GENETIC VARIATION IN COFFEE ACCESSIONS IN KENYA AND INTROGRESSION FROM ROBUSTA TO ARABICA USING RANDOM AMPLIFIED POLYMORPHIC DNA AND MICROSATELLITES

Omingo Daniel Omari (B.Sc.)
Reg.No. I56/21025/2012

A Thesis Submitted in Partial Fulfillment of the Requirements for the Award of the Degree of Masters of Science (Biotechnology) in the School of Pure and Applied Sciences of Kenyatta University

MAY 2015

DECLARATION

This Thesis is my original work, except where due reference is made in the text, and has not been presented for the Award of a degree in this or any other University.

Omingo Daniel Omari (B.Sc.)

Reg.No. I56/21025/2012

Signed……. Date…….

We confirm that the work reported in this thesis was carried out by the candidate under our supervision.

Supervisors

Dr. Stephen Runo
Department of Biochemistry and Biotechnology
Kenyatta University
P.O. Box 43844-00100
Nairobi

Signature ……Date…….....

Dr. Daniel Okun
Department of Biochemistry and Biotechnology
Kenyatta University
P.O. Box 43844-00100
Nairobi

Signature ……..Date……

DEDICATION

This piece of work is dedicated to my parents Moses Omingo and Esther Nyabonyi, my wife Dorca Obwogi and my daughter Alice Maryanne.

ACKNOWLEDGEMENT

First and foremost I wish to acknowledge the almighty GOD for giving me a breath of life, the energy, health and the spirit to persevere through the long and often challenging process of writing this thesis.Secondly, I would like to express my sincere gratitude and appreciationto to my supervisors, Dr. Steven Runo and Dr.Daniel Okun for their encouragement,valuable guidance, unreserved support and willingness to supervise my research and there valuable comments from early stage of proposing the research to the final thesis research results write up, which helped me in the completion of the study. I would like to extend my deepest gratitude to Dr.Chrispine Omondi and Peter Goco for their contribution, guidance, continuous technical support, moral support and commitment throughout my research. I gratefully acknowledge the financial support from the Coffee Research Institute throughout my entire research period. I express my sincere feelings and appreciation to Happiness Msenya, Loise Wambui and Brian Maina for their technical and moral support throughout my research. I am grateful for the support of my fellow students Sammy Muragoli, Allan Kigungo and George Ligabo for their encouragement throughout my research. My heartfelt appreciation also goes to my family members, my father Moses Omingo, my mother Esther Nyabonyi, My sisters Consolata, Dyphina, Janet, Jack, Regina, late Dr.Alice and Linnus for their unconditional, moral and financial support and finally to my wife for her tireless, relentless constant encouragement, care and moral support without which, this research coud not have been asuccess.

TABLE OF CONTENTS

DECLARATION ii
DEDICATION iii
ACKNOWLEDGEMENT iv
TABLE OF CONTENTS v
LIST OF TABLES vii
LIST OF FIGURES viii
LIST OF ABBREVIATIONS AND ACRONYMS ix
ABSTRACT x
CHAPTER ONE 1
INTRODUCTION 1
1.1 Background information of the study 1
1.2 Problem statement and justification 4
1.3 Research questions 5
1.4 Hypotheses 5
1.5 Objectives 5
1.5.1 General objective 5
1.5.2 Specific objectives 6
CHAPTER TWO 7
LITERATURE REVIEW 7
2.1 Coffee growing regions in kenya 7
2.2 Coffee origin and distribution 8
2.3 General description of Coffee 9
2.4 Coffee propagation 10
2.5 Coffee mutants 11
2.6 Coffee improvement 11
2.7 Implication for breeding 13
2.8 Conservation strategies 13
2.9 Economic significance of Coffee in Kenya 14
2.10 Production and marketing of coffee 15
2.11 Evaluation of genetic variation of Coffee using markers 18
2.11.1 Morphological markers 20
2.11.2 Biochemical markers 20

2.11.3 Molecular markers 22
2.11.3.1 PCR-based markers 22
2.11.3.1.2 Simple Sequence Repeats (SSR) 23
2.11.3.1.3 SSR markers in coffee 24
2.11.3.1.4 Random amplified polymorphic DNA (RAPD) 25
2.11. 3.1.5 RAPD markers in coffee 26
2.11.3.1.6 AFLP (Amplified Fragment Length Polymorphism) 27
2.11.3 .1.7 AFLP in Coffee (Amplified Fragment Length Polymorphism) 28
2.11.4 Gene introgression 29
CHAPTER THREE 31
MATERIALS AND METHODS 31
3.1 Plant materials 31
3.2 Genomic DNA extraction 32
3.3 DNA quantification 33
3.4 Amplification of coffee genomic DNA using RAPDS 34
3.6 Data analysis 36
CHAPTER FOUR 37
RESULTS 37
4.1 Genetic diversity of 24 Coffee accessions as revealed by RAPD primers 37
4.2 Genetic diversity of 24 Coffee accessions as revealed by SSR primers 41
4.3 Gene introgression 45
CHAPTER FIVE 47
DISCUSSION, CONCLUSIONS AND RECOMENDATIONS 47
5.1 DISCUSSION 47
5.2 CONCLUSIONS 55
5.3 RECOMMENDATIONS 55
APPENDICES 68

LIST OF TABLES

Table 2.1: Coffee production (Thousand 60-Kilogram Bags).....16

Table 2.2: Total coffee exports (Thousand 60-Kilogram Bags)……..….17

Table 3.1: Status and sources of coffee germplasm used in the study.....31

Table 3.2: RAPD primers used for PCR analysis of 24 coffee accessions.....34

Table 3.3: SSR primers used for PCR analysis of 24 coffee accessions.....35

Table 4.1: Amplification products generated by RAPD primers.....38

Table 4.2: Matrix of genetic distance based on RAPD Markers.....39

Table 4.3:Amplification products generated by SSR primers.....42

Table4.4: Matrix of genetic distance based on SSR markers.....43

Table 4.5: Distribution of Canephora introgreesed fragments in arabica.....46

LIST OF FIGURES

Figure 4.1: RAPD profile generated by primer L-18.....37

Figure 4.2: Dendrogram of RAPD markers using the UPGMA method.....40

Figure 4.3: SSR profile generated by primer Sat 240.....41

Figure 4.4: Dendrogram constructed by cluster analysis using SSR markers.....44

Figure 4.5: Sat 240 showing *Canephora* introgressed fragments.....45

Figure 4.6: Sat 172 showing *Canephora* introgressed fragments.....45

LIST OF ABBREVIATIONS AND ACRONYMS

AFLP	Amplified Fragment Length Polymorphism
CBD	Coffee Berry Disease
CLR	Coffee Leaf Rust
CRS	Coffee Research Station
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
EDTA	Ethylenediaminetetraacetic acid
MATAB	Mixed alkyltrimethylammonium bromide
Nacl	Sodium chloride
PCR	Polymerase Chain Reaction
RAPD	Randomly Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
Rnase	Ribonucleic acid enzyme
SSR	Simple Sequence Repeats
TBE	Tris Boric Ethylenediaminetetraacetic acid
TE	Tris Ethylenediaminetetraacetic acid
UPGMA	Unweighted Pair Group Method using Arithmetic Averages
UV	Ultra violet

ABSTRACT

Coffee provides one of the most widely drunk beverages in the world, and is a very important source of foreign exchange income for many countries. In kenya, coffee production has a significant contribution to Kenya's economy for decades, and a high proportion of the coffee produced is considered the best quality coffee in the world. In coffee, morphological parameters are very often used to discriminate the varieties and hybrids. However, this exercise has some challenges that include the perennial nature of the plant that requires at least 5-7 years for attaining reproductive maturity for evaluation of both vegetative and reproductive characters. Therefore it is critical to identify suitable markers which can identify cultivars/hybrids at early stage of plant growth and at the same time discriminate between different coffee genotypes to fasclitate selection process and thereby speed up the coffee breeding program.Two molecular marker systems, RAPD (Random Amplified Polymorphic DNA) and SSR (Simple Sequence Repeats) were employed for identification of genetic relationship of 24 coffee accessions and to test for gene introgression from *Coffea canephora* into *Coffea arabica* with the objective of providing important information for improvement and *in situ/ex situ* conservation of this species. The total number of bands, the distribution of bands across all species, polymorphic bands, species-specific bands and average number of bands per primer calculated. Genetic dissimilarities was estimated using Pearson dissimilarity. Cluster analysis was performed using Un-weighted Pair Group Method with Arithmetic Averages (UPGMA) using STATSTICA software version 8. A total of 79 bands were detected by 10 RAPD and 50 bands were detected by 13 SSR primers.The polymorphism detected by both markers ranged from 33% to 100% for SSR and 50% to 100% for RAPD with average polymorphism of 65% and 81% respectively. The genetic dissimilarity index among the genotypes ranged from 0.06 to 1 for both SSR and RAPD primers. In this study, UPGMA analysis for RAPD and SSR markers showed some similarities; the 24 coffee accessions clustered according to the three different species namely *C. eugenioides*, *C. canephora* (Robusta) and *C. arabica* (Arabica). Considering that the coffee genotypes evaluated in this study originated from diferent countries, the similarities (for both SSR and RAPD results) observed among Arabica genotypes, attests to the narrow genetic diversity among Arabica coffee.This study confirmed the low genetic diversity in Arabica coffee genotypes evaluated with average dissimilarity index of 0.5 . The study also widened the information on genetic diversity of coffee germplasm available for breeding programmes in Kenya since previous work was biased to commercial cultivars and donors of resistance to diseases .Of the thirteen SSR markers employed in this study, only three markers were able to detect introgressed *Canephora* DNA fragments present in arabica genome (Sat 254, Sat 240 and Sat 172). The percentage of introgressed *Canephora* DNAfragments ranged from 9.1 % to 27.3 %.The results demonstrated that RAPD were suitable for genetic diversity studies in coffee while SSR were suitable for gene introgression studies in coffee.

CHAPTER ONE

INTRODUCTION

1.1 Background information of the study

Coffee is rated the second most important commodity in global trade after petroleum products. It is grown in over 80 countries contributing substantially to their national economies providing a livelihood to about 25 million coffee farming families around the world (Pare, 2002). Coffee belongs to the genus *Coffea* in the Rubiaceae family and is mostly grown in the tropical and subtropical regions (Berthaud and Charrier, 1988). Of the 100 species in the genus *Coffea*, *Coffea arabica* L. (Arabica coffee) and *Coffea canephora* (Robusta coffee) are the two most important commercial species (Davis *et al*., 2006). *C. arabica* is considered to be of high beverage quality and contributes 70 percent of the world coffee production while Robusta contributes 30 percent (Lashermes *et al*., 1997; Lecolier *et al.*, 2009).

Arabica is believed to have originated in southwestern Ethiopia while Robusta coffee originated from central and western equatorial Africa (Ferwerda, 1976). *C. arabica* is predominantly autogamous (Krug and Carvalho, 1951; Wrigley, 1995) and the only *Coffea* species that is tetraploid (2n = 4x = 44) (Krug and Carvalho, 1951 Ferwerda, 1976, Charrier and Berthaud, 1985, Wrigley, 1995, Clarindo and Carvalho, 2008). Genomic analysis using Restriction Fragment Length Polymorphism (RFLP) of chloroplast DNA (cpDNA), which is maternally inherited, supports the notion that *Coffea eugenioides* donated the maternal genome while analysis of ribosomal DNA (rDNA) demonstrated that *Coffea canephora* donated the paternal genome (Lashermes *et al.*, 1995). Cytogenetic analysis established that *C. arabica* is an amphidiploid (allotetraploid; 2n=4x=44) formed by a recent natural hybridization

between the diploids *C. Canephora* and *C. eugenioides* (2n=2x=22) (Lashermes *et al.,* 1999). Due to autogamous nature, *C. arabica,* has a very narrow genetic base (Ferwerda,1976; Vossen,1985). The narrow diversity observed in *C. arabica* is believed to be a consequence of its reproductive biology, origin and evolution (Lashermes *et al.,* 1999; Anthony *et al.,* 2001). The remaining species are diploid with 2n=2x=22 chromosomes and are generally self-incompatible.

Through variety improvement, some of the recommended cultivars or commercialy cultivated cultivars in Kenya include K7, Batian, SL 28, Ruiru 11 and SL 34. Other coffee accessions from Ethiopia, Sudan, Kenya, Tanzania, Angola, India, Reunion, Portugal, Guatamala, South and Central America are conserved ex-situ at Coffee Research Institute (CRI) Ruiru in germplasm fields plots. Inspite of the commercial and social importance of the genus *Coffea*, the genetic relationship between the majorities of coffee species is not extensively studied and their taxonomic status is poorly understood. Understanding the genetic relationship between coffee species is not only important for resolving taxonomic ambiguity but also important for the genetic improvement program and conservation of potential populations.Some of these conserved genotypes have been used as progenitors in the coffee variety development at Coffee Research Institute.

Currently, different methods such as morphological, biochemical and molecular markers are available for estimation of genetic diversity within and among genotypes. The use of morphological techniques in diversity study of plants is limited by the influence of environmental factors and delays associated with phenotypic expression

of various growth stages of the plant (Weising *et al*., 2005). In addition, they are also few in number and require lengthy follow-up during the whole growth stage especially in perennial plants like coffee. In response to the limitation of morphological techniques, the more effective technique based on protein, isozymes, was developed. However, isozyme markers were found to be inappropriate for determining the genetic diversity in *C. arabica* due to fact that analysis of isozymes specifically in *C. arabica* accessions failed to reveal the amount of polymorphism detected using morphological markers (Berthaud and Charrier, 1988; Lopes ,1993; Bustamanate and Polanco, 1999).

A variety of molecular techniques have been developed to measure genetic variation at both interspecific and intraspecific levels in a number of plant species. Recent advances in the field of plant molecular genetics have resulted in the development of a series of DNA markers. DNA-based techniques are in use in different coffee genetic studies. These include the conventional RFLP method (Herrera *et al*., 2001; Crouzillat *et al*., 2004) and the different PCR-based methods such as RAPD (Orozco-Castillo *et al*., 1994), AFLP (Anthony *et al*., 2001a) and microsatellite (SSRs) markers (Lashermes *et al*., 1995; Lashermes *et al*., 1997). The importance of molecular markers for genetic improvement in perennial crops 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 backcross cycles required to restore the quality of the recurrent parent and for simultaneous improvement of different traits (Lashermes *et al.*, 2000b). 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.

1.2 Problem statement and justification

Coffea arabica exhibits low morphological variation in common attributes like leaf shape, leaf size, leaf color, leaf length, leaf width, internode length, plant shape, flower number, seed length, seed width, plant canopy and tree height. Because of high similarity in phenotypic appearance among the majority of Arabica collections, selection of parental lines for inter-varietals hybridization and identification of resultant hybrids at an early stage of plant growth is difficult. Morphological descriptors such as growth habit, leaf type, and floral characteristics and fruit morphology are used in a limited extent to characterize the various species. However, developing morphological descriptors for any particular species/cultivar has severe limitations as these characteristics are influenced by environmental conditions and the long generation time required for the expression of some characters. Uniformity for morphological traits in *C. arabica* could be attributed to the origin of the species and self fertile nature (Lashermes *et al*., 1999). Arabica Coffee originated from a relatively recent hybridization between *Coffea canephora* (Robusta coffee) and *C. eugenioides*. Occurrence of spontaneous hybrids between *C. arabica* and diploid relative species such as *C. canephora*, *C.eugenioides* and *C.liberi*ca is common, especially when these species grow in a close proximity (Cramer, 1957) and it brings about gene introgression within and across species.

