ASSESSMENT OF GENETIC VARIATIONS WITHIN AND BETWEEN POPULATIONS OF MARAMA BEAN [*TYLOSEMA ESCULENTUM* (BURCHELL) SCHREIBER] BASED ON MICROSATELLITES (SSRS) AND INTERGENIC SPACER LENGTH VARIATION MARKERS IN THE NAMIBIAN GERMPLASM

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

OF

THE UNIVERSITY OF NAMBIA

BY

EMMANUEL NEPOLO

JANUARY 2010

Supervisor: Dr. P.M. Chimwamurombe (Department of Biological Sciences, University of Namibia)

Co-supervisor: Prof A. Cullis (Department of Biology, CASE Western Reserve University)

ABSTRACT

Tylosema esculentum (Burchell) L. Schreiber is a herbaceous perennial drought tolerant legume that is endemic to the arid and semi-arid zones of southern Africa. Its agronomic potential is on the basis of its seed's nutritional composition that is comparable with soybean and peanut, the two mostly consumed legumes. In this study SSR markers were used to assess genetic diversity in natural populations of marama in the Namibian germplasm. Additionally, in the present study, the intergenic spacers (IGS) region of the large ribosomal DNA gene was amplified to assess length variation in marama bean. In the present study two to four alleles per primer pair were amplified from the 332 marama individuals among 11 populations. The study revealed that substantial genetic variation exists in marama populations and most of this variation (66.5%) occurred within populations. The direct-count heterozygosities showed a substantial variation of 0.51 among populations. The mean Shannon diversity index for eleven marama populations assessed in this study estimated a higher genetic diversity in marama with a mean total diversity of ($H^{'}$ = 1.526). There were no significant differences ($p > 0.05$) in the mean genetic diversity between populations. However, there were significant differences in the mean genetic diversity attained by different SSR primer pairs ($p < 0.05$). Assessment of length variation in marama bean has disclosed a high intraspecific genetic variation (H'= 1.700) in the nuclear ribosomal DNA repeat units in marama bean. There was no clear indication of population differentiation between the marama populations. This study is the first to document genetic variation present in marama ecotypes in the Namibian germplasm using SSR markers. The evaluation of genetic variation for marama is important for identification of the best strains, for the selection of the genetically diverge parent plants for marama genetic mapping and domestication. This study confirmed the usefulness of SSR in providing a quick and reliable method of assessing genetic variation in plant populations.

Keywords: microsatellite, SSR, marama bean, *Tylosema esculentum,* genetic diversity, rDNA, IGS, length variation

TABLE OF CONTENTS

ABBREVIATIONS

SE -Standard Error

IGS-Intergenic Spacers Regions

rDNA- ribosomal deoxyribonucleic acid

DNA- deoxyribonucleic acid

SSRs- Simple Sequence Repeats

PCR- Polymerase Chain Reaction

RFLP - Restriction Fragment Length Polymorphism

VNTRs -Variable Number of Tandem Repeats

SNPs -Single Nucleotide Polymorphisms

RAPD - Random Amplified Polymorphic DNA

EST -Expressed Sequence Tags

SSLPs -Simple Sequence Length Polymorphisms

STRs- Short Tandem Repeats

ITS-Internal Transcribed Spacers

GPS-Global Positioning System receiver

µl - microlitre

ng - Nanogram

µM- micromole

rpm -revolutions per minute

ml- millilitre

bp - basepair

STRP- Simple tandem repeat polymorphism

°C- Degree Celsius

UPGMA- Unweighted pair Group method using arithmetic Averages

TBE- Tris borate-EDTA buffer

EDTA- Ethylenediaminetetra acetic acid

FIGURES

TABLES

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge the Kirkhouse Trust, United Kingdom (to PC) for funding my MSc. program.

I am grateful to my principle supervisor Dr. P. M. Chimwamurombe who always had faith in me and was a shining beacon of inspiration throughout my study. His astonishing mind is matched only by his genuine concerns for his student's welfare and success. I would also like to acknowledge and express my sincere gratitude to my fellow student Mutsa Takundwa for her magnificent support and contributions to my study and preparation of this thesis. Jean Damascene Uzabakiriho is highly acknowledged for his endless technical support in the laboratory and in the field. Professor I. Mapaure and K. Kunert are thanked for all the positive and creative guidance during my study, I am really grateful.

