

**DIVERSITY, COMBINING ABILITY AND COFFEE BERRY DISEASE  
(*COLLETOTRICHUM KAHAWAE*) RESISTANCE AMONG ETHIOPIAN AND  
TANZANIAN ARABICA COFFEE GENOTYPES**

**DAMIAN JOSEPH MTENGA**

**A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR  
THE DEGREE OF DOCTOR OF PHILOSOPHY OF SOKOINE UNIVERSITY OF  
AGRICULTURE. MOROGORO, TANZANIA.**

**2016**

## EXTENDED ABSTRACT

### Background

Coffee is the most important stimulant beverage in the world and a major form of cash income for millions of smallholder farmers. The world's coffee trade is dominated by two types Arabica and Robusta. Arabica represents 60-70% whereas Robusta represents 30-40% of global exports. Coffee is one of Tanzania's primary agricultural export commodities accounting for about 5% of total exports value, and generating export earnings averaging USD 100 million per annum over 30 years by 2012. Insufficient information of genetic diversity at molecular level of germplasms and breeding materials existing in the Institute, absence of fast, convenient and high precision screening tools/diagnostic markers for coffee berry diseases (CBD) resistance, narrow genetic base of released improved coffee varieties and little knowledge base of genetics of arabica coffee traits leads to low efficiency in development and release of improved superior arabica coffee varieties in a reduced time. The main objective of this study was to generate information, which will be used to address the above problems. The specific objectives were: to determine the genetic diversity of Tanzanian coffee germplasm and the Ethiopian *Coffea arabica* collection maintained in Tanzania using simple sequence repeats (SSRs) molecular markers; to incorporate coffee berry disease resistance and screen progenies of the crosses between Ethiopian accessions and Tanzanian commercial variety KP423 for Coffee Berry Disease (CBD) resistance genes using SSR markers; to determine the combining ability, heritability and relationships of growth and yield characters of progenies of commercial variety KP423 and selected accessions of Ethiopian collection.

## **Materials and Methods**

Leaf samples were collected from three trees per genotype from Ethiopian collection and other germplasms maintained at Tanzania Coffee Research Institute (TaCRI) Lyamungu for genomic DNA extraction which was carried out at Sokoine University of Agriculture molecular biology laboratory. Thirty SSR primer sets were used in PCR amplification of the extracted DNA. Analysis of Molecular Variance (AMOVA), Principal Component Analysis (PCA) and Cluster Analysis were performed from SSR genetic similarity matrix data using GenStat statistical software version 15.1 (GenStat, 2012). Coffee berry disease (CBD) hypocotyl screening was applied on the progenies and parental genotypes of eleven genotypes selected from the germplasm and Ethiopian collection at TaCRI Lyamungu which were crossed to a susceptible variety KP423. Two gene specific markers Sat 235 and Sat 207 targeting a major CBD resistance gene Ck-1 were used to amplify DNA extracted from leaves of five seedlings of each parental genotypes and crosses (F1s) which survived phenotypic (hypocotyl) CBD resistance screening. For determination of the combining ability, heritability and relationships of growth and yield variables nine genotypes were selected including six from Ethiopian collection, two from the germplasm and a commercial variety KP 423 were used in a half-diallel mating design. Data were collected at Year 1 and 2 after establishment on: Stem girth, Plant height, Length of the longest primary (canopy radius), Number of primaries per plant, Number of berries (flower buds) per cluster, Internode length and Number of bearing primaries per plant. Analysis was based on the fixed effect (model 1) method II by Griffing (1956) while Path coefficients analysis was performed according to Dewey and Lu (1959) using (Sheoran *et al.*, 1998) Statistical Package for Agricultural Research Workers.

## Key Findings

Determination of genetic diversity revealed high observed heterozygosity (0.9993), high percent of polymorphic alleles (80 %) whereas average number of alleles per SSRs locus was 2.5 and polymorphic information content was (0.4128). Principal component analysis identified three significantly different groups with diversity percentages 56.75, 5.69 and 4.66 explained 67.11 % of the total diversity. For CBD resistance gene screening, the genotypes that were clearly amplified by SSR primer Sat 235 and Sat 207 and also showed phenotypic resistance through hypocotyl screening confirmed the presence of the CBD resistance gene *Ck-1* by producing bands similar to the progenitors of CBD resistance. Both general and specific combining ability variances were significant ( $p = 0.01$ ) for all variables except berries per node. Broad sense ( $H^2_{bs}$ ) heritability values were higher than that for narrow sense ( $h^2_{ns}$ ) for all variables. Genotypic correlations between all variables were significant and positive except the correlation between internode length (IL) and number of bearing primaries per plant (NBPP). Similarly phenotypic correlations were positive and significant for all variables except that between number of primaries per plant (NPP) and internode length (IL). Path coefficient analysis reveals that variables length of the longest primary (LLP), stem girth (SG) and internode length (IL) were found to contribute mostly in berries per node. The direct effect due to plant height was found contributing most negatively to berries per node comparing with other variables studied.

## Major Conclusions

High heterozygosity was observed among the genotypes studied with Ethiopian accessions been most diverse. The commercial varieties Bourbon (1) and Kent (2) were grouped in the same group in the PCA (Fig. 5) and clustered in the same cluster in the dendrogram (Fig. 4) revealed low diversity among themselves. On the other hand the

Ethiopian genotypes viz (4, 6, 8) and breeding lines viz (65, 66, 86) were scattered in across the groupings revealed their high heterozygosity. Marker *ssrAY2449* was the most informative (PIC = 0.7390). Presence of the CBD resistance gene *Ck-1* in the genotypes studied was confirmed showing phenotypic hypocotyl resistance and by SSR primer Sat 235 and Sat 207 clear amplification by producing bands similar to the progenitors of resistance. Variables length of the longest primary (canopy radius), internode length and stem girth can be used as an important criteria in selection for superior coffee arabica varieties involving Ethiopian genotypes and commercial variety KP423. The information generated will facilitate selection of the most diverse genotypes for development of superior varieties with broad genetic base for various traits at early growth stages reducing the time for selection and release of superior varieties to growers.

### **Recommendations**

Accessions from the Ethiopian germplasm should be intensively used in the development of new improved arabica coffee varieties with broad genetic base and suitable agronomic characteristics including resistance to coffee berry disease.

Deployment of marker assisted selection for resistance to coffee berry disease should be use microsatellites SAT 235 and SAT 207 to shorten selection cycles and new varieties release time.

Variables such as length of the longest primary (canopy radius), internode length and stem girth should be used as important criteria in selection for superior Arabica coffee varieties involving Ethiopian genotypes and commercial varieties at the early growth stage.

**Implication of this study for Tanzania Coffee Industry**

The diversity of the Breeding materials at Tanzania coffee research institute (TaCRI) is high and can continue to sustain the breeding programme in development of superior coffee varieties with broad genetic base especially involving the Ethiopian Arabica coffee collection. The established use of coffee berry disease (CBD) resistance screening through the use of gene specific markers will facilitate early selection of superior genotypes and segregating populations for resistance to CBD at a shortest possible time reducing the time taken to release varieties using the phenotypic hypocotyl and field screening. Early selection of superior genotypes will also be possible through the use of yield and growth variables as revealed in the study reducing the time taken to release superior varieties. Based on these facts the industry will benefit from the fast and efficient system of releasing new superior varieties which will lead to increased coffee production at reduced cost of managing the CBD, increasing farmer's income and national coffee export forex earnings.

**DECLARATION**

I, DAMIAN JOSEPH MTENGA, do hereby declare to the Senate of Sokoine University of Agriculture that this thesis is my own original work done within the period of registration and that it has neither been submitted nor being concurrently submitted in any other institution.

\_\_\_\_\_  
Damian Joseph Mtenga  
(PhD Candidate)

\_\_\_\_\_  
Date

The above declaration is confirmed

\_\_\_\_\_  
Prof. Paul M. Kusolwa  
(Supervisor)

\_\_\_\_\_  
Date

\_\_\_\_\_  
Prof. S.O.W.M. Reuben  
(Supervisor)

\_\_\_\_\_  
Date

**COPYRIGHT**

No part of this thesis may be reproduced, stored in any retrieval system or transmitted in any form or by any means without prior written permission of the author or Sokoine University of Agriculture in that behalf

## ACKNOWLEDGEMENTS

This work could not have been produced without the help of many people and institutions that in diverse ways supported me technically, morally and financially. It may not be possible to list names of all individuals, institutions and organizations that have contributed or supported this work in different ways. I greatly appreciate and recognize every support or contribution given for this work. I wish to mention some of the contributors; I wish to express my profound gratitude to Prof. Paul M. Kusolwa and Prof. S.O.W.M Reuben for their persistence guidance and availability for consultation during the hard times making my study bearable. Secondly, I wish to express my thanks to Tanzania Coffee Research Institute (TaCRI) my employer for financing this study and provision of site for field experiments. I am deeply indebted to the field staff of TaCRI for assisting me with routine management of the field experiments.

I am grateful to the Sokoine University of Agriculture and in particular the staff of Crop Science and Production Department for creating an enabling environment for smooth academic and laboratory work. Mr. Deo Protas is gratefully acknowledged for great support to make things work out at the molecular study laboratory he is wonderful and will never be forgotten.

Special thanks go to all the great people in the Department of Crop Science and Production at SUA who supported me in one way or the other with pieces of advice, technical and editorial assistance, e-mails and warm smiles to encourage me. I cannot forget the friendship with all the graduate students. Finally, I thank God whom I served day and night for seeing me through to a successful end, may His name alone be forever praised. Amen.

## **DEDICATION**

This work is dedicated to my beloved family who endured with me throughout the period of this study and gave me a lot of encouragement and support.

## TABLE OF CONTENTS

EXTENDED ABSTRACT .....	ii
DECLARATION .....	vii
COPYRIGHT .....	viii
ACKNOWLEDGEMENTS .....	ix
DEDICATION .....	x
TABLE OF CONTENTS .....	xi
LIST OF TABLES .....	xvi
LIST OF APPENDICES .....	xix
LIST OF ABBREVIATIONS AND SYMBOLS .....	xx
CHAPTER ONE .....	1
1.0 Introduction and Background .....	1
1.1 Background.....	1
1.2 Problem Statement.....	2
1.3 Justification of the Study .....	2
1.4 Research Questions .....	3
1.5 Hypotheses .....	3
CHAPTER TWO .....	4
2.0 Literature Review .....	4
2.1 Economic Importance of Coffee .....	4
2.2 Origin and Distribution of <i>Coffea</i> species .....	4
2.3 Morphology and Reproductive Biology of <i>Coffea arabica</i> .....	5
2.3 Significance of Ethiopian <i>Coffea arabica</i> Diversity .....	6

2.4	Arabica Coffee Genetic Improvement.....	7
2.5	DNA Molecular Markers Application in Coffee.....	8
2.6	Microsatellite Makers.....	10
2.7	Coffee Berry Disease ( <i>Colletotrichum kahawae</i> Waller and Bridge).....	11
2.7.1	Economic importance of coffee berry disease.....	11
2.7.2	Symptoms and spread of coffee berry disease.....	11
2.7.3	Inheritance of coffee berry disease resistance.....	12
2.7.4	DNA markers application in coffee diseases studies.....	12
2.8	Coffee Growth and Yield Variables Interrelationships.....	13
2.8.1	Combining ability.....	13
2.8.2	Heritability.....	14
2.8.3	Correlations.....	15
2.8.4	Path coefficient.....	16
	References.....	18
CHAPTER THREE.....		27
Determination of genetic diversity of Tanzanian germplasm and the Ethiopian <i>Coffea arabica</i> collection maintained in Tanzania using simple sequence repeats (SSR) molecular markers.....		27
3.1	Abstract.....	27
3.2	Introduction.....	28
3.3	Materials and Methods.....	29
3.3.1	Plant materials.....	29
3.3.2	Methods.....	29
3.3.2.1	DNA extraction.....	29
3.3.2.2	DNA amplification.....	31

3.3.2.3	Scoring of bands .....	31
3.4	Results .....	32
3.4.1	DNA amplification .....	32
3.4.2	Dendrogram analysis.....	37
3.4.3	Principal component analysis.....	38
3.5	Discussion.....	39
3.5.1	Principal component analysis.....	39
3.5.2	Dendrogram analysis.....	40
3.6	Conclusion.....	40
	References.....	41
CHAPTER FOUR.....		44
	Incorporation of coffee berry disease (CBD) resistance genes and screening of progenies of the crosses between Ethiopian coffee accessions and Tanzanian commercial variety KP423 using SSR markers.....	44
4.1	Abstract.....	44
4.2	Introduction .....	45
4.3	Materials and Methods .....	46
4.3.1	Plant materials .....	46
4.3.2	Microsatellite markers used .....	47
4.4	Phenotypic Screening .....	48
4.5	Screening of Markers of Coffee Berry Disease Resistance.....	49
4.5.1	DNA extraction .....	49
4.5.2	Coffee DNA amplification.....	50
4.6	Results .....	51

4.6.1	Phenotypic screening.....	51
4.6.2	Marker screening.....	54
4.7	Discussion.....	56
4.7.1	Phenotypic screening.....	56
4.7.2	Marker screening.....	56
4.8	Conclusion.....	57
	References.....	58
CHAPTER FIVE .....		61
Combining ability, heritability and relationships of growth and yield variables of progenies of commercial coffee variety KP423 and selected accessions of Ethiopian collection .....		61
5.1	Abstract.....	61
5.2	Introduction .....	62
5.3	Materials and Methods .....	63
5.3.1	Description of the experimental site.....	63
5.3.2	Coffee genotypes.....	63
5.3.3	Hybridization.....	64
5.3.4	Field experimental design data collection and analysis .....	65
5.4	Results .....	66
5.4.1	Genotypic and phenotypic correlations.....	73
5.4.2	Path Coefficient Analysis.....	74
5.5	Discussion.....	78
5.6	Conclusions .....	84
	References.....	86

CHAPTER SIX.....	89
6.0 Conclusions, Recommendations and Future Studies.....	89
6.1 Conclusions .....	89
6.2 Recommendations .....	90
6.3 Future studies.....	91
APPENDICES .....	92

## LIST OF TABLES

Table 1:	Summary amplification of 91 coffee genotypes by 24 SSR markers .....	35
Table 2:	Amplification of 91 coffee genotypes by 24 SSR markers used in the study .....	36
Table 3:	List of coffee genotypes used for coffee berry disease resistance genes screening .....	47
Table 4:	Gene specific SSR markers used in coffee berry disease resistance screening .....	47
Table 5:	Coffee parental genotypes for Coffee Berry Disease resistance mean score on (0-4 scale) .....	52
Table 6:	Coffee F <sub>1</sub> hybrids Coffee Berry Disease resistance mean score on (0-4 scale).....	52
Table 7:	List of coffee genotypes used in the study .....	63
Table 8:	Analysis of variance for growth variables in F <sub>1</sub> coffee hybrids Year 1 (2014) .....	67
Table 9:	Analysis of variance for growth variables in F <sub>1</sub> coffee hybrids Year 2 (2015) .....	68
Table 10:	ANOVA for combining ability and heritability of yield and growth variables of F <sub>1</sub> coffee hybrids Year 1 (2014).....	70
Table 11:	ANOVA for combining ability and heritability of yield and growth variables of F <sub>1</sub> coffee hybrids Year 2 (2015).....	71
Table 12:	Simple correlation of yield and growth variables of F <sub>1</sub> coffee hybrids (1 <sup>st</sup> Year) (2014).....	73
Table 13:	Simple correlation of yield and growth variables of F <sub>1</sub> coffee hybrids (2nd Year) (2015).....	73

Table 14: Genotypic and phenotypic correlation coefficients for growth and yield variables of coffee .....	74
Table 15: Direct and indirect influences of some of coffee growth and yield variables on berries per node .....	76

## LIST OF FIGURES

Figure 1:	A 2 % gel photograph used in scoring coffee genotypes accessions 49 - 91 using SSR marker ZAP25.....	33
Figure 2:	A 2 % gel photograph used in scoring coffee genotypes accessions 1- 48 using SSR marker ZAP25.....	33
Figure 3:	A 2% gel photograph used in scoring coffee genotypes accessions 49-91 using SSR marker R175.....	34
Figure 4:	Dendrogram of 91 genotypes amplification by 30 SSR markers based on genetic distance .....	37
Figure 5:	Principal component analysis showing the main diversity groups. ....	38
Figure 6:	Coffee parental genotypes mean score for Coffee Berry Disease resistance (0-4) scale .....	53
Figure 7:	Coffee F1 hybrids mean score for Coffee Berry Disease resistance on (0-4) scale .....	53
Figure 8:	A 2 % gel photograph showing amplification of coffee parental genotypes, F1 hybrids and a commercial variety KP423 (Table 3) by SSR marker Sat 235.....	55
Figure 9:	A 2 % gel photograph showing amplification of coffee parental genotypes, F1 hybrids and a commercial variety (Table 3) by SSR marker Sat 207. ....	55
Figure 10:	Path diagram and coefficients of factors influencing the number of berries per node of coffee. P's are the direct effects; r's are the genotypic correlation coefficients .....	75

**LIST OF APPENDICES**

Appendix 1:	List of materials from Ethiopian coffee collection and germplasm used for diversity study.....	92
Appendix 2:	List of SSR markers used in coffee diversity study.....	95

**LIST OF ABBREVIATIONS AND SYMBOLS**

AFLP	Amplified Fragment Length Polymorphism
ANOVA	Analysis of variance
BC1	Backcross one
Bp	Base pair
CBD	Coffee berry disease
CENICAFE	Centro de Investigaciones del Café
CLR	Coffee Leaf Rust
CRF	Coffee Research Foundation
CTAB	Cetyltrimethyl ammonium bromide
CWD	Coffee Wilt Disease
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide Triphosphates
FAO	Food and Agriculture Organization
GCA	General Combining Ability
GDP	Gross Domestic Product
HO	Observed heterozygosity
ICO	International Coffee Organization
IRD	Institut de recherche pour le développement
ISSR	Inter-simple sequence repeat
l.e	Low electroendosmosis
MAS	Marker-Assisted Selection
Mg	Magnesium
MgCl <sub>2</sub>	Magnesium Chloride
Mm	Milli mole

NaCl	Sodium Chloride
ng / $\mu$ l	Nanogram per micro litre
ORSTOM	Office de la recherche scientifique et technique outre-mer
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
pH	Hydrogen ion concentration
PIC	Polymorphism Information Content
QTL	Quantitative Trait Loci
RAPD	Random Amplified Polymorphic DNA
RCBD	Randomized Complete Block Design
RFLP	Restriction Fragment Length Polymorphism
rP	Percent of polymorphic alleles
Rpm	Revolutions per minute
SCAR	Sequence Characterized Amplified Region
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
TAE	Tris-Acetate-EDTA
TCB	Tanzania Coffee Board
TPCR	Touch down Polymerase Chain Reaction
USA	United States of America
UV	Ultra violet
wt/vol	Weight by volume
$\mu$ l	micro litre
$\mu$ M	micro molar

## CHAPTER ONE

### 1.0 Introduction and Background

#### 1.1 Background

Coffee is the most important stimulant beverage widely consumed in the world, and is a very important source of foreign exchange for many countries. Worldwide, coffee provides a major form of cash income for millions of smallholder farmers and is a significant source of export earnings for many nations including Tanzania (ICO, 2015). The coffee industry is well-organized and has future markets in London and New York. The world coffee trade is mainly dominated by two types, viz Arabica and Robusta (ICO, 2015). Arabica coffee represents approximately 60-70 % of all coffee exports whereas Robusta coffee represents approximately 30-40 % of global exports (ICO, 2015). Tanzania produces both Arabica and Robusta coffee. The major Arabica coffee growing regions are Arusha, Kilimanjaro, Mbeya, Ruvuma. It is also grown in Morogoro, Kigoma, Manyara, Mwanza, Rukwa, Tanga, Iringa and Mara. Robusta coffee is mainly grown in Kagera region (TCB, 2012). The average acreage under coffee is 229 000 hectares of which 90 % is managed by smallholders while Estates manages the rest 10 % (ICO, 2015). The major production constraints are poor management of pests and diseases (CBD and CLR in Arabica coffee producing areas), low productivity of coffee farms due to aged trees, poor extension services, low farm gate prices, the impact of climate change and weak farm financial services (TCB, 2012). Taking into consideration the constraints above, the sustainable solution is to develop superior arabica coffee varieties that will be resistant to coffee berry disease (CBD) and coffee leaf rust (CLR), high yielding, having excellent cup quality at the shortest possible time. The overall objective this study was to generate information which will aid in development of arabica coffee genotypes that are higher yielding and having broad base of resistance to coffee berry disease from diverse

accessions from Tanzanian and Ethiopian collections maintained at Tanzania Coffee Research Institute.

## **1.2 Problem Statement**

The genetic diversity of Tanzanian cultivated arabica genotypes is narrow. However efforts have been made in development of new improved varieties with broad genetic base from different sources of germplasm. These efforts are limited due to lack of information on the diversity of the existing gene pool of advanced breeding materials and introduced accessions. Coffee berry disease (CBD) is a major disease of arabica coffee attacking berries. Currently, selection of resistant genotypes involves the inoculation of the hypocotyls when they are 3-5cm at 5-6 week after sowing using a standard concentration of the pathogen inoculum. This is followed by individual seedling assessment using a standard disease severity scale. Absence of fast, convenient and high precision screening tools for CBD reduces the precision of selection and extends the time taken to develop and release superior arabica coffee varieties with broad genetic base for CBD resistance to farmers. The genetics and relationships of growth and yield variables of Tanzanian and Ethiopian arabica coffee genotypes maintained in Tanzania are not well known. The current utilization of these genotypes in the breeding programme is based on trial and error leading to little gains and too long field evaluation time. Knowledge of these interrelationships will facilitate selection of appropriate improvement approaches reducing selection cycles and time needed to release superior arabica coffee varieties to farmers.

## **1.3 Justification of the Study**

Establishment of the diversity of existing breeding materials will facilitate proper selection of genotypes for crossing in diverse objectives of the breeding programme.

Knowledge of the existing diversity in the breeding materials will also aid proper planning in exchange and acquisition of new germplasm. Availability of fast, convenient and high precision screening tools/methods at early stages of growth will facilitate early selection of highly resistant genotypes for development of superior varieties with broad genetic base at reduced time. Knowledge of the associations of characteristics of agronomic and economic importance will enable breeders to choose improvement strategies rationally leading to reduced time for development and release of superior arabica coffee varieties to farmers.

#### **1.4 Research Questions**

- i) Is there any genetic diversity in Tanzanian and Ethiopian *Coffea arabica* germplasms maintained in Tanzania?
- ii) Is there a fast, precise and reliable alternative method to hypocotyl inoculation for arabica coffee genotypes resistance to coffee berry disease (CBD) screening at the early growth stage?
- iii) Is the genetics and relationships of growth and yield variables of Tanzanian and Ethiopian genotypes well known?

#### **1.5 Hypotheses**

- i) There is no genetic diversity in Tanzanian and Ethiopian *Coffea arabica* germplasms maintained in Tanzania.
- ii) There is no fast, precise and reliable alternative method to hypocotyl inoculation for arabica coffee genotypes resistance to coffee berry disease (CBD) screening at the early growth stage.
- iii) The genetics and relationships of growth and yield variables of Tanzanian and Ethiopian genotypes is not well known.

