Ksh 3,000/=

GENETIC DIVERSITY IN WEST AFRICAN PEARL MILLET LANDRACES AS REVEALED BY SIMPLE SEQUENCE REPEATS (SSRS) MOLECULAR MARKERS

BY

MARY N. ONGUBO, B.ED SCIENCE (MOI UNIVERSITY)
I56/6237/03

A thesis submitted in partial fulfillment of the requirements for the award of the Degree of Master of Science in Biotechnology in the School of Pure and Applied Sciences of Kenyatta University.

NOVEMBER 2008

Ongubo, Mary N.
Genetic diversity in West African pearl

2009/339399

KENYATTA UNIVERSITY LIBRARY

DECLARATION

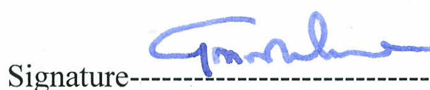
I declare that this thesis is my original work and has not been presented for any other award in any university.

Mary N. Ongubo

Signature---------------------------- Date 1-12-2008

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

Prof. Geoffrey M. Muluvi,
Department of Biochemistry and Biotechnology,
Kenyatta University,
P.O Box 43844,
Nairobi, Kenya

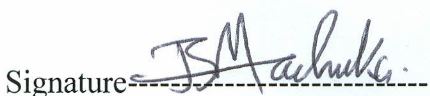


Signature---------------------------- Date 1-12-2008

Dr. Dionysious Kiambi
International Crops Research Institute for Semi-Arid Tropics (ICRISAT)
P.O. Box 30709 - 00100
Nairobi, Kenya

Signature---------------------------- Date--------------------

Prof. Jesse Machuka
Department of Biochemistry and Biotechnology,
Kenyatta University,
P.O Box 43844,
Nairobi, Kenya



Signature---------------------------- Date 1-12-2008

DEDICATION

I dedicate this thesis to my beloved husband James Onsarigo, and my dear children, Edwin Kiage Mecheo, Melody Mong'are Mecheo, and Sheilla Nyanchama Mecheo for toiling with me entirely during my studies and in writing this thesis.

ACKNOWLEDGEMENT

I would like to thank my supervisor at ICRISAT, Dr. Dionysious Kiambi for accepting me as an MSc student and showing confidence in me. Thanks Daktari, for the guidance, inspiration, discussions and moral support during my research work. Thanks also for critically reading all my maniscripts and showing a different perspective of things. I also would like to thank my university supervisors, Professors Geoffrey M. Muluvi who gave me unlimited academic guidance and Jesse Machuka for his invaluable assistance, many fruitful discussions and encouragements during my project work.

I also wish to thank the staff at the International Crops Research Institute for Semi-Arid Tropics (ICRISAT) particularly Dr. Kassa, Micheal Kimani, Mercy Kareichi, and Maggie for their friendly technical support and co-operation. This work could not have been completed without the moral support of my dear friends Bosibori Bett and Evans Nyaboga and Mercy Kitavi. Last but not least, I thank my husband, children Edwin, Melody and Sheila , my sister Jackline and my brother Thomas for always being there and making me happy.

Most of all, I thank God for his faithfulness, for answering my prayers in those lonely hours, watching over this study, my family and everything to do with me. Blessed be the name of the Lord. To all, God mightily bless you for helping me succeed in my various capacities. THANK YOU.

TABLE OF CONTENTS **PAGE**

Title		i
Declaration		ii
Dedication		iii
Acknowledgement		iv
Table of Contents		v
List of Figures		vi
List of Tables		ix
List of Appendices		x
List Abbreviations		xi
Abstract		xii
CHAPTER ONE: INTRODUCTION AND LITERATURE REVIEW		**1**
1.1	Importance and production of pearl millet	1
1.2	History and origin of pearl millet	6
1.3	Distribution of pearl millet	6
1.4	Biology of pearl millet	7
1.5	Constraints to pearl millet production	7
1.6	Germplasm characterization	8
1.7	Measuring genetic distance and prediction of heterosis	9
1.8	Heterotic groupings of pearl millet landraces	10
1.9	Genetic diversity studies	12
1.9.1	Use of morphological markers	12
1.9.2	Use of molecular markers	13
1.9.2.1	Protein markers	14
1.9.2.2	DNA markers	14
1.9.3	Use of SSRs for diversity assessment	16
1.10	Genetic diversity of pearl millet	18
1.11	Problem statement and justification	20

1.12	Hypothesis	21
1.13	Objectives	21
1.13.1	Main objective	21
1.13.2	Specific objectives	21
CHAPTER TWO: MATERIALS AND METHODS		**22**
2.1	Plant materials	22
2.2	Preparation and sampling of materials	24
2.3	DNA isolation	24
2.3.1	DNA extraction	24
2.3.2	Quantification of DNA	26
2.3.3	DNA quality checking	27
2.4	SSR primers	28
2.5	DNA amplification	30
2.6	Gel electrophoresis of PCR products	32
2.7	Capillary electrophoresis	32
2.8	Data analysis	33
CHAPTER THREE: RESULTS		**36**
3.1	Quantification of DNA and quality checking	36
3.2	Optimization of SSR primers	38
3.3.	Polymorphism of microsatellites used to characterize pearl millet accessions	39
3.4	Markers effectiveness in detecting allele variability and polymorphism	41
3.4.1	Marker quality indices, allele variability and polymorphism in the 49 individuals	41
3.4.2	Marker quality indices, allele variability and polymorphism in the 30 bulked accessions	44
3.5	Genetic diversity within 49 individuals of the pearl millet accession	48
3.5.1	Number of alleles	48

3.5.2	Polymorphism, heterozygosity and gene diversity	50
3.6	Genetic relationships of the 49 individuals	51
3.6.1	Genetic distance	51
3.6.2	Genetic similarity	52
3.6.3	Principal Co-ordinate Analysis	54
3.6.4	Genetic structure of the populations	56
3.7	Genetic diversity between 30 bulked pearl millet accession	57
3.7.1	Number of alleles	57
3.7.2	Percentage polymorphism, average heterozygosity and gene diversity	59
3.8	Genetic relationships of 30 bulked accessions of pearl millet	61
3.8.1	Genetic distances	61
3.8.2	Genetic similarity	64
3.8.3	Principal Co-ordinate Analysis based on genetic distance estimates of 30 bulked accessions	64
3.9	Genetic population structure and diversity distribution within and between countries based on 30 accessions of pearl millet	66

CHAPTER FOUR: DISCUSSION **69**

4.1	Polymorphism information content (PIC) of the SSR markers	69
4.2	Genetic diversity and heterozygosity	71
4.3	Population structure	71
4.4	Genetic relationship	75

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS 77

5.1 Conclusion 77

5.2 Recommendation 78

CHAPTER SIX: REFERENCES 79

APPENDICES 90

LIST OF FIGURES

Figure 1	Picture showing pearl millet	2
Figure 2	Global pearl millet production in the year 2005	4
Figure 3	Production contributions of the top 10 millet producing countries in 2005	5
Figure 4	Agarose gel electrophoresis of DNA extraction from fresh leaves of 17 bulked samples and 17 individual	36
Figure 5	Agarose gel showing the optimization of 5 SSR primers	38
Figure6	PSMP 2080 before optimization	38
Figure 7	PSMP 2080 after optimization	39
Figure 8	Polymorphism of the microsatelites amplified by primer PSMP 2008 in 20 individuals and PSMP 2001 in 15 bulks of pearl millet	40
Figure 9	Electrophenograms showing homozygosity	40
Figure 10	Electrophenograms showing heterozygosity	41
Figure 11	Total number of alleles of the ten accessions of pearl millet	48
Figure 12	Dendrogram based on genetic relations of 49 individuals from ten different accessions using UPGMA algorithm method of Darwin version 4.0.	53
Figure 13	Principle co-ordinate analysis of the 49 individuals	54
Figure 14	Allele variation in the 30 pearl millet accessions from West Africa	58
Figure 15	Allele richness in five West African countries	58
Figure 16	Dendrogram of the relationships between 30 accessions from different West Africa countries using the UPGMA method of Darwin ver. 4.0.	64
Figure 17	Principle co-ordinate Multi-Dimensional Scatter plot showing genetic relationships of 30 bulked accessions	65
Figure 18	Dendrogram of distribution of diversity between countries based on 30 accessions using UPGMA algorithm of PowerMarker ver. 3.25	68

LIST OF TABLES

Table 1 Importance of millet production in selected countries of sub-Saharan Africa 5

Table 2 Country of origin, location and other distinguishing characteristics for pearl millet accessions used in this genetic diversity study 23

Table 3 Summary of pearl millet microsatellite primer pairs used in the study 28

Table 4 PCR conditions for each SSR primer 31

Table 5 DNA quantification and quality values using Nanodrop 37

Table 6 Quality indices and characterization of polymorphism detected by 21 markers in 49 individuals 43

Table 7 The number and sizes of alleles detected in 49 individuals across 21 loci 44

Table 8 Quality indices and characterization of polymorphism detected by 21 markers in 30 bulked accessions 46

Table 9 The number and sizes of alleles detected in 30 bulked accessions across 21 loci 47

Table 10 Total number of alleles from 49 individuals in ten pearl millet accessions 49

Table 11 Polymorphism and within accession diversity indices 50

Table 12 Genetic distance matrices between individuals of ten different accessions 51

Table 13 AMOVA summary table calculated according to Weir (1996) 55

Table 14 The Chi-square test for deviations from Hardy-Weinberg equilibrium across the 21 loci 56

Table 15 Percentage polymorphism, average heterozygosity and gene diversity of 30 bulked accessions 60

Table 16 Genetic distance matrices of the 30 bulked accessions calculated according to Dice (1945) 62

Table 17 Genetic distance matrices between countries calculated according to Nei and Li (1979) using PowerMarker ver. 3.25 66

KENYATTA UNIVERSITY LIBRARY

LIST OF APPENDICES

Appendix 1. Percentage polymorphism, average heterozygosity and gene diversity of 30 bulked accessions 90

Appendix 2. Buffers and reagents 91

LIST OF ABBREVIATIONS

AFLP	amplified fragment length polymorphism
Bp	base pair(s)
CTAB	cetyl trimethyl ammonium bromide
DNA	deoxyribonucleic acid
dNTP	2', 3'-deoxyribonucleoside 5'-triphosphate
EDTA	ethylendiamine tetra-acetic acid
ORF	open reading frame
PCR	polymerase chain reaction
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNAse	ribonuclease
Rpm	revolutions per minute
SSR	simple sequence repeats
Taq	Thermophilus aquaticus
TBE	tris borate EDTA
Tris	tris (hydroxymethyl) aminomethane
UV	ultraviolet
OPVs	Open pollinated varieties

ABSTRACT

Pearl millet (*Pennisetum glaucum* [L.] R. Br.) is one of the most important staple food crops in the world, particularly in arid and semi-arid regions. In West Africa, pearl millet covers more than 65% of the total cultivated area. However, genetic diversity of this important crop in West Africa is poorly documented, although it is important for identification of landmarks for *in situ* germplasm conservation and to increase knowledge useful for breeding programs, particularly in heterotic studies. The overall aim of this study was therefore to understand the patterns of diversity within and among populations in order to enhance those populations with continuous genetic variation for association and heterotic studies. Genetic relationships and diversity among and within 30 pearl millet accessions from West Africa was assessed using twenty one Short Sequence Repeats (SSRs) markers. The markers revealed 172 alleles among 30 populations and 171 alleles within 10 populations, with the Polymorphic Information Content (PIC) ranging from 0.39 to 0.82. The mean Nei's genetic diversity across all the 21 loci in the 30 populations was 0.647 while the mean Nei's diversity within the 10 populations studied ranged from 0.402 to 0.835 with an average of 0.681. The partitioning of variance components based on the analysis of molecular variance (AMOVA) for diversity analysis revealed high variability within accessions (84.59%) compared to the variability between accessions (9.56%). The fixation index (Fst) obtained was 0.154 indicating a moderate differentiation among the accessions. This study estimated genetic diversity in a highly cross-pollinated crop-pearl millet and generated data on genetic diversity of 30 accessions from west African that could be used to select potential heterotic groups that can be exploited in breeding programs to increase hybrid vigour.

CHAPTER ONE

1.0 INTRODUCTION AND LITERATURE REVIEW

1.1 Importance and production of pearl millet

Pearl millet (*Pennisetum glaucum* [L.] R. Br.) ranks as the sixth most important cereal in the world. Pearl millet is one of the most extensively cultivated cereals in the world, after rice, wheat, and sorghum, particularly in arid and semi-arid regions (Oelke *et al.*, 1990). It is currently planted on over 14 million ha in Africa and approximately 500 million people depend on it for their survival (Wikipedia, 2006). It is a principal food cereal cultivated in drought-prone semi-arid regions of Africa and Indian subcontinent. In the USA, Australia, Southern Africa, and South America, pearl millet is grown most extensively as a forage crop. There are about 6,000 varieties of millet (Andrews *et al.*, 1996), categorized into several main types. Of all the types, pearl millet produces the largest seeds and is the most common variety used for human consumption (Andrews *et al.*, 1996).

Pearl millet is an excellent dietary source of calcium, iron, manganese, and methionine, an amino acid lacking in the diets of hundred of millions of the poor who live on starchy foods such as cassava, plantain, polished rice, and maize meal (CGIAR, 2005). It has diverse uses especially in preparation of porridge and kasha, soups, breads and stuffings, fermented beverages, and baby food (ARC, 2006). In addition to grain and forage uses, pearl millet crop residues and green plants are a good source of animal feeds, building materials, and fuel for cooking,

particularly in dry land areas. Pearl millet has higher crude protein (CP) levels than corn silage, but it has lower energy content.

Figure 1: A picture showing pearl millet almost ready for harvesting

KENYATTA UNIVERSITY LIBRARY

From a nutritional standpoint, pearl millet is an attractive feed grain. Several studies indicate that metabolizable energy of pearl millet for non-ruminant animals is approximately equal to that of maize (Burton et al., 1974). When compared to maize on a weight basis, pearl millet is 8%-60% higher in crude protein, 40% richer in lysine and methionine, and 30% richer in threonine (Burton et al., 1974). Therefore, substitution of pearl millet for maize reduces the need for high protein feed ingredients and supplemental amino acids. Pearl millet has fewer anti-nutritional factors than most grain crops. In contrast to rye and sorghum, pearl millet grain is low in tannins, which limit palatability and inhibit protein digestion.

Millets are grouped together when reporting global millet production, therefore it is difficult to determine the production levels of each individual genus. As a general estimate, millet production is broken down into 50% - pearl millet (Pennisetum glaucum), 30% - proso (Panicum miliaceum) and foxtail millet (Setaria italica), 10% - finger millet (Eleusine coracana) and 10% - others, including barnyard millet (Echinochloa crusgalli) and kodo millet (Paspalum scrobiculatum) (FAOSTAT, 2005). Figure 2 shows all 2005 millet production records from countries producing at least 1 metric ton annually. Together, these countries produce more than 88% of the global millet. Figure 3 shows production levels of the top 10 millet producing countries in 2005. African farmers produce 40% of the world's millet (Vietmeyer, 1996). Five countries in West Africa (Nigeria, Niger, Mali, Burkina Faso and Senegal) produce 85% of the continent's total millet crop. Sudan accounts for 50% of millet production in eastern and southern Africa. Table 1 shows the relative importance of millet production

in West Africa compared to the other two regions. West Africa is also the only region where millet production has significantly increased over time. However, all of this growth is due to increased area cultivated, and not increased yields.

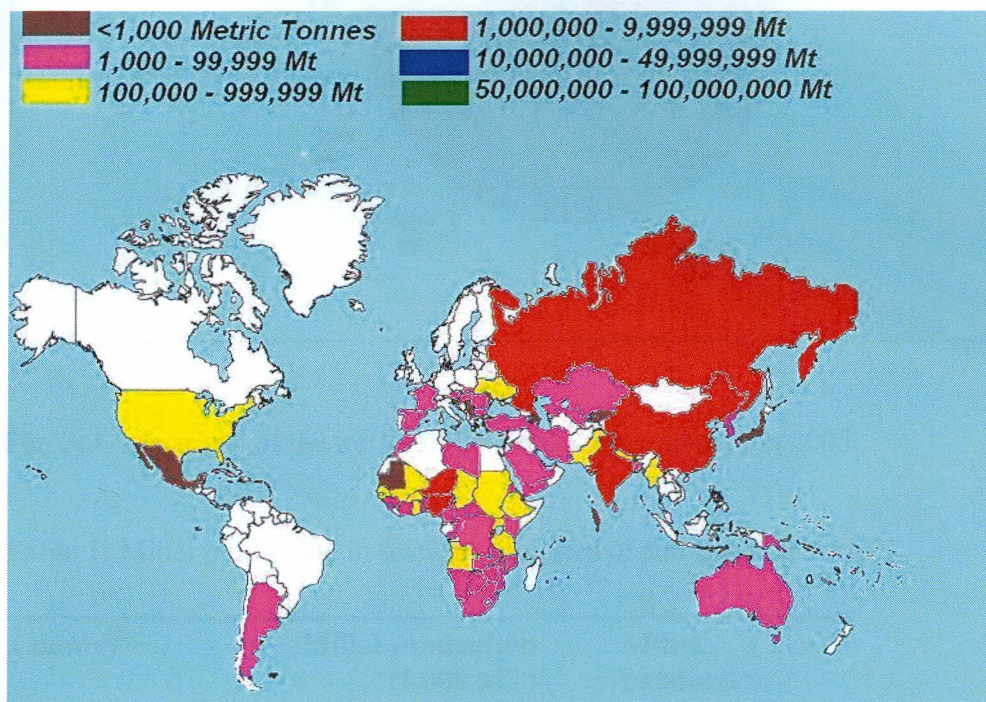


Fig. 2: Global 2005 millet production in countries producing at least 1 Metric ton (FAOSTAT, 2005).

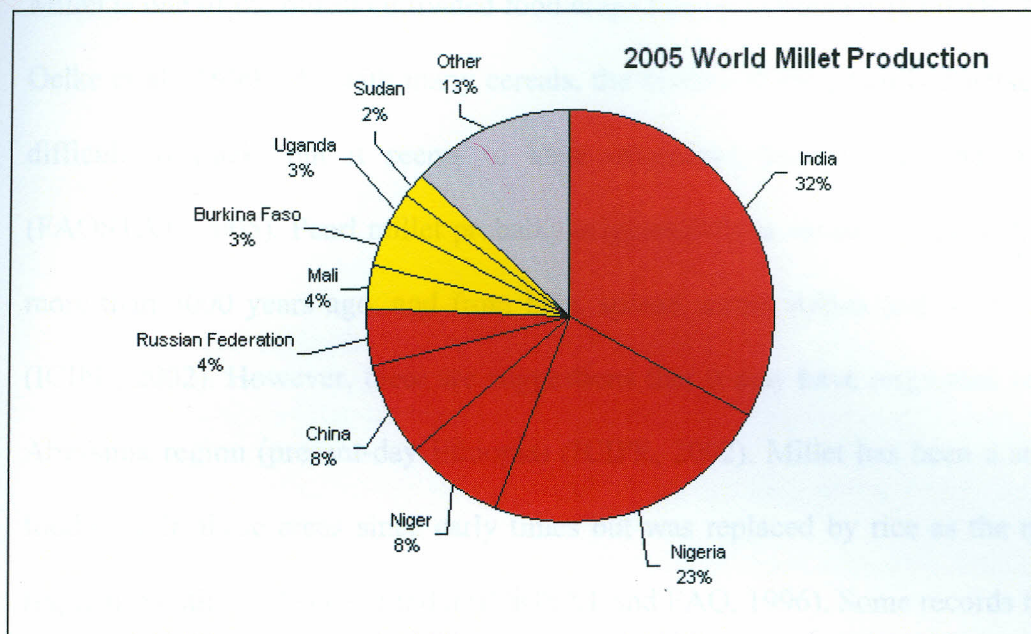


Fig. 3: Comparison of the top 10 millet producing countries in 2005.

Table 1: Millet production in selected countries of sub-Saharan Africa

Country	Millet production (1,000 MT)	Millet production (% of total cereals)
Burkina Faso	734	32
Chad	228	31
Mali	815	38
Niger	1769	91
Nigeria	4952	26
Senegal	667	75
Sudan	385	13
Africa	11740	10

Source: FAOSTAT, 2005.

1.2 History and origin of pearl millet

Millet is one of the oldest cultivated food crops known to humans (CGIAR, 2005, Oelke et al., 1990). As with many cereals, the history of the grain is diverse and difficult to track, but it seems to have originated from China and Africa (FAOSTAT, 2005). Pearl millet probably originated from western tropical Africa more than 3000 years ago, and from here, spread across Africa and South Asia (ICIPE, 2002). However, there are suggestions that it may have originated in the Abyssinia region (present-day Ethiopia) (ICIPE, 2002). Millet has been a staple food crop in these areas since early times but was replaced by rice as the main staple in Southeast Asia and India (ICRISAT and FAO, 1996). Some records from China indicate that millet was grown as early as 4500 BC, or possibly earlier, while other records indicate that several varieties were introduced in China from Africa (Wikipedia, 2006). Other reports suggest that it was grown by the lake dwellers of Switzerland during the Stone Age (Oelke et al., 1990). Millet was introduced to the U.S. in 1875, but was not well accepted for human consumption.