I am deeply grateful to my co-supervisor, Professor Chris Cullis and his family for their support and kindness while I lived in Cleveland. Cory Bickel of Case Reserve Western University is thanked for helping me in the laboratory. I am also grateful to Tiffannie Moss and her family for their generosity in accommodating me during my first stay in Cleveland, Ohio. I thank Linda Bowman for accommodating me during my whole stay in Cleveland and for making my stay a memorable one.

Elikaim is thanked for his unmatched creativity in designing a high-quality map for my thesis. My deepest appreciation goes to the University of Namibia, especially the department of Biological Sciences and to the department of Biochemistry for allowing me to use their equipment and facilities.

To my amazing family John, Onesmus, Aloisius and Theresia thank you for your unconditional love and supports that knows no limits and for always being by my side. My courageous mother, Heneriette Nepolo, is deeply thanked for her invaluable love, care and inspiration.

To my friends Evaristus, Petrina, Celine, Marius, Connie and Jay are acknowledged for their love and support and creative contributions to my study. I wish to thank every person who has come into my life and encouraged, touched and illuminated me through their presence. Ultimately, my sincere gratitude goes to my almighty heavenly God for bestowing in me courage, determination, persistence and guidance to successfully complete my study.

DEDICATION

This thesis is dedicated to those who teach molecular biology and agricultural biotechnology and whose efforts will inspire future generations to utilize under-

utilized crops and provide for human food needs. I also dedicate this work to my entire beloved family and in memory of my beloved father **Simeon. M. Nepolo** who was my beacon of inspiration throughout my life voyages and whose light and love continue to shine through my life.

.

DECLARATIONS

I, Emmanuel Nepolo, declare hereby that this study is a true reflection of my own research, and that this work or part thereof has not been submitted for a degree in any other institution of higher education.

…… Date……..

Emmanuel Nepolo

PUBLICATIONS AND CONFERENCE PROCEEDINGS

PUBLICATION IN JOURNALS

E. Nepolo, M. Takundwa, P. M. Chimwamurombe, C. A. Cullis and K. Kunert **(2009).** A review of geographical distribution of marama bean [*Tylosema esculentum* (Burchell) Schreiber] and genetic diversity in the Namibian germplasm, *African Journal of Biotechnology*, Vol, 8 (10), pp. 2088-2093.

Nepolo E, Chimwamurombe PM, and Cullis CA. Kandawa-Schulz MA.(**2010**) Determining genetic variation based on ribosomal intergenic spacer length variation in Marama bean [*Tylosema esculentum* (Burchell) L. Schreiber] from Omipanda area, Eastern Namibia. *African Journal of Plant Sciences,* Vol, 4 (9), pp 368 -373.

Takundwa M, **Nepolo E**, Chimwamurombe PM, Cullis CA and Kunert K: (**2010**) Development and use of microsatellites markers for genetic variation analysis, in the Namibian germplasm, both within and between natural populations of Marama bean (*Tylosema esculentum*). *Journal of Plant breeding and Crop Science*, Vol 2(8), pp 233- 242.

CONFERENCE PROCEEDINGS

Nepolo E. (**2010**). Simple Sequence Repeats (SSRs) variation in natural populations of marama bean [*Tylosema esculentum* (Burchell) Schreiber] in Namibia. **Power point presentation** presented at the 1st National Research Symposium of the Ministry of Education, Directorate of Research, Science and Technology, from 15th - 17th September 2010, at Safari Court Hotel and Conference Centre, Windhoek.

Nepolo E, Chimwamurombe PM and Cullis CA. (**2009**). Nuclear 5S ribosomal DNA diversity of Marama bean [*Tylosema esculentum* (Burchell) L. Schreiber] in Namibian germplasm- **Power point presentation** presented at the 13th Congress of the Agricultural Scientific Society of Namibia (AGRISSON) from 1st -2nd July 2009, at Oshakati Country Hotel.