## CHAPTER TWO

### 2.0 Literature Review

#### 2.1 Economic Importance of Coffee

Coffee is one of Tanzania's primary agricultural export crop, representing about 5 % of total exports, 24 % of traditional cash crops and generating export earnings averaging USD 100 million per annum (ICO, 2015). More than 90 percent of Tanzanian coffee comes from smallholder farmers. The industry provides direct income to more than 400 000 farmer families and also benefits indirectly the livelihoods of over 2.5 million Tanzanians (TCB, 2012). Tanzania produces both Arabica and Robusta coffee. Tanzania produce an average of 51 000 metric tons per year ranking eighteenth in the world production and fourth in Africa after Ethiopia (406 000 tones), Uganda (204 000 tones) and Ivory Coast (110 000 tones) based on 2015 ICO statistics (ICO, 2015). Most of the coffee produced in Tanzania is exported and internal consumption is only 7 % of the amount produced (ICO, 2015).

#### 2.2 Origin and Distribution of *Coffea* Species

The coffee genus (*Coffea*) comprises 124 species, and is indigenous to the Old World Tropics (Razafinarivo *et al.*, 2012). Out of these species two species are economically important for the production of the beverage coffee: *Coffea arabica* (Arabica coffee) and *C. canephora* (Robusta coffee), and to a lesser extent, *C. liberica* (Liberian or Liberica coffee, or Excelsa coffee) (Davis *et al.*, 2006). The genus is classified into two sub-genera: sub-genus *Coffea* and subgenus *Baracoffea* (Davis *et al.*, 2006). Coffee arabica is native to Ethiopia where it is known to occur in forest, semi-forest, garden and in plantation (Sene and Admassu, 2013). It is confined to the plateau of southwestern Ethiopia and on the Boma plateau of Sudan (Anthony *et al.*, 2002; FAO, 1968).

Coffee arabica is thought to have been smuggled to Yemen from its origins in Ethiopia around the sixth century (Pendergrast, 1999). From Yemen, two genetic bases spread giving rise to most of the present commercial cultivars of Arabica coffee grown worldwide (Anthony *et al.*, 2002). The two sub-populations of wild coffee introduced from Ethiopia to Yemen underwent successive reductions in genetic diversity with the first reduction occurring with the introduction of coffee to Yemen 1500 to 300 years ago (Anthony *et al.*, 2002). In the 18<sup>th</sup> century coffee was introduced to Java, Amsterdam, and La Réunion leading to further reductions in genetic diversity (Anthony *et al.*, 2002). Coffee then spread rapidly to American countries and Indonesia from self-fertilized seeds with intense reduction to genetic diversity (Anthony *et al.*, 2002).

### **2.3 Morphology and Reproductive Biology of *Coffea Arabica***

Arabica coffee is an evergreen shrub of variable size. The tree grows up to 8-10 m high and its branches are opposite, long, flexible and thin. Branches are semi-erect when young and spreading or pendulous when old (Wrigley, 1988). It has white, Jasmine-scented flowers grouped together in the axils of the paired leaves, with two to three cymes making up whorls of 8-15 flowers. Its fruits are sub-globular, ovoid, oblong or squat-shaped and are orange-red to red on ripening (Wrigley, 1988). Bean colour can be yellowish-grey to slate-grey, bluish or grey-green, depending upon the variety, method of preparation and storage condition (Wrigley, 1988). *Coffea arabica* is a tetraploid ( $2n=4x=44$  chromosomes) and self-fertile, whereas all other *Coffea* species are diploid ( $2n=2x=22$  chromosomes) and mostly self-sterile. Arabica coffee is the only known tetraploid ( $2n = 4x = 44$ ) and self-fertile (over 95%) species in the genus (Herrera *et al.*, 2002).

### 2.3 Significance of Ethiopian Coffee Arabica Diversity

The phenotypic variation as well as cultivation under diverse environmental conditions demonstrates the presence of high *C. arabica* genetic diversity in Ethiopia (Bekele, 2005). Chaparro *et al.* (2004) studied the *Coffea arabica* germplasm collection, gathered by the FAO and ORSTOM (IRD) missions to Ethiopia, which is maintained in Colombia by CENICAFE, using RAPD markers where they found that the Ethiopian genotypes diversity differed significantly from the cultivated Caturra variety (Chaparro *et al.*, 2004). This study concluded that the Ethiopian genotypes could be used in the improvement of their cultivated variety (Chaparro *et al.*, 2004).

Moncada and McCouch (2004) when studying diversity in Colombian germplasm collection using 34 fluorescently labeled microsatellite markers found that 55 % of the alleles found in the wild tetraploids were not shared with cultivated *C. arabica* genotypes, supporting the idea that the wild tetraploid ancestors from Ethiopia could be used productively as a source of novel genetic variation to expand the gene pool of elite *C. arabica* germplasm. Aerts *et al.* (2013) genotyped 703 coffee shrubs from unmanaged and managed coffee populations, using 24 microsatellite loci. They also genotyped 90 individuals representing 23 Ethiopian cultivars with resistance to coffee berry disease (CBD). Indeed, all the arabica coffee of the world, outside of Ethiopia, has suffered major diversity bottleneck in the process of diffusion from the center of its origin (Robinson, 2007). Many characteristics of agronomic interest were observed in Ethiopian coffee, such as the incomplete resistance to orange leaf rust *Hemileia vastatrix*, resistance to the nematode *Meloidogyne incognita* and to coffee berry disease caused by *Colletotrichum kahawae* (Van der Vossen, 2001). Ethiopian arabica coffee cultivars are likely to be extremely valuable to other countries in Africa, where coffee berry disease

prevents the cultivation of coffee, particularly by the smallest and poorest farmers, who generally lack both the expertise and the money, to spray their crops (Robinson, 2007).

#### **2.4 Arabica Coffee Genetic Improvement**

The knowledge of the genetic diversity allows the choice of suitable selection strategy (Mistro *et al.*, 2007). There is great bottleneck in the genetic diversity of commercially cultivated *Coffea arabica* outside their area of origin, Ethiopia (Vega *et al.*, 2008) this is a common trend for several crops following domestication process. Efforts have been made through different approaches aimed at widening the genetic base of commercially cultivated arabica coffee through new introductions, collections and hybridization between cultivated varieties and semi wild types (Van der Vossen, 2005). Genetic diversity conserved in coffee field gene banks has been assessed using agromorphological characters and, more recently, by DNA-based genetic markers (Anthony *et al.*, 2002). Currently arabica coffee improvement programmes strive to enlarge the genetic base of their germplasms through the incorporation of disease, insect pest and nematode resistance due to growing incidences of new diseases, pathotypes and lesser known diseases in new coffee growing areas (Van der Vossen, 2005). Other areas of improvement include abiotic stress tolerance due to global climatic changes, deteriorating soil conditions and changing patterns of rainfall (Herrera *et al.*, 2002).

To meet current market demands other areas of improvement include quality, environmental stability and sustainable yield improvement (Van der Vossen, 2005; Herrera *et al.*, 2002). Through these efforts new improved varieties are continuously being developed by different institutions in the world such as IAPAR in Brazil (Sera, 2001), TaCRI in Tanzania (Teri *et al.*, 2004), CRF in Kenya (Gichimu and Omondi, 2010; Van der Vossen and Walyaro, 2009). It is important therefore, to

establish the level of variability achieved as a result of these hybridization programmes carried out in the original germplasm as this information is crucial for proper planning of development of more genetically superior varieties that meet the demands of current markets (Van der Vossen and Walyaro, 2009).

Use of molecular markers assists breeders to choose parents that, when mated, provide populations or gene pools enriched for combination of desirable traits (Aerts *et al.*, 2013). Given the complexity of quantitative traits, many different lines or crosses must be carefully analyzed over years and environments to find out important components of gene interactions (Van der Vossen, 2004; Herrera *et al.*, 2002). Tanzania is one of the countries which benefited from the FAO coffee collection mission in Ethiopia of 1964 where about 196 accessions were received and established in a field germplasm (FAO, 1968). Despite the wide publications of the genetic advantages of this germplasm in other countries it is poorly or not fully exploited in Tanzania due to poor evaluation of its potential (Bertrand *et al.*, 2014).

## **2.5 DNA Molecular Markers Application in Coffee**

Worldwide coffee genetic improvement efforts through conventional methods based on hybridization have given some encouraging results (Aggarwal *et al.*, 2007). However, although successful, they have proven to be too slow and severely constrained due to various factors. These include genetic and physiological makeup (low genetic diversity and ploidy barrier in arabica, and self-incompatibility in cross-species fertilization in Robusta), long generation cycle and requirement of huge land resources for evaluation (Manimekalai *et al.*, 2007). This situation demands new, easy, practical technologies that can provide acceleration, reliability and clear direction to breeding efforts. This allows characterization of cultivated and secondary gene pool for proper utilization of the

available germplasm in genetic improvement programmes. DNA marker tools should help to generate appropriate information for development of improved varieties to meet the new challenges represented by ongoing radical changes in the environment and market (Kochko *et al.*, 2010; Powell *et al.*, 1996). It is important to establish the level of diversity available, as this information is crucial for proper planning of crop improvement strategies for development of superior varieties.

Genetic fingerprinting has been applied in many aspects of crop biology, from analysis of genetic diversity within breeding populations in plants, to differentiation between cultivars, as well as to identification of plants containing a gene of interest. Lashermes *et al.* (1996) studied the inheritance and restriction fragment length polymorphism (RFLP) of chloroplast DNA in the genus *Coffea*. The potential of DNA marker technology is increasingly being utilized in germplasm characterization and genetic diversity analysis. In coffee random amplified polymorphism DNA (RAPD) markers have been used to identify polymorphism in different coffee accessions (Diniz *et al.*, 2005). Amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR) markers were used to assess polymorphism between and within coffee accessions where the average genetic similarity coefficient calculated using SSR markers was much lower than AFLP markers indicating the higher information content of SSR markers (Dessalegn *et al.*, 2009). The first application of micro-satellites in agriculture was in cultivar identification, whereby breeders were able to determine cultivated varieties from traditional varieties and vice versa through the use of microsatellites (Moncada and McCouch, 2004). In a phylogenetic study using AFLP and SSR, Anthony *et al.* (2002) found polymorphism among the sub-spontaneous accessions which was much higher than among the cultivated accessions. Different marker tools were also used in the analysis of foreign genome introgression and in the study of mechanism of introgression into *C. arabica* of a leaf rust

resistance gene from *C. liberica* ( $S_{H3}$ ) where twenty-one markers were strongly associated with the  $S_{H3}$  gene and were grouped together in a single linkage group (Prakash *et al.*, 2004). Restriction fragment length polymorphism (RFLP) and microsatellite loci-markers were studied in two  $BC_1$  populations where it was found that recombination in the (*C. arabica*  $\times$  *C. canephora*) hybrid is not significantly restricted by the genetic differentiation between chromosomes belonging to the different genomes (Herrera *et al.*, 2002). Currently, different molecular marker techniques are being developed for measuring genetic diversity. Dessalegn *et al.* (2009) compared AFLP and SSR markers in Ethiopian arabica coffee while Manimekalai *et al.* (2007) assessed the effectiveness of RAPD, ISSR and SSR markers for analysis of coconut germplasm accessions. The few studies carried out using a limited number of SSR markers desirably demonstrate the SSRs potential utility over other marker approaches in their ability to more efficiently detect the inherent low variability of arabica (Aggarwal *et al.*, 2007; Moncada and McCouch, 2004).

## **2.6 Microsatellite Makers**

Simple Sequence Repeats (SSRs), also known as microsatellites, are DNA sequences that are formed by the arrangement of nucleotides through the combination of one to six base pairs that are repeated in tandem (Kalia *et al.*, 2011). Among different classes of molecular markers, SSR markers are useful for a variety of applications in plant genetics and breeding because of their high efficiency, reproducibility, easy-to-use, high degree of polymorphism, multi-allelic nature, co-dominant inheritance, relative abundance and good genome coverage (Powell *et al.*, 1996). SSRs are used in characterization of genetic variation within natural populations and among breeding lines a crucial step for effective conservation and exploitation of genetic resources for crop improvement programs. SSR markers have been identified for resistance to major coffee arabica diseases.

These included microsatellite markers Sat 225, Sat 229 and Sat 259 for coffee leaf rust (*Hemileia vastatrix* Berk and Br.) Herrera *et al.* (2009) and Sat 207 & Sat 235 for coffee berry disease (*Colletotrichum kahawae*) Gichuru *et al.* (2008).

## **2.7 Coffee Berry Disease (*Colletotrichum kahawae* Waller and Bridge)**

Coffee berry disease (CBD) caused by a fungal pathogen *Colletotrichum kahawae* is a severe and widespread disease in arabica coffee growing areas in Africa and an imminent threat to Arabica coffee cultivation in Asia and America, if the pathogen enters those continents (Pires *et al.*, 2015). Genetic diversity within *C. kahawae* is low but notorious differences in pathogenicity have been described (Pires *et al.*, 2015).

### **2.7.1 Economic importance of coffee berry disease**

Coffee Berry Disease epidemics can quickly destroy 50–80 % of the developing berries (6–16 weeks after anthesis) on susceptible arabica cultivars during prolonged wet and cool weather conditions. Preventive control by frequent fungicide sprays may account for 30–40 % of total production costs. Annual economic damage to arabica coffee production in Africa, due to crop loss by CBD and cost of chemical control, is estimated at US\$ 300–500 million (Van der Vossen and Walyaro, 2009).

### **2.7.2 Symptoms and spread of coffee berry disease**

Crop losses occur upon infection of green berries, with the formation of dark sunken lesions leading to anthracnose-like symptoms that cause their premature dropping and mummification of berries that remain on the branches (Omondi, 1998). Scab lesions can also be observed when infections occur during the less susceptible stages of berry development and on resistant coffee varieties as a defense reaction (Owaka, 2011). Direct loss occurs as a result of flower and young fruit infection. Immature fruit or

berries are most susceptible during their expansion phase that occurs from 4-14 weeks after flowering (Omondi, 1998). Flower cushions, branch bark and mummified berries are considered to be the main sources of primary inoculum. New infections then occur gradually up to the fruit hardening stage, towards the 22<sup>nd</sup> week after flowering (Pires *et al.*, 2015). The pathogen is primarily dispersed by rainfall through splashing, with conidia being transported by droplets caused by the impact of raindrops on the acervuli (Waller, 1972).

### **2.7.3 Inheritance of coffee berry disease resistance**

There is still no consensus among coffee breeders on the type of genetic system, some claiming convincing evidence for oligogenes (1-3 major genes) and others for polygenes determining CBD resistance (Van der Vossen and Walyaro, 2009). Genetic resistance appears to be partial in *C. arabica* and complete in *C. canephora*. Studies on Arabica coffee carried out in Kenya concluded that coffee resistance to CBD appears to be controlled by major genes at three different loci (Van der Vossen and Walyaro, 1980). Rume Sudan, the highly resistant variety carries the dominant *R*- and the recessive *k*-genes. Pretoria variety also has the *k*-gene. The Blue Mountain (K7) variety, a moderately resistant variety carries only the recessive *k*-gene. Hibrido de Timor carries one gene for CBD resistance on the *T*- locus. Van der Graaff (1981, 1983) suggested that CBD resistance is quantitative.

### **2.7.4 DNA markers application in coffee diseases studies**

Three randomly amplified polymorphic DNA (RAPD) markers were shown to be closely associated to the *T* gene in the Hibrido de Timor in the study done to identify markers associated with CBD resistance genes by Agwanda *et al.* (1997). Gichuru *et al.* (2008) identified and mapped eight amplified fragment length polymorphisms (AFLP) and two

simple sequence repeats (SSR) markers linked tightly to the CBD resistant phenotype. In a study for SSR markers linked to CBD resistance genes in F<sub>2</sub> population involving resistant Rume Sudan and susceptible SL28, Kiguongo *et al.* (2014) identified two SSR markers, M 24 (~210bp) and Sat 227 (~200bp) which were putatively linked to two alleles in the resistant material. SSR markers have been identified for resistance to major coffee arabica diseases. These included microsatellite markers Sat 225, Sat 229 and Sat 259 for coffee leaf rust (*Hemileia vastatrix* Berk & Br.) (Herrera *et al.*, 2009) and Sat 207 & Sat 235 for coffee berry disease (*Colletotrichum kahawae*) (Gichuru *et al.*, 2008).

## **2.8 Coffee Growth and Yield Variables Interrelationships**

Knowledge of the genetic variation of a species and the estimation of genetic parameters allows the choice of suitable selection strategy reducing the time required in cultivar development and release (Mistro *et al.*, 2007).

### **2.8.1 Combining ability**

General combining ability (GCA) refers to average performance of individual lines in crosses while specific combining ability (SCA) refers to the performance of crosses relative to the average performance of parents involved in the cross (Woldegiorgis, 2003). GCA describes the average contribution of the parents with good combining ability for specific characters and may be useful in a hybridization programme for improvement of that trait (Woldegiorgis, 2003). SCA is where certain hybrid combinations do relatively better or worse than would be expected on the basis of the average of the performance of the lines involved. Specific combining ability (SCA) consists of dominance and all types of epistatic variances and is regarded as estimates of effects of non-additive gene actions (Falconer and Mackay, 1996). Significant GCA and SCA effects provide information that guides in determination of the efficiency of breeding for improvement of given traits and

they can be used to identify lines to be used as parents in a breeding program for improvement of particular crops. If the ratio of the GCA variances and SCA variances ( $\delta^2 \text{GCA} / \delta^2 \text{SCA}$ ) is equal to 1, the magnitude of additive genetic effects and non-additive effects are of equal importance. Equally true; if the ratio is far less than 1 the magnitude of non-additive effects (dominance, epistatic) is greater on the particular trait (Griffing, 1956). When the GCA variances ( $\delta^2 \text{GCA}$ ) are higher than SCA variances ( $\delta^2 \text{SCA}$ ); suggests the effect of additive gene action is controlling the particular trait and selection based on progeny performance will be effective. When SCA variances are higher than GCA variances; the effect of non-additive gene action is governing the particular trait (dominance and epistasis) hence heterosis (hybrid) breeding will be effective and if both (GCA and SCA) are equal, both are equally important and reciprocal recurrent selection may be useful. The implication of this facts in coffee breeding in the past guided the breeders to properly plan breeding approaches when they were developing arabica (Walyaro, 1980) and robusta (Marandu *et al.*, 2004).

### **2.8.2 Heritability**

Heritability expresses the proportion of total variance that is due to the average effects of genes. Heritability is measured in two ways: broad-sense heritability  $H^2$  and narrow-sense heritability  $h^2$ . Broad-sense heritability  $H^2$  measures the full contribution of genes. It is defined as  $H^2 = V_G/V_P$  where  $V_G$  is the total variance due to genes while narrow-sense heritability  $h^2 = V_A/V_P$  captures the “additive” contribution of genes to the trait (Visscher *et al.*, 2008). Calculation of narrow-sense heritability is important for predicting how a trait will respond to selection. Heritability estimates are useful for comparing the gain from selection under different experimental designs and can be used in designing optimal breeding strategies and also heritability estimates can be used to predict gain from

selection (Visscher *et al.*, 2008; Milligan *et al.*, 1990). Heritability of particular traits in coffee determines the approach to its improvement. Broad sense heritability had been used in improvement of yield in coffee arabica by Walyaro and Van der Vossen and (1979).

### **2.8.3 Correlations**

The girth of the stem at the base of the main stem is genotypically correlated with most of the other components of vigour (height and radius of canopy) and components of yield (rate of node production, berries per node, inter node length and percentage of bearing branches) (Walyaro and Vander Vossen, 1979). This was also true for coffee yield components such as number of berries per node, number of bearing primaries and number of nodes per primary (Walyaro and Vander Vossen, 1979). The highly significant correlation between the total number of berries and the main growth and yield characters indicates that early growth data can be used to measure the yield potential of the material being evaluated (Gichimu and Omondi, 2010).

The study on arabica coffee vegetative character such as stem girth, length of the longest primary and its internode length showed high positive genotypic correlation with yield (Srinivasan, 1980; 1969). Cilas *et al.* (1998) found that morphological traits such as stem diameter, plant height and number of primaries, were genetically correlated to yield and indirect prediction of yield using morphological traits in arabica coffee gave the same value for expected genetic gain as the prediction based on cumulated yield over four years. Xiu-feng *et al.* (2001) carried out a study on relationship between growth characters and yield on coffee arabica and found that plant height, average length of the longest primary and internode length of primary branch had positive and significant correlation with yield. In another study on agronomic performance of arabica coffee

genotypes; plant height was found to be positive and significantly correlated with stem girth and number of primaries (Rodrigues *et al.*, 2014). High genotypic correlation was found on the growth variables stem girth, average length of primary and canopy diameter, for arabica coffee enabling use of these variables and early years' production for development of selection index (Ermias *et al.*, 2009; Walyaro and Van Der Vossen, 1979). Positive and significant correlation was found on growth and yield characters on arabica coffee hybrid variety Ruiru when a study was carried out to assess the performance of this variety in different environments and spacings (Gichimu and Omondi, 2010). Kumar *et al.* (2013) found that in coffee arabica hybrids the coffee plant with its inherent quality characteristics could easily be identified based on its phenotypic parameters such as stem girth and internodal length. They therefore suggest these traits could be utilized as quality markers to produce coffee to the consumer's preferences. Simple correlations identify the mutual associations among the variables under study. Correlation mentions a change in one trait that is accompanied by a change in another. It measures the strength of the linear relationship between two variables. Pearson's correlation coefficient is usually signified by  $r$  and can take on the values from -1.0 to 1.0 where -1.0 is a perfect negative (inverse) correlation, 0.0 is no correlation, and 1.0 is a perfect positive correlation (Griffing, 1956).

#### **2.8.4 Path coefficient**

Path coefficient analysis provides an effective means of partitioning correlation coefficients into its components thus permit a critical examination of specific factors that produce a given correlation. This allows successful utilization of most important components in formulating an effective selection strategy (Gelalcha and Hanchinal, 2013; Dewey and Lu, 1959).

Estimation of genetic parameters of the main selection traits guides selection in breeding programmes and reduces time required in cultivar release (Mistro *et al.*, 2007). Diallel mating designs provide useful genetic information for breeding programs, such as general combining ability (GCA) and specific combining ability (SCA), heritability and interrelationships to help breeders' design appropriate breeding and selection strategies (Johnson and King, 1998). Cilas *et al.* (1998) and Walyaro and Van der Vossen (1979) applied this approach in early selection of superior arabica coffee lines.

## References

- Aerts, R., Berecha, G., Gijbels, P., Hundera, K., Van Glabeke, S., Vandepitte, K., Muys, B., Roldán-Ruiz, I. and Honnay, O. (2013). Genetic variation and risks of introgression in the wild *Coffea arabica* gene pool in south-western Ethiopian montane rainforests. *Evolutionary Applications* 6(2): 243 – 252.
- Aggarwal, R. K., Hendre, P. S., Varshney, R. K., Bhat, P. R., Krishnakumar, V. and Singh, L. (2007). Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theory Applied Genetics* 114(2): 359 – 372.
- Agwanda, C. O., Lashermes, P., Trouslot, P., Combes, M. C. and Charrier, A. (1997). Identification of RAPD markers for resistance to coffee berry disease, *Colletotrichum kahawae*, in arabica coffee. *Euphytica* 97(2): 241 – 248.
- Anthony, F., Combes, M. C., Astorga, C., Bertrand, B., Graziosi, G. and Lashermes, P. (2002). The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theory Applied Genetics* 114: 894 – 900.
- Bekele, Y. D. (2005). Estimation of the level of genetic variation among currently grown *C. arabica* genotypes in Ethiopia. Thesis for Award of PhD Degree at University of the Free State, South Africa, 197pp.
- Chaparro, A. P., Cristancho, M. A., Cortina, H. A. and Gaitán, A. L. (2004). Genetic variability of *Coffea arabica* L. accessions from Ethiopia evaluated with RAPDs. *Genetic Resources and Crop Evolution* 51(3): 291 – 297.
- Cilas, C., Bouharmont, P., Boccara, M., Eskes, A. B. Baradat, P. (1998). Prediction of genetic value for coffee production in *Coffea arabica* from a half-diallel with lines and hybrids. *Euphytica* 104: 49 – 59.