1.3 Distribution of pearl millet

Pearl millet has been used as a cereal for nearly the last 3000 years in Africa and parts of the Near East. The crop is cultivated for both forage and grain. It is grown on about 26 million ha in many countries of southern, eastern, western, and central Africa, a few countries of Asia, particularly in India and in some parts of the Americas, and Australia (Andrews and Bramel-Cox, 1994; Wikipedia, 2006). It is the essential crop in the semi arid regions of the world stretching over 7000

km from China to Somalia (almost 1/6 of the globe at the latitude). African farmers produce 40% of the world's millet (Vietmeyer, 1996).

1.4 The biology of pearl millet

Pearl millet is a tall, robust, erect, annual bunchgrass. Its height varies depending upon the variety and environment, but plants are generally 6-15 ft in height (Barker, 2003). The leaves are grass-like, numerous and slender, measuring about an inch wide and up to more than 6 feet long. Stems are pithy and leaves are long-pointed with finely serrated margins (Barker, 2003). The plant tillers produce an inflorescence with a dense spike-like panicle. The mature panicle is brownish in color, and spikelets are borne in fascicles of two, surrounded by a cluster of bristles. Each spikelet has two florets, one of which is generally staminate. The upper floret is fertile, with the caryopsis (seed) being enclosed by the lemma and palea from which it threshes free during harvest.

1.5 Constraints to pearl millet production

Major production impediments include: fungal diseases, pests, low soil fertility and labor production constraints, especially in Africa. The major fungal diseases include downy mildew (caused by *Sclerospora graminicola*) and ergot. Downy mildew has been a major problem in India because of the cultivation of genetically uniform single cross hybrids that made the crop vulnerable to downy mildew epidemics. Control of these fungal diseases can be achieved through the use of resistant varieties and pulling and burning infected plants. Pearl millet varieties with high resistance to both ergots and downy mildew have been

selected at ICRISAT. Mildew and other forms of seed rots can also be controlled with appropriate fungicides (e.g. Apron 25W) applied as seed treatment during planting. However, in the African Sahel and other semi-arid communities where subsistence agriculture is still being practiced, use of chemical control may not be affordable or accessible for the farmers.

The most important pests include the parasitic weed *Striga hermonthica* (Del.) Benth., birds, grasshoppers and armyworms, which are capable of causing economic damage. The most notorious bird species is *Quelea quelea* and damage is serious in pearl millet since the grains are palatable and smaller in size. Efficient bird scaring and use of awned varieties that prevent severe damage remain the cheapest control options available. Insect problem is generally controlled using the appropriate insecticides such as Sevin Plus (for grasshoppers) and ambush super or DiPel (for armyworms).

1.6 Germplasm characterization

Germplasm characterization refers to the observation, measurement and documentation of heritable plant traits in a collection. The resulting data allows the identification and classification of accessions, building a catalog of descriptors with embedded biological information that is essential to collection management or to direct use in agriculture and crop improvement. Germplasm characterization aims at the description and understanding of the genetic diversity of the organism under study. Today, germplasm is being characterized based mostly on morphological descriptors and molecular markers. Morphological descriptors are reliable, easy to study and relatively low cost to evaluate. However, they have

some limitations as they are influenced by environmental conditions. Due to these limitations, molecular markers are now rapidly being adopted by crop improvement researchers globally as an effective tool for germplasm genetic characterization and basic studies addressing biological components in agricultural production system (Jones *et al.*, 1997; Mohan *et al.*, 1997;). These markers also offer specific advantages in assessment of genetic diversity and in trait-specific crop improvement. Morphological traits cannot, however, be replaced by any of the molecular techniques. The results of the molecular or biochemical studies should be considered as complementary to morphological characterization, and the final aim should be to link polymorphism at the DNA level to phenotypic characteristics.

1.7 Measuring genetic distance and prediction of heterosis

In order to predict mean performance of a hybrid, it is important to select parents that can ensure a sufficient genetic variance. Plant breeders are looking for methods to enable them identify a progeny that can maintain performance standards for most of the agronomic characters and exceed the performance level of a target variety in a few characters. Heterosis is a phenomenon in which an F1 hybrid of two genetically dissimilar (diverse), homozygous parents shows increased vigor over the mid -parent value. The term "better-parent heterosis" refers to hybrid superiority over the better crossing parent. Only certain crosses (F1 hybrids) express heterosis, which is caused by a combination of partial or complete dominance or epistasis (Moll and Stuber, 1974) at loci controlling the trait. For achieving real heterotic yield gains, it is essential to identify

heterotic groups, i.e., genotypes or populations with optimal genetic divergence that produce high yielding progenies after crossing. Parameters used for such an identification of heterotic groups include i) *per se* performance, ii) mitochondrial complementation, iii) general and specific combining ability and iv) genetic diversity determined by geographical origin or through the use of neutral markers (Brummer, 1999).

1.8 Heterotic groupings of pearl millet landraces

Pearl millet is a highly cross-pollinated crop with 75–80% natural outcrossing (Burton, 1974). An extensive survey of pearl millet literature showed 40% average better-parent heterosis for grain yield, ranging from −57 to 424% (Virk, 1988). These two features of the crop satisfy some of the basic requirements for the exploitation of heterosis in cultivar development. Commercial exploitation of heterosis in pearl millet only became possible with the development of a cytoplasmic-nuclear male-sterile line Tift 23A in the USA (Burton, 1965). The first single-cross grain hybrid (HB-1) based on this male-sterile line, released in India in 1965, out-yielded improved local open pollinated varieties (OPVs) by as much as 100% (Athwal, 1965). The extent of grain yield advantage and release of HB-1 stimulated substantial pearl millet grain hybrid research and development activity in India. As a result, more than 18 hybrids are now grown on varying scales, covering more than half of the 10×10^6 ha area under pearl millet in India, and doubling the grain yield over the pre-hybrid era (Rai *et al.*, 1997).

Single-cross hybrids may not provide a good starting point for the exploitation of heterosis in hybrid cultivars in much of Africa, especially western Africa. This is because genetic uniformity of single-cross hybrids makes them more vulnerable to downy

mildew, *Sclerospora graminicola*, as has repeatedly happened in India (Hash *et al.*, 1997; Rai *et al.*, 1997). This problem is likely to be more challenging in western Africa where downy mildew pathotypes have been shown to be relatively more virulent than those in much of India (Singh *et al.*, 1997). Secondly, pearl millet displays a high degree of inbreeding depression. Thus, if seed production is undertaken in the main crop season, characterized by relatively harsher environmental conditions, inbred lines of single-cross hybrids would have poor emergence and vigour, leading to economically unacceptable seed yields in the hybrid seed production plots. Thirdly, inbred parents of single-cross hybrids take considerable time and research investment for development and testing, and the material and manpower resources in the African regions are inadequate to undertake such activities. Thus, the immediate alternative could be to use existing improved OPVs as male parents to breed inter-population hybrids on male-sterile populations, which would address all three issues mentioned above. Because of their adaptation and acceptance by farmers, successful OPVs would make good hybrid parents for inter-population hybrids, thus further capitalizing on OPV breeding. Improved OPVs are already being used as pollen parents to breed top-cross hybrids of pearl millet for African agricultural situations (Anand *et al.*, 1999).

Commercial success of top-cross and inter-population hybrids would depend on the extent of their grain yield advantage over OPVs. Preliminary results from western Africa have shown inter-population hybrids out yielding their improved OPV parents by 27–59% (Lambert, 1983) and 32–45% (Ouendeba *et al.*, 1993). This level of grain yield advantage is of significant value for commercial viability of inter-population hybrids.

1.9 Genetic diversity

Studying the diversity of important crops enables identification of landmarks for *in situ* germplasm conservation, the creation of core accessions for genetic analysis and the extension of knowledge useful for breeding programs (Mariac *et al.*, 2006). The loss of genetic diversity, in part due to the conventional breeding programs associated with modern agricultural practices, has been very dramatic for many cultivated species (Wilkes, 1983). Consequently, the narrow genetic base of the elite germplasm has increased the potential vulnerability to pests and abiotic stress. Better knowledge of the genetic similarity of breeding materials could help to maintain genetic diversity and long-term selection gain. Furthermore, monitoring the genetic variability within the gene pool of elite breeding materials would make crop improvement more efficient by the directed accumulation of favored alleles thus decreasing the amount of material to be screened.

1.9.1 Use of morphological markers

Morphological descriptors are reliable, easy to study and relatively low cost to evaluate. Phenotypic diversity index of morphological characters has been used previously to measure genetic relationships within cereal crop species. Examples include tef (Assefa *et al.*, 1999) and barley (Bekele, 1984; Negassa, 1985; Tolbert *et al.*, 1979; Demissie and Bjomstad, 1996). In the last several years, some researchers have focused their activities on conservation and characterization of pearl millet germplasm. These works have highlighted the existence of a wide

phenotypic and genotypic variation in terms of morphological, phenological and productive variation and climatic adaptation. The conclusion of these efforts is that most of the collected accessions are landraces with high adaptation to specific environmental conditions. Morphological characters have also been used to estimate phenotypic variation among landraces accessions grown in North Shewa and south Welo region of Ethiopia (Teshome *et al.*, 1997). Morphological variation does not reliably reflect the real genetic variation because of genotype-environment interaction and the largely unknown genetic control of polygenically inherited morphological and agronomic traits (Smith *et al.*, 2000).

1.9.2 Use of molecular markers

Molecular markers provide information on the global genetic structure of the species typically by using twenty to fifty neutral markers. Molecular information of germplasm provides the entry point for plant breeding efforts, providing information on responses to biotic and abiotic stresses and farmer and market-preferred characteristics. Molecular markers are rapidly being adopted by crop improvement researchers globally as effective and appropriate tools for basic and applied studies addressing biological components in agricultural production systems (Jones *et al.*, 1997; Mohan *et al.*, 1997). These markers offer several advantages in assessment of genetic diversity and in trait specific crop improvement. Use of markers in applied breeding programs can range from genetic diversity studies and facilitating appropriate choice of parents for crosses to mapping/tagging of gene blocks associated with economically important traits.

1.9.2.1 Protein markers

Protein markers including seed storage proteins, structural proteins and isozymes were among the first group of markers exploited for genetic diversity assessment and genetic linkage map development (Oumar *et al*., 2005). They are the basis for a newly emerging research area called proteomics. They also provide some of the most cost effective tools for data point generation especially when iso-electric focusing equipment is used to precisely distinguish between similar versions of proteins. The major limitations of these markers are that much of the genome does not code for genes, different biochemical procedures are required to visualize allelic differences for enzymes having different functions and many proteins are several post-transcriptional steps removed from underlying DNA sequence polymorphisms and thus can mask variation present at that level (Semagn and Ndjiondjop, 2006).

1.9.2.2 DNA markers

The best markers for genetic diversity studies are largely based on requirements for cost-effectiveness and high through-put to allow analyses of a large number of samples, preferably with co-dominant markers (Westman and Kresovich, 1997). Further, the markers need to be neutral and unlinked in order to give unbiased genetic diversity estimates (Devos and Gale, 1997). They can be used to describe baseline diversity for collection purposes, assess genetic diversity distribution within traditional cropping systems at field, village, and regional levels, and relate this information to farmers' seed management practices. Many studies have aimed

at assessing the genetic diversity in germplasm collections of crops using various types of molecular markers such as Restriction Fragment Length Polymorphism (RFLP) (Dubreuil and Charcosset, 1998), Random Amplified Polymorphic DNA (RAPD) (Fotana *et a*l., 1997) and Amplified Fragment Length Polymorphism (AFLP) (Tohme *et al.*, 1996; Zhu *et al.*, 1998). Recently, microsatellites or SSR (Simple Sequence Repeats) loci, which correspond to tandemly repeated DNA with a very short repeat unit, have been introduced as powerful genetic markers in plants (Morgante and Olivieri, 1993; Powell *et al.*, 1996a).

Diversity Arrays Technology (DArT) is another novel marker system suitable for the analysis of crop genetic diversity (Jaccoud *et al.*, 2001). A DArT marker is a segment of genomic DNA, the presence of which is polymorphic in a defined genomic representation. It was developed to provide a practical and cost-effective whole-genome fingerprinting tool. DArT has three key attributes of interest to plant breeders and scientists studying and managing genetic diversity: (a) it is independent from DNA sequence, (b) the genetic scope of analysis is defined by the user and easily expandable, and (c) the method provides for high-throughput and low-cost data production. It is able to discover hundreds of markers in a single experiment (Jaccourd *et al.*, 2001). DArT markers are typed in parallel, using high-throughput platforms, with a low cost per data point. DArT fingerprints will be useful for accelerating plant breeding, and for the characterization and management of genetic diversity in domesticated species as well as in their wild relatives (Jaccoud *et al.*, 2001).

DArT markers can be used as any other genetic marker. With DArT, comprehensive genome profiles are becoming affordable for virtually any crop, regardless of the

molecular information available for the crop (Jaccoud *et al.*, 2001). DArT genome profiles are very useful for the recognition and management of bio-diversity, for example in germplasm collections. In plant breeding, DArT genome profiles enable breeders to map QTL in one week, thereby allowing them to focus on the most crucial factor in plant breeding: reliable and precise phenotyping (Wittenberg *et al.*, 2004). Once many genomic regions of interest are identified in many different lines, DArT profiles accelerate the introgression of a selected genomic region into an elite genetic background (for example by marker-assisted backcrossing) (Wenzel *et al.*, 2004). Furthermore, DArT profiles can be used to guide the assembly of many different regions into improved varieties. For that purpose, dense genome cover is essential in order to follow many regions simultaneously. Because of the large number of lines to be typed, high throughput and affordability are critical factors in this context

1.9.3 Use of SSRs for diversity assessment

SSRs are abundantly distributed throughout the nuclear genomes of all studied plant species, which makes them useful both for genetic mapping and for the study of natural populations. SSRs have several advantages over other DNA markers such as RFLPs, RAPDs or AFLPs. They are co-dominant and highly informative. They generally display high levels of polymorphism (Beckmann and Soller, 1990; Brown *et al.*, 1996; Senior *et al.*, 1998) and are amenable to automated genotyping strategies. They can also be amplified by PCR and efficiently detect DNA polymorphism (Pejic *et al.*, 1998). Finally, radioisotopes are not required in the detection of SSR markers, because sequence polymorphism

usually can be detected by separation in agarose gels (Burr, 1994). SSRs (microsatellites) have become the molecular markers of choice for a wide range of applications in genetic mapping and genome analysis (Chen *et al.*, 1997; Li *et al.*, 2000), genotype identification and variety protection (Senior *et al.*, 1998), seed purity evaluation and germplasm conservation (Brown *et al.*, 1996), diversity studies (Xiao *et al.*, 1996; Budak *et al.*, 2003), paternity determination and pedigree analysis (Ayres *et al.*, 1997; Bowers *et al.*, 1999), gene and quantitative trait locus analysis (Blair and McCouch, 1997; Koh *et al.*, 1996) and marker-assisted breeding (Ayres *et al.*, 1997; Weiring *et al.*, 1998). For measuring genetic diversity, assigning lines to heterotic groups and genetic fingerprinting, SSRs provide a power of discrimination equal to or greater than that of RFLP in a more cost-effective manner (Smith *et al.*, 1997; Senior *et al.*, 1998). Studies have shown that SSR loci give good discrimination between closely related individuals in some cases even when only a few loci were employed (Powell *et al.*, 1996b; Scott *et al.*, 1999; Kong *et al.*, 2000). The analysis of SSRs has been automated, thereby facilitating data exchange among researchers (Saghai-Maroof *et al.*, 1994; Powell *et al.*, 1996b). SSRs markers have also been shown to provide the highest level of information per single marker when used to detect genetic similarity among maize inbred lines (Pejic *et al.*, 1998).

SSR markers have been useful for integrating the genetic, physical and sequence-based maps in plant species and simultaneously have provided an efficient tool to link phenotypic and genotypic variation (Gupta and Varshney, 2000). They have

been identified in many plant genomes including those of maize (Senior and Heun, 1993), soybean (Akkaya *et al.,* 1992; Morgante and Olivieri, 1993), barley (Saghai-Maroof *et al.,* 1994), sorghum (Brown *et al*, 1996; Dean *et al.*, 1999; Ayan *et al.*, 2000) and pearl millet (Chowdari *et al*, 1998a; ICRISAT, 1999; Qi *et al.*, 2001; 2004 and Budak *et al.*, 2003). The results of using SSR markers in these species suggest that they may provide an outstanding tool for genetic analysis of plant species hence their use in the present studies.

1.10 Genetic diversity in pearl millet

Pearl millet is a highly cross-pollinating species and genetic diversity is distributed both within and among cultivars. Within-cultivar diversity can be very limited in case of single-cross hybrids (Haussmann *et al.*, 2006), but is substantially greater in landraces and improved open-pollinated varieties of pearl millet. Due to the wide range of stressful environments in which it has traditionally been cultivated, pearl millet exhibits a tremendous amount of diversity at both phenotypic and genotypic levels (Liu *et al.*, 1992; 1994a, b). The more complicated distribution of diversity in pearl millet, as well as the higher degree of marker polymorphism, makes genetic diversity studies in this species more difficult than in self-pollinated species. In all crop species, phenotypic estimates of genetic diversity are biased by the environment(s) in which evaluation occurs (Busso *et al.*, 1995). Furthermore, in pearl millet and other cross-pollinated seed-propagated species, these estimates can also be heavily influenced by inbreeding depression that occurs if a closed population structure is

imposed during regeneration (Devos *et al.*, 1995). For genetic diversity assessment in cross-pollinated species, molecular markers offer considerable advantages over methods based on phenotypic evaluation. This is especially true in pearl millet where there is a high degree of marker polymorphism.

The best markers for genetic diversity studies are largely based on requirements for cost-effectiveness and high through-put to allow analyses of a large number of samples, preferably with co-dominant markers. In addition, the markers need to be neutral and preferably unlinked in order to give unbiased genetic diversity estimates (Devos and Gale, 1997). They can be used to describe baseline diversity for collection purposes, assess pearl millet genetic diversity distribution within traditional cropping systems at field, village, and regional levels, and relate this information to farmers' seed management practices. To date, the diversity of pearl millet has been studied using iso-enzyme loci (Tostain *et al.*, 1987; Totstain, 1992; 1994), AFLP markers (Vom Brocke *et al.*, 2003) and RFLP markers (Bhattacharjee *et al.*, 2002). New markers such as SSCP-SNP (Bertin *et al.*, 2005) and microsatellite loci (Qi *et al.*, 2001; 2004; Budak *et al.*, 2003) have recently been developed. However, they have not yet been used to assess the genetic diversity of west African landraces of pearl millet. SSR markers are useful for a variety of applications in plant genetic and breeding because of their reproducibility, multi-allelic nature, codominant inheritance, relative abundance and good genome coverage (Powell *et al*, 1996a, b). SSR markers have been useful for integrating the genetic, physical and sequence-based physical map in plant species and simultaneously have provided an efficient tool to link

phenotypic and genotypic variation (Gupta and Varshney, 2000). Therefore, SSRs are the suitable DNA-marker system for use in the study of genetic diversity of pearl millet.

1.11 Problem statement and justification

Diversity study of pearl millet identifies genetically distinct groups (potential heterotic groups, can also help in tagging resistant genes to diseases , identify QTLs for Marker_Assisted Breeding , all of which ,can be exploited in population crosses to increase hybrid vigour in the progeny. The development of high yielding and stable varieties requires a continuous supply of new germplasm as a source of desirable genes and/or gene complexes. The primary sources of such genes are landraces and weedy and wild relatives of the crop plants. Use of genetically divergent parental materials can also enhance the level of heterozygosity and therefore hybrid vigor in pearl millet crossing pro genies.. However, genetic diversity and population structure of pearl millet has not been fully and systematically exploited in breeding programmes. Therefore there is need to know its genetic diversity and genetic relationships to help maintain it, sustain long-term selection gain and exploit it in breeding and also enhance populations with continuous genetic variation for association mapping and for improvement of germplasm management.

1.12 Hypothesis

i) Genetic diversity among and within pearl millet accessions from West Africa is high and it is partitioned according to geographic origin.

ii) There are significant differences in genetic diversity within and between the pearl millet accessions originating from West Africa

1.13 Objectives

1.13.1 Main objective

To study the genetic diversity and relationships among and within pearl millet landraces from West Africa

1.13.2 Specific Objectives

1. To assess the level and distribution of genetic diversity in pearl millet landraces from West Africa

2. To establish genetic relationships between and within pearl millet landraces form West Africa

CHAPTER TWO

2.0 MATERIAL AND METHODS

2.1 Plant materials

Thirty pearl millet accessions including landraces and improved open-pollinated cultivars were used in this study. Seeds were collected from diverse geographic origins in West Africa. Table 2 shows their origins and some features of the accessions.

