR analyses data were different. On the other hand, dendrograms constructed using only AFLP data and combined data of AFLP and SSR analyses, were similar. Therefore, AFLP markers were more efficient compared to SSR markers for characterisation of evaluated coffee genotypes.

8.2 Introduction

The precision of genetic diversity estimates depends upon the method used. In the past, relationships among genotypes were estimated based on pedigree records, morphological, physiological or biochemical markers. Currently, different molecular marker techniques are developed and available for measuring genetic diversity. These methods detect polymorphism by assaying subsets of the total amount of DNA sequence variation in a genome. They differ in principle, application, type and amount of polymorphism detected and cost, technical skill and time requirements (Karp et al., 1998). Consequently, comparison of the efficiency of different types of molecular markers seems vital and have been carried out by several investigators on a number of plant species (Powell et al., 1996; Milbourne et al., 1997; Russell et al., 1997; Ajmone-Marsan et al., 1998; Das et al., 1999; Garcia-Mas et al., 2000; Maguire et al., 2002; Simioniuc et al., 2002; Uptmoor et al., 2003; Bandelj et al., 2004; Paczos-Grzeda, 2004; Saini et al., 2004; Fossati et al., 2005).

The genetic relationship of different *C. arabica* genotypes was assessed using isozymes (Louarn, 1978 as cited in Ruas et al., 2000), RAPD (Orozco-Castillo et al., 1994; Lashermes et al., 1996b; Anthony et al., 2001; Aga et al., 2003), AFLP (Anthony et al., 2002; Steiger et al., 2002) and SSR (Combes et al., 2000; Rovelli et al., 2000; Anthony et al., 2002; Moncada and McCouch, 2004) markers. Although comparison among molecular marker techniques is important for marker selection and was made on a number of plant species, such type of studies are rare on arabica coffee. Therefore, this study was conducted to estimate the level of genetic diversity among coffee genotypes using combined data of AFLP and SSR markers and to compare AFLP and SSR markers in detecting genetic variation among arabica coffee genotypes.

8.3 Materials and Methods

8.3.1 Genotypes

Twenty-eight arabica coffee genotypes collected from the northwestern and southwestern parts of Ethiopia were used for this study. Genotypes were obtained from Adet Agricultural Research Centre of Amhara Regional Agricultural Research

Institute, Ethiopia. Their description and collection region are given in section 6.3.1 and Table 6.1.

8.3.2 DNA isolation, AFLP and SSR analyses protocols

Genomic DNA was extracted from lyophilized leaves as described in section 6.3.2. Ten AFLP primer combinations and six SSR primer pairs were used and descriptions are given in Tables 6.2 and 7.1, respectively. Protocols for AFLP and SSR analyses are described in sections 6.3.3 and 7.3.2, respectively.

8.3.3 Data analysis

Data matrices for AFLP and SSR analyses were constructed as described in sections 6.3.4 and 7.3.4. Pair-wise genetic similarity coefficients were calculated and dendrograms constructed using NTSYS-pc version 2.02i (Rohlf, 1998) as described in sections 6.3.4 and 7.3.4. From these similarity coefficient matrices, means and variances were calculated. To determine the efficiency of each marker type in detecting genetic variation, the assay efficiency index, AEI (Pejic et al., 1998) (AEI = BP/T, where BP is the total number of polymorphic fragments detected and T is the number of polymorphic primer pairs or primer combinations) and percentage of polymorphic fragments were calculated. Correlations of genetic similarities estimated from AFLP, SSR and combined AFLP and SSR (AFLP+SSR) data were assessed by Spearman`s rank correlation analysis using Number Cruncher Statistical System, NCSS 2000 (Hintze, 1998).

8.4 Results

The two molecular marker techniques differed in the amount of information and polymorphism generated (Table 8.1). The 10 AFLP primer combinations amplified a total of 712 fragments, of which 220 (30.9%) were polymorphic across the 28 arabica coffee genotypes. All 10 primer combinations detected polymorphic fragments with a range of 10 (*Eco*RI-ACC/*Mse*I-CTG) to 39 (*Eco*RI-AAC/*Mse*I-CAA) and an average of 22. The six SSR primer pairs amplified 22 alleles among 28 coffee genotypes (average of 3.7 alleles per primer pair), of which 20 (90.9%) were polymorphic (Table 8.1). As expected, the assay efficiency index of AFLPs was far superior to that of

SSRs (AEI = 22 vs. 3.3). AFLP analysis detected 18 specific fragments for 10 genotypes while SSR analysis identified two specific fragments for two genotypes (Table 8.1). AFLP analysis independently distinguished all 28 coffee genotypes while SSR analysis distinguished 18 genotypes.

Table 8.1 Levels of genetic information generated by AFLP and SSR markers on 28 Ethiopian *C. arabica* genotypes

Parameters	AFLP	SSR
Number of assay units	10.0	6.0
Total number of fragments	712.0	22.0
Number of polymorphic fragments	220.0	20.0
Percentage of polymorphic fragments	30.9	90.9
Mean number of fragments per assay units	71.2	3.7
Assay efficiency index (AEI)	22.0	3.3
Number of genotypes distinguished	28.0	18.0
Number of genotypes with specific band(s)	10.0	2.0

The minimum, maximum and mean genetic similarity estimates between northwestern Ethiopia, southwestern Ethiopia and all genotypes for AFLP, SSR and combined data are given in Table 8.2. Pair-wise genetic similarity coefficients ranged from 0.286 within southwestern Ethiopia to 1.000 within northwestern Ethiopia genotypes using SSR markers. Estimates of similarity coefficients followed the same pattern for all marker types, i.e. higher similarity estimates were observed among northwestern Ethiopia genotypes (means: AFLP = 0.918, AFLP+SSR = 0.913, SSR = 0.603) and lower similarity estimates among southwestern Ethiopia genotypes (means: AFLP = 0.912, AFLP+SSR = 0.904, SSR = 0.503). SSR markers estimated the lowest average similarity value (0.560), AFLP markers the highest (0.915) and AFLP+SSR markers in between (0.909) among all genotypes.

Table 8.2 Minimum, maximum and mean pair-wise genetic similarity coefficients calculated from AFLP, SSR and AFLP+SSR data

Marker		Northwestern Ethiopia Genotypes (21)	Southwestern Ethiopia genotypes (7)	All genotypes (28)
AFLP	Minimum	0.877	0.861	0.860
	Maximum	0.982	0.956	0.982
	Mean	0.918	0.912	0.915
SSR	Minimum	0.318	0.286	0.286
	Maximum	1.000	0.750	1.000
	Mean	0.603	0.503	0.560
AFLP+SSR	Minimum	0.867	0.853	0.842
	Maximum	0.980	0.947	0.980
	Mean	0.913	0.904	0.909

AFLP+SSR = combined AFLP and SSR data

The minimum, maximum, mean and variance of pair-wise genetic similarity coefficients of each genotype for AFLP, SSR and AFLP+SSR analyses are given in Table 8.3. According to AFLP data, AD1491 had the lowest (0.870) and AD5091 the highest (0.930) average genetic similarity coefficients compared to all other genotypes. Using SSR data, AD1491 had the lowest (0.434) and AD3491 the highest (0.654) average genetic similarity coefficients compared to all other genotypes. Genotypes AD1491, AD2991, AD4091, AD4591 and AD4691 had less variable pair-wise genetic similarity coefficients calculated from AFLP data while genotypes AD2291, AD3991 and AD3591 had more variable pair-wise genetic similarity coefficients. According to SSR data, genotypes AD0391, AD2991, AD0791, AD1291 and AD4491 had less variable while genotypes AD2191, AD2691 and AD3591 had more variable pair-wise genetic similarity coefficients. AD1491 had the less variable and the lowest average genetic similarity coefficient compared to all other genotypes in both marker systems. On the other hand, AD5091 and AD3491 had less variable

and the highest average genetic similarity coefficients using AFLP and SSR data, respectively.

Table 8.3 Minimum, maximum, mean and variance of pair-wise genetic similarity coefficients of each genotype calculated from AFLP, SSR and AFLP+SSR data

Genotype	AFLP				SSR				AFLP+SSR			
	Min	Max	Mean	Var	Min	Max	Mean	Var	Min	Max	Mean	Var
AD0391	0.869	0.957	0.921	0.0004	0.357	0.700	0.461	0.0069	0.861	0.950	0.913	0.0004
AD0491	0.871	0.956	0.915	0.0005	0.357	0.818	0.522	0.0091	0.863	0.943	0.908	0.0005
AD0591	0.875	0.960	0.923	0.0004	0.286	0.778	0.470	0.0113	0.863	0.949	0.915	0.0004
AD0791	0.871	0.946	0.916	0.0004	0.385	0.750	0.467	0.0080	0.865	0.943	0.907	0.0004
AD0991	0.861	0.955	0.920	0.0003	0.455	0.800	0.598	0.0090	0.854	0.947	0.915	0.0003
AD1191	0.863	0.951	0.915	0.0004	0.385	1.000	0.583	0.0140	0.853	0.941	0.909	0.0004
AD1491	0.851	0.898	0.870	0.0001	0.286	0.750	0.434	0.0093	0.842	0.887	0.861	0.0001
AD1291	0.851	0.951	0.909	0.0003	0.364	0.778	0.530	0.0086	0.842	0.947	0.903	0.0003
AD1891	0.862	0.941	0.910	0.0003	0.385	1.000	0.650	0.0440	0.854	0.937	0.905	0.0003
AD2191	0.865	0.960	0.916	0.0005	0.357	1.000	0.635	0.0476	0.856	0.959	0.911	0.0006
AD2291	0.860	0.959	0.906	0.0007	0.308	0.700	0.479	0.0112	0.851	0.952	0.899	0.0007
AD2691	0.883	0.965	0.917	0.0004	0.357	1.000	0.635	0.0476	0.874	0.966	0.913	0.0005
AD2991	0.875	0.934	0.914	0.0002	0.357	0.700	0.526	0.0073	0.865	0.929	0.908	0.0002
AD3491	0.858	0.951	0.907	0.0005	0.417	0.778	0.654	0.0153	0.851	0.949	0.903	0.0005
AD3591	0.868	0.982	0.924	0.0006	0.357	1.000	0.635	0.0476	0.859	0.980	0.919	0.0006
AD3791	0.879	0.951	0.915	0.0004	0.385	1.000	0.650	0.0440	0.870	0.950	0.910	0.0005
AD3891	0.883	0.965	0.923	0.0003	0.357	0.900	0.617	0.0312	0.873	0.962	0.917	0.0004
AD3991	0.860	0.982	0.924	0.0007	0.385	1.000	0.650	0.0440	0.852	0.980	0.919	0.0007
AD4091	0.891	0.950	0.925	0.0002	0.333	0.889	0.508	0.0185	0.880	0.941	0.917	0.0002
AD4191	0.868	0.946	0.921	0.0004	0.357	0.778	0.575	0.0192	0.859	0.944	0.915	0.0004
AD4391	0.898	0.965	0.923	0.0003	0.357	1.000	0.635	0.0476	0.887	0.966	0.918	0.0004
AD4491	0.864	0.951	0.911	0.0005	0.308	0.750	0.444	0.0086	0.862	0.941	0.901	0.0005
AD4591	0.861	0.926	0.903	0.0002	0.286	0.889	0.485	0.0167	0.849	0.921	0.896	0.0002
AD4691	0.872	0.941	0.910	0.0002	0.333	0.800	0.506	0.0105	0.864	0.935	0.902	0.0002
AD4791	0.863	0.963	0.918	0.0004	0.385	1.000	0.583	0.0140	0.854	0.956	0.913	0.0005
AD4891	0.860	0.963	0.918	0.0005	0.385	1.000	0.650	0.0440	0.853	0.959	0.913	0.0005
AD4991	0.883	0.947	0.922	0.0003	0.385	0.818	0.522	0.0093	0.875	0.940	0.915	0.0003
AD5091	0.881	0.962	0.930	0.0003	0.357	0.727	0.574	0.0143	0.872	0.958	0.924	0.0003

Min = minimum; Max = maximum; Var = variance; AFLP+SSR = combined AFLP and SSR analyses data

Correspondence of similarities measured using AFLP, SSR and AFLP+SSR data were evaluated using Spearman`s rank correlation approach and correlation coefficients are given in Table 8.4. All approaches were positively and significantly correlated in

estimating genetic similarity among genotypes. Correlation coefficients ranged from 0.217 (between AFLP and SSR) to 0.986 (between AFLP and AFLP+SSR).

Table 8.4 Correlation coefficients among genetic similarity coefficients calculated using AFLP, SSR and AFLP+SSR data

Marker	SSR	AFLP+SSR
AFLP	0.217**	0.986**
SSR		0.350**

** indicates significant correlation at $p \leq 0.01$; AFLP+SSR = combined data

Pair-wise genetic similarity coefficients among genotypes were calculated using combined data of AFLP and SSR analyses and the result is given in Table 8.5. Pair-wise genetic similarity coefficients among genotypes for AFLP data is given in Table 6.6 and for SSR data in Table 7.4. Pair-wise genetic similarity coefficients ranged from 0.842 to 0.980 with an average of 0.909. This value is higher than the value calculated using SSR data (section 7.4) but lower than the value from AFLP data (section 6.4). AD1291 and AD1491 were the most dissimilar while AD3591 and AD3991 were the most similar genotypes of all genotypes evaluated.