- Davis, A. P., Govaerts, R. D., Bridson, M. and Stoffelen, P. (2006). An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). *Botany Journal of Linnaean Society* 152: 465 – 512.
- Dessalegn, Y., Herselman, L and Labuschagne, M. (2009). Comparison of SSR and AFLP analysis for genetic diversity assessment of Ethiopian arabica coffee genotypes. *South African Journal of Plant and Soil* 26(2): 119 – 125.
- Dewey, D. K. and Lu, K. M. (1959). A correlation and path coefficient analysis of components of crested wheat grass seed production. *Agronomy Journal* 51: 515 – 518.
- Diniz, L. E. C., Ruas, C. F., Carvalho, V. P., Torres, F.M., Ruas, E. A., Santos, M. O., Sera, T and Ruas, P. M. (2005). Genetic Diversity Among Forty Coffee Varieties Assessed by RAPD Markers Associated with Restriction Digestion. *Brazil Archive Biology Technical* 48(4): 511 – 521.
- Ermias, H., Bayetta, B. and Singh, H. (2009). Development of selection indices based on morphological traits and average of the first two years yield at early bearing stage of *Coffea arabica* L. In: *Proceedings of International Conference on Coffee Science*, held from 14 - 19 September, 2008 at Campinas, Brazil. pp. 1568 – 1572.
- Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. Longman Publishers, New York. 464pp.
- FAO (1968). *Coffee Mission to Ethiopian 1964-1965*. Food and Agriculture Organization, Rome, Italy. 200pp.

- Gelalcha, S. and Hanchinal, R. R. (2013). Correlation and path analysis in yield and yield components in spring bread wheat (*Triticum aestivum* L.) genotypes under irrigated condition in Southern India. *African Journal of Agricultural Research* 8: 3186 – 3192.
- Gichumu, B. M. and Omondi, C.O. (2010). Morphological Characterization of five newly developed lines of arabica coffee as compared to commercial cultivars in Kenya. *International Journal of Plant Breeding Genetics* 4(4): 238 – 246.
- Gichuru, E. K., Agwanda, C. O., Combes, M. C., Mutitu, E. W., Ngugi, E. C. K., Bertrand, B. and Lashermes, P. (2008). Identification of molecular markers linked to a gene conferring resistance to coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica*. *Plant Pathology* 57(6): 1117 – 1124.
- Griffing, B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Science* 9(4): 463 – 493.
- Herrera P. J. C., Alvarado A. G., Cortina G. H. A., Combes, M. C., Romero G. G. and Lashermes, P. (2009). Genetic analysis of partial resistance to coffee leaf rust (*Hemileia vastatrix* Berk and Br.) introgressed into the cultivated *Coffea arabica* L. from the diploid *C. canephora* species. *Euphytica* 167(1): 57 – 67.
- Herrera, J. C., Combes, M. C., Anthony, F., Charrier, A. and Lashermes, P. (2002). 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*). *Theory Applied Genetics* 114(4): 661 – 668.

- ICO (2015). Sustainability of the coffee sector in Africa. In: *International Coffee Council 115<sup>th</sup> Session* 28 September -2 October 2015, Milan, Italy. 17pp.
- Johnson, G. R. and King, J. N. (1998). Analysis of Half Diallel Mating Designs I – A Practical Analysis Procedure for ANOVA Approximation. *Silvae Genetica* 47: 2 – 3.
- Kalia, R. K., Rai, M. K., Kalia, S., Singh, R. and Dhawan, A.K. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177(3): 309 – 334.
- Kiguongo, A. P. K., Omondi, C. O., Gichuru, E. K and Kasili, R. W. (2014). Analysis of simple sequence repeat markers linked to coffee berry disease resistance genes in a segregating population of arabica coffee (*Coffea arabica* L.). *International Journal of Biotechnology and Food Science* 2(8): 156 – 166.
- Kochko, A., Akaffou, S., Andrade, A. C., Campa, C., Crouzillat, D., Guyot, R., Hamon, P., Ming, R., Mueller, L. A., Poncet, V., Tranchant-Dubreuil, C. and Hamon, S. (2010). Advances in *Coffea* Genomics. *Advances in Botanical Research* 53: 23 – 63.
- Kumar, A., Ganesh, S., Basavraj, K., Mishra, M. and Jayarama, K. (2013). *Morphological Basis for Identification of Cup Quality Characteristics in F1 Hybrids Derived from Coffea arabica L. Crosses. In prospects in Bioscience: Addressing the Issues.* Springer Publishers, India. pp. 173 – 180.
- Lashermes, P., Cros, J., Combes, M. C., Trouslot, P., Anthony, F., Hamon, S.A and Charrier, I. (1996). Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus *Coffea* L. *Theory Applied Genetics* 114(93): 626 – 632.

- Manimekalai, R., Nagarajan, P. and Kumaran, P. M. (2007). Comparison of effectiveness of RAPD, ISSR and SSR markers for analysis of coconut (*Cocos nucifera* L.) germplasm accessions. *Journal of Tropical Agriculture* 18: 217 – 226.
- Milligan, S. B., Gravois, K. A., Bischoff, K. P. and Martin, F. A. (1990). Crop effects on genetic relationships among sugarcane traits. *Crop Science* 30: 927 – 931.
- Mistro, J. C., Fazuoli, L. C., Gallo, P. B., de Oliveira, A. C. B., Toma-Braghini, M and Silvarolla, M.B. (2007). Estimates of genetic parameters in Arabica coffee derived from the Timor hybrid. *Crop Breeding and Applied Biotechnology* 7: 141 – 147.
- Moncada, P. and McCouch, S. (2004). Simple sequence repeats diversity in diploid and tetraploid *Coffea* species. *Genome* 47(3): 501 – 509.
- Omondi, C. O. (1998). Genetic diversity among isolates of *Colletotrichum kahawae* causing coffee berry disease and their interactions with varieties and breeding populations of Arabica coffee. Thesis for Award of PhD Degree at *University of Nairobi, Kenya*, 170pp.
- Owaka, M. 2011. Characterization of the genetic diversity and pathogenicity of *Colletotrichum kahawae*, using random amplified polymorphic DNA (RAPD) analysis. Dissertation for Award of MSc Degree at *Kenyatta University, Nairobi, Kenya*, 63pp.
- Pendergrast, M. (1999). *Uncommon Grounds: The History of Coffee and How It Transformed Our World*. Basic Books, New York. 520pp.

- Pires, A. S., Azinheira, H. G., Cabral, A., Tavares, S., Tavares, D., Castro, M., Várzea, V., Silva, M. C., Abranches, R., Loureiro, J. and P. Talhinhos, P. (2015). Cytogenomic characterization of *Colletotrichum kahawae*, the causal agent of coffee berry disease, reveals diversity in minichromosome profiles and genome size expansion. *Plant Pathology* 2: 1 – 10.
- Prakash, N. S., Marques, D. V., Varzea, V. M. P., Silva, M. C., Combes, M. C. and Lashermes, P. (2004). Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. Arabica* L. *Theory Applied Genetics* 109: 1311 – 1317.
- Powell, W., Machray, G. C. and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends Plant Science* 1(7): 215 – 222.
- Razafinarivo, N. J., Guyot, R., Davis, A. P., Couturon, E., Hamon, S., Crouzillat, D., Rigoreau, M., Dubreuil-Tranchant, C., Poncet, V., De Kochko, A., Rakotomalala, J.J and Hamon, P. (2012). Genetic structure and diversity of coffee (*Coffea*) across Africa and the Indian Ocean islands revealed using microsatellites. *Annals of Botany* 111(2): 229 – 248.
- Robinson, R. A. (2007). The Loss of Resistance in Coffee. In: Return to Resistance. Sharebooks Publishing Canada, 3<sup>rd</sup> Ed, pp. 231-252.
- Rodrigues, W. P., Vieira, H. D., Barbosa, D. H. S. G., Filho, G. R. S. and Partelli, F. L. (2014). Agronomic performance of arabica coffee genotypes in northwest Rio de Janeiro State. *Genetics and Molecular Research* 13(3): 5664 – 5673.
- Sene, K. H. and Admassu, B. (2013). Assessment of Microsatellite markers: An attempt towards documenting genetic diversity of Ethiopian coffee. *Insight Genetics* 1(1): 1 – 6.

- Sera, T. (2001). Coffee Genetic Breeding at IAPAR. *Crop Breeding and Applied Biotechnology* 1(2): 179 – 199.
- Sheoran, O.P., Tonk, D.S., Kaushik, L.S., Hasija, R.C and Pannu, R.S (1998). Statistical Software Package for Agricultural Research Workers. Recent Advances in information theory, Statistics and Computer Applications by D.S. Hooda and R.C. Hasija. Department of Mathematics Statistics, CCS HAU, Hisar (139-143).
- Srinivasan, C. S. (1980). Association of some vegetative characters with initial fruit yield in coffee (*Coffea arabica* L.). *Journal of Coffee Research* 10(2): 21 – 27.
- Srinivasan, C. S. (1969). Correlation studies in coffee, 1: Preliminary studies on correlation between stem girth and ripe cherry yield in some coffee selection. *Indian Coffee* 33(10): 318 – 319.
- TCB (2012). *Tanzania Coffee Industry Development Strategy 2011/2021*. Tanzania Coffee Board, Moshi, Tanzania. 64pp.
- Teri, J. M., Kilambo, D. L., Mtenga, D. J., Nyange, N. E., Nzallawahe, T. S., Chipungahelo, G. S., Kipokola, T. P., Kullaya, I. K. (2004). Improved Arabica varieties for the benefit of Tanzania coffee growers. In: *20<sup>th</sup> International Conference on Coffee Science* 11-15 October 2004 Bangalore, India. pp. 1187 – 1191.
- Van der Graaff, N. A. and Pieters, R. (1983). Resistance to coffee berry disease in Ethiopia. In: *Durable Resistance in Crops. NATO Advanced Science Institutes Series* 55: 317 – 333.

- Van der Graaff, N. A. (1981). Selection of Arabica coffee types resistant to Coffee Berry Disease in Ethiopia. Thesis for Award of PhD Degree at University of Wageningen, Netherlands, 110pp.
- Van Der Vossen, H. Bertrand, B. and Charrier, A. and (2015). Next generation Variety development for sustainable production of arabica coffee (*Coffea Arabica* L.): A review. *Euphytica* 204(2): 243 – 256.
- Van der Vossen, H. A. M. and Walyaro, D. J. (2009). Additional evidence for oligogenic inheritance of durable host resistance to coffee berry disease (*Colletotrichum kahawae*) in arabica coffee (*Coffea arabica* L.). *Euphytica* 165(1): 105 – 111.
- Van der Vossen, H. A. M. (2005). A critical analysis of the agronomic and economic sustainability of organic coffee production. *Experimental Agriculture* 41: 449 – 473.
- Van der Vossen, H. A. M. (2001). Agronomy I: *Coffee breeding Practices*. In: (Edited by Clake, R. J. and Vitzthum, O. G.), *Coffee Recent Development*, Blackwell Science Ltd., London. pp. 184 – 201.
- Van der Vossen, H. A. M. and Walyaro, D. J. (1980). Breeding for resistance to Coffee Berry Disease in *Coffea arabica* L. II. Inheritance of resistance. *Euphytica* 29: 777 – 791.
- Vega, F. E., Ebert, A. W. and Ming, R. (2008). Coffee germplasm resources, Genomics and Breeding. *Plant Breeding Reviews* 30: 415 – 447.
- Visscher, P. M., Hill, W. G. and Wray, N. R. (2008). Heritability in the genomics era - concepts and misconceptions. *Nature Reviews* 9: 255 – 266.

- Waller, J. M. (1972). Water-borne spore dispersal in coffee berry disease and its relation to control. *Annals of Applied Biology* 71(1): 1 – 18.
- Woldegiorgis, T. D. (2003). Genetic variability and combining ability for quality in Ethiopian wheat cultivars. Thesis for Award of PhD Degree at University of Free State, Bloemfontein, South Africa, 137pp.
- Walyaro, D. J. and Van der Vossen, H. A. M. (1979). Early determination of yield potential in arabica coffee by applying index Selection. *Euphytica* 28: 465 – 472.
- Wrigley, G. (1988). *Coffee. Harlow*: Longman Scientific and Technical, London. 639pp.
- Xiu-feng, L., Zhong-liang, P., Shi-gui, Y., Gui-hua, L. and Yu-cang, S. (2001). Relationship between characters and yield of *Coffea arabica*. *Journal of Yunnan Tropical Crops Science and Technology* 3: 14 – 19.

## CHAPTER THREE

### **Determination of genetic diversity of Tanzanian germplasm and the Ethiopian *Coffea arabica* collection maintained in Tanzania using simple sequence repeats (SSR) molecular markers**

#### **3.1 Abstract**

Previous studies have revealed an extremely reduced genetic diversity in cultivated *Coffea arabica* L. The process of diffusion of Coffee arabica and the selection that followed have strongly reduced the genetic diversity present in its area of origin Ethiopia. Diversity was reduced during domestication and the cycles of breeding and selection, due to the homogenization of genetic structures, favoured by the predominant autogamy of the species *C. arabica*. The aim of the current study was to establish the genetic diversity of the Tanzania coffee research institute breeding pool of arabica coffee including the Ethiopian collection maintained at the institute. Ninety-one coffee arabica genotypes were selected from Ethiopian collection, germplasms and from breeding fields at TaCRI Lyamungu. Leaf samples were collected from three trees per genotype for DNA extraction. DNA was extracted following the CTAB method at Sokoine University of Agriculture molecular biology laboratory. Thirty primer sets were used in PCR amplification. The DNA reactions with the 30 SSR primer sets were performed using the touchdown PCR procedure. The reproducibility of the amplification products was checked two times for each primer. Those fragments that were too difficult to score with certainty were excluded from the data analysis. Data analysis was performed using GenStat statistical software. Three significantly different groups were produced from principal component analysis with variation percentages 56.75, 5.69, 4.66 respectively. High genetic diversity was revealed in this study hence the studied materials could be used to improve the current superior coffee varieties and in development of new ones.

### 3.2 Introduction

Genetic diversity of Tanzanian cultivated arabica coffee is limited (Masumbuko and Bryngelsson, 2006; Masumbuko *et al.*, 2003). The limited genetic diversity detected among arabica coffee cultivars is mostly the consequence of its few introductions and its self-pollinating nature. The lack of genetic diversity in the gene pool of arabica coffee limits the potential for germplasm improvement (Lashermes *et al.*, 2000). Efforts have been made through different approaches aimed at widening the genetic base of commercially cultivated arabica coffee through new introductions, collections and hybridization between cultivated varieties and semi wild types. Tanzania is one of the countries that benefited from the FAO coffee collection mission to Ethiopia of 1964 where 196 accessions were received and established in a field germplasm (FAO, 1968). Despite the wide publications of the genetic advantages of this germplasm in other countries, it is has not fully been studied and exploited in Tanzania due to poor evaluation of its potential (Bertrand *et al.*, 2014). Genetic diversity of coffee can be assessed using different techniques including morphological, biochemical and DNA-based molecular markers. Molecular marker techniques have advantages over morphological approaches, which include: they are not subjected to environmental factors or growth stage of the plant, and they exist in abundance, covering the entire genomes. Micro-satellite (SSRs) markers are short tandem repeats of DNA sequence of one to six base pairs. They are highly preferred molecular marker techniques because they are highly polymorphic, reproducible, locus specific and “co-dominant”. SSRs have very high potential utility over other marker approaches due to their ability to more efficiently detect the inherent low variability of arabica (Aggarwal *et al.*, 2007; Moncada and McCouch, 2004).

The objective of this study was to establish the level of genetic diversity of coffee in the currently field maintained germplasm from Ethiopia and other germplasm at the Tanzania Coffee Research Institute using SSR markers to generate information that will facilitate full utilization of the germplasms.

### **3.3 Materials and Methods**

#### **3.3.1 Plant materials**

Ninety one (91) coffee arabica genotypes were selected from Ethiopian collection maintained at Tanzania Coffee Research Institute (TaCRI) and from other germplasms and breeding fields at TaCRI Lyamungu. Forty five accessions were collected from Ethiopian while the rest were obtained from germplasm and breeding fields at Lyamungu (Appendix 1).

#### **3.3.2 Methods**

Coffee leaf samples were collected from three trees per genotype for DNA extraction. Young coffee leaves were picked from the growing tips and lyophilized for 72 hours before DNA extraction.

##### **3.3.2.1 DNA extraction**

The lyophilized coffee leaves were stored at -21°C before DNA extraction. Genomic DNA was then extracted from the leaves using the CTAB method (Diniz *et al.*, 2005) with minor modifications on annealing temperatures. About 500 mg of the lyophilized leaves were ground in eppendorf tubes using a mortar and pestle. 1 ml each of lysis and extraction buffers were added to the powder in the eppendorf tubes. The ground tissue was distributed in two 1.5 ml tubes and incubated at 65°C in a water bath for 20 to 30 min with regular shaking. After incubation, 1 ml of chloroform/isoamyl-alcohol mixture in the

ratio of 24:1 was added to each tube, then mixed gently by shaking and centrifuged at 13 000 rpm for 10 to 15 min in a micro-centrifuge. The supernatants were pipetted out into new 1.5 ml tubes. 20  $\mu$ l of RNase were added to the supernatants and incubated at 37°C in a water-bath for 30 min. An equal volume of isopropyl alcohol was added into each tube and mixed gently by inverting the tubes several times to precipitate DNA. The suspended DNA was centrifuged at 13 000 rpm for 5 min to obtain a DNA pellet and the supernatant carefully removed. The DNA pellets were then washed with 200  $\mu$ l of 70 % ethanol and centrifuged at 13 000 rpm for 3 min. The ethanol was drained by decanting or micro-pipetting and the pellets dried in a vacuum centrifuge for 20 min. The pellets were dissolved in 60  $\mu$ l of Tris-EDTA (TE) buffer and stored at 4°C.

DNA quality and quantity were quantified on 1% agarose gel in 0.5xTAE (Tris-Acetate-EDTA) buffer, stained with ethidium bromide and visualized under UV light. The PCR reactions for the SSR markers were made in 20  $\mu$ L containing 2.0  $\mu$ L of 10x buffer, 150 mM/L of dNTP, 0.1 mM /L of each primer, 25 ng of DNA, 1 mM / L  $MgCl_2$ , 1 U Taq DNA polymerase and the remaining volume was completed with PCR water. The reaction consisted of initial denaturation at 94°C/10 min, followed by 10 cycles of denaturation at 94°C/ 30 s, annealing at 50°C/30 s, and extension at 72°C/1 min. followed by a final extension at 72°C/7 min. All amplified products were confirmed by 2.0% (wt/vol) i.e. (low electroendosmosis) agarose gel electrophoresis in 1 $\times$  TAE (Tris base, acetic acid, and EDTA) buffer at 120V/115mA/15W for 2.30 hours with ethidium bromide post-staining (0.5  $\mu$ l/ml) for 30 min, followed by 10 min of destaining in 1 $\times$  TAE buffer and visualized under UV light. The reproducibility of the amplification products was checked twice for each primer.

### 3.3.2.2 DNA amplification

The amplification reactions with the 30 SSR primer sets (Appendix 2) sourced from literature (Teressa *et al.*, 2010; Chaparro *et al.*, 2004; Baruah *et al.*, 2003; Combes *et al.*, 2000; Rovelli *et al.*, 2000), were performed on a GeneAmp PCR System 9700 thermocycler (Applied BioSystems, USA) using the touch-down PCR procedure (Korbie and Mattick, 2008) as applied in coffee by other authors (Sene and Admassu, 2013; Poncet *et al.*, 2004). The PCR reactions for the SSR markers were made in 20  $\mu$ L containing 2.0  $\mu$ L of 10x buffer, 150 mM/L of dNTP, 0.1 mM /L of each primer, 25 ng of DNA, 1 mM / L MgCl<sub>2</sub>, 1 Unit of Taq DNA polymerase and the remaining volume was completed with PCR water. The reaction consisted of initial denaturation at 94°C/10 min, followed by 10 cycles of denaturation at 94°C/ 30 s, annealing at 50°C/30 s, and extension at 72°C/1 min. followed by a final extension at 72°C/7 min. All amplified products were confirmed by 2.0% (wt/vol) i.e (low electroendosmosis) agarose gel electrophoresis in 1 $\times$  TAE (Tris base, acetic acid, and EDTA) buffer at 120V/115mA/15W for 2.30 hours with ethidium bromide post-staining (0.5  $\mu$ l/ml) for 30 min, followed by 10 min of de-staining in 1 $\times$  TAE buffer and visualized under UV light. The reproducibility of the amplification products was checked twice for each primer.

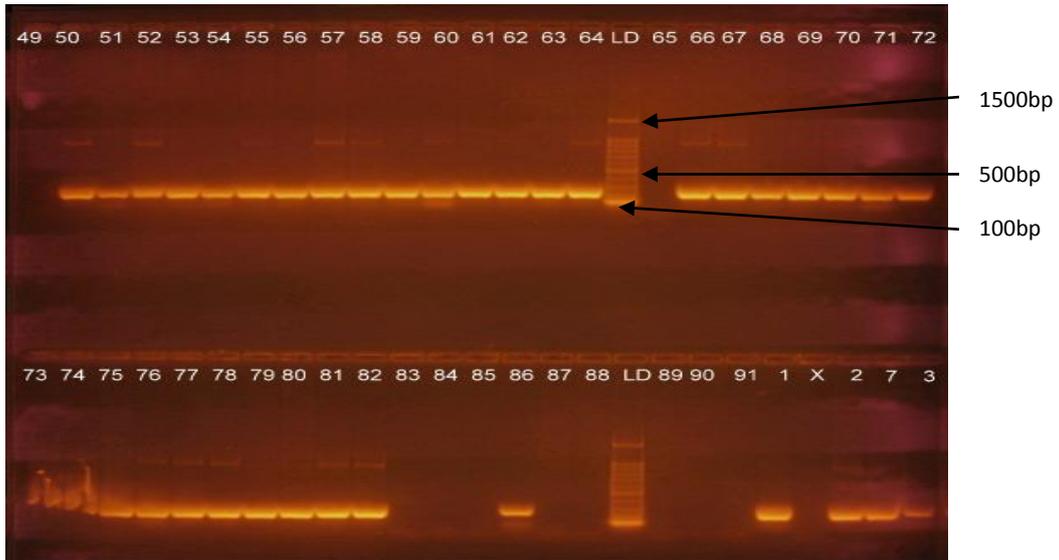
### 3.3.2.3 Scoring of bands

A one hundred base pair DNA ladder was used as a molecular weight marker. Those fragments that were not clear and difficult to read were not included in the analysis. Data analysis was performed using GenStat Release 15.1 (2012) by VSN International Ltd.