Table 2. Characteristics for pearl millet accessions used in this genetic diversity studies

Sample No	Accession ID	Field entry No.2006	Country of origin	Location	Parent
G195	PE02990	327	Senegal	Sagatta	F1
G196	PE03012	328	Senegal	Orefonde	F2
G197	PE05532	329	Mali	Nangal	F3
G198	PE05631	330	Mali	Sanga	F4
G199	PE05996	331	Burkina Faso	Sanga	F5
G200	PE01203	332	Burkina Faso	Banga	F6
G201	PE02801	334	Niger	Deoule	F7
G202	PE02935	335	Niger	Barakan	F8
G203	AON378	336	Nigeria		F9
G204	AON122	337	Nigeria		F10
G205	Souna3	338	Senegal		M1
G206	PE01437	339	Senegal	Vindediabi	M2
G207	PE05393	340	Mali	Aourou	M3
G208	PE05558	341	Mali	Tominian	M4
G209	PE05988	342	Burkina Faso	Tougou	M5
G210	PE05980	343	Burkina Faso		M6
G211	PE02827	344	Niger	Bangou	M7
G212	PE02885	345	Niger	Falouel	M8
G213	AON644	346	Nigeria		M9
G214	AON514	212	Nigeria		M10
G135	PE02987	82	Senegal	Diourbel	P1
G49	PE05344	152	Mali	Korera Kole	P2
G96	PE03942	360	Benin	Birni	P3
G228	Gampela	350	Burkina Faso Inera		P4
G218	ICMV IS 92222	305	Niger ICRISATSC		P5
G191	IP8679	302	Sudan	Darfur	P5
G188	IP12182	347	Nigeria	Gongola	P7
G215	SOSAT-C88	249	Mali ICRISATSC		Improved Variety
G217	ICMV IS89305	202	Niger ICRISATSC		Improved Variety
G129	PE02703		Niger	May Kerawa	Panicle with extra large diameter

2.2 Preparation and sampling of materials

Fifteen to twenty seeds of each accession were planted in a 30mm diameter rehydrated Jiff-belt® pellet soil placed in 6cm diameter wells in plastic trays in a 5 x 7 format in the laboratory. They were placed by the window for access to light during germination. Fresh leaves were collected from one week-old plants, as described by Mace *et al.* (2003). From each accession, there were three bulks of four plants each. Out of the thirty accessions, ten were randomly selected and planted. Five plants from each of the ten accessions were individually sampled for intra-population diversity studies.

2.3 DNA isolation

2.3.1 DNA extraction

Genomic DNA was extracted from fresh leaves of one week old seedlings of each accession using a modified CTAB protocol (Mace *et al.*, 2003). Fully expanded leaves of pearl millet were harvested by cutting them using scissors and transferring them to a pre-cooled mortar and 450µl of pre-heated CTAB buffer was added and leaf particles ground. The macerated leaves was transferred to fresh microfuge tubes and incubated for 10mins at 65°C with occasional mixing. Solvent extraction of the macerated leaves was done by adding 450µl chloroform: isoamylalcohol (24:1) to each tube and inverted twice to mix. The tubes were centrifuged at 12,000 rpm for 10 minutes to separate the phases. A fixed volume of supernatant (400µl) was carefully decanted and transferred to fresh microfuge

UNIVERSITY LIBRARY

tubes. About 0.7 volume of isopropanol (stored at ¯20°C) was added and inverted once to mix and the tubes centrifuged at 12000 rpm for 15 minutes. The supernatant was decanted and pellet air-dried for 30 minutes. Two hundred microliters low salt TE buffer (10mM Tris-HCL,1mM EDTA, 2M Nacl) was added to each sample and then 3μl RNase (10mg/ml) was added to each of the samples and incubated overnight at room temperature (in the dark). A second solvent extraction was done by adding 200μl phenol: chloroform: isoamylalcohol (25:24:1) to each sample and inverted twice to mix then centrifuged at 12,000 rpm for 10 minutes. A fixed volume of 180μl was transferred to fresh microfuge tubes and chloroform: isoamylalcohol (24:1) added to each tube then inverted twice to mix and centrifuged after which a fixed volume of aqueous layer (approximately 180μl) was transferred to fresh microfuge tubes. To purify the DNA, 31.5 ml ethanol: sodium acetate solution (30:1.5) was added to each sample and placed in ¯20°C for 5 minutes and then centrifuged at 12,000 rpm for 5 minutes. The supernatant from each sample was decanted and the pellets washed with 200μl 70% ethanol. The tubes were centrifuged at 12,000 rpm for 5 minutes and the supernatant from each sample decanted and the pellet air-dried for approximately 1hour. The pellet was re-suspended in 100μl low salt TE and stored at 4°C till use. DNA from the three bulks was then pooled into one sample in each accession to be used as samples for between accessions studies and DNA from individual plants was treated as samples for within accessions studies.

2.3.2 Quantification of DNA

DNA was quantified through spectrophotometric absorbance (A) readings at wavelengths (λ) of 260 and 280 nm. Fifteen microlitres of each sample was added to 735 µl of sterile double distilled water and vortexed to give a 1:50 dilution. A nanodrop spectrophotometer was used to read the optical density at 260 mm and 280 mm (OD 260/280) in order to determine the concentration of DNA in µg/ml and also to determine its purity. DNA has been shown to absorb UV light at 260 mm and one optical density (OD) at 260 nm is equivalent to 50 µg/ml for dsDNA and to 40 µg/ml of ssDNA (Sambrook *et al.*, 1989). The ratio A_{260}/A_{280} was used to determine the purity of the DNA samples. It has been shown that if the ratio is between 1.8 and 2.0 the absorption is due to nucleic acids. A ratio less than 1.8 indicates that there may be proteins or other UV absorbers in the sample. A ratio higher than 2.0 indicates that samples may be contaminated with chloroform or phenol (Rojas, 1997). After quantification, samples that had a ratio of less than 1.8 and higher than 2.0 were re-precipitated with ethanol. The precipitate of DNA, which is allowed to form at low temperature ($^{-}20^{0}$C or less) in the presence of moderate concentrations of monovalent cations is recovered by centrifugation and re-dissolved in a low salt TBE buffer (1g NaOH, 108g Tris Base, 55g Boric Acid 7.4g) at the desired concentration. The technique was rapid and was quantitative even with nanogram amounts of DNA.

2.3.3 DNA quality checking

Agarose powder was dissolved in Tris Borate EDTA (TBE) buffer (1% w/v) by slowly boiling in a microwave oven. The agarose was allowed to cool to about 50^{0}C and ethidium bromide was added to the gel at a concentration of 1 mg/ml. Care was taken because ethidium bromide is mutagenic. While the agarose was cooling, the gel tray was prepared by sealing the open edges of a clean, dry glass tray with autoclave tape so as to form a mold to avoid leakage and so that the tray could accommodate the desired thickness of the gel. The edges of the gel tray were sealed with a small quantity of the agarose solution using a Pasteur pipette. When the seal was set, the rest of the warm agarose solution was then poured into the gel tray in which a comb was inserted to form sample slots. The gel was allowed to solidify for 30 minutes before removing the autoclave tape, and then the tray/mould was immersed in the electrophoresis tank containing TBE buffer. The combs were removed and 5 μl of each DNA sample containing 2 μl 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 i.e. in one lane of each gel. DNA was mixed with sample loading buffer in order to allow the solution to sink in the gel wells. The gel was run at a constant voltage of 100 volts until the bromo-phenol blue migrated almost to the end of the gel. Resolution was improved by re-circulating the buffer every 20 min. The gel was then removed from the rig, placed in a UV trans-illuminator and photographed.

2.4 SSR primers

A total of 21 microsatellite primer pairs were used for PCR amplification of the DNA samples (Table 3). SSR markers were chosen based on three criteria: genome position (Mariac et al 2006), repeat size (ranging from di-nucleotide to tri-nucleotide repeats) and the number of previously reported alleles. The forward primers for each of the 21 markers were labeled at the 5' end of the oligonucleotide using fluorescent dyes for screening by capillary electrophoresis on the ABI prism and 3730 genetic analyzer (Applied Biosystems).

Table 3. Summary of the 21 pearl millet microsatellite primer pairs used in the study. The table shows the primer forward and reverse sequences, their motifs , expexted product size , likngage group and the genotypes used in their optimization

Marker Name	Primer_Sequences_Forward	Primer_Sequences_Reverse	Motif	Expected_ Product_ Size	Linkage_Group	Optimised_On_Genotypes
PSMP2001	CATGAAGCCAATTAGGTCTC	ACCATCTGACTTGTTCTTATCC	(CT)8(CA)48	205	5	841B
PSMP2008	GATCATGTTGTCATGAATCACC	ACACTACACCTACATACGCTCC	(GT)37	184	7	863B
PSMP2027	AGCAATCCGATAACAAGGAC	AGCTTTGGAAAAGGTGATCC	(GT)31	239	7	863B
PSMP2030	ACCAGAGCTTGGAAATCAGCAC	CATAATGCTTCAAATCTGCCACAC	(CA)11(GA)10	112	1	841B
PSMP2043	TCATATTCTCCTGTCTAAAACGTC	ACAAATCGTACAAGTTCCACTC	(CA)13(GA)6	189	7	841B
PSMP2076	GGAATAGTATATTGGCAAAATGTG	ATACTACACACTGTAAGCATTGTC	(AC)15	147	4	863B
PSMP2079	AGCCGAAGGCTAATCAACAA	GTGGTCAGCAGCAGATGTAA	(AC)27	194	7	863B
PSMP2080	CAGAATCCCCACATCTGCAT	TGCAACTGAGCGAAGATCAA	(AC)14	175	1	863B
PSMP2085	GCACATCATCTCTATAGTATGCAG	GCATCCGTCATCAGGAAATAA	(AC)11	171	4	863B
PSMP2087	GGAACAGACTCCATACCTGAAA	TACCTGCCTGTGCTGTTAGT	(AC)10	118	7	841B
PSMP2090	AGCAGCCCAGTAATACCTCAGCTC	AGCCCTAGCGCACAACACAAACTC	(CT)12	173	1	863B
PSMP2201	CCCGACGTTATGCGTTAAGTT	TCCATCCATCCATTAATCCACA	(GT)6	365	2	863B
PSMP2208	GAAAGAGCAAACTGAACAATCCC	ACTTTGCCCTGGATGATCCTC	(GT)10	246	5	841B
PSMP2231	TTGCCTGAAGACGTGCAATCGTCC	CTTAATGCGTCTAGAGAGTTAAGTTG	(TG)12GG(TA)4	226	2	841B
PSMP2233	TGTTTTCTCCTCTTAGGCTTCGTTC	TGTTTTCTCCTCTTAGGCTTCGTTC	(TG)9	256	5	863B
PSMP2237	TGGCCTTGGCCTTTCCACGCTT	CAATCAGTCCGTAGTCCACACCCCA	(GT)8	229	2	841B
PSMP2246	CGGATGCTAAATTAACCGAAGC	CCAGCTTGCTTCTGTTGCGTTC	(TG)7imperfect	262	1	841B
PSMP2248	TCTGTTTGTTTGGGTCAGGTCCTTC	CGAATACGTATGGAGAACTGCGCATC	(TG)10	166	6	841B
PSMP2249	CAGTCTCTAACAAACAAACACGGC	GACAGCAACCAACTCCAAACTCCA	(GT)7imperfect	130	3?	841B
PSMP2267	GGAAGGCGTAGGGATCAATCTCAC	ATCCACCCGACGAAGGAAACGA	(GA)16	206	3	863B
ICMP3002	CGAGCCGCCATAGTTGAC	TACACACACATTGCCACACG	(AAG)7	229	6	

2.5 DNA amplification

PCR conditions for each of the 21 SSR markers were optimized and PCR reactions were set up in 10 μl volumes in 96-well PCR plates (ABGene, Rochester N.Y). Each PCR reaction contained 2pmol/μl of primer, 2mM $MgCl_2$, 2mMdNTPs, 0.04U Amplitaq Gold Polymerase (PE-Applied Biosystems) and 1X PCR buffer (KCL, Tris-HCL PH 8.4) (PE-Applied Biosystems). Temperature cycling was carried out using the Gene-Amp PCR system 9700 (PE-Applied Biosystems and touch-down PCR amplification; one 15-min denaturation cycle, followed first by ten cycles of 94°C for 10s, 61°C for 20s (ramp of 1°C per cycle) and 72°C for 30s, then by 31 cycles of 94°C for 10s, 54°C for 20s and 72°C for 30s. After completion of the 31 cycles, a final extension of 30 min at 72°C was included to try and minimize the +A overhang (Smith *et al.*, 1995) (Table 4). The final extension temperature was increased to 30 minutes to reduce the probability of false scoring of stutter bands as alleles. Following amplification, PCR products were stored at 4°C prior to electrophoresis.

Table 4: PCR reaction mixture of x1 for each SSR primer . The table also shows the dye color that is labelled on each primer and how they were combined in co-loading sets.

Primer	Dye color	Allele size	Multiplex set	Repeat motif	DNA (5ng/ µl)	PCR Buffer 910x)	dNTPs(2mM)	Mg++ (25mM)	Taq (5u/µl)	primer (F+R) 2pm/µl	Water	Expected product size (bp)*	Linkage Group	PCR for coloading (ul)
ICMP3002	FAM	130-140	Set 1	(AAG)7	1	1	0.5	0.8	0.04	1	5.66	229	6	0.5
PSMP2043	NED	150-200	Set 1	(CA)13(GA)6	1	1	0.5	0.8	0.04	1	5.66	189	7	0.75
PSMP2001	PET	190-210;275-305	Set 2	(CT)8(CA)48	1	1	0.5	0.8	0.04	1	5.66	205	5	0.5
PSMP2030	NED	100-150	Set 2	(CA)11(GA)10	1	1	0.5	0.8	0.04	1	5.66	112	1	0.5
PSMP2248	FAM	155-170	Set 2	(TG)10	1	1	0.5	0.8	0.04	1	5.66	166	6	0.5
PSMP2201	VIC	350-370	Set 3	(GT)6	1	1	0.5	0.8	0.04	1	5.66	365	2	0.5
PSMP2246	FAM	240-265	Set 3	(TG)7 imperfect	1	1	0.5	0.8	0.04	1	5.66	262	1	0.5
PSMP2008	NED	180-256	Set 4	(GT)37	1	1	0.5	0.8	0.04	1	5.66	184	7	0.75
PSMP2076	FAM	130-165	Set 4	(AC)15	1	1	0.5	0.8	0.04	1	5.66	147	4	0.5
PSMP2087	PET	115-130	Set 4	(AC)10	1	1	0.5	0.8	0.04	1	5.66	118	7	0.5
PSMP2080	PET	130-220	Set 5	(AC)14	1	1	0.5	0.8	0.04	1	5.66	175	1	0.5
PSMP2231	FAM	220-245	Set 5	(TG)12GG(TA)4	1	1	0.5	0.8	0.04	1	5.66	226	2	0.75
PSMP2233	NED	250-265	Set 5	(TG)9	1	1	0.5	0.8	0.04	1	5.66	256	5	0.75
PSMP2085	PET	160-170	Set 6	(AC)11	1	1	0.5	0.8	0.04	1	5.66	171	4	0.5
PSMP2237	FAM	230-260	Set 6	(GT)8	1	1	0.5	0.8	0.04	1	5.66	229	2	0.5
PSMP2249	VIC	135-160	Set 6	(GT)7 imperfect	1	1	0.5	0.8	0.04	1	5.66	130	3?	0.5
PSMP2090	NED	160-190	Set 7	(CT)12	1	1	0.5	0.8	0.04	1	5.66	173	1	0.5
PSMP2208	PET	240-250	Set 7	(GT)10	1	1	0.5	0.8	0.04	1	5.66	246	5	0.75
PSMP2267	VIC	195-240	Set 7	(GA)16	1	1	0.5	0.8	0.04	1	5.66	206	3	0.5
PSMP2027	FAM	230-265	Set 8	(GT)31	1	1	0.5	0.8	0.04	1	5.66	239	7	0.5
PSMP2275	VIC	155-170;265-290	Set 8	(GTT)10	1	1	0.5	0.8	0.04	1	5.66	279	6	0.5

*The expected product size is for the genotypes used for optimization (841B or 863B).

2.6 Gel electrophoresis of PCR products

Agarose gel electrophoresis was carried out as outlined in section 3.3.3. Three microlitres of the PCR product was run on 2% agarose gel to check the amplification of each SSR. To 5 µl of each PCR reaction mix, 3 µl of sample loading buffer was added and mixed by pipetting before loading the resulting mixture in the pre-formed sample wells on the gel. The samples were run alongside 1.0 µl 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.

2.7 Capillary electrophoresis

Forward primers were labeled with FAM, PET, VIC and NED (PE-Applied Biosystems), allowing post-PCR pooling of the 21 primer products into 6 groups of three primer products each and 2 groups of two primer products each, with each primer product in a given group being labeled with a different dye. Co-loading sets were optimized and multiplexed based on dye label, fragment size and fluorescence to reduce the unit cost of high throughput genotyping.

PCR products were co-loaded post-PCR, where a range of 0.5 and 1µl of the FAM, PET and VIC labelled products were mixed with their corresponding µl of formamide made by mixing 1ml of HIDI and 12µl of Genescan™-500LIZ™ size standard (Applied Biosystems) to make up to 10µl volume. DNA fragments were denatured and size-fractioned using capillary electrophoresis on an ABI prism 3730 automatic DNA sequence (PE-Applied Biosystems). The GeneMapper v3.7

software (PE-Applied Biosystems) was used to size peak patterns, using the internal Genescan-500 LIZ size standard and Genotyper 3730 (PE-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. This facilitated the accurate sizing of the microsatellite allele to within ± 0.3 base pairs (Buhariwalla and Crouch, 2004). Size calling, which includes peak detection and fragment size matching were performed using GeneMapper. Bins, which represent a fragment size or base pair range and dye colour that define an allele, were constructed from reference data. Algorithms were used to determine if peaks represented alleles. When a peak from a 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 data base. Allelobin software was used for checking the quality of markers.

2.8 Data analysis

All the SSR markers showed high reproducibility and therefore the 21 markers were included in the analysis. The total number of alleles detected 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. Pair-wise genetic similarities between individual accessions were calculated using PowerMarker v.3.25 (Liu and Muse, 2004). Multi-Dimensional Scatter plots were drawn to visualize the inter-relationship among the accessions. The Nei's genetic

diversity within and among accessions and countries was estimated using the program PowerMarker (Version 3.25; Liu and Muse, 2004) based on estimated allele frequencies, with the following statistics: average number of common alleles per locus (allelic frequencies of at least 5%), number of polymorphic loci, percentages of polymorphic loci, observed heterozygosity (H_{obs}) and average expected gene diversity corrected for small simple sizes (H_{unb}, Nei 1978).

Allele frequency based pair-wise genetic distances between countries and between accessions were calculated using PowerMarker v3.25 (Liu and Muse 2004) based on Rogers distance matrix (Rogers, 1972). The resulting distance matrix was subjected to sequential agglomerative hierarchical nested (SAHN) clustering using un-weighted pair-group method analysis (UPGMA) (Sneath and Sokal, 1973) parameters in PowerMarker v3.25. Distance matrices and dissimilarity indices were also generated using Dice's coefficient features in Darwin ver.4.0 to generate dendrograms showing the clustering of the 49 individuals and 30 bulks . Bootstrapping over loci with 1000 replications was carried out to assess the strength of evidence for the branching patterns in the resulting UPGMA dendrograms. Consensus trees were constructed using the Tree View and Darwin ver. 4.0. ALS_binary program was used to convert the allelic data for both bulks and individuals to binary data. Allele richness was calculated for each accession and individuals from the binary data. Pair-wise genetic similarities between individual accessions were calculated using Darwin ver 4.0 with binary data (single data) as input, based on Dice's genetic similarity coefficient (Dice, 1945). The resulting similarity matrix was subjected to non-metric multi-dimensional scaling (MDS) (Kruskal, 1964a,

b). A scatter plot of the first and the second MDS axes was drawn to visualize the inter-relationship among the accessions for bulks and within the individuals for individuals. In structure analysis, population differentiation test were performed in PowerMarker ver.4.0. The first analysis used a contingency table approach to determine if groups of individuals had a significant difference in allele frequencies for each locus, while the second analysis tested the overall differentiation of groups by using a variant of the Mantel test (MANTEL-STRUCT). Arlequin program was used to calculate the molecular variance (AMOVA) and determine the partitioning of diversity according to Weir. (1996).