In view of the above, molecular markers have been developed as an alternative technique that is quick and reliable in identifying closely related cultivars and that efficiently discriminates coffee hybrids and parents to enable various characters to be confirmed in a hybridization and selection process, and thereby speed up the coffee breeding programs. Noting that seventy percent of the world coffee production is from *C. arabica*; a thorough understanding of the genetic variation in *arabica* cultivars is critical to future Coffee improvement.The results of this study will help in breeding programs assist avoiding duplication of similar germplasm in gene banks (useful in varietal identification).

1.3 Research questions

i) Is there a possibility of gene transfer across and between species of coffee ?

ii) Is there any genetic diversity among various coffee species?

iii) Are the SSR markers more polymorphic than RAPD markers?

1.4 Hypotheses

i) The *C. arabica* does not contain introgressed DNA fragments from *C. canephora*

ii) There is no genetic diversity between the genotypes used under this study

iii) SSR are not more polymorphic than RAPD markers

1.5 Objectives

1.5.1 General objective

To determine genetic diversity among coffee accessions and introgression from *C. Canephora* into *C. arabica* using RAPD and SSR markers in Kenyan accessions.

1.5.2 Specific objectives

i) To identify DNA fragment introgressed from *C. canephora* into *C. arabica*

ii) To determine genetic diversity among genotypes of various coffee accessions used in this study

iii) To determine the most Polymorphic primer between SSR and RAPD markers.

CHAPTER TWO

LITERATURE REVIEW

2.1 Coffee growing regions in kenya

Today, coffee is grown in the highlands districts of Kenya: Kiambu, Muranga, Nyeri, Thika, and Kirinyaga in the former Central Province; Meru North, Meru Central, Meru South, Embu, Machakos and Kitui in the former Eastern Province; Nakuru, West Pokot, Kajiado, Baringo, Kericho, Nandi, Laikipia, Transnzoia, Uasin-Gishu, Keiyo, Marakwet and Kajiado in the former Rift Valley Province; Bungoma, Kakamega, and Busia in the former Western Province; Kisii, Siaya, Kisumu, and South Nyanza in the former Nyanza Province; and Taita in the former Coast Province. The high production zone is a triangle formed by Mt. Kenya, the Aberdare Range and Machakos Town essentially the former Central and Eastern Provinces which account for about 70 per cent coffee producing areas contain about 45 per cent of Kenya's population. Since some of these people are as much as 40 per cent income-dependent on coffee, their lives revolve around the fate of coffee.

Kenya coffee sector is composed of two categories of farms: the plantation sub-sector comprising of about 3,300 farms of which 300 are greater than 25 hectares; and the cooperative sub-sector of some 523 cooperative unions with about 700,000 smallholders cultivating about 120,000 hectares of coffee, equivalent to about 0.2 hectares apiece. It is estimated that a total of 170,000 hectares are under coffee and that 75 per cent of that total is organized around smallholder cooperatives. (Economic Survey, Central Bureau of Statistics, 2006, Nairobi, Kenya).

2.2 Coffee origin and distribution

Kenya produces some of the best coffee in the world. Being the more flavourful *Coffea arabica* rather than coffee *Coffea canephora* (Robusta), the "fully washed mild" belongs to the top quality group called "Colombian milds".This is attributed to the well distributed rainfall; high altitude (1,500–2,000 metres above sea level) and therefore moderate temperatures (averaging20°c, with characteristically high equatorial ultraviolet sunlight diffusing through thick clouds; and deep red volcanic soils. In Kenya, the commercial varieties recommended for cultivation include K7, SL 28, Ruiru 11, Batian and SL 34.

Coffee originated in the Kaffa region of Ethiopia where it grows naturally. It became an item of trade with Yemen in the fifteenth century. French missionaries planted some in Bourbon (now Reunion) Island in 1708, and by 1817 about 3,000 tons were being produced annually.The Bourbon seeds were brought to mainland Tanzania (notably Bagamoyo and Morogoro) in 1863 by the Holy Ghost Fathers of the French Catholic Church who eventually proceeded to plant it at Bura near Taita Hills in Kenya in the early 1890s. At this time, the Protestant Scottish missionaries were experimenting with Mocha seedlings at their various stations in Kenya, including Kibwezi (1893) and Kikuyu. In 1897, Brother Zolanus Zipper of the Holy Ghost missionaries brought seed from Morogoro to plant at the Nairobi mission (St. Austin's Muthangari), added 100 seedlings from Bura the following year and got an acre (less than a half hectare) of flowering crop by 1900. This crop represented varieties of Mocha, with a bronze leaf tip, and Bourbon, with a dark green leaf tip. Due to their cultivation over the years under different conditions, the various coffee varieties seem to have hybridized into a special variety of coffee that was christened "French Mission" coffee. By 1904, the Muthangari station had 5,000 mature

trees, 15,000 by 1910 and 52,000 by 1914. The station supplied seeds and seedlings to other early coffee growers in the country (Economic Survey, Central Bureau of Statistics, 2006, Nairobi, Kenya).

2.3 General description of Coffee

Coffea arabica is a shrub or small tree, and if untended, it may reach a size of 4 to 5 meters. Flowers of *C. arabica* have short corolla, long style and exerted stamen that are typical of the genus *Coffea.* Such floral morphology would permit natural cross-pollination, but nevertheless, *C. arabica* is largely autogamous, and fruit set after self-pollination is 60% or higher (Carvalho *et al*., 1969). *C. eugenioides* is short shrub tree. The flower morphology does not permit for self-pollination. *C. canephora* has a bigger tree when compared to Arabica with broader and larger pale green leaves; they are self-sterile and hence cross pollination is the rule. The fruits mature in 10-11 months (Carvalho, 1988).

Coffea canephora is a coffee plant grown on tropical lowlands (below latitude 10°) up to aheight of 1000m above sea level. *C. canephora* plantsbear high temperatures and greater moisture better than the ones of *C. arabica*. However, they are more sensitive towards cold.

Coffea eugenioides is native to the highlands of East Africa, where it occurs in the eastern part of the Democratic Republic of Congo, Rwanda, Uganda, Kenya and western Tanzania.Uses Beans of *Coffea eugenioides* have a very low caffeine content

and coffee made from them tastes mild and agreeable. *Coffea eugenioides* plays a role in breeding work of *Coffea arabica L.* and *Coffea canephora* especially in broadening the genetic base and lowering the caffeine content.

2.4 Coffee propagation

For *C. Canephora* and *C.eugenioides*, self-incompatible coffee species, vegetative propagation methods are preferred and currently used for multiplying selected varieties of this species. In comparison, most commercially grown *C. arabica* cultivars have been propagated by seed as a standard practice since it is generally believed that Arabica coffee is sufficiently true breeding (Clifford and Willson, 1985; Vossen, 1985). Vegetative propagation methods are applicable to coffee, including cuttings, grafting and tissue culture. Propagation by cuttings is applied when few genotypes need to be propagated in large numbers. Grafting is a preferred method when a small number of plants are needed from each genotype. The advantage of grafting over cuttings is the vigour given to the scion by the rootstock. It is therefore possible to use grafting, as a research tool to rescue weak seedlings or even haploid embryos (Clarke and Macrae, 1988), or to produce the grafting with strong root systems that can be drought tolerant or disease resistant (Bittenbender *et al.,* 2001). In vitro methods can also be used for propagation by two ways: micro-cutting or somatic embryogenesis. This multiplication approach is able to produce a great number of plants but has the limitation of requiring refined techniques and chemical media (Clarke and Macrae, 1988).

2.5 Coffee mutants

C. arabica mutants are very numerous and include variation in leaf shape and color, growth habit, as well as flower, fruit and seed characters (Wrigley, 1988). According to (Clifford and Willson, 1985), many current coffee cultivars are generally believed to be the consequence of spontaneous mutations of major genes influencing plant, fruit and seed characters, rather than the expression of residual heterozygosity. More gene mutations have been found in the tetraploid *C. arabica* than in diploid *C. canephora*. This is probably because *C. canephora* is cross-pollinated, whereas *C. arabica* is largely self-fertilized, so there is more chance for recessive mutants to persist in the homozygous form in *C. arabica*, whereby they are expressed in the phenotype. In addition, a considerably greater number of plants of *C. arabica* are grown in the world, and they have been more extensively studied (Wrigley, 1988). Lashermes *et al.*, (1999) stated that, due to its allotetraploid origin, *C. arabica* has a high level of fixed heterozygosity and therefore the level of internal genetic variability is "twice" that present in its diploid relatives. This variability may account for the success of *C. arabica* as a selected agricultural plant species. Krug and Carvalho (1951) studied more than 40 mutants found in *arabica* coffee and contributed to a much better understanding of its genetics. They clearly demonstrated the diploid mode of inheritance of all characters in this allotetraploid.

2.6 Coffee improvement

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 collectionsand for efficient conservation and utilisation of available genetic resources (Sakiyama, 2000). It assists in maintaining genetic diversity and sustained long-termselection gain in a breeding programme (Chowdhury *et al.,* 2002). Furthermore,evaluation and grouping of landraces of a crop of a certain region is helpful instudying the evolutionary relationships in line with the history of the crop in thatregion (Zeven, 1990).

Coffee, which is a perennial tree crop, is difficult to improve through traditional plant breeding. Conventional coffee breeding methodology faces considerable difficulties due to limitations such as the long generation time of coffee trees, the high cost of field trials and a lack of accuracy in the current breeding strategy (Etienne *et al*., 2002). Molecular analysis of *C. arabica* cultivars could provide knowledge of the levels of genetic variation and the genetic relatedness between genotypes which can improve the efficiency of utilization of current germplasm resources. Furthermore, genetic data are important for designing effective plant breeding programs by influencing the choice of parental genotypes to cross for the development of new populations (Russell *et al*., 1997).

In addition, the development of molecular markers could serve in marker-assisted selection, provided that the markers are linked to any of the important agronomic traits. For instance, RAPD markers have been found to be closely linked to some genes conditioning CBD (coffee berry disease) resistance (Agwanda *et al*., 1997;

Cristancho, 1999). Development of molecular markers for resistance genes to CBD and other major coffee diseases would enable breeders to ensure effective host resistance by gene pyramiding (Cristancho, 1999).

2.7 Implication for breeding

Besides providing information to design conservation strategies, the genetic diversity information could also be used in breeding programs in order to cross genetically diverse parents and maximize the level of variation present in segregating populations. Some coffee breeding centers now emphasize hybrid varieties as the best strategy for further and more rapid increase of plant productivity. In *C. arabica*, 30 to 60% heterosis in yield over the better parent has been observed in Ethiopia (Ameha, 1990). Coffee hybrids were also found to have greater yield stability over location and. combining parents selected from genetically diverse subpopulations increase chances of substantial hybrid vigor. Hybrid vigor for yield noticed in crosses between parents of different origin appears to be the result of accumulation of complementary polygenes dispersed over subpopulations (Van der Vossen, 2001).

2.8 Conservation strategies

The current conservation efforts both at ex-situ field gene banks, and in-situ on farm (landraces) or in its natural forest ecosystem (wild forest coffee population) is very low as compared to the economic importance of coffee, the great threat to its genetic in its populations in Ethiopia which one cannot find anywhere else on the world. According to (Bellachew, 1997), the accessions available in the gene bank are too few to represent the high genetic variability available within the natural coffee populations

in Ethiopia. Unless immediate protective measures are taken at large scale for long-term benefits, the pressure could lead to the total irreversible loss of a significant part of the available genetic resources in less than a couple of decades. This could have a high consequential cost both at national and international level to the coffee production and marketing chain. For the future benefit of coffee economy, it is important to plan and decide a strategy to conserve these populations at its very beginning place. However, conserving the whole populations is practically impossible due to resources limitation. Thus, there is a need to identify and conserve potential populations with the maximum possible genetic diversity, which depends on the availability of genetic diversity information**.** Hence, any effort towards generating information on the genetic pattern of coffee, especially using DNA molecular techniques, is very important.

2.9 Economic significance of Coffee in Kenya

The agricultural sector is the main driver of Kenyas' economy directly contributing 26% of the annual Gross Domestic Product (GDP) (Economic survey, 2010).The sector accounts for 65% of Kenya's total exports.In spite of the many challenges faced by the sector as outlined in the foregoing sections, coffee has remained a major employer in Kenya. Between 2001 and 2005, the estate sub sector accounted for an average of 61,000 employees in any one year equivalent to 19 per cent of total employment in agriculture and forestry activities and about 4 per cent of total employment in Kenya. In terms of gender, 75 per cent of total employment in this commodity chain are males and 25 per cent females. Casuals and parttimers constitute 21 per cent of the total, of whom 20 per cent are men and 24 per cent women.

To the 61,000 employed in the coffee estate sub-sector should be added the persons working in coffee activities either for pay, profit or family gain in the small coffeefarm/cooperative sub-sector either as regular workers or on a seasonal or casual basis. Such activities cover weeding, spraying, harvesting/picking, sorting and transporting coffee to the pulpery. Other workers are employed in coffee factories, milling, marketing and allied activities. As for the smallest units among the smallholders, some 700,000 are self-employed coffee growers. When all are accounted for, close to a million people depend on the coffee sector for their living or employed at some stage in the commodity chain. (Statistical Abstract 2006, and Economic Survey 2006, Kenya). The traditional export destinations for the Kenya coffee have been Germany (30%), Benelux (Belgium, Netherlands, and Luxembourg) (12%) USA and Canada (11%), Sweden (7%), Finland (6%) and UK (6%), (Statistical abstracts, 2008). However, in 2009, new emerging markets were identified which included China, Japan and Russia.

2.10 Production and marketing of coffee

Brazil dominates the world production and has a major influence on the world Coffee industry. Production amounted to approximately 21% of the worlds output in 1992 with Columbia, second largest producer recording an output of 16%, most of which is Arabica. Indonesia is the largest single producer of Robusta coffee in the world followed by Ivory Coast (Omondi, 1998). Highland areas in Kenya, Tanzania, Uganda, Angola, Burundi, Cameroon, Ethiopia,Malawi, Rwanda, Zambia and Zimbabwe also produce Arabica coffee. Robusta coffee isproduced commercially at lower altitudes in Western and Central Africa (Wrigley, 1988).Ivory Coast is the largest African producer with 4.1% of world production.A high proportion of

world‟s coffee is imported by developed countries.The global production of coffee producing countries is highlited in table 2.1 and their output from 2010-2013.