### **3.4 Results**

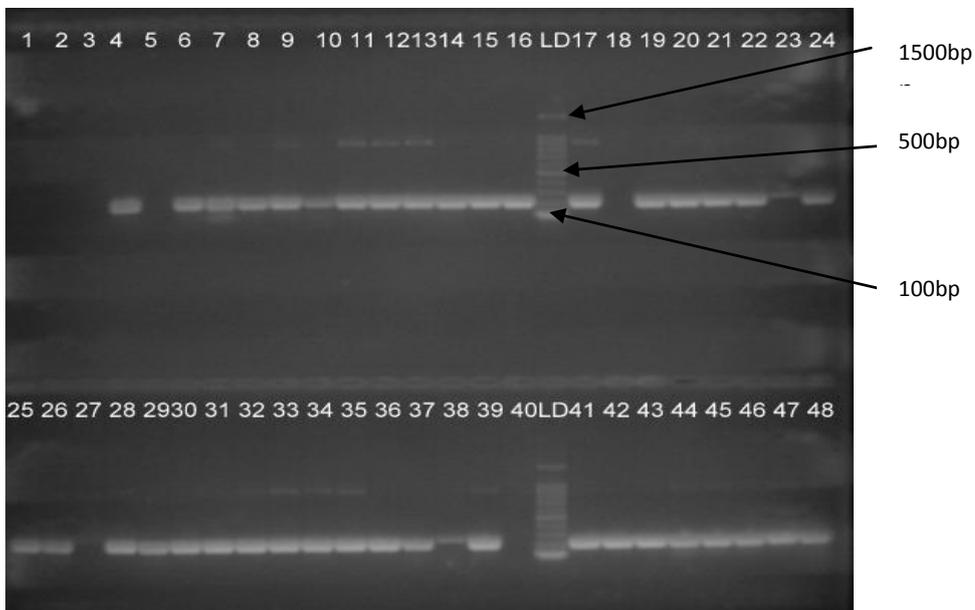
#### **3.4.1 DNA amplification**

Some of the results of genotype-allele amplifications by different microsatellite markers are presented in Figs. 1, 2 and 3. In Figure 1, microsatellite SSR ZAP25 could not amplify genotype numbers 49, 65, 83, 84, 85, 87, 88, 89, 90, 91 from the list (Appendix 1). The amplification of coffee genotype numbers 73 and 74 produced poor bands which could not be scored. Control genotypes of known amplification numbers 1, X, 2, 7 and 3 were included as negative or positive checks (Appendix 1). Genotype numbers 1, 2, 7 were amplified while 3 produced poor band and x was not amplified. In figure 2, using the same microsatellite SSR ZAP25, coffee genotypes 1, 2, 3, 5, 18, 27 and 40 were not amplified while genotypes 23 and 38 produced poor bands (Fig. 2). Microsatellite marker SSR R175 did not produce amplification on coffee genotype numbers 49, 83, 84, 85, 87 and 89, 90 and 91 from the list (Appendix 1). The same microsatellite amplified control numbers 1, 2, and 3 while control numbers x and 7 were not amplified indicating that this microsatellite was polymorphic.



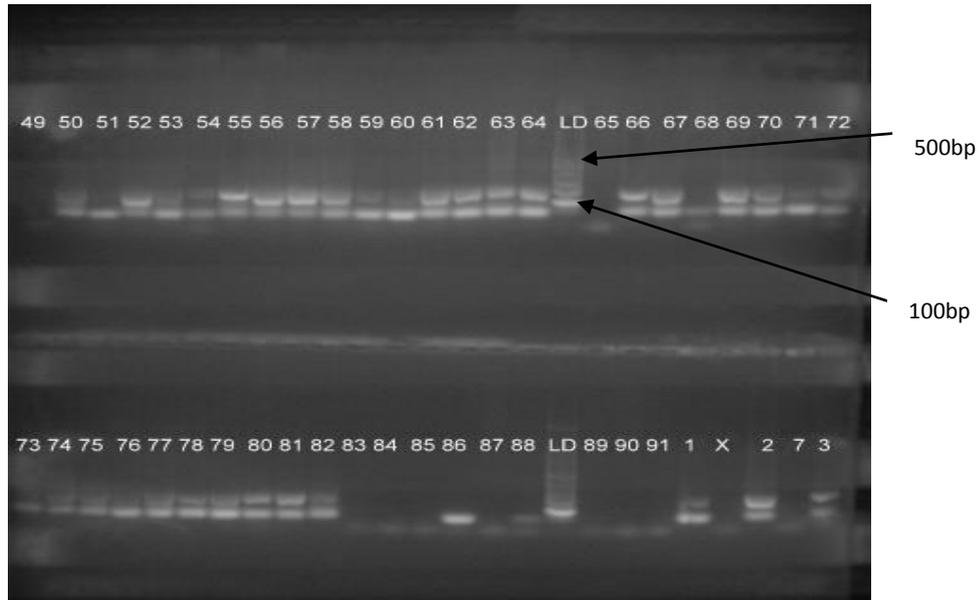
**Figure 1: A 2% gel photograph used in scoring coffee genotypes accessions 49 - 91 using SSR marker ZAP25.**

Lanes 49-91 represent the coffee genotypes, LD represents a 100-bp ladder used to estimate the amplification range of individual coffee genotypes. Lane 1, X, 2, 7 and 3 represent control samples from known genotypes with established amplification ranges.



**Figure 2: A 2 % gel photograph used in scoring coffee genotypes accessions 1- 48 using SSR marker ZAP25.**

Lanes 1 – 48 represent the coffee genotypes and LD represents a 100bp ladder used to estimate the amplification range of individual genotypes.



**Figure 3: A 2 % gel photograph used in scoring coffee genotypes accessions 49-91 using SSR marker R175.**

Lanes 49-91 represent the coffee genotypes, LD represents a 100-bp ladder used to estimate the amplification range of individual coffee genotypes. Lane 1, X, 2, 7 and 3 represent control samples from known coffee genotypes with established amplification ranges.

Results indicates that six SSR markers (CMA055, CMA059, R209, R268, CM08 and CM16) were monomorphic while 24 SSR markers were polymorphic (80 %) and they gave good amplification of the coffee genotypes. The monomorphic markers were not used in the analysis. The results (Table 1) revealed that the total number of alleles was 60 with average of 2.5 alleles per SSR marker. Alleles per marker ranged from 2 to 8 (Table 2) with the highest number generated by the marker *ssrAY2449*. Gene diversity ranged from 0.4999 to 0.7753 with mean of 0.5271 (Table 2). The marker which gave the highest coffee gene diversity was *ssrAY2449* with the value of 0.7753 followed by *ssrCM2*, *ssrCM11* and *ssrR175* with gene diversity of 0.6352, 0.6170 and 0.6156 respectively.

The marker that gave the lowest gene diversity value was *ssrA8837* which gave coffee gene diversity of 0.4999. The observed heterozygosity of the coffee materials studied was generally very high with a mean of 0.999 and a range of 0.9841 to 1.000. The marker with the lower heterozygosity level (0.9841) was *ssrA8837* (Table 2). The polymorphic information content generated by the SSR markers for the analyzed coffee genotypes ranged from 0.3749 to 0.7390 with a mean of 0.4128. The marker that generated the highest polymorphic information was *ssrAY2449* followed by *ssrCM2*, *ssrCM11* and *ssrR175* with polymorphic values 0.739, 0.5723, 0.5436 and 0.5425, respectively.

**Table 1: Summary amplification of 91 coffee genotypes by 24 SSR markers**

Character analyzed	
Total alleles produced	60
Average number of alleles per SSRs locus	2.5
Percent of polymorphic alleles (rP)	80.0
Observed heterozygosity (HO),	0.9993
Gene diversity	0.5271
Polymorphic information content (PIC)	0.4128

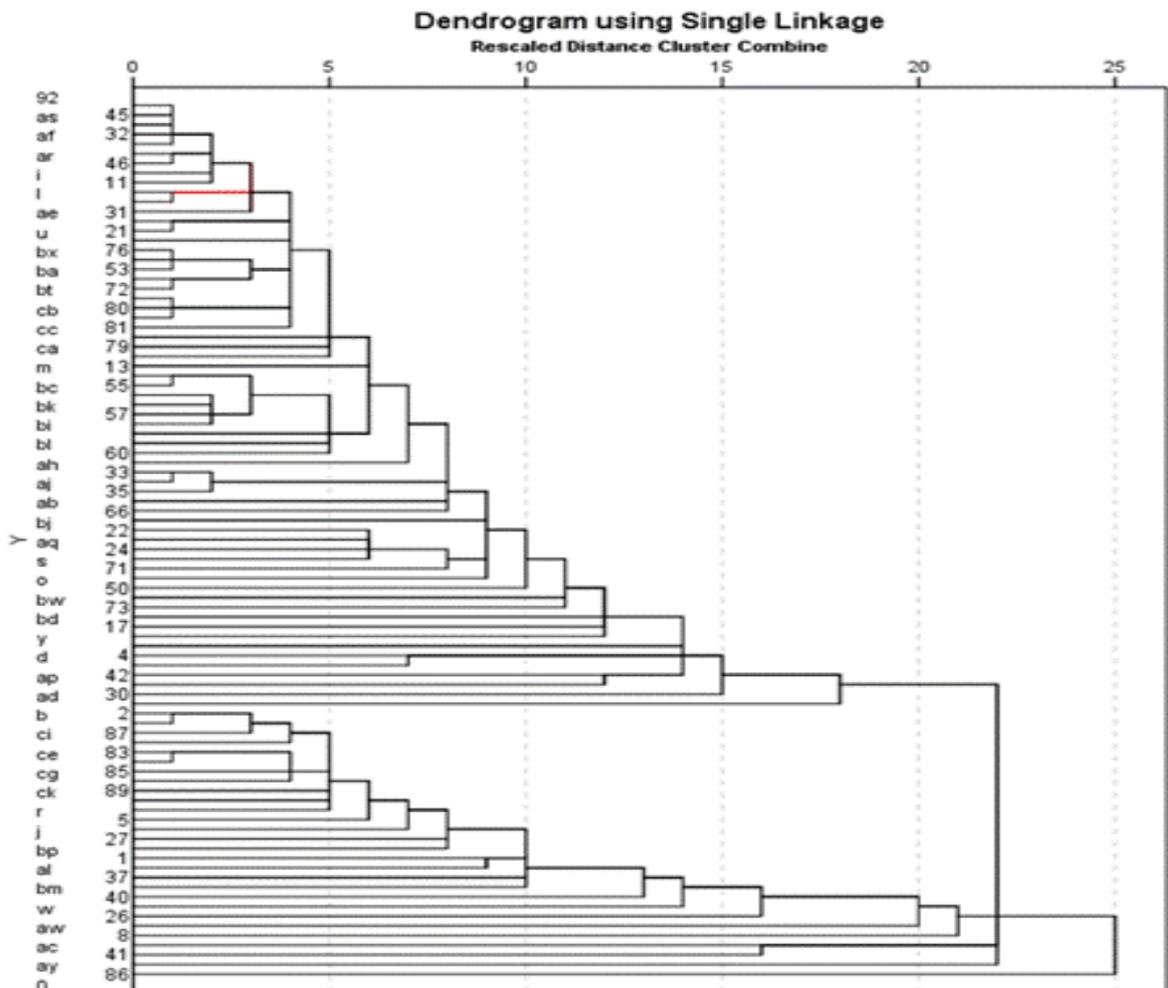
**Table 2: Amplification of 91 coffee genotypes by 24 SSR markers used in the study**

Marker name	Marker ID	Genotype No	Sample Size	No. of obs.	Allele No	Availability	Gene Diversity	Heterozygosity	PIC
ssrA8783	1	1.0000	91.0000	53.0000	2.0000	0.5824	0.5000	1.0000	0.3750
ssrA8837	2	2.0000	91.0000	63.0000	2.0000	0.6923	0.4999	0.9841	0.3749
ssrA8847	3	1.0000	91.0000	65.0000	2.0000	0.7143	0.5000	1.0000	0.3750
ssr AY2434	4	2.0000	91.0000	67.0000	3.0000	0.7363	0.5074	1.0000	0.3860
ssrAY2449	5	6.0000	91.0000	68.0000	8.0000	0.7473	0.7753	1.0000	0.7390
ssrZAP25	6	1.0000	91.0000	71.0000	2.0000	0.7802	0.5000	1.0000	0.3750
ssrCMA008	7	1.0000	91.0000	67.0000	2.0000	0.7363	0.5000	1.0000	0.3750
ssrCMA151	10	1.0000	91.0000	72.0000	2.0000	0.7912	0.5000	1.0000	0.3750
ssrCMA198	11	1.0000	91.0000	71.0000	2.0000	0.7802	0.5000	1.0000	0.3750
ssrCMA199	12	1.0000	91.0000	68.0000	2.0000	0.7473	0.5000	1.0000	0.3750
ssrCMA233	13	1.0000	91.0000	63.0000	2.0000	0.6923	0.5000	1.0000	0.3750
ssrCMA263	14	1.0000	91.0000	48.0000	2.0000	0.5275	0.5000	1.0000	0.3750
ssrR105	15	1.0000	91.0000	47.0000	2.0000	0.5165	0.5000	1.0000	0.3750
ssrR126	16	1.0000	91.0000	58.0000	2.0000	0.6374	0.5000	1.0000	0.3750
ssrR175	17	3.0000	91.0000	61.0000	4.0000	0.6703	0.6156	1.0000	0.5425
ssrR278	20	1.0000	91.0000	43.0000	2.0000	0.4725	0.5000	1.0000	0.3750
ssrR325	21	1.0000	91.0000	63.0000	2.0000	0.6923	0.5000	1.0000	0.3750
ssrR338	22	1.0000	91.0000	29.0000	2.0000	0.3187	0.5000	1.0000	0.3750
ssrR339	23	1.0000	91.0000	70.0000	2.0000	0.7692	0.5000	1.0000	0.3750
ssrCM2	24	3.0000	91.0000	28.0000	4.0000	0.3077	0.6352	1.0000	0.5723
ssrCM5	25	1.0000	91.0000	51.0000	2.0000	0.5604	0.5000	1.0000	0.3750
ssrCM6	26	1.0000	91.0000	7.0000	2.0000	0.0769	0.5000	1.0000	0.3750
ssrCM11	28	2.0000	91.0000	67.0000	3.0000	0.7363	0.6170	1.0000	0.5436
ssrCM17	30	1.0000	91.0000	62.0000	2.0000	0.6813	0.5000	1.0000	0.3750
	<b>Mean</b>	<b>1.5000</b>	<b>91.0000</b>	<b>56.7500</b>	<b>2.5000</b>	<b>0.6236</b>	<b>0.5271</b>	<b>0.9993</b>	<b>0.4128</b>
	Range	1.0-6.0		7.0-72.0	2.0 – 8.0	0.31- 0.79	0.499-0.775	0.98-1.00	0.37-0.74

NB: Markers with poor performance have been excluded ssrCMA055 (8), ssrCMA059 (9), ssrR209 (18), ssrR268 (19), ssrCM8 (27), ssrCM16 (29).

### 3.4.2 Dendrogram analysis

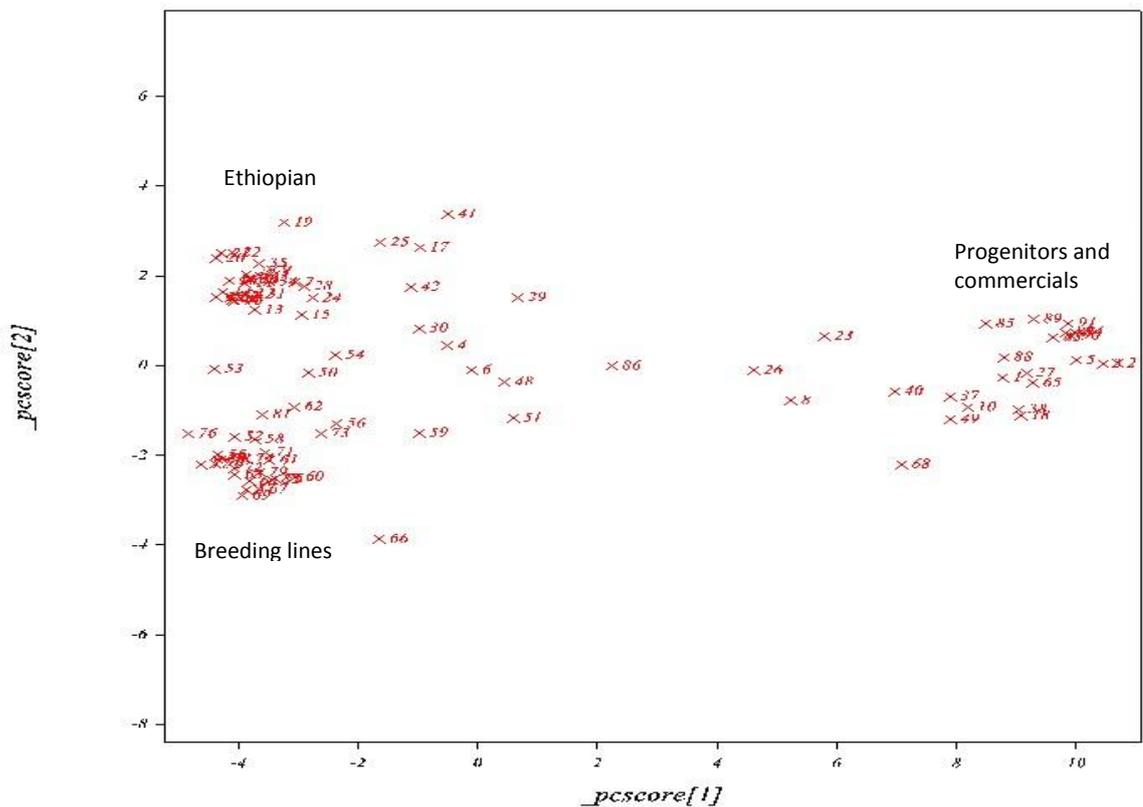
Single linkage dendrogram of the 91 genotypes studied (Fig. 4) using 30 SSR generated three main clusters. The first cluster was composed of the Ethiopian coffee genotypes while the second consisted of mainly the improved coffee breeding lines. The last cluster comprised the mixture of Ethiopian coffee genotypes, breeding lines, progenitors and commercial varieties



**Figure 4: Dendrogram of 91 genotypes amplification by 30 SSR markers based on genetic distance**

### 3.4.3 Principal component analysis

In Principal component analysis of the general diversity of the coffee materials studied, three significantly different groups were formed with the following diversity percentages 56.75, 5.69, 4.66 (Fig. 5). These three groups explained 67.11 % of the coffee diversity. Group one was comprised of the Ethiopian coffee genotypes, while the second group was composed of the improved coffee breeding lines. The third group included mixtures of Ethiopian coffee types, progenitors and improved coffee lines. Some genotypes that belonged to various coffee groups are widely scattered across the quadrants (Fig. 5).



**Figure 5: Principal component analysis showing the main diversity groups.**

The numbers in the figure correspond to the coffee accessions as listed in Appendix 1.

### 3.5 Discussion

Microsatellites markers (SSR) showed good efficiency in assessing genetic diversity and relationship among coffee genotypes (Teressa *et al.*, 2010; Aga; 2005; Moncada and McCouch, 2004). Masumbuko and Bryngelsson (2006) and Masumbuko *et al.* (2003) found a narrow genetic base in the Tanzanian cultivated Arabica coffee when using random amplified polymorphic DNA (RAPD) and inter simple sequence repeat (ISSR) molecular markers. Even though, the overall genetic diversity of *Coffea arabica* is believed to be low, the populations in Ethiopia its place of origin and diversity, have a lot of genetic variability the fact that has been supported by many studies based on different techniques (Teressa *et al.*, 2010).

#### 3.5.1 Principal component analysis

From current study, the principal component analysis (PCA) revealed three main groups. These groups were the Ethiopian coffee genotypes, breeding lines and commercial genotypes (Fig. 5). Some individuals within the groups did not cluster closely together indicating that genetic variation exists within each group. Similar findings were reported in previous studies (Teressa *et al.*, 2010; Hue, 2005). The three main components explained most of the diversity in the studied coffee genotypes amounting to 67.11 % (Fig. 5). The unexplained variation probably suggests that more SSR markers were needed to sufficiently assess genetic diversity of the studied coffee genotypes. Masumbuko and Bryngelsson (2006) when using ISSR in studying diversity of Tanzanian cultivated arabica coffee PCA managed to explain 43 % of the coffee genetic variation.

### **3.5.2 Dendrogram analysis**

Dendrogram analysis revealed two main groups and the third one was not very clear. This implied that PCA was more informative compared to the dendrogram in this study and these coffee arabica genotypes. Similarities were observed between these two analysis genotype 86 a breeding line was more diverse compared with the others as revealed in both analyses. However both the PCA and the dendrogram revealed that the Ethiopian genotypes were the most widely scattered implying that they were the most diverse. Previous studies by Teresa *et al.* (2010) found similar results when studying the genetic diversity of the Ethiopian arabica coffee genotypes.

### **3.6 Conclusion**

The present study showed high heterozygosity among the coffee genotypes studied. Number of alleles identified was high ranging from 2-8 with an average of 2.5 per marker. Mean polymorphic information content was moderate while Ethiopian genotypes were the most variable group. The commercial varieties Bourbon (1) and Kent (2) were grouped in the same group in the PCA (Fig. 5) and clustered in the same cluster in the dendrogram (Fig. 4) revealed low diversity among themselves. On the other hand the Ethiopian genotypes example (4, 6, 8) and breeding lines example (65, 66, 86) were scattered in across the groupings revealed their high heterozygosity. This information will facilitate selection of superior genotypes for development of superior varieties with broad genetic base for various traits. Marker *ssrAY2449* was the most informative.

## References

- Aga, E. (2005). Molecular Genetic Diversity Study of Forest Coffee Tree (*Coffea arabica* L.) Populations in Ethiopia: Implications for conservation and breeding. Alnarp: *Acta Universitatis Agriculturae Sueciae* 79: 1652 – 6880.
- Aggarwal, R. K., Hendre, P. S., Varshney, R. K., Bhat, P. R., Krishnakumar, V and Singh, L. (2007). Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theory Applied Genetics* 114(2): 359 – 372.
- Baruah, A., Naik, V., Hendre, P. S., Rajkumar, R., Rajendrakumar, P. and Aggarwal, R. K. (2003). Isolation and characterization of nine microsatellite markers from *Coffea arabica* L., showing wide cross-species amplifications. *Molecular Ecology Notes* 3: 647 – 650.
- Chaparro, A. P., Cristancho, M. A., Cortina, H. A. and Gaitán, A. L. (2004). Genetic variability of *Coffea arabica* L. accessions from Ethiopia evaluated with RAPDs. *Genetic Resources and Crop Evolution* 51(3): 291 – 297.
- Combes, M. C., Andrzejewski, S., Anthony, F., Bertrand, B., Rovelli, P., Graziosi, G. and Lashermes, P. (2000). Characterization of microsatellites loci in *Coffea arabica* and related coffee species. *Molecular Ecology* 9: 1171 – 1193.
- Diniz, L. E. C., Ruas, C. F., Carvalho, V. P., Torres, F.M., Ruas, E. A., Santos, M. O., Sera, T. and Ruas, P. M. (2005). Genetic diversity among forty coffee varieties assessed by RAPD markers associated with restriction digestion. *Brazil Archieve Biology Technical* 48(4): 511 – 521.
- FAO (1968). *Coffee Mission to Ethiopian 1964-1965*. Food and Agriculture Organization, Rome, Italy. 200pp.