CHAPTER THREE

3.0 RESULTS

3.1 Quantification of DNA and quality checking

DNA was quantified using a Nanodrop and confirmed by agarose gel against Lambda DNA standard. DNA quantities ranged from 3.74 to 8.11 ng/ul while purity ranged from 1.79 to 2.1 O.D ratio. It was then diluted to give a concentration of about 5ng/µl (Figure 4).

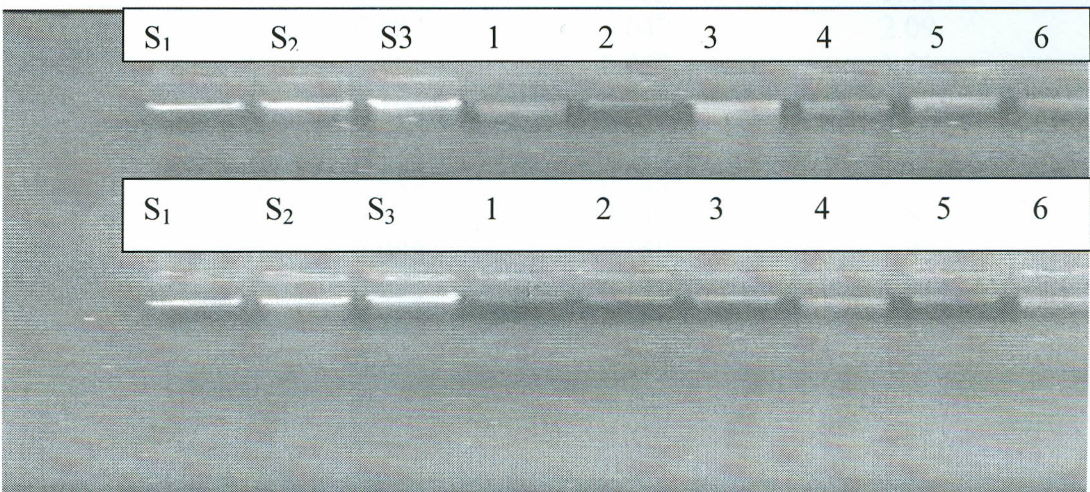


Figure 4: Agarose electrophoresis of DNA extraction from fresh leaves of 6 individuals (upper lane 1,2,3,4,5 and 6) and 6 bulked samples (lower lane 1, 2, 3, 4, 5 and 6) after dilution to 5ng. S_1, S2 and S_3 stands for standard 1 standard 2 and standard 3 respectively .

Table 5. DNA quantification and purity values using Nanodrop.This table shows concentration values of DNA (ng/ul) and the purity (260/280)

Sample ID	**ng/ul**	**A_{260}**	**A_{280}**	**260/280**
1	8.11	0.162	0.086	1.89
2	4.61	0.092	0.174	1.89
3	5.19	0.104	0.208	2
4	4.11	0.082	0.044	1.86
5	7.54	0.151	0.077	1.95
6	8.16	0.163	0.083	1.98
7	4.81	0.096	0.191	1.99
8	4.75	0.095	0.188	1.98
9	4.68	0.094	0.197	2.1
10	4.48	0.09	0.164	1.82
11	4.04	0.081	0.152	1.86
12	4.44	0.089	0.042	2.09
13	4.8	0.096	0.053	1.81
14	5.29	0.106	0.204	1.92
15	5.47	0.109	0.058	1.88
16	5.73	0.115	0.221	1.93
17	5.71	0.114	0.215	1.89
18	4.41	0.088	0.159	1.81
19	4.72	0.094	0.179	1.9
20	5.04	0.101	0.189	1.87
20	5.14	0.103	0.191	1.85
22	5.62	0.112	0.200	1.79
23	5.74	0.115	0.216	1.88
24	4.78	0.096	0.051	1.88
25	6.66	0.133	0.247	1.86
26	6.21	0.124	0.236	1.9
27	6.11	0.122	0.227	1.86
28	3.74	0.075	0.146	1.95
29	5.55	0.111	0.210	1.89

3.2 Optimization of SSR primers

Each of the 21 SSR markers was successfully optimized and PCR reactions were set up in 10 μl volumes in 96-well PCR plates (ABGene, Rochester N.Y) (Figure 5)

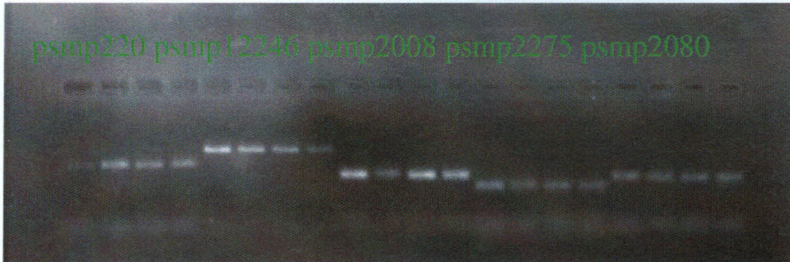


Figure 5: Agarose gel showing the optimization of 5 SSR primers (PSMP 2201, PSMP 2246, PSMP 2008, PSMP 2275 and PSMP 2080). This figure shows how the primers were polymorphic in their amplification. They amplified different sizes of the PCR product

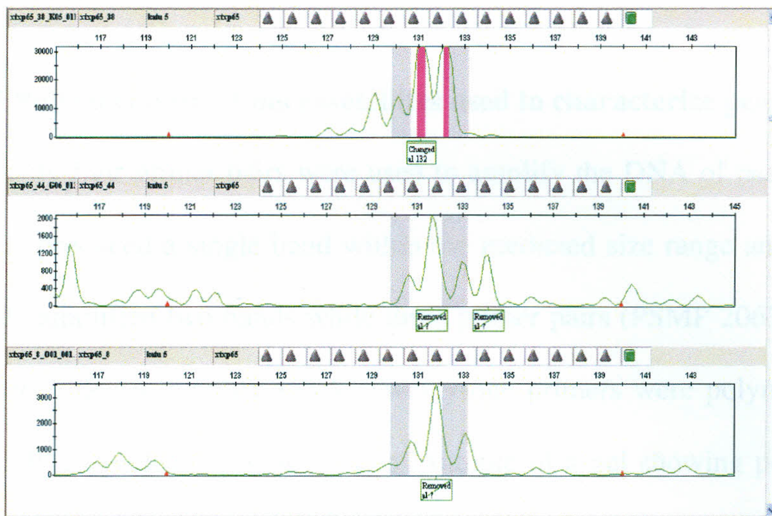


Figure 6. PSMP 2080 before optimization. The pink colour shows that the PCR product was in excess when loaded for genotyping

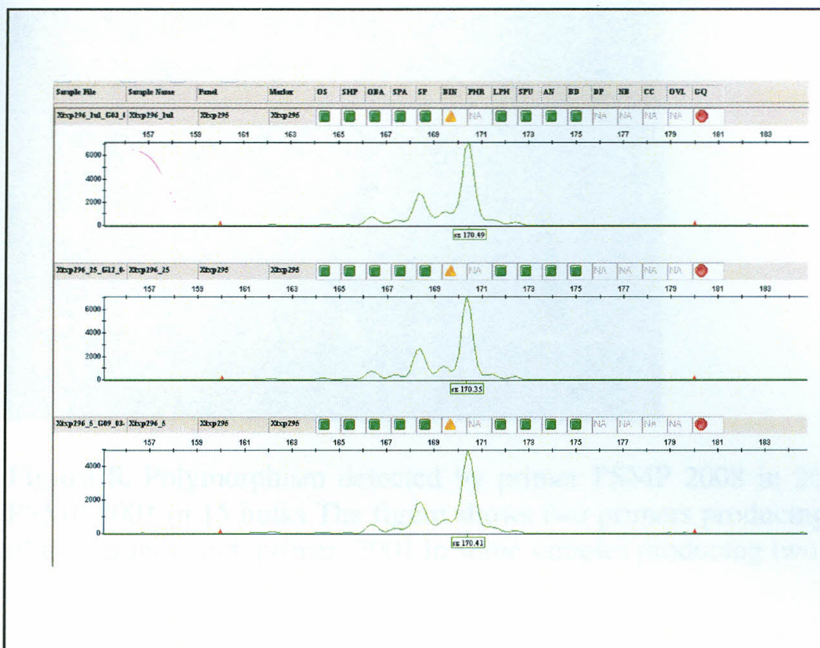


Figure 7. PSMP 2080 after optimization. After optimization, the pink colour disappears and the peaks are sharp indicating the alleles hae been scored well

3.3 Polymorphism of microsatellites used to characterize pearl millet accessions

Twenty-four primer pairs were used to amplify the DNA of pearl millet. Twenty primer pairs produced a single band within the predicted size range and one primer pair PSMP 2001 amplified two bands while three primer pairs (PSMP 2063, PSMP2071 and PSMP 2079) failed to amplify DNA. Twenty-one primers were polymorphic for the 30 bulks and 49 individuals. Figure 8 is an example of a gel showing polymorphism by primers PSMP 2001 and PSMP 2008 in some individuals and bulks of pearl millet. A singe peak denoted homozygous genotypes while two clear peaks indicated heterozygous genotypes (Figures 9and 10).

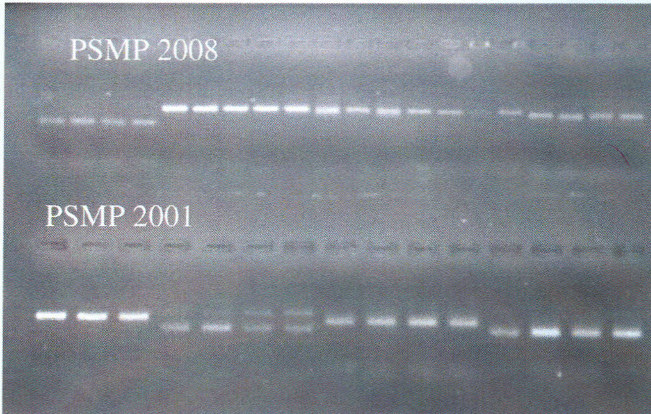


Figure 8. Polymorphism detected by primer PSMP 2008 in 20 individuals and primer PSMP 2001 in 15 bulks.The figure shows two primers producing different PCR products of the samples and primer 2001 in some samples producing two PCR products.

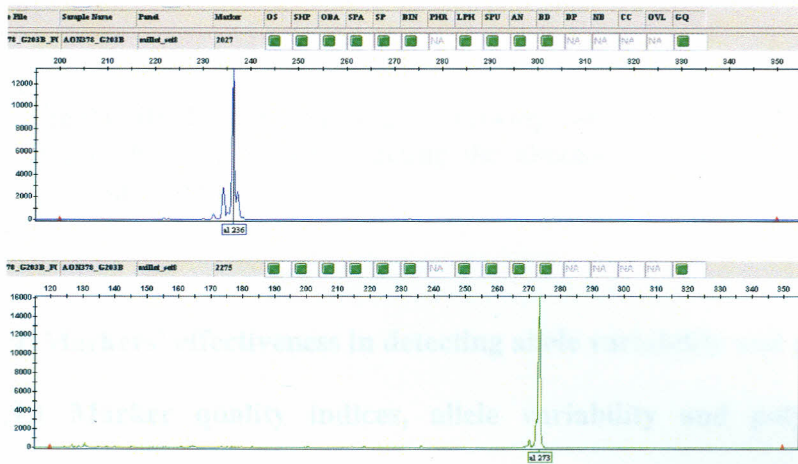


Figure 9: Electrophenograms showing homozygosity. The figure shows that the two different primers detected one allele in each sample

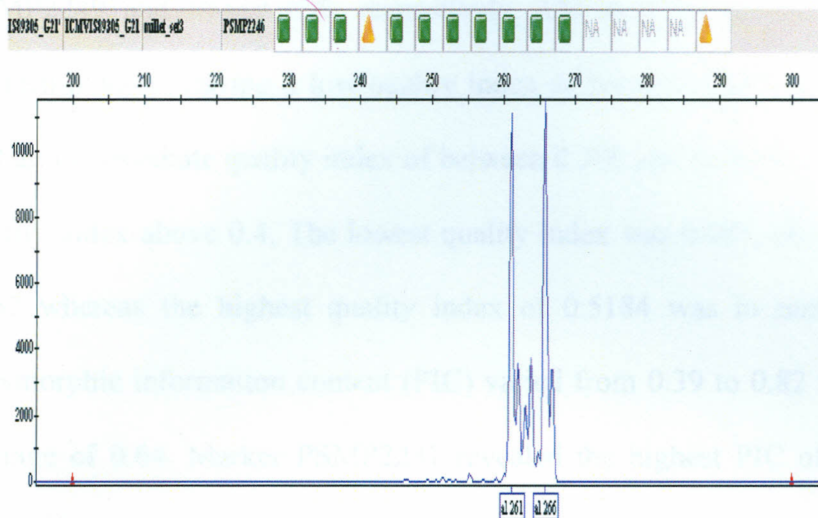


Figure 10: Electrophenograms showing heterozygosity. The figure shows a primer detecting the alternative of the same allele in a sample

3.4 Markers' effectiveness in detecting allele variability and polymorphism

3.4.1 Marker quality indices, allele variability and polymorphism in the 49 individuals

In total, 171 alleles were detected among the 49 individuals using 21 primer pairs. The total number of alleles per locus ranged from 4 to 21 with an average of 8 alleles per locus (Table 6). Primer pair PSMP 2027 amplified the highest number of alleles (21) while primer pairs PSMP 2246 and PSMP 2249 amplified 4 alleles each, which was the lowest number of alleles. The maximum size range of 80 bp was observed with primer PSMP 2027 whereas the lowest of 6bp was observed with primer pair PSMP 2246 (Table

whereas the lowest % of abundant alleles was observed in primer PSMP 2027 (23.6%). The highest and lowest numbers of rare alleles at <=5% were 5 with primers PSMP2030, PSMP2080, PSMP2027 and PSMP2231 and 1 with primers PSMP2076, PSMP2085, PSMP2249 and PSMP2090, respectively. The average quality index was 0.29, with thirteen markers having a low quality index of between 0.087 and 0.295, three markers had an intermediate quality index of between 0.300 and 0.3954, while five markers had a quality index above 0.4. The lowest quality index was 0.087, observed in marker PSMP 2267 whereas the highest quality index of 0.5184 was in marker ICMP 3002. The polymorphic information content (PIC) varied from 0.39 to 0.82 for individuals with an average of 0.64. Marker PSMP2231 revealed the highest PIC of 0.8174 while Marker PSMP2201 had the lowest PIC of 0.3864.

Table 6: Quality indices and characterization of polymorphism detected by 21 SSR markers in 49 pearl millet individuals

Marker name	Quality Index	Total no of alleles	Allele size range	Abundant Allele (%)	Rare Allele(s) (<=5%)	PIC values
ICMP3002	0.5184	5	180-234	228 (49.30)	180,225,234	0.432
PSMP2043	0.3662	9	179-199	179 (32.00)	181,189,193,199	0.7453
PSMP2001	0.2443	9	197-213	205 (38.71)	201,203,211,213	0.6783
PSMP2030	0.1672	15	108-146	110 (29.87)	108,114,116,122,125	0.7403
PSMP2248	0.2235	6	158-174	168 (47.76)	158,174	0.593
PSMP2201	0.2393	7	332-368	366 (69.49)	332,354,364,368	0.3864
PSMP2246	0.4621	4	259-265	263 (55.22)	261,265	0.4666
PSMP2008	0.2491	10	181-257	189 (28.07)	181,183,199,257	0.805
PSMP2076	0.1512	6	148-166	160 (44.62)	150	0.7589
PSMP2087	0.1904	8	117-133	121 (40.54)	127,129,131,133	0.8064
PSMP2080	0.2713	11	113-187	177 (30.43)	133,159,181,183,185	0.5683
PSMP2233	0.4462	9	255-271	259 (32.20)	269,271	0.4547
PSMP2085	0.2097	5	167-177	169 (38.81)	177	0.6095
PSMP2237	0.2957	6	231-259	257 (40.30)	235,253,259	0.5987
PSMP2249	0.3075	4	130-156	152 (44.87)	156	0.5642
PSMP2090	0.2225	8	174-188	178 (29.87)	176	0.7582
PSMP2267	0.087	5	206-216	206 (57.83)	214,216	0.7762
PSMP2208	0.4476	6	245-255	247 (53.73)	245,255	0.6522
PSMP2027	0.342	21	226-306	238 (23.61)	226,228,230,232,242	0.6473
PSMP2231	0.486	12	222-248	244 (32.43)	222,228,232,234,236	0.8174
PSMP2275	0.1105	5	268-283	274 (51.67)	268,271,283	0.4804
Mean	0.2875	8.142				0.6352

Table 7; Number and size of alleles detected in 49 peark millet individuals with 21 SSR markers

Marker name	Alleles																				
ICMP3002	180	219	225	228	234																
PSMP2043	179	181	183	185	187	189	191	193	199												
PSMP2001	197	199	201	203	205	207	209	211	213												
PSMP2030	108	110	112	114	116	118	120	122	124	126	128	130	136	142	146						
PSMP2248	158	164	166	168	170	174															
PSMP2201	332	346	354	362	364	366	368														
PSMP2246	259	261	263	265																	
PSMP2008	181	183	185	187	189	191	193	195	199	257											
PSMP2076	148	150	160	162	164	166															
PSMP2087	117	119	121	123	127	129	131	133													
PSMP2080	113	143	159	173	175	177	179	181	183	185	187										
PSMP2233	255	257	259	261	263	265	267	269	271												
PSMP2085	167	169	171	175	177																
PSMP2237	231	235	253	255	257	259															
PSMP2249	130	152	154	156																	
PSMP2090	174	176	178	180	182	184	186	188													
PSMP2267	206	210	212	214	216																
PSMP2208	245	247	249	251	253	255															
PSMP2027	226	228	230	232	234	236	238	240	242	244	252	256	258	260	276	278	280	290	302	304	306
PSMP2231	222	224	226	228	232	234	236	238	242	244	246	248									
PSMP2275	268	271	274	280	283																

3.4.2. Marker quality indices, allele variability and polymorphism in the 30 bulked samples

In total, 172 alleles were detected among the 30 bulked samples using 21 primer pairs. The total number of alleles per locus ranged from 2 to 25, with an average of 8 alleles per locus (Table 8). Primer pair PSMP 2027 amplified the highest number of alleles (25) while primer pair PSMP 2275 amplified 2 alleles which was the least alleles. The maximum size range of 80 bp was observed with primer PSMP 2027 whereas the lowest of 4 bp was observed with primer pair PSMP 2246 (Table 9). The highest % of abundant alleles was observed in primers PSMP 2267 (57.45%) whereas the lowest % of abundant alleles was observed in primer PSMP 2027 (15.97%). The highest and lowest numbers of rare alleles at <=5% were 5 and 1, respectively with different primers. There were no rare alleles observed from primer pairs PSMP 2201, PSMP 2246, PSMP 2249 and PSMP 2208.

The average quality index was 0.28 with fourteen markers having a low quality index of between 0.084 and 0.2995. Two markers had an intermediate quality index of between 0.30 and 0.3954 while five markers had a quality index above 0..4. The lowest quality index was 0.082 observed from marker PSMP 2267 whereas the highest quality index of 0.5173 was from marker ICMP 3002 (Table 9). The polymorphic information content (PIC) varied from 0.31 to 0.84, with an average of 0.59. Marker PSMP2231 revealed the highest PIC of 0.8174 while Marker PSMP2201 had the lowest PIC of 0.3864.

Table 8. Quality indices and characterization of polymorphism detected by 21 SSR markers in 30 pearl millet DNA bulks

Marker Name	Quality index	Total no of alleles	Allele size range	Abundant Allele(%)	RareAlleles (s) (<=5%)	PIC Values
ICMP3002	0.5173	3	221-233	227 (50.91)	233	0.3739
PSMP2043	0.2996	10	175-199	187 (23.53)	175,183,193,199	0.7721
PSMP2001	0.1614	5	199-209	205 (67.44)	203	0.3116
PSMP2030	0.1729	19	108-146	112 (28.42)	108,114,116,118,124	0.665
PSMP2248	0.2234	8	158-174	168 (39.44)	158,162,172,174	0.5592
PSMP2201	0.1794	4	332-366	366 (45.00)	None	0.6229
PSMP2246	0.4889	3	259-263	263 (49.09)	None	0.46
PSMP2008	0.2135	10	181-201	189 (22.22)	181,193,199,201	0.7843
PSMP2076	0.1629	9	146-168	160 (38.36)	146,150,158,168	0.8438
PSMP2087	0.1804	6	117-133	121 (32.91)	125,133	0.6232
PSMP2080	0.2626	11	113-189	179 (22.61)	113,185,187,189	0.5
PSMP2231	0.4377	13	220-244	242 (20.18)	220,230,238,240	0.378
PSMP2233	0.4947	8	255-269	257 (24.68)	255,269	0.5704
PSMP2085	0.2026	6	167-177	171 (35.80)	173	0.6134
PSMP2237	0.3335	8	231-259	231 (37.70)	235,237,239,249,259	0.5424
PSMP2249	0.2887	3	130-154	152 (42.03)	None	0.7732
PSMP2090	0.2302	10	174-192	178 (27.08)	190,192	0.6689
PSMP2267	0.0842	5	206-218	206 (57.45)	210,216,218	0.5514
PSMP2208	0.5018	4	247-253	247 (50.00)	None	0.6213
PSMP2027	0.3779	25	229-309	239 (15.97)	229,233,245,247,249	0.8013
PSMP2275	0.1105	2	274-280	274 (54.17)	None	0.3705
Mean	0.2820	8.19				0.5908

Table 9. Number and size of alleles detected in the 30 pearl millet DNA bulks with 21 SSR markers.

