Collection, Morphological and Molecular Characterization of Papaya Germplasm in Kenya.

George Ochieng' Asudi

A Thesis Submitted in Partial Fulfillment for the Degree of Master of Science in Biotechnology in the Jomo Kenyatta University of Agriculture and Technology

2010

DECLARATION

This thesis is my own original work and has not been submitted for a degree in any other university.

Signature……………………………………Date…………………………

George Ochieng' Asudi

This thesis has been submitted for examination with our approval as university supervisors.

Signature……………………………………Date…………………………

Dr. Fredah K. Rimberia Wanzala

JKUAT, Kenya

Signature……………………………………Date…………………………

Dr. Aggrey B. Nyende

JKUAT, Kenya

Signature……………………………………Date…………………………

Dr. Elijah M. Ateka

JKUAT, Kenya

DEDICATION

To my late beloved mother, Hellen Asudi, may her soul rest in peace.

ACKNOWLEDGEMENTS

I would like to express my deep appreciation and gratitude to my supervisors Dr. F. K. Rimberia Wanzala, Dr. A. B. Nyende and Dr. E. M. Ateka for their guidance, encouragement, advice, patience and moral support during the period of research work. My deep appreciation also goes to the research, production and extension (RPE) division of Jomo Kenyatta University for Agriculture and Technology (JKUAT) for funding this research through the papaya research project based in Horticulture Department, JKUAT.

Thanks to Prof. L. S. Wamocho of Masinde Muliro University of Science and Technology, Department of Sugar Technology, for assisting and guiding me during germplasm collection. I am also grateful to Dr. L. A. Oyugi of Jomo Kenyatta University for Agriculture and Technology Department of Commerce and Economics studies for guiding with the questionnaire writing and data analysis.

I would also like to thank my friends and papaya farmers in Coast, Central, Nyanza, Rift Valley, Eastern and Western provinces for their hospitability during our visits to their farms without which, this work could not have been accomplished. I wish to express my gratitude to Ms. H. W. Karuri for her assistance with laboratory work and statistical analysis. My very special thanks are extended to Biosciences Eastern and Central Africa (BECA) staff, M. Kitavi, and L. Muthui for their constant guidance and technical expertise in genotyping analysis at International Livestock Research Institute (ILRI). I would like to express a deep sense of gratitude to Mr. P. Mwirigi for his help with morphological statistical analysis. I also express my gratitude to all my fellow laboratory

mates and students at Institute for Biotechnology Research (IBR), Makonde, Lily, Julia, Hul, Doris, Kevo, and Miriti for their help and suggestions. A special word of thanks to Mr. F. Ombwara for the encouragement and helping with my work during field visits. I would wish to express my appreciation to my relatives for their love, patience and constant support which made me competent to achieve my goals in life. A special thanks to Silas, Peter, Evans and Fred. Finally, to God for he has brought me this far.

TABLE OF CONTENTS

DECLARATION ..... **i**

DEDICATION ..... **ii**

ACKNOWLEDGEMENTS ..... **iii**

TABLE OF CONTENTS ..... **v**

LIST OF TABLES ..... **ix**

LIST OF FIGURES ..... **x**

LIST OF PLATES ..... **xi**

LIST OF APPENDICES ..... **xii**

LIST OF ABBREVIATIONS ..... **xiii**

ABSTRACT ..... **xv**

CHAPTER ONE **1**

1.0 INTRODUCTION **1**

1.1 General introduction 1

1.2 Statement of the problem and justification 4

1.3 General objective 6

1.3.1 Specific objectives 7

1.4 Hypothesis 7

CHAPTER TWO **8**

2.0 LITERATURE REVIEW **8**

2.1 Description of papaya 8

2.2 Origin and distribution of papaya 9

2.3 Uses and composition and papaya 10
2.4 The genome of papaya 10
2.5 Papaya production in Kenya 11
2.6 Papaya production constraints 12
2.7 Papaya morphology 13
2.9 Characterization of papaya germplasm 15
2.9.1 Use of morphological traits 15
2.9.2 Use of molecular markers 16
2.10 Gel electrophoresis 20
2.11 Capillary electrophoresis 21
CHAPTER THREE 23
3.0 COLLECTION AND DOCUMENTATION OF KENYAN PAPAYA GERMPLASM 23
3.1 Introduction 23
3.2 Materials and methods 24
3.2.1 Collection sites of papaya germplasm 24
3.2.2 Sampling 25
3.2.3 Data collection 26
3.2.4 Data analysis 26
3.3 Results 27
3.4 Discussion 37
3.5 Conclusion and recommendations 40

CHAPTER FOUR **42**

4.0 MORPHOLOGICAL CHARACTERIZATION OF KENYAN PAPAYA GERMPLASM **42**

4.1 Introduction 42

4.2 Materials and Methods 43

4.2.1 Morphological characterization of papaya germplasm 43

4.2.2 Data collection 43

4.2.3 Data analysis 44

4.3 Results 45

4.3.1 Morphological diversity in the papaya germplasm in Kenya 45

4.3.2 Principal component analysis 46

4.3.3 Cluster analysis 48

4.4 Discussion 51

4.5 Conclusion and recommendations 53

CHAPTER FIVE **55**

5.0 CHARACTERIZATION OF KENYAN PAPAYA GERMPLASM USING SSR MARKERS **55**

5.1 Introduction 55

5.2 Materials and methods 56

5.2.1 Sample collection 56

5.2.2 DNA extraction from papaya 56

5.2.3 DNA quantity and quality determination 58

5.2.4 Polymerase chain reaction (PCR) with SSR primers 59
5.2.5 Agarose gel electrophoresis of amplified products 62
5.2.6 Capillary electrophoresis.. 62
5.2.7 Fragment analysis 63
5.2.8 Data analysis.. 64
5.3 Results 64
5.3.1 Polymorphism of microsatellites used to characterize papaya accessions 64
5.3.2 Marker quality indices, allele variability, polymorphism and observed heterozygosity in the 42 accessions 65
5.3.3 Genetic variability within the 42 Papaya accessions 66
5.3.3.1 Number and frequency of alleles 66
5.3.3.2 Polymorphism of the 42 papaya accessions 69
5.3.4 Genetic relationships among the 42 papaya accessions.................................. 69
5.3.5 Phylogenetic analysis... 72
5.3.6 Principal component analysis (PCA) based on genetic distance estimates of 42 papaya accessions 74
5.4 Discussion... 75
5.4 Conclusion and recommendations 78
CHAPTER SIX 79
6.0 GENERAL DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS 79
7.0 REFERENCES.. 83
APPENDICES... 93

LIST OF TABLES

Table 1. Locations of the various farms studied during germplasm collection 25

Table 2. Occurrence of accessions in the sampled districts................................. 27

Table 3. Distribution of papaya accessions in the districts.................................. 29

Table 4. Papaya accessions collected and their major morphological descriptors 30

Table 5. Scales of production of papaya in the studied districts........................... 33

Table 6. Papaya propagation materials in the sampled districts 34

Table 7. Constraints to papaya production reported in Kenya............................. 36

Table 8. Principal component analysis (PCA) performed using the XLSTAT 2008 statistical package showing the correlations of the first nine principal components with the variables observed on papaya individuals. 46

Table 9. PCR optimization conditions of the seven SSR primers......................... 60

Table 10. *Carica papaya* microsatellite primers used in the study......................... 61

Table 11. Quality indices and polymorphism detected by 7 SSR markers in the 42 papaya accessions 66

Table 12. Frequency of alleles detected in 42 papaya accessions with 7 SSR markers.. 68

Table 13. Number of accessions and their corresponding percentage polymorphism across the seven SSR loci.. 69

Table 14. Similarity matrix among papaya accessions 71

LIST OF FIGURES

Figure 1. Distribution of variates in PC1 and PC2. PC1 accounts for 20.98% of the variation while PC2 accounts for 12.28%).. 47

Figure 2. Correlation among characters associated with the first and second Principal Components. 48

Figure 3. Dendrogram constructed based on morphological characters of 60 papaya accessions from six provinces 50

Figure 4. Electrophenogram showing homozygosity of accession MF1 using SSR marker mCpCIR3.. 65

Figure 5. Electrophenogram showing heterozygosity of papaya accession IMA8 using SSR marker mCpCIR1.. 65

Figure 6. UPGMA dendrogram of 42 Kenyan papaya accessions based on genetic similarity matrix calculated from SSR markers 73

Figure 7. Principal component analysis of 42 papaya accessions using the seven SSR markers 74

LIST OF PLATES

Plate 1. Constraints to papaya production observed in Kenya 36

Plate 2. Morphological diversity observed in the Kenyan papaya germplasm............. 45

LIST OF APPENDICES

Appendix 1. Questionnaire used to garner information on Kenyan papaya germplasm...... 93

Appendix 2. Papaya morphological characters used in phenotypic evaluation 95

Appendix 3. Allele sizes (base pairs) of 42 papaya accessions... 97

LIST OF ABBREVIATIONS

AB	Applied Biosystems
ABSPII	Agricultural Biotechnology Support Project II
AFLP	Amplified fragment length polymorphism
OGTR	Australian Government Office of the Gene Technology Regulator
CTAB	Cetyl trimethyl ammonium bromide
CIMMYT	International Maize and Wheat Improvement Center
DNA	Deoxyribonucleic acid
dNTP	2', 3'-deoxyribonucleoside 5'-triphosphate
EDTA	Ethylenediamine tetra-acetic acid
GPS	Global positioning system
IBPGR	International Board for Plant Genetic Resources
ILRI	International Livestock Research Institute
JKUAT	Jomo Kenyatta University of Agriculture and Technology
NRC	National Research Council
NTSYS	Numerical taxonomy multivariate analysis system
OECD	Organization for Economic Co-operation and Development
PCA	Principal Component Analysis
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphic DNA
RNAse	Ribonuclease
RFLP	Restriction fragment length polymorphism

SSR	Simple sequence repeats
Taq	Thermophilus aquaticus
TBE	Tris borate EDTA
Tris	Tris (hydroxymethyl) aminomethane
UPGMA	Unweighted pair group method of arithmetic averages

ABSTRACT

Papaya is one of the major fruit crops of the tropical regions of the world. It shows considerable phenotypic variation in morphological and horticultural traits that can be utilized in its genetic improvement. The objectives of this research were to collect, document and characterize the Kenyan papaya germplasm.

A survey conducted to garner information on papaya germplasm and production in six Kenyan provinces was based upon a structured questionnaire supplemented by oral interviews. The survey was carried out between June and September 2008. The papaya germplasm was collected from Coast, Nyanza, Western, Rift Valley, Eastern and Central provinces of Kenya and characterized in the field using IBPGR (International Board of Plant Genetic Resources) morphological descriptors based on fruit, flower, stem and leaf characteristics. The morphological characteristics were recorded and morphological data from sixty accessions submitted to principal component and Neighbor-Joining cluster analysis. The genetic diversity of 42 papaya accessions from the above named six provinces of Kenya was also investigated using seven simple sequence repeats (SSR) markers. The amplified DNA fragments were screened by capillary electrophoresis on the ABI 3730 genetic analyzer and analyzed using the Genemapper v3.7 software. Cluster and principal component analysis (PCA) were done using NTSYS-pc v2.2.

Data from the field survey indicated that majority of Kenyan papaya growers have medium sized farms (0.2-0.8 ha), upon which they conduct mixed cropping. While growers on small (less than 0.2 ha) and medium sized farms indicated no varietal

preferences, the growers on large farms (larger than 0.8 ha) preferred specific varieties of known performance and traits such as Solo, Sunrise, US, and Redlady. Most farmers produced papaya for both subsistence and market. Constraints to papaya production included seedling sex paradox, pests such as stink bugs and spider mites, viral diseases (papaya ringspot virus), and the lack of clean planting material.

Accessions from Coastal, Western, Rift Valley and Nyanza provinces showed the widest morphological diversity while those from Eastern and Central provinces showed the least diversity. In total, 7 SSR markers used in the analysis were highly polymorphic among the accessions and the polymorphic information content (PIC) varied from 0.75 to 0.852 with an average of 0.81. The number of alleles within the 42 papaya accessions across the seven loci ranged from 8 to 18 with an average of 11.93. The phylogenetic analysis clustered the 42 accessions into two main clusters A and B. Cluster A had four sub-clusters while cluster B had no sub-clusters. Microsatellite markers used therefore, showed limited genetic diversity among papaya accessions. However, the Coast province accessions showed the widest diversity, as they were scattered all over the dendrogram.

The morphological and genetic differences among the genotypes revealed by their clustering into distinct groups suggest the presence of different sources of variations among the papaya accessions. This could be attributed to their diversity, geographical locations and also due to exchange of plant genetic resources among farmers within and between the provinces. The high morphological diversity observed within the accessions

points to ample possibilities of obtaining desirable trait combinations in Kenyan papaya. Knowledge of this diversity together with the documentation of cultural practices is also fundamental for the future improvement of Kenyan papaya germplasm. The wide diversity observed among accessions from Coastal, Western, Rift Valley and Nyanza provinces can be utilized in the selection of promising parents in hybrid variety, inbred line development and estimating the potential of genetic gain in a breeding programme. There is also need for proper conservation of the different accessions reported as they could serve as raw material for the genetic improvement of different characters of the crop through recurrent selection after hybridisation.

Key words: Collection, diversity, documentation, germplasm, morphological markers, papaya, SSR markers.

CHAPTER ONE

1.0 INTRODUCTION

1.1 General introduction

Papaya is a medium sized fruit crop with a potential to produce fruits throughout the year. It is native to tropical America (Samson, 1986: Nakasone and Paull, 1998), but is now cultivated fruit crop in tropical and subtropical regions worldwide. The major producers of papaya are Brazil, Indonesia, India, Mexico, Hawaii, Zaire, and Nigeria (Samson, 1986). Until recently, Tanzania was the main producer of papain (Nakasone and Paull, 1998) and the chief importer was the United States (Samson, 1986). It is also a very important crop of Kenya both for local and export markets (Imungi and Wabule, 1990). It is polygamous species with three basic sex types: female, male, and hermaphrodite (Storey, 1938; Hofmyer, 1938). Both females and hermaphrodites are commercially useful for fruit production, while the males are solely used as a pollen source. The ripe fruits are popularly used as dessert or processed into jam (Samson, 1986: Nakasone and Paull, 1998), puree or wine while the green fruits can be cooked as vegetables. Latex from green fruits contains papain, a proteolytic enzyme, which is used in pharmaceutical, cosmetics and garment industries (Ali and Lazan, 1998; Nakasone and Paull, 1998). Thus papaya is a good source of nutritious food as well as income for the producers.

Due to its nature as a single stemmed tree it can be economically produced in any size of land from kitchen gardens to large plantations. The genus *Carica* is endowed with enormous morphological variation and is also adapted to a wide range of environments. Under optimum conditions, the growth and development of papaya proceeds at a fast rate (Samson, 1986).

Papaya was introduced into Kenya during the colonial times. The main variety, 'Solo' was introduced from Hawaii. Other varieties like Cavite, 417, 418, 455 and 457 were introduced from Philippines, India (2nd-4th) and Indonesia, respectively (Kamau *et al*., 1993). Redlady, sunrise papaya, US, mountain, honey dew, Kapoho solo, Kitale, Malindi, Waimanalo, PP1, Kiru, Higgins, Wilder, and Heas 7812 (Imungi and Wabule, 1990; Griesbach, 1992), have also been reported in Kenya. There are two main types of papaya cultivars; dioecious cultivars better adapted to subtropical climates and the hermaphrodite cultivars adapted to more tropical areas. The dioecious plant also requires adequate pollination in order to set fruit. In tropical and sub-tropical climates, fruit set occurs throughout the year (Martins, 2003). Phenotypic variation among these varieties is enormous ranging from small-sized fruits in solo variety to relatively large fruits found in the Kiru variety. Fruits are oblong to nearly spherical in female trees and pear-shaped, cylindrical or grooved if formed from hermaphrodite trees (Samson, 1986). In addition, wild germplasm grows spontaneously in many places (d'Eeckenbrugge *et al*., 2007) including Kenya and therefore exhibit wide morphological variation. However, despite this considerable phenotypic variation in morphological and horticultural traits there are no published studies of morphological diversity in the species of Kenyan

papaya. A better knowledge of the morphological and genetic diversity of Kenyan papaya is therefore, desirable.

The extent of genetic variation in a species and its distribution among populations is determined by factors such as the mating system, the demographic history, the effective population size and the extent of gene flow by migration or through seed dispersal between populations (Weising *et al*., 2005). Knowledge of this genetic variation in genotypes of any crop is necessary to estimate the potential of genetic gain in a breeding program and for germplasm conservation (Sakiyama, 2000). A good understanding of genetic relationships is also critical for the effective organization and management of papaya germplasm. DNA-based genetic markers are increasingly being utilized in varietal identification, cultivar development, quality control of seed production, measurement of genetic diversity for conservation management, and intellectual property protection (Smith *et al*., 2000). Previously reported methods to characterize papaya include morphological, isozyme, random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP) and PCR-RFLP markers (Stiles e*t al*., 1993; Morshidi, 1998; Aradhya *et al*., 1999; Kim et al., 2002; Van Droogenbroeck *et al*., 2002, 2004). Most of these studies showed limited genetic diversity in the common papaya (Ocampo *et al*., 2006a). Each of these classes of molecular markers also has characteristics affecting comprehensiveness of genome coverage, discrimination ability, reproducibility, speed, and cost of data generation and scoring that impact applications of the individual technologies by plant breeders and conservators (Smith *et al*., 2000).

Simple sequence repeats (SSR) or microsatellites are flanked by conserved sequences which confer specificity to each microsatellite locus and development of primers for SSRs requires significant resources (Estoup and Turgeon, 1996). However, once SSRs are developed, they can be used to provide profiles that are highly discriminative among genotypes (Senior and Heun, 1993; Smith et al., 1997). Because of their multiallelic nature, microsatellite markers are most efficient for the discrimination of genotypes on a per-locus basis. They can also be mapped to discrete loci (Senior *et al*., 1996), the locations of which are stably inherited within a species, in contrast to the loci revealed by arbitrarily primed methods such as RAPDs. The SSRs used for variety profiling can also provide high discrimination, with excellent reproducibility at less cost than for RFLP (Smith *et al*., 2000).

1.2 Statement of the problem and justification

Production of papaya (*Carica papaya* L.) is particularly attractive for fruit crop diversification in Kenya because of its yield potential, high demand on the local market and potential for export. However, the production in Kenya encounters serious problems including diseases such as *Papaya ringspot potyvirus* (PRSV) (ABSPII, 2004) and bacterial canker caused by *Erwinia papaya,* cause extremely severe impacts in papaya production thus making papaya responding to market standards impossible (Webb, 1985; Gardan *et al.,* 2004). These problems also limit commercial production in most papaya growing areas (Nakasone and Paull, 1998).

Papaya is commercially propagated by seed (Griesbach, 1992) and a lot of changes have occurred as a result of open pollination resulting in loss of identity of the introduced varieties. Indeed, it is possible that new genotypes have arisen from open pollination. Thus, it is difficult to distinguish papaya varieties from different regions of Kenya. Additionally, other wild relatives of the commercial papaya like *Vasconcella spp* are available in some parts of the Kenya highlands. However, little documentation exists on the collection, characterization and documentation of the Kenyan papaya germplasm. Therefore, there is need to document and characterize the existing papaya germplasm in the country in order to inform future improvement programs of this crop.

Genetic diversity is the basis for crop or plant improvement. Information regarding the genetic diversity of available germplasm is vital to devise plant breeding programmes as well as to maintain genetic diversity in a given gene pool. Genetic diversity can be estimated using morphological, biochemical and DNA-based markers. Of these, the morphological method is the oldest and is considered as the first step in the description and classification of germplasm (Smith and Smith, 1989). Morphological characterization is normally accomplished by use of morphological descriptors (IBPGR, 1988). These descriptors are, however, greatly influenced by environment and are subject to individual bias (CIAT, 1993). Thus, more accurate methods of characterization such as DNA based genetic markers are necessary for characterization.

Papaya germplasm in different parts of the world have reportedly been characterized using morphological, isozyme, RAPD, RFLP, AFLP and PCR-RFLP markers (Stiles *et*

al. 1993; Morshidi, 1998; Aradhya *et al*. 1999; Kim *et al*. 2002; Van Droogenbroeck *et al*. 2002, 2004). Most studies showed limited genetic diversity in the common papaya (Ocampo *et al.*, 2006a) and supported the recent separation in two genera by Badillo (2000). Thus genetic characterization is a sure way of differentiating between the available papaya varieties and establishing genetic relationships within the cultivated germplasm. A good understanding of genetic relationships is also critical for the effective organization and management of papaya germplasm. Genetic differences exhibited as presence or absence of polymorphisms between accessions can be combined with phenotypic analyses to augment germplasm characterization. For genetic analysis of papaya genotypes, the SSR markers produce easily scorable, unique alleles, and allele combinations which make them an ideal system for cultivar identification. They have been used successfully for determining the genetic relationship between *Carica papaya* L. cultivars elsewhere (Ocampo *et al.*, 2006a) but not for the Kenyan papaya germplasm. Other PCR-based fingerprints such as RAPD are more discriminatory but often have much less reproducibility (Ghislain *et al*., 2004). This study was therefore designed to determine the extent and distribution of the available diversity in Kenya through collection of papaya germplasm, characterization and its documentation.

1.3 General objective

- To collect, document and characterize papaya germplasm existing in Kenya for increased utility.

1.3.1 Specific objectives

- To collect and document papaya germplasm in Kenya.
- To morphologically characterize the papaya germplasm in Kenya.
- To characterize the papaya germplasm using molecular (SSR) markers.

1.4 Hypothesis

- Papaya germplasm is not widespread in Kenya.
- Phenotypic differences do not exist within the Kenyan papaya germplasm.
- Genetic differences do not exist within the Kenyan papaya germplasm.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Description of papaya

Papaya is a perennial plant species that flowers as early as 3 months after planting and produces fruit within 9 months. It is a trioecious with an intriguing sex determination system and three basic sex types; male, female, and hermaphrodite (Storey, 1938; Hofmyer, 1938; Janick, 1986). The tree, fast growing, and herb-like, is normally single-stemmed, but occasionally lateral branches result from the wounding. Fruit is borne in the axils of the large leaves that open in sequence up the stem as the tree grows. The fruit may be round, pyriform or oval, and ranges in weight from 0.5-7 kg. The skin color of ripe fruit is orange while the flesh is pale orange or light red.