- Hue, T. T. M. (2005). Genetic variation in cultivated coffee (*Coffea arabica* L.) Accessions in Northern New South Wales, Australia. Dissertation for Award of MSc Degree at South Cross University, 117pp.
- Korbie, D. J. and Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols* 3(9): 1452 – 1456.
- Lashermes, P., Andrzejewski, S., Bertrand, B., Combes, M. C., Dussert, S., Graziosi, G., Trouslot, P. and Anthony, F. (2000). Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). *Theory Applied Genetics* 114(100): 139 – 146.
- Masumbuko, L. I. and Bryngelsson, T. (2006). Inter Simple Sequence Repeat (ISSR) Analysis of Diploid Coffee Species and Cultivated *Coffea arabica* L. from Tanzania. *Genetic Resources and Crop Evolution* 53(2): 357 – 366.
- Masumbuko, L. I., Bryngelsson, T., Mneney, E. E. and Salomon, B. (2003). Genetic diversity in Tanzanian Arabica coffee using random amplified polymorphic DNA markers. *Hereditas* 139: 56 – 63.
- Moncada, P. and McCouch, S. (2004). Simple sequence repeats diversity in diploid and tetraploid *Coffea* species. *Genome* 47(3): 501 – 509.
- Poncet, V., Hamon, P., Minier, J., Carasco, C., Hamon, S. and Noirot, M. (2004). SSR cross-amplification and variation within coffee trees (*Coffea* spp.). *Genome* 47: 1071 – 1081.
- Rovelli, P., Mettullo, R., Anthony, F., Anzueto, F., Lashermes, P. and Graziosi, G. 2000. Microsatellites in *Coffea Arabica* L. In: *Proceedings of the 3<sup>rd</sup> International Seminar on Biotechnology in the Coffee Agro-Industry*, Londrina, Brazil. pp. 123 – 133.

- Sene, K. H. and Admassu, B. (2013). Assessment of Microsatellite markers: An attempt towards documenting genetic diversity of Ethiopian coffee. *Insight Genetics* 1(1): 1 – 6.
- Teressa, A., Crouzillat, D., Petiard, V and Brouhan. P. (2010). Genetic diversity of Arabica coffee (*Coffea arabica* L.) collections. *Ethiopian Journal of Applied Sciences and Technology* 1(1): 63 – 79.
- Van Der Vossen, H. Bertrand, B. and Charrier, A. and (2015). Next generation Variety development for sustainable production of arabica coffee (*Coffea Arabica* L.): A review. *Euphytica* 204(2): 243 – 256.
- VSN International (2012). *GenStat for Windows*. (15<sup>th</sup> Edition).

## CHAPTER FOUR

### **Incorporation of coffee berry disease (CBD) resistance genes and screening of progenies of the crosses between Ethiopian coffee accessions and Tanzanian commercial variety KP423 using SSR markers**

#### **4.1 Abstract**

Coffee berry disease (CBD), caused by *Colletotrichum kahawae*, is a major constraint for Arabica coffee cultivation in Africa. Several previous studies have revealed molecular markers associated with its resistance. CBD is a mature stage disease that attacks berries at different developmental stages. The current study was aimed at using some of these markers at early stage screening for resistance to the disease on parents and progenies developed from the Tanzanian commercial variety and Ethiopian collection accessions. Eleven genotypes were crossed to a susceptible variety KP423, and the parents and their F1 progenies were used in the study. Physiological screening were applied on the hypocotyls of Parents and F1 seedlings using the procedure by Van der Vossen *et al.*, 1976. Marker screening was used on the DNA extracted from their leaves using gene specific markers Sat 235 and Sat 207. The genotypes containing the resistance gene were expected to show phenotypic resistance to CBD and to show banding patterns similar to the resistance donors while those lacking the genes were expected to show phenotypic susceptibility to CBD and similar banding pattern to susceptible commercial variety KP423. Presence of resistance gene was observed in the genotypes amplified by SSR marker Sat 235 which produced polymorphic bands similar to resistance donors except for the genotype 103 (Rume Sudan cross), 109 (commercial variety KP423) and 110 (PNI 088). The SSR marker Sat 207 produced similar bands to resistance donor genotypes with the exception of 103 (Rume Sudan cross) and 109 (commercial variety KP423. Resistant

and susceptible genotypes were clearly differentiated by gene specific markers SSR Sat 235 and Sat 207. Therefore these markers can be used in screening for resistance to CBD at early stages of development of genotypes involving Ethiopian and Tanzanian coffee arabica.

#### **4.2 Introduction**

In perennial crops, the use of MAS can greatly decrease the time taken to select for resistance and make it possible to rapidly identify genotypes combining several disease resistance genes (Collard and Mackill, 2008). With numerous plant diseases, selection progress is hampered by the fact that it is difficult to ensure uniform exposure to inoculum in the field in the screening process. Moreover, production and application of artificial inoculum can be difficult or expensive. Under these circumstances, marker assays are more reliable or simpler than tests for resistance and offer the possibility of screening breeding materials at a large scale with acceptable costs (Melchinger, 1990).

Coffee production in Tanzania is constrained by several factors. However the major challenge in the industry is the management of diseases and insect pests. In Arabica coffee, coffee berry disease (CBD) and coffee leaf rust (CLR) diseases are the major challenges while in Robusta production the big challenge is combating coffee wilt disease (CWD) (TCB, 2012). Epidemics of CBD can quickly destroy 70–80 % of the developing berries (6–16 weeks after anthesis) on susceptible arabica cultivars during prolonged wet and cool weather conditions (Van der Vossen and Walyaro, 2009). Preventive control measures by frequent fungicide sprays may account for 30–40 % of total production costs. Annual economic damage to arabica coffee production in Africa, due to crop loss by CBD and the cost of chemical control, is estimated at US\$ 300–500 million (Van der Vossen and Walyaro, 2009).

There is a wide publication on the availability of CBD resistance in some Ethiopian coffee genotypes. However, there is no information on their utilization anywhere outside Ethiopia (Van der Vossen *et al.*, 2015). Tanzania maintains 196 accessions of Ethiopian coffee origin which were collected and distributed by the FAO coffee collection mission of 1968 (FAO, 1968). Ethiopian arabica coffee cultivars are likely to be extremely valuable to other countries in Africa, where coffee berry disease prevents the cultivation of coffee, particularly by the smallest and poorest farmers (Robinson, 2007). More often traits of interest in perennial crops such as disease resistance can be observed and screened only at late stages of development and require assessment over a number of years at different locations. Disease resistance is typically a quantitative trait for which marker assisted selection (MAS) procedures can be efficiently developed.

The objective of this study was to incorporate CBD resistance from progenitors of resistance and screen progenies of the crosses between Ethiopian accessions and Tanzanian commercial variety KP423 for coffee berry disease (CBD) resistance genes using SSR markers.

### **4.3 Materials and Methods**

#### **4.3.1 Plant materials**

Eleven coffee genotypes (Table 3) were crossed to a susceptible commercial variety KP423, and the parents and their F1 progenies were screened phenotypically for coffee berry disease resistance using hypocotyl test (Van der Vossen *et al.*, 1976).

**Table 3: List of coffee genotypes used for coffee berry disease resistance genes screening**

Genotypes	Description	Code
F45/64/2049 x KP423	F1 hybrid	101
F90/64/4660 x KP423	F1 hybrid	102
Rume Sudan VC298 x KP423	F1 hybrid	103
F45/64/2061 x KP423	F1 hybrid	104
F24/64/902 x KP423	F1 hybrid	105
F24/64/886 x KP423	F1 hybrid	106
Catimor PRO 127 x KP423	F1 hybrid	107
F89/64/4650 x KP423	F1 hybrid	108
KP 423	Commercial	109
F24/64/886	Ethiopian	110
Catimor PRO 127	Progenitor	111
F89/64/4650	Progenitor	112
Rume Sudan VC298	Progenitor	113
PNI 088	Progenitor	114
Sarchimor	Progenitor	115
HdT VCE 1593	Progenitor	116
F90/64/4660	Ethiopian	117
F45/64/2049	Ethiopian	118
F45/64/2061	Ethiopian	119
F24/64/902	Ethiopian	120

#### 4.3.2 Microsatellite markers used

Two gene specific SSR markers for coffee berry disease (CBD) resistance - Sat 235 and Sat 207 from literature (Gichuru *et al.*, 2008) and which were sourced from Eurofins MWG Operon Ebersberg, Germany were used with primer sequences described in Table 4.

**Table 4: Gene specific SSR markers used in coffee berry disease resistance screening**

Marker name	Primer (5' - 3') Sequence	PCR size range (bp)	Reference
1 Sat 207	F GAAGCCGTTTCAAGCC R CAATCTC TTTCCGATGCTCT	190-250	Gichuru <i>et al.</i> , 2008
2 Sat 235	F TCGTTCTGTCATTAATCGTCAA R GCAAATCATGAAAATAGTTGGTG	97	Gichuru <i>et al.</i> , 2008

#### 4.4 Phenotypic Screening

Ripe coffee berries were harvested from the selected eleven genotypes (parents) while at the same time mature and ripe F1 crosses were harvested and processed separately maintaining the identity of the genotypes (parents) and hybrids. Two hundred seeds of each coffee genotype and F1 hybrid were sown out, with the parchment removed, in moist sterilized sand in plastic boxes with closely fitting transparent lids, keeping the boxes at normal room temperatures (20-24°C). The seedlings were ready for inoculation when they had a hypocotyl stem of 3-5 cm long at the 5-6<sup>th</sup> weeks after sowing while the cotyledons were still enclosed in the testae. The experiment was arranged in the laboratory in a completely randomized design with four replications. Each replicate was represented by 50 hypocotyl seedlings of the crosses and a line of 10 seedlings of the susceptible N 39 and resistant PNI 088 control pre-germinated alongside the crosses in a plastic box containing sterilized sand.

The boxes measuring 15 cm wide, 22 cm long and 15 cm deep were filled to half-depth with sand. All seedlings were inoculated the same day with conidia suspensions from 10 days old cultures standardized to  $2 \times 10^6$  conidia/ml following the procedure of Van der Vossen *et al.* (1976). The control seedlings were only used to observe pathogenicity of the isolate on the susceptible and resistant host. Just before inoculation the lids were removed from the plastic box and the seedlings were sprayed with the standard CBD inoculum ( $2 \times 10^6$  spores/ml) by means of a small hand atomizer held at 10 cm from the hypocotyls and the boxes were immediately closed again. The inoculation was repeated after 48 hours. For successful infection a temperature of about 22-24°C was maintained for the first four days, while relative humidity in the boxes was maintained at 100 % by use of lids tightly covering the plastic boxes containing the inoculated hypocotyls.

This was followed by an incubation period at lower constant temperature (19-20°C) with the lids removed from the boxes to allow for normal humidity. The first symptoms of CBD infection became noticeable within one week after the first inoculation, but an incubation period of at least two weeks was required for a full expression of the coffee genotype disease reaction. At the end of the incubation period (i.e. four weeks after inoculation), coffee seedlings were individually scored for CBD disease symptoms developed on the hypocotyl stem using a scale of 0-4 (Van der Graaff, 1978), where (0 = No of symptoms, 1 = one or two small brownish chlorotic lesions, 2 = coalescence brownish lesions, 3 = abundant black lesions, 4 = dead hypocotyls). Mean score data were subjected to analysis of variance using GenStat Release 12.1 statistic software by VSN International Ltd and mean separation was done as per Fisher (1958).

#### **4.5 Screening of Markers of Coffee Berry Disease Resistance**

Young coffee leaves were picked from the growing tips of 10 seedlings of the 11 genotypes and their F1 hybrids which were found resistant to CBD in the phenotypic CBD resistance screening. Leaf samples from the included susceptible commercial variety KP423 were picked from the seedlings of the same age but these seedlings were not subjected to phenotypic CBD resistance screening. These samples were lyophilized 72 hours before DNA extraction.

##### **4.5.1 DNA extraction**

The lyophilized leaves were stored at -21°C before DNA extraction. Genomic DNA was then extracted from the lyophilized leaves following the CTAB method (Diniz *et al.*, 2005) with minor modifications. About 500 mg of the lyophilized leaves were ground in eppendorf tubes using mortar and pestle. One ml each of lysis and extraction buffers was added to the powder in the eppendorf tube. The ground tissue was distributed in two 1.5

ml tubes and incubated at 65°C in a water bath for 20 to 30 min with regular shaking. After incubation, 1 ml of chloroform/isoamyl-alcohol mixture in the ratio of 24:1 was added to each tube, then mixed gently by shaking and centrifuged at 13 000 rpm for 10 to 15 min in a micro-centrifuge. The supernatants were pipetted out into new 1.5 ml tubes. 20 µl of RNase were added to the supernatants and incubated at 37°C in a water-bath for 30 min. An equal volume of isopropyl alcohol was added into each tube and mixed gently by inverting the tubes several times to precipitate DNA. The suspended DNA was centrifuged at 13 000 rpm for 5 min to obtain a DNA pellet and the supernatant carefully removed. The DNA pellets were then washed with 200 µl of 70 % ethanol and centrifuged at 13 000 rpm for 3 min. The ethanol was drained by decanting or micro-pipetting and the pellets dried in a vacuum centrifuge for 20 min. The pellets were dissolved in 60 µl of Tris-EDTA (TE) buffer and stored at 4°C. DNA quality and quantity were quantified on 1% agarose gel in 0.5xTAE (Tris-Acetate-EDTA) buffer, stained with ethidium bromide and visualized under UV light.

#### **4.5.2 Coffee DNA amplification**

Extracted coffee genomic DNA was amplified on GeneAmp PCR System 9700 thermocycler Applied BioSystems USA using two microsatellite primers designed by Eurofins MWG Operon LLC USA with the following forward and reverse sequence Sat 235 TCGTTCTGTCATTAAATCGTCAA and GCAAATCATGAAAATAGTTGGTG and Sat207 GAAGCCGTTTCAAGCC and CAATCTCTTTCCGATGCTCT respectively sourced from literature (Geleta *et al.*, 2012; Gichuru *et al.*, 2008). Touchdown PCR procedure (Poncet *et al.*, 2004) was followed. The PCR reactions for the SSR markers were made in 20 µL containing 2.0 µL of 10x buffer, 150 mM/L of dNTP, 0.1 mM /L of each primer, 25 ng of DNA, 1 mM / L MgCl<sub>2</sub>, 1 Unit Taq DNA polymerase and the remaining volume was completed with PCR water (Poncet *et al.*, 2004).

The reaction consisted of initial denaturation at 94°C/10 min, followed by 10 cycles of denaturation at 94°C/ 30s, annealing at 50°C/30s, and extension at 72°C/1 min. followed by a final extension at 72°C/7 min. All amplified products were confirmed by 2.0% (wt/vol) l.e (low electroendosmosis) agarose gel electrophoresis in 1× TAE (Tris base, acetic acid, and EDTA) buffer at 120V/115mA/15W for 2.30 hrs with ethidium bromide post-staining (0.5 µl/ml) for 30 min, followed by 10 min of destaining in 1× TAE buffer and visualized on UV. The reproducibility of the amplification products was checked twice for each primer.

## **4.6 Results**

### **4.6.1 Phenotypic screening**

Coffee berry disease resistance phenotypic screening mean data on parents are presented in (Table 5). There is significant ( $p<0.05$ ) difference among the parental genotypes for resistance to CBD. Four groups were deduced from the mean separation. The commercial variety KP423 had a mean score of 3.56 being the highest infection of CBD in the parental genotypes. The Ethiopian coffee genotype F24/64/T2061 had a mean score of 1.256 significantly ( $p<0.05$ ) differing from the rest of the parental coffee genotypes. The performance of the F1 hybrids of the cross between the parental genotypes and a CBD susceptible variety KP423 (Table 6) revealed significant ( $p<0.05$ ) difference on coffee berry disease resistance mean score among the hybrids. Five coffee hybrids groups were separated. The commercial variety KP423 had the highest CBD mean score (3.680).

**Table 5: Coffee parental genotypes for Coffee Berry Disease resistance mean score on (0-4 scale)**

Genotype	Description	Mean
SARHIMOR	Progenitor	0.08a
VC 298	Progenitor	0.08a
F90/64/T4660	Ethiopian	0.104a
PNI088	Progenitor	0.184a
HDT VCE 1593	Ethiopian	0.34ab
F45/64/T2049	Ethiopian	0.384ab
F89/64/T4650	Ethiopian	0.384ab
PRO127	Progenitor	0.62bc
F24/64/902	Ethiopian	0.912cd
F24/64/T886	Ethiopian	0.9296cd
F24/64/T2061	Ethiopian	1.256d
KP423	Commercial variety	3.560e

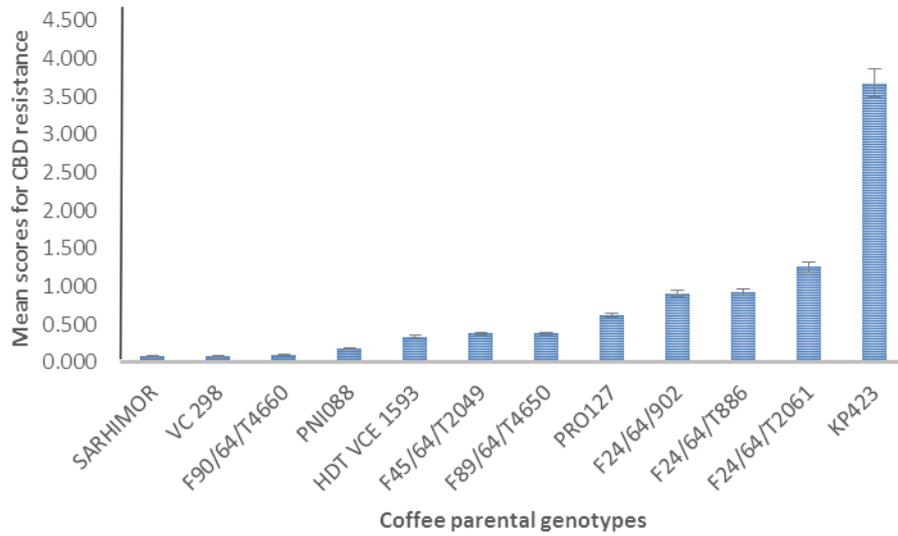
Mean values on the same column having the same letter(s) are not significantly different at  $p < 0.05$  according to Fisher. Disease severity scale (0-4) where 0 most resistant and 4 very susceptible

**Table 6: Coffee F<sub>1</sub> hybrids Coffee Berry Disease resistance mean score on (0-4 scale)**

Genotype	Description	Mean
VC 298 x KP423	F1 hybrid	0.112a
F90/64/T4660 x KP423	F1 hybrid	0.136a
VCE1593 x KP423	F1 hybrid	0.24ab
F89/64/T4650 x KP423	F1 hybrid	0.336bc
PNI 088 x KP423	F1 hybrid	0.408cd
SARCHIMOR x KP423	F1 hybrid	0.416cd
F45/64/T2049 x KP423	F1 hybrid	0.432cde
PRO 127 x KP423	F1 hybrid	0.464cde
F24/64/T2061 x KP423	F1 hybrid	0.504def
F24/64/902 x KP423	F1 hybrid	0.56ef
F24/64/T886 x KP423	F1 hybrid	0.64f
KP423	Commercial variety	3.680g

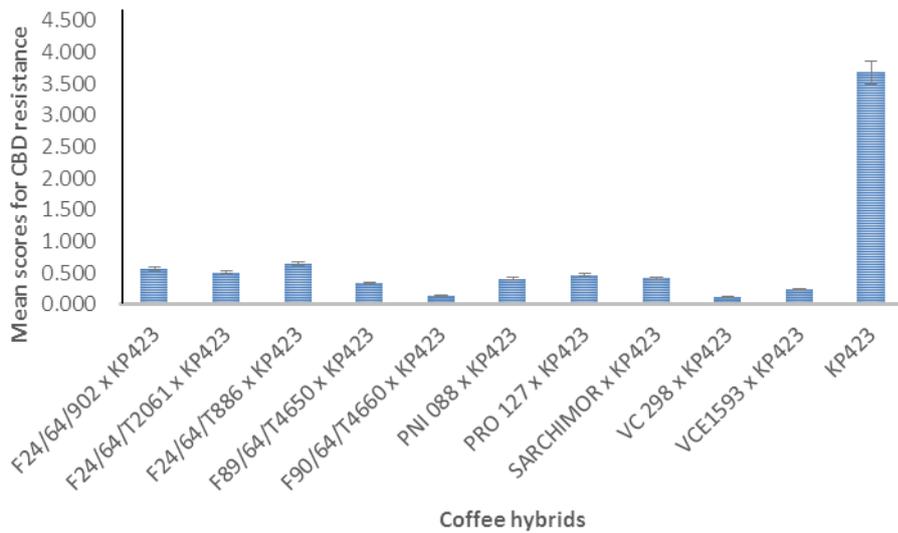
Mean values on the same column having the same letter(s) are not significantly different at  $p < 0.05$  according to Fisher.

Disease severity scale (0-4) where 0 most resistant and 4 very susceptible.



**Figure 6: Coffee parental genotypes mean score for Coffee Berry Disease resistance (0-4) scale**

Disease severity scale (0-4) where 0 most resistant and 4 very susceptible.

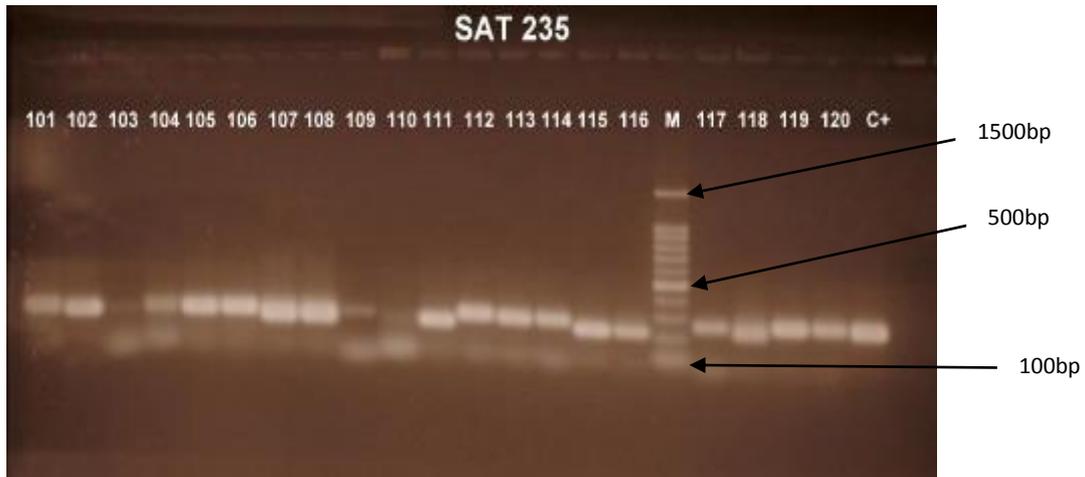


**Figure 7: Coffee F1 hybrids mean score for Coffee Berry Disease resistance on (0-4) scale**

Disease severity scale (0-4) where 0 most resistant and 4 very susceptible.