Table 2.1: Coffee production. Report by United States Department of Agriculture Foreign Agricultural Service Circular Series December 2013

List of coffee producing countries	The amount of coffeeProduced by respective countries in a 1000- 60 kg bags				
	2010/11	2011/12	2012/13	June 2013	Dec 2013
Angola	25	25	30	30	30
Bolivia	125	140	125	145	145
Brazil	54500	49200	56100	53700	53100
Burundi	235	210	235	200	225
Cameroon	715	735	615	700	700
Central African republic	27	10	10	10	10
Colombia	8525	7655	9925	9000	10000
Congo(Kinshasa)	265	255	230	220	220
Costa Rica	1575	1775	1675	1425	1425
Cote d'Ivoire	1600	1600	1750	2000	1900
Cuba	120	125	125	100	100
Dominican Republic	500	500	475	450	450
Ecuador	650	630	590	575	575
El Salvador	1860	1200	1250	800	1000
Ethiopia	6125	6320	6325	6350	6350
Ghana	57	70	25	50	50
Guatemala	3960	4410	4210	3885	3885
Guinea	325	425	175	400	300
Haiti	300	300	300	300	300
Honduras	3975	5600	4600	5000	5000
India	5035	5230	5303	5200	5125
Indonesia	9325	8300	10500	9200	9500
Jamaica	23	20	15	20	20
Kenya	680	850	900	900	900
Lagos	500	450	460	400	450
Liberia	4	5	5	5	5
Madagascar	550	550	525	500	500
Malawi	25	25	25	25	25
Malaysia	1100	1450	1400	1500	1500
Mexico	4000	4300	4500	3800	3800
Nicaragua	1740	2100	1925	1500	1700
Nigeria	30	40	30	35	35
Panama	87	80	80	80	80
Papua New Guinea	865	1400	825	1050	840
Paraguay	25	25	25	25	25
Peru	4100	5200	4300	4100	3850
Philippines	460	455	455	450	450
Rwanda	317	245	260	275	275

Sierra Leone	45	90	70	80	80
Sri Lanka	35	30	35	30	30
Tanzania	1050	565	1100	1200	1200
Togo	615	560	200	400	400
Uganda	3212	3075	3350	3500	3500
United States	108	92	97	100	100
Venezuela	625	700	600	800	800
Vietnam	19415	26000	26500	24800	28500
Yemen	145	150	150	150	150
Zambia	7	10	5	5	5
Zimbabwe	10	8	8	5	5
TOTAL	140447	144040	153268	146325	150465

Consumption in the producing countries and other developing states is small with the largest individual importer being U.S.A, which imports 23% of the world‟s total production. European Economic Community as a whole imports 39% of the world's total production.The quantity of coffee exported by various countries is shown in table 2.2.

Table 2.2: Total Coffee exports. Report by United States Department of Agriculture Foreign Agricultural Service Circular Series December 2013

List of coffee exporting countries	Amount of coffee exported by respective countries in a 1000 -60 kg bags			
	2013\14	2011/12	2012/13	June 2013
Angola	5	8	8	10
Bolivia	66	87	60	75
Brazil	35010	29843	30660	31040
Burundi	235	205	240	195
Cameroon	650	670	540	625
Central Africa Republic	22	7	5	5
Colombia	8385	7360	8855	8625
Costa Rica	1255	1455	1400	1200
Code d'Ivoire	985	1620	1680	1800
Cuba	5	11	10	10
Dominican Republic	48	47	40	40
Ecuador	1250	1550	1650	1750
El Salvador	1772	1130	1150	750
Ethiopia	3235	3140	3280	3300
European Union	330	240	285	400

Ghana	57	70	25	50
Guatemala	3725	3915	3800	3500
Guinea	300	380	160	370
Haiti	7	9	10	10
Honduras	3900	5290	4400	4800
India	5515	5223	5255	5230
Indonesia	9720	7450	8900	7400
Jamaica	18	17	10	15
Kenya	650	800	825	850
Laos	400	350	360	300
Liberia	4	5	5	5
Madagascar	105	105	100	100
Malawi	24	24	24	24
Malaysia	1675	1950	2200	2350
Mexico	2460	3365	3550	3190
Nicaragua	1665	1780	2070	1450
Nigeria	2	10	5	5
Panama	37	30	30	35
Papua New Guinea	850	1350	775	1000
Paraguay	5	5	5	5
Peru	3880	5140	4100	3950
Rwanda	317	245	620	275
Sierra Leone	25	69	50	60
Tanzania	1005	525	730	960
Thailand	1207	750	820	775
Togo	615	560	200	400
Uganda	3150	3000	3200	3400
Venezuela	100	100	80	80
Vietnam	18640	24495	24200	23600
Yemen	20	20	20	20
Zambia	7	10	5	5
Zimbabwe	6	3	3	3
TOTAL	113419	114487	116095	114092

2.11 Evaluation of genetic variation of Coffee using markers

According to Jump *et al.* (2008), there is heavy reliance on plant genetic diversity for future crop security in agriculture and industry. However, they observed that genetic diversity for natural populations receives less attention. Like it is for many crops, evaluation of the genetic diversity and available resources within the genus *Coffea* is an important step in coffee breeding (Cubry *et al.*, 2008). As new coffee varieties are

continuously being developed through hybridization, there is a need to determine the level and sources of genetic variation within and between new and existing coffee varieties (Gichimu and Omondi, 2010a). Genetic consistency within varieties is also essential to quality assurance for any agricultural product. Hue (2005) reported that morphological variability in coffee plantations is adverse to the product quality. Reduced genetic diversity is also reported to compromise the ability of populations to evolve to cope with environmental changes and thus reducing their chances of long-term persistence (Frankham *et al.,* 2002).

Genetic variation of coffee can be assessed using different techniques that range from the traditional morphological techniques to the modern DNA-based molecular markers. The use of morphological techniques in coffee diversity studies is limited by the influence of environmental factors and the long period it takes for some characters to be expressed during the various growth stages of the plant (Weising *et al*., 2005). In addition, they are also few in numbers and require lengthy follow-up during the whole growth stage especially in perennial plants like coffee. In response to the limitation of morphological techniques, the more effective technique based on protein (isozymes), was developed. However, it was found to be inappropriate for determining the genetic variation in *C. arabica*due to fact that analysis of isozymes specifically in *C. arabica* accessions failed to reveal the amount of polymorphism detected using morphological markers (Berthaud and Charrier, 1988; Lopes, 1993; Bustamanate and Polanco, 1999). DNA-based marker techniques are more efficient, precise and reliable for discriminating between closely related species and cultivars and in determining gene introgression within and between species (Etienne *et al.*,2002).

2.11.1 Morphological markers

Morphological markers are a classical method to distinguish variation based on the observation of the external morphological differences such as the size and shape of the leaf and of the plant form, the color of the shoot tip, the characteristics of the fruit, the angle of branching and the length of the internodes. However, assessing polymorphism with morphological markers can be inefficient since they are generally dominant traits. Moreover, they often exhibit epistatic interactions with other genetic traits and can also be influenced by the environment (Vienne, 2003). The sharing of physical features is also often accepted as an indication of relatedness. Besides, assessment of morphological characters in perennial plants such as coffee, often require a lengthy and expensive evaluation during the whole vegetative growth (Weising *et al*., 2005).Morphological variation among*C.arabica* in kenya is low (Gichimu and Omondi, 2010a).

2.11.2 Biochemical markers

Enzymes are the basic tools of cellular chemistry and were introduced as markers in the early 1970s (Glaubitz and Moran, 2000). 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. Isozymes generally exhibit Mendelian inheritance, co-dominant expression.Complete 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 (comigration) may not be homologous (Morell *et al.*, 1995). Isozyme studies in plants have demonstrated that pattern and band intensities differ by tissue types and developmental stages (Montarroyos *et al.*, 2003).

Although isozymes are 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). 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; Lopes, 1993; Bustamanate and Polanco, 1999). The attempt of Paillard *et al.*, (1996) to construct a genetic map for coffee based on isozymes was not successful due to the low polymorphism level.

2.11.3 Molecular markers

Molecular markers have been replacing or complementing traditional morphological and agronomic characterization, since they are virtually unlimited, cover the whole genome, are not influenced by the environment and less time consuming. 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. Each molecular marker has its advantages and drawbacks. Therefore, application of molecular marker techniques to genetic variation must take into account whether or not the data derived from the technique will provide the right type of information for answering the question being addressed (Karp *et al*., 1997). Plant DNA polymorphism assays are powerful tools for characterizing and investigating germplasm resources and genetic relatedness (Powell *et al*., 1996). These include non-PCR-based DNA markers such as restriction fragment length polymorphisms (RFLP) and PCR based DNA markers.These techniques include RAPD (random amplified polymorphic DNA) (Welsh and McClelland, 1990; Williams *et al*., 1990), AFLP (amplified fragment length polymorphism) (Vos *et al*., 1995), SSR (simple sequence repeat) or microsatellite (Morgante and Olivieri, 1993).

2.11.3.1 PCR-based markers

The Polymerase Chain Reaction (PCR) has become essential in studies of molecular ecology and population genetics research in the brief time since its invention. With this technique, defined DNA segments can be amplified to microgram quantities from as little as a single template molecule (Hoelzel and Green, 1998). A wide variety of

PCR-based marker techniques have been developed during the last decade, each with various advantages and different shortcomings (McGregor *et al.,* 2000). These PCR-based markers differ in principle, application, complexity, informativeness, cost and time requirements (Mignouna *et al.,* 2003). The choice of marker system often depends on the crop investigated (Powell *et al.,* 1996a; Milbourne *et al.,* 1997; Russell *et al.,* 1997). With the development of the polymerase chain reaction (PCR), many PCR-based DNA molecular techniques have been, and still are being developed for plant genome analysis. These techniques include RAPD (random amplified polymorphic DNA) (Welsh & McClelland, 1990; Williams *et al.,* 1990), AFLP (amplified fragment length polymorphism) (Vos *et al.,* 1995), ISSR (inter-simple sequence repeat) (Zietkiewicz *et al.,* 1994), SSR (simple sequence repeat) or microsatellite (Morgante and Olivieri, 1993).

2.11.3.1.2 Simple Sequence Repeats (SSR)

Simple sequence repeats (SSRs) are tandem repeated motifs of 1-6 bases (Beckman and Weber, 1992; Hancock, 1999; Gupta *et al*., 1999) found in all prokaryotic and eukaryotic genomes. They are present in both coding and non-coding regions and are usually characterized by a high degree of length polymorphism. Microsatellites are surprisingly common in the vicinity of genes, and tri-nucleotide repeats preferably occur in exons). Slippage of DNA polymerase during DNA replication and failure to repair mismatches is considered as a common mechanism for creation of hyper variability of microsatellites (Levinson and Gutman, 1987).

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) and easy to assay using polymerase chain reaction (PCR) (Kuleung *et al.,* 2004). Their relative abundance, multiallelic nature, codominant inheritance, high reproducibility and good genome coverage make them powerful molecular markers for use in genetic studies (Weber and May, 1989).The major drawback of microsatellites is the necessity of sequence information for primer design that they need to be isolated de novo from most species being examined for the first time (Beckman and Weber, 1992) .

2.11.3.1.3 SSR markers in coffee

Microsatellites were applied to identify *C. arabica, C. canephora* and related coffee species (Combes *et al.,* 2000); to assess polymorphism among 16 *C. arabica* and four *C. canephora* accessions and to identify DNA introgression fragments from *C. canephora* in four *C. arabica* lines (Anthony *et al.,* 2000). SSRs and AFLPs were appropriate markers for studying introgression in coffee. The microsatellites also showed low genetic variation in *C. arabica* compared to that detected by RAPDs (Combes *et al.,* 2000). Similarly, nine SSR markers were developed by Baruah *et al.,* (2003) to identify polymorphism in *C. arabica, C. canephora* and 17 species of *Coffea* and the related genera Pilanthus also revealed very low polymorphism across the 45 Arabica genotypes. The SSR primers of Rovelli *et al.* (2000), isolated from two genomic libraries of *C. arabica,* showed polymorphisms among the *C. arabica* accessions studied and were able to discriminate between the two chromosome sets derived from the diploid donor ancestral plants.

Moncada and Couch, (2004) used thirty-four fluorescently labeled microsatellite markers to assess genetic diversity among five diploid species and 23 various cultivated and wild accessions of tetraploid *C. arabica* from Colombia. The results showed higher genetic polymorphism in comparison with previous reports. The combined use of SSR and RFLP or AFLP markers were useful in study the gene introgression into *C. arabica* by way of triploid hybrids (*C. arabica* x *C. canephora*) (Herrera *et al.,* 2002b) and by way of tetraploid interspecific hybrids (*C. arabica* x *C. canephora*) (Herrera *et al.,* 2002a). The results found that SSRs and AFLPs were appropriate markers for studying introgression in coffee and to analyze the introgression of DNA fragments from *C. canephora* and *C. liberica* into *C. arabica* (Lashermes *et al.,* 2000, Prakash et *al.,* 2002), Gichuru et *al.,* 2008 Lashermes, *et al.,* 2010).

2.11.3.1.4 Random amplified polymorphic DNA (RAPD)

RAPD was first used by Williams *et al.* (1990) to examine human DNA samples. This method is based on the fact that using short arbitrary primer sequences; they can by chance anneal on random sequences within the genome in close proximity and in opposite orientation to be amplified in a PCR programme. RAPD marker technique is quick, easy and requires no prior sequence information (Welsh and McClelland, 1990). This technique employs single primers with 10 arbitrary nucleotide sequences and at least 50% GC content. Randomly Amplified Polymorphic DNA (RAPD) is a PCR-based marker system, jointly described by (Williams *et al*., 1990) and (Welsh and McClelland, 1990).

Amplification of genomic DNA using single primers of arbitrary nucleotide sequence in low stringency conditions, results in multiple amplification products from loci distributed throughout the genome (Welsh and McClelland, 1990; Williams *et al*., 1990). RAPD markers became popular because of their simplicity, applicability to any genome, no sequence information requirement and relatively small DNA quantities required . Besides, RAPD is sensitive to slight changes in reaction conditions, which interfere with the reproducibility of banding patterns between separate experiments, PCR instrumentation, and laboratories (Penner *et al*., 1993). Because reproducibility mainly depends on appropriate optimization of PCR components, it is advisable to determine optimal RAPD conditions empirically by performing a set of pilot experiments. Given that the outcome of RAPD experiments is influenced by many interacting variables, complete optimization can only be achieved if each component is tested independently and across a wide concentration range

2.11. 3.1.5 RAPD markers in coffee

RAPD makers have been used to confirm the relationships within the genus Coffea (Orozco-Castillo *et al.,* 1996), to construct a linkage map in coffee (Lashermes *et al.,* 1996a), to detect markers for resistance to coffee berry disease (Agwanda *et al.,* 1997) and coffee leaf rust (Rani *et al.,* 2000), and to study genetic diversity amongst wild accessions (Orozco-Castillo *et al.,* 1994; Lashermes *et al.,* 1996b; Anthony *et al.,* 2001) and cultivated varieties (Lashermes *et al.,* 1996b; Masumbuko *et al.*,2003; Sera *et al.,* 2003; Crochemore *et al.,* 2004).

Lashermes *et al.* (1996b) employed RAPD markers to study genetic diversity between cultivated and wild accessions of *C. arabica* and found the RAPD method appeared to be effective in resolving genetic variation and in grouping germplasm in *C. arabica*. The study also confirmed the narrow genetic base of commercial cultivars of *C. arabica* and only 12 of 140 RAPD primers detected polymorphisms and were not able to distinguish the cultivars within the same type groups, either Bourbon or Typical.

RAPD markers were used to investigate the genetic diversity in Tanzanian cultivated *C.arabica* and found variability in these accessions and that these accessions clustered according to geographical locations (Masumbuko *et al.,* 2003). However, only ten out of 100 decamer primers exhibited polymorphism and gave reproducible banding patterns.

2.11.3.1.6 AFLP (Amplified Fragment Length Polymorphism)

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 *etal.,* 1996). 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 agarose and acrylamide gel electrophoresis or capillary electrophoresis 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.