Papaya is commercially propagated by seed (Griesbach, 1992; Nakasone and Paull, 1998). Seeds from hermaphrodite trees always segregate into hermaphrodites and females at the ratio of 2:1 and the sex types of the plants can be determined only by inspection of the flowers. Therefore, it is a general practice for farmers to plant three to five seedlings in one hill, allowing them to grow for 4 to 6 months until the sex types are identified, and then to remove the undesired plants to develop the orchards with only hermaphrodite plants.

2.2 Origin and distribution of papaya

Papaya is a member of the family *Caricaceae,* order *Brassicales* sharing a common ancestor with Arabidopsis. It is in the class Magnoliopsida and subclass *Dicotyledoneae* (Janick, 1986; Nakasone and Paull, 1998). It is indigenous to tropical America but is now grown throughout the tropics. The common papaya is the only species of the genus *Carica*, which very probably originated in Mesoamerica. More precisely, the Caribbean coast of Central America appears to be the most probable centre of origin of papaya (Manshardt and Zee, 1994). Until recently, the family *Caricaceae* was thought to comprise 4 genera and 31 species. Three genera namely; *Carica*, *Jacaratia* and *Jarilla* have their origins from tropical America while one genus *Cylicomorpha* was believed to have originated from equatorial Africa (Nakasone and Paull, 1998). However, a recent taxonomic revision proposed that some species formerly assigned to *Carica* were more appropriately classified in the genus *Vasconcellea* (OGTR, 2008). Accordingly, the family's classification has been revised to comprise *Cylicomorpha* and five South and Central American genera (*Carica*, *Jacaratia*, *Jarilla*, *Horovitzia* and *Vasconcellea*) (OGTR, 2008), with *Carica papaya* the only species within the genus *Carica* (OGTR, 2008). The highland papayas, *Vasconcellea,* are considered the nearest relatives to *Carica papaya* although the relationship is not close (Aradhya *et al*., 1999; Van Droogenbroeck *et al*., 2002). A more recent study (Van Droogenbroeck *et al*., 2004) actually suggested that there are two lineages within *Caricaceae* family and that some members of *Vasconcellea* are more closely allied to *Carica papaya* than others having implications for the successful use of *Vasconcellea* species in hybrid breeding

programmes (OGTR, 2008). The plant also known as pawpaw is a widespread fruit crop throughout Kenya where enough water is available for it to be cultivated. It is grown widely in Kenya on small-holdings as a domestic fruit and small-scale cash crop.

2.3 Uses and composition and papaya

Papaya is a source of papain, a proteolytic enzyme that is used as a beer clarifier and in meat tenderizers. Papain is obtained by collecting and drying the latex exuded from scratches in the surfaces of slightly immature fruit. It is also used to pretreat red blood cells prior to cross matching and to dissolve cartilage-like substances that often develop in disks between vertebrae (Janick, 1986). Latex is also involved in defense of the plant against a wide range of pests and herbivores (El Moussaoui *et al.*, 2001). However, the ripe fruit contains no latex (Villegas, 1997), possibly because the latex-producing cells cease functioning with age. The papaya fruit is popular in many parts of the world as a dessert and is a rich source of vitamins A and C (Imungi and Wabule, 1990). It is also processed into juice, puree and jam (Nakasone and Paull, 1998). The green fruit is rich in potassium, calcium and phosphorous and is cooked as a vegetable in South East Asia (Manshardt, 1992, Nakasone and Paull, 1998). In Kenya, papaya is mainly consumed fresh as dessert or processed into jam or wine.

2.4 The genome of papaya

Papaya is diploid with nine pairs of chromosomes and a small genome size of 372 Mbp (Storey, 1941; Arumuganathan and Earle, 1991) which is advantageous in genetic

analysis (Chen *et al.*, 2007). On the basis of segregation ratios from crosses among three sex types, Storey (1938) and Hofmeyr (1938) proposed that sex determination in papaya is controlled by a single gene with three alleles: M, M^h, and m. The male individuals (Mm) and hermaphrodite individuals (M^hm) are heterozygous, whereas female individuals (mm) are homozygous recessive. The genotypes with homozygous dominant alleles, MM, M^hM^h, and MM^h, are lethal, resulting in a 2:1 segregation of hermaphrodite to female from self-pollinated hermaphroditic seeds and a 1:1 segregation of male to female or hermaphrodite to female from cross-pollinated female seeds.

2.5 Papaya production in Kenya

In Kenya, many varieties of papaya are grown and eaten (Imungi and Wabule, 1990). Although, isolated trees can be seen almost all over the arable parts of the country, the main growing areas are in Embu, Machakos, Meru, Murang'a and Kisii districts. Here, the fruits are mainly either intercropped with other crops or planted along farm boundaries, with most farmers growing not more than just a few trees (Imungi and Wabule, 1990) harvesting throughout the year, and selling the fruit to exporters, in towns and in rural village markets (Martins, 2003).

Papaya is sold and eaten locally as a fresh fruit, with much demand from the numerous hotels, local grocery stores and the town markets. Papaya is dried and exported as part of a dried fruit mixture. The milky latex produced by the unripe fruit is harvested and used in the production of papain, which is a proteinase (Nakasone and Paull, 1998). Papain is a proteolytic enzyme that digests proteins and is used a meat tenderizer, as a digestive

medicine in the pharmaceutical industry and in the brewing industry (Nakasone and Paull, 1998). Coastal people use the latex from unripe fruit to ease the pain and remove the spines and stinging cells of sea-urchins and jellyfish (Martins, 2003). The seeds of *Carica papaya* are dried and exported to health food stores and have shown promise in male contraception (Lohiya *et al*., 2005).

2.6 Papaya production constraints

Papaya producers in Kenya and many other producing countries encounter agronomic, production and marketing challenges. These challenges include unreliable methods of selecting the required sex of seedlings at planting time, lack of disease-free planting materials, lack of improved varieties and outbreak of diseases such as *Papaya ringspot potyvirus* (PRSV) (ABSPII, 2004). In commercial orchards, the fruit tree is cultivated for 3-4 years, after which trees become too tall for economical harvesting and yields are reduced (Griesbach, 1992). *Papaya ringspot potyvirus* (PRSV) and other viruses are common. They affect papaya production and quality, subsequently displaying stunted growth, deformed and inedible fruit, and eventually, cause plant mortality (ABSPII, 2004). Often these fruits are not marketable. As a result papaya is being wiped out at a fast rate (personal observation), and the devastation is already becoming obvious in both research institutions and farmers' fields.

Fungal pathogens of papaya are numerous (Ploetz *et al.,* 1994). *Phytophthora palmivora* cause root, stem and fruit rot which is especially severe during wet seasons. Powdery mildew is caused by *Oidium caricae* which affects the underside of leaves and petioles

(Nakasone and Aragaki, 1973; Nakasone and Paull, 1998). Other fungal pathogens cause papaya leaf spot and include *Cercospora spp, Colletotrichum spp, Curvularia spp, Gloeosporium spp,* and *Corynespora spp*. The major pathogenic nematodes are the reniform nematodes (*Rotylenchus spp*.) and the root knot nematode, *Meloidogyne spp* (Nakasone and Aragaki, 1973; Nakasone and Paull, 1998).

The cultivation of papayas in response to modern export standards has also been made impossible because of the extremely severe impact of a bacterial canker caused by an *Erwinia spp* (Webb, 1985; Gardan, 2004). Similar bacterial diseases have been observed in other places in the world. Root rots can cause rapid death of papaya plants (Nakasone and Paull, 1998). Anthracnose appears on the fruit as irregular, water-soaked spots that later enlarge, darken and become sunken in the rind. Anthracnose is favored by wet weather conditions (Samson, 1986; Griesbach, 1992). Serious pests include fruit flies and mites (N'Guetta, 1994; Borge and Basedow, 1997). Common whitefly (Samson, 1986; Morton, 1987) affects papaya, leading to sooty mold on the foliage and fruit. Sweet potato whitefly also is reported to prefer papaya. Aphids are important only as vectors of PRSV (Nakasone and Paull, 1998).

2.7 Papaya morphology

Papaya germplasm shows considerable phenotypic variation for many horticulturally important traits, including fruit size, fruit shape, flesh color, flavor and sweetness, length of juvenile period, plant stature, stamen carpellody, and carpel abortion (Kim *et al*., 2002). In addition, commercial papaya cultivars may be inbred gynodioecious lines,

typified by the Hawaiian Solo lines, out-crossing dioecious populations, such as the Australian papaws from southern Queensland; F1 hybrids, including the Tainung series (Taiwan), Eksotica II (Malaysia), and Rainbow (Hawaii); or occasionally even clones, such as Hortus Gold in South Africa. (Kim *et al.*, 2002). Solo refers to a group of small-fruited, high sugar content, commercial cultivars developed in Hawaii and originating via introduction from Barbados in 1910 (Storey, 1969).

Many landraces and cultivars present hermaphrodite plants, which are generally preferred for production. They display considerable phenotypic variation for many morphological and horticultural traits (Ocampo *et al.*, 2006b). However, there are few precise data on morphological diversity in papaya in the literature. Somsri (1999) investigated the possibility of using morphological traits to predict sex type. He found plant height, plant height at first flower, leaf shape, shape of petiole sinus and number of nodes to first flower useful for cultivar identification. Plant height and number of nodes at first flower appeared useful for distinguishing between female, hermaphrodite and male hybrid plants, contradicting earlier results presented by Nakasone and Storey (1955) who showed no such effect. The results of Somsri (1999) have also been recently confirmed by Ocampo et al. (2006b) who showed a marked influence of sex on the quantitative traits evaluated, except on those related to stem size (plant height, internode length and stem diameter).

Fruits are technically, classified as fleshy berries (Villegas, 1997) sometimes called pepo-like berries since they resemble melons by having a central seed cavity. They are

borne axillary on the main stem, usually singly but sometimes in clusters (OGTR, 2008). The fruit has a smooth exocarp (peel) and thick, fleshy mesocarp. Fruits production in papaya plants may follow either cross pollination (out-crossing), self-pollination or parthenocarpy, depending on whether dioecious or gynodioecious lines are planted and the particular cultivar that is grown (Nakasone and Paull, 1998).

2.9 Characterization of papaya germplasm

2.9.1 Use of morphological traits

Phenotypic identification of plants is based on morphological traits recorded in the field. It has been used as a powerful tool in the classification of genotypes and to study taxonomic status. Certification of new varieties is based on the genetic stability of a particular crop. However, traditionally these assessments depended on botanical traits (Stegemann, 1984; Zacarias, 1997). The distinctness uniformity stability (DUS) test is carried out as an observation trial, lasting several years, during which primarily morphological characteristics are recorded (Weising *et al*., 2005). The identification of sufficient distinctness from cultivars in the reference collections is becoming a problem in the major horticultural and agricultural crop species. Therefore, the potential use of molecular methods for cultivar description and identification is now being researched for example by Community Plant Variety Office (UPOV) (Weising *et al*., 2005).

Most characteristics of agronomic importance are controlled by multiple genes and are subject to varying degrees of environmental modifications and interactions, hence are

ambiguous and have limited use for cultivar identification. Newbury and Ford-Lloyd (1993) and Hardon et al, (1994) reported that although agro-morphological characters are often influenced by environmental conditions, the method is still useful and easy to apply for classification, estimating diversity and registration of cultivars. Camussi et al. (1985) also indicated that morphological data showing continuous distributions or that are polygenically controlled, may be particularly useful in inter-group classification below species level. Morphological characterization has been used for various purposes including identification of duplicates, studies of genetic diversity patterns, and correlation with characteristics of agronomic importance. It involves a lengthy survey of plant growth that is costly, labor intensive and vulnerable to environmental conditions (CIAT, 1993). Moreover, morphological markers may be poorly suited for progeny analysis due to dominance effects (Werlemark *et al.,* 1999). Papaya germplasm is generally distinguished on the basis of morphological traits and has a wide variability of botanical characteristics. These features are not very distinct and sometimes quite variable, but have to be used by local farmers for basic identification of plant material.

2.9.2 Use of molecular markers

The use of DNA markers is widespread among plant geneticists because of the substantial amount of useful information that can be gathered from these markers. DNA markers are popular tool for examining genetic diversity of organisms and generating gene map for tagging traits of interest, for germplasm conservation or genetic enhancement (Weising *et al*., 2005). They have been defined as specific pieces of DNA

whose phenotypic expression is usually discerned, used to identify to an individual or a cell that carries it, or as a probe to mark a nucleus, chromosomes or locus (King and Stansfied, 1990). In plant breeding, superior cultivars can be detected by identifying quantitative traits loci (QTL) manifested with DNA markers. The markers provide a linkage framework and an estimate of similarity and difference among individuals (Stuber *et al*., 1999). Based upon the principles of marker assisted techniques, a gene or genes conferring traits of interest are expected to link with sets of markers. Thus, selection can be targeted to the molecular markers rather than for the trait itself (Karp and Edwards, 1997).

Biochemical markers

For the generation of molecular markers based on protein polymorphisms, the most frequently used technique is the electrophoretic separation of proteins, followed by specific staining of a distinct protein subclass. Although some earlier studies focused on seed storage protein patterns, the majority of protein markers are derived from allozymes (Weising *et al*., 2005). The term 'biochemical markers' was first introduced by Markert and Moller (1959) often referred to as allozyme or isozyme markers. Sometimes the terms isozymes and allozymes, incorrectly, are treated as interchangeable. Isozymes are enzymes that convert the same substrate, but are not necessarily products of the same gene. They may be active at different life stages or in different cell compartments. Allozymes are isozymes that are encoded by orthologous genes, but differ by one or more amino acids due to allelic differences (Weising *et al*., 2005).

Allozyme markers have more success distinguishing genotypes than use of morphological markers (Crozier, 1993). They are user friendly, require low cost of chemicals and labor and the markers are codominant; that is, both alleles in a diploid organism are usually clearly identifiable, and heterozygote can be discriminated from homozygote (Weising *et al.*, 2005). However, they are difficult to work with due to their limited amount of polymorphism, low levels of reproducibility and are influenced by tissue type and developmental stage of the plant (Zacarias, 1997) as they are unevenly distributed throughout the genome (Neilsen and Scandalios, 1994). Ocampo et al. (2006b) used nine isozyme systems to characterize papaya and found only four of the systems were polymorphic, with isozyme variation appearing lower than expected from morphological observations. In the past, allozymes have been used rather extensively for the discrimination of genotypes, but are now superseded by DNA markers because the latter usually detect much higher levels of polymorphism (Weising *et al.*, 2005). SSR have several advantages over other molecular markers allowing the identification of many alleles at a single locus, they are evenly distributed all over the genome, they are co-dominant, little DNA is required and the analysis can be semi-automated and performed without the need of radioactivity (Chen *et al.,* 2007).

SSR markers

Microsatellites also known as simple sequence repeats consist of tandemly reiterated short DNA sequence motifs spread throughout the genomes of most organisms (Estoup and Turgeon, 1996; Weising *et al.*, 2005). The tandem repeats of di- to tetra-nucleotide

sequence motifs are flanked by unique sequences which are ubiquitous, abundant, and well distributed in eukaryotic genomes (Wang *et al*., 1994; Tautz, 1989; Cardle *et al*., 2000; Morgante *et al*., 2002). They are very polymorphic due to the high mutation rate affecting the number of repeat units. Such length-polymorphisms can be easily detected on high resolution gels (for example, sequencing gels) by running PCR fragments obtained using a unique pair of primers flanking the repeat.

In recent years, SSRs have become one of the more popular molecular markers with applications in many fields as massive amounts of genomic sequences become available (Chen *et al.,* 2007). The DNA sequences flanking SSR's are conserved (Estoup and Turgeon, 1996), allowing the selection of PCR primers that will amplify the intervening SSR in all genotypes of the target species.

Although most early studies focused on dinucleotide repeats, other types of microsatellites such as mononucleotide have also proved to be useful (Weising *et al*., 2005). A dinucleotide microsatellite-enriched genomic library obtained from *Carica papaya* L. revealed polymorphism in papaya (Ocampo *et al*., 2006a). Twenty-four microsatellites gave polymorphism in *Carica papaya* and only four among the seven species of *Vasconcellea* confirming the divergence between both genera. Ocampo et al. (2005) also reported that microsatellite markers were more polymorphic in papaya than AFLP or RAPD.

2.10 Gel electrophoresis

The choice of the gel electrophoresis system to be used, and of its various components, depends on the expected size of the amplification product(s), on the resolution required to clearly see the difference in size among the amplified products and, to a lesser extent, on the intensity of the amplified products (CYMMT, 2005). The general rules involve using agarose gels for STSs due to the larger fragment sizes. For SSRs used for genetic diversity purposes, polyacrylamide gel electrophoresis (PAGE) is always used due to the required higher resolution. However, SSRs used in mapping studies, screening parental lines for polymorphisms begins on agarose gels and rerun on polyacrylamide gels (CYMMT, 2005).

During gel electrophoresis, agarose concentration of 2% is used for SSRs because of their smaller fragments (CYMMT, 2005; Weising *et al*., 2005). Migration distance and ratio of better quality agarose to normal quality agarose are the factors involved in the resolution of the differences in amplification product sizes. The larger the distance, the better the resolution (CYMMT, 2005). The buffer is used to prepare the gel and run it for better resolution. This buffer can be re-used once or twice with no problem since the running time is usually short. The electrophoresis tanks include gel trays where the various tooth combs are inserted, depending on the difference in size of the amplification products. For very small differences, 2 combs (12.5 cm migration distance) become necessary, but if the difference is large, 8 combs, or 3 cm migration distance, are

enough. Polyacrylamide gel electrophoresis is used when higher band resolution is required (CYMMT, 2005).

2.11 Capillary electrophoresis

The capillary electrophoresis uses fluorescence-labeled PCR primers and real-time laser scanning with an automated DNA sequencing device combined with specific fragment analysis software to visualize PCR-generated fragments (Blair *et al.*, 2002). The fragments are either resolved in sequencing gels or in capillaries allowing both high-precision microsatellite genotyping and high throughput (Mansfield *et al.*, 1996; Wenz *et al.*, 1998). An additional advantage of capillary electrophoresis is that every sample is run separately, and therefore a spillover between wells is impossible. In sequencers that can detect two or more dyes, for example, Applied Biosystems, markers labeled with different fluorochromes can be multiplexed in a single lane. One dye may be attached to an internal size marker that allows the computer to generate a calibration curve for automated allele sizing and quantification. This obviates problems of lane-to-lane and gel-to-gel variation, for example, band shifts and smiling effects, but can still yield inaccurate size estimates (Haberl *et al.*, 1999).

An automated DNA sequencer will allow much more accurate fragment length determinations than agarose gels, and usually provides single base-pair resolution. However, electrophoresis artifacts may still occur. For example, overloading was reported to cause error in microsatellite analysis on an ABI 377 sequencer (Fernando *et al.*, 2001). One must also be aware that any change in running conditions or the

particular fluorescent label of a primer may have a slight influence on the fragment mobility. Therefore, a set of standard samples should be included in the analysis (Weising *et al*., 2005).

CHAPTER THREE

3.0 COLLECTION AND DOCUMENTATION OF KENYAN PAPAYA GERMPLASM

3.1 Introduction

The pawpaw (*Carica papaya* L.), also known as the papaya, papita or tree melon (Samson, 1986; Nakasone and Paull, 1998), is a popular fruit in Kenya (Griesbach, 1992). It belongs to the family *Caricaceae* (Nakasone and Paull, 1998). A related member of this family is the mountain papaya (*Vasconcellea cundinamarcensis*), a native plant of Colombia and Ecuador. In Kenya it performs well from 1500-2000m (Griesbach, 1992). Papaya is cultivated from sea level at elevations from approximately 12-1500m (Griesbach, 1992), where warm climatic conditions prevail and there is good soil moisture. Papaya has never been found wild, but it is probable that it originated in Southern Mexico and Costa Rica from where it was taken to West Indies (Storey, 1941; Purseglove, 1968; Nakasone and Paull, 1998). It was taken by the Spaniards to Manila in the mid-16th century and reached Malacca shortly afterwards (Purseglove, 1968). From there it was taken to India. It was reported in Zanzibar in the 18th century and in 1874 in Uganda (Purseglove, 1968). It has now spread to all tropical and subtropical countries (Purseglove, 1968) including Kenya. It is also possible for papaya to have been moved from Uganda to Kenya through regional trade.

The extent and distribution of the available inter and intra-specific diversity in papaya is inadequately documented in Kenya. The Kenyan papaya is not only endowed with enormous morphological variation and horticultural traits, but also with adaptation to a wide range of environments (Griesbach, 1992; d'Eeckenbrugge, 2007). This great diversity is a genetic reservoir to papaya breeders for achieving different breeding objectives.

The fruit crop is therefore worth documenting as different cultural and agronomic practices may be incorporated in the future development of Kenyan papaya. Collection also constitutes an important step in sustainable development. Thus the main objective of this study was to collect and document papaya germplasm, to investigate the different papaya cultural and agronomic practices and to garner information on current status of papaya production in Kenya.

3.2 Materials and methods

3.2.1 Collection sites of papaya germplasm

The papaya germplasm collection was carried out to garner information on current status of papaya production in Kenya and available germplasm between June and September 2008. Each field surveyed was mapped by Global Positioning System (GPS) receiver (Table 1), prior to data gathering to provide accurate information on the locations. The collection sites were first selected in terms of geographical locations and distances to cover the approximate ecological ranges of papaya so that valid generalizations could be

drawn from the findings (Table 1). The germplasm was collected from six major papaya producing provinces of Kenya namely, Coast (Kilifi and Taita), Nyanza (Rongo and Nyamira), Western (Bungoma, Kakamega and Vihiga,), Rift Valley (Baringo, Keiyo and Nakuru), Eastern (Tharaka and Embu), and Central (Kirinyaga and Maragua) (Table 1).