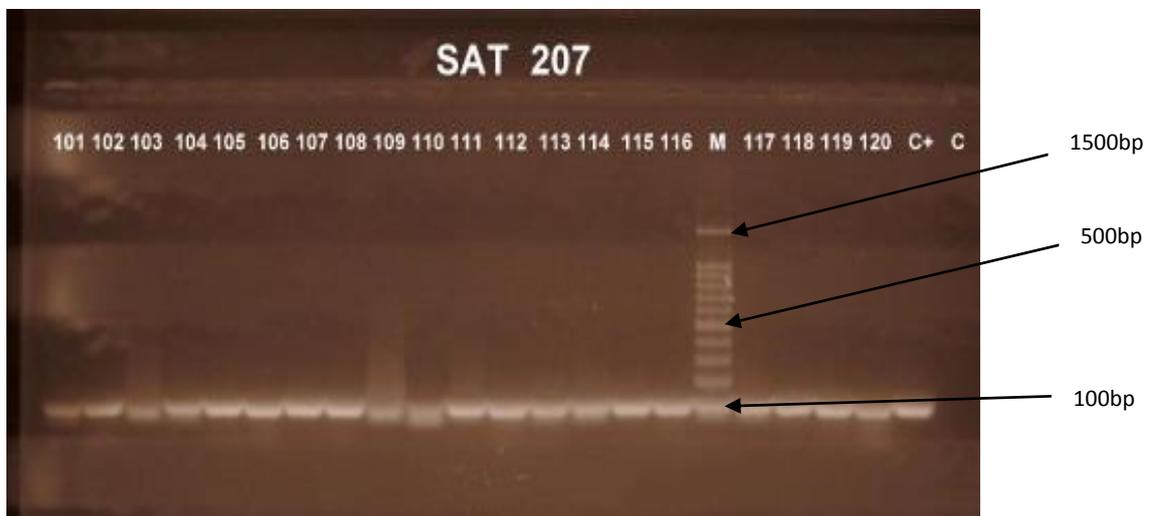
#### 4.6.2 Marker screening

The results of amplification of the 20 genotypes including parents and F1 hybrids (Fig. 8) by the SSR marker SAT 235 shows that the following coffee genotypes and hybrids were amplified F45/64/2049 x KP423 (101), F90/64/4660 x KP423 (102), F45/64/2061 x KP423 (104), F24/64/902 x KP423 (105), F24/64/886 x KP423 (106), Catimor PRO 127 x KP423 (107), F89/64/4650 x KP423 (108), Catimor PRO 127 (111), F89/64/4650 (112), Rume Sudan VC298 (113), PNI 088 (114) Sarchimor (115), HdT VCE 1593 (116), F90/64/4660 (117), F45/64/2049 (118), F45/64/2061 (119) and F24/64/902 (120). The coffee genotypes Rume Sudan VC298 x KP423 (103), KP 423 (109) and F24/64/886 (110) were not clearly amplified. The Ethiopian F1 coffee hybrids and parental genotypes that were amplified by SSR marker SAT207 (Fig. 9) included F45/64/2049 x KP423 (101), F90/64/4660 x KP423 (102), F45/64/2061 x KP423 (104), F24/64/902 x KP423 (105), F24/64/886 x KP423 (106), F89/64/4650 x KP423 (108), F90/64/4660 (117), F45/64/2049 (118), F45/64/2061 (119) and F24/64/902 (120) while Ethiopian parental genotype F24/64/886 (110) was not amplified (Fig. 9).



**Figure 8:** A 2% gel photograph showing amplification of coffee parental genotypes, F1 hybrids and a commercial variety KP423 (Table 3) by SSR marker Sat 235.

Lanes 101–120 represent coffee genotypes. Lane M represents a 100-bp ladder and lane C+ represent a control of a known amplification.



**Figure 9:** A 2% gel photograph showing amplification of coffee parental genotypes, F1 hybrids and a commercial variety (Table 3) by SSR marker Sat 207.

Lanes 101–120 represent genotypes. Lane M represents a 100-bp ladder and lane C+ and C represents controls of a known amplification.

## **4.7 Discussion**

### **4.7.1 Phenotypic screening**

A marker genetically linked to a Coffee Berry Disease resistance locus can be used to predict the presence of the resistant or the susceptible allele. Phenotypic CBD screening revealed that all coffee parental genotypes and F1 hybrids were medium to highly resistant to CBD (Tables 5 and 6) according to the disease severity scale of 0-4 by Van der Graaff (1978). One Ethiopian coffee parental genotype F24/64/T2061 (Table 5) scored a mean of 1.256 which indicates medium resistance to CBD. The commercial coffee variety KP423 clearly showed its susceptibility to CBD by scoring a mean of 3.68 and 3.68 respectively (Table 5 and 6). Phenotypic hypocotyl screening for CBD resistance have been applied in previous studies by Kiguongo (2015); Gichuru *et al.* (2008); Van der Graaff (1978) and Van der Vossen *et al.* (1976) where they were able to differentiate resistant and susceptible coffee arabica genotypes.

### **4.7.2 Marker screening**

The genotypes containing the resistance gene were expected to show phenotypic resistance to CBD and to show banding patterns similar to the resistant coffee genotypes and resistance donors (progenitors). The coffee genotypes lacking the gene were expected to show phenotypic susceptibility to CBD and similar banding pattern to susceptible commercial coffee variety KP423 (Gichimu *et al.*, 2014). However coffee genotype 103 a hybrid between Rume Sudan and a commercial variety KP423, showed a band pattern similar to the susceptible coffee variety KP423 (109) while phenotypically displayed a good resistance CBD. This is due to the fact that the resistance donor Rume Sudan coffee genotype is having its resistance gene on different loci and could not match the marker Sat 235 used in this screening. This fact is supported by two works done separately by Kiguongo *et al.* (2014) and Omondi and Pinard (2007) who found the resistance gene in

the coffee genotype Rume Sudan using SSR markers M24 and M24 and Sat 227 respectively. Similarly the Ethiopian coffee genotype F24/64/886 (110) displayed the pattern similar to the susceptible commercial coffee variety KP423 but phenotypically it displayed good resistance. This implied that its resistance gene was probably located in a different loci that could not match marker Sat 235. To achieve more comprehensive results it is necessary to do marker screening for resistance to CBD using more markers. In marker assisted selection (MAS) in other perennial crop grapes, the allele linked to resistance were found to be unique in size and were not shared with the susceptible cultivars when they were screened with two flanking microsatellite markers. The distinctiveness of these resistant alleles allowed the use of MAS to optimize the breeding of Pierce's Disease-resistant grape cultivars (Riaz *et al.*, 2009).

#### **4.8 Conclusion**

The presence of the coffee berry disease resistance genes was revealed in all the studied coffee genotypes amplified by SSR marker Sat 235 and Sat 207. This was confirmed by production of bands similar to the progenitors of CBD resistance. Coffee genotypes Rume Sudan hybrid (103) and PNI 088 (110) showed that their CBD resistance genes were located in a different loci as they were not amplified by markers Sat 235 and Sat 207. This findings indicated that marker screening can be used in coffee berry disease resistant genotypes selection at early stages of growth hence reducing the time of selection cycle.

## References

- Collard, B. C. Y. and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society of London. Biological Sciences* 363: 557 – 572.
- Diniz, L. E. C., Ruas, C. F., Carvalho, V. P., Torres, F. M., Ruas, E. A., Santos, M. O., Sera, T and Ruas, P. M. (2005). Genetic diversity among forty coffee varieties assessed by RAPD markers associated with restriction digestion. *Brazil Archive Biology Technical* 48(4): 511 – 521.
- FAO (1968). *Coffee Mission to Ethiopian 1964-1965*. Food and Agriculture Organization, Rome, Italy. 200pp.
- Fisher, R. A. (1958). *Statistical Methods for Research Workers*. (13<sup>th</sup> Ed.), Oliver and Boyd Edinburgh, London. 120pp.
- Geleta, M., Herrera, I., Monzón, A. and Bryngelsson, T. (2012). Genetic Diversity of Arabica Coffee (*Coffea arabica* L.) in Nicaragua as estimated by simple sequence repeat markers. *The Scientific World Journal* 2012: 1 – 11.
- Gichimu, B. M., Gichuru, E. K., Mamati, G. E and Nyende, A. B. (2014). Occurrence of Ck-1 gene conferring resistance to Coffee Berry Disease in *Coffea arabica* cv. Ruiru 11 and its parental genotypes. *Journal of Agricultural and Crop Research* 2(3): 51 – 61.
- Gichuru, E. K., Agwanda, C. O., Combes, M. C., Mutitu, E. W., Ngugi, E. C. K., Bertrand, B. and Lashermes, P. (2008). Identification of molecular markers linked to a gene conferring resistance to coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica*. *Plant Pathology* 57(6): 1117 – 1124.

- Kiguongo, A. P. K., Omondi, C. O., Gichuru, E. K and Kasili, R. W. (2014). Analysis of simple sequence repeat markers linked to coffee berry disease resistance genes in a segregating population of arabica coffee (*Coffea arabica* L.). *International Journal of Biotechnology and Food Science* 2(8):156 – 166.
- Melchinger, A. E. (1990). Use of molecular markers in breeding for oligogenic disease resistance. *Plant Breeding* 104: 1 – 9.
- Omondi, C. O. and Pinard, F. (2007). Screening populations of Arabica coffee for molecular markers linked to Coffee Berry Disease resistance. In: *21<sup>st</sup> International Conference on Coffee Science*, held from 11-15 September 2007, Montpellier, France. pp. 1377 – 1379.
- Poncet, V., Hamon, P., Minier, J., Carasco, C., Hamon, S. and Noirot, M. (2004). SSR cross-amplification and variation within coffee trees (*Coffea* spp.). *Genome* 47: 1071 – 1081.
- Riaz, S., Tenschler, A. C., Graziani, R., Krivanek, A. F., Andrew, M. and Walker, M. A. (2009). Using marker-assisted selection to breed pierce's disease-resistant grapes. *American Journal of Enology and Viticulture* 60(2):199 – 207.
- Robinson, R. A. (2007). The Loss of Resistance in Coffee. *Return to Resistance*. (3<sup>rd</sup> Ed.), Sharebooks Publishing Canada. pp. 231 – 252.
- TCB (2012). *Tanzania Coffee Industry Development Strategy 2011/2021*. Tanzania Coffee Board, Moshi, Tanzania. 64pp.
- Van der Graaff, N. A. (1978). Selection for resistance to coffee berry disease in arabica coffee in Ethiopia. Evaluation of selection methods. *Netherlands Journal of Plant Pathology* 84(6): 205 – 215.

- Van Der Vossen, H. Bertrand, B. and Charrier, A. and (2015). Next generation Variety development for sustainable production of arabica coffee (*Coffea Arabica* L.): A review. *Euphytica* 204(2): 243 – 256.
- Van der Vossen, H. A. M. and Walyaro, D. J. (2009). Additional evidence for oligogenic inheritance of durable host resistance to coffee berry disease (*Colletotrichum kahawae*) in arabica coffee (*Coffea arabica* L.). *Euphytica* 165(1): 105 – 111.
- Van der Vossen, H. A. M., Cook, R. T. A. and Murakaru, G. N. W. (1976). Breeding for resistance to Coffee berry disease caused by *Colletotrichum coffeanum* Noack (sensu Hindorf) in *Coffea arabica* L. I. Methods of preselection for resistance. *Euphytica* 25: 733 – 745.
- VSN International (2012). *GenStat for Windows*. (15<sup>th</sup> Edition).

## CHAPTER FIVE

### **Combining ability, heritability and relationships of growth and yield variables of progenies of commercial coffee variety KP423 and selected accessions of Ethiopian collection**

#### **5.1 Abstract**

Knowledge of the associations of characteristics of agronomic importance enable breeders to choose rational crop improvement strategies. Estimation of genetic variables of important selection traits guides selection in breeding programmes reducing time required in selection and variety release. In coffee, growth variables stem girth, plant height, inter-node length and canopy radius were found heritable on young coffee trees. In determining the combining ability, heritability and relationships of growth and yield characters of Ethiopian genotypes and Tanzanian commercial variety KP423, nine genotypes were selected including six from Ethiopian collection maintained at Tanzania Coffee Research Institute, two from the germplasm and a commercial variety KP 423 were studied in a half-diallel mating design ( $9 \times 9$ ). Forty five F1s were produced. The experiment was established using a randomized complete block design with three replications and three plants per plot. Data were collected at Year 1 and 2 after establishment on: Stem girth, plant height, length of the longest primary, number of primaries per plant, number of berries per cluster, internode length and number of bearing primaries per plant. Analysis was based on the fixed effect (model 1) method II by Griffing (1956) while Path coefficients analysis was performed according to Dewey and Lu (1959). Analysis was done using Sheoran *et al.* (1998) Statistical Package. Both general and specific combining ability variances were significant ( $p=0.01$ ) for all variables except berries per node. Genotypic correlation coefficients ( $r$ ) were found to be

positive and highly significant for all variables. Similarly phenotypic correlations were positive and significant for all variables except that between number of primaries per plant and internode length. Broad sense heritability values were higher than narrow sense heritability for all variables. Variables length of the longest primary, stem girth and internode length were found contributing mostly to berries per node. The direct effect due to plant height was found contributing most negatively to berries per node. Variables length of the longest primary, internode length and stem girth may be used as important criteria in selection for superior coffee arabica varieties involving Ethiopian genotypes and commercial variety KP423.

## **5.2 Introduction**

Knowledge of interrelationships and associations of characters of agronomic and commercial interest provides an efficient tool to be considered in breeding programmes. This approach allows the choice of suitable selection strategy reducing the time required in cultivar development and release (Mistro *et al.*, 2007). In coffee, most of the agronomic characters are polygenic and they display different levels of interrelationships (Sureshkumar *et al.*, 2013). Estimation of genetic parameters of the main selection traits guides selection in breeding programmes. In coffee, growth characters especially stem girth, plant height, inter-node length on stem and primaries, and canopy radius had high repeatability. Such characters are heritable, even with a single measurement taken on young coffee trees (Walyaro, 1983).

The objective of this work was to determine the combining ability, heritability and relationships of growth and yield variables of progenies of commercial variety KP423 and selected accessions of Ethiopian collection.

## 5.3 Materials and Methods

### 5.3.1 Description of the experimental site

The trial was established at an experimental field of the Tanzania Coffee Research Institute, Lyamungu in 2013. The experimental site is located at 3°14'58'' S and 37°14'36''E and at an altitude of 1275 meters above sea level. The area receives an average of 1250 mm rainfall per year.

### 5.3.2 Coffee genotypes

Nine coffee genotypes including six selected from Ethiopian collection maintained at the Tanzania Coffee Research Institute, two progenitors from the germplasm and a commercial coffee variety KP423 were used in this study (Table 7). They were selected on the basis of their good agronomic performance from the records. The selected coffee genotypes were crossed in a half-diallel mating design (9 × 9). Forty five progenies (36 F1 hybrids and 9 selfs) were produced.

**Table 7: List of coffee genotypes used in the study**

S/No.	Accession	Origin
1	F45/64/2061	Ethiopian
2	F45/64/2049	Ethiopian
3	F24/64/902	Ethiopian
4	F24/64/886	Ethiopian
5	F90/64/4660	Ethiopian
6	F89/64/4650	Ethiopian
7	KP423	Commercial variety
8	VC 298	Progenitor
9	PRO 127	Progenitor

### **5.3.3 Hybridization**

Selection of male and female parent trees from the selected coffee genotypes was done in November 2012 followed by uniform intensive management of the selected trees to ensure synchronization of flowering for male and female selected trees. Flower blossoming in coffee is stimulated wherever there is a certain increase in the availability of soil moisture preceded by water stress period of at least 6 weeks. Emasculation of the female mother tree flowers was effected by hand removal of the stamen 1-4 days prior to anthesis. The emasculated flower parts were then isolated by pollen barriers made from pollen proof muslin cloth. The stigmas of un-pollinated flowers remained receptive for at least nine days (Walyaro and Van der Vossen, 1977). Branches selected for pollen harvest were also isolated as emasculated flowers to avoid contamination of the pollen until the pollen was harvested. Pollen was harvested from male tree flowers between 6 am – 8 am of the pollination day by using sterilized camel brushes. To avoid contamination of pollen from different coffee genotypes, brushes were sterilized using methylated spirit after every genotype.

Pollination was done early in the morning on the day of anthesis by conveying it to the stigma of the emasculated flower by help of camel hair brushes. The pollinated female tree flowers remained isolated for seven days from the day of pollination to ensure no contamination from other sources of pollen. After pollination the label was attached on the branches which included information on: pollen source (male genotype), female genotype, date of pollination, number of emasculated flowers, and name of the pollinator (person). The crosses were cared in the field by ensuring that the trees were properly nourished, pests were controlled and fruits were picked and processed separately and immediately sown in polybags with the proper identity labels. The seedlings were cared at the nursery for six months before field planting.

#### 5.3.4 Field experimental design data collection and analysis

The experiment was conducted using a randomized complete block design with three replications and three trees per plot. The planting and field maintenance followed recommended procedures for coffee (TaCRI, 2011). Data was recorded on stem diameter (G), measured in centimeter, five centimeters from the ground level using a vernier-calliper, plant height (H), measured in centimeter, from the ground level to the top tip using three meter ruler. Length of the longest primaries (LP), were measured in centimeters on the three longest selected branches per tree between rows. The number of primaries (NP) was obtained by counting the total number of primaries per coffee tree. The berries per node (BN) were obtained by counting and recording as an average of the eight most heavily bearing nodes per tree. Internode length (IL) was obtained by measuring with a graduated ruler between two nodes, on the longest primary. The number of bearing primaries (NBP) per coffee tree was obtained by counting all the primaries with berries or flower buds. Data were recorded on three trees per plot, at the end of the first and second year after field establishment.

Data were collected for two seasons and data analysis was done using Sheoran *et al.*, (1998) Statistical Package for Agricultural Research Workers online software to assess the performance variation of the genotypes. From the ANOVA, the mean squares were used to evaluate levels of significant differences among genotypes. Analysis was based on the fixed effect (model 1) method II as recommended by Griffing (1956) when using parents and  $F_1$ s. The following model was used.

$$Y_{ijk} = \mu + g_i + g_j + s_{ij} + e_{ijk}$$

Where  $\mu$  is the grand mean,  $g_i$  and  $g_j$  are the general combining abilities for  $i^{\text{th}}$  and  $j^{\text{th}}$  parents, respectively;  $s_{ij}$  is the specific combining ability between  $i^{\text{th}}$  and  $j^{\text{th}}$  parent and  $e_{ijk}$  is the residual effect.

Combining ability analysis was performed where the mean squares for combining abilities GCA and SCA effects, variances and standard errors were determined. Heritability values were obtained using GCA and SCA values based on the following equations.

$$h^2_n = 2 \sigma_g^2 / (2 \sigma_g^2 + \sigma_s^2 + \sigma_e^2) \quad \text{Narrow sense}$$

$$h^2_b = (2 \sigma_g^2 + \sigma_s^2) / (2 \sigma_g^2 + \sigma_s^2 + \sigma_e^2) \quad \text{Broad sense}$$

Where:  $\sigma_g^2$ ,  $\sigma_s^2$  and  $\sigma_e^2$  are respectively the GCA, SCA and within-cross variances.

Path coefficients analysis was performed according to Dewey and Lu (1959).

#### **5.4 Results**

Study of the relationships of coffee growth variables stem girth, plant height, length of the longest primary and number of primaries per coffee plant showed positive and significant variations. Similarly yield variables berries per node, internode length and number of bearing primaries per coffee plant revealed that there were positive and significant variations for coffee hybrids during the first year (Table 8). Second year coffee growth and yield variables data showed that there were significant variations in growth and yield variables except for the yield variable coffee berries per node (Table 9).

**Table 8: Analysis of variance for growth variables in F<sub>1</sub> coffee hybrids Year 1 (2014)**

Character	Source of Variation	DF	SS	MS	F-calc.	Significance.
Stem girth	Replications	2	2.771			
	Treatments	44	9.776	0.222	1.989	0.00317
	Error	88	9.831	0.112		
	Total	134	22.378			
Plant height	Replications	2	2 902.41			
	Treatments	44	40 284.32	915.553	5.426	0.00000
	Error	88	14 848.18	168.729		
	Total	134	58 034.91			
Length of longest primary	Replications	2	1 919.48			
	Treatments	44	7 861.02	178.66	3.243	0.00000
	Error	88	4 848.05	55.091		
	Total	134	14 628.54			
Number of primaries	Replications	2	236.784			
	Treatments	44	690.012	15.682	1.867	0.00664
	Error	88	739.293	8.401		
	Total	134	1 666.09			
Berries per node	Replications	2	2 220.45			
	Treatments	44	1 671.21	37.982	2.681	0.00004
	Error	88	1 246.76	14.168		
	Total	134	5 138.43			
Internode length	Replications	2	6.477			
	Treatments	44	55.765	1.267	2.200	0.00086
	Error	88	50.695	0.576		
	Total	134	112.937			
Number of bearing primaries	Replications	2	161.762			
	Treatments	44	954.555	21.694	2.133	0.0013
	Error	88	894.882	10.169		
	Total	134	2 011.20			

**Table 9: Analysis of variance for growth variables in F<sub>1</sub> coffee hybrids Year 2 (2015)**

Character	Source Variation.	DF	SS	MS	F-calc.	Significance
Stem girth	Replications	2	1.622			
	Treatments	44	15.474	0.352	2.453	0.00018
	Error	88	12.617	0.143		
	Total	134	29.713			
Plant height	Replications	2	4 473.282			
	Treatments	44	54 487.930	1 238.362	3.564	0.00000
	Error	88	30 577.512	347.472		
	Total	134	89 538.725			
Length of longest primary	Replications	2	1 122.236			
	Treatments	44	10 419.943	236.817	3.092	0.00000
	Error	88	6 740.407	76.596		
	Total	134	18 282.586			
Number of primaries	Replications	2	199.994			
	Treatments	44	2 397.900	54.498	2.491	0.00014
	Error	88	1 925.413	21.880		
	Total	134	4 523.306			
Berries per node	Replications	2	37.557			
	Treatments	44	181.417	4.123	0.715	0.88991
	Error	88	507.648	5.769		
	Total	134	726.622			
Internode length	Replications	2	0.106			
	Treatments	44	38.132	0.867	2.914	0.00001
	Error	88	26.168	0.297		
	Total	134	64.406			
Number of bearing primaries	Replications	2	142.226			
	Treatments	44	1 138.145	25.867	2.014	0.00272
	Error	88	1 130.215	12.843		
	Total	134	2 410.586			

Analysis of variance (ANOVA) for combining ability and heritability of growth variables in F<sub>1</sub> coffee hybrids for the first year data (Table 10) showed that general and specific combining abilities were significant ( $p = 0.05$ ) for stem girth, plant height, length of the longest primary and internode length.

Significant ( $p = 0.01$ ) specific combining ability (SCA) was observed in coffee variables the number of primaries per coffee plant, average number of berries per node, and number of bearing primaries per plant. The ratio of variances due to general combining ability to variances due to specific combining ability ( $\delta^2_{GCA} / \delta^2_{SCA}$ ) was close to one for most variables except number of bearing primaries per coffee plant (0.813), number of berries per node (0.743) and internode length (1.759).