2.11.3 .1.7 AFLP in Coffee (Amplified Fragment Length Polymorphism)

AFLP markers were used to detect the introgression of *C. canephora* genetic material in *C. arabica* (Lashermes *et al.,* 2000a), the introgression of *C. liberica* genetic material in *C. arabica* (Prakash *et al.,* 2002), to construct a genetic linkage map in coffee (Lashermes *et al.,* 2001; Pearl *et al.,* 2004), to identify the origin of cultivated

C. arabica (Anthony *et al.,* 2002a) and to study genetic polymorphisms in *C. arabica* (Steiger *et al.,* 2002; Anthony *et al.,* 2002b). AFLP also showed a low genetic diversity in *C. arabica* (Anthony *et al.,* 2002a). Genome introgression of *C. arabica* and *C. canephora* into accession of Híbrido de Timor was studied using AFLP molecular marker technique and found that introgression of *C.canephora* genome ranged from 8% (in Catimor 3) to 25% (in Sachimor) (Lashermes *et al.,* 2000). The advantages of the AFLP technology include no need of prior DNA sequence information, and the possibility of applying high stringency during PCR, which ensures high reproducibility of the method. Although it is a very powerful approach, it has a number of limitations such as dominance of markers, clustering of some markers in distinct genomic regions, limited levels of polymorphism in some cultivated species and the requirement of good quality DNA to ensure complete restriction (Weising *et al.,* 2005).

2.11.4 Gene introgression

Gene introgression is the persistent incorporation and integration of genes from onepopulation or species into the gene pool of another one (Lashermes *et al.,*2000a), using AFLP (amplified fragment length polymorphism) markers recently estimated that the approximate amounts of introgressed materials in many introgressed *arabica* lines ranged from 8% to 27% of the *C. canephora genome*. Genome introgression of *C. arabica* and *C.canephora* into accession of Híbrido de Timor was studied using AFLP molecular marker technique. The small portion of the genome of *C. canephora* was introgressed into Híbrido de Timor which gave resistance to coffee leaf rust and. Genome introgression of *C. arabica* and *C. canephora* into accession of Híbrido de Timor studied using AFLP molecular marker technique found out that introgression of

C. canephora genome ranged from 8% (in Catimor) to 25% (in Sachimor) (Lashermes *et al.,* 2000). The transfer of desirable genes from wild relatives to cultivated species through wide crosses is one of the proven breeding strategies for crop improvement (Hajjar and Hodgkin, 2007). Gene exchange is possible due to the meiotic recombination which allows segments from the parental chromosomes to recombine into new genetic entities that are passed onto the next generation Stebbins, (1950). Inherent problems of interspecific hybridization such as hybrid instability, infertility, non-Mendelian segregations, and low levels of inter-genomic crossing-over are important limitations to crop improvement (Stebbins, 1958).

CHAPTER THREE

MATERIALS AND METHODS

3.1 Plant materials

A total of 24 coffee genotypes consisting of *C. arabica*, *C. canephora* and *C. eugenioides* (Table 3.1) were analyzed in this study. The coffee trees of these genotypes are available in the commercial fields, experimental sites and museum plots at Coffee Research Institute (CRI) located at Ruiru-Kenya.The study was conducted at Coffee Research Institute .

Table 3.1: Status and Sources of coffee germplasm used in the study.

Genotype	Status	Source
Rume sudan	Gene bank accession	Sudan
Arabusta	Gene bank accession	Kitale,Kenya
Canephora	Gene bank accession	Uganda
BA	Gene bank accession	India
Ruiru11 cd 93	Advanced selection	Kenya
Blue Mountain	Gene bank accession	Kenya
Typica	Gene bank accession	Kenya
SL28	Commercial variety	Kenya
Mundo novo	Gene bank accession	Brazil
Robarbica	Gene bank accession	Kenya
Catura	Gene bank accession	Kenya
Hybrido de timor	Gene bank accession	Timor
Catimor (line 86)	Breeders material	Colombia
Devermachy	Gene bank accession	Kenya
Ruiri11 Cd 80	Advanced selection	Kenya
Ruiru11Cd50	Advanced selection	Kenya
Sarchmore	Gene bank accession	Costa Rica
Bourbon	Gene bank accession	Kenya

Mokka	Gene bank accession	Brazil
Colombia	Gene bank accession	Guatamala
Batian	Advanced selection	Kenya
C.eugeniodes	Gene bank accession	Kenya
Erecta	Gene bank accession	NAL Kenya
Drought resistant (DR 1)	Gene bank accession	Tanzania\

3.2 Genomic DNA extraction

Young disease-free leaves were picked from mature coffee trees at Coffee Research Institute. The DNA of the genotypes was extracted according to the method described by (Diniz *et al.*, 2005), using mixed alkyl trimethyl ammonium bromide (MATAB). The harvested leaves were wiped with 70% ethanol and 0.5g weighed and placed in a motar. Liquid nitrogen was added and the leaves crushed to fine powder by use of a pestle. Lysis and extraction buffers (Appendix 5) were added to the powder (1ml each) and crushing continued. The mixture was transferred to a 2ml eppendorf tubes and incubated at 62°C in a water bath for 30-45 minutes with regular shaking at interval of ten minutes. After incubation, 1 ml of chloroform/isoamyl-alcohol mixture, (24:1) was added to each eppendorf tubes and vigorously shaken and then centrifuged at 13000 rpm for ten minutes in a desktop micro-centrifuge. The supernatants were carefully pipetted out into new 2 ml eppendorf tubes. Ten µl of RNase (10 mg/ml) was added to the supernatants and incubated at 37°C in a waterbath for 30 minutes to get rid of RNA. A volume of isopropyl alcohol equal to the volume of each 24supernatant was added into each bottle, and mixed gently by inverting the tubes several times to precipitate DNA. The suspended DNA was centrifuged at 14000 rpm for ten minutes and a DNA pellet was obtained and the supernatant was carefully discarded. The DNA

pellets were then washed with 200µl of 70% ethanol and centrifuged at 13000 rpm for five minutes. The ethanolwas drained by decanting or micro-pipetting, and the pellets air dried for one hour. The pellets were dissolved in 200 µl of TE (Tris-EDTA)

3.3 DNA quantification

Quality of DNA was read at optical density of 260-280 nm using aspectophotometer. DNA which withvalues between 1.8-2.0 were selected. For DNA quantification, One percent agarose gel was prepared in 0.5x TAE (Tris acetate ethylenediaminetetraaceticacid) by weighing 0.7grams of agarose in 70mls 0.5 x TAE. The weight of the original volume was recorded. The solution was heated in a microwave at short intervals of 15-30sec with occasional shaking until it was clear.Someevaporation occurred during heating, therefore the solution wasweighted again and water was added to obtain the original volume and wasleft to cool to about 55°C. The gel was poured on the casting tray of the mini-electrophoresis unit and bubbles removed.The combs were fixed and the gel allowed to set.After solidification, the combs were removed and the gel placed in the gel tank and 0.5x TAE buffer added to cover the gel .The DNA samples were prepared by taking: 2µl of DNA sample, 6µl of distilled water, 2µl of 6x loading buffer making a total volume of 10µl. The agarose gel was loaded with 10µl of lambda and sample DNA preparations and run at 50Vfor 45 minutes. The gel was stained in 1mg/ml ethidium bromide (50ul of 10mg/ml ethidium bromide in 500ml distilled water) for 20 minutes. The gel was rinsed in water for 20 minutes and placed into uv trans illuminator for imaging.The lambda preparation table was be used to estimate the quantity of DNA (Appendix 1) .

3.4 Amplification of coffee genomic DNA using RAPDS

The method described by (Lashermes *et al*.,1996) and modified by (Agwanda *et al*., 1997) was used for RAPD analysis.Ten arbitrary decamer oligonucleotides were selected based on laboratory results at Coffee Research Institute Ruiru.The primers were obtained from Coffee Research Institute Molecular Laboratory. A total of 4 ng of each DNA sample was used in PCR reactions for RAPD markers amplification. A reaction mix was prepared to include: 2.5 ul of buffer (10x), 2.5 ul of MgCl2 (25 mM), 3.5 ul of dNTPs (500 uM), 2.0 ul RAPD primer (10 µM), 0.5 ul of Taq polymerase 5u / µl. Reaction was incubated in a thermocycler set for the following amplification conditions using the primers described in Table 3.2. Initial denaturation 95ºC for 5minutes. loop 1 (45 cycles) : denature 94ºC for one minute, anneal at 34ºC for one minute and extension 72ºC for 1 minute 30 seconds, final extension 72ºC for 10 minutes and was held at 4 ºC. The RAPD products were electrophoresed in 1.8% agarose gel and then visualized in a UV trans-illuminator after staining in ethidium bromide solution

Table 3.2: RAPD primers used for PCR analysis of 24 coffee accessions

Primer	Sequence
OPN-18	5-'GGT GAG GTC A-3'
OPL-18	5'-ACC ACC CAC C-3'
OPM-04	5'-GGC GGT TGT C-3'
OPI-07	5'-CAG CGA CAA G-3'
OPJ-19	5'-GGA CAC CAC T-3'
OPY-10	5'-CAA ACG TGG G-3'
OPX-20	5'-CCC AGC TAG A-3'
OPY-15	5'-AGT CGC CCT T-3'
OPI-20	5'-AAA GTG CGG G-3'
OPX-16	5'-CTC TGT TCG G-3'

3.5 Amplification of coffee genomic DNA using SSR

Thirteen microsatellites were selected based on laboratory results at Coffee Research Institute Ruiru. Methodology described by (Combes *et al.*, 2000) was used.The primers were obtained from Coffee Research Institute Molecular Laboratory. A total of 100 ng of each DNA sample was used in PCR reactions for SSR markers amplification. A reaction mix was prepared to include: 2.5 ul of buffer (10 x) , 2.5 ul of MgCl(25 mM), 3.5 ul of dNTPs (500 uM) , 2ul of SSR (10 µM) Reverse and forward primer, 0.5 ul of Taq polymerase 5u / µl. Reaction was incubated in a thermocycler set for the following amplification conditions using the primers described in table 3.3. Initial denaturation 95ºC for 5minutes. loop 1 (35 cycles) : denature 94ºC for one minute, anneal at 60ºC for one minute and extension 72ºC for 10 minute 30 and was held 4 ºC. The amplified products were electrophoresed in 2.3% agarose gel and then visualized in a UV trans-illuminator after staining in ethidium bromide solution

Table 3.3: SSR primers used for PCR analysis of 24 coffee accessions

Locus	Reverse prime	Forward primer
Sat11	CCACACAACTCTCCTCATTC	ACCCGAAAGAAAGAACCAA
Sat32	CTGGGTTTTCTGTGTTCTCG	AACTCTCCATTCCCGCATTC
Sat207	CAATCTCTTTCCGATGCTCT	GAAGCCGTTTCAAGCC
Sat227	ATCCAATGGAGTGTGTTGCT	TGCTTGGTATCCTCACATTCA
Sat235	GCAAATCATGAAAATAGTTGG TG	TCGTTCTGTCATTAAATCGTCAA
Sat240	GGTAAATCACCGAGCATCCA	TGCACCCTTCAAGATACATTCA
Sat255	GGGAAAGGGAGAAAAGCTC	AAAACCACACAACTCTCCTCA
Sat283	GTGTGTGATTGTGTGTGAGAG	GCACACACCCATACTCTCTT
Sat254	AAGTGTGGGAGTGTCTGCAT	ATGTTCTTCGCTTCGCTAAC
Sat229	TTTAATGGGCATAGGGTCC	GGCTCGAGATATCTGTTTAG
M24	TTCCTCCATGCCCATATTG	TTCTAAGTTGTTAAACGAGACGCTTA
Sat172	TCAAAGCAGTAGTAGCGGATG	ACGCAGGTGGTAGAAGAATG
Sat262	GCCGGGAGTCTAGGGTTCTGT G	CTGCGAGGAGGAGTTAAAGATACCAC

3.6 Data analysis

The SSR and RAPD amplified bands were scored for the presence (1) or absence (0) of amplified products to create a binary matrix. The total number of bands, the distribution of bands across all species, polymorphic bands and average number bands per primer was calculated using STATSTICA software version 8. DNA introgression was done using the formula no / ne*100 where no is nunber of observed introgressed fragments while ne is the number of expected introgression.Genetic dissimilarities indices was estimated using Pearson dissimilarity. Cluster analysis was performed using Un-weighted Pair Group Method with Arithmetic Averages (UPGMA) (Sneath and Sokal, 1973) using STATSTICA software version 8.

CHAPTER FOUR

RESULTS

4.1 Genetic diversity of 24 Coffee accessions as revealed by RAPD primers

Ten RAPD primer combinations were used to amplify DNA. Among the ten RAPD primers tested, the primers showed amplification and produced clear bands that could be scored. The total number of fragments observed among the coffee genotypes based on the 10 RAPD primers was 79. The agarose gel result is shown in figure 4.1

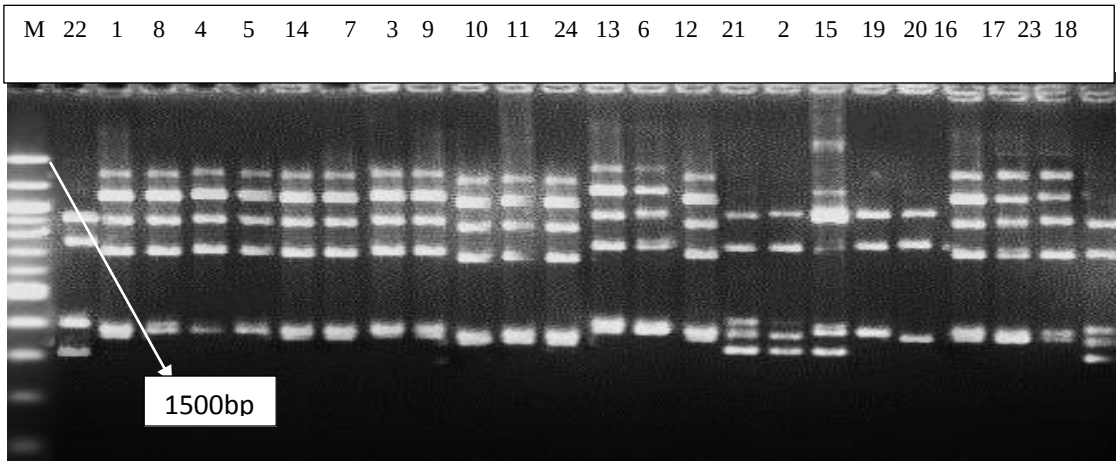

Figure 4.1: RAPD profiles generated by primer L-18.M is a 100 base pair marker while lanes 1–24 are coffee accessions (1-SL28; 2-BA; 3-Batian; 4-Erecta; 5-Ruiru 11 cd 93; 6-Catmor 86; 7-Ruiru 11 cd 50; 8-Dr1; 9-HDT; 10-Rume Sudan; 11-Mundonovo; 12-Devermarchy;13-Catura; 14-Colombia; 15-Eugenioides; 16-Mokka; 17-Typica; 18-Blue Mountain; 19-Ruiru 11 cd 80; 20-Bourbon; 21- Arabusta; 22-Canephora; 23-Robarbica; 24.-Sarchmore)

The number of amplified bands per primer varied from 3.0 to 12.0.The total number of polymorphic fragments produced was 65.The average bands produced by the ten primes was 7.9 and recording average polymorphic bands of 6.5. Percent polymorphism ranged from 50% to 100%, with a mean of 81% polymorphism. Primer I-7 produced the lowest number (3) bands while primer L-18 produced the highest

number (12) bands. In terms of polymorphism, primer L-18) produced highest number of polymorphic bands (12) while 1-7 produced lowest number of polymorphic bands (3). The distribution of bands revealed by ten RAPD primers is shown in (Table 4.1).