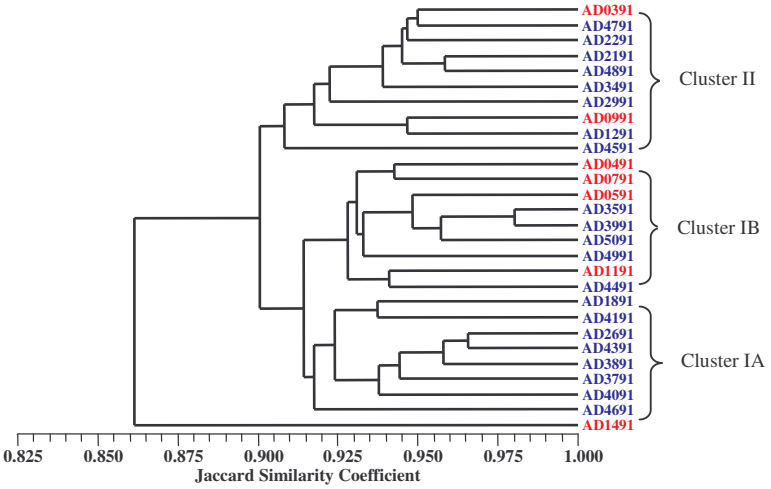
Three dendrograms (Fig.8.1a-c) were constructed using genetic similarity coefficients calculated from AFLP, SSR and a combination of AFLP and SSR data. Cophenetic coefficients were high for all approaches (AFLP = 0.87, SSR = 0.89 and AFLP+SSR = 0.87), indicating a good fit of the original data to clustering. Cluster analysis performed using AFLP and AFLP+SSR data grouped the 28 coffee genotypes into two main clusters and a singleton. AD1491 was the singleton in both cases. Genotypes assigned to each cluster and sub-clusters were similar (Fig. 8.1a and b). All genotypes were distinguished using either AFLP data alone or a combination of AFLP and SSR data. On the other hand, cluster analysis performed using SSR data classified genotypes into five clusters. Ten genotypes in three groups were not independently distinguished (Fig. 8.1c). Comparing all three approaches, most clusters contained coffee genotypes from both regions. Therefore, genotypes were not clustered according to collection regions.

Table 8.5 Pair-wise genetic similarity coefficient matrix for 28 *C. arabica* genotypes calculated from AFLP+SSR data

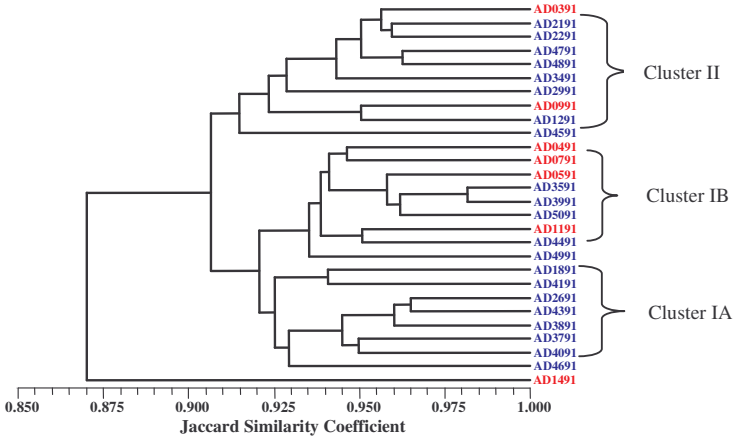
Genotype	AD0391	AD0491	AD0591	AD0791	AD0991	AD1191	AD1491	AD1291	AD1891	AD2191	AD2291	AD2691	AD2991	AD3491
AD0491	0.901													
AD0591	0.909	0.943												
AD0791	0.903	0.943	0.925											
AD0991	0.947	0.914	0.922	0.913										
AD1191	0.912	0.928	0.925	0.928	0.922									
AD1491	0.861	0.863	0.863	0.865	0.854	0.853								
AD1291	0.919	0.901	0.918	0.901	0.947	0.895	0.842							
AD1891	0.896	0.895	0.909	0.911	0.903	0.914	0.854	0.893						
AD2191	0.947	0.883	0.894	0.885	0.927	0.898	0.856	0.912	0.908					
AD2291	0.950	0.875	0.887	0.874	0.914	0.883	0.851	0.913	0.876	0.952				
AD2691	0.898	0.903	0.914	0.897	0.894	0.899	0.874	0.882	0.912	0.911	0.887			
AD2991	0.916	0.886	0.903	0.894	0.914	0.888	0.865	0.907	0.892	0.925	0.921	0.915		
AD3491	0.930	0.888	0.897	0.885	0.918	0.894	0.851	0.900	0.885	0.940	0.935	0.898	0.929	
AD3591	0.900	0.930	0.947	0.927	0.915	0.935	0.859	0.911	0.928	0.909	0.886	0.922	0.896	0.888
AD3791	0.892	0.903	0.911	0.902	0.894	0.896	0.870	0.884	0.921	0.905	0.883	0.950	0.912	0.887
AD3891	0.912	0.909	0.914	0.908	0.914	0.908	0.873	0.904	0.921	0.922	0.898	0.953	0.915	0.896
AD3991	0.901	0.935	0.949	0.931	0.917	0.934	0.852	0.907	0.929	0.908	0.882	0.926	0.895	0.888
AD4091	0.913	0.906	0.926	0.914	0.914	0.914	0.880	0.910	0.915	0.905	0.898	0.933	0.924	0.902
AD4191	0.909	0.913	0.922	0.916	0.913	0.921	0.859	0.909	0.937	0.913	0.888	0.931	0.911	0.885
AD4391	0.907	0.906	0.923	0.908	0.905	0.911	0.887	0.892	0.921	0.916	0.892	0.966	0.917	0.903
AD4491	0.900	0.930	0.930	0.924	0.915	0.941	0.862	0.903	0.891	0.879	0.876	0.885	0.883	0.892
AD4591	0.905	0.876	0.889	0.886	0.912	0.883	0.849	0.896	0.887	0.906	0.907	0.904	0.921	0.901
AD4691	0.903	0.908	0.908	0.902	0.907	0.907	0.864	0.903	0.894	0.887	0.883	0.920	0.917	0.887
AD4791	0.950	0.903	0.908	0.902	0.939	0.916	0.854	0.909	0.894	0.942	0.944	0.900	0.920	0.943
AD4891	0.937	0.893	0.901	0.892	0.920	0.897	0.853	0.908	0.913	0.959	0.936	0.916	0.924	0.949
AD4991	0.913	0.927	0.921	0.929	0.912	0.929	0.875	0.894	0.919	0.903	0.884	0.922	0.907	0.893
AD5091	0.913	0.935	0.949	0.929	0.926	0.925	0.872	0.922	0.915	0.914	0.907	0.927	0.909	0.911

Table 8.5 (Contd.)

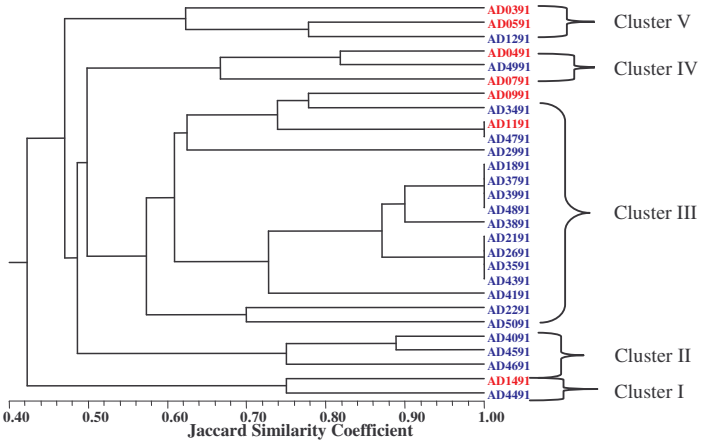
Genotype	AD3591	AD3791	AD3891	AD3991	AD4091	AD4191	AD4391	AD4491	AD4591	AD4691	AD4791	AD4891	AD4991
AD3791	0.927												
AD3891	0.922	0.935											
AD3991	0.980	0.929	0.926										
AD4091	0.925	0.941	0.938	0.924									
AD4191	0.944	0.934	0.928	0.943	0.926								
AD4391	0.930	0.947	0.962	0.929	0.938	0.934							
AD4491	0.932	0.890	0.888	0.930	0.905	0.901	0.890						
AD4591	0.891	0.895	0.901	0.887	0.913	0.888	0.898	0.867					
AD4691	0.898	0.911	0.935	0.900	0.923	0.919	0.923	0.884	0.894				
AD4791	0.910	0.891	0.908	0.908	0.909	0.899	0.908	0.902	0.915	0.893			
AD4891	0.911	0.904	0.924	0.915	0.907	0.909	0.921	0.889	0.913	0.886	0.956		
AD4991	0.940	0.924	0.921	0.933	0.933	0.929	0.927	0.909	0.908	0.906	0.909	0.908	
AD5091	0.957	0.926	0.929	0.958	0.933	0.934	0.935	0.936	0.893	0.908	0.917	0.921	0.939



a. AFLP+SSR data



b. AFLP data



c. SSR data

Blue for northwestern Ethiopia and red for southwestern Ethiopia genotypes

Fig. 8.1 Dendrograms of 28 *C. arabica* genotypes constructed by cluster analysis using combined AFLP and SSR (a), AFLP (b) and SSR (c) data

8.5 Discussion

Given the proliferation of molecular marker techniques, comparisons between these techniques seem inevitable. In this study, the efficiency of two molecular marker systems, AFLP and SSR, were compared using 28 arabica coffee genotypes. These systems differed both in technique and type and amount of polymorphism detected. The level of polymorphism was higher for SSR (90.9%) compared to AFLP (30.9%), indicating the hypervariability of SSR. Similar results were reported by Milbourne et al. (1997) for potato, Russell et al. (1997) for barley, Jakse et al. (2001) for hop, Almanza-Pinzon et al. (2003) for wheat, Uptmoor et al. (2003) for sorghum and Bandelj et al. (2004) for olive. Although AFLP did not offer the highest overall percentage of polymorphism, it was more efficient in revealing many polymorphic fragments per primer combination (higher multiplex ratio). The average number of amplified fragments per primer combination for AFLP was 71.2, compared to 3.7 for SSR. The efficiency of AFLP is associated with simultaneous analysis of large numbers of fragments rather than the overall percentage of polymorphism detected. This is consistent with findings of other studies (Powell et al., 1996; Russell et al., 1997; Pejic et al., 1998; Almanza-Pinzon et al., 2003; Uptmoor et al., 2003; Bandelj et al., 2004). This observation probably reflects two major differences between the two molecular marker types. Firstly, as mostly detecting dominant markers, AFLP detect two alleles per locus, which reduces the maximum level of heterozygosity to 0.5 even for polyploid crops like *C. arabica*. Secondly, SSR regions are well known to exhibit much higher level of mutation compared to other parts of the genome (Milbourne et al., 1997; Maguire et al., 2002). Therefore, SSR would generally reveal a higher percentage of polymorphism than AFLP.

Estimates of genetic similarities among 28 arabica coffee genotypes, based on AFLP markers, were higher (mean: 0.915) compared to those calculated for SSR (0.560). This low value is a reflection of the high information content provided by SSR markers and lead to the suggestion that SSR markers better differentiate pairs of genotypes that show a low level of genetic variation between them. Similar results have been reported by Russell et al. (1997) for barley, Uptmoor et al. (2003) for sorghum and Bandelj et al. (2004) for olive. Our finding was in agreement with previously established facts about these marker types. These two marker types

consistently identified the presence of much higher genetic diversity among coffee genotypes collected from southwestern Ethiopia compared to those from northwestern Ethiopia. Hence, both marker types are similar in detecting genetic variation among coffee genotypes.

Spearman`s rank correlation analysis indicated concordance of genetic similarity estimates using AFLP and SSR analyses. The correlation coefficient was positive and statistically significant, but low (0.217). This could be since a small number (6) of SSR primer pairs was used. Similar results were reported for mangrove species by Maguire et al. (2002) and for sorghum by Uptmoor et al. (2003). The correlation coefficient between genetic similarity coefficients calculated from AFLP and combined data was much higher (0.986) than the correlation coefficient calculated between genetic similarity coefficients estimated using SSR data and combined data (0.350). This could be since the 10 AFLP primer combinations contributed 32 fold information to the combined data compared to the six SSR primer pairs.

Cluster analysis using AFLP and SSR data resulted in different dendrograms for the 28 arabica coffee genotypes. Similar results were reported by Anthony et al. (2002) using 37 AFLP and six SSR primer combinations on coffee. Likewise, Bandelj et al. (2004) observed dissimilar dendrograms for the SSR and AFLP markers on olive. Dendrograms constructed during this study using AFLP analysis data and combined data were almost identical. In both dendrograms, genotypes were clustered into two main clusters and a singleton. AD1491 did not cluster with any genotype on both dendrograms, showing very low genetic similarities with all other genotypes. Dendrograms constructed using SSR markers and a combined data were different. AFLP as well as combined data distinguished all analysed genotypes. SSR markers failed to independently distinguish 10 of the genotypes. Similarly, Anthony et al. (2002) could not identify nine of the genotypes using the same six SSR primer pairs. This observation reflected the superiority of AFLP markers compared to the available SSR markers in detecting genetic variation among coffee genotypes. It also illustrated the danger of using a small number of SSR markers to differentiate genotypes in genetic diversity studies.

AFLP analysis covers large parts of the genome (almost the whole genome) whereas SSRs detect variation only at pre-determined loci/alleles, usually in the non-coding regions of the genome (Milbourne et al., 1997). AFLP will give a better coverage and better estimates of genetic diversity. On the other hand, SSRs are co-dominant markers, thus distinguishes heterozygous and homozygous individuals while AFLPs are mostly dominant markers, which could not usually distinguish the genetic difference present between homozygous and heterozygous individuals. Therefore, both marker types have advantages and disadvantages. Use of combined data gives more accurate results (Almanza-Pinzon et al., 2003).

For all approaches, genotypes did not cluster according to collection regions. This illustrated both the presence of diverse coffee genotypes in each region as well as close genetic similarity between some coffee genotypes collected from the two regions. This result together with knowledge on horticultural characteristics of each genotype may contribute to coffee genetic improvement as well as germplasm conservation programmes in northwestern Ethiopia.