Table 1. Locations of the various farms studied during germplasm collection

Farm names	District	Province	Latitude	Longitude	Elevation
Migingo	Kilifi	Coast	S03.91324	39.73970	12
Kilifi Institute	Kilifi	Coast	S03.92096	39.44260	37
Salehe-in	Kilifi	Coast	S03.92476	39.84260	25
Khosla farm	Kilifi	Coast	S03.90316	39.75990	12
Imani	Taita	Coast	S03.59904	38.73290	613
Voi	Taita	Coast	S03.39423	38.56310	582
Voi	Taita	Coast	S03.42621	38.55370	589
Manyani	Taita	Coast	S03.09620	38.99020	552
Marigat	Baringo	Rift Valley	N00.46673	35.99300	1025
Marigat	Baringo	Rift Valley	N00.46885	36.00450	1011
Cheptebo	Keiyo	Rift Valley	N00.47220	35.60330	1232
Sacred Training Institute	Bungoma	Western	N00.58966	34.53640	1434
Vihiga	Vihiga	Western	N00.01008	34.74950	1455
Rapogi	Rongo	Nyanza	S00.44360	34.34110	1380
Nyasaoro	Rongo	Nyanza	S00.45200	34.38170	1561
Nyakongo	Nyamira	Nyanza,	S00.59200	34.91910	2019
Molo	Nakuru	Rift Valley	S00.14206	35.44330	2425
Molo	Nakuru	Rift Valley	S00.92290	36.07100	1913
Kaunu	Tharaka	Eastern	S00.16746	37.80440	897
Kaunu	Tharaka	Eastern	S00.16739	37.80590	891
Kianamothi	Tharaka	Eastern	S00.18659	37.81490	874
Embu Training Institute	Embu	Eastern	S00.51275	37.45750	1488
Mwea	Kirinyaga	Central	S00.72724	37.42610	1124
Kibirigwe	Kirinyaga	Central	S00.53487	37.18250	1431
Kimirine	Maragwa	Central	S00.46060	37.09120	1289

3.2.2 Sampling

A targeted sampling procedure which targets special segments of the population such as species of plant was followed to define the sampling units. Consultation with the

respective District Agricultural Officers and farmers knowledgeable with the sampled areas enabled accurate identification of the farmers growing papaya.

3.2.3 Data collection

Data was collected through personal interviews with members in each household or in institutes responsible for management of papaya fields as respondents using structured and semi-structured questionnaires (Appendix 1). The semi-structured questionnaires enabled full consideration of the open-ended questions such as how farmers evaluate and identify the different cultivars and landraces. The number of varieties in each farm was recorded on-farm, where each farmer was asked to distinguish, name and describe the different varieties grown. Data was also collected on total farm size, cultivation practices, the proportion of the land occupied by papaya, the source of planting materials, average time to fruit maturity, fruit yield per year, proportion of papaya fruits sold, uses of the papaya fruits, stems and leaves and constraints to papaya production (Appendix 1).

3.2.4 Data analysis

Descriptive statistics of the data collected from survey were performed using Statistical Package for the Social Sciences (SPSS). The statistical significance of the differences between the distribution of papaya accessions in the districts, scales of production and the constraints to papaya production in Kenya was assessed by analysis of variance (ANOVA), with SAS statistical software version 9.1. The data was square rooted

transformed for homogeneity and subjected to least significant differences (LSD) to separate the means.

3.3 Results

Distribution of papaya germplasm in Kenya

The papaya germplasm collected consisted of 65 papaya accessions, representing 29 accessions from Coast, 4 from Central, 10 from Eastern, 9 from Nyanza, 5 from Western and 8 accessions from Rift Valley provinces of Kenya (Table 2). The variation of papaya germplasm among the districts with respect to distribution in each area and abundance of accessions is clearly evident from corresponding number of accessions recorded in each district surveyed which could be partly related to their better representation in the collection and partly due to the number of named cultivars (Table 2).

Table 2. Occurrence of accessions in the sampled districts

District	Accessions collected	No. of samples
Kirinyaga	Local, Papayi	2
Maragua	Papayi	2
Taita	Papayi	17
Kilifi	Papayi, Redlady, Sunrise, US	12
Tharaka	Local, Sunrise	8
Embu	Sunrise	2
Rongo	Apoyo	8
Nyamira	*Vasconcellea cundinamarcensis*	1
Keiyo	Honey dew, Solo	2
Baringo	Kiru, Solo, Sunrise solo	4
Nakuru	*Vasconcellea cundinamarcensis*	2
Bungoma	Kiru	1
Vihiga	Papayi	4

Different papaya accessions were also observed in the fields. The survey highlighted the existence and use of a considerable number of vernacular names such as ‘apoyo’ in Rongo district, ‘local’ in Tharaka and Kirinyaga districts, and ‘papayi’ in Kilifi, Taita, Vihiga, Nakuru, Kirinyaga and Maragua districts (Table 2). Some papaya accessions had known varietal names thus could be differentiated from others while others did not have known varietal names. ‘Red-lady’, ‘US’, and 'Sunrise’ cultivars were identified in Coast province (Table 2). ‘Sunrise’, ‘Solo’, ‘Honey dew’, 'Papayi’, and ‘Kiru’ were identified in Rift Valley province (Table 2). ‘Sunrise’ cultivar was identified in Central and Eastern provinces (Table 2). From the survey, Kilifi district had the highest number of varieties reported according to local names and known varietal names namely papayi, red-lady, sunrise and US (Table 2). Embu, Vihiga, Nakuru, Maragua, Bungoma and Rongo had the least number of known varieties. In total Coast and Rift Valley provinces had the highest number of varieties with each province recording four of the total varieties collected. *Vasconcellea cundinamarcensis* was found growing in Nakuru and Nyamira districts in Rift Valley and Nyanza provinces respectively (Table 2).

The most frequently recorded accessions were Papayi and Sunrise (Table 3). An accession locally named as ‘Papayi’ was found widely cultivated in five districts including Kilifi, Taita, Vihiga, Maragua, and Kirinyaga districts while Sunrise which is a commercial cultivar was found in four districts namely Baringo, Tharaka, Embu and Kilifi districts. The least distributed accessions included Apoyo, Sunrise solo, Honey dew, Redlady and US found only in one district each in Rongo, Baringo, Keiyo and Kilifi districts respectively. Analysis of variance done to compare the occurrence of the

various papaya accessions in the districts indicated significant differences in the papaya accessions (P= 0.03). Papayi and sunrise showed the highest distributions in the sampled districts.

Table 3. Distribution of papaya accessions in the districts

Accessions	**Occurrence of the different accessions**
Papayi	*0.87 ± 0.04^{a}
Sunrise	0.82 ± 0.04^{ab}
Solo	0.80 ± 0.04^{abc}
Kiru	0.78 ± 0.04^{abc}
Local	0.75 ± 0.03^{bc}
Sunrise solo	0.76 ± 0.03^{bc}
Apoyo	0.76 ± 0.03^{bc}
Vasconcellea cundinamarcensis	0.76 ± 0.03^{bc}
Redlady	0.73 ± 0.02^{c}
US	0.73 ± 0.02^{c}
Honeydew	0.73 ± 0.02^{c}
$LSD_{(0.05)}$	0.09
CV%	22.09

*Means±SE represents frequency of occurrence of the accessions in each district. Means with the same letter are not significantly different from each other. The frequency was square rooted transformed for homogeneity and subjected to least significant differences ($LSD_{0.05}$).

Table 4. Papaya accessions collected and their major morphological descriptors

	Local name	Code	District	Major morphological characteristics
1	Papayi	MIG1	Kilifi	1, 16, 21, 24, 27, 28,30, 33, 35, 38 40, 46
2	Papayi	MIG2	Kilifi	1, 9, 21, 24, 27, 29, 30, 32, 35, 38, 40, 45
3	Papayi	MIG3	Kilifi	1, 9, 20, 24, 27, 28, 30, 34, 35, 38, 40, 46
4	Papayi	MIG4	Kilifi	1, 9, 21, 24, 27, 29, 30, 34, 35, 38, 40, 46
5	Papayi	ST1	Kilifi	1, 10, 20, 24, 27, 28, 30, 32, 35, 38, 40, 45
6	Papayi	ST2	Kilifi	1, 10, 20, 24, 27, 28, 30, 33, 35, 38, 40, 44
7	Papayi	KLF2	Kilifi	2, 10, 20, 24, 27, 28, 30, 33, 35, 38, 40, 44
8	US	KOS1	Kilifi	2, 5, 21, 23, 25, 29, 30, 33, 35, 38, 39, 43, 44
9	Redlady	KOS2	Kilifi	2, 13, 21, 23, 27, 28, 30, 32, 35, 38, 39, 42, 44
10	Sunrise	KOS3	Kilifi	1, 11, 21, 23, 25, 28, 30, 33, 36, 37, 39, 42, 46
11	US	KOS4	Kilifi	2, 5, 21, 23, 25, 28, 30, 33, 35, 37, 39, 42, 44
12	Papayi	IMA1	Taita	2, 10, 21, 24, 25, 28, 30, 34, 35, 38, 40, 45
13	Papayi	IMA2	Taita	2, 10, 21, 24, 27, 28, 30, 34, 35, 38, 40, 46
14	Papayi	IMA3	Taita	1, 19, 21, 24, 26, 28, 30, 33, 35, 38, 40, 44
15	Papayi	IMA4	Taita	2, 16, 21, 24, 25, 28, 30, 32, 35, 38, 40, 44
16	Papayi	IMA6	Taita	1, 19, 21, 24, 25, 28, 30, 32, 35, 38, 40, 44
17	Papayi	IMA7	Taita	1, 10, 21, 22, 25, 28, 30, 33, 35, 38, 40, 44
18	Papayi	IMA8	Taita	2, 10, 21, 24, 25, 28, 30, 33, 35, 38, 40, 44
19	Papayi	VOI1	Taita	1, 11, 21, 24, 25, 28, 30, 33, 35, 38, 40, 45
20	Papayi	VOI2	Taita	1, 11, 21, 24, 25, 28, 30, 33, 35, 38, 40, 45
21	Papayi	VOI4	Taita	1, 11, 21, 24, 25, 28, 30, 32, 35, 38, 40, 44
22	Papayi	VB1	Taita	1, 10, 21, 24, 27, 29, 30, 32, 35, 38, 40, 44
23	Papayi	VB2	Taita	2, 10, 21, 24, 27, 29, 30, 32, 35, 38, 40, 44
24	Papayi	VB3	Taita	2, 10, 21, 24, 27, 29, 30, 32, 35, 38, 40, 45
25	Papayi	MAN1	Taita	1, 11, 21, 24, 27, 28, 30, 33, 35, 38, 40, 45
26	Papayi	MAN2	Taita	1, 17, 21, 24, 27, 28, 30, 33, 35, 38, 40, 45
27	Papayi	GAV1	Vihiga	2, 7, 20, 24, 25, 28, 30, 33, 35, 38, 40, 43, 45
28	Papayi	GAV2	Vihiga	2, 9, 20, 24, 26, 29, 30, 33, 35, 38, 40, 43, 45
29	Papayi	GAV3	Vihiga	1, 12, 20, 24, 26, 28, 30, 33, 35, 38, 40, 45
30	Papayi	GAV4	Vihiga	1, 6, 20, 24, 26, 28, 30, 33, 35, 38, 40, 45
31	Kiru	SCC	Bungoma	2, 14, 20, 24, 27, 29, 31, 33, 35, 38, 40, 45
32	Apoyo	RAP1	Rongo	1, 8, 21, 24, 26, 28, 30, 32, 35, 38, 40, 44
33	Apoyo	KIZ1	Rongo	2, 7, 21, 24, 26, 28, 30, 33, 35, 38, 40, 45
34	Apoyo	KIZ2	Rongo	1, 14, 21, 24, 26, 28, 30, 33, 35, 38, 40, 43, 45
35	Apoyo	KIZ3	Rongo	1, 11, 21, 24, 27, 28, 30, 33, 35, 38, 40, 43, 45
36	Apoyo	KIZ4	Rongo	1, 10, 20, 24, 25, 28, 30, 33, 35, 38, 40, 45
37	Apoyo	KIZ5	Rongo	2, 18, 20, 24, 25, 29, 30, 32, 35, 38, 40, 43, 44
38	Apoyo	KIZ6	Rongo	1, 4, 21, 24, 25, 28, 30, 33, 35, 38, 40, 43, 44
39	Apoyo	KIZ7	Rongo	1, 15, 20, 24, 26, 28, 30, 33, 35, 38, 40, 44
40	Papayi (*V. c*)	KAN1	Nakuru	2, 14, 21, 24, 26, 28, 30, 32, 35, 38, 39, 41, 44
41	Papayi (*V. c*)	MUT1	Nakuru	2, 14, 21, 24, 26, 28, 30, 32, 35, 38, 39, 41, 45

42	Sunrise	REB1	Tharaka	2, 10, 21, 24, 25, 29, 30, 32, 35, 38, 40, 43, 44
43	Sunrise	REB2	Tharaka	1, 19, 21, 24, 26, 29, 30, 33, 36, 38, 40, 43, 44
44	Papayi	REB3	Tharaka	1, 19, 21, 24, 25, 29, 30, 33, 36, 38, 43, 44
45	Sunrise	REB4	Tharaka	2, 16, 21, 24, 25, 29, 30, 32, 36, 38, 43, 44
46	Local	THK2	Tharaka	1, 6, 21, 24, 25, 29, 30, 33, 35, 38, 43, 44
47	Sunrise	TMJ1	Tharaka	2, 10, 20, 24, 25, 29, 30, 32, 36, 38, 42, 44
48	Sunrise	TMJ2	Tharaka	1, 19, 20, 24, 26, 29, 30, 33, 35, 38, 41, 44
49	Sunrise	EMB1	Embu	2, 14, 21, 24, 27, 29, 30, 32, 35, 38, 41, 44
50	Sunrise	EMB2	Embu	1, 14, 21, 24, 25, 29, 30, 33, 35, 38, 41, 44
51	Papayi	MF1	Kirinyaga	1, 19, 20, 24, 26, 29, 30, 32, 35, 38, 42, 44
52	Local	MF2	Kirinyaga	2, 9, 21, 24, 27, 29, 30, 32, 35, 38, 41, 44
53	Papayi	KIB1	Kirinyaga	2, 14, 20, 24, 25, 28, 30, 33, 35, 38, 41, 45
54	Papayi	MR1	Maragua	1, 14, 20, 24, 26, 29, 30, 34, 36, 38, 42, 44
55	Papayi	MR2	Maragua	2, 9, 20, 24, 26, 29, 30, 32, 35, 38, 41, 44
56	Sunrise solo	MRG1	Baringo	2, 14, 21, 24, 26, 28, 30, 33, 35, 38, 46
57	Kiru	PKR1	Baringo	1, 11, 20, 24, 26, 29, 30, 33, 35, 38, 45
58	Solo	PKR2	Baringo	1, 14, 21, 24, 26, 29, 30, 34, 35, 38, 46
59	Honey dew	CHP1	Keiyo	1, 11, 20, 24, 26, 29, 30, 34, 35, 38, 41, 45
60	Solo	CHP2	Keiyo	2, 9, 21, 24, 26, 28, 30, 33, 35, 38, 41, 45
61	Papayi	KLF1	Kilifi	1, 24, 28, 30, 33, 35, 38, 40, 41, 45
62	Papayi	IMA5	Taita	3, 28, 30, 32, 35, 38, 40, 45
63	Papayi	VOI3	Vihiga	1, 24, 28, 30, 33, 35, 38, 40, 44
64	Local	THK1	Tharaka	3, 29, 30, 33, 36, 38, 44
65	Kiru	PKR3	Baringo	3, 29, 31, 33, 35, 38, 45

V. c = *Vasconcellea cundinamarcensis*

Key of the morphological characteristics

Sexual type		**Uniformity of fruit**		**Flower color**	
Female	1	Not uniform	20	Yellow	35
Hermaphrodite	2	Uniform	21	White	36
Male	3	**Skin color at maturity**		**Stem color**	
Fruit shapes		Green	22	Light grey	37
Blossom and tapered	4	Orange	23	Greenish	38
Club	5	Yellow	24	**Height to first fruit**	
Cylindrical	6	**Fruit skin texture**		Low bearing (<1.0m)	39
Cylindrical elongated	7	Intermediate	25	High bearing (>1.5m)	40
Cylindrical lengthened	8	Ridged	26	**Fruit flesh color**	
Elliptic	9	Smooth	27	Yellow	41
Elongated	10	**Tree habit**		Deep yellow to orange	42
Globular	11	Multiple stems	28	Reddish	43
High round	12	Single stem	29	**Petiole length**	
Oblong	13	**Mature petiole color**		Generally small	44
Oval	14	Green	30	Intermediate	45
Oval elongated	15	Pale green	31	Generally large	46
Pear-shaped	16	**Flower size**			
Plum-shaped	17	Generally small	32		
Reniform	18	Generally intermediate	33		
Round	19	Generally large	34		

Scale of production and utilization of papaya in Kenya

Of the farms visited, the proportion of land under papaya production was small to large ranging from <0.8ha to >2ha (Table 3.5). Responses indicated that a majority (45%) of the sampled Kenyan papaya growers have small sized farms measuring <0.8 ha upon which they conduct mixed cropping mainly producing papaya fruits for subsistence. The fruits are mainly intercropped (62.5%) or planted along farm boundaries, with most farmers growing not more than just a few trees. Imungi and Wabule, (1990) reported similar results in Embu, Machakos, Meru, Murang'a, and Kisii districts and indicated that the fruits were either intercropped with other crops or planted along farm boundaries, with farmers growing not more than just a few trees. This can be attributed to among other factors preference by farmers for other crops such as tea, coffee, banana, maize and vegetables.

Large scale and medium scale farmers included mainly commercial farmers and research institutes found in Kilifi, Tharaka, and Baringo districts. Small scale farmers were found in all the districts visited (Table 5). While growers on small (less than 0.8 ha) and medium sized farms (0.8-2ha) indicated no varietal preferences, the commercial growers and institutes on large farms (>2ha) preferred specific varieties of known performance and traits. Preferred varieties included: solo, US, red-lady, and sunrise papaya which were mainly found in the Kilifi, Tharaka and Embu districts because of good quality fruits. Horticultural Crops Development Authority (HCDA), (2008) also reported that good quality papaya fruits were produced in the lower warm parts of the country (from

12-900m above the sea level) mainly in Coast and Eastern provinces to which the three above districts fall in confirming the suitability of such areas for papaya production. Small scales of production varied significantly (p=0.0058) among the districts while there were no significant differences in medium (p=0.0536) and large scale (p=0.9580) categories.

Table 5. Scales of production of papaya in the studied districts

Districts	Small scale (<0.8ha)	Medium scale 0.8-2ha	Large scale >2ha
Baringo	*0.71 ± 0.00^{b}	1.05 ± 0.17^{ab}	0.88 ± 0.17^{a}
Keiyo	0.71 ± 0.00^{b}	1.22 ± 0.00^{a}	0.84 ± 0.00^{a}
Taita	0.97 ± 0.15^{ab}	0.84 ± 0.13^{ab}	0.71 ± 0.13^{a}
Bungoma	1.22 ± 0.00^{a}	0.71 ± 0.00^{b}	0.71 ± 0.00^{a}
Embu	1.22 ± 0.00^{a}	0.71 ± 0.00^{b}	0.71 ± 0.00^{a}
Kilifi	0.84 ± 0.13^{b}	0.97 ± 0.15^{ab}	0.84 ± 0.13^{a}
Kirinyaga	0.71 ± 0.00^{b}	1.22 ± 0.00^{a}	0.71 ± 0.00^{a}
Maragua	0.71 ± 0.00^{b}	1.22 ± 0.00^{a}	0.71 ± 0.00^{a}
Nakuru	1.22 ± 0.00^{a}	0.71 ± 0.00^{b}	0.71 ± 0.00^{a}
Nyamira	1.22 ± 0.00^{a}	0.71 ± 0.00^{b}	0.71 ± 0.00^{a}
Rongo	1.22 ± 0.00^{a}	0.71 ± 0.00^{b}	0.71 ± 0.00^{a}
Tharaka	0.88 ± 0.17^{ab}	0.88 ± 0.17^{ab}	0.88 ± 0.17^{a}
Vihiga	1.22 ± 0.00^{a}	0.71 ± 0.00^{b}	0.71 ± 0.00^{a}
LSD(0.05)	0.36	0.41	0.39
CV%	19.11	23.4	25.89

*Means±SE represents frequency of occurrence of the scale of production in the studied farms in each district. Means with the same letter are not significantly different. Data was square root transformed and subjected to least significant difference test ($LSD_{0.05}$).

The papaya fruits were produced for the purpose of subsistence and for market, with most farmers (45.83%) selling more than 75% of their produce. The results reported 0.5 to 4 tonnes yield per day by large scale producers (>2ha). Fruits were picked when the green color changes into yellow halfway up the fruits, after which they are sized, graded and packed for local consumption and export. Research institutes were involved in

papaya research mainly in developing improved lines for the farmers. Fruits were mainly used as salads and desserts, and in jam making while the papaya leaves were used mainly as compost and animal fodder. Horticultural Crops Development Authority (HCDA), (2008) also reports similar reason for production for fresh consumption and processing. However, processing was not reported in this study.

Propagation of papaya

Papaya is propagated by seeds. Vegetative propagation is possible but not economical for commercial propagation. All farmers (100%) reported seed as the main planting material in all the farms visited (Table 6). Small and medium scale farmers obtained seeds from healthy-looking, ripe fruits while large scale farmers obtained seeds from commercial outlets, or imported seeds. Exchange of planting materials was also common among farmers. Seeds were sown in polyethylene bags in most studied farms before transplanting.

Table 6. Papaya propagation materials in the sampled districts

District of collection	Planting material	Source of planting material
Kirinyaga	Seeds	Local market
Maragua	Seeds	Local market
Taita	Seeds	Local market
Kilifi	Seeds	Local market, import (USA)
Tharaka	Seeds	Local market, import (Hawaii), K.A.R.I
Embu	Seeds	Local market, K.A.R.I-Thika
Rongo	Seeds	Local market
Keiyo	Seeds	Local market, K.A.R.I-Baringo
Baringo	Seeds	Local market, import, K.A.R.I-Baringo
Nakuru	Seeds	Local market
Bungoma	Seeds	Local market
Vihiga	Seeds	Local market

Constraints to papaya production

Constraints to production reported by farmers included difficulties in selection of the sex of the seedlings. In Tharaka, for instance, the farmers cut several trees as a result of occurrence of male trees in their farms. Farmers in Kirinyaga, Kilifi, and Taita districts planted at least two plants per hole to minimize this problem and at first flowering, a vigorous plant of the desired sex was kept and the others removed. Other farmers used seed color and floatation method to differentiate among the sexes. For example the seeds that were not black were said to be males. In floatation method, male seeds floated on water while female and hermaphrodite seeds did not. However, this is subjected to further confirmation by research. Other constraints reported by the farmers included insufficient water, tree height, diseases such as papaya ringspot virus (Plate 1), yellowing of leaves, powdery mildews, pests such as spider mites, unpredictable weather and market, destruction of plants by wild animals such as monkeys and elephants and lack of improved varieties.