Table 4.1: Amplification products generated by RAPD primers

Primer	Number of bands	Polymorphic bands	% Polymorphism
OPI-07	3	3	100
OPX-20	8	6	75
OPJ-19	9	8	88
OPY-15	10	8	80
OPI-20	8	7	88
OPX-16	8	6	75
OPY-10	6	3	50
OPM-04	8	8	100
OPN-18	7	4	57
OPL-18	12	12	100
TOTAL	79	65	
RANGE	3-12	3-12	50-100
AVERAGE	7.9	6.5	81

To estimate the genetic diversity in the evaluated germplasm, amplified data from RAPD marker system was used for calculation of genetic distance matrices generated by the Person's dissimilarity index. The highest values for genetic distances were obtained between Sarchmore and SL28 and between Sarchmore and Batian with ascore of 1 implying that these varieties were genetically quite distinct. The lowest

genetic distance values for the 24 genotypes was between Sarchmore and Robarbicawith ascore of 0.07 implying high genetic resemblance of the two genotypes.The results are shown in (Table 4.2).

Table 4.2: Matrix of genetic distance between 24 genotypes based on RAPD Markers (1-24 represent the coffee genotypes)

	1	2	3	4	5	6	7	8	9	10	11	12
1	0.00											
2	0.47	0.00										
3	0.42	0.63	0.00									
4	0.54	0.60	0.49	0.00								
5	0.47	0.58	0.52	0.54	0.00							
6	0.63	0.52	0.63	0.61	0.58	0.00						
7	0.61	0.71	0.45	0.66	0.39	0.46	0.00					
8	0.45	0.61	0.55	0.50	0.34	0.42	0.43	0.00				
9	0.50	0.66	0.50	0.50	0.45	0.53	0.48	0.33	0.00			
10	0.42	0.63	0.53	0.54	0.36	0.50	0.40	0.36	0.30	0.00		
11	0.42	0.52	0.57	0.64	0.42	0.62	0.55	0.38	0.33	0.30	0.00	
12	0.45	0.66	0.55	0.66	0.39	0.63	0.53	0.38	0.38	0.30	0.33	0.00
13	0.55	0.55	0.66	0.71	0.61	0.46	0.37	0.59	0.65	0.51	0.55	0.59
14	0.50	0.55	0.50	0.63	0.39	0.54	0.38	0.44	0.44	0.42	0.38	0.49
15	0.37	0.47	0.42	0.52	0.37	0.55	0.35	0.41	0.41	0.38	0.36	0.35
16	0.26	0.52	0.42	0.54	0.26	0.47	0.34	0.23	0.29	0.15	0.26	0.24
17	0.55	0.55	0.55	0.62	0.39	0.47	0.38	0.49	0.49	0.41	0.49	0.54
18	0.71	0.50	0.55	0.64	0.50	0.39	0.26	0.48	0.53	0.50	0.60	0.48
19	0.61	0.55	0.61	0.60	0.50	0.52	0.43	0.48	0.48	0.40	0.49	0.37
20	0.95	0.58	1.01	0.84	0.83	0.57	0.78	0.75	0.81	0.80	0.86	0.84
21	0.53	0.63	0.42	0.58	0.31	0.38	0.24	0.35	0.35	0.27	0.41	0.40
22	0.45	0.60	0.39	0.64	0.34	0.39	0.26	0.32	0.37	0.34	0.39	0.37
23	0.77	0.82	0.77	0.88	0.77	0.75	0.63	0.79	0.57	0.70	0.78	0.79
24	1.12	0.85	1.07	0.97	1.01	0.91	0.96	1.03	0.81	1.01	1.00	1.01
	13	14	15	16	17	18	19	20	21	22	23	24
13	0.00											
14	0.29	0.00										
15	0.61	0.67	0.00									
16	0.24	0.24	0.56	0.00								
17	0.36	0.32	0.65	0.34	0.00							
18	0.36	0.29	0.72	0.41	0.31	0.00						
19	0.26	0.39	0.67	0.38	0.42	0.43	0.00					
20	0.28	0.34	0.66	0.36	0.31	0.31	0.21	0.00				
21	0.62	0.66	0.85	0.63	0.66	0.49	0.56	0.54	0.00			
22	0.82	0.86	1.01	0.90	0.90	0.73	0.76	0.77	0.45	0.00		
23	0.28	0.18	0.56	0.23	0.27	0.31	0.31	0.30	0.68	0.88	0.00	
24	0.24	0.15	0.56	0.20	0.24	0.31	0.31	0.26	0.67	0.87	0.07	0.00

The pattern of genetic relationships among genotypes was assessed using UPGMA method of cluster analysis. Dendrogram representing most probable genetic relationship between cultivars is presented in (Fig.4.2). The genotypes separated into three main clusters namely: *C. arabica*, *C. canephora* and *C. eugenioides*. The results indicated that although the arabica varities did originate from accessions from Ethiopia, Sudan, Kenya, Tanzania, Angola, India, Reunion, Portugal, South and Central America they all clustered together thus exhibiting low genetic variation within arabica coffee

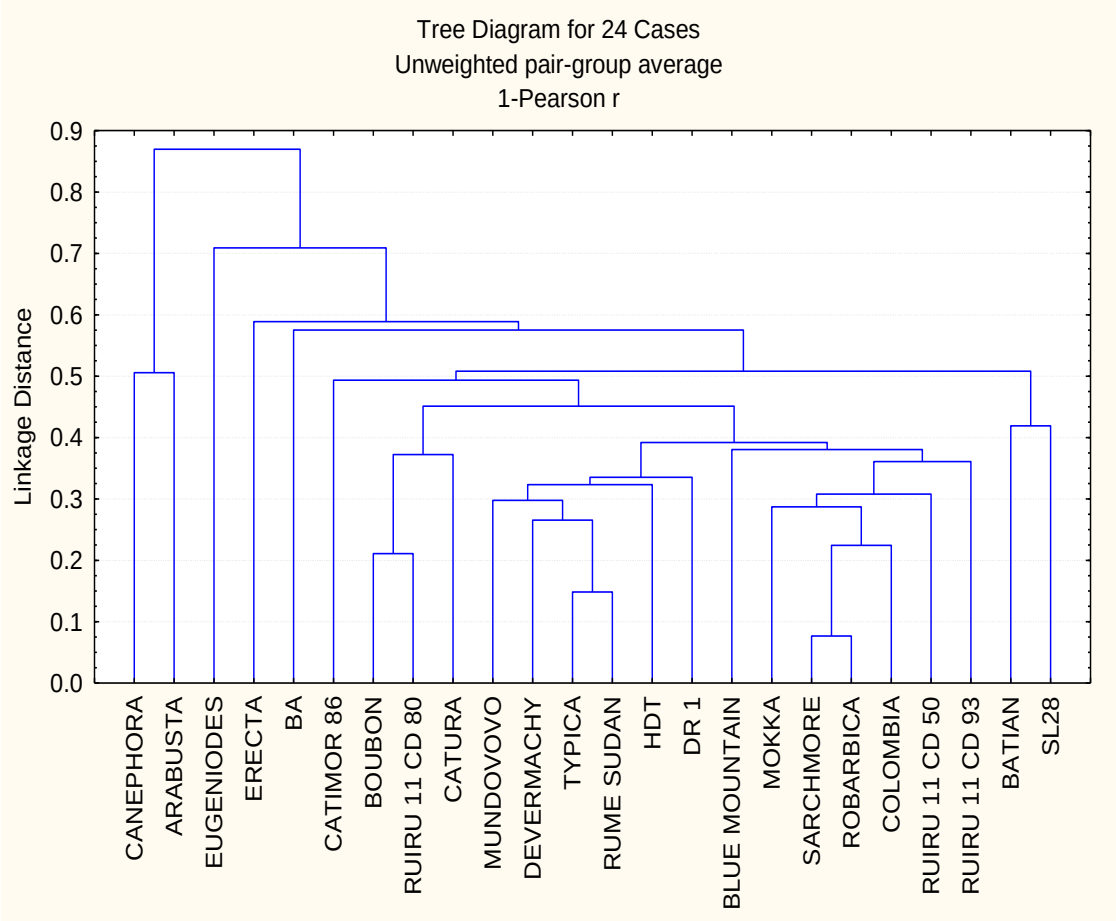

Figure 4.2: Dendrogram of 24 coffee genotypes based on genetic distance obtained from RAPD markers using the UPGMA method.

4.2 Genetic diversity of 24 Coffee accessions as revealed by SSR primers

Thirteen SSR primer combinations were used to amplify DNA. Among the thirteen SSR primers tested, all the primers showed amplification and produced clear bands that could be scored. The total number of fragments observed among the coffee genotypes based on the 13 SSR primers was 50. Due to polyploidy nature of *Coffee arabica*, the determination of heterozygosity and homozygosity is challenging. In addition, available software only supports diploid species analysis. Therefore, microsatellite data was formatted as dominant data in which each allele was treated as a locus and scored as present (1) and absent (0) (Medini *et al.*, 2005; Montemurro *et al.*, 2005) . The agarose gel result is shown in figure 4.3 .

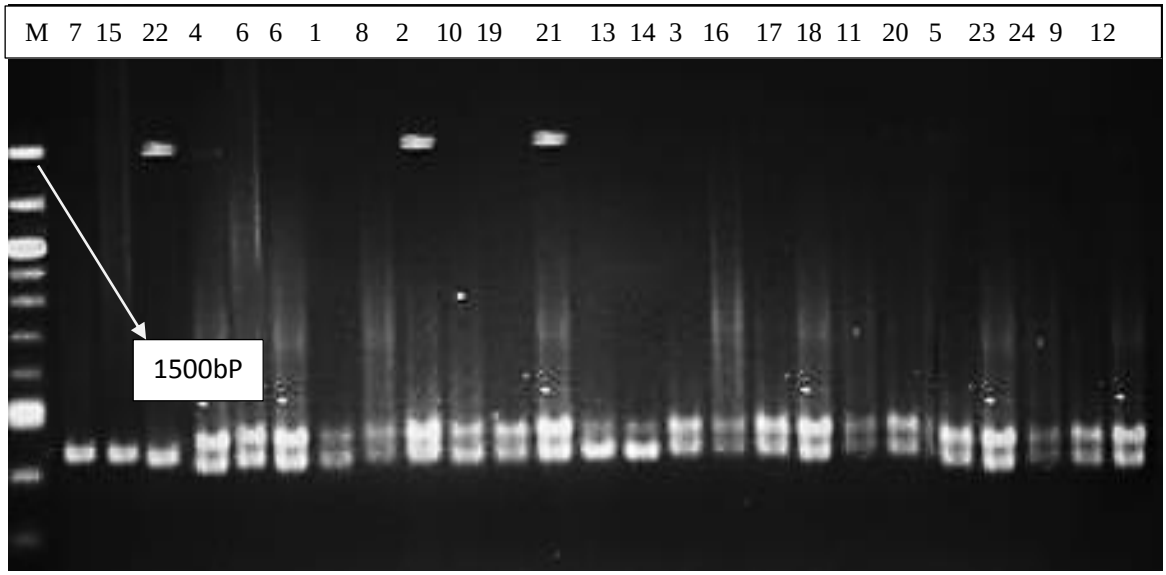

Figure 4.3: Relationships between 24 genotypes generated by UPGMA) Based on SSR Markers (Sat 240). M is a 100 base pair marker while lanes 1–24 are coffee accessions.accessions (1-SL28; 2-BA; 3-Batian; 4-Erecta; 5-Ruiru 11 cd 93; 6-Catmor 86; 7-Ruiru 11 cd 50; 8-Dr1; 9-HDT; 10-Rume Sudan; 11-Mundonovo; 12-Devermarchy; 13-Catura; 14-Colombia; 15-Eugenioides; 16- Mokka; 17-Typica; 18-Blue Mountain; 19-Ruiru 11 cd 80; 20-Bourbon; 21-Arabusta; 22-Canephora; 23-Robarbica; 24-Sarchmore)

A total of 50 alleles were amplified among 24 coffee genotypes using thirteen SSR primers. Of these amplified alleles, 33 were polymorphic. The number of amplified alleles per primer varied from 2.0 to 6.0 with an average of 3.8 alleles.The distribution of bands across the three coffee species is shown (Table 4.3)

Table 4.3: Amplification products generated by SSR primers

Primer	Number of alleles	Polymorphic alleles	% polymorphism
Sat254	6	6	100
Sat 235	3	3	100
Sat11	2	1	50
Sat32	2	1	50
Sat207	3	3	100
Sat227	6	5	83
Sat240	3	2	66
Sat255	4	2	50
Sat283	6	2	33
M24	4	3	75
Sat229	4	2	50
Sat 172	3	1	33
Sat262	4	2	50
TOTAL	50	33	
RANGE	2-6	1-6	33-100
AVERAGE	3.8	2.5	65

To estimate the genetic diversity in the evaluated germplasm, amplified data from SSR marker system was used for calculation of genetic distance matrices generated by the Pearson's dissimilarity index. The highest values for genetic distances were obtained between *C. canephora* and *C. eugenioides* and betweeen *Canephora* and Batian with a score of 1 implying that these varieties were genetically quite distinct. The lowest values for genetic distance were recorded between Sarchmore and Robarbica and between and Sarchmore and Colombia with a score of 0.06 implying that these varieties were genetically fairly similar. The genetic distance matrices between the twenty four coffee accessions are represented in the table 4.4.

Table 4.4: Matrix of genetic distance between 24 coffee genotypes based on SSR marker

	1	2	3	4	5	6	7	8	9	10	11	12
1	0.00											
2	0.37	0.00										
3	0.36	0.50	0.00									
4	0.37	0.45	0.41	0.00								
5	0.45	0.50	0.52	0.41	0.00							
6	0.52	0.37	0.52	0.41	0.46	0.00						
7	0.48	0.57	0.39	0.41	0.42	0.40	0.00					
8	0.40	0.52	0.54	0.26	0.41	0.38	0.41	0.00				
9	0.37	0.49	0.44	0.30	0.44	0.35	0.45	0.26	0.00			
10	0.31	0.45	0.41	0.33	0.37	0.38	0.31	0.36	0.26	0.00		
11	0.36	0.37	0.49	0.34	0.39	0.42	0.45	0.37	0.31	0.24	0.00	
12	0.34	0.45	0.47	0.39	0.41	0.44	0.44	0.36	0.29	0.23	0.24	0.00
13	0.42	0.37	0.49	0.38	0.46	0.29	0.29	0.44	0.44	0.34	0.36	0.34
14	0.42	0.47	0.46	0.34	0.39	0.39	0.29	0.38	0.34	0.28	0.29	0.31
15	0.65	0.76	0.78	0.66	0.71	0.63	0.66	0.75	0.56	0.66	0.72	0.70
16	0.31	0.39	0.41	0.33	0.37	0.38	0.34	0.39	0.36	0.29	0.27	0.29
17	0.32	0.47	0.45	0.34	0.32	0.35	0.35	0.24	0.24	0.21	0.29	0.24
18	0.48	0.47	0.52	0.41	0.45	0.40	0.40	0.48	0.41	0.34	0.41	0.40
19	0.55	0.44	0.46	0.44	0.49	0.36	0.23	0.51	0.48	0.38	0.48	0.41
20	0.44	0.42	0.47	0.39	0.47	0.38	0.35	0.43	0.36	0.29	0.37	0.26
21	0.70	0.51	0.86	0.65	0.75	0.50	0.69	0.72	0.65	0.67	0.66	0.69
22	0.97	0.78	0.98	0.85	0.98	0.81	0.92	0.99	0.81	0.87	0.86	0.89
23	0.37	0.45	0.37	0.26	0.34	0.28	0.25	0.36	0.30	0.23	0.31	0.29
24	0.34	0.42	0.34	0.33	0.34	0.28	0.24	0.33	0.29	0.23	0.27	0.23
	13	14	15	16	17	18	19	20	21	22	23	24
13	0.00											
14	0.29	0.00										
15	0.61	0.67	0.00									
16	0.24	0.24	0.56	0.00								
17	0.36	0.32	0.65	0.34	0.00							
18	0.36	0.29	0.72	0.41	0.31	0.00						
19	0.26	0.39	0.67	0.38	0.42	0.43	0.00					
20	0.28	0.34	0.66	0.36	0.31	0.31	0.21	0.00				
21	0.62	0.66	0.85	0.63	0.66	0.49	0.56	0.54	0.00			
22	0.82	0.86	1.01	0.90	0.90	0.73	0.76	0.77	0.45	0.00		
23	0.28	0.18	0.56	0.23	0.27	0.31	0.31	0.30	0.68	0.88	0.00	
24	0.24	0.15	0.56	0.20	0.24	0.31	0.31	0.26	0.67	0.87	0.06	0.00

The UPGMA algorithm was used for grouping all cultivars based on their genetic distances / increasing dissimilarities. Dendrograms representing most probable genetic relationship between cultivars are presented in figure 4.4.The cluster dendogram constructed was used to estimate genetic diversity among 24 coffee accessions.