Marker name	Alleles																							
ICMP3002	221	227	233																					
PSMP2043	175	179	181	183	185	187	189	191	193	199														
PSMP2001	199	203	205	207	209																			
PSMP2030	108	110	112	114	116	118	120	122	124	126	128	130	132	134	136	140	142	144	146					
PSMP2248	158	162	164	166	168	170	172	174																
PSMP2201	332	346	362	366																				
PSMP2246	259	261	263																					
PSMP2008	181	183	185	187	189	191	193	195	199	201														
PSMP2076	146	148	150	158	160	162	164	166	168															
PSMP2087	117	119	121	123	125	133																		
PSMP2080	113	143	173	175	177	179	181	183	185	187	189													
PSMP2231	220	222	224	226	228	230	232	234	236	238	240	242	244											
PSMP2233	255	257	259	261	263	265	267	269																
PSMP2085	167	169	171	173	175	177																		
PSMP2237	231	235	237	239	249	255	257	259																
PSMP2249	130	152	154																					
PSMP2090	174	176	178	180	182	184	186	188	190	192														
PSMP2267	206	210	212	216	218																			
PSMP2208	247	249	251	253																				
PSMP2027	229	231	233	235	237	239	241	243	245	247	249	251	253	255	259	261	265	267	271	273	275	279	289	307
PSMP2275	274	280																						

3.5 Genetic diversity within ten pearl millet accessions

3.5.1 Number of alleles

To assess within-accession genetic diversity, five plants from each of ten randomly selected accessions had been individually sampled and genotyped. The total number of alleles within the ten accessions across the 21 loci ranged from 107 to 158 per accession. The highest number of alleles (158) was observed in accessions G197 and G215, whereas the lowest was 107, which was observed in accession G211 (Figure 11).

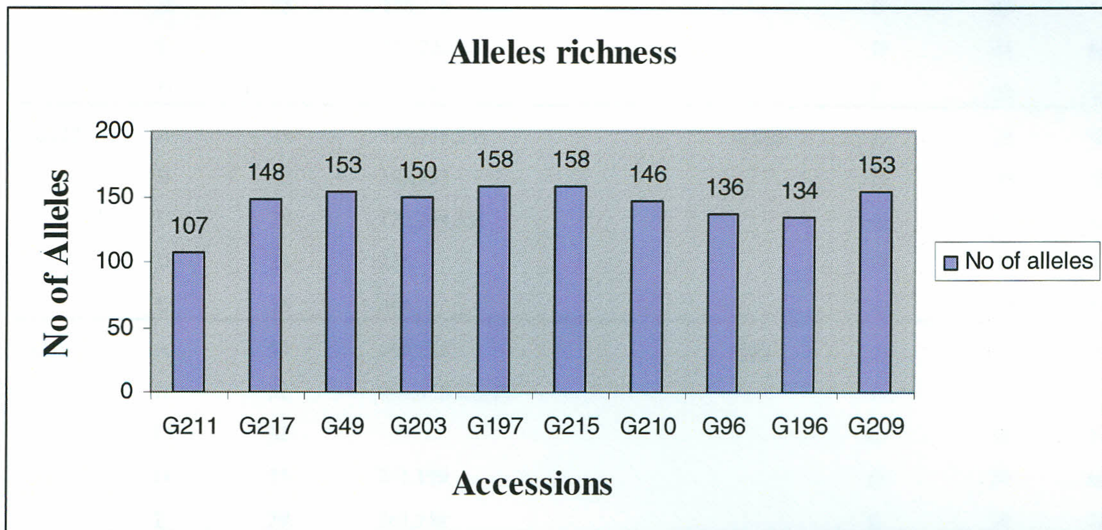


Figure 11. Total number of alleles in the ten accession of pearl millet. The figure shows how the number of alleles in each accession varied from one another

The difference between the highest and the lowest total number of alleles in each individual of the ten different accessions ranged from 4 to 9 with an average of 5.9. The highest total number of alleles was 36 from individuals G203A of accession G203 while the lowest was 22 from individual G96C of accession G96 (Table 10).

same accession was observed in accession G203 with 9 alleles. The difference between the highest and the lowest total number of alleles was least within individuals of accessions G211, G197 and G215 with 4 alleles.

Table 10. Total number of alleles in 49 individuals from ten pearl millet accessions genotyped with 21 SSR markers. The table shows the total number of alleles and their corresponding rare alleles in the ten accessions.

Accession	Sample Identity	Total alleles	Rare Allele(s) (<=5%)	Accession	Sample Identity	Total alleles	Rare Allele(s) (<=5%)
G211				**G215**	A	34	228,234
	A	29	265,222		B	31	158
	B	26	180,176		C	30	176
	C	25	242,222		D	31	None
	D	27	181,181		E	32	133,235
G217	A	28	177,245,255	**G210**	A	25	None
	B	32	193,332		B	30	189
	C	29	122,174,332		C	27	201,183
	D	33	255		D	32	116,159
	E	26	201		E	32	189,193
G49	A	27	108,232	**G96**	A	29	234,
	B	33	354,368,129,177		B	28	225,271
	C	31			C	22	235,271
	D	33	211,199		D	29	None
	E	29	261,259		E	28	185
G203	A	36	174,364,368,183,253,000,000	**G196**	A	29	114,261,183,232
	B	30			B	28	213,183,236
	C	30	122,257,230		C	27	183,236
	D	27	156		D	26	114,261,131,226,234,00
	E	27	225,228		E	24	364,232
G197	A	32	150,127,232,234	**G209**	A	28	176
	B	34	216		B	34	189
	C	30	269		C	30	203
	D	31	113,216,283		D	30	269,271
	E	31	199,181,214		E	31	None

3.5.2 Polymorphism, heterozygosity and gene diversity of the ten accessions of pearl millet

Polymorphism within the ten accessions ranged from 90.5 % to 100%. Four accessions namely G217 from Niger ICRISATSC, G49 from Mali, G203 from Nigeria and G209 from Burkina Faso showed 100% polymorphism across the 21 loci. The lowest level of polymorphism (90.5%) was observed in accessions G210 from Burkina Faso and G196 from Senegal. Polymorphism Information Content (PIC) in the ten accessions varied from 0.40 to 0.53, with an average of 0.47. The highest PIC was observed in accessions G209 and G49, with a value of 0.53 whereas the lowest was observed in accession G210 with a value of 0.40. The average Nei's gene diversity (*H*) varied from 0.45 to 0.59 with a mean of *H*= 0.53 across the ten accessions. The highest gene diversity was *H*= 0.59 and lowest *H*= 0.45 in accessions G209 and G210, respectively. Average heterozygosity ranged from 0.29 to 0.53 with a mean of 0.43 across the ten accessions. The highest average heterozygosity was observed in accession G197 with a value of 0.53 while the lowest was 0.29 in accession G211 (Table 11).

Table 11: Polymorphism and within accession diversity indices The table shows how polymorphism, average gene diversity, average heterozygosity and average PIC varied within the accessions.

Accession	Polymorphism (%)	Average gene Diversity	Average Heterozygosity	Average PIC
G211	95.2	0.51	0.29	0.44
G217	100	0.51	0.47	0.45
G49	100	0.58	0.50	0.53
G203	100	0.54	0.46	0.47
G197	95.2	0.57	0.53	0.52
G215	95.2	0.52	0.51	0.45
G210	90.5	0.45	0.40	0.4
G96	95.2	0.53	0.35	0.47
G196	90.5	0.53	0.32	0.47
G209	100	0.59	0.49	0.53

3.6 Genetic relationships among the 49 individuals

3.6.1 Genetic distance

The Rogers distance matrix revealed that genetic distance varied from 0.23 (between G196C and G196B) to 0.88 (between G196D and G96B). The highest genetic distance of 0.79 was observed within accession G196 between individual G196D and G196B while the lowest was 0.12 within accession G203, between individuals G203D and G203B (Table 12).

Table 12. Genetic distance matrices between individuals of ten different accessions

Accession / individual	196A	196B	196C	196D	196E
196B	0.5	****			
196C	0.4737	0.225	****		
196D	0.625	0.7857	0.725	****	
196E	0.4444	0.5789	0.5278	0.7632	****

Accession / individual	203A	203B	203C	203D	203E
203B	0.619	****			
203C	0.5714	0.5952	****		
203D	0.6429	0.119	0.5476	****	
203E	0.7143	0.4762	0.6429	0.4762	****

Accession / individual	210A	210B	210C	210D	210E
210B	0.425	****			
210C	0.5	0.5714	****		
210D	0.45	0.5476	0.4762	****	
210E	0.5	0.3571	0.4524	0.4048	****

Accession / individual	209A	209B	209C	209D	209E
209B	0.619	****			
209C	0.6667	0.5714	****		
209D	0.6053	0.5263	0.6053	****	
209E	0.5952	0.6667	0.6429	0.5789	****

Accession / individual	217A	217B	217C	217D	217E
217B	0.5238	****			
217C	0.4211	0.4474	****		
217D	0.3333	0.5476	0.4474	****	
217E	0.5789	0.5789	0.5556	0.5263	****

Accession / individual	215A	215B	215C	215D	215E
215B	0.5	****			
215C	0.5714	0.4762	****		
215D	0.5476	0.5476	0.4286	****	

3.6.2 Genetic similarities

Cluster analysis (CA) partitioned the accessions according to their geographic origin, breeding groups and pedigrees (Figure 12). Four clusters were generated, with two large clusters having 19 (Cluster C) and 18 (Cluster A) individuals. The smallest clusters had five individuals each (Clusters B and D). Cluster A had individuals from accessions

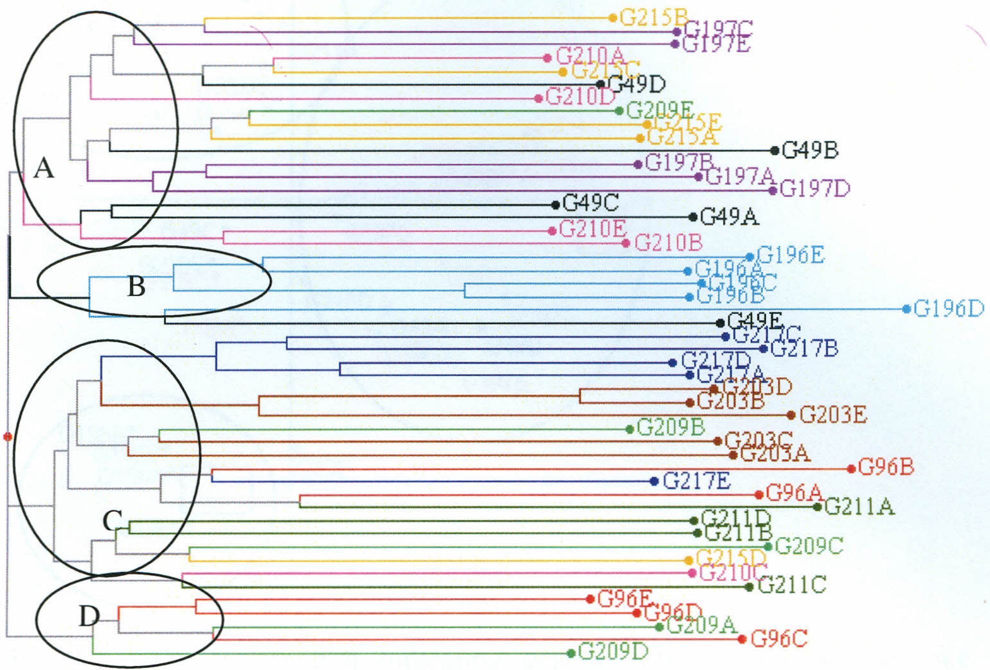


Figure 12: Dendrogram based on genetic relations of 49 individuals from ten different accessions using UPGMA algorithm method of Darwin Version 4.0. (Population numbers are as indicated in Table 12, each color shade represents individuals from the same accession).

A scatter plot of the first and second axes of non-metric multi-dimensional scaling (MDS) revealed three clusters of inter-relationships within accessions. The first axis explained 10.2 % of the variation. Most of individuals from Mali grouped together with a few individuals of accession G210 and G209 from Burkina Faso and G96 from Benin. Individuals of accessions G 211, G203 and G217 from Niger and Nigeria grouped together and a few individuals of accessions G209 and G96 (group B). All individuals of accession G196, except one, from Senegal grouped together (group C) (Figure 13).

Figure 13. An MDS scatter plot indicating genetic relationships between the 49 individuals. Accessions are labeled according to their countries of origin.

3.6.3 Population structure

There was a clear genetic differentiation both among countries, among populations within countries and within the populations using significance tests based on 1000 permutations calculated according to Weir (1996) in Arlequin. The examination of the hierarchical partitioning of genetic variation by AMOVA demonstrated that genetic differentiation was significant at $p<0.003$ using the exact value for population differentiation. The genetic variation was higher among countries (with a variance component of 0.436) than among populations within countries, with variance components of 0.713. The variation was highest within populations, with a variance component of 6.309. Of the total diversity, 5.85 % was attributed to country differences, 9.56% was attributed to population differences within the countries while 84.59% was attributed to differences

was significant at p<0.003 using the exact value for population differentiation. The genetic variation was higher among countries (with a variance component of 0.436) than among populations within countries, with variance components of 0.713. The variation was highest within populations, with a variance component of 6.309. Of the total diversity, 5.85 % was attributed to country differences, 9.56% was attributed to population differences within the countries while 84.59% was attributed to differences within populations. The Fst value was 0.154, indicating a significant level of differentiation. The value of F, which is equivalent to F_{IT} and estimates the overall (total population) inbreeding, was 0.38 indicating a minimal level of inbreeding while the value of Theta (P) was ˉ0.011 showing a slight excess of heterozygotes.

Table 13: AMOVA Calculated according to *Weir, B.S. 1996.*

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among countries	5	100.861	0.43602 Va	5.85
Among populations within countries	4	52.975	0.71325 Vb	9.56
Within populations	88	555.225	6.30938 Vc	84.59
Total	97	709.061	7.45864	

The Hardy-Weinberg equilibrium assumes a random mating system. Random mating can be assessed by computing the deviation of the observed heterozygosity from Nei's unbiased estimate of gene diversity (*H*). The degree of heterozygosity can also be

determined by calculating the percentage of heterozygosity in a population. The mean Chi-square exact p-value test for deviation of the observed heterozygosities from *H* (Nei's unbiased estimate of gene diversity) using the Markov Chain Method showed that there was no significant (*P* <0.2057) deviation from Hardy-Weinberg equilibrium when the entire data set was analyzed (Table 14).

Table 14. The Chi-square test for deviations from Hardy-Weinberg equilibrium across the 21 loci.

Marker	ChiSquare value	ChiSquare d.f.	ChiSquare p-value	Exact p-value
ICMP3002	11.0587	1	0.0009	0.0030
PSMP2043	50.5892	28	0.0056	0.0000
PSMP2001	6.9815	10	0.7272	0.5690
PSMP2030	16.2694	36	0.9981	0.7620
PSMP2248	38.6213	15	0.0007	0.0000
PSMP2201	15.7242	6	0.0153	0.0570
PSMP2246	20.6467	3	0.0001	0.0000
PSMP2008	35.1115	28	0.1666	0.4700
PSMP2231	46.1709	55	0.7959	0.7190
PSMP2090	24.4773	15	0.0574	0.0190
PSMP2208	8.6982	6	0.1913	0.3590
PSMP2267	5.6998	3	0.1272	0.0810
PSMP2085	23.7006	6	0.0006	0.0000
PSMP2237	15.5825	10	0.1122	0.0040
PSMP2249	22.7075	3	0.0000	0.0000
PSMP2080	26.0185	15	0.0378	0.0260
PSMP2233	12.8332	15	0.6152	0.4800
PSMP2076	14.1528	21	0.8629	0.1540
PSMP2087	17.3299	10	0.0674	0.0000
PSMP2027	34.4019	36	0.5447	0.3340
PSMP2275	1.4748	1	0.2246	0.2840

The level of heterozygosity in the 10 populations analyzed was quite high ranging from 0.29 in population G211 to 0.53 in population G197 (Table 11). The difference between the level of heterozygosity and Nei's unbiased estimate of gene diversity (H) was small.

3.7 Genetic diversity among 30 bulked pearl millet accessions

3.7.1 Number of alleles

The total number of alleles in each of the 30 bulked accessions varied from 42 to 63 with an average of 52.57. The highest number of alleles (63) was observed in accessions G200 and G218 whereas the lowest (42) were observed in accession G204. Out of the 5 accessions from Niger, 3 namely G202, G211 and G129 had the same number of alleles (50 alleles). Four accessions namely G205, G207, G228 and G191 originating from different countries (i.e Senegal, Mali, Burkina Faso and Sudan, respectively) had the same number of alleles (49 alleles). The accessions with the highest and lowest number of alleles (G218 and G204, respectively) originated from Nigeria (Figure 14). The country with the highest number of alleles was Nigeria with 23% and the lowest was Senegal with 17% (Figure 15).

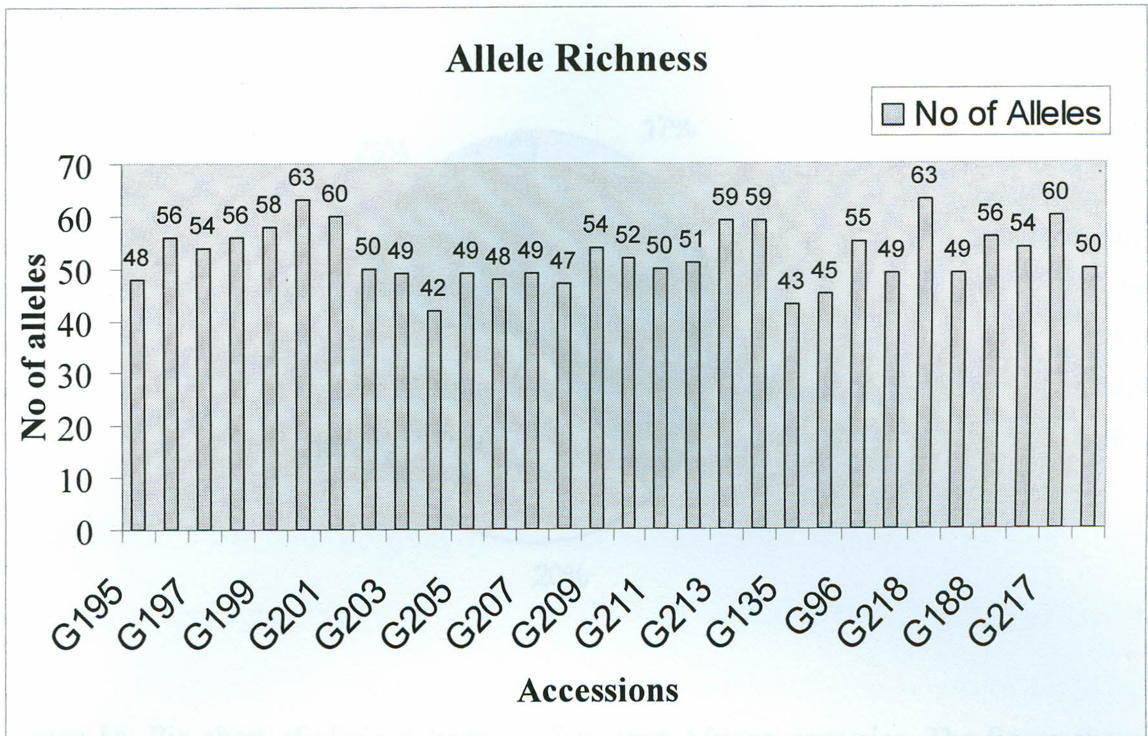


Figure 14: Allele richness in the 30 pearl millet accessions from West Africa. The table shows how the number of alleles in each accession varied between them.