Most respondents reported diseases (37.5%), sex ratio (15.6%) and pests (15.6%) as the major challenges in papaya production in Kenya. PRSV was common (20%) and was reported to have to caused a reduction in papaya production in Nakuru district. Powdery mildews (Table 7) were common in Central, Eastern, Western, Nyanza and Rift Valley provinces. The yellowing and drying of leaves at the tips were also observed in Nyanza and Rift Valley provinces. Several control methods were reported including spraying the plants with pesticides and biological control using herbal concoctions.

Constraints to papaya production in Kenya varied significantly ($p=0.005$) among the districts indicating the occurrence of different constraints in different districts.

Table 7. Constraints to papaya production reported in Kenya.

Constraint	(%) of occurrence	District in which the constraint was observed
Papaya ringspot virus (PRSV)	$*1.01\pm0.05^{a}$	Maragua, Kirinyaga, Embu, Tharaka, Kilifi, Taita and Nakuru
Sex ratio	0.93 ± 0.05^{ab}	Keiyo, Kilifi, Rongo, Taita and Tharaka
Spider mites	0.92 ± 0.05^{ab}	Tharaka Embu, Kirinyaga, Maragua, Taita
Powdery mildews	0.90 ± 0.05^{b}	Kirinyaga, Tharaka, Rongo, Baringo, Bungoma
Insufficient water	0.85 ± 0.04^{bc}	Baringo, Keiyo and Taita
Tree height	0.85 ± 0.04^{bc}	Taita, Kirinyaga and Kilifi
Market	0.76 ± 0.03^{dc}	Baringo
Root infections	0.76 ± 0.03^{dc}	Tharaka
Lodging	0.74 ± 0.02^{d}	Kirinyaga
Deformity in fruits	0.74 ± 0.02^{d}	Embu
Wild animals (Monkeys)	0.72 ± 0.02^{d}	Taita
Wild animals (Elephants)	0.72 ± 0.02^{d}	Taita
$LSD_{(0.05)}$	0.09	
CV%	24.08	

*Means±SE represents frequency of occurrence of the constraint in the district. Means with the same letter are not significantly different. Data was square root transformed and subjected to least significant difference test ($LSD_{0.05}$).

(a) **(b)** **(c)** **(d)**

Plate 1. Male papaya tree bearing many male flowers (a), papaya field with very tall trees (b), and PRSV infected leaf(c) and fruit (d).

3.4 Discussion

Complete documentation of plant's germplasm captures and preserves the germplasm's history and it is the basis of breeding and research. The use of vernacular names such as 'apoyo' and 'papayi' to name papaya may be due to the existence of different ethnic groups in Kenya. Each locality has its own unique set of names for different cultivars with even very different cultivars sometimes referred to by the same name. Thus there is need for sound characterization of papaya germplasm in Kenya.

The highland papaya (*Vasconcellea cundinamarcensis*) was found growing in altitudes greater than 2000m in Nakuru and Nyamira districts (Table 2). The only remarkable difference of this papaya from *Carica papaya* is in the fruit itself, which only grows to a length of up to 10cm. It develops a deep golden color and has acidic flavor. Propagation of this papaya by seed is easy but, since the seedlings develop into either male or female plants, this has to be considered when planting an orchard (Griesbach, 1992). The fruits are yellow in flesh, yellow when ripe and are less succulent. This papaya performs well in Kenya but has had little commercial importance (Griesbach, 1992). Some species of *Vasconcellea* are also reported to possess other traits such as high level of proteolytic enzymes, cold tolerance and disease resistance that could be useful in breeding programmes for papaya improvement (NRC, 1989; Scheldeman and Van Damme, 2002). Therefore, there is need to exploit this fruit tree in the improvement of *Carica papaya* in Kenya.

Kilifi and Baringo in Coast and Rift Valley provinces respectively, reported many papaya varieties. This may be due to the existence of both commercial farmers and research institution (K.A.R.I) that import and conserve papaya germplasm in such areas.

The type of the planting material used by all the farmers was seed (Table 6). Clonal propagation of papaya by grafting, layering or rooting of side-shoots is possible. However, clonal propagation does not appear to be feasible for large scale (commercial) planting (Griesbach, 1992; Nakasone and Paull, 1998), thus the preference to use of seeds as planting material. Tissue culture methods have also been demonstrated to be feasible for papaya (Nakasone and Paull, 1998). Papaya plantations were established mainly from seedlings, usually prepared from farmers' nurseries. Growers also imported the seeds from countries such as Hawaii.

The ripe *Carica papaya* fruits were used as desserts and in fruit salads in Kenya, reported in all farms surveyed and jam and juice making by some farmers. HCDA, (1987) reports similar reasons for papaya production in Kenya. Papaya fruit is also a source of vitamins A and C (Nakasone and Paull, 1998). These benefits together justify the scale of papaya production in Kenya from small to large scale farms (Table 4). Papaya industry in Kenya is relatively small and there are a few major growers with plantations in the >2ha category (Table 4). Papaya is also easy to propagate requiring less to maintain and easily intercropped with other economically important crops such as bananas, coffee, maize, tea and vegetables (Imungi and Wabule, 1990).

Infestations of papaya leaves by pests such as mites (Singh, 1990; Chay-Prove *et al*., 2000) often occur during long dry spells (Griesbach, 1992). Mites were found mostly on papaya leaves in all farms sampled. They caused scaring and discoloration of the papaya fruits, thus greatly reducing papaya fruit's market value. Elephants and monkeys were the chief mammalian predator of papaya especially in Taita district, Coast province (Table 7). They destroyed papaya trees mostly by knocking them down thus causing reduction in yield.

Diseases reported were mostly caused by fungi or viruses and were found destroying plants, reducing yields and thus impairing marketability of fruits. The major fungal pathogens of papaya included powdery mildew and other fungal and bacterial pathogens that were not identified in the study. In case of powdery mildew, the fungus was found mainly on the underside of the leaves where it develops typical powdery, white growth. Severely infected leaves turned yellow and dropped prematurely. Cool and cloudy weather favors the infection (Griesbach, 1992). Papaya ring-spot virus (PRSV) significantly reduced crop productivity in Nakuru district and other parts of Kenya. Early symptoms include yellowing and vein clearing in young leaves and sometimes severe blistering and leaf distortion (OECD, 2003). Dark concentric rings and spots or "C"-shaped markings (Plate 1) develop on the fruit which may turn tan-brown as the fruit ripens.

Sex selection (Plate 1) in papaya still remains a challenge in papaya production with most farmers planting more seeds in one hole, using seed color or floatation method to

encounter this problem. Dioecious papaya (Griesbach, 1992) segregate into females which bear the fruit and unproductive male plants. Dioecy is therefore a contributing factor in sex selection challenge in papaya production. This can be corrected by reintroduction of Hawaiian cultivars which are usually hermaphrodite and normally all the plants bear fruits. The practice of growing plants from seed also gives rise to variation, and vegetative propagation is seen as a means of maximizing profitability and uniformity (OECD, 2005) and has been used with varying success (Sookmark and Tai, 1975; OECD, 2005). Micro-propagation is also feasible for papaya (Nakasone and Paull, 1998) and have several advantages over seed propagation including, reduced time to produce new varieties, ease of maintaining genetic uniformity and production of plants that are all the same sex (Hansen, 2005). Despite these stated advantages of vegetative propagation, commercial propagation of papaya continues to be done largely via seed and probably reflects the ease of processing and lower start-up costs associated with seed propagation (OGTR, 2008).

3.5 Conclusion and recommendations

Many papaya accessions were identified by farmers only in their ethnic languages such as papayi and apoyo were often different from each other and were reportedly grown mainly by small and medium scale farmers. Some cultivars reported earlier such as Kitale, Cavite, Malindi, PP1, 77, 116, 273, 417, 418, 455 and 457 (Imungi and Wabule, 1990; Kamau *et al.*, 1993) which could not be traced in the fields could be due to inability to maintain the same cultivar's name and also due to changes that occur as a

result of open pollination in papaya. This shows the need for everyday record keeping for papaya germplasm in Kenya that will result into the conservation of papaya genetic resources. Besides, farmers in the research area possess considerable knowledge about the diversity present in papaya and the attributes of each farm. They also possess considerable knowledge about the challenges facing papaya production in Kenya. Hence their knowledge needs to be integrated in programs geared towards papaya improvement. Consequently, a thorough analysis of the indigenous knowledge systems, farmers' participation in designing and implementing conservation as well as improvement programs is critical to bringing practical solutions to problems of immediate concern to them.

The major challenges such as pests (aphids, whiteflies and mites), diseases (*Papaya Ringspot Virus*, powdery mildews, bacteria canker), and sex selection paradox are worth investigating. Thus, future research should focus on clean planting materials and breeding against pests and diseases. The practice of growing plants from seed also gives rise to variation, and vegetative propagation is seen as a means of maximizing profitability and uniformity (OECD, 2005) and should be exploited as a means of controlling the sex selection paradox in papaya production.

CHAPTER FOUR

4.0 MORPHOLOGICAL CHARACTERIZATION OF KENYAN PAPAYA GERMPLASM

4.1 Introduction

Carica papaya belonging to the *Caricaceae* family is a medium sized fruit crop with a potential to produce fruits throughout the year. Morphologically, its stem is hollow between the nodes, except in young plants; consisting mainly of wood parenchyma and bears large triangular scars. The peltate, leaves are arranged in a 2/5 spiral, having large long hollow petioles and large, deeply-lobed blades except in some cultivars (Samson, 1986).

Female flowers, 3-5 cm long occur alone or in small groups in the leaf axils while the ovary is 2-3 cm long and has five fan-shaped stigmas on top. The male flowers, with ten stamens each, are found on long hanging panicles. Bisexual flowers have either five or ten stamens and some of these become carpelloid (fruit-like), in which case the fruits have a 'cat-face' appearance and are unmarketable (Nakasone and Paull, 1998). Different types of hermaphrodite flowers may occur on the same tree depending on the season or on the age of the tree. Male trees are also variable: sometimes a fruit is found at the end of a long panicle. A complete change of sex may take place when an old male tree is cut back: sprouts bearing female flowers (and later fruits) may appear (Samson, 1986). There is a difference between pure males and sex reverting males (Teaotia and

Singh, 1967). The fruit is a large, fleshy, hollow berry. Fruits formed from female flowers are oblong to nearly spherical, but if formed from bisexual flowers they are pear-shaped, cylindrical or grooved. Marketable fruits weigh from 0.5- 2 kg (Samson, 1986) and are 10-20 cm long. The thin green skin turns yellow at the bottom when maturity sets in. The flesh is yellow to orange, in some cultivars and has a pleasant flavor. Around the cavity lie a thousand or more black seeds, but seedless fruits occur too. Many landraces and cultivars present hermaphrodite papaya, which are generally preferred for production. They display considerable phenotypic variation for many morphological and horticultural traits (Ocampo *et al*., 2006b). As a first step therefore, a germplasm collection from Kenya was gathered and its morphological diversity assessed.

4.2 Materials and Methods

4.2.1 Morphological characterization of papaya germplasm

Morphological characterization was done in the field during the collection using a descriptor list from IBPGR (1988). Fifteen descriptors (Appendix 2) were used to characterize 60 accessions (Table 4).

4.2.2 Data collection

The data collected from the papaya germplasm surveys included stem diameter, internode length, petiole length and leaf length. Plant internode length was determined as the mean length of five internodes measured 10cm above the ground. The color of

mature leaf petioles was determined by visual appraisal based on the guidelines of IBPGR, (1988). Length of mature petioles was determined as the average length of five mature leaf petioles measured from every observed petiole. Leaf length represented the average of 5 leaves, and measured from base of middle leaflet midrib to tip. The habit of the trees were also recorded as either single or multiple stemmed (Appendix 2). Flower bud length was determined as the length measured from observed flowers. The color of flowers was also recorded as determined by visual appraisal on fully developed open flowers (Appendix 2).

The fruits used for quality analysis were harvested and collected depending on the availability at the time of collection in the fields. Data for fruit length, fruit weight, fruit width, flesh thickness and cavity volume were taken. Longitudinal sections of the harvested fruits per tree were made, and then the fruit length was determined from pole to pole of the fruits. Fruit width was determined from the equator of the sectioned fruit.

4.2.3 Data analysis

Only the female and hermaphrodite plants were included in the analysis of the morphological diversity. Morphological data was submitted to principal component analysis (PCA), using the XLSTAT 2008 statistical package. Cluster analyses were carried out on the principal components with Eigenvalues of 0.988 to 3.147 using the Neighbor Joining method (Nei, 1973) or hierarchic ascendant analysis and Euclidian average distance.

4.3 Results

4.3.1 Morphological diversity in the papaya germplasm in Kenya

Kenyan papaya accessions display a wide variation in all aspects of the fruit tree characteristics, including leaf size, leaf shape, leaf color, fruit size, shape of the fruit, flesh color, stem pigmentation and tree habit (Plate 2a-h).

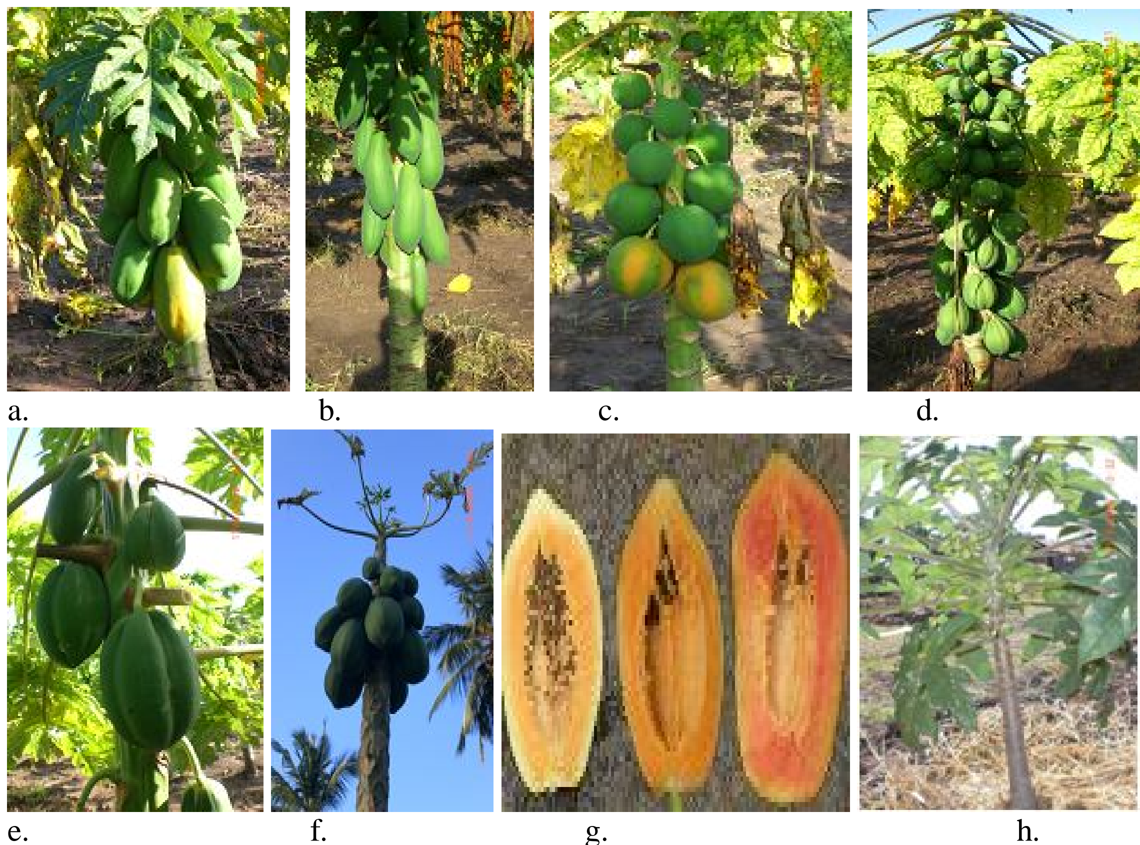

Plate 2a-h: Morphological diversity observed in the Kenyan papaya germplasm; (a) female papaya with lengthened cylindrical shaped fruits, (b) papaya with club shaped fruits, (c) round shaped fruits, (d and e) papaya with ridged with fruits (f) old papaya trees with fruits at higher heights (h) three fruits with yellow and red flesh and highland papaya found growing in the Kenyan highlands (Nakuru and Nyamira).

4.3.2 Principal component analysis

In the principal components analysis (PCA), the first seven principal components took into account 72.9% of the total variance (Table 8). The first principal component was related to fruit length, fruit diameter, tree habit, petiole length, flower color and leaf size, the second one to fruit shape, fruit skin texture, flower length and sexual or tree type, the third one to fruit diameter and uniformity of fruit distribution, the fourth one to petiole color, the fifth one to uniformity of fruit distribution and fruit skin color at maturity and the sixth one to fruit skin color at maturity. The seventh principal component was related to internode length.

Table 8. Principal component analysis (PCA) performed using the XLSTAT 2008 statistical package showing the correlations of the first nine principal components with the variables observed on papaya individuals.

Principal Component	Eigenvalue	Variability (%)	Cumulative (%)
PC1	3.147	20.98	20.98
PC2	1.842	12.28	33.26
PC3	1.417	9.45	42.71
PC4	1.272	8.48	51.19
PC5	1.142	7.61	58.80
PC6	1.128	7.52	66.32
PC7	0.988	6.59	72.90

The distribution of accessions based on the PC-1 and PC-2 shows the phenotypic variation among the accessions and how widely dispersed they are along both axes (Figure 1). The two components explain a cumulative variability of 33.26%. Based on the distribution of variates, the accessions of IMA4 and IMA7 are the most distantly related to that group while the second group shows MIG4 and KOS3 to be the least similar to the group. The most distant in the third quarter are the accessions REB2 and

MR1. The last quarter is made up of KIZ5 and REB1 (Figure 1) that are least similar to the group.

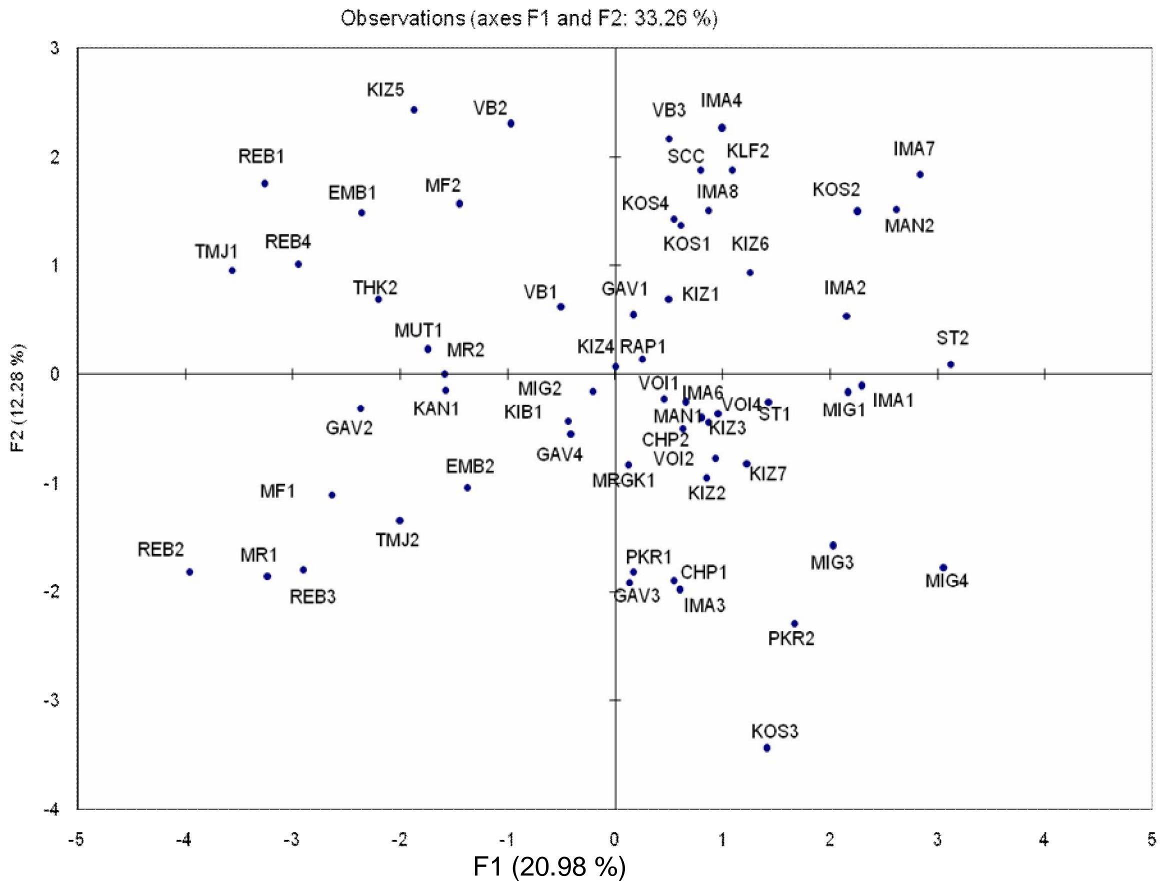

Figure 1. Distribution of variates in PC1 and PC2. PC1 accounts for 20.98% of the variation while PC2 accounts for 12.28%).

The correlation among characters showed four main clusters of characters (Figure. 2). The first cluster comprised traits associated with fruit shape, uniformity of fruits distribution, fruit skin color at maturity, flower color and fruit length; the second cluster comprised traits associated with the tree habit, fruit diameter, internode length, stem diameter, petiole length and flower length: the third cluster comprised characters related

to petiole color and fruit skin texture and the fourth cluster comprised only sexual or tree type.