8.6 Conclusions

The efficiency of AFLP and SSR markers were compared for detecting genetic variation among 28 arabica coffee genotypes collected from Ethiopia. The two marker systems were positively and significantly correlated in estimating genetic similarities among genotypes. Both marker systems similarly identified the presence of low levels of genetic variation among coffee genotypes collected from northwestern Ethiopia compared to those from southwestern Ethiopia. However, the study revealed the superiority of AFLP markers in detecting genetic variation among arabica coffee genotypes. This result could be in part attributed to the small number of SSR loci included in this study. Hence, identification of additional polymorphic SSR loci for arabica coffee is suggested as a future area of investigation. AFLPs are mostly dominant markers but scan almost the entiree genome whereas SSRs are co-dominant markers and detect variation only at a pre-determined site of the genome. Use of combined data of both marker systems gives more accurate genetic similarity estimates compared to individual marker techniques.

CHAPTER 9

CONCLUSIONS AND RECOMMENDATIONS

Diversity in genetic resources is the basis for genetic improvement. It is essential to fulfil the ever-changing needs of humans as well as to cope with unforeseen biotic and abiotic stresses. Genetic resources will have little value unless it is efficiently conserved and properly utilised. Its efficient utilisation as well as conservation depends on the availability of reliable genetic diversity information. In the present study, the genetic diversity of arabica coffee genotypes collected from the northwestern and southwestern parts of Ethiopia was evaluated using morphological characters, quality attributes, green bean biochemical composition and molecular markers. In addition, the correlation of different agro-morphological characters, coffee quality traits and green bean biochemical compounds was investigated.

Results of morphological diversity analysis demonstrated the presence of substantial variability among evaluated *C. arabica* genotypes for 27 agro-morphological characters. These genotypes may serve as sources of desirable genes for the genetic improvement of different characters of arabica coffee. Cluster analysis revealed the superiority of quantitative morphological characters compared to qualitative characters for characterisation of coffee genotypes. Use of a combination of qualitative and quantitative characters is recommended, since quantitative characters are more sensitive to environmental influences and the growth stage of the plant.

Correlation analysis among 18 quantitative agro-morphological characters revealed positive correlations between average green bean yield per tree and all other characters. Average green bean yield per tree showed statistically significant correlations with percentage of bearing primary branches, bean weight, canopy and trunk diameters, tree height, internode lengths of orthotropic and primary branches, bean length and thickness and fruit and petiole lengths. Tree height, canopy diameter and bean weight are recommended for indirect selection of average green bean yield per tree due to effective measurement and high heritability. Of all the agro-morphological characters evaluated in this study, number of secondary branches per

tree, canopy diameter, plant height, average green bean yield per tree and bean weight had high genetic coefficients of variation, broad sense heritability and genetic advance. Selection progress for these five traits should therefore be effective in these genotypes.

Results of cup quality and green bean physical characters variability studies indicated the presence of considerable variation among evaluated genotypes. Of all the genotypes tested, AD0691 and AD1691 had both desirable cup quality and green bean physical characters and are recommended as desirable gene sources for cup quality and green bean physical characters improvement. Cup quality and green bean physical character variation were observed among genotypes collected from the same as well as between different regions. This indicated the opportunity of genetic improvement of cup quality and green bean physical characters using genotypes collected from each region. Associations among cup quality attributes (acidity, body, flavour and overall standard of liquor) were positive and statistically significant. Correlations between cup quality attributes and green bean physical characters were positive but mostly statistically not significant. Therefore, green bean physical characters are not good indicators for cup quality improvement. Both desirable cup quality and green bean physical characters can be simultaneously selected due to positive correlation.

Green bean biochemical composition variability analysis demonstrated the presence of significant variation among evaluated coffee genotypes for green bean caffeine, chlorogenic acids, sucrose and trigonelline contents. Six genotypes namely AD0291, AD0591, AD2491, AD2691, AD2791 and AD2891 had caffeine contents less than 1.0% and are recommended as sources of genes for the development of coffee varieties with relatively low caffeine content. Genotypes were not clustered according to collection regions, demonstrating the presence of substantial green bean biochemical composition variation among genotypes collected within as well as between regions. Green bean caffeine content had negative and statistically significant correlations with all desirable cup quality attributes, indicating that low green bean caffeine content and desirable cup quality can be selected simultaneously. Green bean sucrose content had positive correlations with desirable cup quality attributes and green bean physical characters. Higher green bean sucrose content and desirable cup

quality and green bean physical characters can therefore be simultaneously selected in these *C. arabica* genotypes.

Ten AFLP primer combinations were evaluated for detecting genetic variation among *C. arabica* genotypes. Of these primer combinations, *Eco*RI-ACA/*Mse*I-CAA, *Eco*RI-ACC/*Mse*I-CAG, *Eco*RI-ACT/*Mse*I-CAG and *Eco*RI-AAC/*Mse*I-CAA were more efficient and recommended for arabica coffee genetic diversity analysis in the future. All evaluated coffee genotypes were independently distinguished and 18 AFLP markers which were specific to 10 genotypes were identified. These specific AFLP markers can be used for fingerprinting of these 10 arabica coffee genotypes. Conversion of these specific AFLP markers to SCAR markers is suggested as a future area of work. Similar to cup quality and green bean biochemical compounds, AFLP markers did not classify coffee genotypes according to collection regions. AFLP analysis demonstrated the presence of genetic variation among coffee genotypes collected from the same region and the presence of close genetic similarity between some coffee genotypes collected from different regions. Generally, results of this study indicated the efficiency of AFLP analysis for genetic diversity analysis as well as for cultivar fingerprinting in *C. arabica*.

SSR based genetic diversity analysis indicated the presence of genetic variation among some coffee genotypes collected in the same as well as different regions. Higher levels of genetic diversity were detected among coffee genotypes collected from southwestern Ethiopia compared to northwestern Ethiopia. Three alleles linked to northwestern Ethiopia genotypes were identified. Replacement of northwestern Ethiopia coffee genotypes with a few southwestern Ethiopia coffee genotypes, as it is currently practiced, without any conservation attempt, seems hazardous and may result in considerable genetic erosion. Ten of the evaluated coffee genotypes were not independently distinguished using SSR markers. This could be due to the small number of polymorphic SSR primers used. The number of polymorphic SSR primers currently available for characterisation of *C. arabica* is limited. Development of additional polymorphic SSR primers for effective characterisation of arabica coffee genotypes is suggested as a future research area.

Among AFLP and SSR markers, AFLP markers were superior in detecting genetic variation among arabica coffee genotypes. AFLPs are mostly dominant markers indicating the inability to detect genetic differences present between homozygous and heterozygous individuals. SSR markers detect genetic variation at specific sites of the genome and do not consider the entire genome while analysing genetic diversity among genotypes. Because both these methods have advantages and disadvantages the use of a combination of AFLP and SSR markers is a more reliable approach for studying genetic diversity.

All analyses demonstrated the presence of genetic variation among genotypes. Methods were similar but not identical in clustering genotypes into groups. Use of combined data from different diversity assessment methods is recommended to generate more reliable information. All approaches detected genetic variation among genotypes collected within as well as between regions. Results also indicated the availability of diverse coffee genetic resources even in marginal coffee production areas of the country, like in northwestern Ethiopia. Evaluated genotypes should be effectively conserved for immediate as well as unforeseen future uses and the national coffee genetic resource collection and conservation effort should address marginal coffee production areas.

Four genotypes suitable for high density planting (AD0891, AD0991, AD1091 and AD1191), two genotypes with desirable cup quality and green bean physical characters (AD0691 and AD1691), and six genotypes with relatively low caffeine content (AD0291, AD0591, AD2491, AD2691, AD2791 and AD2891) were identified. In addition, genetically dissimilar as well as similar genotypes were identified using AFLP and SSR markers. Information regarding the correlation of morphological as well as quality characters was generated. This will assist to direct future coffee breeding programmes of the centre.

The relationship between genetic distance and hybrid performance and identification of molecular markers linked to low caffeine content, better cup quality, compact growth habit and CBD resistance are suggested as future areas of investigation.

CHAPTER 10

SUMMARY

The genetic diversity of *C. arabica* genotypes collected from the northwestern and southwestern parts of Ethiopia was evaluated using morphological, biochemical and molecular (AFLP and SSR) markers. The objectives of the study were to (1) estimate the level of genetic variation among currently grown *C. arabica* genotypes, (2) estimate the level of associations among agro-morphological characters, coffee quality traits and green bean biochemical compounds, (3) compare the level of genetic variation between coffee genotypes collected from the northwestern and southwestern parts of Ethiopia and (4) compare the efficiency of AFLP and SSR markers in detecting genetic variation in *C. arabica*.

Results of diversity analysis using nine qualitative and 18 quantitative morphological characters indicated the presence of substantial variability among evaluated coffee genotypes. Genotypes were significantly different for all 18 quantitative morphological characters. Higher Shannon-Weaver diversity indices ranging from 0.401 to 0.989 were recorded for nine qualitative morphological characters. Cluster analysis using qualitative, quantitative and a combination of qualitative and quantitative characters classified genotypes into two groups. Qualitative characters failed to independently distinguish seven of the genotypes. Principal component analysis grouped 18 quantitative characters into 15 principal components and the first four explained 82.3% of the entire variability among genotypes. Average green bean yield per tree and bean length were important for the variation explained in three of the first four principal components.

Correlation analysis performed among 18 agro-morphological characters indicated positive associations between average green bean yield per tree and all other characters. Average green bean yield per tree had statistically significant correlations with percentage of bearing primary branches per tree, bean weight, canopy and trunk diameters, tree height, bean length and thickness, internode lengths of orthotropic and primary branches and fruit and petiole lengths. Broad sense heritability varied from

38% for bean thickness to 94% for bean weight and number of secondary branches per tree. Average green bean yield per tree, canopy diameter, tree height and bean weight had higher genetic coefficients of variation and broad sense heritabilities. These characters were positively and significantly correlated, allowing simultaneous improvement of these traits.

Genotypes were evaluated for variability in four cup quality traits (acidity, body, flavour and overall standard) and four green bean physical characters (bean shape, size, uniformity and weight). Considerable variation was observed among genotypes both for cup quality as well as green bean physical characters. Of all the genotypes, AD0691 and AD1691 had both desirable cup quality and green bean physical characters. Cluster analysis performed using cup quality and green bean physical characters, classified genotypes into two groups. Genotypes were not clustered according to collection regions. All cup quality traits were positively and significantly correlated and correlation coefficients ranged from 0.69 to 0.93. Positive correlations were observed among desirable cup quality and green bean physical characters indicating that cup quality and green bean physical characters can be improved simultaneously.

Genotypes were evaluated for green bean caffeine, chlorogenic acids, sucrose and trigonelline content variability. Analysis of variance indicated significant differences among genotypes for all biochemical compounds. Green bean caffeine, chlorogenic acids, sucrose and trigonelline contents ranged between 0.91-1.32%, 2.34-4.67%, 5.30-8.98% and 1.04-1.71%, respectively on dry matter basis. Cluster analysis classified coffee genotypes into eight groups and genotypes were not clustered according to collection regions. Coffee genotypes collected from northwestern Ethiopia were more diverse in green bean caffeine, chlorogenic acids and sucrose contents compared to southwestern Ethiopia genotypes. On average, southwestern Ethiopia coffee genotypes had higher green bean caffeine, chlorogenic acids, sucrose and trigonelline contents compared to northwestern Ethiopia genotypes. Green bean caffeine content showed negative and statistically significant associations with all desirable cup quality attributes and correlation coefficients ranged from -0.305 to -0.407. Desirable cup quality traits and low green bean caffeine content can be selected simultaneously. Green bean sucrose content showed positive associations with all

desirable cup quality and green bean physical characters. Higher green bean sucrose content, desirable cup quality as well as green bean physical characters can be improved together.

The genetic diversity of coffee genotypes was investigated using 10 AFLP primer combinations. Results revealed differences among AFLP primer combinations in detecting genetic variation among genotypes. Of the 10 primer combinations *Eco*RI-ACA/*Mse*I-CAA, *Eco*RI-ACC/*Mse*I-CAG, *Eco*RI-ACT/*Mse*I-CAG and *Eco*RI-AAC/*Mse*I-CAA were superior in detecting genetic variation. Almost all primer combinations were positively correlated in estimating pair-wise genetic similarity coefficients, indicating the similarity of primer combinations in detecting genetic variation among genotypes. All genotypes were independently distinguished and pair-wise genetic similarity coefficients ranged from 0.851 to 0.982 with an average of 0.915. Ten genotypes had 18 specific AFLP markers which could be utilised for genotype fingerprinting. AD1291 and AD1491 were the most dissimilar while AD3591 and AD3991 were the most similar genotypes. Genotypes were classified into two groups using UPGMA method of cluster analysis. Genotypes did not cluster according to collection regions. Results unveiled the presence of genetic variation among genotypes collected in each region and the presence of close genetic similarity among some coffee genotypes collected from different regions. The genetic diversity among coffee genotypes from southwestern Ethiopia was higher compared to those from northwestern Ethiopia.

Genetic diversity analysis performed using six SSR primer pairs indicated the presence of considerable genetic variation among some of the genotypes collected from northwestern and southwestern Ethiopia. Twenty polymorphic SSR markers were amplified of which two were specific to genotypes AD1491 and AD2991. Using 20 polymorphic SSR markers, 64.3% of the genotypes were independently distinguished. Pair-wise genetic distances ranged from 0.286 to 1.000. The lowest pair-wise genetic similarity coefficient was recorded between AD0591 and AD1491 as well as AD1491 and AD4591. Genotypes from southwestern Ethiopia were more diverse compared to those from northwestern Ethiopia.

Molecular markers differ in the amount of information generated per PCR reaction. The efficiency of 10 AFLP primer combinations and six SSR primer pairs in detecting genetic variation was compared using 28 *C. arabica* genotypes. AFLP primer combinations amplified 220 and SSR primer pairs 20 polymorphic alleles among evaluated genotypes. AFLP markers independently distinguished all evaluated coffee genotypes whereas SSR markers distinguished 64.3% of the genotypes. Moreover, 18 genotype specific AFLP markers compared to two SSR genotype specific markers were identified. The assay efficiency index of AFLP markers (22) was far superior to that of SSR markers (3.3). AFLP markers surpassed SSR markers in detecting genetic variation among evaluated arabica coffee genotypes. Results of this study indicated the presence of good congruence between the two genetic markers in estimating pair-wise genetic similarity coefficients among genotypes.