The broad sense heritability ( $H^2_{bs}$ ) for all variables studied ranged from 0.852 to 0.946 while narrow sense heritability ( $h^2_{ns}$ ) ranged from 0.354 to 0.537. The broad sense heritability values were generally higher than the narrow sense heritability (Table 10). Analysis of variance data for the second year (Table 11) for combining ability and heritability of growth variables in  $F_1$  coffee hybrids revealed that general (GCA) and specific combining ability (SCA) were significant for stem girth, plant height, length of the longest primary, number of primaries per coffee plant, internode length and number of bearing primaries per coffee plant. The ratio of variances due to general combining ability to variances due to specific combining ability ( $\delta^2_{GCA} / \delta^2_{SCA}$ ) was close to one for most variables except number of bearing primaries per coffee plant (1.556), berries per node (0.720) and number of primaries per coffee plant (0.796). Broad sense heritability ( $H^2_b$ ) ranged from 0.647-0.912 while narrow sense heritability ( $h^2_n$ ) ranged from 0.236-0.501.

**Table 10: ANOVA for combining ability and heritability of yield and growth variables of F<sub>1</sub> coffee hybrids Year 1 (2014)**

Source	Df	Stem Girth (G) (cm)	Plant height (H) (cm)	Longest Primary Length (LP) (cm)	Number of Primaries (NP) (nos)	Number of berries per node (B/N) (nos)	Internode Length (IL) (cm)	Number of bearing Primaries (NBP) (nos)
GCA	8	0.277*	1,023.122**	190.087**	16.490	29.615	1.959**	18.261
SCA	36	0.210**	891.649**	176.117**	15.502*	39.842**	1.114**	22.457**
Error	88	0.112	168.729	55.091	8.401	14.168	0.576	10.169
$\delta^2_{GCA} / \delta^2_{SCA}$		1.318	1.147	1.079	1.064	0.743	1.759	0.813
Narrow sense ( $h^2_{ns}$ )		0.462	0.491	0.451	0.408	0.354	0.537	0.359
Broad sense ( $H^2_{bs}$ )		0.872	0.946	0.910	0.852	0.875	0.897	0.853

\*Significantly different at P=0.05, \*\* significantly different at P=0.01

**Table 11: ANOVA for combining ability and heritability of yield and growth variables of F<sub>1</sub> coffee hybrids Year 2 (2015)**

Source	Df	Stem girth (G) (cm)	Plant height (H) (cm)	Longest Primary length (LP) (cm)	Number of primaries (NP) (nos)	Number of berries per node (B/N) (nos)	Internode length(IL) (cm)	Number of bearing primaries (NBP) (nos)
GCA	8	0.323*	1 162.044**	282.783**	45.052*	3.125	0.965**	36.551**
SCA	36	0.358**	1 255.322**	226.607**	56.597**	4.345	0.845**	23.493*
Error	88	0.143	347.472	76.596	21.880	5.769	0.297	12.843
$\delta^2_{GCA}/\delta^2_{SCA}$		0.902	0.926	1.248	0.796	0.720	1.142	1.556
Narrow sense ( $h^2_{ns}$ )		0.392	0.420	0.483	0.365	0.236	0.458	0.501
Broad sense ( $H^2_{bs}$ )		0.875	0.912	0.912	0.870	0.647	0.903	0.883

\*Significantly different at P=0.05, \*\* significantly different at P=0.01

First year correlation data (Table 12) revealed that there were positive and significant correlations between variables stem girth vs plant height, length of the longest primary, number of primaries per coffee plant, stem girth vs berries per node, stem girth vs internode length, and number of bearing primaries per coffee plant. Similarly, there were positive and significant correlations between plant height vs length of the longest primary, number of primaries per coffee plant, berries per node, internode length and bearing primaries per coffee plant. Likewise there were positive and significant correlations between variables length of longest primary vs number of primaries per coffee plant, berries per node, internode length and number of bearing primaries per coffee plant. Also positive and significant correlations were observed for variables number of primaries per coffee plant vs berries per node and bearing primaries per coffee plant. Positive and significant correlations were also observed on variables berries per node vs internode length and number of bearing primaries per coffee plant. There were no significant correlations on variables internode length vs number of primaries per coffee plant and number of bearing primaries vs internode length. Correlation data of the growth and yield characters for the second year (Table 13) showed that stem girth was positive and significantly correlated with plant height, length of the longest primary, number of primaries per plant, average berries per node, average internode length and number of bearing primaries per plant. Plant height was found to be positive and significantly correlated with the length of the longest primary, average internode length and number of bearing primaries per plant. Length of the longest primary was found to be positive and significantly correlated with average internode length and number of bearing primaries per coffee plant. Non-significant correlations were observed on variables plant height vs number of primaries per coffee plant and number of berries per node. Similarly non-significant correlations were observed on variables length of the longest primary vs number of primaries per coffee plant and berries per node. Negative and non-significant

correlation was observed on variable number of primaries per coffee plant vs internode length.

**Table 12: Simple correlation of yield and growth variables of F<sub>1</sub> coffee hybrids (1<sup>st</sup> Year) (2014)**

	Stem girth	Plant height	Length of primary	Number of primaries	Berries per node	Internode length	Bearing primaries
Stem girth							
Plant height	0.664**						
Length of primary	0.755**	0.887**					
Number of primaries	0.774**	0.445**	0.492**				
Berries per node	0.645**	0.462**	0.618**	0.577**			
Internode length	0.474**	0.769**	0.814**	0.172 <sup>NS</sup>	0.381**		
Bearing primaries	0.687**	0.350*	0.460**	0.838**	0.571**	0.179 <sup>NS</sup>	

\* Significantly different at P = 0.05, \*\* significantly different at P = 0.01

**Table 13: Simple correlation of yield and growth variables of F<sub>1</sub> coffee hybrids (2nd Year) (2015)**

	Stem girth	Plant height	Length of primary	Number of primaries	Berries per node	Internode length	Bearing primaries
Stem girth							
Plant height	0.754**						
Length of primary	0.739**	0.826**					
Number of primaries	0.498**	0.184 <sup>NS</sup>	0.072 <sup>NS</sup>				
Berries per node	0.360*	0.180 <sup>NS</sup>	0.241 <sup>NS</sup>	0.156 <sup>NS</sup>			
Internode length	0.503**	0.701**	0.818**	-0.182 <sup>NS</sup>	0.169 <sup>NS</sup>		
Bearing primaries	0.627**	0.451**	0.402**	0.665**	0.144 <sup>NS</sup>	0.062 <sup>NS</sup>	

\* Significantly different at P = 0.05, \*\* significantly different at P = 0.01

#### 5.4.1 Genotypic and phenotypic correlations

Genotypic (above) and phenotypic (below) correlations between the variables are shown in Table 14. All genotypic correlations between variables were significant and positive except the correlation between internode length (IL) with number of bearing primaries per coffee plant (NBPP). Phenotypic correlations were positive and significant between pairs of all variables except number of primaries per plant (NPP) with internode length (IL) which was positive but not significantly so.

**Table 14: Genotypic and phenotypic correlation coefficients for growth and yield variables of coffee**

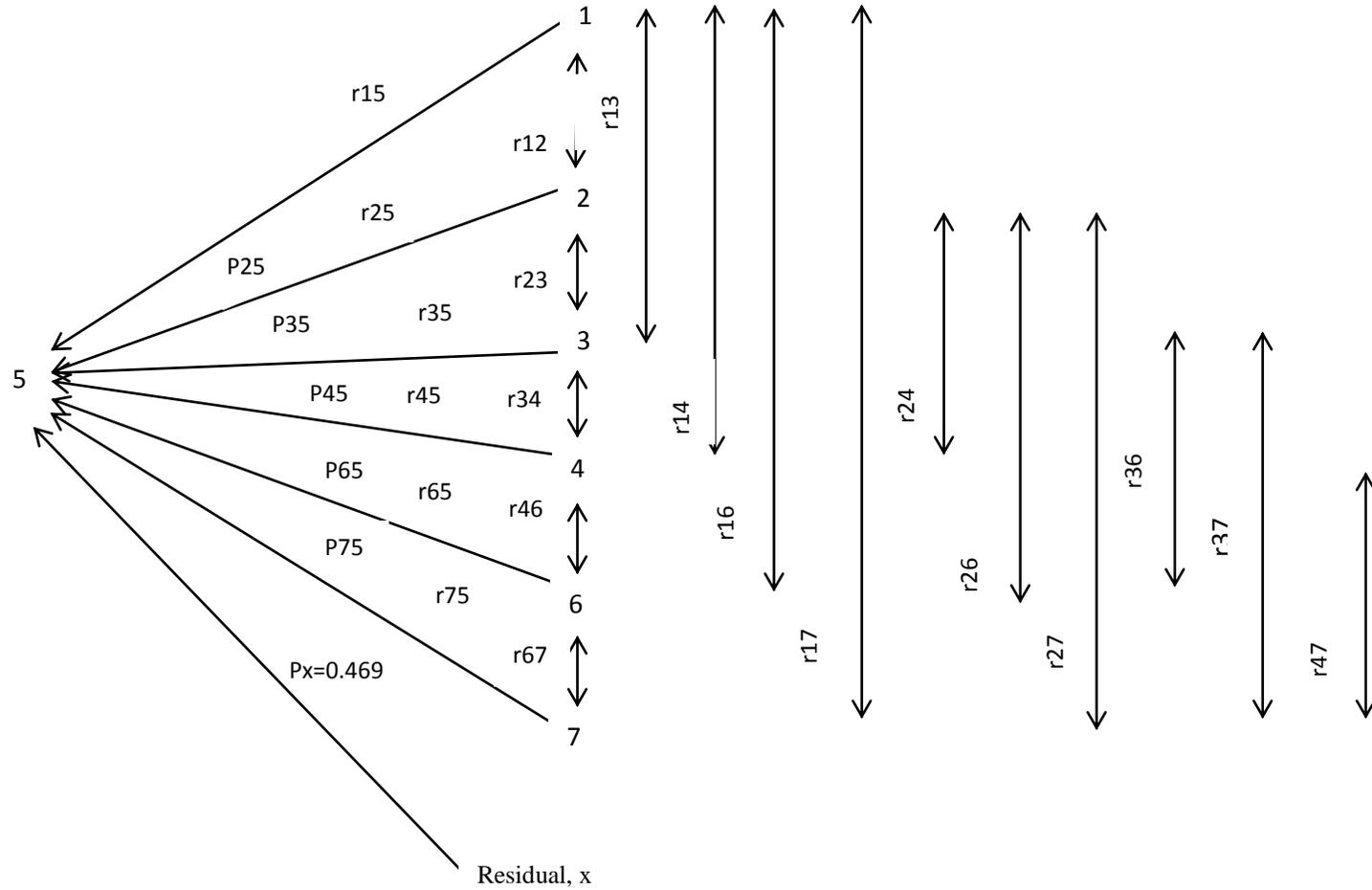
Variable		1	2	3	4	5	6	7
1. SG								
2. PH	rG	0.679**						
	rP	0.684**						
3. LLP	rG	0.687**	0.922**					
	rP	0.787**	0.857**					
4. NPP	rG	0.841**	0.354**	0.327**				
	rP	0.731**	0.515**	0.535**				
5. BPN	rG	0.538**	0.403**	0.583**	0.389**			
	rP	0.558**	0.413**	0.602**	0.431**			
6. IL	rG	0.543**	1.006**	1.070**	0.197*	0.483**		
	rP	0.458**	0.606**	0.645**	0.151NS	0.284**		
7. NBPP	rG	0.648**	0.306**	0.343**	0.792**	0.248**	0.110NS	
	rP	0.678**	0.433**	0.537**	0.751**	0.451**	0.256**	

Genotypic coefficient of correlation ( $r_G$ ) is shown on the top and the phenotypic correlation coefficient ( $r_P$ ) is shown on the bottom of each cell corresponding to the variables in a row. SG = stem girth, PH = plant height, LLP = length of the longest primary, NPP = number of primaries per plant, BPN = berries per node, IL = internode length, NBPP = number of bearing primaries per plant.

\*Significantly different at  $P = 0.05$  \*\* significantly different at  $P = 0.01$

#### 5.4.2 Path Coefficient Analysis

Fig10 and Table 15 show the direct and indirect effects based on genotypic correlations on analysis of growth and yield variables of coffee. Since number of berries per node (BPN) has been found in previous studies to be representative of yield (Walyaro, 1983), it was considered as a dependent variable in this investigation and the rest as independent variables.



**Figure 10: Path diagram and coefficients of factors influencing the number of berries per node of coffee. P's are the direct effects; r's are the genotypic correlation coefficients**

**Legend:**  
 P's are the direct effects; r's are the genotypic correlation coefficients

$P_{15} = 0.356, P_{25} = -1.170, P_{35} = 0.934, P_{45} = 0.286, P_{65} = 0.434, P_{75} = -0.220, r_{12} = 0.679, r_{13} = 0.687, r_{14} = 0.841, r_{15} = 0.538, r_{16} = 0.543, r_{17} = 0.648, r_{23} = 0.922, r_{24} = 0.354, r_{25} = 0.403, r_{26} = 1.006, r_{27} = 0.306, r_{34} = 0.327, r_{35} = 0.583, r_{36} = 1.070, r_{37} = 0.343, r_{45} = 0.389, r_{46} = 0.197, r_{47} = 0.792, r_{65} = 0.483, r_{67} = 0.110, r_{75} = 0.248$

**Table 15: Direct and indirect influences of some of coffee growth and yield variables on berries per node**

Variable	r
<b>1. Stem girth <math>r_{15}</math></b>	0.538**
Direct effect $P_{15}$	0.356
Indirect effect via plant height $r_{12}P_{25}$	-0.794
Indirect effect via longest primary $r_{13}P_{35}$	0.642
Indirect via No. primary per plant $r_{14}P_{45}$	0.241
Indirect via internode length $r_{16}P_{65}$	0.237
Indirect via No. bearing primaries per plant $r_{14}P_{45}$	-0.143
<b>Total</b>	<b>0.538</b>
<b>2. Plant height <math>r_{25}</math></b>	0.403**
Direct effect $P_{25}$	-1.17
Indirect effect via stem girth $r_{12}P_{15}$	0.242
Indirect effect via longest primary $r_{23}P_{35}$	0.861
Indirect via No. primaries per plant $r_{24}P_{45}$	0.101
Indirect via internode length $r_{26}P_{65}$	0.437
Indirect via No. bearing primaries per plant $r_{27}P_{75}$	-0.067
<b>Total</b>	<b>0.403</b>
<b>3. Length of the longest primary <math>r_{35}</math></b>	0.583**
Direct effect $P_{35}$	0.934
Indirect effect via stem girth $r_{13}P_{15}$	0.245
Indirect effect via plant height $r_{23}P_{25}$	-1.079
Indirect via No. primaries per plant $r_{34}P_{45}$	0.094
Indirect via internode length $r_{36}P_{65}$	0.464
Indirect via No. bearing primaries per plant $r_{37}P_{75}$	-0.075
<b>Total</b>	<b>0.583</b>
<b>4. Number of primaries per plant <math>r_{45}</math></b>	0.389**
Direct effect $P_{45}$	0.286
Indirect effect via stem girth $r_{14}P_{15}$	0.299
Indirect effect via plant height $r_{24}P_{25}$	-0.414
Indirect effect via longest primary $r_{34}P_{35}$	0.305
Indirect via internode length $r_{64}P_{65}$	0.085
Indirect via No. bearing primaries per plant $r_{74}P_{75}$	-0.174
<b>Total</b>	<b>0.389</b>
<b>5. Internode length <math>r_{65}</math></b>	0.483**
Direct effect $P_{65}$	0.434
Indirect effect via stem girth $r_{16}P_{15}$	0.199
Indirect effect via plant height $r_{26}P_{25}$	-1.177
Indirect effect via longest primary $r_{36}P_{35}$	0.999
Indirect via No. primaries per plant $r_{46}P_{45}$	0.056
Indirect via No. bearing primaries per plant $r_{67}P_{75}$	-0.024
<b>Total</b>	<b>0.487</b>
<b>6. Number of bearing primaries per plant <math>r_{75}</math></b>	0.240**
Direct effect $P_{75}$	-0.22
Indirect effect via stem girth $r_{17}P_{15}$	0.230
Indirect effect via plant height $r_{27}P_{25}$	-0.358
Indirect effect via longest primary $r_{37}P_{35}$	0.320
Indirect via No. primaries per plant $r_{47}P_{45}$	0.226
Indirect via internode length $r_{67}P_{65}$	0.047
<b>Total</b>	<b>0.240</b>
<b>7. Residual effect <math>P_{x5}</math></b>	0.469

### **Stem girth (SG) vs number of berries per node (BPN)**

The significant genetic correlation between stem girth with number of berries per node ( $r = 0.538^{**}$ ) was predominantly due to the direct effect (0.356) and the indirect effect of stem girth via length of the longest primary (0.642) and to a lesser extent via primaries per plant (0.241) and internode length (0.237) (Table 15). However, these positive influences were reduced to a low but significant correlation by the negative indirect effect through plant height (-0.794). The latter was due to the negative direct effect (-1.17) of plant height on berries per node while the correlation between stem girth and plant height was positive (Table 15).

### **Plant height (PH) vs berries per node (BPN)**

The significant and positive correlation ( $r = 0.403^{**}$ ) between plant height and berries per node was largely due to the indirect effect of plant height via length of the longest primary (0.861) and via internode length (0.437) and to a lesser extent via stem girth (0.242). The overall values were brought down to a low though positive and significant correlation ( $r = 0.403^{**}$ ) between plant height and berries per node by the high negative direct effect (-1.17) of plant height on berries per node (Table 15).

### **Length of the longest primary (LLP) vs berries per node (BPN)**

The significant total correlation ( $r = 0.583^{**}$ ) between length of the longest primary with berries per node was predominantly due to direct effect ( $r = 0.934$ ) and indirect effect (0.464) via internode length and to a lesser extent via stem girth (0.245). On the other hand the indirect effect of length of the longest primary via plant height (-1.079) was negative and high (Table 15).

### **Number of primaries per plant (NPP) vs berries per node (BPN)**

The significant positive correlation between the number of primaries per plant (0.389\*\*) with berries per node was mainly due to the positive indirect effect via length of the longest primary (0.305), stem girth (0.299) and to a lesser extent, the direct effect (0.286). On the other hand there was a negative indirect effect via plant height (-0.414) and to a lesser extent via the number of bearing primaries per coffee plant (-0.174) (Table 15).

### **Internode length (IL) vs berries per node (BPN)**

There was a significant positive correlation between internode length with berries per node ( $r = 0.483^{**}$ ) mainly due to the indirect effect via length of the longest primary (0.999) and the independent effect (0.434). On the other hand there was a negative indirect effect via plant height (-1.177) (Table 15).

### **Number of bearing primaries per plant (NBPP) vs berries per node (BPN)**

Results showed that the number of bearing primaries per plant had low though positive and significant correlation with berries per node (0.240\*\*). This was accompanied by positive though low indirect effects via stem girth (0.230), length of the longest primary (0.320) and number of primaries per plant (0.226). On the other hand, the indirect effect through plant height was negative (-0.358) together with the direct effect (-0.22). Generally the relationships of bearing primaries per plant with berries per node were low compared to other variables (Table 15).

## **5.5 Discussion**

Results of study on combining ability, heritability and relationships of growth and yield variables of  $F_1$  coffee hybrids derived from a commercial coffee variety KP423 and selected accessions of Ethiopian coffee collection are discussed. Estimates of genetic

variances and derived statistics of important agronomic traits are essential for efficient plant breeding programs (Milligan *et al.*, 1990).

### **Combining ability**

Analysis of field data for year 1 (Table 10), showed positive and significant effects for specific combining ability (SCA) for variables stem girth, coffee plant height, length of the longest primary, number of primaries per coffee plant, number of berries per node, internode length and number of bearing primaries per coffee plant. General combining ability (GCA) was found to be positive and significant for variables stem girth, coffee plant height, length of the longest primary and internode length while general combining ability for variables number of primaries per plant, number of berries per node and number of bearing primaries per coffee plant were positive but not significant. This suggests that both specific and general combining ability have influence on most of the studied variables. However, the quotient  $\delta^2\text{GCA} / \delta^2\text{SCA}$  showed that for most of the studied variables it is close to 1 hence the influence of both additive and non-additive gene effects are influencing most of the variables studied.

Data analysis for year 2 (Table 11) indicated positive and significant effects for general and specific combining ability for variables stem girth, coffee plant height, longest primary per coffee plant, number of primaries per coffee plant, internode length and number of bearing primaries per coffee plant. The variable average number of berries per node showed positive but not significant effect for both GCA and SCA. This observation almost agreed with the finding of the first year data suggesting that both general and specific combining ability were important in selection for the studied variables in arabica coffee. This observation was also found in previous studies involving arabica coffee (Cilas *et al.*, 1998; Walyaro, 1983).

## Heritability

In this study from the first and second year data (Tables 10 and 11) the general trend showed that broad sense heritability ( $H^2_b$ ) was higher than the narrow sense heritability ( $h^2_n$ ) for variables studied suggesting that heritability of variables stem girth, coffee plant height, length of the longest primary, number of primaries per plant, number of berries per node, internode length and number of bearing primaries per coffee plant were governed by total genetic effect. Categorizing the narrow sense heritability ( $h^2_n$ ) into categories according to Robinson *et al.* (1949) (0-25 %) as low, (25 %-50 %) as medium and (>50 %) as high only variable internode length had high value (53.7 %). The rest of the studied variables had medium values (35.4 % - 49.1 %) (Table 10). Likewise for the second year data (Table 11) narrow sense heritability was low for variable number of berries per node (23.6 %) and high for the variable number of bearing primaries per coffee plant (50.1 %). The rest of the studied variables had medium values (36.5 % - 48.3 %) (Table 11). This implies that in most of the studied variables with the exception of internode length and number of bearing primaries per coffee plant, probably had other genetic phenomenon such as epistasis or dominance which determined effects on genotype hence less predictability. Walyaro (1983) found that in coffee arabica; growth characters such as stem girth, plant height, internode length on primaries and canopy radius had high heritability. Studies on arabica coffee growth characteristics found that canopy diameter, stem girth, average length of the primary branch and average internode length had higher estimate values of heritability in broad sense at early stages of production justifying their use in indirect selection of superior genotypes at early stage for further advancement (Habte *et al.*, 2007). These findings from the previous studies are in agreement with the findings of this study from the first and second year data (Tables 10 and 11).

## Correlations

Simple correlation results from this study for the first and second years (Tables 12 and 13) showed that for the first year, variables stem girth, coffee plant height, length of the longest primary, number of primaries per coffee plant, number of berries per node, internode length and number of bearing primaries per coffee plant were positively correlated with each other. The positive correlations observed between the various variables indicated the possibility of gaining through selection in more than one variable simultaneously (Ferrão *et al.*, 2008).

Genotypic and phenotypic correlations (Table 14) for most of the growth and yield variables studied showed positive and significant correlations with each other indicating the possibility of simultaneous improvement of these variables (Cilas *et al.*, 1998). Variable internode length vs number of bearing primaries per tree showed non-significant genotypic correlation. Non-significant phenotypic correlation was observed on internode length vs number of primaries per tree. Stem girth was found genotypically and phenotypically positive and significantly correlated with berries per node. These findings were also revealed in previous studies in arabica coffee by Wardiana and Pranowo (2014) and Cilas *et al.* (1998). Similar results were recorded by Marandu *et al.* (2004) in robusta coffee. Because internode length vs bearing primaries per plant (NBPP vs IL) and number of primaries per plant vs internode length (NPP vs IL) were consistently not significantly correlated, these cannot with certainty be improved/selected simultaneously in breeding work (Tables 12 and 13).