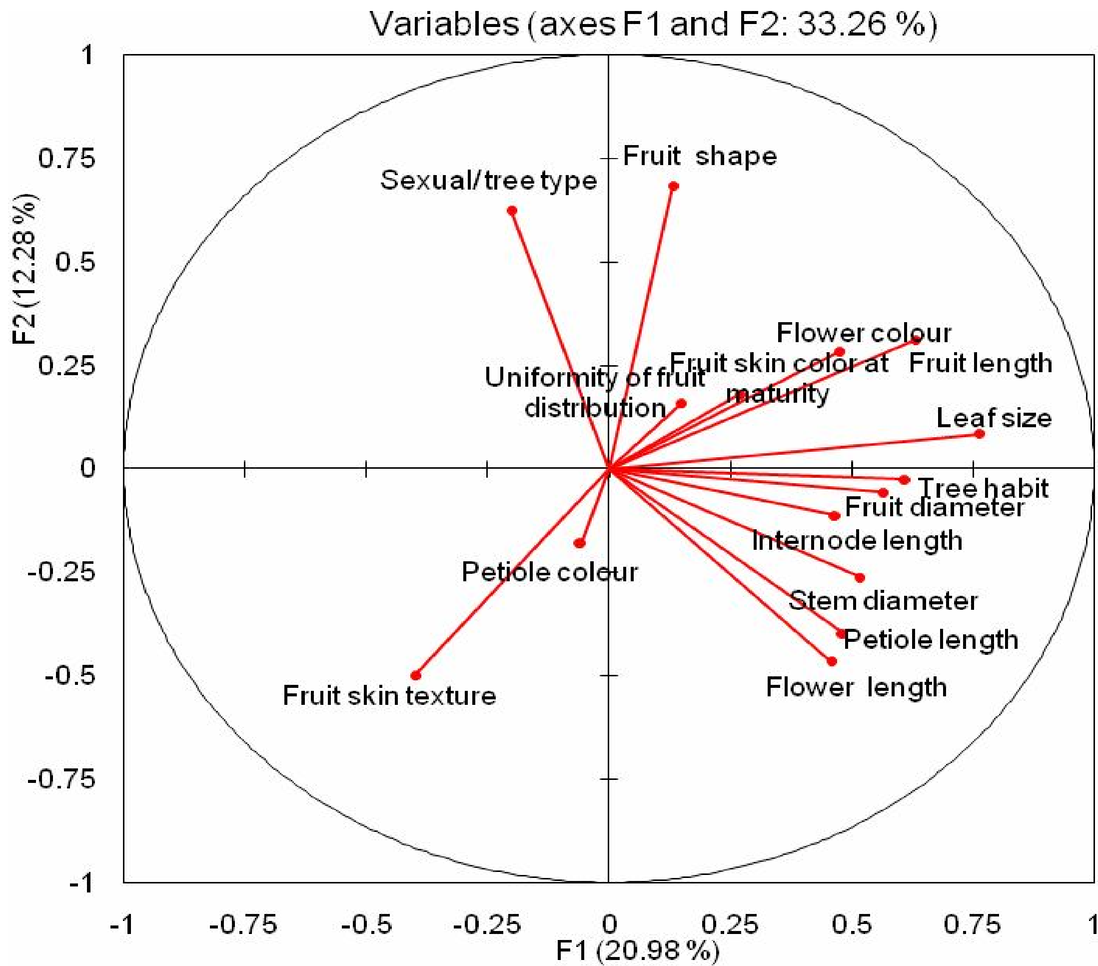

Figure 2. Correlation among characters associated with the first and second Principal Components. The closer the attributes are to each other in the PCA plot, the higher the correlation (i.e. the smaller the angle between the attributes, the higher the correlation).

4.3.3 Cluster analysis

The agglomerative hierarchical clustering dendrogram illustrates the relationship among the accessions (Figure 3). At 4.899 level of dissimilarity, almost all the sixty accessions were distinct from each other while at 20 levels and above, larger numbers of the

accessions were similar to each other. The cluster analysis separated the sixty accessions as different genotypes with Euclidean dissimilarity distance ranging from 4.9 to 57.

The dendrogram (Figure 3) was divided into four main branches, C1, C2, C3 and C4 based on the major morphological characters associated with them. The first cluster represented five accessions, the second one contained twenty five accessions, the third one contained twelve accessions and the fourth one contained eighteen accessions. At higher similarity levels, the above clusters were further divided into smaller sub-clusters.

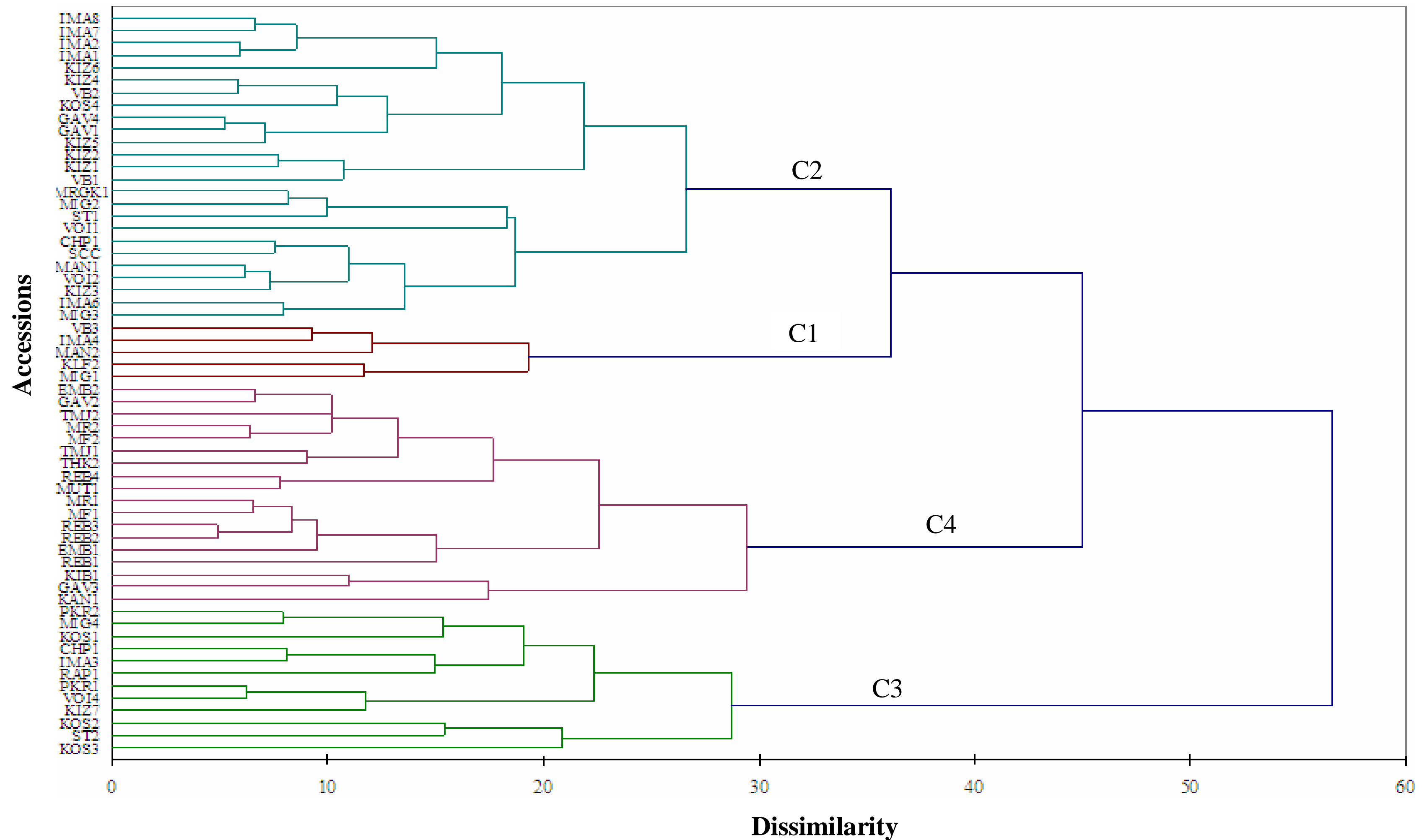

Figure 3. Dendrogram constructed based on morphological characters of 60 papaya accessions from six provinces (Coast, Central, Eastern, Rift Valley, Nyanza and Western) of Kenya using the Neighbor Joining method (Nei, 1973) and Euclidian average distance. C1, C2, C3 and C4 are the clusters (1-4) generated in the cluster analysis.

4.4 Discussion

Morphological analysis based on fifteen selected papaya descriptors (IBPGR, 1988) has shown significant diversity within the cultivated papaya. As expected, not all the accessions collected do correspond to distinct cultivars. US, Redlady, Mountain, Papayi, Apoyo, Local, Sunrise, Kiru, Solo, Honey dew, Sunsise solo were recorded (Table 4). Each locality had its own unique set of names for different cultivars, with even very different cultivars sometimes referred to by the same name. This linguistic polymorphism constitutes an obstacle to reliable identification of cultivars and therefore their eventual use for different research programmes (Dansi *et al.,* 1999). This shows, as expected, some challenges in the utilization of morphological characters in papaya classification (Dansi *et al.,* 1999).

Morphological traits were the major criteria used by farmers in differentiating and naming the papaya accessions. In the different production zones, factors influencing farmers' variety of choice and determining the level of diversity that is maintained included social, cultural, economic, abiotic and biotic factors.

The assumption in the multivariate analysis is based on the use of genotypes with maximum genetic divergence (Bhatt, 1970). Knowledge of correlations among characters is useful in designing an effective breeding programme for any crop. There are several reasons for using indirect selection. Sometimes the main character is expressed late or measurement of the indirect character is much easier than for the direct

character. Moreover, complex plant characters such as yield are quantitatively inherited and influenced by genetic effects, as well as by genotype and environment interaction. Therefore, identification and use of highly correlated characters are appropriate. Strong correlations were observed between traits related to fruit characteristics (fruit length, fruit shape, fruit diameter and fruit skin texture), petiole length and leaf size, confirming the results of Ocampo et al. (2006b) who showed correlation between these traits. This indicates presence of diverse variable arrangements at the individual genotype level pointing to ample possibilities of obtaining desirable trait combinations in specific cultivars. Broad trait variation apparent among the tested papaya accessions entails the presence of ample opportunities and prospects for genetic improvement of these characters through selection either directly or following recombination through intraspecific hybridisation of desirable genotypes. This would be crucial in meeting the demand of the farmers, researchers and consumers of this fruit crop.

Substantial morphological variation within and between the various accessions may be attributed to pollination, sexual recombination, and perhaps mutation followed by intensive selection by isolated human communities in diverse environments (Martin, 1976). Problems in pollination, fruit set and production are intimately associated with sex expression resulting from genotype-environment interactions. Cultivar and environmental differences have also produced a wide array of modified forms so the number and types of modifications have varied in reports by various researchers (Nakasone and Paull, 1998). Consequently, special breeding programs and experimental

designs are needed to distinguish genotypic from phenotypic variation (Weising *et al*., 2005).

From the dendrogram, however, the first cluster (C1) comprised all accessions from Taita and Kilifi representing Coast province. The second (C2) and the third (C3) clusters were characterized by cultivars from Rift Valley, Nyanza, Western and Coast provinces showing possibility of frequent exchange of planting materials among farmers from different zones. The fourth cluster (C4) was, however, characterized by cultivars from Central, Eastern, Rift Valley and Western provinces. In the dendrogram, obtained using quantitative and qualitative traits, a small part seems to be of geographic location of collection, as indicated by first cluster (C1). In both cases, the germplasm from Coast and Rift Valley provinces presented the highest variation, being scattered all over the tree, with little differentiation of accessions from other provinces.

Apart from showing great phenotypic variation among the accessions, the study has also found characteristics such as the fruit shape, flesh, texture, color, and fruit shape that can be useful as markers for classifying the accessions and can be employed in achieving papaya breeding objectives.

4.5 Conclusion and recommendations

This study has shown significant variation in morphological traits among the papaya accessions studied. The results demonstrated that traits that best discriminate between the accessions included fruit shape, uniformity of fruits distribution, fruit skin color at

maturity, flower color, fruit length, tree habit, fruit diameter, internode length, stem diameter, petiole length, flower length, petiole color, fruit skin texture and sexual type. The analysis using morphological characters revealed considerable amount of diversity among 60 papaya accessions that can be used in selecting diverse parents in breeding programme. This is also crucial in utilizing the genetic potential of these genotypes for improvement of traits needed for adaptation to various conditions. However, there is the need for complementing the similar work with other techniques such as DNA genetic markers to further accurately classify papaya germplasm existing in Kenya.

CHAPTER FIVE

5.0 CHARACTERIZATION OF KENYAN PAPAYA GERMPLASM USING SSR MARKERS

5.1 Introduction

Genetic diversity is the basis for genetic improvement. Information regarding the genetic diversity of available germplasm is vital to devise efficient plant breeding programmes as well as to maintain genetic diversity in a given gene pool. Genetic diversity can be estimated using morphological, biochemical and DNA-based markers. Morphological markers are often influenced by prevailing environmental conditions (CIAT, 1993). In addition, morphological markers are not suitable for perennial crops as it takes a long time to generate data. DNA-based markers offer consistent results regardless of the cropping conditions and type and age of sampled tissue (Sakiyama, 2000). These characteristics of DNA-based markers make it suitable for papaya research.

Papaya germplasm shows considerable phenotypic variation for many horticulturally important traits, including fruit size, fruit shape, flesh color, flavor and sweetness, length of juvenile period, plant stature, stamen carpellody, and carpel abortion. Their diversity has been investigated with morphological, isozyme, RAPD, RFLP, and AFLP markers (Stiles *et al*., 1993; Morshidi 1998; Aradhya *et al*., 1999; Kim *et al*., 2002; Van Droogenbroeck *et al*., 2002, 2004). Most of these studies showed limited genetic diversity in the common papaya and supported the recent separation in two genera by Badillo (2000).

Microsatellite markers exhibit high levels of polymorphism and such markers have been developed for *Carica papaya* (Ocampo *et al*., 2006a) and have been used successfully for determining the genetic relationship between *Carica papaya* cultivars elsewhere but not for the Kenyan papaya. The markers are also available for germplasm diversity, pedigree analysis, cultivar identification, as well as genetic mapping studies. Therefore, the objective of the present study was to assess the genetic diversity existing in Kenyan papaya germplasm using SSR markers.

5.2 Materials and methods

5.2.1 Sample collection

Young and healthy, newly developed shoots from the apex were collected from plants growing in the fields, immediately placed in labeled polyethylene bags which were kept in cool boxes (approximately 10°C) containing dry ice and then transported to the laboratory for DNA extraction. The materials were frozen and later used for DNA extraction. Forty two out of the total sixty five papaya accessions collected (Table 4) were used for SSR analysis.

5.2.2 DNA extraction from papaya

Total DNA was isolated from frozen leaves of the 42 papaya accessions (Appendix 3) by the method described by Doyle and Doyle (1990). Leaf tissue (100mg) from each accession was ground in liquid nitrogen with a pestle and pre-chilled mortar into a fine powder. The powder was transferred to 2.0ml eppendorf tube, 700µl 2X CTAB

(cetyltrimethylammonium bromide) buffer (distilled H_2O, 1M Tris-HCl (pH 8.0), 5M NaCl, 0.5M EDTA, 2% CTAB and Mercaptoethanol) was added and then vortexed gently. The homogenate was incubated at 65°C in a water bath for 45 minutes, while shaking tubes every 15 minutes. The samples were then cooled at room temperature for five minutes. Solvent extraction was done by adding 700µl of chloroform: isoamyl alcohol (24:1) to each tube, vortexed briefly and gently to avoid shearing of the DNA and then inverted severally before centrifuging for 10 minutes at 13,200 rpm to separate the phases. The aqueous top layer was carefully removed and transferred to a new, labeled eppendorf tube (2.0ml). The chloroform: isoamyl alcohol waste was however disposed of properly in a labeled waste container. To each tube, 50µl of 10% CTAB (in 0.7 M NaCl) was added, vortexed gently and mixed thoroughly. About 700µl of chloroform: isoamyl alcohol (24:1) was added to each tube, vortexed briefly and gently and then inverted severally before centrifuging for 10 minutes at 13,200 rpm to separate the phases. The aqueous top layer was carefully removed and transferred to a new, labeled eppendorf tube (2.0ml). An equal volume of $4^{o}C$ isopropanol (500µl) was added to each tube, tubes inverted severally and then let to sit at $-20^{o}C$ for 15 minutes before spinning at 13,200 rpm for 20 minutes to get the supernatant. The supernatant was decanted and the DNA pellet washed with 1 ml of 70% ethanol (for 3 minutes) and spun at 13,200 rpm for 30 minutes. The ethanol was discarded and the DNA pellet washed again in 1ml of 90% ethanol then spun for 30 minutes at 13,200 rpm. The ethanol was discarded, the tubes inverted and air dried for 30 minutes. The DNA was dissolved in

100µl of low salt TE buffer, 1-2µl of DNAse-free RNAseA (10mg/ml) added to each sample and then incubated at 37 ^{o}C for 1 hour. The DNA was then stored at -20°C.

5.2.3 DNA quantity and quality determination

DNA quantity and quality were determined using the agarose gel electrophoresis and a spectrophotometry. Agarose powder (1g) was dissolved in 100 ml of Tris Borate EDTA (TBE) buffer (1% w/v) by slowly boiling in a microwave oven. The mixture was allowed to cool to about 50 ^{o}C and ethidium bromide added to the gel at a concentration of 1mg/ml. While the agarose was cooling, the gel tray was prepared by sealing the open edges of a clean, dry glass tray with autoclavable tape so as to form a mold to avoid leakage and so that the tray could accommodate the desired thickness of the gel. The warm agarose solution was then poured into the gel tray in which a comb was inserted to form the wells. The gel was allowed to cool for 30 minutes before removing the autoclave tape, and immersing the gel in the electrophoresis tank containing 1X TBE buffer. The combs were removed and 8µl of each DNA sample containing 3µl of loading solution was loaded to the wells of the gel to the top. DNA lambda digested with E*cor*I and Hind III restriction enzymes was used as a molecular weight marker that was run in parallel on one lane of the gel. The gel was run at a constant voltage of 100 volts until the bromo-phenol blue migrated almost to the end of the gel. The gel was then removed from the rig, placed in a UV trans-illuminator and photographed.

Quantification of DNA in solution was done by measuring the absorbance of light (260 nm) in a spectrophotometer. 2µl of each sample was added to 98µl TE, mixed well, and

OD260 and OD280 read to determine purity (CYMMT, 2005). After UV quantification, the concentration of each DNA sample was adjusted to a concentration of choice with TE of 10ng/ µl, and stored at 4°C (CYMMT, 2005).

5.2.4 Polymerase chain reaction (PCR) with SSR primers

A total of seven microsatellite primer pairs used to amplify the extracted DNA samples (Table 10) were selected randomly from a dinucleotide microsatellite-enriched library among 24 markers that gave polymorphism in papaya (Ocampo *et al.*, 2006a). The forward primers for each of the seven markers were labeled at the 5^1 end of the oligonucleotide using fluorescent dyes for screening by capillary electrophoresis on the ABI prism and 3730 genetic analyzer (Applied Biosystems). The fluorescent capillary based dyes were 6FAM (Blue), PET (Red), VIC (Green), and NED (Yellow) (Table 10). After screening of the seven pairs of SSR markers, all were found to amplify scorable and reproducible banding profiles. PCR amplifications were performed with a Gene-Amp PCR system 9700 (Applied Biosystems) in a 20 µL final volume containing 10 ng of genomic DNA, 10X PCR buffer with $MgCl_2$ (Gene script), 2.5 mM dNTP, 5/µl *Taq* DNA polymerase and 0.5µM of each primer (Table 9). The microtubes were placed in a thermal cycler (a Gene-Amp PCR system 9700 (Applied Biosystems)) and the thermocycling reactions done in the following scheme: denaturation at 94 ^{o}C for 5 min, 35 cycles of 30 s at 94 ^{o}C; 1 min. between 46 ^{o}C and 52 ^{o}C (Table 10); 45 s at 72 ^{o}C; and a final elongation for 4 min. at 72 ^{o}C to reduce the probability of false scoring of stutter bands as alleles.

Table 9. PCR optimization conditions of the seven SSR primers

	1st PCR trial	2nd PCR trial	Final optimized PCR conditions
PCR buffer with $MgCl_2$ (10X)	3 µl	3 µl	1.5 µl
dNTP (2.5 mM)	3 µl	3 µl	1.5 µl
Forward primer	1.5 µl	2 µl	0.5 µl
Reverse primer	1.5 µl	2 µl	0.5 µl
Taq DNA polymerase (5/µl)	0.3 µl	0.4 µl	0.25 µl
Sterile water	20.3 µl	18.6 µl	14.75 µl
Template DNA (10ng)	0.5 µl	0.5 µl	1 µl
Total volume	30 µl	30 µl	20 µl
	No positive result	Positive result with primer dimers	

Table 10. *Carica papaya* microsatellite primers used in the study

Marker name	Dye	Forward primers 5^1-3^1	Reverse primers 5^1-3^1	Repeat motif	Tm (^oC)	Expected allele size
mCpCIR1	PET	GCATTACTTATCATCGTCC	CTATCCTTGGCGTCTT	$(CT)_{18}$…$(GA)_3$	48	314–364
mCpCIR2	PET	GTCTATCTACCTCCCA	GAGTGTTATCATAGTCTACA	$(TC)_{24}$	52	260–284
mCpCIR3	NED	GAACTCACCTACACGAACT	ACTTCTACCACCGGC	$(TC)_{14}$	50	188–210
mCpCIR8	6FAM	ATGGCTGAAGACAACTC	CTCAATAGCCCAATAACA	$(CT)_{20}$…$(AC)_5$	46	283–293
mCpCIR10	VIC	CAGCAGAAAACAAGGG	GGGTTCCGGTTTAGTT	$(TA)_4$…$(AG)_{18}$	46	341–349
mCpCIR18	NED	ATGGGATTTTAGAGGTG	GTATGAGGGAATGGAAA	$(CT)_9$ …$(CT)_9$	50	291–295
mCpCIR23	NED	CGCATTGTTATTGACT	ACCTACAGGGCCTAC	$(TC)_8$	50	281–283

Source: (Ocampo *et al*., 2006a)

5.2.5 Agarose gel electrophoresis of amplified products

Following amplification, PCR products were stored at $4^{o}C$ prior to electrophoresis. Agarose gel preparation was carried out as outlined in section 5.2.3. The PCR product (8µl) was run on 1% agarose gel. To 5µl of each PCR product, 3µl of sample loading buffer (bromo-phenol blue) was added and mixed by pipetting before loading the resulting mixture in the preformed sample wells on the gel. The samples were run alongside 3µl 1kb DNA ladder at 100 volts for 45 minutes. After the run, the gel was viewed under UV light and photographed. Each amplified SSR fragment was visualized as a distinct band.