OPSOMMING

Die genetiese diversiteit van *C. arabica* genotipes wat versamel is in die noordwestelike en suidwestelike dele van Ethiopië is geëvalueer deur gebruik te maak van morfologiese, biochemiese en molekulêre (AFLP en SSR) merkers. Die doel van hierdie studie was (1) om die vlak van genetiese variasie tussen *C. arabica* tipes wat tans geplant word, te bepaal, (2) om die vlak van assosiasie tussen agro-morfologiese eienskappe, koffie kwaliteits eienskappe en groenboon biochemiese eienskappe te bepaal, (3) om die vlak van genetiese variasie tussen genotipes van noordwestelike en suidwestelike Ethiopië te bepaal en (4) om die effektiwiteit van AFLP en SSR merkers te vergelyk in die bepaling van genetiese variasie in *C. arabica.*

Resultate van die diversiteits analise met die gebruik van nege kwalitatiewe en 18 kwantitatiewe morfologiese eienskappe het ‘n hoë vlak van diversiteit tussen koffie genotypes getoon. Genotipes was betekenisvol verskillend vir al 18 kwantitatiewe morfologiese eienskappe. Hoër Shannon-Weaver diversiteits indekse wat wissel van 0.401 tot 0.989 is vir die nege kwalitatiewe morfologiese eienskappe gevind. Tros analise wat gedoen is deur gebruik te maak van kwalitatiewe, kwantitatiewe eienskappe en ‘n kombinasie van die twee, het die genotipes in twee groepe verdeel. Kwalitatiewe eienskappe kon sewe van die genotipes nie van mekaar onderskei nie. Hoof komponent analise het die 18 kwantitatiewe eienskappe in 15 hoof komponente verdeel, waar die eerste vier 82.3% van die totale variasie tussen genotipes verklaar het. Gemiddelde groenboon opbrengs per boom en boon lengte was belangrik vir drie van die vier eerste hoof komponente.

Korrelasie analise van 18 agro-morfologiese eienskappe het aangetoon dat daar positiewe assosiasie tussen groenboon opbrengs per boom en alle ander eienskappe voorkom. Die gemiddelde groenboon opbrengs per boom was betekenisvol gekorreleer met persentasie draende primêre takke per boom, boon gewig, boon massa en boom en stam deursnee, boom hoogte, boon lengte en dikte, internode lengte en vrugte en petiool lengte. Breë sin oorerflikheid het gewissel van 38% vir boon dikte tot 94% vir boon massa asook getal sekondêre takke per boom. Gemiddelde groenboon opbrengs per boom, boom deursnee, boom hoogte en boon

massa het hoër genetiese koeffisiente van variasie en breë sin oorerflikhede gehad. Hierdie eienskappe was ook betekenisvol positief gekorreleer, dus kan hierdie eienskappe gelyktydig verbeter word.

Genotipes is vir variasie in vier koppie kwaliteit eienskappe (suurheid, volheid, geur en algemene standaard) en vier groenboon eienskappe (boon vorm, grootte, univormiteit, en massa) geëvalueer. Groot variabiliteit is tussen koppie kwaliteit en groenboon fisiese eienskappe gevind. Van al die genotipes, het AD0691 en AD1691 beide goeie koppie kwaliteit en groenboon eienskappe getoon. Tros analise wat gedoen is met koppie kwaliteit en groenboon fisiese eienskappe, het die genotipes in twee groepe verdeel. Genotipes het nie volgens versamelingsareas gegroepeer nie. Alle koppie kwaliteits eienskappe was positief betekenisvol gekorreleer en korrelasie koeffisiënte het van 0.69 tot 0.93 gewissel. Daar was ook positiewe korrelasies tussen goeie koppie kwaliteit en groenboon fisiese eienskappe, daarom kan koppie kwaliteit en groenboon fisiese eienskappe gelyktydig verbeter word.

Genotipes is vir goenboon kaffeïen, chlorogeniese sure, sukrose en trigonelien inhoud variasie geëvalueer. Analise van variasie het betekenisvolle verskille tussen genotipes vir alle biochemiese stowwe aangetoon. Groenboon kaffeïen, chlorogeniese sure, sukrose en trigonelien inhoud het onderskeidelik tussen 0.91-1.32%, 2.34-4.67%, 5.30-8.98% en 1.04-1.71% gewissel op ‘n droë massa basis. Tros analise het koffie genotipes in agt groepe verdeel en weer eens is genotipes nie volgens gebied van versameling gegroepeer nie. Koffie genotipes van noordwes Ethiopië was meer divers vir groenboon kafeïen, chlorogeniese sure en sukrose inhoud in vergelyking met genotipes van suidwes Ethiopië. In die algemeen het koffie genotipes van suidwes Ethiopië hoër groenboon kafeïen, chlorogeniese sure, sukrose en trigoneliensuur inhoud gehad as genotypes van noordwes Ethiopië. Groenboon kafeïen inhoud was betekenisvol negatief gekorreleer met alle gewenste koppie kwaliteit eienskappe en korrelasies het gewissel van -0.305 tot -0.407. Gewenste koppie kwaliteit eienskappe en lae kafeïen inhoud kan dus gelyktydig geselekteer word. Groenboon sukrose inhoud was positief met gewenste koppie kwaliteit en groenboon fisiese eienskappe geassosieer. Sukrose inhoud, goeie koppie kwaliteit en groenboon fisiese eienskappe kan dus gelyktydig verbeter word.

Die diversiteit van koffie genotipes is met 10 AFLP voorvoerder kombinasies geëvalueer. Resultate het verskille tussen die vermoë van voorvoerder kombinasies om genetiese variasie tussen genotipes te identifiseer, getoon. Van die 10 voorvoerder kombinasies, was *Eco*RI-ACA/*Mse*I-CAA, *Eco*RI-ACC/*Mse*I-CAG, *Eco*RI-ACT/*Mse*I-CAG en *Eco*RI-AAC/*Mse*I-CAA die effektiefste om genetiese variasie te bepaal. Feitlik alle voorvoerder kombinasies was positief gekorreleer in hulle bepaling van paargewyse genetiese similariteits koeffisiënte. Alle genotypes kon van mekaar onderskei word en koeffisiënte het van 0.851 tot 0.982 gewissel met 'n gemiddeld van 0.915. Agtien spesifieke AFLP merkers is vir 10 van die genotypes gekry en kan vir genotipiese vingerafdruk analises gebruik word. AD1291 en AD1491 asook AD3591 en AD3991 het die minste en meeste ooreensteming respektiewelik gehad. Genotipes is in twee groepe verdeel met die UPGMA metode van tros analise, maar genotipes het nie volgens versamelings areas gegroepeer nie. Dit het gedui op die genetiese variasie tussen genotipes van dieselfde area asook sterk genetiese verwantskape tussen genotipes van verskillende areas. Die genetiese diversiteit tussen genotipes van suidwes Ethiopië was hoër as die van noordwes Ethiopië.

Genetiese diversiteits studies wat met ses SSR voorvoerder pare gedoen is, het groot genetiese variasie tussen genotipes wat in suidwes en noordwes Ethiopië versamel is, aangedui. Twintig polimorfiese SSR merkers is geamplifiseer waarvan twee spesifiek vir genotipes AD1491 en AD2991 was. Met die gebruik van 20 polimorfiese SSR merkers kon 64.3% van die genotipes onafhanklik onderskei word. Paar gewyse genetiese afstande het van 0.286 tot 1.000 gewissel. Die laagste paar gewyse genetiese koeffisiënt van ooreenstemming was tussen AD0591 en AD1491 en ook tussen AD1491 en AD4591. Die genetiese diversiteit tussen genotipes van suidwes Ethiopië was hoër as die van noordwes Ethiopië.

Molekulêre merkers het gewissel in die hoeveelheid inligting wat per PCR reaksie gegenereer is. Die effektiwiteit van 10 AFLP voorvoerder kombinasies en ses SSR voorvoerder pare is in 28 *C. arabica* genotipes vergelyk. AFLP voorvoerder kombinasies het 220 polimorfiese fragmente geamplifiseer, terwyl SSR voorvoerder pare 20 polimorfiese allele het in die geëvalueerde genotipes geamplifiseer. AFLP merkers het al die genotipes onafhanklik onderskei terwyl SSR merkers onderskeid tussen 64.3% van die genotypes getref het. Verder is 18 genotipe spesifieke AFLP

merkers identifiseer terwyl SSR net twee genotipe spesifieke merkers getoon het. Die evaluasie effektiwiteits indeks van AFLP merkers (22) was baie hoër as die van die SSR merkers (3.3). Daarom was AFLP merkers baie beter as SSR merkers om genetiese variasie tussen die geëvalueerde arabica koffie genotipes te identifiseer. Tog het die resultate van hierdie studie getoon dat daar goeie ooreenkoms tussen die twee genetiese merker sisteme bestaan om paar-gewyse genetiese similariteits koeffisiënte tussen genotipes te bepaal.

REFERENCES

Aerts, R.J. and T.W. Baumann, 1994. Distribution and utilization of chlorogenic acid in coffee seedlings. Journal of Experimental Botany 45: 497-503.

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

Aggarwal, R.K., D.S. Brar, S. Nandi, N. Huang and G.S. Khush, 1999. Phylogenic relationships among *Oryza* species revealed by AFLP markers. Theoretical and Applied Genetics 98: 1320-1328.

Agwanda, C.O., 1999. Flavour: an ideal selection criterion for the genetic improvement of liquor quality in arabica coffee. In the proceedings of 18th International Scientific Colloquium on Coffee, pp. 383-389. Helsinki, Finland.

Agwanda, C.O., P. Baradat, A.B. Eskes, C. Cilas and A. Charrier, 2003. Selection for bean and liquor qualities within related hybrids of arabica coffee in multi-local field trials. Euphytica 131: 1-14.

Agwanda, C.O., P. Lashermes, P. Trouslot, M.C. Combes and A. Charrier, 1997. Identification of RAPD markers for resistance to coffee berry disease, *Colletotrichum kahawae*, in arabica coffee. Euphytica 97: 241-248.

Ajmone-Marsan, P., P. Castiglioni, F. Fusari, M. Kuiper and M. Motto, 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theoretical and Applied Genetics 96: 219-227.

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan, 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132: 1131-1139.

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

Akkaya, M.S., R.C. Shoemaker, J.E. Specnt, A.A. Bhagwat and P.B. Cregan, 1995. Integration of simple sequence repeat DNA markers in to soybean linkage map. Crop Science 35: 1439-1445.

Allard, R.W., 1960. Principles of Plant Breeding. Wiley, New York.

Almanza-Pinzon, M.I., M. Khairallah, P.N. Fox and M.L. Warburton, 2003. Comparison of molecular markers and coefficients of parentage for the analysis of genetic diversity among spring bread wheat accessions. Euphytica 130: 77-86.

Alonso-Blanco, C., A.J. Peeters, M. Koornneef, C. Lister and C. Dean, 1998. Development of an AFLP based linkage of *Ler* Col and *Cvi Arabidopsis thaliana* ecotypes and construction of a *Ler*/*Cvi* recombinant inbred line populations. Plant Journal 14: 259-271.

Alvarez, A.E., C.C.M. Van de Wiel, M.J.M. Smulders and B. Vosman, 2001. Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon.* Theoretical and Applied Genetics 103: 1283-1292.

Ameha, M., 1980. Yield assessment of indigenous coffee collections grown at Jimma research station. Ethiopian Journal of Agricultural Science 2: 69-77.

Ameha, M., 1982. Heterosis in crosses of indigenous coffee (*Coffea arabica* L.) selected for yield and resistance to coffee berry disease. I. At first bearing stage. Ethiopian Journal of Agricultural Science 4: 33-43.

Ameha, M., 1986. The significance of Ethiopian coffee genetic resources to coffee improvement. In the proceedings of an international symposium on the conservation and utilization of Ethiopian germplasm, pp. 247-253. Addis Ababa, Ethiopia, 13-16 October, 1986.

Ameha, M. and B. Bellachew, 1983. Heterosis in crosses of indigenous coffee selected for yield and resistance to coffee berry disease. II. First three years. Ethiopian Journal of Agricultural Science 5: 13-21.

Anonymous, 1991. User`s guide to MSTAT-C. Michigan State University, East Lansing.

Anonymous, 2004a. Facts about caffeine. www.daap.ca/facts on caffeine.html.

Anonymous, 2004b. Caffeine; facts. www.xs4all.nl/~4david/caffeine.html.

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

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

Anthony, F., M.N. Clifford and M. Noirot, 1993. Biochemical diversity in the genus *Coffea* L.: chlorogenic acids, caffeine and mozambioside contents. Genetic Resources and Crop Evolution 40: 61-70.

Anzueto, F., B. Bertrand, J.L. Sarah, A.B. Eskes and B. Decazy, 2001. Resistance to *Meloidogyne incognita* in Ethiopian *Coffea arabica* accessions. Euphytica 118: 1-8.

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. Theoretical and Applied Genetics 106: 1341-1352.

Ardelean, M., M. Cordea, D. Pamfil, N.W. Blackhall, S.C. Andras, M.R. Davey and J.B. Power, 2004. Revealing genetic diversity of three sections (*Pharbitis*, *Quamoclit* and *Batatas*) of Ipomoea genus by means of RAPD analysis. Acta Horticulturae 651: 149-153.

Ashihara, H. and C. Crozier, 1999. Biosynthesis and metabolism of caffeine and related purine alkaloids in plants. In: J.R. Callow (Ed.), Advances of Botanical Research, pp. 117-205. Academic press, London.