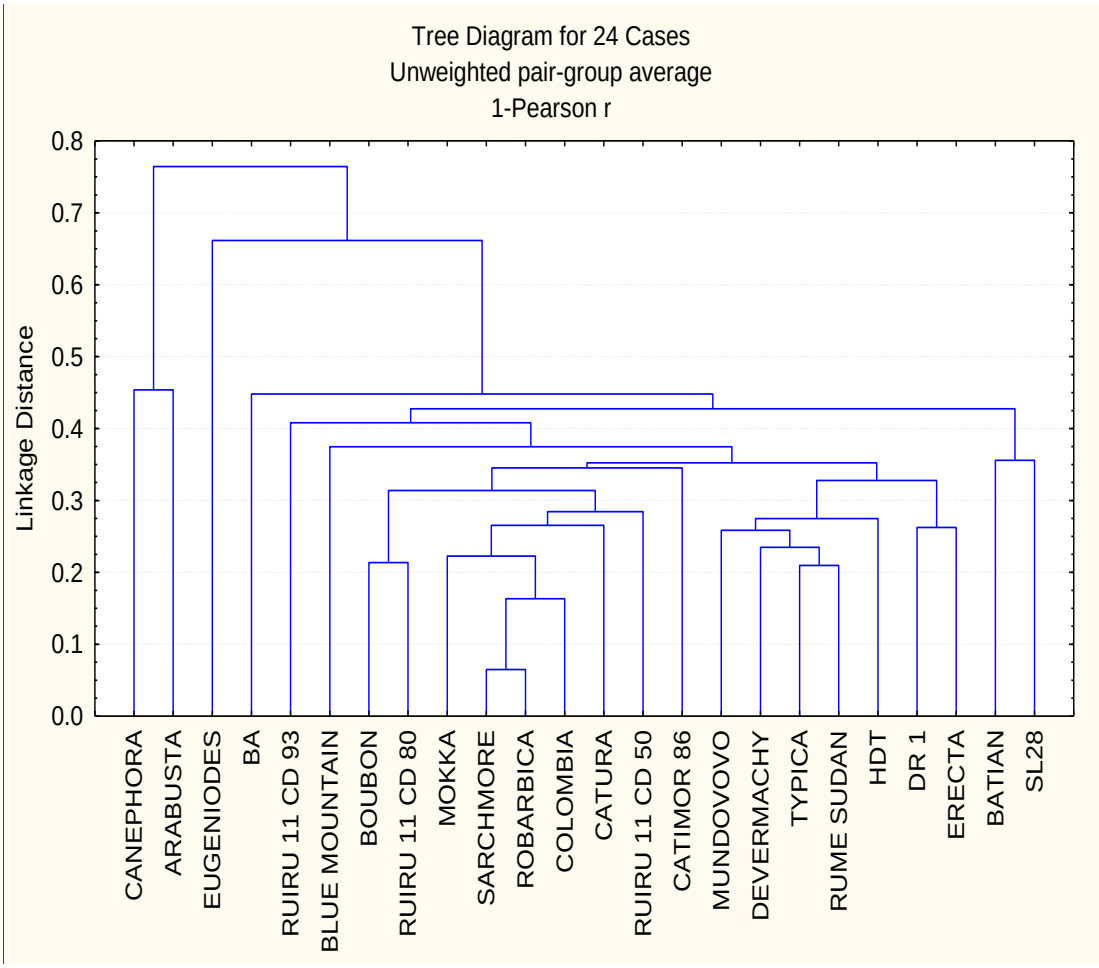

Figure 4.4: Dendrogram of 24 coffee genotypes constructed by cluster analysis using SSR markers

4.3 Gene introgression

Out of 13 SSR primer assayed , only three that detected *Canephora* introgressed fragments present in arabica coffee.The introgressed fragments are presented in the figures (4.5), (4.6) and table 4.5.

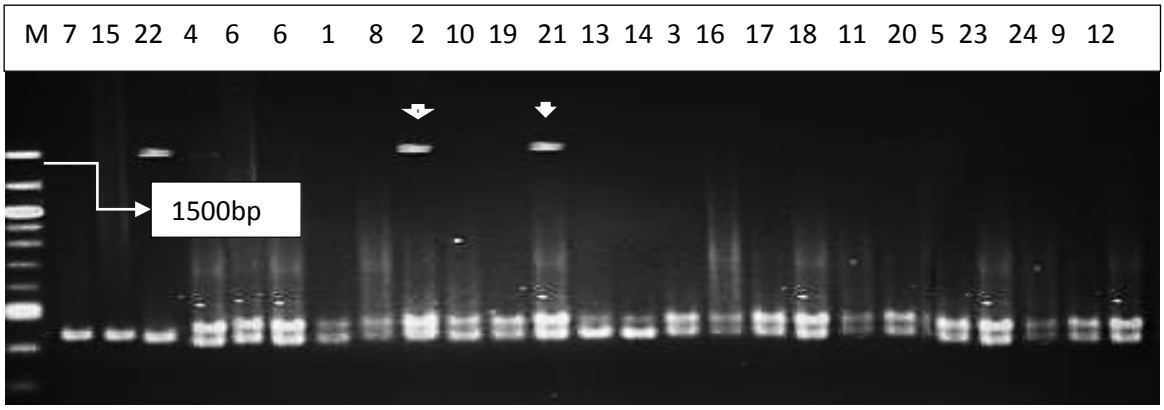

Figure 4.5 : Sat 240 showing two *Canephora* introgressed fragments.Arrows indicate introgressed fragments.(1-SL28; 2-BA; 3-Batian; 4-Erecta; 5-Ruiru 11 cd 93; 6-Catmor 86; 7-Ruiru 11 cd 50; 8-Dr1; 9-HDT; 10-Rume Sudan; 11-Mundonovo;12-Devermarchy; 13-Catura; 14-Colombia; 15-Eugenioides; 16-Mokka; 17-Typica; 18-Blue Mountain; 19-Ruiru 11 cd 80; 20-Bourbon; 21-Arabusta; 22-Canephora; 23-Robarbica; 24.-Sarchmore)

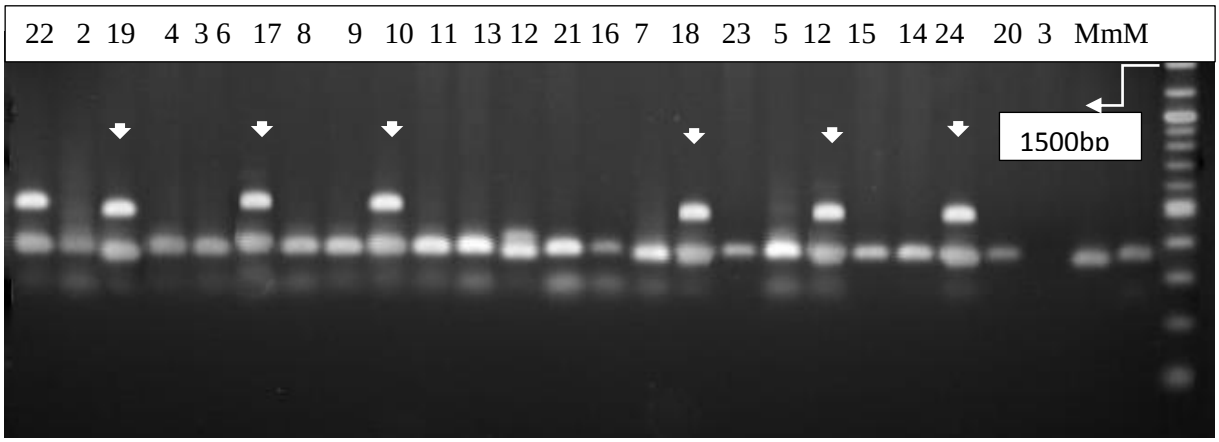

Figure 4.6: Sat 172 showing six *Canephora* introgressed fragments.Arrows indicate introgressed fragments. (1-SL28; 2-BA; 3-Batian; 4-Erecta; 5-Ruiru 11 cd 93; 6-Catmor 86; 7-Ruiru 11 cd 50; 8-Dr1; 9-HDT; 10-Rume Sudan; 11-Mundonovo; 12-Devermarchy; 13-Catura; 14-Colombia; 15-Eugenioides; 16-Mokka; 17-Typica; 18-Blue Mountain; 19-Ruiru 11 cd 80; 20-Bourbon; 21-Arabusta; 22-Canephora; 23-Robarbica; 24.-Sarchmore)

The distribution of introgressed *Canephora* fragments are shown in table 4.5 below.

Table 4.5: *Canephora* introgressed fragments in arabica

Genotype	Canephora Introgressed Fragments in Arabica		
	Sat 254	Sat 172	Sat 240
Arabusta	0	0	1
BA	0	0	1
Batian	0	0	0
Erecta	0	0	0
Ruiru 11 cd 93	1	1	0
Catmor line 86	0	1	0
Ruiru 11 cd 50	1	1	0
Dr1	1	0	0
HDT	0	1	0
Rume Sudan	0	0	0
Mundonovo	0	0	0
Devermarchy	1	0	0
Sarchmore	1	0	0
Catura	0	0	0
Colombia	0	1	0
Robarbica	0	0	0
Mokka	0	0	0
Typica	0	0	0
Blue Mountain	0	0	0
Ruiru 11 cd 80	1	1	0
Bourbon	0	0	0
SL 28	0	0	0
Total	6	6	2
% Introgression	27.3	27.3	9.1

CHAPTER FIVE

DISCUSSION, CONCLUSIONS AND RECOMENDATIONS

5.1 DISCUSSION

The transfer of desired traits from related wild diploid coffea species into the cultivated allotetraploid *C. arabica* is essential in coffee breeding to develop pest/disease-resistant cultivars. The present work is an attempt to gain insights into alien introgression in *C. arabica*. Efficient use of the genetic resources available in wild diploid species is essential for the continued improvement of arabica coffee varieties (Van der Vossen, 2001; Lashermes *et al*., 2009). However, breeding programmes face considerable difficulties due to the long generation time of coffee trees, the high cost of field trials, and the imprecise breeding strategies.New insights into the introgression of *C. arabica* are particularly valuable to improve current breeding methodology.

To identify introgressed fragments from *C.canephora* into *C. arabica,* only *C. canephora* fragments that were arabica specific were considered as introgressed fragments as amplified by SSR markers (evenly distributed throughout eukaryotic nuclear, chloroplasts, and mitochondrial genomes and are inherited as Mendelian co-dominant markers). From the total of 13 SSRs used in this study, the detected alleles were assigned according to their putative species origin (*C. canephora*) for 3 SSRs. For 10 SSRs, it was difficult to assign alleles to their putative species origin due to size overlapping for the alleles of the two genomes (*C. canephora* or *C. eugenioides*) . The result indicated higher allelic values in the *C. eugenioides* genome part in arabica as compared to *C. canephora* genome. The number of alleles per

marker ranged from one to four for *C. eugenioides* genome while it ranged from one to six for *C. canephora* genome. The presence of a large part of Eugenioides genome in the arabica coffee could be attributed to the origin of arabica species.

The three markers that were able to detect introgressed *Canephora* fragments present in arabica were Sat 254, Sat 240, and Sat 172. The percentage of introgressed *Canephora* fragments ranged from 9.1 % to 27.3 %. Sat 254 and Sat 172 revealed 27.3 % while Sat 240 revealed 9.1 % . Sat 254 revealed *Canephora* introgresssed fragments in Ruiru 11 F1 hybrids (CD,93,CD50,CD80), Devermachy, Tanganyika drought resistance (Dr1) and Sarchmore. The size of introgressed DNA fragments was 1000bp. Sat 172 revealed *Canephor*a introgressed fragments in Ruiru 11 F1 hybrids (CD93,CD50,CD80), Colombia,Hibrido de Timor (HdT) and Catimor line86. The size of introgressed DNA fragments was 500bp. Both primers revealed Canephora introgressed fragments in Ruiru 11 F1 hybrids indicating that *Canephora* was one of the parents of Ruiru 11 F1 hybrids or the parents of Ruiru 11 F1 hybrids were derivatives of *Canephora*. Sat 240 revealed Canephora introgressed fragments in Arabusta and BA (indegeneous coffee from India). The size of introgressed DNA fragments was 1500bp. These results are in agreement with previous research. Genome introgression of C. *canephora* into accession of Híbrido de Timor studied using AFLP molecular marker technique found out that introgression of *C. canephora* genome ranged from 8% (in Catimor 3) to 25% (in Sachmor) (Lashermes *et al.,* 2000), using AFLP (amplified fragment length polymorphism) markers (Lashermes *et al.,* 2000a) recently estimated that the approximate amounts of introgressed materials in many introgressed *arabica* lines ranged from 8% to 27% of the *C. canephora* genome.

Analysis of crop genetic diversity is very important for breeding and conservation programs, and molecular markers offer an approach to unveil the genetic diversity among different species and cultivars based on nucleic acid polymorphisms. Extent of distribution, areas sampled and plant characteristics such as mode of reproduction, breeding behaviour and generation time are some of the important parameters that determine the level of genetic variability revealed in a species (Bhat *et al.,* 1999). Diversity in genetic resources is the basis for genetic improvement. 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. The knowledge of the genetic diversity is important for efficient management of germplasm and utilisation of material in breeding programmes.

In this study, two markers, RAPD and SSR were simultaneously used to investigate the genetic diversity among 24 coffee accessionss consisting of (1) *C. canephora*, (1) *C. eugenioides* and (22) arabica. However, Results of the present study using both SSR and RAPD demonstrated the presence of low genetic variation within *C. arabica* genotypes as compared to the diploid species. Furthemore the results did indicate high genetic similarity between *C. arabica* and *C. eugenioides* as compared to *C. canephora* thus supporting the previous research (Lashermess *et al.,* 2009 which did confirm *C. eugeniodes* as the maternal parent of *C. arabica*.