5.2.6 Capillary electrophoresis

The amplified DNA fragments were screened by capillary electrophoresis on the ABI 3730 genetic analyzer (Applied Biosystems). The capillary electrophoresis runs were post- PCR co-loaded in three groups based on the strength of the dyes and fragment sizes. The first set of group included mCpCIR10, mCpCIR2 and mCpCIR18. The second set of group included mCpCIR1 and mCpCIR3. The third set of group included mCpCIR8 and mCpCIR23.

PCR products were co-loaded post-PCR, where 0.75µl of 6-FAM and VIC and 1µl of the PET and NED labeled products were mixed with their corresponding 9 capillary electrophoresis cocktail (prepared by mixing 1ml of HiDi formamide and 12µl of the Genescan™ -500LIZ™ size standard (Applied Biosystems) for a half plate reactions.

The preparation was sealed and the DNA fragments denatured at 95 ^{o}C for 3-5 minutes using ABI PCR machine.

The DNA fragments were size-fractioned using capillary electrophoresis on an ABI prism 3730 automatic DNA sequencer (Applied Biosystems). The GeneMapper v3.7 software (Applied Biosystems) was used to size peak patterns, using the internal Genescan^{TM} -500LIZ^{TM} size standard and Genotyper 3730 (Applied Biosystems) for allele calling. Genotyping was carried out by capillary electrophoresis using the ABI PRISM 3730 (Applied Biosystems), a fluorescent based capillary detection system that uses polymer as the separation matrix at the International Livestock Research Institute (ILRI). This facilitated the accurate sizing of the microsatellite allele to within ±0.3 base pairs (Buhariwalla and Crouch, 2004). Primers were optimized by running different ratios of PCR products and choosing the one giving the best signal profile that is, signal/noise ratio and relative fluorescent units (RFU).

5.2.7 Fragment analysis

The amplified fragments were analyzed using the GeneMapper v3.7 software (Applied Biosystems). Size calling, which includes peak detection and fragment size matching were performed using the GeneMapper v3.7. Bins, which represent a fragment size or base pair range and dye color that define an allele, were constructed from reference data. Algorithms were used to determine if peaks represented alleles. When a peak from a given data sample matches the location of a bin, the software made an allele call. Alleles were automatically assigned allele calls based on the bin definitions. The results were

stored in the GeneMapper v3.7 database. Allelobin software was used for checking the quality of the markers.

5.2.8 Data analysis

Seven SSR markers were included in the subsequent analysis. The total number of alleles, the number of common alleles with allelic frequencies of at least 5% and the polymorphism information content (PIC) values (Bostein *et al*., 1980; Smith *et al*., 2000) were determined for each SSR marker. The SSR data was analyzed using the SIMQUAL (Similarity for Qualitative Data) routine to generate simple matching coefficients. These similarity coefficients were used to construct dendrogram using UPGMA and employing the Sequential, Agglomerative, Hierarchical and Nested clustering (SAHN) from the Numerical Taxonomy and Multivariate Analysis (NTSYS-pc) version 2.1T (Rohlf, 2000).

5.3 Results

5.3.1 Polymorphism of microsatellites used to characterize papaya accessions

The seven pairs of SSR primers used were polymorphic across all loci and had a polymorphic information index ranging from 0.75 to 0.852. A single peak denoted homozygous genotypes while two clear peaks indicated heterozygous genotypes (Figures 4 and 5).

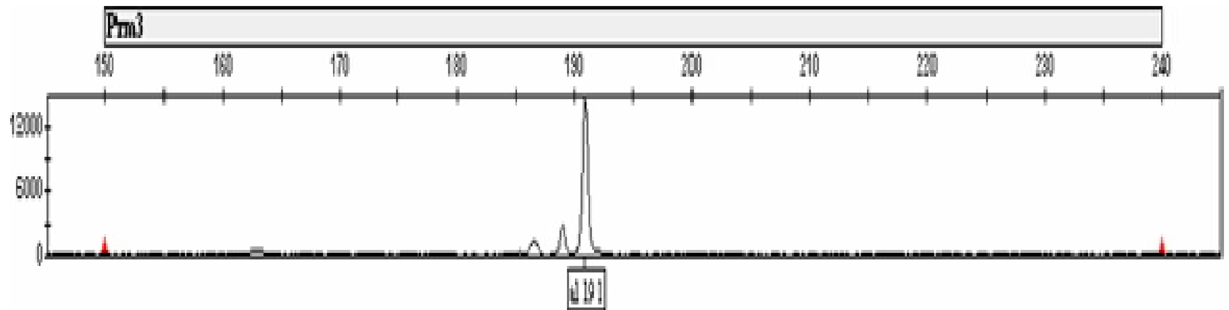

Figure 4. Electrophenogram showing homozygosity of accession MF1 using SSR marker mCpCIR3. The X-axis and Y-axis represent allele sizes and peak intensities, respectively.

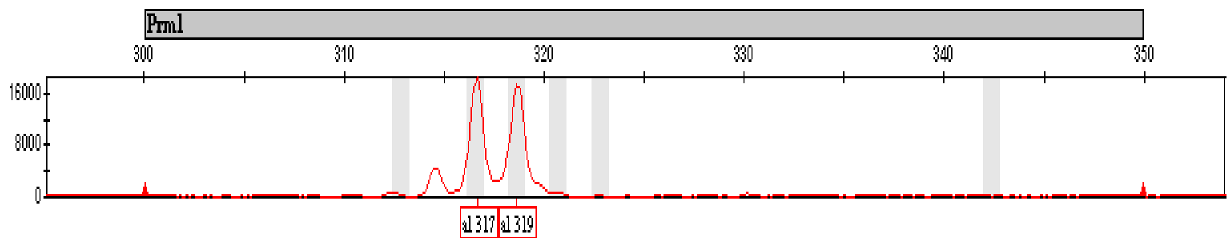

Figure 5. Electrophenogram showing heterozygosity of papaya accession IMA8 using SSR marker mCpCIR1. The X-axis and Y-axis represent allele sizes and peak intensities, respectively.

5.3.2 Marker quality indices, allele variability, polymorphism and observed heterozygosity in the 42 accessions

Seven microsatellite markers were used to characterize 42 papaya genotypes. A total of 24 alleles were detected. The number of alleles per locus ranged from 2 for mCpCIR10, mCpCIR8 and mCpCIR1 to 5 for mCpCIR18 and mCpCIR23 with an average of 3.42 (Table 11). The maximum size range of 350bp was observed with primer mCpCIR10 whereas the lowest of 163 bp was with primer pair mCpCIR3 (Table 11). The highest % abundant alleles was observed in primer mCpCIR3 (41.86%) whereas the lowest % abundant alleles was observed in primer mCpCIR2 (24.72%). None of the markers had rare alleles at <=5%. The lowest quality index was 0.3942, observed in mCpCIR3,

whereas the highest quality index of 0.5302 was in marker mCpCIR18 with an average being 0.4615. The polymorphic information content (PIC) varied from 0.75 to 0.852 for varieties with an average of 0.81. Marker mCpCIR2 had the highest PIC value of 0.852 while marker mCpCIR3 had the lowest PIC of 0.75. The observed heterozygosity ranged from 0.48 to 0.88 with a mean of 0.62 across the seven SSR loci. The highest observed heterozygosity was in marker mCpCIR2 with a value of 0.88 while the lowest was 0.48 in marker mCpCIR1 (Table 11).

Table 11. Quality indices and polymorphism detected by 7 SSR markers in the 42 papaya accessions

Marker name	**Quality index**	**Total no. of alleles**	**Allele size range**	**Abundant allele (%)**	**Rare allele (<=5%)**	**[a]PIC values**	**[b]Observed heterozygosity**
mCpCIR1	0.4184	2	313-342	30.95	None	0.822	0.48
mCpCIR2	0.4938	4	242- 300	24.72	None	0.852	0.88
mCpCIR3	0.3942	4	163- 211	41.86	None	0.75	0.62
mCpCIR8	0.4766	2	283- 300	28.57	None	0.816	0.6
mCpCIR10	0.4792	2	337- 350	28.57	None	0.846	0.64
mCpCIR18	0.5302	5	288- 305	33.3	None	0.804	0.55
mCpCIR23	0.4381	5	271- 293	34.04	None	0.791	0.57
Mean	**0.4615**	**3.42**		**31.71**		**0.81**	**0.62**

[a]$PIC=1-\sum(pi^2)$ where Pi is the frequency of the ith allele detected
[b]Frequency at which heterozygous individuals occur in a population at a given locus.

5.3.3 Genetic variability within the 42 Papaya accessions

5.3.3.1 Number and frequency of alleles

The 42 accessions used were classified into two major groups based on genetic distance from each other. The number of alleles within the 42 papaya accessions across the seven loci ranged from 8 to 18, with an average of 11.93 (Table 12). The highest number of

alleles (18) was observed in genotype KOS3, whereas the lowest (8) was observed in sample MIG9 (Appendix 3). These two genotypes were collected in Coast province and appeared in different classes.

The allele frequencies ranged from 0.01 in all seven SSR markers used (i.e. mCpCIR1, mCpCIR2, mCpCIR3, mCpCIR8, mCpCIR10, mCpCIR18, mCpCIR23) to 0.43 in marker mCpCIR3. The highest allele frequency was observed in allele 191bp at locus mCpCIR3 (Table 12).

Table 12. Frequency of alleles detected in 42 papaya accessions with 7 SSR markers

Marker	[a]Alleles (base pairs)	Allele frequencies
mCpCIR1	313, 315, 316, 325	0.01
	317, 323	0.15
	318	0.05
	319	0.32
	321, 322	0.12
	334	0.02
	342	0.04
mCpCIR2	163, 193, 209, 242, 269,272,280, 298,300	0.02
	266, 296	0.08
	295	0.04
	211, 276, 294	0.17
mCpCIR3	192, 194, 197, 198, 199, 200, 201, 262, 267, 277, 293	0.01
	274	0.25
	191	0.43
	190	0.15
	189	0.09
mCpCIR8	297, 300	0.05
	295	0.23
	294	0.19
	283, 287, 290, 293	0.01
	292	0.02
	289	0.29
	288	0.08
	285	0.04
mCpCIR10	338, 344, 345, 347, 349, 350	0.02
	341, 339, 346	0.12
	343	0.06
	342	0.29
	340	0.17
	337	0.01
mCpCIR18	288, 289, 291, 295, 303, 305	0.01
	297, 301	0.02
	292, 300	0.08
	299	0.03
	296	0.33
	293, 294	0.2
mCpCIR23	272, 292, 293	0.02
	273, 278, 291	0.01
	289	0.07
	287	0.04
	284	0.26
	283	0.37
	271, 281	0.09

[a]Allele variants at a specific SSR locus

5.3.3.2 Polymorphism of the 42 papaya accessions

Polymorphism within the 42 papaya accessions ranged from 33.33% to 75.00% with an average of 49.61% (Table 13). One papaya accession KOS3 from Coast showed the highest polymorphism of 75.00% across the seven loci. The lowest level of polymorphism based on the seven SSR primers was observed in one MIG9 also from Coast.

Table 13. Number of accessions and their corresponding percentage polymorphism across the seven SSR loci

No. of genotypes	Papaya accessions	% Polymorphism
1	MIG9	33.33
3	THK2, KOS4, KLF1	37.50
6	KOS1, VB2, ST2, JKU3, RAP1, KLF2	41.67
1	PKR1	41.83
9	JKU1, KIZ3, MF1, MIG6, MAN2, EMB1, MF1, IMA5, VB1	45.83
9	KIZ4, MUT, CHP2, IMA6, MIG2, KIB1, MIG1, ST1, JKU2	50.00
5	VOI1, IMA2, IMA8, MF2, PKR4	54.17
3	THK1, CHP1, MRGK2	58.33
1	IMA1	62.50
2	KOS2, IMA7	66.67
1	MIG4	66.70
1	KOS3	75.00

5.3.4 Genetic relationships among the 42 papaya accessions

The genetic similarity among the 42 papaya accessions ranged from 0.764 to 0.933 with an average of 0.843. The genetic similarity matrix showed that most papaya accessions were closely related. About 96.9% of the pair-wise comparisons among papaya accessions exhibited genetic similarity greater than 0.802; less than 4% (3.1%) showed

genetic similarity lower than 0.802. There was strong genetic similarity between the papaya accessions IMA6 (Papayi) and MIG1 (Papayi) and between JKU2 (Papayi) and KIZ3 (Apoyo) (Figure 6) all of which were classified in cluster A. KOS3 (Red-lady), MIG4 (Papayi) and IMA7 (Papayi) accessions showed the lowest similarities with the other accessions (Table 14). KOS3 (Red-lady) showed the least similarity with all accessions with an average genetic similarity of 0.797 ranging from 0.764 to 0.845. The two highland papayas (*Vasconcellea cundinamarcensis*) (JKU3 and MUT) had an average genetic similarity of 0.840.

Table 14: Similarity matrix among papaya accessions

	MUT	KIZ3	IMA1	IMA6	IMA2	MIG2	VB2	KOS2	MIG6	MIG1	KOS3	IMA7	KOS4	JKU3	JKU2	MIG4
MUT	1.000															
KIZ3	0.851	1.000														
IMA1	0.811	0.851	1.000													
IMA6	0.851	0.851	0.797	1.000												
IMA2	0.872	0.845	0.791	0.885	1.000											
MIG2	0.878	0.878	0.824	0.905	0.912	1.000										
VB2	0.838	0.851	0.838	0.865	0.845	0.851	1.000									
KOS2	0.831	0.804	0.831	0.845	0.811	0.831	0.831	1.000								
MIG6	0.845	0.885	0.791	0.858	0.865	0.845	0.845	0.811	1.000							
MIG1	0.838	0.851	0.824	0.932	0.858	0.878	0.838	0.885	0.845	1.000						
KOS3	0.777	0.777	0.764	0.791	0.797	0.804	0.804	0.784	0.797	0.804	1.000					
IMA7	0.811	0.797	0.797	0.838	0.845	0.851	0.797	0.804	0.804	0.851	0.831	1.000				
KOS4	0.851	0.892	0.824	0.865	0.872	0.892	0.838	0.818	0.858	0.838	0.804	0.878	1.000			
JKU3	0.865	0.824	0.851	0.851	0.831	0.865	0.811	0.845	0.804	0.865	0.791	0.838	0.865	1.000		
JKU2	0.865	0.932	0.838	0.865	0.885	0.905	0.851	0.818	0.858	0.851	0.777	0.797	0.892	0.838	1.000	
MIG4	0.811	0.784	0.770	0.811	0.791	0.811	0.784	0.804	0.791	0.811	0.845	0.838	0.824	0.797	0.797	1.000

5.3.5 Phylogenetic analysis

The unweighted pair group method of arithmetic averages (UPGMA) analysis clustered the 42 papaya accessions into two main clusters (Figure 6) namely A and B. The simple matching coefficient ranged from 0.81 to 0.91. Cluster 'A' had 39 papaya accessions across all the six provinces. Cluster 'B' contained three papaya accessions KOS3, MIG4 and IMA7 all of which were collected in Coast province. Cluster A, sub-cluster 'i' contained 25 accessions all of which were collected across the six provinces. This sub-cluster also had six accessions KOS4 (US), KOS2 (Red-lady), CHP1 (Honey dew), CHP2 (Solo), PKR1 (Kiru) and EMB1 (Sunrise) identified with specific varietal names. Cluster A, sub-cluster 'ii' contained two accessions (JKU3 and MUT) all of which were mountain papayas collected in JKUAT and Nakuru respectively. Cluster A, sub-cluster 'iii' contained 9 papaya accessions all of which were from five different provinces. Cluster A, sub-cluster 'iv' contained 3 papaya accessions all of which were collected in Coast Province and were only identified as 'papayi' as the specific varietal names were not known. In the dendrogram obtained with the SSR markers, the structure is weak (Figure 6). However, there are some little indications of a geographic component. Thus, three accessions clustered in group B and three accessions clustered .in cluster A, sub-cluster 'iv' are all from Coast province. The Coast province accessions show the widest diversity, as they are scattered all over the dendrogram. This gives similar results of the dendrogram obtained using morphological characteristics (Figure 3). However, the diversity remains limited in the dendrogram obtained using SSR markers (Figure 6).

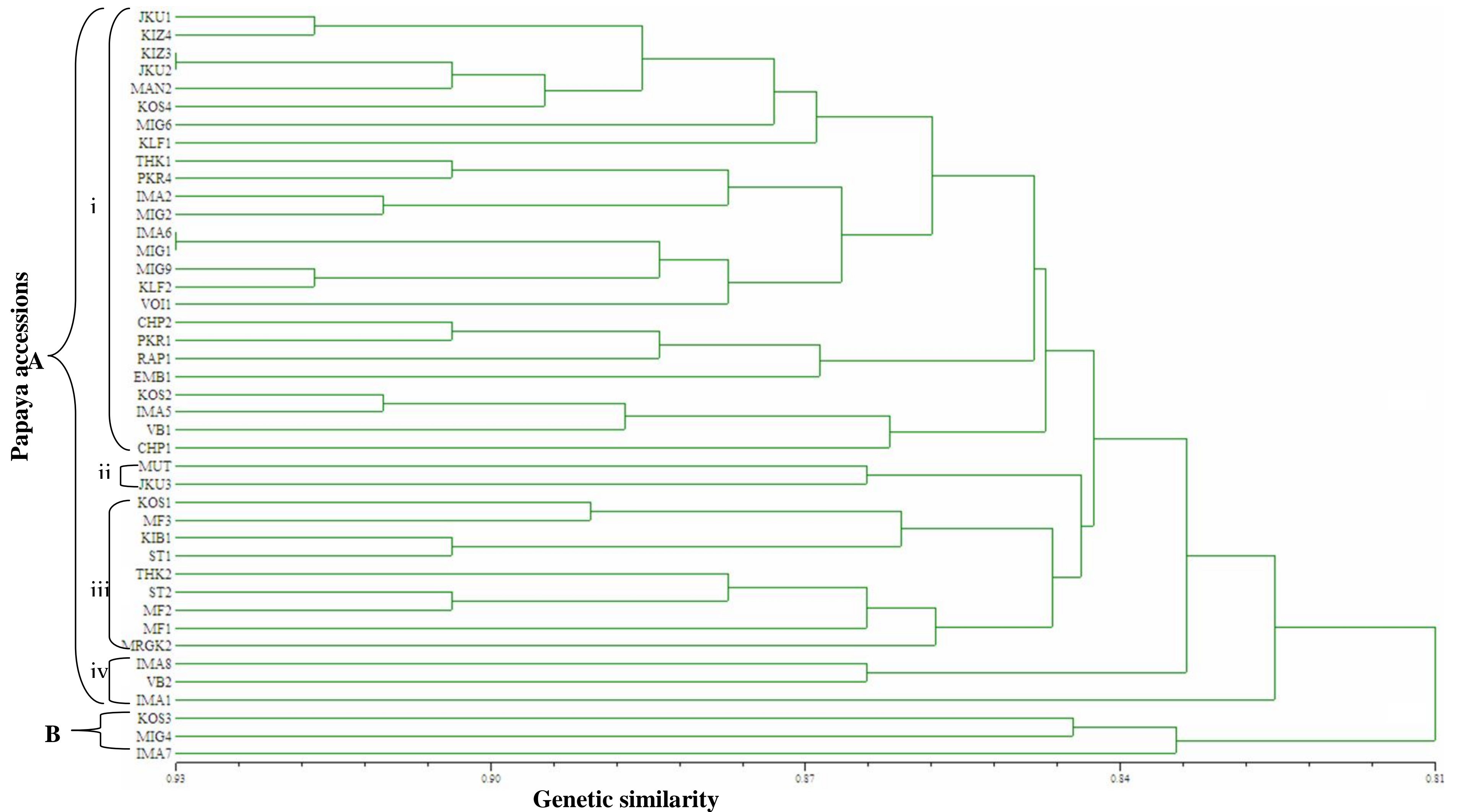

Figure 6. UPGMA dendrogram of 42 Kenyan papaya accessions based on genetic similarity matrix calculated from SSR markers. There were two main clusters, A with 39 accessions while and B with 3 accessions.

5.3.6 Principal component analysis (PCA) based on genetic distance estimates of 42 papaya accessions

The first and second principal components accounted for 16.21% and 23.82% of the total variation, respectively (Figure 7). The PCA scatter plot which gives the spatial representations of genetic distances among accessions, revealed two major cluster groups (Figure 7). Generally the PCA scatter plot detected trends similar to the clustering illustrated in the dendrogram (Figure 6).