Assefa, K., S. Ketema, H. Tefera, H.T. Nguyen, A. Blum, M. Ayele, G. Bai, B. Simane and T. Kefyalew, 1999. Diversity among germplasm lines of the Ethiopian cereal tef (*Eragrostis tef* (Zucc.) Trotter). Euphytica 106: 87-97.

Ayana, A., E. Bekele and T. Bryngelsson, 2000. Genetic variation in wild sorghum (*Sorghum bicolour* sp. *verticilliforum* (L) Moench) germplasm in Ethiopia assessed by RAPD. Hereditas 132: 249-254.

Ayele, M., H. Tefera, K. Assefa and H.T. Nguyen, 1999. Genetic characterisation of two *Eragrostis* species using AFLP and morphological traits. Hereditas 130: 33-40.

Bahrman, N., J. Le Gouis, D. Hariri, L. Guilbaud and L. Jestin, 1999. Genetic diversity of old French six-rowed winter barley varieties assessed with molecular, biochemical and morphological markers and its relation to BaMMV resistance. Heredity 83: 568-574.

Baker, J.H.A., M. Matthes, G. M. Arnold, K.J. Edwards, I. Ahman, S. Larsson and A. Karp, 1999. Characterization of genetic diversity in potential biomass willows (*Salix* sp.) by RAPD and AFLP analyses. Genome 42: 173-183.

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

Barbosa, A.M.M., I.O. Geraldi, L.L. Benchimol, A.A.F. Gracia, C.L. Souza Jr. and A.P. Souza, 2003. Relationship of intra- and inter-population tropical maize single

cross hybrid performance and genetic distances computed from AFLP and SSR markers. Euphytica 130: 87-99.

Barre, P., S. Akaffou, J. Louarn, A. Charrier, S. Hamon and M. Noirot, 1998. Inheritance of caffeine and heteroside contents in an interspecific cross between a cultivated coffee species *Coffea liberica* var. *dewevrei* and a wild species caffeine-free *C. pseudozanguebariae*. Theoretical and Applied Genetics 96: 306-311.

Barrett, B.A. and K.K. Kidwell, 1998. AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. Crop Science 38: 1261-1271.

Baumann, T.W. and H. Gabriel, 1984. Metabolism and excretion of caffeine during germination of *Coffea arabica* L. Plant and Cell Physiology 25: 1431-1436.

Baumann, T.W., M.R. Sondahl, S.S.M. Waldhauser and J.A. Kretschmar, 1998. Non-destructive analysis of natural variability in bean caffeine content of laurina coffee. Phytochemistry 49: 1569-1573.

Becker, J., P. Vos, M. Kuiper, F. Salamini, and M. Heun, 1995. Combined mapping of AFLP and RFLP markers in barley. Molecular and General Genetics 249: 65-73.

Bellachew, B., 2001. Arabica coffee breeding for yield and resistance to coffee berry disease (*Colletotrichum kahawae* Sp.nov.). A PhD degree thesis submitted to the University of London.

Bellachew, B., B. Atero and F. Tefera, 2000. Breeding for resistance to coffee berry disease in arabica coffee: progress since 1973. In the proceedings of the workshop on control of coffee berry disease in Ethiopia, pp. 85-98. Ethiopian Agricultural Research Organization (EARO), Addis Ababa.

Bellachew, B., B. Atero and G. Temesgen, 1998. Description and production recommendations for new cultivars of arabica coffee, Research report no. 34. IAR, Addis Ababa.

Bellachew, B., M. Ameha and D. Mekonen, 1993. Heterosis and combining ability in coffee (*Coffea arabica* L.). In the proceedings of the 15th International Scientific Colloquium on Coffee, pp. 234-242. Montpellier, France.

Berhanu, A., 1998. Mythological and historical background of coffee. Kaffa coffee 1: 13-20.

Berthou, F. and P. Trouslot, 1977. L`analyse du polymorphisme enzymatique dans le genre *Coffea*: adaptation d`une me`thode d`electrophorese en serie. 8th Conference of ASIC, Abidjan : 373-383.

Berthaud, J. and A. Charrier, 1988. Genetic resources of *Coffea*. In: R.J. Clarke and R. Macrae (Eds.), Coffee Agronomy, pp. 1-41. Elsevier Applied Science, London.

Bicchi, C.P., A.E. Binello, G.M. Pellegrino and A.C. Vanni, 1995. Characterisation of green and roasted coffees through the chlorogenic acid fraction by HPLC-UV and principal component analysis. Journal of Agriculture and Food Chemistry 43: 1549-1555.

Biratu, T., 1997. Effects of caffeine on coffee berry disease (*Colletotrichum kahawae*). In the proceedings of the 17th International Scientific Colloquium on Coffee. Nairobi, Kenya.

Birhane, T., 2002. Ethiopian coffee on a new road to compete in the world market. Merewa pp. 9-13.

Blears, M.J., S.D. De Grandis, H. Lee and J.T. Trevors, 1998. Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. Journal of Industrial Microbiology and Biotechnology 21: 99-114.

Bond, J.M., R.J. Mogg, G.R. Squire and C. Johnstone, 2004. Microsatellite amplification in *Brassica napus* cultivars: cultivar variability and relationship to a long-term feral population. Euphytica 139: 173-178.

Botstein, D., R.L. White, M. Skolnick and R.W. Davis, 1980. Construction of a genetic map in man using restriction fragment length polymorphism. American Journal of Human Genetics 32: 314-331.

Breyne, P., D. Rombaut, A. Van Gysel, M. Van Montagu and T. Gerats, 1999. AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. Molecular and General Genetics 261: 627-634.

Breyne, P., W. Boerjan, T. Gerats, M. Van Montagu and A. Van Gysel, 1997. Applications of AFLP in plant breeding, molecular biology and genetics. Belgian Journal of Botany 129: 107-117.

Briard, M., A. Horvais and J.Y. Peron, 2002. Wild seakale (*Crambe maritime* L.) diversity as investigated by morphological and molecular markers. Scientia Horticulturae 95: 1-12.

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.

Bruna, S., L. De Benedetti, A. Mercuri, T. Schiva, G. Burchi, N. Pecchioni and C. Agrimonti, 2004. Use of RAPD markers for the genetic characterization of *Limonium* Species. Acta Horticulturae 651: 155-160.

Burton, W.A., V.L. Ripley, D.A. Potts and P.A. Salisbury, 2004. Assessment of genetic diversity in selected breeding lines and cultivars of Canola quality *Brassica juncea* and their implications for Canola breeding. Euphytica 136: 181-192.

Camussi, A., E. Ottaviano, T. Calinski and Z. Kaczmarek, 1985. Genetic distances based on quantitative traits. Genetics 111: 945-962.

Carr, J., M. Xu and J.W. Dudley, 2003. AFLP analysis of genetic variability in New Guinea impatiens. Theoretical and Applied Genetics 106: 1509-1516.

Carvalho, A., 1988. Principles and practices of coffee plant breeding for productivity and quality factors: *Coffea arabica.* In: R.J. Clarke and R. Macrae (Eds.), Coffee Agronomy, pp. 129-165. Elsevier Applied Science, London.

Cerna, F.J., S.R. Cianzio, A. Rafalski, S. Tingey and D. Dyer, 1997. Relationship between seed yield heterosis and molecular marker heterozygosity in soybean. Theoretical and Applied Genetics 95: 460-467.

Cervera, M.T., J. Gusamao, M. Steenackers, J. Peleman, V. Storme, A. Vanden Boerck, M. Van Montagu and W. Boerjan, 1996. Identification of AFLP molecular markers for resistance against *Melampsora larici-populina* in Populus. Theoretical and Applied Genetics 93: 733-737.

Chaparro, A.P., M.A. Cristancho, H.A. Cortina and A.L. Gaitan, 2004. Genetic variability of *Coffea arabica* L. accessions from Ethiopia evaluated with RAPDs. Genetic Resources and Crop Evolution 51: 291-297.

Charcosset, A. and L. Moreau, 2004. Use of molecular markers for the development of new cultivars and the evaluation of genetic diversity. Euphytica 137: 81-94.

Charrier, A. and J. Berthaud, 1985. Botanical classification of coffee. In: M.N. Clifford and K.C. Wilson (Eds.), Coffee botany, biochemistry and production of beans and beverage, pp. 13-47. Croom Helm, London.

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.

Chifra, W., N. Wolde-Tekle and G. Tadesse, 1998. The excellence of Ethiopian coffee and the quality control mechanisms in place. Kaffa Coffee 1: 36-40.

Cho, Y.G., T. Ishii, S. Temnykh, X. Chen, L. Lipovich, S.R. McCouch, W.D. Park, N. Ayres and S. Cartinhour, 2000. Diversity of microsatellites derived from genomic libraries and gene bank sequences in rice. Theoretical and Applied Genetics 100: 713-722.

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

Cilas, C., C. Montagnon, B. Bertrand and C. Godin, 2000. Wood elasticity of several *Coffea canephora* Pierre clones: A new trait to be included in selection schemes. Agronomie 20: 439-444.

Clarke, R.J., 1985. The technology of converting green coffee into the beverage. In: M.N. Clifford and K.C. Willson (Eds.), Coffee botany, biochemistry, and production of beans and beverage, pp. 375-393. Croom Helm, London.

Clifford, M.N., 1985. Chemical and physical aspects of green coffee and coffee products. In: M.N. Clifford and K.C. Willson (Eds.), Coffee botany, biochemistry, and production of beans and beverage, pp. 305-374. Croom Helm, London.

Clifford, M.N. and J.R. Ramirez-Martinez, 1991. Phenols and caffeine in wet-processed coffee beans and coffee pulp. Food Chemistry 40: 35-42.

Clifford, M.N. and T. Jarvis, 1988. The chlorogenic acids content of green robusta coffee beans as possible index of geographic origin. Food Chemistry 29: 291-298.

Clifford, M.N., T. Williams and D. Bridson, 1989. Chlorogenic acids and caffeine as possible taxonomic criteria in *Coffea* and *Psilanthus*. Phytochemistry 28: 829-838.

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.

Coste, R., 1992. Coffee: the plant and the product. Macmillan, Hong Kong.

Cregan, P.B., T. Jarvick, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. Vantoai, D.G. Lohnes, J. Chung and J.E. Specht, 1999. An integrated genetic linkage map of the soybean genome. Crop Science 39: 1464-1490.

Cros, J., M.C. Gavalda, N. Chabrillange, C. Recalt, C. Duperray and S. Hamon, 1993. Variation in the total nuclear DNA content in African *Coffea* species (*Rubiaceae*). In the proceedings of the 15th International Scientific Colloquium on Coffee, Montpellier, France.

CTA, 2003. The profile of Ethiopian coffee. Coffee and Tea Authority, Addis Ababa.

Das, S., J. Rajagopal, S. Bhatia, P.S. Srivastava and M. Lakshmikumaran, 1999. Assessment of genetic variation within *Brassica campestris* cultivars using AFLP and RAPD DNA markers. Journal of Biosciences 24: 433-440.

Dean, R.E., J.A. Dahlberg, M.S. Hopkins, S.E. Mitchell and S. Kresovich, 1999. Genetic redundancy and diversity among "Orange" accessions in the U.S. National sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Science 39: 1215-1221.

De Maria, C.A.B., L.C. Trugo, F.R. Aquino Neto, R.F.A. Moreira and C.S. Alviano, 1996. Composition of green coffee water-soluble fractions and identification of volatiles formed during roasting. Food Chemistry 55: 203-207.

De Maria, C.A.B., L.C. Trugo, R.F.A. Moreira and C.C. Werneck, 1994. Composition of green coffee fractions and their contribution to the volatile profile formed during roasting. Food Chemistry 50: 141-145.

Dessalegn, Y., 2002. Coffee bearing unusual flowers and berries. AgriTopia 17: 12-14.

Dias, J.C., A.A. Monteiro and M.B. Lima, 1993. Numerical taxonomy of Portuguese Tronchuda cabbage and Galega kale landraces using morphological characters. Euphytica 69: 51-68.

Dirlewanger, E., P. Cosson, M. Tavaud, M.J. Aranzana, C. Poizat, A. Zanetto, P. Arus and F. Laigret, 2002. Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics 105: 127-138.

Diwan, N. and P.B. Cregan, 1997. Automated sizing of fluorescent labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theoretical and Applied Genetics 95: 723-733.

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.

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

Dussert, S., F. Engelmann, N. Chabrillange, F. Anthony, M. Noirot and S. Hamon, 1997. *In vitro* conservation of coffee (*Coffea* spp.) germplasm. In: M.K. Razdan

and E.C. Cocking (Eds.), Conservation of plant genetic resources *in vitro*, pp. 287-305. Science Publishers, Enfield, New Hampshire.

Elhoumaizi, M.A., M. Saaidi, A. Oihabi and C. Cilas, 2002. Phenotypic diversity of date-palm cultivars (*Phoenix dactylifera* L.) from Morocco. Genetic Resources and Crop Evolution 49: 483-490.

Ellis, R.P., J.W. McNicol, E. Baird, A. Booth, P. Lawrence, B. Thomas and W. Powell, 1997. The use of AFLPs to examine genetic relatedness in barley. Molecular Breeding 3: 359-369.

Erschadi, S., G. Haberer, M. Schoniger and R.A. Torres-Ruiz, 2000. Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphism (AFLP). Theoretical and Applied Genetics 100: 633-640.

Eskes, A. B., 1983. Incomplete resistance to coffee leaf rust (*Hemileia vastatrix*). A PhD thesis, Wageningen Agricultural University.

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

FAO, 1968a. FAO coffee mission to Ethiopian 1964 –1965. Food and Agriculture Organization of the United Nations, Rome.

FAO, 1968b. World coffee survey report. Food and Agriculture Organization of the United Nations, Rome.

Fazuoli, L.C., M.P. Maluf, O.G. Filho, H.M. Filho and M.B. Silvarolla, 2000. Breeding and Biotechnology of coffee. In: T. Sera, C.R. Soccol, A. Pandey and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 27-45. Kluwer Academic Publishers, Dordrecht.