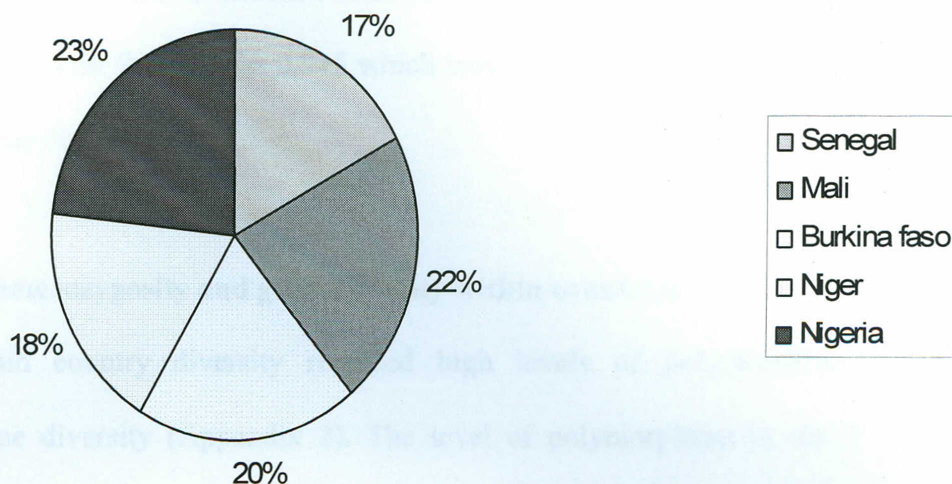


Figure 15: Pie chart of allele richness in five west African countries. The figure shows the variation of total number of alleles in the west African countries in percentage.

3.7.2 Allele frequencies, average heterozygosity and gene diversity of 30 bulked accessions

The diversity parameters in the 30 bulked accessions across the 21 SSR loci are provided in Table 14. The major allele frequency ranged from 0.817 in marker PSMP 2001 to 0.233 in PSMP 2231, with an average 0.46. The highest number of alleles per marker was 11 in PSMP 2231 and the lowest was 2 in PSMP 2275. The average number of alleles per marker was 5. The observed levels of heterozygosity and Nei's unbiased estimate of gene diversity (*H*) for all microsatellites was quite high. The mean Nei's unbiased estimate of gene diversity varied between H=0.858 in PSMP 2231 which was the highest to H= 0.324 in PSMP 2001, the lowest. The mean Nei's unbiased estimate of gene diversity across the 21 loci was H=0.647. The observed heterozygosity level per microsatellite in

the populations was high with the exception of PSMP 2001. The levels ranged from 0.933 in five of the markers to 0.333 in marker PSMP 2001. The mean heterozygosity for all the 30 accessions across the 21 loci was 0.815 which was higher than the mean Nei's unbiased estimate of gene diversity (*H*).

3.7.3 Polymorphism, heterozygosity and gene diversity within countries

The analysis of within country diversity revealed high levels of polymorphism, heterozygosity and gene diversity (Appendix 2). The level of polymorphism in the 7 countries ranged from 76.2% in Benin and Sudan to 100% in Mali, Niger, Nigeria and Burkina Faso. Benin and Sudan had one accession each while the other countries had 5 or more accessions and this may be the cause of differences. The mean Nei's unbiased (estimate of gene diversity (*H*) within countries was variable, ranging from the highest *H*=0.62) in Niger to the lowest (*H*= 0.38) in Sudan. The average of within country diversity was H=0.53. The average heterozygosity was quite high with the exception of Benin. The highest level of heterozygosity was 0.89 observed in Niger accessions.

Table 15: Percentage polymorphism, average heterozygosity and gene diversity of 30 bulked accessions from 7 countries of West Africa.

Country	Polymorphism (%)	Average Gene diversity	Average Heterozygosity
Benin	76.2	0.40	0.00
Mali	100	0.59	0.78
Nigeria	100	0.59	0.86
Sudan	76.2	0.38	0.76
Burkina Faso	100	0.61	0.85
Niger	100	0.62	0.89
Senegal	90.5	0.53	0.71

3.8 Genetic relationships of 30 bulked accessions of pearl millet

3.8.1 Genetic distances

The Roger's modified genetic distance was calculated to determine the relationships among the 30 populations originating from the seven countries and pairwise genetic similarities between individuals' accessions were assessed based on Dice's genetic similarity coefficients (Table 15). The most distant accessions were G202 and G205 from Niger and Senegal respectively. The closest accessions were G213 and G214 both of them from Nigeria. The maximum dissimilarity value was 0.62 between accessions G202 and G205 whereas the minimum dissimilarity value was 0.20 between accessions G213 and G214 (Table 16).

Table 16:Genetic distance matrices of the 30 bulked accessions calculated according to Dice (1945)

	G195	G196	G197	G198	G199	G200	G201	G202	G203	G204	G205	G206	G207	G208	G209	G210	G211	G212	G213	G214	G135	G49	G96	G228	G218	G191	27	28	29
G195	0.34																												
G196	0.47	0.49																											
G197	0.4	0.44	0.4																										
G198	0.51	0.47	0.46	0.35																									
G199	0.42	0.44	0.44	0.31	0.37																								
G200	0.39	0.41	0.47	0.4	0.44	0.37																							
G201	0.47	0.45	0.42	0.4	0.44	0.36	0.47																						
G202	0.4	0.46	0.45	0.33	0.36	0.38	0.35	0.42																					
G203	0.38	0.42	0.46	0.39	0.5	0.45	0.47	0.43	0.4																				
G204	0.42	0.46	0.55	0.47	0.55	0.54	0.45	0.62	0.48	0.56																			
G205	0.52	0.44	0.55	0.42	0.49	0.5	0.48	0.57	0.48	0.56	0.3																		
G206	0.42	0.54	0.48	0.49	0.46	0.45	0.39	0.56	0.44	0.56	0.37	0.48																	
G207	0.43	0.43	0.52	0.38	0.47	0.42	0.42	0.46	0.47	0.55	0.5	0.49	0.44																
G208	0.39	0.41	0.41	0.35	0.36	0.35	0.28	0.37	0.41	0.42	0.44	0.41	0.4	0.37															
G209	0.44	0.42	0.57	0.35	0.4	0.37	0.36	0.45	0.39	0.45	0.49	0.44	0.49	0.37	0.3														
G210	0.39	0.47	0.46	0.36	0.41	0.27	0.4	0.34	0.36	0.41	0.52	0.55	0.47	0.48	0.33	0.39													
G211	0.47	0.53	0.52	0.46	0.43	0.42	0.46	0.43	0.51	0.51	0.54	0.49	0.5	0.41	0.33	0.44	0.43												
G212	0.42	0.47	0.47	0.43	0.47	0.41	0.39	0.43	0.42	0.47	0.48	0.5	0.44	0.43	0.36	0.48	0.32	0.42											
G213	0.48	0.47	0.5	0.44	0.42	0.34	0.38	0.45	0.43	0.5	0.48	0.5	0.43	0.45	0.35	0.46	0.34	0.45	0.2										
G214	0.35	0.42	0.4	0.51	0.49	0.45	0.38	0.48	0.41	0.52	0.34	0.51	0.31	0.46	0.34	0.4	0.45	0.48	0.44	0.44									
G135	0.53	0.48	0.49	0.37	0.46	0.41	0.43	0.39	0.43	0.49	0.49	0.44	0.49	0.28	0.39	0.42	0.45	0.42	0.42	0.48	0.52								
G49	0.43	0.48	0.44	0.41	0.49	0.45	0.39	0.42	0.41	0.47	0.5	0.54	0.45	0.43	0.34	0.45	0.34	0.43	0.36	0.45	0.45	0.52							
G96	0.44	0.5	0.45	0.33	0.47	0.42	0.42	0.44	0.43	0.44	0.48	0.48	0.44	0.39	0.39	0.31	0.46	0.48	0.53	0.51	0.48	0.41	0.44						
G228	0.32	0.41	0.52	0.39	0.49	0.35	0.32	0.45	0.35	0.41	0.38	0.44	0.43	0.44	0.32	0.39	0.27	0.46	0.31	0.31	0.43	0.48	0.37	0.44					
G218	0.42	0.44	0.5	0.47	0.48	0.41	0.39	0.45	0.41	0.45	0.47	0.42	0.47	0.5	0.34	0.45	0.41	0.4	0.44	0.48	0.42	0.4	0.39	0.43	0.41				
G191	0.37	0.41	0.45	0.39	0.42	0.38	0.36	0.4	0.3	0.45	0.41	0.42	0.45	0.5	0.35	0.46	0.34	0.4	0.36	0.41	0.39	0.47	0.44	0.48	0.26	0.41			
G188	0.45	0.5	0.31	0.42	0.45	0.42	0.46	0.4	0.43	0.54	0.46	0.51	0.4	0.41	0.33	0.49	0.42	0.5	0.42	0.45	0.33	0.41	0.46	0.45	0.47	0.42	0.45		
G215	0.38	0.44	0.45	0.42	0.5	0.37	0.34	0.44	0.41	0.49	0.46	0.48	0.39	0.42	0.3	0.42	0.33	0.44	0.38	0.42	0.43	0.48	0.38	0.39	0.31	0.43	0.32	0.45	
G217	0.37	0.49	0.52	0.42	0.48	0.36	0.36	0.46	0.4	0.48	0.47	0.49	0.41	0.46	0.4	0.47	0.32	0.47	0.34	0.39	0.47	0.41	0.43	0.51	0.29	0.43	0.28	0.5	0.33

The distance matrix (Table 16) of the pair-wise genetic distances between accessions was subjected to sequential agglomerative hierarchical nested (SAHN) using unweighted pair-group analysis (UPGMA) using Dice's indices as implemented in Darwin 4.0. The hierarchical cluster analysis (HCA) partitioned the accessions in accordance with their origin from different breeding groups that were used in factorial and diallel crosses. Three major clusters were generated from Nei and Li genetic distance matrices. Cluster A contained accessions G129, G217, G202, G212 and G201 from Niger, G 188, G218, G214 and G213 from Nigeria, G191 from Sudan, G96 from Benin and G200, G209 from Burkina Faso (Figure 16). Cluster B consisted of accessions G198, G49, G208, G207, G217 and G197 from Mali, G228, G210 and G199 from Burkina Faso, G135, G206 and G205 from Senegal and G203 from Nigeria. Cluster C contained accessions G196, G195 from Senegal and G204 from Nigeria (Figure 17).

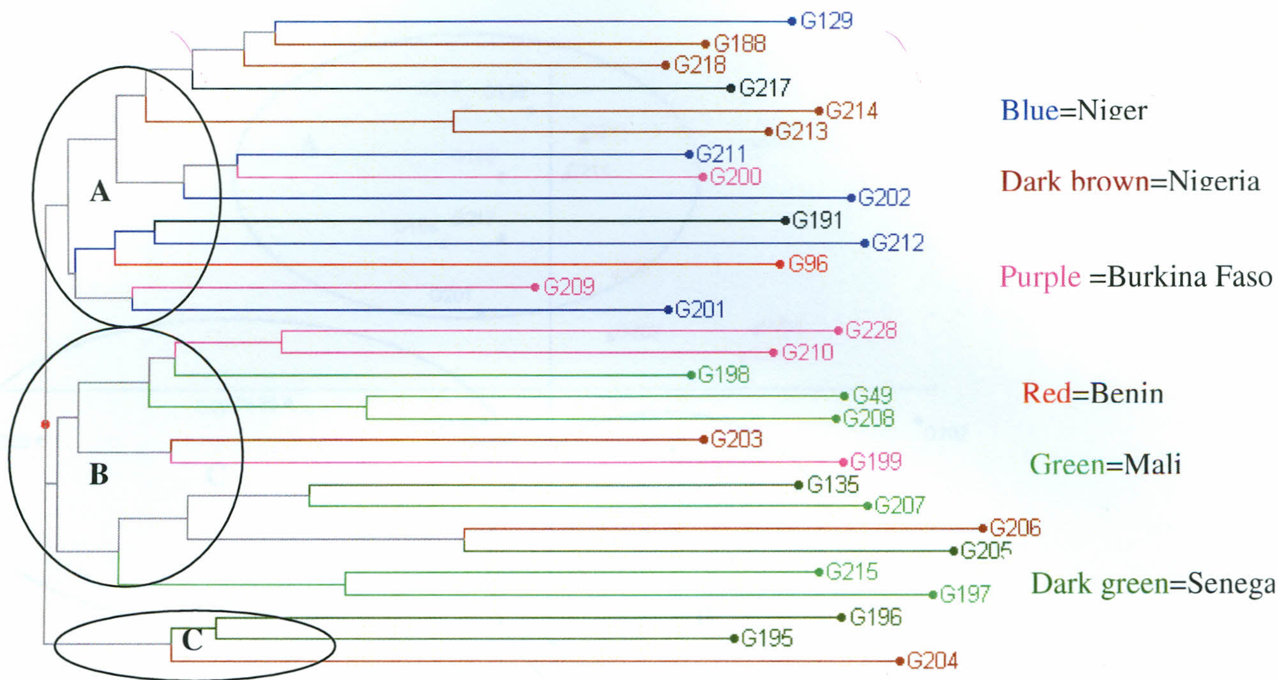


Figure 16: A dendrogram of the relationships between 30 accessions from different West Africa countries using the UPGMA method of Darwin Ver. 4.0.

3.8.2 Principal Co-ordinate Analysis based on genetic distance estimates of 30 bulked accessions

The patterns of cluster analysis were confirmed by principal co-ordinate analysis (PCoA). A scatter plot of the first and second axes of non-metric multi-dimensional scaling (MDS) revealed three clusters of inter-relationships among accessions (Figure 17).The first Eigen vector explained 12.9 % variation. The analysis showed that the accessions generally clustered on the basis of the geographical origins.Accessions from Senegal clustered together in group C but were widely distributed. Materials from Niger and

Nigeria grouped together (in group A), while those from Mali and Burkina Faso grouped together in B.

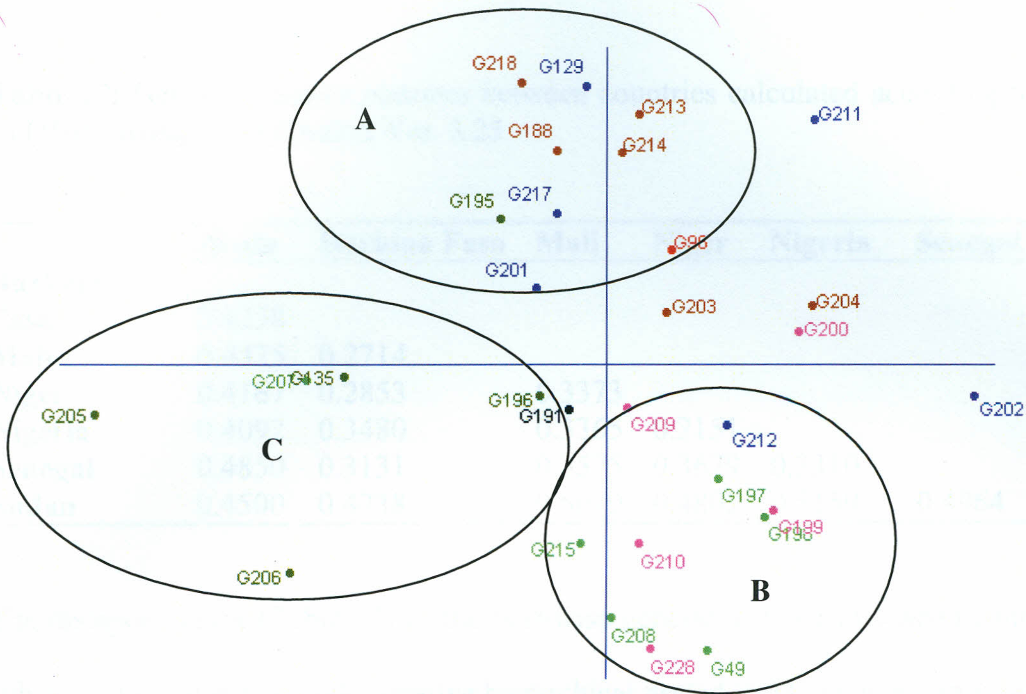


Figure 17: Principle co-ordinate scatter plot showing genetic distance estimates of 30 bulked accessions. The figures illustrates how the different accessions grouped into three different clusters. The clustering depended on the geographical origin of the accessions.

Mali and Burkina Faso are also quite close and the Senegal accessions are generally further from the other West African accessions, which may be attributed to geographical distance and/or climatic differences.

Table 17: Genetic distance matrices between countries calculated according to Nei and Li (1979) using PowerMarker Ver. 3.25

	Benin	Burkina Faso	Mali	Niger	Nigeria	Senegal	Sudan
Burkina Faso	0.4238						
Mali	0.4375	0.2714					
Niger	0.4167	0.2853	0.3373				
Nigeria	0.4092	0.3480	0.3365	0.2151			
Senegal	0.4850	0.3131	0.3575	0.3679	0.3310		
Sudan	0.4500	0.4738	0.5079	0.4802	0.5159	0.4964	

The distance matrix (Table 17) of the pair-wise genetic distances between countries was subjected to sequential agglomerative hierarchical nested (SAHN) using unweighted pair-group analysis (UPGMA) using Dice's indices as implemented in Darwin 4.0 resulting in five distinct clusters (clusters A, B, C, D, E, F)). The largest cluster (Number B comprised Senegal, Mali, Burkina Faso, Niger and Nigeria (Figure 18). However Niger and Nigeria were closely related and formed a sub-cluster. Mali and Burkina Faso were also closely related and formed another sub-cluster. Senegal was more distinct than all the other countries in this cluster. The second cluster (number A comprised of accessions from Niger and Nigeria ICRISAT stations. The accessions from Benin, Sudan and Mali ICRISAT separated clearly as independent clusters. .

Mali and Burkina Faso are also quite close and the Senegal accessions are generally further from the other West African accessions, which may be attributed to geographical distance and/or climatic differences.

Table 17: Genetic distance matrices between countries calculated according to Nei and Li (1979) using PowerMarker Ver. 3.25

	Benin	Burkina Faso	Mali	Niger	Nigeria	Senegal	Sudan
Burkina Faso	0.4238						
Mali	0.4375	0.2714					
Niger	0.4167	0.2853	0.3373				
Nigeria	0.4092	0.3480	0.3365	0.2151			
Senegal	0.4850	0.3131	0.3575	0.3679	0.3310		
Sudan	0.4500	0.4738	0.5079	0.4802	0.5159	0.4964	

The distance matrix (Table 17) of the pair-wise genetic distances between countries was subjected to sequential agglomerative hierarchical nested (SAHN) using unweighted pair-group analysis (UPGMA) using Dice's indices as implemented in Darwin 4.0 resulting in five distinct clusters (clusters A, B, C, D, E, F)). The largest cluster (Number B comprised Senegal, Mali, Burkina Faso, Niger and Nigeria (Figure 18). However Niger and Nigeria were closely related and formed a sub-cluster. Mali and Burkina Faso were also closely related and formed another sub-cluster. Senegal was more distinct than all the other countries in this cluster. The second cluster (number A comprised of accessions from Niger and Nigeria ICRISAT stations. The accessions from Benin, Sudan and Mali ICRISAT separated clearly as independent clusters. .

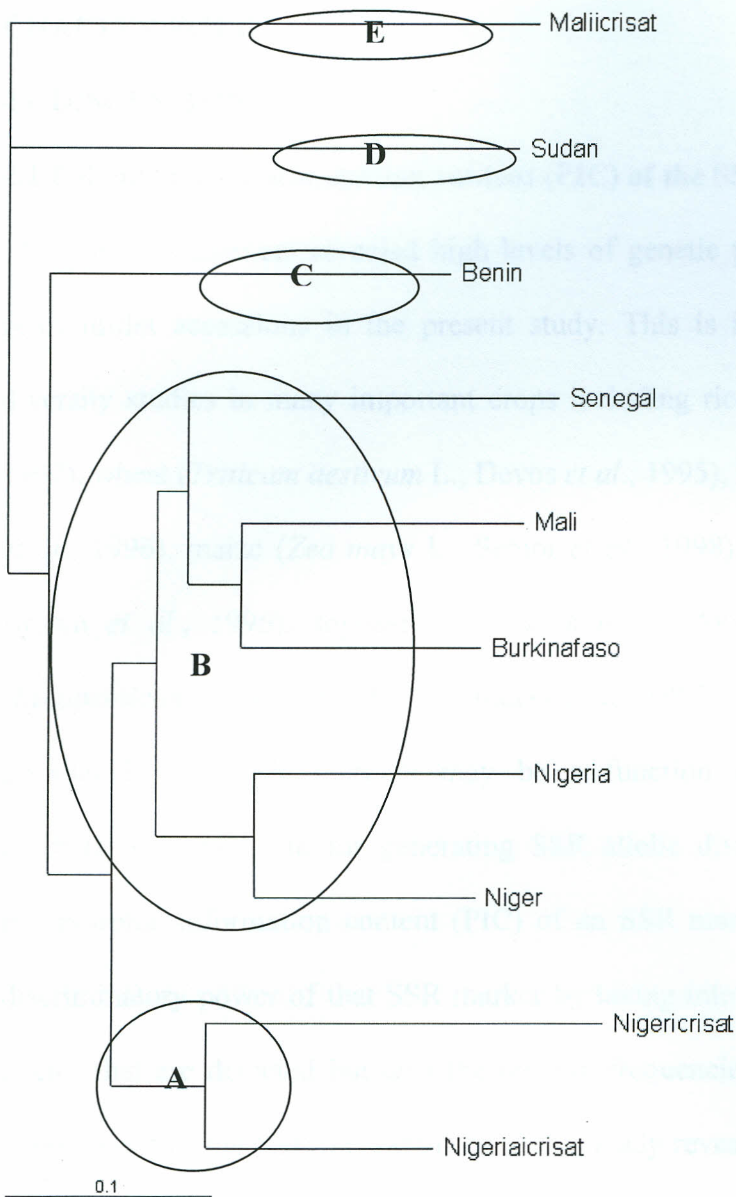


Figure 18: Dendrogram of distribution of diversity between countries based on 30 accessions using Neighbor joining algorithm of Power Marker ver. 3.25. The figure illustrates a clustering of accessions into different groups. The grouping depended on the geographical origins of the accessions.