Using thirteen SSR primers, a total of 50 alleles were amplified among 24 coffee genotypes. Of these amplified alleles, 33 were polymorphic. The number of

amplified alleles per primer varied from 2.0 to 6.0 with an average value of 3.8 alleles and recording average polymorphic bands of 2.5. percent. Polymorphism ranged from 33% to 100%, with a mean of 65% polymorphism. These results are relatively similar to previous results. Anthony *et al. (*2002) reported an average number of 4.7 alleles per primer using only six primers in arabica coffee collections containing four Typica, five Bourbon and 10 subspontaneous derived accessions. Using 34 SSRs, Moncada and McCouch (2004) reported an average of 2.5 and 1.9 amplified alleles per primer in 11 wild coffee genotypes and 12 cultivated Arabica coffee, respectively, with the number of alleles ranging from one to eight. Maluf *et al.* (2005) also reported an average number of 2.87 alleles in 28 cultivated Arabica lines using 23 SSRs . Teressa *et al.* (2010) reported total of 209 alleles for 32 SSR markers across 133 Arabica accessions. Out of 209 alleles, 200 alleles were polymorphic for all samples. The number of observed alleles per SSRs varied from two to fourteen with an average of 6.5 alleles for all arabica collection. The reason for such difference could be due to the sample size and the type of coffee genotypes used in the previous studies as compared to the present study. The other reason could be the number of SSRs used and their genome coverage (Teressa *et al.*, 2010) .

The number of alleles produced in this study by thirteen SSR primers was low as compared to RAPD. This illustrated the inadequacy of the used SSR primers to scan different parts of the genome and topinpoint genetic differences between coffee genotypes. Hence, a wider range of informative SSR primers need to be adopted for successful fingerprinting of *C. arabica* genotypes.

The UPGMA method of cluster analysis classified genotypes into four main clusters.The first main cluster contained genotypes *Canephora* and Arabusta, the second cluster contained Eugenioides and the third cluster BA. Eugenioides and Canephora were genetically the most divergent genotypes from all other genotypes with dissimilarity of 1.01. The fourth main cluster contained 20 *C.arabica* genotypes and was further divided into three sub-clusters. The first sub-cluster contained Bourbon derived coffee genotypes while the second sub-cluster contained typica derived genotypes while the third subcluster contained SL28 and Batian.

The number of amplified bands in RAPD primers varied from 3.0 to 12.0. The total number of polymorphic fragments produced was 65. The average bands produced by the ten primers was 7.9 and recording average polymorphic bands of 6.5. Percent polymorphism ranged from 50% to 100%, with a mean of 81% polymorphism. Relatively similar results were obtained from previous work. Lashermes *et al.* (1993) used 23 primers amplified 118 fragments ,number of fragments ranging from 1 to 8 and average of 4.3 fragments per primer. Agwanda *et al.* (1997) reported average of 8 fragments in *C. arabica* with number of fragments ranging from 1 to 15, Anthony *et al.* (2001) used RAPD (150 decamers) on wild and semi wild cultivars of ArabicaTotal = 118 ,16 polymorphic primers with 29 polymorphic amplicons out of atotal of 106. Terezinha *et al.* (2002) used fifty two primers on forty coffee species and reported the average number of polymorphic bands was 6.69 per primer among all genotypes, and 1.27 among Arabica coffee genotypes. Anthony *et al. (*2002) reported an average number of 4.7 alleles per primer using only six primers in Arabica coffee collections containing four Typica, five Bourbon and 10 subspontaneous derived accessions. Aga *et al.* (2003) used twelve RAPD primers

to assess genetic diversity among 144 genotypes representing 16 *C. arabica* populations. The number of polymorphic bands detected with each primer ranged from 2 to 9 with a mean of 6.2 bands per primer. Banding patterns ranged in percentage polymorphism from 37 % to 73 % with an overall mean of 56 % for the populations analyzed, Moncada and McCouch (2004) reported an average of 2.5 and 1.9 amplified alleles per primer in 11 wild coffee genotypes and 12 cultivated Arabica coffee, respectively, with the number of alleles ranging from one to eight.Tshilenge *et al.* (2009) established high variability in the Co`ngolese *C. canephora* using 7 primers. Kathurima *et al.* (2012) using 14 RAPD primers analysed 24 coffee genotypes and reported 83 amplified fragments with arange of 2 to 12 fragments per primer and 35 polymorphic fragments. Lal and Chaturvedi (2013) used a total ten RAPD primers to check the genetic variation in sixteen different accessions of *C. roseus*. Out of these ten, four primers gave satisfactory and reproducible bands.Total 266 bands were observed. From these total bands, 234 were polymorphic while, 32 bands were monomorphic. This resulted in total polymorphism of 87.96%. Bigirimana *et al.* (2013) using six primers reported atotal bands of 19,range between 2-5 and average of 3.2. The reason for such difference could be due to the small sample size and the type of coffee genotypes used in the previous studies as compared to the present study.

The UPGMA method of cluster analysis classified genotypes into four main clusters.The first main cluster contained genotypes *Canephora* and Arabusta, the second cluster contained Eugenioides, the third cluster Erecta and BA and the fourth cluster contained 20 *C. arabica* and was further divided into three sub-clusters. The first sub-cluster contained Bourbon derived coffee genotypes while the second sub-cluster contained typica derived genotypes while the third subcluster contained

SL28 and Batian.Sarchmore and SL28 were genetically the most divergent genotypes from all other genotypes with dissimilarity of 1.12.

The polymorphism detected by both markers ranged from 33% to 100% for SSR and 50% to 100% for RAPD with average polymorphism of 65% and 81% respectively. These results demonstrated that RAPD were suitable for genetic diversity studies in coffee accessions. Considering that the coffee genotypes evaluated in this study originated from different countries (Kenya, Tanzania, Costa Rica, India, Portugal, Brazil, Sudan , Uganda Guatemala and Colombia), the similarities (for both SSR and RAPD results) observed among Arabica genotypes, attests to the narrow genetic diversity among Arabica coffee as reported in other studies (Lashermes *etal.*,1993). This could be attributed to the allotetraploid origin, reproductive biology , evolution of *C. arabica* and may also be explained by the high level of homozigosity as *C. arabica* is a self-pollinated species (Lashermes*et al.,* 1995, 1999). From the general analysis, the 24 coffee accessions clustered according to the three different species namely *C. eugenioides*, *C. canephora* (Robusta) and *C. arabica* (Arabica). This is in agreement with prevoius research. Kathurima *et al.* (2012) reported coffee accessions clustered according to the three different species namely Eugenioides, Robusta and Arabica, Gimase *et al.* (2014) reported coffee genotypes clustered according to the two different species of origin namely Robusta and Arabica. Thus, for rapid improvement in breedingwork, widening of the existing genetic diversity through interspecific hybridisation is desirableand by having more introductions especially from the centre of diversity (Antony *et al.*, 2001), initiate hybridisation programmes to create variability and use of diploid species as a source of desirable genes (Lashermes *et al.*, 1995, 1999), diversifying the

genetic base and increasing the number of varieties released for production with different genetic composition is vital, which helps to reduce looses due to disease out break and other constraints.Similar observation was made by Lashermes *et al.* (1993) and Agwanda *et al.* (1997).

5.2 CONCLUSIONS

Results of the present study using both SSR and RAPD demonstrated the presence of low genetic variation within *C.arabica* genotypes as compared to the diploid species (*C.canephora* and *C.eugenioides*. The narrow genetic base in arabica coffee may also be explained by the high level of homozygosity as *C. arabica* is a self-pollinated species. Thus this call for widening of the existing genetic base by introducing more accessions in Kenyan cultivated Arabica coffee as well as exploring the wild coffee arabica.

A combination of different markers may provide more reliable information about genetic diversity compared to the use of a single marker because errors presented by one marker could be minimized using other markers.Among the two marker types, RAPD marker system appeared to be more informative in detecting genetic diversity compared with SSR marker and was, therefore, favored for further genetic analysis.

In this study SSR markers were able to detect *C.canephora* introgressed fragments in Arabica thus indicating the possible use of SSR markers in marker assisted selection to screen for *C.canephora* introgressed fragments in Arabica. The present study has revealed SSR marker approach is highly efficient and reproducible not only for identification of hybrids but also useful for marker inheritance and hybrid progeny analysis in coffee.

5.3 RECOMMENDATIONS

1) From the study, RAPD markers are more polymorphic than the SSR makers hence there is need to be employ more RAPD primers for coffe genetic studies to increase the genome coverage and the genetic information revealed.

2) The availabe SSR markers produced fewer alleles compared to previous results thus showing the inadequacy 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 employed for successful fingerprinting of *C.arabica* genotypes.
3) There is need to determine the nature of the introgressed *C.canephora* fragments revealed by the three SSR markers in Arabica hybrids and their agronomic significance if any.
4) The number of polymorphic SSR primers currently available for characterisation of Coffee is limited. Development of additional polymorphic SSR primers for effective characterisation of Arabica coffee genotypes is suggested as a future research.

REFERENCE

Aga, E., Bryngelsson, T., Bekele, E., and Salomon, B. (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.

Agwanda, C., Lashermes, P., Trouslot, P., Combes, P. and Charrier, A. (1997). Identification of RAPD markers for resistance to Coffee Berry Disease, Colletotrichum kahawae, in Arabica coffee. *Euphytica*, 97: 241–248.

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.

Ameha, M. (1990). Heterosis and Arabica coffee breeding in Ethiopia. *Plant Breeding Abstrac*t, 60: 594–598.

Anthony F, Astorga-Domian CG, Quiros O, Bertrand B, Etienne H, Topart P, Lashermes P (2000). Genetic diversity of wild and cultivated coffees (Coffea arabica) revealed by molecular markers. Proceedings of the 19th Latin American Symposium on Coffee, Oct. 2-6, Memory, San Jose, pp: 251-261.

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

Anthony, F., Combes, M.C., Herrera, J.C., Prakash, N.S., Berthand, B. and Lashermes, P. (2001a).Genetic Diversity and introgression analysis in coffee (*Coffea arabica* L.) using molecular markers. ASIC 19th. Trieste, Italy.

Anthony, F., Combes, M.C., Herrera, JC, Prakash, N.S., Berthand, B. and Lashermes, P. (2001b). Genetic Diversity and introgression analysis in coffee (*Coffea arabica* L.) using molecular markers. ASIC 19th. Trieste, Italy.

Baruah, A., B. Naik, P. S. Hendre, R. Rajkumar, R. Rajendrakumar, and R. Aggarwal. (2003). Isolation and characterization of nine microsatellite markers from *Coffea arabica* L.showing wide cross-species amplifications. *Molecular Ecology Notes*, 3: 647–650.

Beckman, J.S., and Weber, J.L. (1992). Survey of human and rat microsatellites. *Genomics*, 12: 627-631.

Bellachew, B. (1997). Arabica coffee breeding in Ethiopia: A Review. ASIC 17th. Nairobi, Kenya.

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

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

Bhat, K.V., Babrekar, P.P & Lakhanpaul, S. 1999. Study of genetic diversity in Indian and exotic sesame (Sesamum indicumL.) germplasm using random amplified polymorphic DNA (RAPD) markers. *Euphytica*, 110: 21–23.

Bigirimana, J. N.Kiarie, J.W. Muthomi, D. Gahakwa, N.A. Phiri, E.K. Gichuru and D.J. Walyaro (2013). Genetic Diversity among Disease Resistant Coffee Varieties and Cultivars in Rwanda Based On RAPD and SSR Markers. *Journal of Renewable Agriculture*, 1(6): 106-112

Bittenbender, H. C., Schmitt, D. P., Serracin, M., and Cavaletto, C. G. (2001). Fukunaga, a Coffee Rootstock Resistant to the Kona Coffee Root-Knot Nematode. New Plants for Hawaii. College of Tropical Agriculture & HumanResources. University of Hawaii at Manoa.

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.

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.

Bustamanate, J.P., and Polanco, D.L. (1999). Molecular characterization of coffee (*Coffea arabica* L.) genotypes. *Anales-de-Botanica-Agricola*, 6:21-28.

Kathurima, Kenji G. M.2, Muhoho S. M.2, Boulanger R.3, Gichimu B. M.1 and Gichuru E. K.1 (2012). Genetic diversity among commercial coffee varieties, advanced selections and museum collections in Kenya using molecular markers. *International Journal of Biodiversity and Conservation*, 4(2). pp. 39-46.

Carvalho A (1952) Taxonomia de *Coffea arabica* L. caracteres morfologicas dos haploids. *Bragantia*, 12: 201–212

Carvalho, A. (1988). Principles and Practice of Coffee Plant Breeding for Productivity and Quality Factors: Coffea arabica. In Clarke, R. J., and Macrae, R., eds. Coffee. *Elsevier Applied Science*, Pp. 129-160.

Carvalho, A., Ferwerda, F.P., Frahm-Leliveld, J.A., Medina, D.M., Mendes, A.J.T. and Monaco, L.C. (1969). Coffee (Coffea arabica L. & C. canephora Pierre ex Froehner). In: F.P. Ferwerda & F. Wit (eds.) O*utlines of perennial crop breeding in the tropics*, Veenman & Zonen, Wageningen, pp. 189–241.

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

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.

Clarindo, w. R. and Carvalho (2008). First Coffea arabica karyogram showing that this species is a true allotetraploid. *Plant Systematics and Evolution*, 274: 237-241.

Clarke, R. J., and Macrae, R. (1988). *Coffee – Volume* 4: Agronomy. Elsevier Applied Science.

Clifford, M. N., and Willson, K.C. (1985). *Coffee: Botany, Biochemistry and Production of Beans and Beverage*. Croom Helm Ltd, Sydney.

Combes, M.C., Andrzejewski, S, Anthony, F, Bertrand, B, Rovelli, P, Graziosi, G. and Lashermers, P. (2000). Characterization of microsatellite loci in Coffea Arabica and related coffee species. *Molecular Ecology*, 9:1171-1193.

Cramer, P.J.S. (1957). Review of literature of coffee research in Indonesia. SIC Editorial International American Institute of Agricultural Sciences, Turrialba, Costa Rica.

Cristancho, M. (1999). Genetic diversity of the cenicafe germplasm collection. Third International Seminar on Biotechnology in the Coffee Agro industry, Londrina, Brazil.

Crochemore, M. L., Nunes, L. M., Andrade, G. A., Molinari, H. B. C., and Vasconcellos, M. E. (2004). Varietal Identification of Coffee Seeds by RAPD Technique. *Brazilian Archives of Biology and Technology*, 47:7-11.

Cros, J., Combes, M.C., Trouslot, P., Anthony, F., Hamon, S., Charrier, A., Lashermes, P. (1998): Phylogenetic analysis of chloroplast DNA variation in CoffeaL. *Molecular Phylogenetics and Evolution*, 9: 109-117.

Crouzillat, D., Rigoreau, M, Bellanger, L., Priyono, P., Mawardi, S., Ahrudi, S.Y.,McCarthy, J., Tanksley, S., Zaenudin, I. and Petiard, V. (2004). A Robusta consensus genetic map using RFLP and microsatellite markers for the detection of QTL. ASIC 20th. Bangalore, India.

Cubry, P., Musoli, P., Legnate, H., Pot, D., de Bellis, F., Poncet, V., Anthony, F., Dufour, M., Leroy, T. (2008). Diversity in coffee assessed with SSR markers: Structure of the genus Coffea and perspectives for breeding. *Genome*, 51: 50-63.