Feldman, J.R., W.S. Ryder and J.T. Kung, 1969. Importance of non-volatile compounds to the flavour of coffee. Journal of Agricultural and Food Chemistry 17: 733-739.

Florin, B., P. Ducos, L. Firmin, M.C. Meschine, C. Thierry, V. Petiard and A. Deshayes, 1995. Preservation of coffee somatic embryos through desiccation and cryopreservation. In the proceedings of the 16th International Scientific Colloquium on Coffee, pp. 542-547. Kyoto, Japan.

Fossati, T., I. Zapelli, S. Bisoffi, A. Micheletti, L. Vietto, F. Sala and S. Castiglione, 2005. Genetic relationships and clonal identity in collection of commercially relevant poplar cultivars assessed by AFLP and SSR. Tree Genetics and Genomes In press.

Fuentes, J.L., F. Escobar, A. Alvarez, G. Gallego, M.C. Duque, M. Ferrer, J.E. Deus and J.M. Tohme, 1999. Analysis of genetic diversity in Cuban rice varieties using isozyme, RAPD and AFLP markers. Euphytica 109: 107-115.

Garcia-Mas, J, M. Oliver, H. Gomez-Paniagua and M.C. de Vicente, 2000. Comparing AFLP, RAPD and RFLP markers for measuring genetic diversity in melon. Theoretical and Applied Genetics 101: 860-864.

Geleta, L.F., M.T. Labuschagne and C.D. Viljoen, 2004. Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. Plant Breeding 123: 467-473.

Gepts, P., 1991. Biotechnology sheds light on bean domestication in Latin America. Diversity 7: 49-50.

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

Goulao, L., T. Valdiviesso, C. Santana and C.M. Oliveira, 2001. Comparison between phonetic characterisation using RAPD and ISSR markers and phenotypic data of cultivated chestnut (*Castanea sativa* Mill.). Genetic Resources and Crop Evolution 48: 329-338.

Granati, E., V. Bisignano, D. Chiaretti, P. Crino and G.B. Polignano, 2003. Characterisation of Italian and exotic *Lathyrus* germplasm for quality traits. Genetic Resources and Crop Evolution 50:273-280.

Guzman, F.A., H. Ayala, C. Azurdia, M.C. Duque and M.C. de Vicente, 2005. AFLP assessment of genetic diversity of capsicum genetic resources in Guatemala: Home gardens as an option for conservation. Crop Science 45: 363-370.

Haarer, A.E., 1962. Modern coffee production. Leonard Hill Books Limited, London.

Hagen, M.J. and J.L. Hamrick, 1998. Genetic variation and population genetic structure in *Trifolium pratense.* Journal of Heredity 89: 178-181.

Halward, T., T. Stalker, E. LaRue and G. Kochert, 1992. Use of single primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). Plant Molecular Biology 18: 315-325.

Hamza, S., W.B. Hamida, A. Rabai and M. Harrabi, 2004. SSR-based genetic diversity assessment among Tunisian winter barley and relationship with morphological traits. Euphytica 135: 107-118.

Harborne, J.B., 1988. Introduction to ecological biochemistry. Academic Press, London.

Harding, R.M., A.J. Boyce and J.B. Clegg, 1992. The evolution of tandemly repetitive DNA: recombination rules. Genetics 132: 847-859.

Hardon, J.J., B. Vosman and T.L. Hintum, 1994. Identification of genetic resources and their origin: The capabilities and limitations of modern biochemical and legal systems commission on plant genetic resources. First extraordinary session. Rome, 7-11 November 1994.

Harlan, J.R., 1992. Crops and man. Second edition, American society of Agronomy, Madison.

Herrera, J.C., M.C. Combes, F. Anthony, A. Charrier and P. Lashermes, 2002. Introgression into the allotetraploid coffee (*Coffea arabica L.*): segregation and recombination of the *C. canephora* genome in the tetraploid interspecific hybrid (*C. arabica* x *C. canephora*). Theoretical and Applied Genetics 104: 661-668.

Herselman, L., 2003. Genetic variation among South African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica 133: 319-327.

Herselman, L., R. Thwaites, F.M. Kimmins, B. Courtois, P.J.A. Van der Merwe and S.E. Seal, 2004. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. Theoretical and Applied Genetics 109: 1426-1433.

Hintze, J.L., 1998. NCSS-2000 Statistical System for Windows. Number Cruncher Statistical Systems, Kaysville, UT.

Hormaza, J.I., 2002. Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. Theoretical and Applied Genetics 104: 321-328.

Hu, J., M. Nakatani, A.G. Lalusin and T. Fujimura, 2004. New microsatellite markers developed from reported *Ipomoea trifida* sequences and their application to sweet potato and its related wild species. Scientia Horticulturae 102: 375-386.

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. Theoretical and Applied Genetics 105: 699-707.

Hulupi, R., P. Rahardjo and D.S. Mawardi, 2004. Inheritance of bean abnormality in arabica coffee. www.iccri.net/m-97-4.htm.

ICO, 2004. All about coffee. http://www.ico.org/aico/wwd.htm.

Iketani, H., T. Manabe, N. Matsuta, T. Akihama and T. Hayashi, 1998. Incongruence between RFLPs of chloroplast DNA and morphological classification in east Asian pear (*Pyrus* Spp.). Genetic Resources and Crop Evolution 45: 533-539.

Innan, H., R. Terauchi, G. Kahl and F. Tajima, 1999. A method for estimating nucleotide diversity from AFLP data. Genetics 151: 1157-1164.

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

ITC, 1992. Coffee: An exporter`s guide. International Trade Centre, Geneva.

ITC, 2002. Coffee: An exporter's guide. International Trade Centre, Geneva.

ITC, 2004. Bitter or better future for coffee producers? The magazine of the International Trade Centre, 2: 9-13. International Trade Centre, Geneva.

Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. Bull Soc Vaud Sci Nat 44: 223-270.

Jacoby, A., M.T. Labuschange and C.D. Viljoen, 2003. Genetic relationships between South African *Solanum retroflexum* Dun. and other related species measured by morphological and DNA markers. Euphytica 132: 109-113.

Jakse, J., K. Kindlhofer and B. Javornik, 2001. Assessment of genetic variation and differentiation of hop genotypes by microsatellite and AFLP markers. Genome 44: 773-782.

Johns, M.A., P.W. Skroch, J. Nienhuis, P. Hinrichsen, G. Bascur and C. Munoz-Schick, 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Science 37: 605-613.

Johnson, H.W., H.F. Robinson and R.E. Comstock, 1955. Estimates of genetic and environmental variability in soybeans. Agronomy Journal 47: 314-318.

Jones, C.J., K.J. Edwards, S. Castaglione, M.O. Winfield, F. Sala, C. Van de Wiel, G. Bredemeijer, B. Vosman, M. Matthes, A. Maly, R. Brettschneider, P. Bettini, M. Buiatti, E. Maestri, A. Malcevschi, N. Marmiroli, R. Aert, G. Volckaert, J. Rueda, R. Linaacero, A. Vazque and A. Karp, 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding 3: 381-390.

Kalendar, R., T. Grob, M. Regina, A. Suoniemi and A. Schulman, 1999. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theoretical and Applied Genetics 98: 704-711.

Karp, A., P.G. Isaac and D.S. Ingram, 1998. Molecular tools for screening biodiversity. Chapman and Hall, London.

Keim, P., B.W. Dier, T.C. Olson and R.C. Shoemaker, 1990. RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics 126: 735-742.

Kim, J.H., H. Joung, H.Y. Kim and Y.P. Lim, 1998. Estimation of genetic variation and relationship in potato (*Solanum tuberosum* L.) cultivars using AFLP markers. American Journal of Potato Research 75: 107-112.

Kochert, G., 1994. RFLP technology. In: R.L. Phillips and I.K. Vasil (Eds.), DNA based markers, pp. 8-38. Kluwer Academic Publishers, Dordrecht.

Kubik, C., M. Sawkins, W.A. Meyer and B.S. Gaut, 2001. Genetic diversity in seven perennial ryegrass (*Lolium perenne* L.) cultivars based on SSR markers. Crop Science 41: 1565-1572.

Kuleung, C., P.S. Baenziger and I. Dweikat, 2004. Transferability of SSR markers among wheat, rye and triticale. Theoretical and Applied Genetics 108: 1147-1150.

Ky, C.L., J. Louarn, B. Guyot, A. Charrier, S. Hamon and M. Noirot, 1999. Relations between and inheritance of chlorogenic acid contents in an interspecific cross between *Coffea pseudozanguebariae* and *Coffea liberica* var. *dewevrei*. Theoretical and Applied Genetics 98: 628-637.

Ky, C. L., J. Louarn, S. Dussert, B. Guyot, S. Hamon and M. Noirot, 2001. Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild *Coffea arabica* L. and *C. canephora* P. accessions. Food Chemistry 75: 223-230.

Ky, C.L., M. Noirot and S. Hamon, 1997. Comparison of five purification methods for chlorogenic acids in green coffee beans (*Coffea* Sp.). Journal of Agriculture and Food Chemistry 45: 786-790.

Ky, C.L., S. Doulbeau, B. Guyot, S. Akaffou, A. Charrier, S. Hamon, J. Louarn and M. Noirot, 2000. Inheritance of coffee bean sucrose content in the interspecific cross *Coffea pseudozanguebariae* x *Coffea liberica dewevrei*. Plant Breeding 119: 165-168.

Lage, J., M.L. Warburton, J. Crossa, B. Skovmand and S.B. Andersen, 2003. Assessment of genetic diversity in synthetic hexaploid wheats and their *Triticum dicoccum* and *Aegilops tauschii* parents using AFLPs and agronomic traits. Euphytica 134: 305-317.

Lagercrantz, U., H. Ellegran and L. Anderson, 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Research 21: 1111-1115.

Lanteri, S., E. Saba, M. Cadinu, G.M. Mallica, L. Baghino and E. Portis, 2004. Amplified fragment length polymorphism for genetic diversity assessment in globe artichoke. Theoretical and Applied Genetics 108: 1534-1544.

Lashermes, P., J. Cros, M.C. Combes, P. Trouslot, F. Anthony, S. Hamon and A. Charrier, 1996a. Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus *Coffea L.* Theoretical and Applied Genetics 93: 626-632.

Lashermes, P., J. Cros, P. Marmey and A. Charrier, 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., M.C. Combes, J. Robert, P. Trouslot, A. D`Hont, F. Anthony and A. Charrier, 1999. Molecular characterisation and origin of the *Coffea arabica* L. genome. Molecular and General Genetics 261: 259-266.

Lashermes, P., M.C. Combes, P. Topart, G. Graziosi, B. Bertrand and F. Anthony, 2000b. Molecular breeding in coffee (*Coffea arabica* L.). In: T. Sera, C.R. Soccol, A. Pandey and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 101-112. Kluwer Academic Publishers, Dordrecht.

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

Lashermes, P., S. Anderzewski, B. Bertrand, M.C. Combes, S. Dussert, G. Graziosi, P. Trouslot and F. Anthony, 2000a. Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). Theoretical and Applied Genetics 100: 139-146.

Lefort, F., C. Echt, R. Streiff and G.G. Vendramin, 1999. Microsatellite sequences: a new generation of molecular markers for forest genetics. Forest Genetics 6: 5-10.

Leloup, V., A. Louvrier and R. Liardon, 1995. Degradation mechanisms of chlorogenic acids during roasting. In the proceedings of the 16th International Scientific Colloquium on Coffee, pp. 420-428. Kyoto, Japan.

Leroy, T., C. Montagnon, A. Charrier and A.B. Eskes, 1993. Reciprocal recurrent selection applied to *Coffea canephora* Pierre. I. Characterisation and evaluation of breeding populations and value of intergroup hybrids. Euphytica 67: 113-125.

Li, C.Y., 2000. RAPD analysis of genetic variation in *Eucalyptus microtheca* F. Muell. Populations. Hereditas 132: 151-156.

Lin, J.J., J. Kuo, J. Ma, J.A. Saunders, H.S. Beard, M.H. Macdonald, W. Kenworthy, N. Ude and F. Matthews, 1996. Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Molecular Biology Report 14: 156-159.

Litt, M. and J.A. Luty, 1989. A hypervariable microsatellite revealed by *in vitro* amplification of dinucleotide repeat within the cardiac muscle action gene. American Journal of Human Genetics 44: 397-401.

Lombard, V., C.P. Baril, P. Dubreuil, F. Blouet and D. Zhang, 2000. Genetic relationships and fingerprinting of rapeseed cultivars by AFLP: Consequences for variety registration. Crop Science 40: 1417-1425.

Lucchin, M., G. Barcaccia and P. Parrini, 2003. Characterisation of a flint maize (*Zea mays L. convar. mays*) Italian landrace: I. Morpho-phenological and agronomic traits. Genetic Resources and Crop Evolution 50: 315-327.

Magoma, G.N., M. Obanda, M. Imbuga and S.G. Agong, 2000. The use of catechins as biochemical markers in diversity studies of tea (*Camellia sinensis*). Genetic Resources and Crop Evolution 47 (2): 107-114.

Maguire, T.L., R. Peakall and P. Saenger, 2002. Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. Theoretical and Applied Genetics 104: 388-398.

Markert, C.L. and F. Moller, 1959. Multiple forms of enzymes: tissue ontogenetic and species specific patterns. Proceedings of National Academy of Science 45: 753-763.

Mars, M. and M. Marrakchi, 1999. Diversity of pomegranate (*Punica granatum* L.) germplasm in Tunisia. Genetic Resources and Crop Evolution 46: 461-467.

Martin, M.J., F. Pablos and A.G. Gonzalez, 1998. Discrimination between Arabica and Robusta green coffee varieties according to their chemical composition. Talanta-Oxford 46: 1259-1264.

Martinello, G.E., N.R. Leal, A.T. Amaral Junior, M.G. Pereira and R.F. Daher, 2001. Comparison of morphological characteristics and RAPD for estimating genetic diversity in *Abelmoschus* Sp. Acta Horticulturae 546: 101-104.