CHAPTER FOUR

4.0 DISCUSSION

4.1 Polymorphism information content (PIC) of the SSR markers

Microsatellite markers revealed high levels of genetic polymorphism among the tested pearl millet accessions in the present study. This is in agreement with SSR marker diversity studies in many important crops including rice (*Oryza sativa* L., Chen *et al.*, 1997), wheat (*Triticum aestivum* L., Devos *et al.*, 1995), barley (*Hordeum vulgare* L., Liu *et al.*, 1996), maize (*Zea mays* L., Senior *et al.*, 1998), sorghum (*Sorghum bicolor* L., Brown *et al.*, 1996), soybean (*Glycine max* L., Akkaya *et al.*, 1992) and tomato (*Lycopersicon esculentum* Mill., Smulders *et al.*, 1997). The high level of polymorphism associated with SSR markers may be a function of unique replication slippage mechanism responsible for generating SSR allelic diversity (Pejic *et al.*, 1998). The polymorphic information content (PIC) of an SSR marker provides an estimate of the discriminatory power of that SSR marker by taking into account not only the number of alleles that are detected but also the relative frequencies of those alleles (Smith *et al.*, 2000). Most of the SSR markers used in this study revealed a high discriminatory power (PIC of greater than 0.5) and hence were highly informative. This is in agreement with previous studies by Qi *et al.* (2004) and Mariac *et al.* (2006), who found a PIC of greater than 0.5 with the same primers. Weber (1990) reported that there was a significant

relationship between the repeat length and the degree of polymorphism in human microsatellites.

Microsatellite markers have been used to investigate genetic diversity of a large number of cultivars in rice (Yeng *et al.*, 1994), wheat (Plaschke *et al.*, 1995) and maize (Senior *et al.*, 1998). The number of alleles amplified per primer pair ranged from 3 to 25 in rice, 3 to 16 in wheat and 2 to 23 for maize. In the present study 2 to 25 alleles per primer were amplified from the 79 samples of pearl millet. Thus the level of microsatellite polymorphisms in pearl millet is relatively high and similar to other out-crossing crops.

One possible reason is that the materials used in the present study were from diverse geographical areas and thus had a relatively wide genetic base. In addition, West Africa is considered a centre of origin and diversity for pearl millet. In this study, several unique/rare alleles were observed in both individuals and bulks of pearl millet. The presence of many unique alleles may be explained by the relatively high rate of mutation in SSR loci (Henderson and Petes, 1992). Such alleles are important because they may be diagnostic for particular regions of the genome specific to a particular type of pearl millet. The number of rare alleles within landrace accessions was high reflecting high variability within the accessions. This is probably due to the cross-pollinated nature of the crop. High levels of allelic variability and heterozygosity in individuals was observed in this study. This could be attributed to the outbreeding nature of pearl millet, where the proportion of heterozygous loci is likely to be high.

4.2 Genetic diversity and heterozygosity

Genetic diversity is measured in terms of number of alleles per locus and their frequencies and Nei's unbiased estimate of gene diversity (*H*) (Nei, 1987). High genetic diversity was observed in this study with both individuals and bulks displaying different number of alleles per locus (up to 8 alleles). The present study used materials from Senegal, Mali, Burkina Faso, Niger, Benin, Nigeria, and Sudan. Niger, the 2nd largest pearl millet producer in Africa, after Nigeria, showed the highest morphological diversity of pearl millet in West Africa (Tostain, 1994). Nigeria clustered with Niger probably because it neighbors Niger and there might have been gene flow between the two countries through seed exchanges between farmers and this may have been the main mechanism of shaping molecular diversity in landraces at the regional scale (Dje *et al.*, 1999; Tsegaye *et al.*, 1996). Nigeria's genetic diversity was H=0.59 which was close to Niger's (H=0.62).

4.3 Population structure

Analysis of Molecular Variance (AMOVA) is carried out to assess overall distribution of diversity. Diversity among populations within countries was rather low, while within-population diversity was high; this could be explained by the fact that landraces within the country have not been subjected to deliberate selection. The same was found when AFLPs were used to study diversity of pearl millet in Niger (Mariac *et al.*, 2006). A

similar situation has also been observed in Rajasthan-India (Vom Brocke *et al.*,2003) where frequent importation of seed lots by farmers explain the very low differentiation between pearl millet populations cultivated in distant villages. Eco-geographical adaptation, coupled with human selection for diversified needs, has led to substantial differentiation within pear millet landraces. Also the allogamous breeding behavior contributes to within-population diversity (Busso *et al.*, 2000). Genetic variation was higher among countries. This was because a landrace is a population that is conspicuously variable between and within sites (Busso *et al.*, 2000). A study of pearl millet in West Africa showed that differentiation among morphological categories within a geographical origin explained 6.98% differentiation among accessions (Mariac *et al.*, 2006). In the present study, differentiation among accessions within countries was 9.56%. This was slightly higher probably because the materials that were used in this study were collected from several different sites within each country while in the previous study, the materials were from only two locations within the same country.

Genetic variation within populations was very high. Landraces of pearl millet display a much larger within-population genetic variation (Rai *et al.*, 1999) than between-population variation. Considerable within-population variation has been also detected by RFLPs in the assessment of genetic diversity of pearl millet landraces from the core collection at ICRISAT gene bank (Bhattacharjee *et al.*, 2002). Equally, Tostain *et al.*, 1987 reported that accessions from Niger showed very high intra-population variation. For cross-pollinating crops, some genetic studies indicate that mixture and replacement

serves the purpose of protecting the genetic viability of the seed (Berthaud *et al.*, 2002). In the present study, the within-population diversity accounted for 84.59% of the variation, which was very close to the 86.86% observed by Berthaud *et al.* (2002) using AFLPs.

Such high levels of within-landrace genetic diversity point to the potential use of these landraces in breeding programs. Presence of genetic variation is a precondition for achieving selection gains. Given the high intra-population diversity of the studied landraces, they could be used directly as initial materials in a recurrent population improvement program.

The observed high intra-population genetic variation may also reflect the value of within-cultivar diversity for adaptation to variable climates. West Africa pearl-millet-growing areas are characterized by a highly variable beginning of the rainy season and appearance of drought periods at any time during the growing season. In such unpredictably variable, harsh environments, within-cultivar diversity may enhance the capacity of populational buffering and therefore yield stability of the crop (B.I.G. Haussmann, pers. communication). More investigations are needed to better understand the role of intra-population diversity in adaptation to climate variability and possibly future climate change.

The inbreeding level was minimal while the heterozygosity was excess. In an out-breeding species such as pearl millet, the proportion of heterozygous loci is likely to be high (Busso *et al.*, 2000). The 'male' component in a mixed planting regime in small fields on small farms is expected to be extremely variable (Busso *et al.*, 2000). Moreover, the pollen cloud is likely to include pollen from wild relatives and volunteer plants growing nearby as well as from cultivated plants growing in the same and adjacent fields. This heterogeneity was reflected in the excess heterozygotes in this study with the Theta (P) value being negative. The accessions studied had been previously multiplied at ICRISAT-Niger through sibbing, i.e., collecting and mixing pollen from 30-40 panicles, and distributing it on other 30-40 panicles where the female part was flowering and anthers had not yet appeared (protogynie). These sibbed panicles were harvested; always at least 35-40 sibbed panicles per accession were collected to assure appropriate effective population size (i.e. at least 70 individuals contributing to the next generation). The sibbing can lead to excess heterozygotes compared to HWG equilibrium, because it excluded selfing, whereas in a random mating population selfings can occur at a probability of 1/N, with N being the number of individuals in the population (Haussman *et al.*, 2006)

There was moderate but significant (p>0.003) population differentiation in this study. However, founder effects, and the fact that plants from different generations and even from different geographical origins (because of seed flows due to farmers' exchanges) probably contribute to the seed bank, could be causative factors for the maintenance of a

low, but significant differentiation. Quantitative-genetic models have shown that population differentiation can be high at the trait (phenotypic) level and low at the allelic frequency for the majority of QTLs underlying the trait, especially when both gene flow and diversifying selection are strong (Le Corre and Kremar, 2003).

4.4 Genetic relationships

The principal component analysis clustered the accessions according to their geographical origins. A study of pearl millet in West Africa equally showed that domesticated and weedy pearl millet clustered according to their geographic origin (Mariac *et al.*, 2006).

Allozyme analysis on wild and domesticated pearl millet samples from throughout Africa showed that diversity is patterned according to geographic distances at a large scale (Pilat-Andre` 1992). Busso *et al.* (2000) observed that genetic distances are related to geographical separation.

In this study, landraces from Senegal, Mali, Burkina Faso clustered together as well as landraces from Niger and Nigeria. Materials from Sudan and Benin, as well as improved cultivars clustered independently from the other accessions. Landraces from Niger and Nigeria clustered particularly close, most probably because Nigeria is neighboring Niger and there could have been seed flows as this has been reported before by Le Corre and Kremar (2003) hence sharing a common genetic pool. .

The observed clusters give first indications on potential heterotic groups within the studied materials. However, these data need to be combined with field studies of heterosis or hybrid superiority and combining ability in diallel or factorial crosses among landraces belonging to the same *versus* different clusters. Studying hybrid vigor in putative intra- *versus* inter-pool crosses will help to determine the optimal genetic diversity among parental populations to maximize heterozygosity (without disturbing integrity of the genome) and therefore hybrid vigor (B.I.G Haussmann, pers. communication). In fact, different sets of factorial and diallel crosses among the landrace accessions analyzed in the present study are currently being evaluated in multi-location field trials in West Africa to study patterns of hybrid vigor and combining ability in the various crossing progenies, and to finally link the phenotypic observations to marker data and genetic distances obtained in the present study.

The first axis of the PCA explained 12.19% of the overall variation in the bulks and 10.2% in the individuals. The variation was comparatively high and significant since the different accessions presented a unique allele combination. This was expected since pearl millet is an out-crossing crop. A similar finding was observed by Mariac *et al* (2006) in his study of wild pearl millet. In contrast, Matsuoka *et al.* (2002) observed low percentages of variation explained on the two axes (4.25% on first axis and 1.6% on second axis) which he explained as common while using a high number of alleles from different microsatellite loci.

CHAPTER FIVE

5.0 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In this study, the genetic diversity of pearl millet landraces was evaluated using SSR markers. The genetic parameters estimated from SSR data indicated high levels of genetic diversity in the landraces of pearl millet from West Africa. The genetic diversity obtained was higher within accessions than between accessions. The partitioning of variance components based on the analysis of molecular variance (AMOVA) for diversity analysis revealed a low variability between accessions (9.56%) but a high variability within accessions (84.59%). The high level of within-accession diversity is expected since pearl millet is predominantly an out-crossing species.

Geographical separation was associated with overall genetic differences. A dendrogram based on the dissimilarity matrix obtained using UPGMA algorithm method of Darwin Version 4.0 delineated the 49 individuals into four major clusters and the 30 accessions into three major clusters. Cluster analysis (CA) partitioned the accessions in accordance with their origin from different breeding groups and geographical locations. This study has added significant knowledge and understanding of the population structure and genetic diversity of pearl millet in West Africa and partitioning among countries, between and within populations. Niger accessions had the highest within-country genetic diversity .

5.2 Recommendation

This study recommends that all the accessions should be used in breeding programs because they have displayed a high level of genetic diversity. The clusters are potential heterotic groups which can be exploited in population crosses in breeding programs to improve their hybrid vigour. In previous studies, morphological markers have been used to study the genetic diversity of pearl millet landraces (Bhattacharjee *et al*., 2006). In this study SSR markers were used to determine the genetic diversity of pearl millet landraces. Therefore, there is need to assess the level of correlation between morphological and molecular traits among the 30 accessions of pearl millet from West Africa.

CHAPTER SIX

6.0 REFERENCES

Agricultural Research Corporation ARC Sudan. (Accessed 3/30/06) (http://www.arcsudan.org/millet.htm) Millet Research Program.

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

Almekinders, C. J. M., Louwaars, N. P., de Bruijn G. H. 1994. Local seed systems and their importance for an improved seed supply in developing countries. *Euphytica* 78: 207 – 216.

Anand Kumar, K.B., Ouendeba, S., Boureima, S. C., Gupta, S. and Ouattara, 1999. Current status of hybrid pearl millet in West Africa.Proc. West African Hybrid Sorghum and Pearl millet Seed Workshop, 28 September-2 October 1998, Niamey, INRAN/INTSORMIL/ICRISAT,in press.

Andrews, D. J., Hanna, W. W., Rajewski, J. F., Collins, V. P. 1996. Advances in grain pearl millet: Utilization and production research. p. 170-177. In: J. Janick (ed.), Progress in new crops. ASHS Press, Alexandria, VA.

Andrews D. J., Bramel-Cox K. 1994. breeding cultivars for sustainable crop production in low-input dry land agriculture in tropics. In: Buxton D.A (ed.), International Crop science, Volume 1 CSSA, madison, W. L., pp.211-222.

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

Athwal, D. S. 1965: Hybrid bajra-1 marks a new era. *Indian Farmg* 15: 6 - 7.

Ayana, A., Bryngelssou, T., Bekele, E. 2000. Genetic variation of Ethiopian and Erithrean Sorghum (Sorghum. Bicolor (L.) Moench) germplasm assessed by random amplied polymorphic DNA (RAPD). *Genetic Resourse Crop Evolution* 47: 471- 482.

Ayres, N. M., McClung, A. M., Larkin, P. D., Bligh, H. F. J., Jones, C. A., Park, W. D. 1997. Microsatellite and a single nucleotide polymorphism differentiate apparent amylase classes in an extended pedigree of US rice germplasm. *Theoretical Applied Genetics* 94: 773 – 881.

Baker, R. D. 2003. Millet Production, New Mexico State University.

Bekele, E. 1984. Analysis of regional patterns of phenotypic diversity in the Ethiopian tetraploid and hexaploid wheats. *Hereditas* 100: 131-154.

Beckmann, J.,S, Soller, M. 1990. Towards a united approach to genetic mapping of eukaryotes based on sequence tagged microsatellite sites. *Bio/technology* 8: 930 - 932.

Berthaud J., Pressoir G., Ramirez_corona F., Bellon M. R. 2002. farmres manegement of maize landraces diversity. A case study in Oaxaca, and beyond. Proceedings of the inernational symposium on biosafety of geneticlly modified organisms, Beijig, china, october 10-16.

Bertin, I., Zhu, J. H., Gale, M. D., 2005. SSCP_SNP in pearl millet – a new markrer system for comparative genetics. *Theoretical Applied Genetics* 110: 467-1472.

Bhattacharjee, R., Bramel, P. J., Hash, C. T., Koleshikova-Allen, M. A., Khairwal, I. S., 2002. Assessment of genetic diversity within and between pearl millet landraces. *Theoretical Applied Genetics* 105: 666 – 673.

Bhattacharjee, R., Khairwal, I. S., Bramel, P. J., Reddy, K. N. 2006. Establishment of a pearl millet [*Pennisetum glaucum* (L.) R. Br.] core collection based on geographical distribution and quantitative traits. *Euphytica.*

Blair, M., McCouch, S. R. 1997. Microsatellite and sequence-tagged site markers diagnostic for the bacterial blight resistance gene, Xa-5. *Theoretical Applied Genetics* 95: 174-184.

Bowers,J., Boursiquet, J.,This, P., Chu, K., Johansson, H., Meredith, C. 1999. Historical genetics. The parentage of chardonnay, Gamay and other wine grapes of Northeastern France. *Science* 285: 1562-1565.

Bostein, D., White, R. L. Solnick, M., Davis, R. W. 1980. Construction of a genetic linkage map in Man using restriction fragment Length polymorphisms. *American Journal of Human Genetics* 32: 314 – 331.

Bramel, P. J. 2004. *Assessing the Risk of losses of Biodiversity in Traditional Cropping Systems: A case study of Pigeonpea in Andra Pradesh.* International Crops Research Institute for the Semi-Arid tropics pantacheru, 502 324, Andhra Pradesh, India. Pg 168.

Bryan, G. J., Collins, A. J., Stephenson, P., Orry, A., Smith, J. B., Gale, M. D. 1997. Isolation and characterization of of microsatellites from hexaploid bread wheat. *Theoretical Applied Genetics* 93: 190 – 198.

Brown, S. M., Hopkins, M. S., Mitchell, S. E., Senior, M. L., Wang, T. Y., Duncan, R. R., Gonzalez-Candelas, F., Kresovich, S. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum *[Sorghum bicolour* (L.) Moench]. *Theoritical Aapplied Genetics* 93: 190-198.

Brown, J. H., Lomolino, M. V. 1998. Biogeography, Sinauer Associates, Inc.USA.

Bruken, J. N. 1977. A systematic study of Pennisetum sect. pennisetum (Graminae). *American Journal of Botany* 64: 161-176.

Brummer, E.C., 1999. Capturing heterosis in forage crop cultivar development. *Crop Science* 39: 934 - 954.

Budak, H., Pedraza F., Cregan P. B., Baenzinger, P. S., Dweikat, I. 2003. Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of Pearl millet germplasm. *Crop Science* 43: 2284 – 2290.

Buhariwalla, H. K., Crouch, J. H. 2004. Optimization of marker screening protocol to assess the degree and distribution of genetic diversity in landraces of pigeonpea. *In*:

Burr, B. 1994. Some concepts and new methods for molecular mapping in plants. In: Philips RL, Vasil IK (eds) DNA-based markers in plants. *Kluwer Academic Publishers*, Pordrecht, the NetherLands, pp 1-7.

Burton, G. W. 1965. Pearl millet Tift 23A released. *Crops and Soils* 17: 19.

Burton, G. W. 1974: Factors affecting pollen movement and natural crossing in pearl millet.*Crop Science* 14: 802 - 805.

Busso, C. S., Liu, C. J., Hash, C. T., Witcombe, J. R., Devos, K. M., de Wet, J .M. J. Gale, M. D. 1995. Analysis of recombination rate during male and female gametogenesis in pearl millet (*Pennisetum glaucum*) using RFLP markers. *Theoretical Applied Genetics* 90: 242 – 246.

Busso C. S., Devos K. M., Ross G., Mortomore M., Adams W.M., Ambrose M. J., Alldricks, gale M. D. 2000. Genetic diversity within and among landraces of pearl millet (pennisetum glaucum under farmer management in west-africa. Genetic resource crop Evolution 47:561-568.

Chen, X., Terunyay, S., Xu, Y., Cho, Y. G. McCouch, S. R. 1997. Development of a microsatellite framework map providing genome-wide coverage in rice (Oryza savita L.). *Theoretical Applied Genetics* 95: 553 -567.

Consultative Group on International Agricultural Research (CGIAR). 2005. *Research &Impact: Areas of Research: Millet: Areas of Research: Millet.*

Chowdari, K.V., Davierwala, A. P., Gurpta, V. S., Ranjekar, P.K. Govilla P. 1998a. Genotype identification and assessment of genetic relationships in pearl millet (*Pennisetum glaucum* (L.) R. Br.) using microsatellites and RAPDS. *Theoretical Applied Genetics* 97: 154 - 162.

Dawson, L. K., Powell, W. 1999. Genetic variation in the Afromontane tree *Prunus africana,* an endangered medicinal species. *Molecular Ecology* 8: 151-156.

Dean, R. E., Dahlbrg J. A., Hopkins, M. S., Mitchell, S. E., Kresovitch, S. 1999. Genetic reduodancy and diversity among orange accessions in the U.S National Sorghum collection as assessed with simple sequence repeat SSR markers.*Crop Science*. 39: 1215 - 1221.

Demisssie, A. Bjornstad, A. 1996. Phenotypic diversity of Ethiopian barley in relation to geographical regions, altitudinal range and agro-ecological zones: as an aid to germplasm collection and preservation strategy. *Hereditas* 124: 17 - 29.