Davis, A.P., R. Govaerts, D.M. Bridson and P. Stoffelen. (2006). An annotated checklist of the genus Coffea L. (Rubiaceae). *Botanical Journal of Linnean Society*, 152: 465-512.

Diniz LFC, Ruas, C.F., Carvalho, V.P., Torres, F.M, Ruas, E.A. (2005). Genetic Diversity Among Forty Coffee Varieties Assessed by RAPD Markers Associated with Restriction Digestion. *Brazilian Archives of Biology and Technology*, 48(4): 511-521.

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.

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

Economic survey, Central Bureau of Statistics, 2010, Nairobi, Kenya)

Economic Survey, Central Bureau of Statistics, 2006, Nairobi, Kenya).

Etienne, H., Anthony, F., Fernandez, S., Lashermes, P., and Bertrand, B. (2002). Biotechnological applications for the improvement of coffee (*Coffea Arabica* L.). *In Vitro Cellular and Development Biology-Plant*, 38:129-138.

Ferwerda, F. P. (1976). Coffee. In N.W. Simpsons, ed. *Evolution of crop plants*, Longman, London. Pp. 257-260.

Frankham, R., Briscoe DA, Ballou JD (2002). Introduction to Conservation Genetics. 4th Edition., Cambridge University Press, New York, USA.

Gichimu, B.M. and Omondi CO (2010a). Early performance of five newly developed lines of Arabica coffee under varying environment and spacing in Kenya. *Agriculture and Biology Journal of North America*, 1: 32-39.

Gichuru, E.K, Agwanda CO Combes MC Mutitu EW, Ngugi ECK, Bertrand B, Lashermes P (2008). Identification of molecular markers linked to a gene conferring resistance to coffee berry disease Colletotrichum kahawae in *Coffea arabica*. *Plant Pathololog*y, 57: 1117-1124.

Glaubitz, J. C. and Moran, G. F. (2000). Genetic tools: the use of biochemical and molecular markers. In Young, A., Boshier, D, and Boyle, T., eds. *Forest Conservation Genetics-Principles and Practice.* CSIRO Publishing and CABI Publishing. Pp. 39-59.

Gupta, P., Varshney, R. K., Sharme, P. C., and Ramesh, B. (1999). Review: Molecular markers and their applications in wheat breeding. *Plant Breeding,* 118:369-390.

Hajjar,R ., Hodgkin T (2007) the use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica*, 156:1–13

Hancock, J. M. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms, Goldstein, D. B., and Schlotterer, C., eds. Microsatellites-Evolution and Applications. Oxford University Express

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

Herrera, J.C, Combes, M.C., Anthony, F. and Lashermes, P. (2001). Efficient use of coffee genetic resources: Molecular analyses of genome interactions in the arabusta hybrid (*Coffea arabica* X *C. canephora*). ASIC 19th. Trieste, Italy

Hoelzel, A. R., and Green, A. (1998). PCR protocols and population analysis by direct DNA sequencing and PCR-based DNA fingerprinting. In Hoelzel, A. R., ed. Molecular genetic Analysis of Populations- A practical approach. Oxford University Press. Pp.201-233.

Hue, T.T.M. (2005). Genetic variation in cultivated coffee (*Coffea arabica* L.) accessions in Northern New South Wales, Australia. Master's Thesis, Southern Cross University.

Jones, C. J., Edwards, K. J.,Castaglione, S., Winfield, M. O., Sala, F., Wrel, C. v. d.,Bredemeijer, G., Vosman, B., Matthes, M., Daly, A.,Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volkaert, G., Rueda, J., Linacero, R., Vazquez, A., and Karp, A. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular breeding*, 3: 381-390.

Jump, A.S., Marchant R., Penuelas, J. (2008) Environmental change and the option value of genetic diversity. *Trends Plant Science*, 14: 51-58.

Karp, A., Kresovich, S., Bhat, K.V., Ayad, W.G. and Hodgkin, T. (1997). Molecular tools in plant genetic conservation: a guide to the techniques. *International Plant Genetic Resources Institute Technical Bulletin,* 2: 5–47.

Krug, C. A., and Carvalho, A. (1951). The genetics of Coffea. *Advance in Genetics,* 4:127-158.

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.

Lal and Chaturvedi, (2013). Random amplified polymorphic DNA (RAPD) fingerprinting for evaluation of intraspecific catharanthus roseus populations collected fromg ujarat, india. *International journal for biotechnology and plant agricultural science*, 2(2): 373-385.

Lashermes, P., Combes, M.C., Topart, P., Graziosi, G., Bertrand, B. and Anthony, F., (2000b). Molecular breeding in coffee (Coffea arabicaL.). 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., Agwanda, C.O., Anthony, F., Combes, M.C., Trouslot, P. and Charrier, A.(1997). Molecular marker-assisted selection: A powerful approach for coffee improvement. ASIC 17th. Nairobi, Kenya.

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

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

Lashermes, P., Combes, M. C., Prakash, N. S., Trouslot, P., Lorieux, M., and Charrier, A. (2001). Genetic linkage map of Coffea canephora: Effect of segregation distortion and analysis of recombination rate in male and female meiosis. *Genome,* 44: 589-596.

Lashermes, P., Combes, M.C, Ansaldi C, Gichuru, E.K., Noir, S. (2010). Analysis of alien introgression in coffee tree (*Coffea Arabica* L.). *Molecular breeding*, 27: 223-232.

Lashermes, P., Combes, M.C., Cros, J., Trouslot, P., Anthony, F. and Charrier, A. (1995). Origin and genetic diversity of *Coffea arabica* L. based on DNA molecular markers. ASIC, 16th. Kyoto.

Lashermes, P., Comes, M.C., Trouslot, P., Anthony, F., Hamon, S. and Charrier, A. (1999). Molecular characterization and origin of the Coffea Arabica L. genome. *Molecular Genomics and Genetics*, 261:259-266.

Lashermes, P., Cros, J., Combes, M.C., Trouslot, P., Anthony, F., Hamon, S. and Charrier, A. (1996b). Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus Coffea L. *Theoretical Application & 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., Trouslot, P., Anthony, F., Comes, M.C. & Charrier, A. (1996a). Genetic diversity for RAPD markers between cultivated and wild accessions of Coffea arabica. *Euphytica*, 87: 59–64.

Lecolier, A., P. Besse, A. Charrier, T.N. Tchakaloff and M.Noirot. (2009). Unravelling the origin of Coffea arabica 'Bourbon pointu' from La Reunion: a historical and scientific perspective. *Euphytica*, 168: 1-10.

Levinson, G. and Gutman, G.A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology Evolution*, 4: 203–221.

Lin, J.-J., J. Kuo and J. Ma, 1996. Effect of different primer combinations on the resolution of ALFP in plants with small genomes. *Focus*, 18**:** 68–69.

Lopes, C. R. (1993). Phylogenetic studies of some species of the genus Coffea. II-Numerical analysis of isoenzymatic data. Revista Brasileira de Genetica (*Brazilian Journal of Genetics*) 16:393-407.

M. P. Maluf, M. Silvestrini, O. G. Filho and C. A. Colombo, (2005).Genetic diversity of cultivated Coffea Arabica inbred lines assessed by RAPD, AFLP and SSR markers systems. *Scientia Agricola*, 62: 366–373.

Masumbuko, L. I., Bryngelsson, T., Mneney, E. E., and Salomon, B. (2003). Genetic diversity in Tanzanian arabica coffee using random amplified polymorphic DNA (RAPD) markers. *Hereditas*, 139: 56–63.

Medini, M., Hamza, S., Rebai, A., and Baum, M. (2005). Analysis of genetic diversity inTunisian durum wheat cultivars and related wild species by SSR and AFLP markers. *Genetic Resources and Crop Evolution*, 52:21–31

McGregor, C. E., Lambert, C. A., Greyling, M. M., Louw, J. H., and Waenich, L. (2000). A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (Solanum tuberosum L.) *Euphytica*, 113:135-144.

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.

Mignouna, H. D., Abang, M. M., and Fagbemi, S. A. (2003). A comparative assessment of molecular marker assays (RAPD, SSR and AFLP) for white yam germplasm characterization. *Applied Biology*, 142:269-276.

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

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

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.

Morgante, M. and Olivieri, A.M. (1993). PCR-amplified microsatellites as markers in plant genetics. *Plant Journal*, 3: 175–182.

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

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

Pare, S. (2002). Expériences du commerce équitable du café Recherche et caféiculture,

Pearl, H. M., Nagai, C., Moore, P. H., and Steiger, D. L. (2004). Construction of a genetic map for arabica coffee. *Theoretical & 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.

Penner, G.A., Bush, A., Wise, R., Kim, W., Domier, L., Kasha, K., Laroche, A., Scoles, G., Molnar, S.J. & Fedak, G. 1993. Reproducibility of Random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR Methods Applications*, 2: 341–345.

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

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

Rani, V., Singh, K.P., Shiran, B., Nandy, S., Goel, S., Devarumath, R.M., Sreenath, H.L. & Raina, S.N. (2000). Evidence for new nuclear and mitochondrial genome organizations among high-frequency somatic embryogenesis-derived plants of allotetraploid Coffea arabica L. (Rubiaceae). *Plant Cell Reports*, 19: 1013–1020.

Rovelli, P., Mettulio, R, Anthony, F, Anzueto, F, Lashermes, P, and Graziosi, G. (2000). Microsatellite in Coffea arabica L. pp.123–133. In Sera T, Soccol CR, Pandey A and Roussos S (eds) *Coffee biotechnology and quality*, Kluwer Academic Publishers, Dordrecht, Netherlands.

Russell, J. R., Fuller, J. D., and Macaulay, M. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical & Applied Genetics*, 95:714-722.

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.

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

Sneath, P.H.A., Sokal, R.R. (1973): Numerical taxonomy, the principles and practice of numerical classification; W.H. Freeman & Co, San Francisco, California, P.573.

Statistical Abstract 2006, p. 253 and Economic Survey (2006), Kenya, p. 69).

Statistical abstracts, (2008) kenya.

Stebbins GL (1958) The inviability, weakness and sterility in interspecific hybrids. *Advance Genetics*, 9:147–215.

Stebbins, G.L., (1950) Variation and evolution in plants.Colombia Univ. Press, New York

Steiger, D.L., Nagai, C., Moore, P.H., Morden, C.W., Osgoog, R.V. and Ming, R. (2002). AFLP analysis of genetic diversity within and among Coffea Arabica cultivars. *Theoretical & Applied Genetics,* 185:189-215.

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

Teressa, A., D.Crouzillat, V. Petiard and Pier. (2010) .Genetic diversity of Arabica coffee (Coffea arabicaL.) *collection. Ejast,* 1(1): 63-79

Terezinha, T. Cabral; Ney, S. Sakiyama; L. Zambolim; A.A. Pereira; E.Gonçalves and, Sakiyama (2002). *Crop Breeding and Applied Biotechnology*, 2: 121-129.

Tshilenge, P., Nkongolo, K. K., Mehes, M. and Kalonji, A. (2009). Genetic variation in Coffea canephoraL. (var. Robusta) accessions from the founder gene pool evaluated with ISSR and RAPD. *African Journal of Biotechnology*, 8: 380–390.

United StatesDepartment ofAgricultureForeign Agricultural Service Circular Series December (2013)

Van der Vossen, H.A.H. (2001). Agronomy I: Coffee breeding Practices. In: R: J. Clake and O: G. Vitzthum (eds) *Coffee Recent Development*, Blackwell Science Ltd, London, pp.184–201.

Vienne, D. d., Santoni, S., and Falque, M. (2003). Principal Sources of Molecular Markers in Vienne, D. d., ed. Molecular Markers in Plant Genetics and Biotechnology. *Science publishers*, Inc., Plymouth, UK. Pp. 3-41.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Plot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research,* 23: 4407–4414.

Vossen, H. A. M. v. d. (1985). Coffee selection and breeding. In Clifford, M. N., and Willson, K. C., eds. *Coffee Botany, Biochemistry and Production of Beans and Beverage,* Croom Helm, London-New York-Sydney. Pp. 48-96.

Weber, J.L. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. *Genetics,* 44, 388–396.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005) . DNA Fingerprinting in Plants: Principles, Methods, and Applications. 2nd edition. CRC Press, Taylor and Francis Group.

Welsh, J. and McClelland, M. (1990) . Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, 18: 7213–7218.

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

Wrigley, G. (1988) . Coffee. Longman Scientific and Technical.

Wrigley, G. (1995) . Coffee-Coffea spp. (Rubiaceae). In J. Smartt, and Simmonds, N. W.,eds. *Evolution of crop plants*, Longman Scientific and Technical, Harlow, England. Pp. 438-443.

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.

APPENDICES

Appendix 1: Preparation of DNA samples for loading

DNA samples	2ul
Water	6ul
Loading buffer 6X	2ul
Total volume	10ul

Appendix 2: The lambda preparation table used to estimate the quantity of DNA

		Lambda 3	Lambda 4
	Size Kb	Band amount (total=250ng)	Band amount (total=125ng)
1	21,23	100	50
2	5,15	50	25
3	4,98		
4	4,27	22,5	11
5	3,52(low intensity)	18	9
6	2,02	10	5
7	1,91	10	5
8	1,58	7,5	3,5
9	1,38	7	3,5
10	0,95(950bp)	5	2,5
11	0,83(830bp)	4	2
12	0,56(560bp)	3	
13	0,125(125bp)		

Appendix 3: SSR MASTER MIX

REAGENT	1 Reaction (RXN)
dd H_20	6.65
10x buffer(15Mm $MgCl_2$)	2.5
$MgCl_2$(100m M)	2.5
dNTPs (500 µ M)	3.75
SSR (10 µM)Reverse primer SSR (10 µM)Reverse primer	1.0 1.0
Taq polymerase 5u / µl	0.4
TOTAL	18 µL
DNA (10ng/ µl)	10.0

Appendix 4 : RAPD MASTER MIX

REAGENT	1 Reaction (RXN)
dd H_20	13.85
10x buffer(15Mm $MgCl_2$)	2.5
$MgCl_2$(100m M)	2.5
dNTPs(500 µ M)	3.75
RAPDS(10 µM)	1.0
Taq 5u / µl	0.4
TOTAL	24µL
DNA (1ng/ µl)	4.0

Appendix 5: Preparation of DNA extraction buffers solutions

Buffer A (for 100 ml) (Lysis buffer)

Reagent	Concentration	Quantity
Sorbitol	0.35M	6.38 g
Tris-HCL	0.20M	20 mls
EDTA	0.05M	1.49g
Top up with double distilled water		

NB. Adjust pH 8.0 with conc. HCL.

Preparation of extraction buffer stock solution

Buffer B (100ml) (extraction buffer)

Reagent	Concentration	Quantity
Sodium Chloride(AR)	2.0M	8.77 g
Tris-HCL(approx. 1M)	0.20M	20 mls
EDTA	0.05M	1.49 g
Sarcosil	5 %	9.5 mls
Top up with double distilled water		

The Buffer B is viscose and during preparation it could be maintained under agitation and heating.

Preparation of working buffer solution

The working buffer solutions are prepared prior to use by adding the following reagents in the already prepared stock buffers respectively

Reagent	Buffer A	Buffer B
mixed alkyl trimethyl ammonium bromide (MATAB)	**-**	**2%**
Sodium disuiphite	**-**	**1 %**
Polyvinyl pyloridone (PVP)	**2 %**	