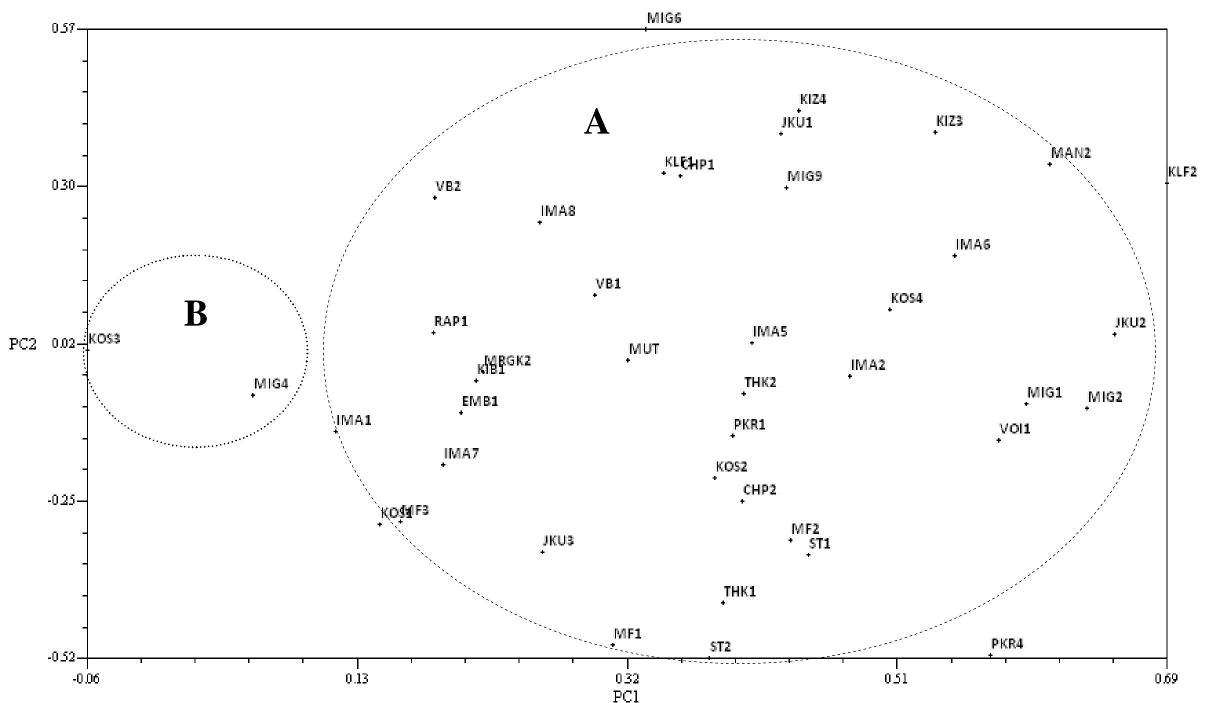

Figure 7. Principal component analysis of 42 papaya accessions using the seven SSR markers. PC1 and PC2 represent 16.21% and 23.82% of the total variation, respectively.

5.4 Discussion

Assessment of genetic diversity and relationships of accessions is of interest not only for germplasm conservation but also for breeding purposes (Sakiyama, 2000). The present study is the first genetic characterization of Kenyan papaya germplasm using microsatellite markers. Microsatellite markers have shown high levels of genetic polymorphism in many important crops including *Oryza sativa* L. (Bligh *et al*., 1999), *Sorghum bicolar* (Smith *et al*., 2000), *Vitis vinifera* (Di Gaspero *et al*., 2000), *Helianthus annuus* L. (Yu *et al*., 2000), *Cucumis melon* L. (Danin-Poleg et al., 2001), cassava cultivars (Moyib *et al*., 2007) sweet potato (Karuri *et al*., 2009) The present study showed that microsatellite markers were highly polymorphic in papaya accessions assessed in this study (Table 11). It has been suggested that high levels of polymorphism in microsatellite markers are related to the mechanism of mutations and the high rate at which they occur (Ashiley and Dow, 1994). The polymorphic information content (PIC) of an SSR marker provides an estimate of the discriminatory power of that SSR marker by taking into account the number of alleles that are detected and also the relative frequencies of those alleles (Smith *et al*., 2000). Microsatellite variation was clearly detected in the germplasm studied. The number of alleles varied greatly with all the markers having PIC values greater than 0.75. This amount of variation within each accession might best be explained by occasional outcrossing. Natural outcrossing in hermaphrodite papaya fields immediately adjacent to a potential source of contaminating pollen may affect about 5% of seed, as demonstrated using the β-glucuronidase (GUS)

transgene as a marker in assays of seeds from non-transgenic papaya trees surrounding transgenic plants (Kim et al., 2002).

The level of microsatellite polymorphisms in papaya is relatively high considering the present study of 8 to 18 alleles per primer pair were amplified from the 42 papaya accessions (Table 12). This may be due to pollination and wide geographical diversity from which the papaya accessions were collected. Moderate to high levels of heterozygosity were also observed in this study, and varied greatly across the seven loci, ranging from 0.48 to 0.88. This may be due to the sex in papaya which is determined by a single gene with three alleles, in which males and hermaphrodites are heterogametic, while females are homogametic (Hofmyer, 1967 and Storey, 1976). These results indicate present a high level of heterozygosity, which can be related to the plant dioecy and to the low level of selection in papaya varieties and landraces of Kenya.

Papaya (*Carica papaya* L. 2n=18) with a small genome size of 372 Mbp (Arumuganathan and Earle 1991), is a polygamous angiosperm with male, female and hermaphroditic forms (Storey 1938; Hofmyer, 1938). In this study, individual accessions were observed to contain between 2 and 5 alleles at any one locus. This is similar with the previous findings of between 2 and 8 alleles (Ocampo *et al*., 2006a).

Pair-wise comparisons of genetic similarity among the Kenyan papaya accessions revealed limited genetic diversity which may be the consequence of inbreeding from a limited gene pool. This limited genetic variation detected among the accessions does not correspond to the wide range of morphological characteristics observed in the field.

These characteristics of agronomic importance are controlled by multiple genes, a few of which can significantly change plant morphological traits (Camussi *et al*., 1985). Most of these traits are sufficiently variable, and improvement consists of selecting desirable recombinants from segregating populations (Kim *et al*., 2002). Our SSR study did not identify any of these, but could help point to related materials.

The genetic differences among the accessions revealed by their clustering into distinct groups suggest the presence of different sources of relationships among the papaya accessions. The dendrogram analysis (Figure 6) based on the seven SSR markers (Table 10) showed a very limited geographic structuration of diversity. In both cases, the germplasm from Coast, and to a lesser extent from Rift Valley, presented the highest variation, being scattered all over the tree, with little or no differentiation from most remaining accessions, while some accessions from Coast regrouped in particular clusters. Generally, the relationships among accessions in the cluster could be attributed to their diversity, geographical locations and also due to exchange of plant genetic resources among farmers within and between the provinces. Principal component analysis (Figure 7) is one of the most important methods of ordination analysis. It constructs a new set of orthogonal coordinate axes such that the projection of points onto them has maximum variance (Weising *et al*., 2005). In this study, the PCA analysis further provides information about associations between accessions, which are useful to formulate better strategies for breeding. The result of the principal component analysis was consistent with the results of the UPGMA cluster analysis. It is therefore thought

that combining these accessions from different groups would be helpful in papaya improvement considering their diversity.

5.4 Conclusion and recommendations

While there is considerable phenotypic variation of 60 accessions (Figure 3) of Kenyan *Carica papaya* germplasm, DNA finger printing of 42 Kenyan papaya accessions has indicated a limited genetic variation (Figure 6) confirming reports of Kim *et al*., (2002), who showed similar results while using 63 accessions from different countries. The genetic relationships of the 42 papaya accessions are much closer to each other in the cluster analysis. The results of genetic diversity study provide estimates on the level of genetic variation among diverse materials that can be used in germplasm management, varietal protection, and papaya improvement. Although identical samples were found between 'MIG1' and 'IMA6', and between 'JKU2 and KIZ3', all the other papaya accessions were distinctively separated by seven SSR markers (Figure 6). The amount of genetic diversity found with SSR markers is sufficient to distinguish between breeding lines for varietal protection. The estimates of genetic similarity are particularly useful in choosing widely divergent parents with desirable traits for genetic mapping and selection.

CHAPTER SIX

6.0 GENERAL DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

Diversity in genetic resources is the basis for genetic improvement. It is essential to fulfill the ever-changing needs of humans as well as to cope with unforeseen biotic and abiotic stresses (Bekele *et al*., 2005). Genetic resources will have little value unless it is efficiently conserved and properly utilized. Its efficient utilisation as well as conservation depends on the availability of reliable genetic diversity information (Bekele *et al*., 2005). In the present study, the genetic diversity of papaya accessions collected from six major papaya growing provinces (Coast, Central, Eastern, Western, Nyanza and Rift Valley) of Kenya was evaluated using morphological characters, quality attributes and molecular markers.

Complete documentation captures history while characterization consists of recording those characters which are highly heritable, can be easily seen by the eye and are expressed in all environments (IBPGR, 1988). In addition, the extent of variation and the occurrence of papaya germplasm was also investigated and documented. Many accessions of papaya both local and imported were reportedly cultivated in Kenya and showed high diversity in both morphological and horticultural traits (Table 4) that can be exploited in further papaya improvement and conservation. In general, the number of accessions described in this study area, their role as a source of fruit and cash income as well the existing local classification system reflect the significance of papaya in the local subsistence agriculture. In addition Highland papaya was identified and collected in

some parts of Kenya. This species had the widest distribution mostly covering the mountainous zones of Nyamira, Molo and Nyeri. The major constraint in their propagation still remains growth conditions which require higher altitudes of up to 2500m (Scheldeman *et al*., 2006). Compared to their better- known lowland cousin, *Carica papaya*, highland papayas are generally smaller and have distinct texture, taste and aroma (Scheldeman *et al*., 2006). The species therefore deserves special attention in Kenya and shows potential as genetic resources in papaya breeding programs (Scheldeman *et al*., 2006). However, knowledge of the nature and potential of these fruit crops is lacking in Kenya and needs to be addressed by assessing the existing diversity of the species of Kenya and to analyze their agronomic practices.

Results of morphological diversity analysis demonstrated the presence of substantial variability among evaluated papaya accessions for 15 morphological characters. These accessions may serve as sources of desirable genes for the genetic improvement of different characters of papaya. Use of a combination of qualitative and quantitative characters is recommended, since quantitative characters are more sensitive to environmental influences and the growth stage of the plant. The PCA done gives a representation when numerous characters were considered simultaneously showing a considerable variation in the morphological characteristics and horticultural traits of papaya. Cluster analysis, however, sufficiently provided a clear and more informative display of the relative positions of the accessions. Similar results have been reported earlier for papaya (Ocampo *et al.,* 2006b).

SSR markers done for 42 papaya accessions revealed a limited genetic diversity in papaya with accessions from Coast and Rift Valley to a lesser extent presenting the highest variation as compared to accessions from other provinces indicating high levels of diversity in papaya in such areas. Simple sequence repeats can therefore be used to provide profiles that are highly discriminative among cultivars for many species. The knowledge about the genetic relationships of genotypes provides useful information to address breeding programmes and germplasm resource management (Rolda´n-Ruiz *et al.*, 2001). In this study, morphological data analysis of the papaya accessions was coupled with molecular analyses (SSR markers) to investigate the genetic relationships. The accessions showed diverse morphological traits and distinct SSR markers patterns.

The range of genetic distance based on morphological traits was on average higher than SSR markers, which may reflect the influence of the environment on the performances of the materials as suggested by Salem et al. (2008). Therefore, the DNA markers and morphological traits will not necessarily gain closely matching results (Mart´nez *et al.*, 2005; Vollmann *et al.*, 2005). Semagn (2002) suggested two reasons for low correlation between DNA markers and morphological as well as protein data: (a) DNA markers cover a larger proportion of the genome, including coding and noncoding regions, than the morphological markers and (b) DNA markers are less subjected to artificial selection compared with morphological markers. Mart´nez *et al.*, (2005) believed that the correspondence between different methods might be improved by analysing more morphological characters and DNA markers.

This study using microsatellite markers and morphological characters revealed considerable amount of diversity among studied accessions. The data can be used in selecting diverse parents in breeding programme and in maintaining genetic variation in the germplasm and is crucial in utilizing the genetic potential of these genotypes for improvement of traits needed for adaptation to various conditions. In many cases these resources are threatened with loss through genetic erosion due to social, political and economic challenges in the country. It is therefore, recommended that genetic conservation and improvement based on the selected materials should be encouraged as these could lead to reversal in this erosion. Such activities would also maximize the use of these resources and provide the opportunity to explore the potential use of the resource for other purposes. The results will also further supplement efforts needed for the sustainable conservation of the genetic diversity and improvement of papaya together with its wild relatives.

7.0 REFERENCES

Ali, Z. M. and Lazan, H. (1998). Papaya. In: Shaw, P. E., H. T. Chan Jr and S. Nagy (Eds.). Tropical and sub tropical fruits. Agscience Inc., Florida (U.S.A.), pp 401-445.

Applied Biosystems (AB). (2004). Microsatellite analysis on the Applied Biosystems 3130 Series Genetic Analyzers. Applied Biosystems. Available from: *www.appliedbosystems.com*

Agricultural Biotechnology Support Project II (ABSPII). (2004). Papaya Ringspot Virus-resistant (PRVR) Papaya 'Why genetically engineer virus resistance into papaya?' Fact Sheet. Cornell University. 9/10/2004.

Aradhya, M., Manshardt, R. Zee, F. and Morden, C. (1999). A phylogenetic analysis of the *Carica* sp. Caricaceae based on restriction fragment length variation in a *cp*DNA intergenic spacer region. *Genetic Resources and Crop Evolution* 46: 579–586.

Arumuganathan, K. and Earle, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 9: 208–218.

Ashiley, M. V. and Dow, B. D. (1994). The use of microsatellite analysis in population biology, background, methods and potential applications. In: Schierwater, B. Streit, B. Wagner, G. P. and Desalle, R. (Eds), Molecular Ecology and Evolution, Approaches and Applications. Birkha user Verlag, Basel, Switzerland, pp. 185-201.

Badillo, V. M. (2000). *Carica* L. vs. *Vasconcella* St.-Hil. (*Caricaceae*) con la rehabilitación de este último. *Ernstia* 10 (2), 74–79.

Bekele, Y. D. (2005). Estimation of the level of genetic variation among currently grown *C. arabica* genotypes in Ethiopia. PhD thesis. University of the Free State, Bloemfontein, South Africa.

Bhatt, G.M. (1970). Multivariate analysis approach to selection of parents for hybridization aiming at yield improvement in self-pollinating crops. *The Australian Journal of Research* 21:1-7.

Blair, M. W., Hedetale, V. and McCouch, S. R. (2002). Fluorescent-labeled microsatellite panels useful for useful for detecting allelic diversity in cultivated rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 105:449-457.

Bligh., H. F. J. Blackhall, N. W. Edwards, K. Y. and McClung. A. M. (1999). Using amplified length polymorphism to identify cultivars of brown and white milled rice.

Crop science 3a(6):1715-1721.

Borge, M. N. R and Basedow, T. (1997). A survey on the occurrence and flight periods of fruit fly species (Diptera: Tephritidae) in a fruit growing area in Southwest Nicaragua, 1994/95. *Bulletin of Entomological Research* 87 (4) 405–412.

Botstein, D., White, R. L. Skolnick, M. and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32(3): 314- 331.

Buhariwalla, H. K. and Crouch, J. H. (2004). Optimization of marker screening protocol to assess the degree and distribution of genetic diversity in landraces of pigeon pea. In: Bramel, P. J. (Ed), Assessing the risk of losses in biodiversity in traditional cropping systems. ICRISAT, Patancheru, pp. 57- 74.

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

Cardle, L., Ramsay, L. Milbourne, D. Macaulay, M. and Marshall, D. (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics* 156: 847–854

Chay-Prove, P., Ross, P. O'Hare, P. Macleod, N. Kernot, I. Evans, D. Grice, K. Vawdrey, L. Richards, N. Blair, A. Astridge, D. (2000). *Agrilink Series: Your Growing Guide to Better Farming. Papaw Information Kit*. Queensland Horticulture Institute and Department of Primary Industries, Qld, Nambour, Qld.

Chen, C., Yu, Q. Hou, S. Li, Y. Eustice, M. Skelton, L. R. Veatch, O. Herdes, R. E. Diebold, L. Saw, J. Feng, Y. Qian, W. Bynum, L. Wang, L. Moore, P. H. Paull, R. E. Alam, M. and Ming, R. (2007). Construction of a Sequence-Tagged High-Density Genetic Map of Papaya for Comparative Structural and Evolutionary Genomics in Brassicales: *Genetics* 177: 2481–2491.

Crozier, W. W. (1993). Electrophoretic evidence of genetic diversity in the red winter wheat regions of the United States. *Proceedings of the National Academy of Science,* USA. 85: 5583-5586

Danin-Poleg., Y. Reis N, Baudraci, S. Arnas, S. Pitrat, M. Staub, J. E. Oliver, M. Arus, P. Devicente, C. M. Katzir, N. (2001). Simple sequence repeats in cucumis mapping and map margin. *Genetics* 43:963-974.

Dansi., A. Mignouna, H. D. Zoundjihe´kpon, J. Sangare, A. Asiedu, R. and Quin, F. M. (1999). Morphological diversity, cultivar groups and possible descent in the cultivated

yams (*Dioscorea cayenensis/ Dioscorea rotundata)* complex in Benin Republic. *Genetic Resources and Crop Evolution* 46: 371–388.

Di-Gaspero., G. Peterlunger, E. Testolin, R. Edwards, K. J. Clipriani, G. (2000). Conservation of microsatellite loci within the genus *Vitis*. *Theoretical and Applied Genetics* 101:303-308.

Doyle, J. J. and Doyle, J. T. (1990). A rapid total DNA preparation procedure for fresh plant tissues. *Focus* 19: 11–15.

d'Eeckenbrugge, C. G., Restrepo, M. T. and Jimenez, D. (2007). Morphological and isozyme characterization of common papaya in Costa Rica. In: Proc. 1st IS on papaya. Chan, Y. K. and Paull, R. E. (Eds). *Acta Horticulturae* 740, ISHS 109-120.

El Moussaoui, A., Nijs, M. Paul, C. Wintjens, R. Vincentelli, J. Azarkan, M. Looze, Y. (2001). Revisiting the enzymes stored in the laticifers of *Carica papaya* in the context of their possible participation in the plant defense mechanism. *Cell and Molecular Life Sciences* 58: 556-570.

Estoup, A and Turgeon, J. (1996). Microsatellite isolation protocol, version 1, November 1996.

Fernando, P., Evans, B. J., and Melnick, D. J. (2001). Electrophoresis artifacts- a previously unrecognized cause of error in microsatellite analysis. *Molecular Ecology Notes* 1:325-328.

Gardan L., Christen, R. Achouak, W. Prior, P. (2004). *Erwinia papayae* sp. Nov., a pathogen of papaya (*Carica papaya*), *International Journal of Systematic and Evolutionary Microbiology* 54 107–113.

Ghislain, M., Grunberg, W. and Benavides, J. (2004). Application of molecular markers for gene pool division and heterosis estimation under drought conditions in sweet potato, CIP, Lima, Peru.

Griesbach, J. (1992). A guide to propagation and cultivation of fruit-trees in Kenya, Technical Cooperation – Federal Republic of Germany (GTZ), Eschborn, pp. 1-97.

Haberl, M and Tautz, D. (1999). Comparative allele sizing can produce inaccurate allele size differences for microsatellites. *Molecular Ecology Notes* 8: 1347-1350.

Hansen, V. (2005). Papaya breeding and variety development. Report No. FR99018, Horticulture Australia Limited, Sydney, Australia.

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

Horticultural Crops Development Authority (HCDA), (1987). Annual report. Horticultural Crops Development Authority. Nairobi, Kenya.

Horticultural Crops Development Authority (HCDA), (2008). Validation report. Horticultural Crops Development Authority. October, 2008. Nairobi, Kenya.

Hofmeyr, J. D. J. (1938). Genetic studies of *Carica papaya* L. *South African Journal of Science* 35: 300-304.

Hofmeyr, J. D. J. (1967). Some genetic and breeding aspects of *Carica papaya*. *Agronomia Tropical* 17: 345–351

Imungi, J. K. and Wabule, M. N. (1990). Some chemical characteristics and availability of vitamin A and C from Kenyan varieties of papaya (*Carica papaya* L.). *Ecology of Food and Nutrition* 24: 115-120.

International Board for Plant Genetic Resources (IBPGR). (1988). Descriptors for Papaya. Rome. ISBN 92-9043-129-6, p. 1-33.

International Center for Tropical Agriculture (CIAT). (1993). Biotechnology Research Unit. *Annual report*, Cali, Colombia.

International Maize and Wheat Improvement Center (CIMMYT). (2005). Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. 3rd Edition. Mexico, D.F: CIMMYT. ISBN: 968-6923-30-6.

Janick, J. (1986). Horticultural Science (4th Edition). W. H. Freeman and Company (New York). p. 639- 641.

Kamau, J. K., Wabule, M. N. Gathungu, C. N. and Muli, S. (1993). Evaluation of some F1 pawpaw hybrids. In: Wamocho, L. S. Olembo, S. A. H. Mugai, E. N. Shiomi, S. and Murakami, M. (Eds) Proceedings of a seminar, ‘Meeting the national seed and planting material needs of Kenya agriculture for the year 2000’, JKUCAT in conjunction with JICA. p 122-128. Nairobi Kenya.

Karp, A. and Edwards, J. (1997). DNA markers: a global overview. In: Caetano-Anolles, G. and Gresshoff, P. M. (Eds) DNA markers – protocols, applications, and overviews. Wiley-Liss, Inc., New York. pp. 1-13.

Karuri., H. W. Ateka, E. M. Amata, R. Nyende, A. B. and Muigai, A. W. T. (2009). Characterization of Kenyan sweet potato genotypes for SPVD resistance and high dry matter content using simple sequence repeat markers. *African Journal of Biotechnology* 8 (10): 2169-2175.

Kim, M. S., Moore, P. H. Zee, F. Fitch, M. M. M. Steiger, D. L. Manshardt, R. M. Paull, R. E. Drew, R. A. Sekioka, T. and Ming, R. (2002). Genetic diversity of *Carica papaya* as revealed by AFLP markers. *Genome* 45: 503–512.

King, R. C. and Stansfied, W. D. (1990). A dictionary of genetics (4th Ed). Oxford University Press (New York) pp. 1-406.

Lohiya, N. K., Mishra, P. K. Pathak, N. Manivannan, B. Bhande, S. S. Panneerdoss, S. Sriram, S. (2005). Efficacy trial on the purified compounds of the seeds of *Carica papaya* for male contraception in albino rat. *Reproductive Toxicology* 20: 135–148.

Mansfield, E. S., Vainer, M. Enad, S. Baeker, D. L. Harris, D. Rappaport, E. and Fortina, P. (1996). Sensitivity, reproducibility, and accuracy in short tandem repeat genotyping using capillary array electrophoresis. *Genome Research* 6: 893-903

Manshardt, R. (1992). Papaya. In: Hammerschlag, F. A. and Litz, R. E. (Eds.) Biotechnology of perennial fruit crops. CAB international, University Press Cambridge. p 489-511.