Maughan, P.J., M.A. Saghai Maroof, G.R. Buss and G.M. Huestis, 1996. Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. Theoretical and Applied Genetics 93: 392-401.

Mazzafera, P. and A. Carvalho, 1992. Breeding for low seed caffeine content of coffee (*Coffea* L.) by interspecific hybridization. Euphytica 59:55-60.

McGregor, C.E, M.M. Greyling and L. Warnich, 2000. The use of simple sequence repeats (SSRs) to identify commercially important potato (*Solanum tuberosum* L.) cultivars in South Africa. South African Journal of Plant and Soil 17: 177-179.

McMillin, D.E., 1983. Plant isozymes: A historical perspective. In: S.D. Tanksley and J.J. Orton (Eds.), Isozymes in plant genetics and breeding, Part A, pp. 3-13. Elsevier Science Publishers, Amsterdam.

Mengesha, M.H., 1975. Crop germplasm diversity and resources in Ethiopia. In: O.H. Frankel and J.G. Hawkes (Eds.), Crop genetic resources for today and tomorrow, pp. 449-453. Cambridge University Press, London.

Meyer, E., 1965. Notes on the wild *Coffea arabica* from Southwestern Ethiopia, some historical considerations. Economic Botany 19: 136-151.

Milbourne, D., R. Meyer, J.E. Bradshaw, E. Baird, N. Bonar, J. Provan, W. Powell and R. Waugh, 1997. Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Molecular Breeding 3: 127-136.

Miller, J.C. and S.D. Tanksley, 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theoretical and Applied Genetics 80: 437-448.

Mitchell, H.W., 1988. Cultivation and harvesting of the arabica coffee tree. In: R.J. Clarke and R. Macrae (Eds.), Coffee agronomy, pp. 43-90. Elsevier Applied Science, London.

Miyashita, N.T., A. Kawabe and H. Innan, 1999. DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism analysis. Genetics 152: 1723-1731.

Mohan, M., S. Nair, A. Bhagwat, T.G. Krishna, M. Yano, C.R. Bhatia and T. Sasaki, 1997. Genome mapping, molecular markers and marker assisted selection in crop plants. Molecular Breeding 3: 87-103.

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

Montagnon, C., B. Guyot, C. Cilas and T. Leroy, 1998. Genetic parameters of several biochemical compounds from green coffee, *Coffea canephora*. Plant Breeding 117: 576-578.

Montagnon, C. and P. Bouharmont, 1996. Multivariate analysis of phenotypic diversity of *Coffea arabica* L. Genetic Resources and Crop Evolution 43: 221-227.

Montarroyos, A.V.V., M.A.G. de Lima, E.O. dos Santos and J.G.E. de Franca, 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.

Moreno, G., E. Moreno and G. Cadena, 1995. Bean characteristics and cup quality of the Colombian variety (*Coffea arabica*) as judged by international tasting panels. In the proceedings of the 16th International Scientific Colloquium on Coffee, pp. 574-583. Kyoto, Japan.

Moschetto, D., C. Montagnon, B. Guyot, J.J. Pierrot, T. Leroy, and A.B. Eskes, 1996. Studies on the effect of genotype on cup quality of *Coffea canephora*. Tropical Science 36: 18-31.

Mosjidis, J.A., S.L. Greene, K.A. Klingler and A. Afonin, 2004. Isozyme diversity in wild red clover populations from Caucasus. Crop Science 44: 665-670.

Moufida, S. and B. Marzouk, 2003. Biochemical characterization of blood orange, sweet orange, lemon, bergamot and bitter orange. Phytochemistry 62: 1283-1289.

Muschler, R.G., 2001. Shade improves coffee quality in a sub-optimal coffee zone of Costa Rica. Agroforestry systems 51:131-139.

Narasimhaswamy, R.L., 1962. Some thoughts on the origin of *Coffea arabica*. Turrialba 4: 1-5.

Nathanson, J.A., 1984. Caffeine and related methylxanthines: possible naturally occurring pesticides. Science 226: 184-187.

Newbury, H.J. and B.V. Ford-Lloyd, 1993. The use of RAPD for assessing variation in plants. Plant Growth Regulation 12: 43-51.

Nissim, Y., F. Jinggui, S. Arik, P. Neta, L. Uri and C. Avner, 2004. Phenotypic and genotypic analysis of a commercial cultivar and wild populations of *Anemone coronaria*. Euphytica 136: 51-62.

Noir, S., F. Anthony, B. Bertrand, M.C. Combes and P. Lashermes, 2003. Identification of a major gene (*Mex-1*) from *Coffea canephora* conferring resistance to *Meloidogyne exigua* in *Coffea arabica*. Plant Pathology 52: 97-103.

Ogita, S., H. Uefuji, Y. Yamaguchi, N. Koizumi and H. Sano, 2003. Producing decaffeinated coffee plants. Nature 423: 823.

Onguso, J.M., E.M. Kahangi, D.W. Ndiritu and F. Mizutani, 2004. Genetic characterization of cultivated bananas and plantains in Kenya by RAPD markers. Scientia Horticulturae 99: 9-20.

Orozco-Castillo, K., J. Chalmers, R. Wagh and W. Powell, 1994. Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theoretical and Applied Genetics 87: 934-940.

Ortiz, R., S.F. Mohamed, S. Madsen, J. Weibull and J.L. Christiansen, 2001. Assessment of phenotypic variation in winter barley in Scandinavia. Acta Agriculturae Scandinavia 51: 151-159.

Owuor, O.J.B., 1988. An assessment of the cup quality of the new disease resistant *Coffea arabica* cultivar Ruiru 11 in Kenya. Acta Horticulturae 224: 383-388.

Paczos-Grzeda, E., 2004. Pedigree, RAPD and simplified AFLP-based assessment of genetic relationships among *Avena sativa* L. cultivars. Euphytica 138: 13-22.

Paillard, M., P. Lashermes and V. Petiard, 1996. Construction of a molecular linkage map in coffee. Theoretical and Applied Genetics 93: 41-47.

Peakall, R., S. Gilmore, W. Keys, M. Morgante and A. Rafaske, 1998. Cross-species amplification of soybean (*Glycine max*) simple sequence repeat (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Molecular Biology and Evolution 15: 1275-1287.

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. Theoretical and Applied Genetics 108: 829-835.

Pejic, I., P. Ajmone-Marsan, M. Morgante, V. Kozumplick, P. Castiglioni, G. Taramino and M. Motto, 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. Theoretical and Applied Genetics 97: 1248-1255.

Pence, V.V., 1995. Cryopreservation of recalcitrant seeds. In: Y.P.S. Bajaj (Ed.), biotechnology in agriculture and forestry: Cryopreservation of plant germplasm, pp. 29-50. Springer, Berlin.

Perez de la Vega, M., 1994. Biochemical characterisation of populations. In: M.D. Hayward, N.O. Bosemark and I. Romagosa (Eds.), Plant breeding: principles and prospects, pp. 184-200. Chapman and Hall, London.

Petracco, M., 2000. Organoleptic properties of Espresso coffee as influenced by coffee botanical varieties. In: T. Sera, C.R. Soccol, A. Pandey and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 347-355. Kluwer Academic Publishers, Dordrecht.

Pinto, L.R., M.L.C. Vieira, C.L. de Souza Jr. and A.P. de Souza, 2003. Genetic diversity assessed by microsatellites in tropical maize populations submitted to a high-intensity reciprocal recurrent selection. Euphytica 134: 277-286.

Pinto-Maglio, C.A.F. and N.D. Da Cruz, 1998. Pachytene chromosome morphology in *Coffea* L. II. *Coffea arabica* L. complement. Caryologia 51: 19-35.

Portis, E., C. Comino, A. Lenzi, P. Lombardi, R. Tesi and S. Lanteri, 2004. Amplified fragment length polymorphism for variety identification and genetic diversity assessment in Oleander (*Nerium oleander* L.). Euphytica 136: 125-137.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski, 1996. The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding 2: 225-238.

Prabhuji, S.K., G.C. Srivastava, S.J.H. Rizvi and S.N. Mathur, 1983. 1,3,7-Trimethylaxanthine (caffeine); a new natural fish fungicide. Experientia 39: 177-179.

Prakash, N.S., D.V. Marques, V.M.P. Varzea, M.C. Silva, M.C. Combes and P. Lashermes, 2004. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. Theoretical and Applied Genetics 109: 1311-1317.

Prakash, N.S., M.C. Combes, N. Somanna and P. Lashermes, 2002. AFLP analysis of introgression in coffee cultivars (*Coffea arabica* L.) derived from a natural interspecific hybrid. Euphytica 124: 265-271.

Prasad, M., R.K. Varshney, J.K. Roy, H.S. Balyan and P.K. Gupta, 2000. The use of microsatellite for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theoretical and Applied Genetics 100: 584-592.

Qian, S.G., Y.L. Zheng, Y. Yan and J.L. Liu, 1999. Polymorphism of 8 allozymes at 17 loci from maize synthesized populations and correlated with quantitative traits. Acta Agronomy 25: 759-765.

Quero-Garcia, J., J.L. Noyer, X. Perrier, J.L. Marchand and V. Lebot, 2004. A germplasm stratification of taro (*Colocasia esculenta*) based on agro-morphological descriptors, validation by AFLP markers. Euphytica 137: 387-395.

Rafalski, J.A., 1997. Randomly amplified polymorphic DNA (RAPD) analysis. In: G. Caetano-Anolles and P.M. Gresshoff (Eds.), DNA markers protocols, applications and overviews, pp. 75-83. Wiley-Vch, New York.

Raina, S.N., Y. Mukai and M. Yamamoto, 1998. *In situ* hybridization identifies the diploid progenitor species of *Coffea arabica* (Rubiaceae). Theoretical and Applied Genetics 97: 1204-1209.

Rajagopal, J., L. Bashyam, S. Bhatia, D.K. Khurana, P.S. Srivastava and M.L. Kumaran, 2000. Evaluation of genetic diversity in the Himalayan poplar using RAPD markers. Silvae Genetica 49:60-66.

Raju, K.R., S. Vishveswara and C.S. Srinivasan, 1978. Association of some characteristics with cup quality in *Coffea canephora x Coffea arabica* hybrids. Indian Coffee 42: 195-197.

Rakotomalala, J.J.R., E. Cros, M.N. Clifford and A. Charrier, 1992. Caffeine and theobromine in green beans from *Mascarocoffea.* Phytochemistry 31: 1271-1272.

Ramsay, L., M. Macaulay, R. Cardle, M. Morgante, S. Gegli-Ivanissevich, E. Maestri, W. Powell and R. Waugh, 1999. Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plantation Journal 17: 415-425.

Razdan, M.K. and E.C. Cocking, 1997. Biotechnology in conservation of genetic resources. In: M.K. Razdan and E.C. Cocking (Eds.), Conservation of plant genetic resources *in vitro*, pp.3-25. Science Publishers, Enfield, New Hampshire.

Renganayaki, K., J.C. Read and A.K. Fritz, 2001. Genetic diversity among Texas bluegrass genotypes (*Poa arachnifera* Torr.) revealed by AFLP and RAPD markers. Theoretical and Applied Genetics 102: 1037-1045.

Roche, D., 1995. Coffee genetics and quality. In the proceedings of the 16th International Scientific Colloquium on Coffee, pp. 584-588. Kyoto, Japan.

Rohlf, F.J., 1998. NTSYS-pc: numerical taxonomy and multivariate analysis system. Version 2.02i. Exter Software, Setauket, New York.

Rotondi, A., M. Magli, C. Ricciolini and L. Baldoni, 2003. Morphological and molecular analyses for characterization of a group of Italian olive cultivars. Euphytica 132: 129-137.

Rovelli, P., R. Mettulio, F. Anthony, F. Anzueto, P. Lashermes and G. Graziosi, 2000. Microsatellites in *Coffea arabica* L. In: T. Sera, C.R. Soccol, A. Pandey and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 123-133. Kluwer Academic Publishers, Dordrecht.

Ruas, P.M., L.E.C. Diniz, C.F. Ruas and T. Sera, 2000. Genetic polymorphism in species and hybrids of *Coffea* revealed by RAPD. In: T. Sera, C.R. Soccol, A. Pandey and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 187-195. Kluwer Academic Publishers, Dordrecht.

Ruiz, R., J. Dendauw, E. Van Bockstaele, A. Depicker and M. De Loose, 2000. AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* Sp.). Molecular Breeding 6: 125-134.

Russell, J.R., J.D. Fuller, M. Macaulay, B.G. Hatz, A. Jahoor, W. Powell and R. Waugh, 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theoretical and Applied Genetics 95: 714-722.

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen and R.W. Allard, 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proceedings of National Academy of Science, USA. 81: 8014-8019.

Saini, N., N. Jain, S. Jain and R.K. Jain, 2004. Assessment of genetic diversity within and among Basmati and non-Basmati rice varieties using AFLP, ISSR and SSR markers. Euphytica 140: 133-146.

Sakiyama, N.S., 2000. DNA markers for coffee tree breeding. In: T. Sera, C.R. Soccol, A. pandy and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 179-185. Kluwer Academic Publishers, Dordrecht.

Sanchez-Perez, R., D. Ruiz, F. Dicenta, J. Egea and P. Martinez-Gomez, 2005. Application of simple sequence repeat (SSR) markers in apricot breeding: molecular characteristion, protection and genetic relationships. Scientia Horticulturae 103: 305-315.

Selvakumar, M. and M.S. Sreenivasan, 1989. Studies on morphology and quality of Ethiopian arabica coffee. Journal of Plantation Crops 16: 321-324.

Sensi, E., R. Vignani, M. Scali, E. Masi and M. Cresti, 2003. DNA fingerprinting and genetic relatedness among cultivated varieties of *Olea europaea* L. estimated by AFLP analysis. Scientia Horticulturae 97: 379-388.

Shan, F., H. Clarke, G. Yan, J.A. Plummer and K.H.M. Siddique, 2004. Development of DNA fingerprinting keys for discrimination of *Cicer echinospermum* (P.H. Davis) accessions using AFLP markers. Australian Journal of Agricultural Research 55: 947-952.