Some relationships were not stable across years, suggesting that they are waived by environmental changes hence unpredictable. Such relationships include plant height vs number of primaries per plant and berries per node, length of the longest primary vs

number of primaries and berries per node, berries per node vs internode length and bearing primaries per plant (Tables 12 and 13).

Some relationships are predominantly due to genetic causes since the phenotypic correlation was not significant while the genotypic correlation was significant. This was typical for number of primaries per plant vs internode length (NPP vs IL) suggesting that the underlying cause could be genetic linkage or pleiotropy (Table 14). On the other hand, number of bearing primaries per plant vs internode length (NBPP vs IL) had significant and positive phenotypic correlation with non-significant genetic correlation implying that the relationship is largely due to environmental factors (Table 14).

### **Path analysis**

Path analysis (Fig. 10) and (Table 15) showed that stem girth interacted negatively with plant height in influencing berries per node while, stem girth interacted positively with length of the longest primary in influencing berries per node. Stem girth could be a good indicator for berries per node since it was significantly and positively correlated with berries per node while the direct effect (0.356) was high and positive. Had it not been for the negative indirect influence (-0.794) of stem girth via plant height, the correlation between stem girth and berries per node would have been higher.

Studies done previously established that stem diameter was an important character in selection of coffee plants, because it was positively correlated with yield characters in robusta coffee (Marandu *et al.*, 2004). Similarly Montagnon *et al.* (2001) found that selection for vigorous coffee plants can be done based on stem girth and plant height. Results of the present investigation also indicate the importance of stem girth in selection for yield. Positive relationship between variable stem diameter and others was found

influencing positively the yields in robusta coffee (Nikhila *et al.*, 2008). Wardiana and Pranowo (2014) in their study involving *Coffea arabica* concluded that variable stem girth can be used as positive selection criteria for high yielding coffee genotypes at early generations.

Holding other variables constant, plant height had a detrimental independent contribution on berries per node. However, due to the favourable (positive) interaction of plant height with other variables viz length of the longest primary and internode length these compensated the relationship to a positive one between plant height and berries per node pointing the importance of compensation mechanisms in a system of variable interrelations.

Holding other variables constant, the length of the longest primary would have a positive independent contribution on berries per node. The length of the longest primary interacted positively with internode length and to a lesser extent stem girth in influencing berries per node. On the other hand the unfavorable (negative) interaction between the length of the longest primary with plant height (-1.079) compensated the positive influences to a low but positive and significant relationship ( $r = 0.583^{**}$ ) of length of the longest primary with berries per node. Had it not been for the large negative indirect effect via plant height (-1.079), the correlation between length of the longest primary with berries per node would have been higher than the observed. Studies done by Nikhila *et al.* (2008) established that positive relationship in length of the primary branch influenced positively the yields in robusta coffee.

The positive relationship between primaries per plant with berries per node was predominantly due to the independent contribution and the positive interactions between

primaries per plant with stem girth, plant height and length of the longest primary. However primaries per plant interacted negatively with plant height and to a lesser extent bearing primaries per plant in influencing berries per node and the former was due to the negative direct influence of plant height on berries per node. Had it not been for these negative interactions, the correlation between primaries per plant with berries per node would have been higher.

The significant positive correlation between internode length with berries per node was largely caused by the independent effect. Holding other variables constant, internode length would still increase berries per node. However, the strong unfavorable interaction (-1.177) between internode length with plant height was compensated by the strong positive indirect effect (0.999) via length of the longest primary to a low though positive and significant correlation between internode length with berries per node.

Holding other variables constant, the number of bearing primaries per plant would reduce berries per node due to its negative independent effect (-0.22). The low but positive and significant relationship ( $r = 0.240^{**}$ ), of the number of bearing primaries with berries per node was largely due to indirect effect via length of the longest primary (0.320), stem girth (0.230) and number of primaries per plant (0.226). Compensation mechanisms through the direct effect and indirect effect via plant height (-0.358) contributed in reducing the positive indirect effects of other variables to a low though positive and significant relationship of bearing primaries per plant with berries per node.

## 5.6 Conclusions

The current study demonstrates that combining ability and correlations as estimated by a half diallel and path coefficient analysis were effective in identifying variables for yield

improvement in coffee. Stem girth, length of the longest primaries and internode length may be used in development of superior Arabica coffee varieties involving the commercial variety KP423 and selected Ethiopian genotypes. Positive and significant general combining ability and specific combining ability were observed for all studied variables except for variable berries per node. This suggested the importance of both additive and non-additive components in inheritance of these variables. Broad sense heritability values were generally higher than the narrow sense heritability for all studied variables across years. Variable stem girth had positive and significant correlations with all studied variables across years. Most of the studied variables had positive and significant genotypic and phenotypic correlations. Path analysis revealed that variable length of the longest primary per plant contributed most positively to yield while variable plant height contributed most negatively to yield.

### **5.7 Recommendations**

- i. Growth and yield variables should be used in early years for selection of superior lines for advanced evaluation of superior varieties of coffee.
- ii. In superior coffee varieties development and evaluation for release to growers it is important to consider both genetic and environmental factors
- iii. Variables stem girth, length of the longest primaries and internode length should be used in selection of superior Arabica coffee varieties involving the commercial variety KP423 and selected genotypes from Ethiopian arabica coffee collection.

## References

- Cilas, C., Bouharmont, P., Boccara, M., Eskes, A. B. Baradat, P. (1998). Prediction of genetic value for coffee production in *Coffea arabica* from a half-diallel with lines and hybrids. *Euphytica* 104: 49 – 59.
- Dewey, D.K. and Lu, K.M. (1959). A correlation and path coefficient analysis of components of crested wheat grass seed production. *Agronomy Journal* 51: 515 – 518.
- Ferrão, R. G., Ferreira, A., Cruz, C. D., Cecon, P. R., Ferrão, M. A. G., Almeida da Fonseca, A. F., Carneiro, P. C. and Silva, M. F. (2008). Inter-trait relations for direct and indirect selection in coffee. *Crop Breeding and Applied Biotechnology* 8: 271 – 278.
- Griffing, B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biology Science* 9(4): 463 – 493.
- Habte, E., Bellachew, B. and Singh, H. (2007). Early bearing stage evaluation of *Coffea arabica* L. germplasm for yield and morphological characters. In: *21st International Conference on Coffee Science*, Montpellier, France, held on 11-15 September, 2006. pp. 981 – 984.
- Marandu, E. F. T., Reuben, S. O. W. M and Misangu, R. N. (2004). Genotypic correlations and paths of influence among components of yield in selected Robusta coffee (*Coffea canephora* L) clones. *West African Journal Applied Ecology* 5:11 – 20.
- Milligan, S. B., Gravois, K. A., Bischoff, K. P. and Martin, F. A. (1990). Crop effects on genetic relationships among sugarcane traits. *Crop Science* 30: 927 – 931.

- Mistro, J. C., Fazuoli, L. C., Gallo, P. B., de Oliveira, A. C. B., Toma-Braghini, M. and Silvarolla, M. B. (2007). Estimates of genetic parameters in Arabic coffee derived from the Timor hybrid. *Crop Breeding and Applied Biotechnology* 7: 141 – 147.
- Montagnon, C., Flori, A and Cilas, C. (2001). A new method to assess competition in coffee clonal trials with single-tree plot in Cote d'Ivoire. *Agronomy Journal* 93: 227 – 231.
- Nikhila, K. R., Sureskhumar, V. B., Mohanan, K. V. and Santaran, A. (2008). Association of agronomic characters in robusta coffee (*Coffea canephora* Pieree ex Froehner). *International Journal of Plant Breeding and Genetic* 2(1): 47 – 50.
- Robinson, H. F., Comstock, R. E. and Harvey, P. H. (1949). Estimates of heritability and the degree of dominance in corn. *Agronomy Journal* 41: 353 – 359.
- Sheoran, O. P., Tonk, D. S., Kaushik, L. S., Hasija, R. C. and Pannu, R. S. (1998). Statistical software package for agricultural research workers. *Recent Advances in information theory, Statistics and Computer Applications*. (Edited by Hooda, D. S. and R. C. Hasija, R. C.), Department of Mathematics Statistics, Hisar. pp. 139 – 143.
- Sureshkumar, V. B., Nikhila, K. R., Prakash, N. S. and Mohanan, K. V. (2013). Interrelationship and association of characters in robusta coffee (*Coffea canephora* var. *robusta*). *Agriculture, Forestry and Fisheries* 2(2): 98 – 104.
- TaCRI (2011). Coffee research for development 2001-2011. *A Decade of Contributing to a Profitable and Sustainable Coffee Industry in Tanzania*. (Edited by Teri, J. M., Kilambo, D. L., Maro, G. P., Magesa, J. M. and Van der, Vossen, H. A. M.), Tanzania Coffee Research Institute, Kilimanjaro. pp. 9 – 71.

- Walyaro, D. J. (1983). Considerations in breeding for improved yield and quality in Arabica coffee (*Coffea arabica* L.). Thesis for Award of PhD Degree at Agricultural University, Wageningen, 121pp.
- Walyaro, D. J. and Vossen, H. A. M. (1977). Pollen longevity and artificial cross-pollination in *Coffea arabica* L. *Euphytica* 26(1): 225 – 231.
- Wardiana, E and Pranowo, D. (2014). Selection of vegetative and generative characters of arabica coffee by using sequential path analysis and structural equation models. *Journal Littri* 20(2): 77 – 86.

## CHAPTER SIX

### 6.0 Conclusions, Recommendations and Future Studies

The current study met its overall objective by generating adequate information which will enable the breeders to develop arabica coffee varieties with broad genetic base with resistance to coffee berry disease at a reduced time. High genetic diversity revealed in the germplasm material is not reflected in the commercially cultivated varieties hence a need to utilize the revealed diversity in development of new superior varieties. Possibility of using marker tool in screening for CBD resistance will enhance and reduce time needed for conventional screening hence reduce time of evaluation and release of new superior arabica coffee varieties. The revealed associations of growth and yield variables will enable early selection of superior breeding lines for further field evaluation hence reduce time needed for new superior variety release.

### 6.1 Conclusions

1. The present study revealed high heterozygosity among the arabica coffee genotypes studied. Number of alleles identified was high ranging from 2-8 with an average of 2.5 per marker. Mean polymorphic information content was moderate while Ethiopian genotypes were the most variable group.

Microsatellite marker *ssrAY2449* was the most informative

2. The presence of the coffee berry disease (CBD) resistance genes was revealed in all the studied arabica coffee genotypes amplified by SSR marker Sat 235 and Sat 207. This was confirmed by production of bands similar to the progenitors of CBD resistance. This findings implied that marker screening can be used in coffee berry

disease resistant genotypes selection at early stages of growth hence reducing the time of selection cycle.

3. The current study demonstrates that combining ability, correlations and heritability as estimated by a half diallel and path coefficients analysis was effective in identifying variables such as stem girth, length of the longest primaries and internode length which may be improved in development of superior Arabica coffee varieties involving the commercial variety KP423 and selected Ethiopian genotypes. Variables such as length of the longest primary (canopy radius), internode length and stem girth can be used as selection criteria in selection of superior coffee arabica varieties involving Ethiopian coffee genotypes and commercial coffee variety KP423.

## **6.2 Recommendations**

- Ethiopian arabica coffee collection is a source of high genetic diversity and a valuable source of genes for broadening the genetic base of the commercially cultivated coffee arabica. Hence should be fully utilized in the breeding programme.
- Microsatellite marker *ssrAY2449* should be used in future studies when a need arises in differentiating improved, germplasm or commercially grown varieties involving the studied genotypes.
- Microsatellite markers *Sat 235* and *Sat 207* should be used in early coffee berry disease resistance screening as they were able to differentiate resistant and susceptible genotypes.
- Variables stem girth, length of the longest primary or canopy radius and internode length should be used in early selection for superior lines for further evaluation in evaluation trials of crosses involving commercial varieties and Ethiopian coffee arabica.

### **6.3 Future studies**

1. Further study on the materials studied may be done by use of more precision marker tools like SNP or sequencing to reveal more diversity.
2. Diseases resistance screening at early growth stage is very important especially for perennial crop like coffee. Therefore it is important to conduct a more detailed study involving more markers for the two major diseases of coffee arabica coffee berry disease (CBD) and coffee leaf rust (CLR).
3. More yield variables should be involved in interrelationships study. Current study relied more on growth stage variables.

## APPENDICES

**Appendix 1: List of materials from Ethiopian collection and germplasm used for diversity study**

S/No.	Accessions	Description	Code number
1	F3/64 (E236)	Ethiopian	4
2	F4/64 (E237)	Ethiopian	5
3	F5/64 (E238)	Ethiopian	6
4	F6/64 (E239)	Ethiopian	7
5	F14/64 (E247)	Ethiopian	8
6	F15/64 (E248)	Ethiopian	9
7	F16/64(E249)	Ethiopian	10
8	F17/64 (E250)	Ethiopian	11
9	F18/64(E251)	Ethiopian	12
10	F19/64 (E252)	Ethiopian	13
11	F20/64 (E253)	Ethiopian	14
12	F21/64 (E253)	Ethiopian	15
13	F22/64 (E255)	Ethiopian	16
14	F23/64 (E256)	Ethiopian	17
15	F24/64 (E257)	Ethiopian	18
16	F40/64 (E273)	Ethiopian	19
17	F44/64 (E89)	Ethiopian	20
18	F45/64 (E87)	Ethiopian	21
19	F50/64 (NA)	Ethiopian	22
20	F52/64 (E98)	Ethiopian	23
21	F59/64 (E126)	Ethiopian	24
22	F73/64 (E579)	Ethiopian	25
13	F74/64 (E7)	Ethiopian	26
24	F81/64 (E38)	Ethiopian	27
25	F83/64 (E52)	Ethiopian	28
26	F88/64 (E67)	Ethiopian	29
27	F89/64 (E68)	Ethiopian	30
28	F90/64 (E71)	Ethiopian	31
29	F99/64 (E114)	Ethiopian	32
30	F102/64(E117)	Ethiopian	33
31	F103/64 (E118)	Ethiopian	34
32	F104/64 (E123)	Ethiopian	35
33	F105/64 (E128)	Ethiopian	36
34	F106/64 (E129)	Ethiopian	37

S/No.	Accessions	Description	Code number
35	F121/64 (E178)	Ethiopian	38
36	F130/64 (E224)	Ethiopian	39
37	F147/64 (E442)	Ethiopian	40
38	F154/64 (E449)	Ethiopian	41
39	F167/64 (E464)	Ethiopian	42
40	F187/64 (E16)	Ethiopian	43
41	F188/64 (E22)	Ethiopian	44
42	F190/64 (E37)	Ethiopian	45
43	F193/64 (E125)	Ethiopian	46
44	F195/64 (E556)	Ethiopian	47
45	F196/64 (E579)	Ethiopian	48
46	Kent KP 162 (H66)	Commercial	1
47	Bourbon N39	Commercial	2
48	Bourbon 197	Commercial	3
49	Kent KP423	Commercial	74
50	Rume sudan VC298	Progenitor	49
51	Rume sudan VC299	Progenitor	50
52	Rume sudan VC510	Progenitor	51
53	Hybrid de Timor VCE 1589	Progenitor	52
54	Hybrid de Timor VCE 1587	Progenitor	53
55	Hybrid de Timor VCE 1593	Progenitor	54
56	Catimor-PRO 127	Progenitor	89
57	Catimor-PNI088	Progenitor	90
58	Catimor-PNI 086	Progenitor	91
59	Sarchimor	Progenitor	88
60	RS x HdT1343	Breeding line	62
61	RS x Illubabor 855	Breeding line	66
62	HdT x N39 x HdT	Breeding line	63
63	RS x KP423	Breeding line	67
64	(N39xHDT) x RS 510	Breeding line	65
65	(N39xHDT) x KP423	Breeding line	64
66	(N39xHDT)x(KP423xHDT)	Breeding line	70
67	(N39xHDT)x(BMJxS6Cioiccie)	Breeding line	69
68	(N39xHDT)x(RS x Geisha)	Breeding line	68
69	45/996 RSxN39 F2	Breeding line	81

S/No.	Accessions	Description	Code number
70	45/1004 RSxN39 F2	Breeding line	79
71	45/901 RSxN39 F2	Breeding line	80
72	N39 x VCE1593 BOURBON x HDT	Breeding line	82
73	N39 x VC298 Bourbon x Rume	Breeding line	84
74	OP715 x VC1217 (Bourbon x Geisha) x Rume	Breeding line	85
75	KP423 xVCE1593 Kent x HdT	Breeding line	83
76	N39xVCE1593 Bourbon x HdT F2	Breeding line	86
77	KP423 x VCE 1593 Kent x HdT F2	Breeding line	87
78	(HdTxN39) x (SL28xRS)	Breeding line	71
79	(N39xHdT) x (HdTxN39) x RS	Breeding line	73
80	PADANG x (HdTxN39) x RS	Breeding line	72
81	PNI 088 x N39 x RS	Breeding line	61
82	PNI 088 x (SL34 x HDT) x KENT x RS	Breeding line	57
83	PNI 088 x (PADANG x (HDT x N39)) x RS VC506	Breeding line	55
84	PNI 088 x SL34 x (HDT x N39)x RS	Breeding line	60
85	PRO127 x N39 x RS VC298 Selfed F2	Breeding line	58
86	PRO127 x (N39 x HDT) x HDT	Breeding line	59
87	PNI 086 x N39 x RS VC298 Selfed F2	Breeding line	56
88	(PADANG x (KENT x HdT) x PNI 089	Breeding line	78
89	N39 x VC298 (HdTxN39)xSL28x(N39xRS)x PNI 089	Breeding line	75
90	(N39 x HdT) x (N39 x HdT) x RS x PNI 089	Breeding line	77
91	PNI 086 x (N39 x RS) F2	Breeding line	76

**Appendix 2: List of SSR markers used in diversity study**

S/No	Marker name	Primer (5' - 3') Sequence	PCR size range (bp)	Reference
1	ssrA8783	F CTTCGTATGGTTGTCTGTGT	106-126	Rovelli <i>et al.</i> , 2000
		R AATGATAGGAGGCACTTGAC		
2	ssrA8837	F AAAAGTGAGCACGTCATGTG	148-165	Rovelli <i>et al.</i> , 2000
		R GCGTGAGAGGGACCAT		
3	ssrA8847	F GCACACATGAAAAAGATGCT	159-192	Rovelli <i>et al.</i> , 2000
		R GATGGACAGGAGTTGATGG		
4	ssr AY2434	F CGCAAATGTTTATGTCAATC	178-199	Cristancho <i>et al.</i> , 2002
		R GCAACTTATGAGCCTAATCC		
5	ssrAY2449	F CGAAAATATGCTGCCCATG	273-294	Cristancho <i>et al.</i> , 2002
		R CCGAACCATAAGGTGTGAC		
6	ssrZAP25	F GCGAAATCTTCTCCCTCCC	185-193	Combes <i>et al.</i> , 2000
		R CCGTCCTTTTCTCGAACTC		
7	ssrCMA008	F CATTCTGGTCCCTGATGCTCT	106-128	Teressa <i>et al.</i> , 2010
		R TCATTCACTTATTAACGTCCATC		
8	ssrCMA055	F TTGAGCAAAAACCCTATTCC	82-97	Teressa <i>et al.</i> , 2010
		R TAAACCCAAAAAGACCACAA		
9	ssrCMA059	F GATGGACAGGAGTTGATGGT	129-165	Teressa <i>et al.</i> , 2010
		R TTTTAACACTCATTTTGCCAAAT		
10	ssrCMA151	F GCCAGAAGAAGCTGGATGAC	168-177	Teressa <i>et al.</i> , 2010
		R ACCGTCCTTTTCTCGAACT		
11	ssrCMA198	F AGCAACTCCAGTCCCTCAGGT	195-236	Teressa <i>et al.</i> , 2010
		R TGGAAGCCCGCATATAGTTT		
12	ssrCMA199	F CATGCCATCATCAATCCAT	122-153	Teressa <i>et al.</i> , 2010
		R CTAGCTAGCTGGATCAGTACCC		
13	ssrCMA233	F CAACGAGATAACTGGCAGGTC	255-270	Teressa <i>et al.</i> , 2010
		R CAAACCAATATTAGGAATAAAGAAC		
14	ssrCMA263	F TGCTTGGTATCCTCACATCA	178-200	Teressa <i>et al.</i> , 2010
		R ATCCAATGGAGTGTGTTGCT		
15	ssrR105	F CACCAATCCACTGACAATG	187-222	Teressa <i>et al.</i> , 2010
		R TCCCTGCCAACACACTTC		
16	ssrR126	F GCACAATCACTCCCAAAG	206-242	Teressa <i>et al.</i> , 2010
		R TGACGGCCTACTACTTACAG		
17	ssrR175	F GCAGTGACGCAGCAATG	214-217	Teressa <i>et al.</i> , 2010
		R AAAAGGAGAGCCAAAGCAGT		
18	ssrR209	F CGGGGGTAAAAAGATTGTAA	161-173	Teressa <i>et al.</i> , 2010
		R TTGGTGGGAGGGGAGTA		
19	ssrR268	F GTATCCCACAATGAAATCAC	131-147	Teressa <i>et al.</i> , 2010
		R AGTAGAATTTTCAACATATAAG		
20	ssrR278	F TGTAGATTTGAAACCCAATC	123-141	Teressa <i>et al.</i> , 2010
		R AAGTCTCGACAAGTTTTGAC		
21	ssrR325	F CCTTGTGTGGGGAATGTC	224-262	Teressa <i>et al.</i> , 2010
		R GGCTGTTCTGGGCTTTGTG		

S/No	Marker name	Primer (5' - 3') Sequence	PCR size range (bp)	Reference
22	ssrR338	F CGAAGGCTGTCAACAACTGG	221-233	Teresa <i>et al.</i> , 2010
		R GGGATAAACAAGTTAAAGGA		
23	ssrR339	F ATTATGCTCGCTGGGCTGTT	215-226	Teresa <i>et al.</i> , 2010
		R TGGGATCACTCCTGTGTCGC		
24	ssrCM2	F TGTGATGCCATTAGCCTAGC	177-229	Baruah <i>et al.</i> , 2003
		R TCCAACATGTGCTGGTGATT		
25	ssrCM	F GTAACCACCACCTCCTCTGC	167-230	Baruah <i>et al.</i> , 2003
		R TGGAGGTAACGGAAGCTCTG		
26	ssrCM6	F GCTAAGTTCAATTGCCCTGT	210-232	Baruah <i>et al.</i> , 2003
		R GGGTTAATTTGATTGCGTGA		
27	ssrCM8	F GCCAATTGTGCAAAGTGCT	159-183	Baruah <i>et al.</i> , 2003
		R ATTCATGGGGCCTTTGTCTT		
28	ssrCM11	F AATCACCTTCGCAAACCAAC	200-254	Baruah <i>et al.</i> , 2003
		R CCGAACGCAATATCTTATGC		
29	ssrCM16	F TGGGGAAAAGAAGGATATAGACAAGAG	95-129	Baruah <i>et al.</i> , 2003
		R GAGGGGGGCTAAGGAATAACATA		
30	ssrCM17	F CCAGCCTTTTCACAATTCTCACCC	308-390	Baruah <i>et al.</i> , 2003
		R TGCCCCCTAGATATGGTACAAGCTTTC		

Source: Eurofins MWG Operon Ebersberg, Germany.