Depeige, A., Goubely, C., Lenio,r, A., Cocherel, S., Picard G., Raynal M., Grellet, F., Delseny, M. 1995. Identification of the most represented repeated motifs in microsatellites loci. *Theoretical Applied Genetics* 91: 160 - 168.

Dogget, H. Parsada, R. K. E. 1995. In Smartt J. and Simmonds N. W. (eds), Evolution of Crops Plants, 2nd ed. Longman, UK, pp. 173-180.

Dogget, H. 1988. Sorghum, 2nd edition. Longman, UK.

Devos, K.M., Pittaway, T. S., Busso, C. S., Gale, M. D., Witcombe, J. R., Hash, C. T. 1995. Molecular tools for the pearl millet nuclear genome. *International Sorghum and Millets Newsletter* 36: 64-66.

Devos, K. M., and Gale, M. D. 1997. Comparative genetics in the grasses. *Plant Molecular Biology* 35: 3.-.15.

Dice, L.R. 1945. Measures of the amount of ecological associations between species. *Ecology* 2: 297-302.

Dje Y., Forcioli D., Ater M., Lefebvre C. vekemans X. 1999. Assessing from North-western Moroco using allozyme and microsatellite markers. *Theoretical Applied Genetics* 99:157-163

Dubreuil, P., Charcosset, A. 1998. Genetic diversity within and among maize populations: A copmarision between isozyme and nuclear RFLP loci. *Theoretical Applied Genetics.* 96: 577 – 587.

Food Agricultural Organization (FAO). 1995. Production Yearbook. Vol.49. FAO, Roe, Italy.

Food Agricultural Organization Statistics (FAOSTAT) 2004 http://appsfoorg/defaulthtm

Food Agricultural Organization Statistics (FAOSTAT) 2005 http://appsfoorg/defaulthtm

Grist, S. A., Firgaira, F. A., Morley, A. A. 1993. Dinucleotide repeat polymorphisms isolated by the polymerase chain reaction. *Biological Techniques* 15: 304 – 309.

Gupta, P. K. Varshney, R. K. 2000. The development and use of microsatellites markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica* 113: 163 - 185.

Huttner, E., Wenzel, P., Akbari, M., Caig, V., Carling, J., Cayla, C., Evers, M., Jaccoud, D., Peng, K., Patarapuwandol, S., Uszynski, G., Xia, L., Yang, S., Kilian, A. 2004. Diversity Array Technology: A novel Tool for Harnessing the Genetic Potential of Orphan Crops.

Hartl, D. L. 1987. A primer of population genetics, 2nd Edition. *Sinauer.Associates Inc. Publishers*. Sunderland, Massachusetts.

Hash, C. T., Witcombe, J. R., Thakur, R. P., Bhatnagar, S. K., Singh, S. D., Wilson, J. P. 1997. Breeding for pearl millet disease resistance. *Proceedings International Conference on Genetic Improvement of Sorghum and pearl millet*, 22–27 September. 1996, Lubbock, 337 - 372.

Haussmann B.I.G., A. Boubacar A., S.S. Boureima S.S., and Y. Vigouroux. 2006. Multiplication and preliminary characterization of West and Central African pearl millet landraces. International Sorghum and Millet Newsletter 47: 110-112.

House, L. R., 1995. Sorghum: One of the world's great cereals. *African crop science Journal* 3: 135 -142.

Henderson, S. T., Petes, T. D. 1992. Insability of simple sequence DNA in Saccharomyces cerevisae. *Molecular Cell Biology* 12 : 2749 – 2757.

ICIPE. 2002. Mbita Point Research and Training Centre. Management Approaches: Host plants: millets.

International Crops Research Institute for the Semi-Arid Tropics and Food and Agriculture organization of the United Nations. 1996. *The world sorghum and millet economics: facts, trends and outlook*. International Crops Research Institute for the Semi-Arid tropics (ICRISAT), Patancheru, India p.68.

Jaccoud, D., Peng. K, Feinstein, D. Kilian. A. 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research*, 29(4): 25.

Jones, C. J, Edwards, K. J, Castaglione, S., Winfield, M. O., Sala, F., Vande, W. C. 1997. Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3: 381-390.

Kiambi, D.,K., Ford-Lioyd, B. V., Newbury, H. J., Maxted, N. 2001. Mapping the geographic distribution of molecular diversity of AFLPs using *O.Longistaminata* as case study. PhD, Thesis, University of Birmingham.

Koh, H. J., Heu, M. H., McCouch, S. R . 1996. Molecular mapping of the gene controlling the super giant embryo character in rice (Oryza sativa L.). *Theoritical Applied Genetics* 93: 257 - 261.

Kong, L., Dong L., Hart, G. E. 2000. Characteristics, linkage map position, and allelic differentiation of sorgum bicolor(L.) Moench) DNa simple-sequence repeats (SSRs). *Theoretical Applied Genetics* 101: 438 - 448.

Kruskal J. B. 1964a. Multidimensional scaling by optimasing goodness of fit to a non-metric hypothesis. *Psychometrika* 29:1-27.

Kruskal J. B.1964b. Non-metric multidimensional scaling: a numerical method. *Psychometrika* 29: 28-42.

Kumar K. A., Andrews, D. J., Jain, R. P., Singh ,S. D. 1984. ICMA-1 and ICMB-1 pearl Millet Parental lines with A1 cytoplasmic genic male-sterity system. *Crop Science* 24: 832.

Lambert, C. 1983. L'IRAT et l'amelioration du mil.*L'Agron. Trop.* 38: 78 - 88.

Le Corre V., Kremer A. 2003. genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* 164:1205-1219.

Li, X. C., Wu, J. L. 1998. SSR heterogenic patterns of parents for marking and predicting heterosis in rice breeding. *Molecular Breeding* 4: 263 – 268.

Li, C. D. Rossnagel, B. G., Scoles, G. J. 2000. The development of oat microsatellite markers and their use in identifying Arena species and oat cultivars. *Theoretical Applied Genetics (In press).*

Liu, C. J., Witcombe, T. S., Pittaway, M., Hash, C. S., Gale, M. D. 1992. Restriction length polymorphism in pearl millet, *Pennisetum glau*cum. p. 233-241 *in*

Liu, C. J., J. R., Witcombe, C. T., Hash, C. S., Busso, T. S., Pittaway, M., Gale, M D. 1994a. Pg. 57-64 *in* J. R. Witcombe and R. R. Duncan (eds.): Use of molecular markers in sorghum and pearl millet breeding for developing countries. Overseas Development Administration: London, U.K.

Liu, C. J., Devos, K. M., Witcombe, J. R, Pittaway, T. S., Gale, M. D. 1996. The effect of genome and sex on recombination rates in pennisetum species, *Theoretical Applied Genetics* 91: 902 – 908.

Liu K., Muse S. 2004. Powermarker: new genetic data analysis software version 3.7 free program distributed by the author over the internet from http://www.powermarker.net.

Mohan, M., Nair, S., Bhagwat, A., Krishma, T. G., Yano, M. 1997. Genome mapping, molecular markers and marker-Assisted Selection in Crop plants. *Molecular Breeding* 3: 87 - 103.

Louette, D., Smale, M. 2000. Farmers seed selection practices and maize variety characteristics in a traditional Mexican community. *Euphytica* 113: 25 – 41.

MacArthur, R. H., Wilson, E. O. 1963. Ann equilibrium theory of insular zoogeography. *Evolution* 17: 373 – 387.

Mace, E. S., Buhariwalla, H. K., Crouch, J. H. 2003. A high throughput DNA extraction protocol for molecular breeding programs. *Plant molecular Biology* 21: 459a-459h.

Maquet, A., Zoro, B. I., Delvaux, M., Wathelet, B., Baudoin, J. P. 1997. Genetic structure of a lima bean base collection using alloenzyme markers. *Theoretical Applied Genetics* 95: 980 - 991.

Mariac, C., Luong, V., Kapran, I., Mamodou, A., Segnard, F., Deu, M., Chantereau, J., Gerad, B., Ndjeunga, J., Bezancon, G., Pham, J. L., Vigouroux, Y. 2006. Diversity of wild and cultivated Pearl millet accessions [(pennisetum glaucum (L.) R. Br.)] In Niger assessed by microsatellite markers. *Theoretical Applied Genetics* 114: 49 – 58.

Matsouka Y., Mitchell S. E., Kresovichs, Goodman M., Doebly J. 2002. Microsatellites in zea; variability, patterns of mutations and use for evolutionary studies. *Theoretical Applied Genetics* 104:436-480.

Moll, R. H., Stuber, C. W. 1974. Quantitative genetics- empirical results relevant to breeding. In: *Advanced Agronnomy Academic Press*, San Francisco, pp: 290 - 313.

Morden, C. W., Doebley, J., Schert, K. F. 1989. Allozyme variation in old world races of sorghum bicolor (Poaceae). *American Journal of Botany* 76: 247 - 255.

Morgante, C. W., Olivieri, A. M. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* 3: 175 - 182.

Negassa, M. 1985. Patterns of phenotypic diversity in the Ethiopian barley collection and the Arsi-bale highland as a centre of origin of barley. *Hereditas* 102: 139 - 150.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University press, New York.

Nei M. 1978. Estimatin of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583 - 590.

Oelke, E. A., Oplinger, E. S., Putnam, D. H., Durgan, B. R.., Doll, J. D., Undersander, D. J. 1990. Millets. *In Alternative Field Crops Manual.*

Ouendeba, B., Ejeta, G., Nyquist, W. E., Hanna, W. W., Kumar, A. 1993. Heterosis and combining ability among African pearl millet landraces. *Crop Science* 33: 735 - 739.

Oumar, I., Farhat, C., Sawsen, A. O., Mohamed, B., Yannis, K., Aly, R. 2005. Allozyme variation among some Pearl millet (*Pennisetum glaucum* L.) cultivars collected from Tunisia and West Africa. *Genetic Resources and Crop Evolution* 52: 1087 – 1097.

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

Pilat-Andre, J. S., Paris. S., Lamy, F., Sarr, A. 1992. Etude de la region Adh du mil (pennisetum typhoides) par utlisation des techniques RFLP. In: mounolou. J.C. and Sarr, A. (Eds), Comlexes d'especes, flux de genes et resources genetiques des plantes. Colloque International , 8-10 Janvier, 1992,pp.58-71. Publications du Bureau des Resource.

Plaschke, J., Ganal, M. W., Roder, M. S. 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theoretical Applied Genetics* 91: 1001 – 1007.

Poncet, V., Lamy, F., Enjalbert, J., Jolytt-Sarr, A, Robert, T. 1998. Genetic analysis of the domestication syndrome in pearl millet. *Heredity* 81 (6): 648 - 659.

Powell, W., Margante, M., Andre, C., Hanafers, M., Vogels, S., Tingey, S., Rafalski A. 1996b. The comparison of RFLP, RAPD AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2: 225 – 238.

Qi. X., Lindup, S., Pittaway, T. S., Allouis, S., Gale, M.D., Devos, K. M. 2001. Development of simple Sequence repeat markers from bacterial artificial chromosomes without subcloning Biotechniques 31:355-362.

Qi, X., Pitaway, T. S., Lindup, S., Liu, H., Waterman, E., Padi, F.K., Hash, C. T., Zhu, J., Gale, M. D., Devos, K.M. 2004. An integrated genetic map and a new set of simple sequence repeat markers for pearl,, *pennisetum* glaučum. *Theoretical Aplied genetics* 109:1485-1493

Rai, K. N. K., AnandKumar, D. J., Andrews, S. C., Gupta, B., Ouendeba, 1997: Breeding pearl millet for grain yield and stability.Proceedings of International Conference on Genetic Improvement of Sorghum and pearl millet, 22–27 September. 1996, Lubbock,71 83.

Rai, K. N. Murty, D. S., Andrews, D. J., Bramel-Cox P.J. 1999. Genetic enhancement of pearl millet and sorghum for the Semi-Arid tropics o f Asia and Africa.

Rogers J. S. 1972. Measuers of genetic similarity and genetic distance. Student genetics VII university textbook Publication 7213:145-153.

Saghai-Maroof, M. A., Biyashev, R. M., Yang, G. P., Zhang, Q., Allard R. W. 1994. Extraordinarily polymorphic microsatellites DNA in barley: Species diversity, chromosomal locations and population dynamics. *Proceedings Natural Academy Science* (USA) 91: 5466 - 5470.

Sambrook, J., Fritsch, E. F., Maniatis, T. 1989. Molecular Cloning: a laboratory manual, 2nd edn. Cold spring Harbor laboratory, Cold Spring Harbor, New York.

Scott, I., Paglia, G., Magri, F., Morgante, M. 1999. Microsatellite markers as a tool for detection of intra-and inter-population genetic structures. In: Gillet Em (ed) molecular tools for biodervisity which DNA marker for which purpose?
(http:// webdoc.sub.gwdg.de/ebook/y/1999/whidimarker/index.html)

Semagn, K., Bjornstad, A., Ndjiondjop, M. N. 2006. Progress and prospects of marker assisted backcrossing as a tool in crop breeding programs. *African Journal of Biotechnology* 5 (25): 2588 – 2603.

Senior, M. L., Murphy, J. P., Goodman, M. .M., Stubber, C. W. 1998. Utility of SSR for determining genetic similarities and relationships in maize using an a garose gel system *Crop Science* 38: 1088 – 1098.

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

Singh, S. D., Wilson, J. P., Navi, S. S., Talukdar, B. S., Hess, D. E., Reddy, K. N. 1997. Screening Techniques and Sources of Resistance to Downy Mildew and Rust in pearl millet. *Information Buletin No. 48.*ICRISAT, Patancheru

Smith, J. S. C., Kresovich, S., Hopkins, M. S., Mitchell, S. E., Dean, R. F., Woodman 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. R., Carpten, J. D., Brownstein, M. J., Ghost, S., Magnuson, V. L., Gilbert D. A., Trent, J. M., Collins, F. S. 1995, Approach to genotyping errors caused by non-templated nucleotide addition by Taq DNA polymerase. *Genome Research* 5: 312 - 317.

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

Smulders, M. J. M., Bredemeijer, G., Rus-Kortekaas, W., Arens, P., Vosman, B. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical Applied Genetics* 97: 264 – 272.

Sneath, P. H. A., Sokal, R.,R. 1973. Numerical Taxonomy. The principles and Practice of Numerical Classification, XV. Freeman, San Francisco. CA. pp573.

Sperling, L., Scheidegger, V., Buruchara, R. 1996. Designing seed systems with small farmers: Principles derived from bean research in the great lakes region of Africa. Agricultural research and extention network paper no. 60. *Overseas development Agency* London.

Teshome, A., Baum, B. R., Fahring. 1997. [(sorghum bicolor) (L.) moench] Landrace variation and classification in North Shewa and South Welo, Ethiopia. *Euphytica* 97: 255 – 263.

Tohme, J., Gonzalez, D. O., Beebe, S., Duque, M. C. 1996. AFLP analysis of gene pools of wild bean core collection. *Crop Science* 36: 1375 - 1384.

Tolbert, D. M., Qualset, C. O., Jain, S. K., Craddock, J. C. 1979. A diversity analysis of a world collection of barley. *Crop Science* 19: 789 - 794.

Tostain, S. 1994a. Isozymic classification of pearl millet (pennisetum glaucum, poaceae) landraces from Niger(west Africa).*Plant System Evolution*. 193:81-93

Tsegaye S., Tesemma T., Belay G. 1996. Rlationships among tetraploid wheat (Triticum turgidum L.) landrace populations revealed by isozyme markers and agronomic traits. *Theoretical applied genetics* 93:600-603.

Virk, D. S. 1988: Biometrical analysis in pearl millet - a review. *Crop Improvement* 15: 1-29.

Vom Brocke, K., Christinks, A., Weltzien, E R., Preterl, T., Geiger, H. H. 2003. Farmers' seed systems and management practices determined pearl millet genetic diversity patterns in semi-arid regions of India. *Crop Science* 43: 1680 – 1689.

Weber, J. L. 1990. Informativeness of human (dC-dA)n. (dG-dT)n polymorphism. *Genomics* 7: 524 – 530.

Weiring, K., Winter, P., Huttel, B., Kahl, G. 1998. Microsatellite markers for molecular breeding. *Journal of Crop production* 1: 113-143.

Weir, B.S. 1996. Genetic data analysis. II. 2nd ed. Sinauer Associates, Inc, Sunderland, MA.

Westman, A. L., Kresovitch, S. 1997. Use of molecular marker techniques for description of plant genetic variation. In: Callow J. A, Ford-Lloyd B. V., Newbury H. J. (eds) plant genetic resources. Conservation and use. CAB Int, Wallingford, UK, pp 9 - 48.

Wenzel, P., Caig, V., Carling, J., Evers, C. C. M., Jaccoud, D., Peng, K., Patarapuwandol, S., Uszynski, G., Xia, L., Yang, S., Huttner, E., Kilian, A. 2004. Diversity Array technology, a novel tool for harnessing crop genetic diversity .

Wenzel, P., J. Carling, D., Kudrna, D., Jaccoud, E., Huttner, A., Kleinhofs, A. Kilian. 2004. Diversity Arrays Technology (DArT) for whole genome-profiling of barley. *Proceeding Nataral Academy Science U S A 101*(26): 9915 - 9920.

Wilkes, G. 1983. Current status of crop plant germplasm. *CRC Critical reviews of plant science* 1: 133 - 181.

Wikipedia contributors, 'Millet', Wikipedia, The Free Encyclopedia, 3rd April 2006, 15: 43 UTC, [accessed 3 April 2006].

Wittenberg, A. H. J., Kilian, A., Schouten, H. J. 2004. DArT, a promising genetic

fingerprinting technology. *Plant and Animal Genomes XII Conference*, San Diego, 10-14 January 2004 (www.intl-pag.org/12/abstracts/W36_PAG12_168.html).

Xiao, J., Li, J., Yuan, L., McCouch, S. R., Tanksley, S. D. 1996. Genetic diversity and its relationships to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theoretical Applied genetics* 92; 637 – 643.

Zhu, J.,. Gale M.D., Quarrie S., Jackson M.T., Bryan G. J. 1998. AFLP markers for the study of rice biodiversity. *Theoretical Applied Genetics* 96: 602 – 611.

6.0 APPENDICES

Appendix 1. Percentage polymorphism, average heterozygosity and gene diversity of 30 bulked accessions

Marker	Major.Allele.Frquency	GenotypeNo	AlleleNo	GeneDiversity	Heterozygosity	PIC
ICMP3002	0.5333	3.0000	2.0000	0.4978	0.8000	0.3739
PSMP2043	0.3000	12.0000	8.0000	0.8000	0.9000	0.7721
PSMP2001	0.8167	6.0000	5.0000	0.3239	0.3333	0.3116
PSMP2030	0.4667	12.0000	9.0000	0.7022	0.8667	0.6650
PSMP2248	0.4667	8.0000	6.0000	0.6283	0.9000	0.5592
PSMP2201	0.4464	6.0000	4.0000	0.6792	0.7500	0.6229
PSMP2246	0.5000	6.0000	3.0000	0.5578	0.8333	0.4600
PSMP2008	0.2833	16.0000	8.0000	0.8089	0.8333	0.7843
PSMP2231	0.2333	22.0000	11.0000	0.8583	0.9333	0.8438
PSMP2090	0.4667	8.0000	6.0000	0.6739	0.9333	0.6232
PSMP2208	0.6500	4.0000	4.0000	0.5361	0.7000	0.5000
PSMP2267	0.6852	3.0000	3.0000	0.4520	0.6296	0.3780
PSMP2085	0.4828	5.0000	4.0000	0.6379	0.9310	0.5704
PSMP2237	0.4000	7.0000	5.0000	0.6761	0.7333	0.6134
PSMP2249	0.5000	4.0000	3.0000	0.6161	0.9667	0.5424
PSMP2080	0.2667	12.0000	6.0000	0.8028	0.9333	0.7732
PSMP2233	0.3833	11.0000	6.0000	0.7156	0.8333	0.6689
PSMP2076	0.5167	9.0000	7.0000	0.6156	0.8333	0.5514
PSMP2087	0.4167	9.0000	5.0000	0.6789	0.9333	0.6213
PSMP2027	0.2500	15.0000	9.0000	0.8244	0.9333	0.8013
PSMP2275	0.5667	3.0000	2.0000	0.4911	0.6000	0.3705
Mean	0.4586	8.6190	5.5238	0.6465	0.8148	0.5908

Appendix 2. Buffers and reagents

3 M Nacl
2M NH_4OAC
70% ETOH
500 mM Tris-HCL
500 mM EDTA
1 M Tris-HCL
10 M NaOH
3 M Sodium Acetate
TE Buffer (pH 8)
Chloroform:Isoamylalcohol 24:1
Phenol:chloroform:Isoamylalcohol 25:24;1

KENYATTA U...VE... ..TY LIBRARY