Shannon, C.E. and W. Weaver, 1949. The mathematical theory of communication. University of Illinois Press, Urbana.

Shattuck-Eidens, D.M., R.N. Bell, S.L. Neuhausen and T. Helentjaris, 1990. DNA sequence variation within maize and melon: observations from polymerase chain reaction amplification and direct sequencing. Genetics 126: 207-217.

Shetty, M., K. Subanaya and P.G. Shivananda, 1994. Antibacterial activity of tea (*Camella sinensis* L.) and coffee (*Coffea arabica* L.) with special reference *Salmonella typhimurium.* Journal of Communicable Diseases 26: 140-150.

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

Silvarolla, M.B., P. Mazzafera and M.M. Alves de Lima, 2000. Caffeine content of Ethiopian *Coffea arabica* beans. Genetics and Molecular Biology 23: 213-215.

Simioniuc, D., R. Uptmoor, W. Friedt and F. Ordon, 2002. Genetic diversity and relationships among pea cultivars revealed by RAPDs and AFLPs. Plant Breeding 121: 429-435.

Singh, R.K. and B.D. Chaudhary, 1985. Biometrical methods in quantitative genetic analysis. Kalyani publishers, New Delhi.

Sivetz, M., 1963. Coffee processing technology: Aromatization-properties-brewing-decaffeination-plant design. The Avi Publishing Company, Westport, Connecticut.

Smith, R.F., 1985. A history of coffee. In: M.N. Clifford and K.C. Willson (Eds.), Coffee botany, biochemistry and production of beans and beverage, pp. 1-12. Croom Helm, London.

Smith, J.S.C. and O.S. Smith, 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.

Smith, J.S.C., S. Kresovich, M.S. Hopkins, S.E. Mitchell, R.E. Dean, W.L. Woodman, M. Lee and K. Porter, 2000. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Science 40: 226-232.

Smulders, M.J.M., Y. Noordijk, W. Rus-Kortekaas, G.M.M. Bredemeijer and B. Vosman, 2003. Microsatellite genotyping of carnation varieties. Theoretical and Applied Genetics 106: 1191-1195.

Sosinski, B., M. Gannavarapu, L.D. Hager, L.E. Beck, G.J. King, C.D. Ryder, S. Rajapakse, W.V. Baird, R.E. Ballard and A.G. Abbott, 2000. Characterisation of microsatellite markers in peach (*Prunus persica* (L.) Batsch). Theoretical and Applied Genetics 101: 421-428.

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

Struss, D and J. Plieske, 1998. The use of microsatellite markers for detection of genetic diversity in barley populations. Theoretical and Applied Genetics 97: 308-315.

Sylvain, P.G., 1955. Some observations on *Coffea arabica* L. in Ethiopia. Turrialba 5: 37-53.

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

Tachida, W. and M. Iizuka, 1992. Persistence of repeated sequences that evolve by replication slippage. Genetics 131: 471-478.

Tadesse, W. and E. Bekele, 2003. Phenotypic diversity of grass pea (*Lathyrus sativus* L.) in relation to geographical regions and altitudinal range. Genetic Resources and Crop Evolution 50: 497-505.

Tanksley, S.D. and T.J. Orton, 1983. Isozymes in plant genetics and breeding. Elsevier, Amsterdam.

Tanya, P., P. Srinives, T. Toojinda, A. Vanavichit, B.K. Ha, J.S. Bae, J.K. Moon and S.H. Lee, 2001. Evaluation of genetic diversity among soybean genotypes using SSR and SNP. Korean Journal of Crop Science 46: 334-340.

Taramino, G., R. Tarchini, S. Ferrario, M. Lee and M.E. Pe, 1997. Characterisation and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theoretical and Applied Genetics 95: 66-72.

Taramino, G. and S. Tingey, 1996. Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39: 277-287.

Tatineni, V.R., G. Cantreil, and D.D. Davis, 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Science 36: 186-192.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA. Nucleic Acids Research 17: 6463-6471.

Teketay, D., 1999. History, botany and ecological requirements of coffee. Walia 20: 28-50.

Temnykh, S., W.D. Park, N. Ayres, S. Cartinhour, N. Hauch, L. Lipovich, Y.G. Cho, T. Ishii and S.R. McCouch, 2000. Mapping and genome organization of microsatellite sequences in rice (*Orysa sativa* L.). Theoretical and Applied Genetics 100: 697-712.

Tomkins, J.P., T.C. Wood, L.S. Barnes, A. Westman and R.A. Wing, 2001. Evaluation of genetic variation in daylily (*Hemerocallis* Sp.) using AFLP markers. Theoretical and Applied Genetics 102: 489-496.

Tonguc, M. and P.D. Griffiths, 2004. Genetic relationships of Brassica vegetables determined using database derived simple sequence repeats. Euphytica 137: 193-201.

Tramontano, W.A. and D. Jouve, 1997. Trigonelline accumulation in salt stressed legumes and the role of other osmoregulators as cell cycle control agents. Phytochemistry 44: 1037-1040.

Tressl, R., M. Holzer and H. Kamperschoer, 1982. Bildung von Aromastoffen in Rostkaffe in Abhangigkeit vom Gehalt an Aminosauren und reduzierenden Zuckern. Proceedings of International Congress. ASIC 10, 279-292.

Tsegaye, Y., O. Getachew and Z. Tesfaye, 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.

Turpeinen, T., T. Vanhala, E. Nevo and E. Nissila, 2003. AFLP genetic polymorphism in wild barley (*Hordeum spontaneum*) populations in Israel. Theoretical and Applied Genetics 106: 1333-1339.

Tyrka, M., P. Dziadczyk and J.A. Hortynski, 2002. Simplified AFLP procedure as a tool for identification of strawberry cultivars and advanced breeding lines. Euphytica 125: 273-280.

Ude, G., M. Pillay, D. Nwakanma and A. Tenkouano, 2002. Analysis of genetic diversity and sectional relationships in *Musa* using AFLP markers. Theoretical and Applied Genetics 104: 1239-1245.

Upadhyaya, H.D., 2003. Phenotypic diversity in groundnut (*Arachis hypogaea* L.) core collection assessed by morphological and agronomic evaluations. Genetic Resources and Crop Evolution 50: 539-550.

Upadhyaya, H.D., R. Ortiz, P.J. Bramel and S. Singh, 2002. Phenotypic diversity for morphological and agronomic characteristics in chickpea core colloction. Euphytica 123: 333-342.

Uptmoor, R., W. Wenzel, W. Friedt, G. Donaldson, K. Ayisi and F. Ordon, 2003. Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from South Africa by RAPDs, AFLPs and SSRs. Theoretical and Applied Genetics 106: 1316-1325.

Van de Wouw, M., N. Maxted and B.V. Ford-Lioyd, 2003. Agro-Morphological characterisation of common vetch and its close relatives. Euphytica 130: 281-292.

Van der Voort, J.R., P. Wolters, R. Folkertsma, R. Hutten, P. Van Zandvoort, H. Vinke, K. Kanyuka, A. Bendahmane, E. Jacobsen, R. Janssen and J. Bakker, 1997. Mapping of the cyst nematode resistance locus Gpa2 in potato using a strategy based on co-migrating AFLP markers. Theoretical and Applied Genetics 95: 874-880.

Van der Vossen, H.A.M., 1985. Coffee selection and breeding. In: M.N. Clifford and K.C. Willson (Eds.), Coffee botany, biochemistry and production of beans and beverage, pp. 49-96. Croom Helm, London.

Van der Vossen, H.A.M. and D.J. Walyaro, 1983. Breeding for resistance to coffee berry disease in *Coffea arabica* L. II. Inheritance of the resistance. Euphytica 29: 777-791.

Van Droogenbroeck, B., P. Breyne, P. Goetghebeur, E. Romeijn-Peeters, T. Kyndt and G. Gheysen, 2002. AFLP analysis of genetic relationships among papaya and its wild relatives (*Caricaceae*) from Ecuador. Theoretical and Applied Genetics 105: 289-297.

Vaz-Patto, M.C., Z. Satovic, S. Pego and P. Fevereiro, 2004. Assessing the genetic diversity of Portuguese maize germplasm using microsatellite markers. Euphytica 137: 63-72.

Viniegra-Gonzalez, G., 2000. Biotechnology and the future of coffee production. In: T. Sera, C.R. Soccol, A. Pandey and S. Roussos (Eds.), Coffee biotechnology and quality, pp. 1-16. Kluwer Academic Publishers, Dordrecht.

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

Waldhauser, S.S.M. and T.W. Baumann, 1996. Compartmentation of caffeine and related purine alkaloids depends exclusively on the physical chemistry of their vacuolar complex formation with chlorogenic acids. Phytochemistry 42: 985-996.

Walyaro, D.J.A., 1983. Considerations in breeding for improved yield and quality in arabica coffee (*Coffea arabica* L.). A PhD thesis, Wageningen Agricultural University.

Walyaro, D.J. and H.A.M. Van der Vossen, 1979. Early determination of yield potential in arabica coffee by applying index selection. Euphytica 28: 465-472.

Wang, Y.H., C.E. Thomas, and R.A. Dean, 1997. A genetic map of melon (*Cucumis melo* L.) based on amplified fragment length polymorphism (AFLP) markers. Theoretical and Applied Genetics 95: 791-798.

Wang, Z.J., L. Weber, G. Zhong and S.D. Tanksley, 1994. Survey of plant short tandem DNA repeats. Theoretical and Applied Genetics 88: 1-6.

Waugh, R., N. Bonar, E. Baird, B. Thomas, A. Graner, P. Hayes and W. Powell, 1997. Homology of AFLP products in three mapping populations of barley. Molecular and General Genetics 255: 311-321.

Weber, J.L. and P.E. May, 1989. Abundant class of human DNA polymorphisms which can be typed using polymerase chain reaction. American Journal of Human Genetics 44: 388-396.

Weeden, N.F., 1989. Genetics of plant isozymes. In: D.E. Soltis and P.S. Soltis (Eds.), Isozymes in plant biology, pp. 46-72. Chapman and Hall, London.

Weising, K., H. Nybom, K. Wolff and W. Meyer, 1995. DNA fingerprinting in plants and fungi. CRC Press, Boca Raton.

Wellman, F.L., 1961. Coffee: botany, cultivation and utilization. Leonard Hill Books Limited, London.

Welsh, J. and M. McClelland, 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18: 7213-7218.

Wendel, J.F. and N.F. Weeden, 1989. Visualization and interpretation of plant isozymes. In: D.E. Soltis and P.S. Soltis (Eds.), Isozymes in plant biology, pp. 5-45. Chapman and Hall, 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. Theoretical and Applied Genetics 98: 557-563.

Westman, A.L. and S. Kresovich, 1999. Simple sequence repeats (SSR) based marker variation in *Brassica nigra* gene bank accessions and weed populations. Euphytica 109: 85-92.

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

Willson, K.C., 1999. Coffee, Cocoa and Tea. CABI publishing, Wallingford.

Woelore, W.M., 1995. Parchment arabica coffee storage. In the proceedings of the 16th International Scientific Colloquium on Coffee, pp. 565-573. Kyoto, Japan.

Woldemariam, T., M. Denich, D. Teketay and P.L.G. Vlek, 2002. Human impacts on *Coffea arabica* L. genetic pools in Ethiopia and the need for its *in situ* conservation. In: J.M.M. Engels, V.R. Rao, A.H.D. Brown and M.T. Jackson (Eds.), Managing plant genetic diversity, pp. 237-247. IPGRI, Rome.

Wondimu, M., 1998. The genetic diversity of Ethiopian coffee. Kaffa Coffee 1: 25-30.

Wood, A.J., 1999. Composition of salt-induced osmotic adjustment and trigonelline accumulation in two soybean cultivars. Biologia-plantarum 42: 389-394.

Wrigley, G., 1988. Coffee. Tropical Agriculture Series. Longman Scientific and Technical, Longman group UK Limited, England.

Wu, X., K. Skog and M. Jogerstad, 1997. Trigonelline, a naturally occurring constituent of green coffee beans behind the mutagenic activity of roasted coffee. Mutation research, genetic toxicology and environmental mutagenesis 391: 171-177.

Xu, W.J., B.W. Wang and K.M. Cui, 2004. RAPD and SCAR markers linked to sex determination in *Eucommia ulmoides* Oliv. Euphytica 136: 233-238.

Yee, E., K.K. Kidwell, G.R. Sills and T.A. Lumpkin, 1999. Diversity among selected *Vigna angularis* (Azuki) accessions on the basis of RAPD and AFLP markers. Crop Science 39: 268-275.

Yemane-Berhan, Y., 1998. Coffee production in Ethiopia. Kaffa Coffee 1: 31-35.

Yu, C.Y., S.W. Hu, H.X. Zhao, A.G. Guo and G.L. Sun, 2005. Genetic distances revealed by morphological characters, isozymes, proteins, and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.). Theoretical and Applied Genetics 110: 511-518.

Zeid, M., C.C. Schon and W. Link, 2004. Hybrid performance and AFLP-based genetic similarity in fababean. Euphytica 139: 207-216.

Zeng, J., Y. Zou, J. Bai and H. Zheng, 2003. RAPD analysis of genetic variation in natural populations of *Betula alnoides* from Guangxi, China. Euphytica 134: 33-41.

Zeven, A.C., 1990. Classification of landraces and improved cultivars of rivet wheat (*Triticum turgidum*) and bread wheat (*T. aestivum*) from Great Britain and described in 1934. Euphytica 47: 249-258.

Zewdie, Y. and A.C. Zeven, 1997. Variation in Yugoslavian hot pepper (*Capsicum annuum* L.) accessions. Euphytica 97: 81-89.

Zoghlami, A. and M. Zouaghi, 2003. Morphological variation in *Astragalus hamosus* L. and *Coronilla scorpioides* L. populations of Tunisia. Euphytica 134: 137-147.