Manshardt, R. M. and Zee, F. T. P. (1994). Papaya germplasm and breeding in Hawaii, *Fruit Varieties Journal* 48 (3) 146–152.

Markert, C. L. and Moller, F. (1959). Multiple forms of enzyme: tissue, ontogenetic and species specific patterns. *Proceedings of the National Academy of Sciences* 45: 753-763

Mart´nez, L., Cavagnaro, P. and Masuelli, R. (2005). Evaluation of diversity among Argentine grapevine (*Vitis vinifera* L.) varieties using morphological data and AFLP markers. *Elect. J. Biotechnol* 6: 37-45.

Martin, F. W. (1976). Tropical yams and their potential. Series part 3. *Dioscorea alata*. *Tropical Agriculture* (Trinidad). 54:1-13.

Martins, D. J. (2003). Pollination Ecology of Papaya (Carica papaya) on small-holder farms in Kenya. *Case studies on conservation of pollination services as a component of agricultural biological diversity* African Pollinator Initiative Environment Liaison Centre International Nairobi Kenya. 1-7.

Morgante, M., Hanafey, M. and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics* 30: 194–200.

Morshidi, M. (1998). Genetic control of isozymes in *Carica papaya* L. *Euphytica* 103: 89-94.

Morton, J. (1987). Papaya. In: Fruits of warm climates. Morton, F. J. and Miami, F. L. (Eds) 336–346.

Moyib., O. K. Odunola, O. A. and Dixon, A. G. O. (2007). SSR markers reveal genetic variation between improved cassava cultivars and landraces within a collection of Nigerian cassava germplasm. *African Journal of Biotechnology* Vol. 6(23): 2666-2674.

Nakasone, H. Y. and Aragaki, M. (1973). Tolerance to *Phytophthora* fruit and root rot in *Carica papaya* L. *Proceedings of the American Society for Horticultural Science* 66: 168-182

Nakasone, H. Y. and Storey, W. B. (1955). Studies on the inheritance of fruiting height of *Carica papaya* L. *Proceedings of the American Society of Horticultural Science*. 66: 168-182.

Nakasone, H. Y. and Paull, R. E. (1998). Papaya. In: Tropical fruits. CAB International, Wallingford. Oxon. UK. pp. 239-269.

National Research Council (NRC). (1989). Lost crops of the Incas: little-known plants of the Andes with promise for worldwide cultivation. National Academy Press, Washington.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences* USA, 70: 3321–3323.

Neilsen, G. and Scandalios, J. G. (1974). Chromosomal location by use of trisomics and new alleles of an endopeptidase in *Zea mays*. *Genetics* 77:679-686.

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

N'Guetta, K. (1994). Inventory of insect fruit pests in northern Cote d'Ivoire. Fruits. 49 (5/6): 430-431, 502-503.

Ocampo J. P., Dambier, D. Ollitrault, P. D'Eeckenbrugge G. C. Brottier, P. and Risterucci A. (2005). *Proceedings of the Interamerican Society for Tropical Horticulture*

48: 90-93. Annual Meeting of the Interamerican Society for Tropical Horticulture. 50, 2004-10-07/2004-10-11, Limon, Costa Rica.

Ocampo, J. P., Dambier, D. Ollitrault, P. D'Eeckenbrugge, G. C. Brottier, P. Froelicher, Y. and Risterucci, A. (2006a). Microsatellite markers in *Carica papaya* L.: isolation, characterization and transferability to *Vasconcellea* species. *Molecular Ecology Notes* 6: 212–217

Ocampo, J. P., d'Eeckenbrugge, G. C. Bruyère, S. DE Bellaire, L. L. and Ollitrault, P. (2006b). Organization of morphological and genetic diversity of Caribbean and Venezuelan papaya germplasm. *Fruits* 61: 25–37.

Office of the Gene Technology Regulator (OGTR). (2008). The biology of *Carica papaya* L. (papaya, papaw, paw paw). Australian Government Department of Health and Ageing. 2:1-55.

Organization for Economic Co-operation and Development (OECD). (2003). Draft Consensus Document on the Biology of *Carica papaya* (L.) (Papaya). *Report No*. 5 February 2003, OECD, France.

Organization for Economic Co-operation and Development (OECD). (2005). Consensus document on the biology of papaya (*Carica papaya*). OECD Environment, Health and Safety Publications, Series on Harmonization of Regulatory Oversight in Biotechnology No. 33, France.

Ploetz, R. C., Zentmyer, G. A. Nishijima, W. T. Rohrbach, K. G. and Ohr, H. D. (1994). Compendium of Tropical Fruit Diseases. *American Phytopathological Society*.

Purseglove, J. W. 1968. Tropical crops – Dicotyledons. Longman, London, pp. 45–51

Rohlf, F. J. (2000). NTSYS-pc: Numerical taxonomy and multivariate analysis system. Applied Biostatistics, Exerter Publishing Ltd., New York.

Rolda´n-Ruiz., I. van Eeuwijk, F. A. Gilliland, T. J. Dubreuil, P. Dillmann, C. Lallemand, J. De Loose, M. and Baril, C. P. (2001). A comparative study of molecular and morphological methods of descri.ing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. *Theoretical and Applied Genetics* 103: 1138-1150.

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

Salem., K. F. M. El-Zanaty, A. M. and Esmail, R.M. (2008). Assessing Wheat (*Triticum*

aestivum L.) Genetic Diversity Using Morphological Characters and Microsatallite Markers. *World Journal of Agricultural Sciences* 4 (5): 538-544.

Samson, J. A. (1986). Tropical fruits (2nd Ed). Longman Scientific and Technical, New York. pp. 256- 269.

Scheldeman, X. and Van Damme, P. (2002). Highland papayas in Southern Ecuador: Need for conservation actions. *Acta Horticulturae* 575: 199-205.

Scheldeman, X., Willemen, L. d'Eeckenbrugge, C. G. Peeters, R. E. Restrepo, M. T. Motoche, R. J. Jiménez, D. Lobo, M. Medina, C. I. Reyes, C. Rodgriguez, D. Ocampo, J. A. Van Damme, P. and Goetgebeur, P. (2006). Distribution, diversity and environmental adaptation of highland papayas (*Vasconcellea* spp.) in tropical and subtropical America. Springer Science + Business Media B.V. 2006.

Semagn, K. (2002). Genetic relationships among ten endod types as revealed by a combination of morphological, RAPD and AFLP markers. *Hereditas* 137: 149-156.

Senior M. L., Chin E. C. L. Lee, M. Smith, J. S. C. Stuber, C. W. (1996). Simple sequence repeat markers developed from maize sequences found in the GENBANK database: Map construction. *Crop Science* 36:1676-1683.

Senior, M. L. and Heun, M. (1993). Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome* 36:884-889.

Singh, I. (1990). Papaya. Oxford and IBH Publishing Co. Pvt Ltd, New Delhi, India. pp. 224-225.

Smith J.S.C., Chin, E. C. L. Shu, H. Smith, O. S. Wall, S. J. Senior, M. L. Mitchell, S. E. Kresovich, S. Ziegle, J. (1997). An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. *Theoretical and Applied Genetics* 95:163-173.

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

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

Somsri, S. (1999). Improvement of papaya (*Carica papaya* L.) for South-East Queensland: investigation of sex-type and fruit quality, Australia. *New Crop Newsletter* 11 (25.2).

Sookmark, S. and Tai, E. A. (1975). Vegetative propagation of papaya by budding. *Acta Horticulturae* 49:85-90.

Stegemann, H. (1984). Retrospect on 25 years of cultivar identification by protein patterns and prospects for the future. *Seed Science and Technology* 15: 20-31.

Stiles, J. I., Lemme, C. Sondur, S. Morshidi, B. M. Manshardt, R. (1993). Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. *Theoretical and Applied Genetics* 85, 697–701.

Storey, W.B. (1938). Segregations of sex types in solo papaya and their application to the selection of seed. *Proceedings of the American Society of Horticultural Science* 35:83-85.

Storey, W. B. (1941). Papaya production in the Hawaiian Islands. Part 1. The botany and sex relationships of the papaya. *Bulletin No*. 87, Hawaii Agricultural Experiment Station, University of Hawaii, Honolulu 87: 5–22.

Storey, W. B. (1969). Papaya. In: Outlines of perennial crop breeding in the tropics. Ferwerda, F. P. Wit, F. Veenman, H. and Zonen, N. Wageningen, V. (Eds) The Netherlands. pp. 389–408.

Storey, W. B. (1976). Papaya. In: Evolution of crop plants. N.W. Simmonds (Ed) Longman, London. pp. 21–24.

Stuber, C. W., Polacco, M. and Senior, L. (1999). Synergy of empirical breeding, marker assisted selection, and genomics to increase crop yield potential. *Crop Science* 39: 1571-1581.

Tautz, D. (1989). Hypervariability of simple sequences as general source for polymorphic DNA markers. *Nucleic Acids Research* 17: 6463–6472.

Teaotia, S. S. and Singh, R. N. (1967). Seasonal variation in sex expression of papaya. *Indian Agriculturalist* 11, 45-9.

Van Droogenbroeck, B., Breyne, P. Goetghebeur, P. Romeijn-Peeters, E. and Kyndt, T. Gheysen, G. (2002). AFLP analysis of genetic relationships among papaya and its wild relatives (Caricaceae) from Ecuador. *Theoretical and Applied Genetics* 105, 2876–2297.

Van Droogenbroeck, B., Kyndt, T. and Maertens, I. (2004). Phylogenetic analysis of the highland papayas (Vasconcellea) and allied genera (Caricaceae) using PCR-RFLP. *Theoretical and Applied Genetis* 108, 1473–1486.

Villegas, V. N. (1997). Edible fruits and nuts - *Carica papaya* L. In: Verheij, E. W. M and Coronel, R. E. (Eds) volume 2. Wageningen University, Netherlands.

Vollmann., J. Grausgruber, H. Stift, G. Dryzhyruk, V. and Lelley, T. (2005). Genetic diversity in camelina germplasm as revealed by seed quality characteristics and RAPD polymorphism. *Plant Breeding* 124:446-453.

Wang, Z., Weber, J. L. Zhong, G. and Tanksley, S. D. (1994). Survey of plant short tandem DNA repeats. *Theoretical and Applied Genetics* 88: 1–6.

Webb, R. R. (1985). Epidemiology and control of bacterial canker of papaya caused by an *Erwinia* sp. on St. Croix US Virgin Islands. *Plant Dis* 69 (4) 305–309.

Weising, K., Nybom, H. Wolff, K. and Kahl, G. (2005). DNA fingerprinting in plants: principles, methods, and applications (2nd Ed). CRC Press, London.

Wenz, H. M., Robertson, J. M. Menchen, S, Oaks, F. Demorest, D. M. Scheibler, D. Rosenblum, B. B. Wike, C. Gilbert, D. A. and Efcavitch, J. W. (1998). High-precision genotyping by denaturing capillary electrophoresis. *Genome Research* 8: 69-80.

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

Yu., J. K. Mangot, J. Thompson, L. Edwards, K. J. Slabaugh, M. B. Knapp, S. J. (2000). Allelic diversity of simple sequence repeats among elite inbred lines of cultivated sunflower. *Genome* 45: 652-660.

Zacarias, A. M. (1997). Identification and genetic distance analysis of cassava (Manihot esculenta Crantz) cultivars using RAPD fingerprinting. MSc thesis. Department of plant breeding, Faculty of Natural and Agricultural Sciences, University of the Free State, Bloemfontein, South Africa.

APPENDICES

Appendix 1: Questionnaire used to garner information on Kenyan papaya germplasm

AREA

District ____

Date ____

Location _________________________ Sub location ______________________

Village ____

HOUSEHOLD

Farmer's name ____

Age of the farmer________________________ Gender______________________

Marital status____

PAPAYA PRODUCTION

Main crop grown by the farmer_______________________________________

Farm production: Subsistence ( ) Subsistence + market ( ) Only market ( )

What proportions of papayas do you sell? 0-25% ( ) 25-50% () 50-75 % ( ) >75 % ( )

FARM DESCRIPTION

Size (acreage) of the farm: 0-2 ( ) 2-4 ( ) 4-8 ( ) >8( )

Proportions of the farm under papaya 0-2 % ( ) 2-4 ( ) 4-8 ( ) >8( )

Nature of farm production: Intercrop Main crop Abandoned

How is land preparation done? By hand ( ) Animal ( ) Power machine ( )

What is the normal planting time in the year? ______________________________

What is the source of planting material? Seeds ( ) Tissue culture ( )

Others (specify) ____

Type of seeds used ____

Spacing of the papayas ____

Field management e.g. weeding, fertilizer application (Inorganic ( ) organic ( )

How long do they take to mature? _______________________________________

What are the maturity indices? ____

Number of harvests per year Once ( ) 1-2 ( ) Continuous ( )

Yields per plant ____

Uses of papaya fruit____

Uses of Papaya plant ____

Other uses ____

Major constraints to papaya production (in order of importance i.e. 1, 2, 3…): Pests ( ) Diseases ( ) Market for produce ( ) others ( )

CHARACTERIZATION DESCRIPTORS- fruit

Farm no.	Local name/ variety	Fruit length	Fruit diameter	Fruit weight	Flesh density	Fruit shape	Uniformity of fruit distribution	Flesh color at maturity	Fruit skin texture

CHARACTERIZATION DESCRIPTORS-stem/leaf

Internode length	Stem diamet er	Tree habit	Petiole sinus shape	Petiole colour	Petiole length	Flower length	Flower color	Leaf shape	Sexual type

Appendix 2: Papaya morphological characters used in phenotypic evaluation

A	Local name/ vernacular name		**I**	Petiole length	1. Generally small 2. Generally intermediate 3. Generally large
B	Tree type/ Sexual type	1. Female 2. Hermaphrodite 3. Male	**J**	Flower size	1. Generally small 2. Generally intermediate 3. Generally large
C	Tree stem diameter	Measured at 10cm from the ground	**K**	Fruit diameter	Measured in terms of length and width
D	Tree habit	1. Single stem 2. Multiple stems	**L**	Fruit skin texture	1. Smooth 2. Intermediate 3. Rough (ridged)
E	Tree height	1. Short (<1m) 2. Intermediate 3. Tall (>1m)	**M**	Color of mature petiole	1. Pale green 2. Normal green 3. Dark green 4. Shades of red purple
F	Internode length	Measured 10cm from the ground	**N**	Uniformity of fruit distribution	0 Not uniform + Uniform
G	Color at maturity (fruit skin color)	1. Yellow 2. Deep yellow to orange 3. Red/ purple 4. Yellowish green 5. Green	**O**	Length of mature leaf (cm)	Average of 5 leaves, and measured from base of middle leaflet midrib to tip
H	Fruit flesh color	1. Light yellow 2. Bright yellow 3. Deep yellow to orange 4. Reddish orange 5. Scarlet	**P**	Fruit shape	1. Globular 2. Round 3. High round 4. Elliptic 5. Oval 6. Oblong 7. Oblong-ellipsoid 8. Oblong-blocky 9. Elongate 10. Lengthened cylindrical 11. Pear shaped (pyriform) 12. Club 13. Blossom end tapered 14. Acron (heart shaped) 15. Reniform 16. Turbinate inferior 17. Plum shaped

Appendix 3: Allele sizes (base pairs) of 42 papaya accessions

Code	Genotypes	Locus											
		mCpCIR1		mCpCIR2				mCpCIR3				mCpCIR8	
CHP1	Honey dew	316	319	276	295	0	0	190	200	0	0	289	297
CHP2	Solo	321	323	274	294	0	0	190	209	0	0	289	295
EMB1	Sunrise	322	322	274	294	0	0	190	193	0	0	294	294
IMA1	Papayi	321	321	269	280	0	0	189	191	209	211	289	294
IMA2	Papayi	323	323	266	274	294	0	191	211	0	0	288	295
IMA5	Papayi	319	319	274	276	0	0	197	211	0	0	289	289
IMA6	Papayi	319	342	294	296	0	0	191	211	0	0	295	295
IMA7	Papayi	317	321	266	274	0	0	191	193	0	0	294	297
IMA8	Papayi	317	319	267	269	298	300	189	191	0	0	295	295
JKU1	Papayi	319	319	276	296	0	0	191	211	0	0	294	294
JKU2	Papayi	319	323	274	294	0	0	189	191	0	0	289	295
JKU3	Papayi	318	318	295	295	0	0	191	191	0	0	285	295
KIB1	Papayi	317	323	272	274	0	0	189	191	0	0	300	300
KIZ3	Apoyo	319	319	276	294	0	0	189	191	0	0	289	294
KIZ4	Apoyo	319	319	276	296	0	0	189	191	0	0	294	294
KLF1	Papayi	321	321	266	276	0	0	211	211	0	0	294	294
KLF2	Papayi	319	319	274	296	0	0	190	211	0	0	289	295
KOS1	Sunrise	321	323	242	294	0	0	191	191	0	0	300	300
KOS2	US	321	334	274	276	0	0	191	211	0	0	289	289
KOS3	Redlady	322	322	266	274	0	0	190	199	0	0	287	288
KOS4	US	317	317	266	294	0	0	191	191	0	0	294	294
MAN2	Papayi	319	319	266	276	0	0	189	191	0	0	294	297
MF1	Mountain	317	321	276	293	0	0	191	191	0	0	283	290
MF3	Mountain	319	322	274	276	0	0	191	191	0	0	293	294
MF2	Local	317	334	242	294	0	0	191	191	0	0	288	288
MIG1	Papayi	319	319	294	296	0	0	191	211	0	0	289	295
MIG2	Papayi	315	342	262	274	0	0	191	191	0	0	289	295
MIG4	Papayi	317	317	280	298	0	0	190	192	0	0	295	297
MIG6	Papayi	319	319	266	276	295	0	190	211	0	0	292	292
MIG9	Papayi	319	319	274	274	0	0	190	190	0	0	289	289
MRGK2	Sunrise solo	322	322	274	296	0	0	189	191	0	0	289	294
MUT	Papayi	313	318	276	276	0	0	191	201	0	0	285	295
PKR1	Kiru	321	323	274	294	0	0	190	190	0	0	288	295
PKR4	Papayi	323	325	274	294	0	0	191	191	0	0	289	295
RAP1	Apoyo	318	323	276	294	0	0	190	190	0	0	288	288
ST1	Papayi	317	323	272	274	0	0	191	191	0	0	289	295
ST2	Papayi	317	317	274	274	0	0	191	191	0	0	285	289
THK1	Local	323	323	274	294	0	0	191	194	0	0	289	295
THK2	Local	319	342	276	276	0	0	163	163	0	0	289	289
VB1	Papayi	319	319	276	294	0	0	211	211	0	0	295	297
VB2	Papayi	322	322	277	300	0	0	198	211	0	0	289	289
VOI1	Papayi	317	319	274	276	296	0	191	211	0	0	289	295

		Locus											
Code	**Genotypes**	**mCpCIR10**		**mCpCIR18**					**mCpCIR23**				
CHP1	Honey dew	342	346	293	296	0	0	0	273	283	0	0	0
CHP2	Solo	339	339	296	296	0	0	0	271	284	0	0	0
EMB1	Sunrise	343	346	293	293	0	0	0	271	284	0	0	0
IMA1	Papayi	344	350	300	300	0	0	0	281	284	0	0	0
IMA2	Papayi	340	342	294	297	0	0	0	283	283	0	0	0
IMA5	Papayi	342	346	293	293	0	0	0	281	284	0	0	0
IMA6	Papayi	340	346	294	296	0	0	0	283	283	0	0	0
IMA7	Papayi	340	343	294	296	0	0	0	272	281	284	292	0
IMA8	Papayi	346	349	300	300	0	0	0	283	283	0	0	0
JKU1	Papayi	341	341	294	296	0	0	0	283	292	0	0	0
JKU2	Papayi	339	342	293	293	0	0	0	283	283	0	0	0
JKU3	Papayi	343	343	293	296	0	0	0	281	284	0	0	0
KIB1	Papayi	341	341	294	296	0	0	0	287	289	0	0	0
KIZ3	Apoyo	341	345	293	293	0	0	0	283	283	0	0	0
KIZ4	Apoyo	337	340	294	296	0	0	0	283	293	0	0	0
KLF1	Papayi	342	347	296	296	0	0	0	283	283	0	0	0
KLF2	Papayi	342	342	296	296	0	0	0	283	283	0	0	0
KOS1	Sunrise	344	346	293	293	0	0	0	289	289	0	0	0
KOS2	US	342	346	294	296	0	0	0	271	284	287	289	291
KOS3	Redlady	347	349	294	296	299	303	305	271	281	284	287	0
KOS4	US	339	343	296	296	0	0	0	283	283	0	0	0
MAN2	Papayi	340	342	296	296	0	0	0	283	283	0	0	0
MF1	Mountain	338	338	291	294	0	0	0	284	284	0	0	0
MF3	Mountain	340	342	292	294	0	0	0	272	284	0	0	0
MF2	Local	340	342	293	293	0	0	0	281	284	0	0	0
MIG1	Papayi	340	346	296	296	0	0	0	271	284	0	0	0
MIG2	Papayi	340	342	292	296	0	0	0	283	283	0	0	0
MIG4	Papayi	342	342	288	296	299	301	0	271	281	284	287	0
MIG6	Papayi	341	341	294	295	0	0	0	283	283	0	0	0
MIG9	Papayi	339	339	294	301	0	0	0	283	283	0	0	0
MRGK2	Sunrise solo	341	345	292	294	296	0	0	284	293	0	0	0
MUT	Papayi	342	342	296	300	0	0	0	278	283	0	0	0
PKR1	Kiru	340	340	296	296	0	0	0	283	289	0	0	0
PKR4	Papayi	340	342	292	296	0	0	0	281	284	0	0	0
RAP1	Apoyo	339	339	293	293	0	0	0	271	283	0	0	0
ST1	Papayi	342	342	289	292	0	0	0	284	289	0	0	0
ST2	Papayi	340	342	297	299	0	0	0	284	284	0	0	0
THK1	Local	340	342	294	294	0	0	0	271	281	284	289	0
THK2	Local	342	342	292	296	0	0	0	284	284	0	0	0
VB1	Papayi	342	346	294	294	0	0	0	271	284	0	0	0
VB2	Papayi	346	350	300	300	0	0	0	283	283	0	0	0
VOI1	Papayi	339	339	294	296	0	0	0	284	284	0	0	0