Nepolo E, Chimwamurombe PM and Cullis CA. (**2009**).Genetic diversity in Marama bean [*Tylosema esculentum* (Burchell) L. Schreiber] in Namibian germplasm revealed by variation in ribosomal intergenic spacer regions- **Poster presentation** presented at the 2nd ***Tropical Crop Biotechnology Conference*** from 22nd -25th July 2009,Hazyview,Mpumalanga,South Africa. **This poster was awarded the 1st prize as the best poster presented in the presentation category.**

CHAPTER 1: INTRODUCTION

1.1 General Introduction

Biodiversity can be defined as the total variation found within all living organisms and their habitats (Carvalho, 2004, p. 1). According to Carvalho (2004, p. 1), biodiversity can be assessed at three different levels: communities (environment), species, and genes. Community diversity can be defined as the different biological communities and their associations with the physical environment (Primack, 2004, p. 11). According to Primack (2004, p. 11) species diversity is all the species on earth, including bacteria and protists as well as the species of the multi-cellular kingdoms (plants, fungi and animals), while genetic diversity is the genetic variation within species, both among geographically separate populations and among individuals within single populations. When biodiversity is assessed at the species level, emphasis is on observing differences among individuals or populations of that particular species. This can be referred to as the genetic diversity of the species. Thus, genetic diversity can be considered as a form of biodiversity. Genetic diversity is associated with the degree of differentiation among the genomes of individuals in a population. The genetic material corresponds to the DNA, genic or cytoplasmic, called the genotype. Expression of genes contained in the DNA is the result of interaction between the environment and the genotype, and the resulting form of the individual is called the phenotype. Genetic diversity is important because it is the raw material that will enable evolution and adaptation of the species to an ever changing environment, making it essential for the long-term survival of a species.

Therefore, studies of genetic diversity are important because they are a tool for genetic identification of the possibilities for improvement and facilitating the efficient use of the available germplasm of a species. Since early in the history of the world, humans have exploited the genetic diversity of plants, primarily as sources of food, and to improve their landraces and cultivars (Callow Ford-Lloyd, and Newbury, 1997, p. 102). A quarter of century ago, the world was preoccupied with an impending food crisis. Birth rates and life expectancies were increasing, especially in the developing world. The situation was particularly menacing in more populous regions where mass starvation was predicted on an unprecedented scale (Callow et al., 1997, p. 202). The expanding population has taken an additional and serious toll on the environment, because soils have been eroded, forests have been depleted, and biodiversity has been lost.

In response to the crisis, a few far-sighted scientists set out to tackle the increasing problem of food production (Callow et al., 1997, p. 212). The solution to tackle food shortages is through the application of genetics, to develop new varieties of the world's most important food crops that would be both high yielding and more responsive to inputs such as fertilizers and irrigation (Karp, Seberg and Buiatti, 1996, p. 143). Therefore, people living in harsh regions or environments that are marginal for agriculture have to improve and increase their awareness of under-utilized crops which thrive well in these extreme environments to address or compliment food security.

While molecular genetics will not feed people, it can help plant breeders to improve crop varieties that could be used by farmers. New genetic methods are greatly accelerating the pace of improvements to food crops and for the introduction of new crop species to satisfy the growing need for improvements to food crops and introduction of new species to meet the demands of an ever increasing world population and a desire for improved standards of living (Callow et al., 1997, p. 213).

These techniques offer renewed hope for solving the intricate problems that lie at the heart of today's concerns for poverty, nutrition and the environment (Callow et al., 1997, p. 213). They allow humans to better understand the composition and functioning of genomes, and to transfer useful genes among widely different group of organisms.

By such means, humans are learning how to more accurately tailor food crops to meet new pest and disease challenges and produce crops that can better withstand the rigour of stressful environments (Stephen, 2002, p. 8). Moreover, through these techniques, humans are learning how to conserve the biological resources that will underpin their ability to continue to meet new food challenges in the future, allow for a better understanding of the nature and distribution of genetic diversity, which provide for long-term conservation of biological diversity (Stephen, 2002, p. 12).

However, while enormous resources have been expended in recent decades on grasses like wheat, corn, sorghum and barley, among the legumes only soybean and peanuts (groundnuts) have received much attention. Of all plants used for food in agriculture only the grasses are more important than the legumes (USA National Academy of Science, 1979, p. 1). Of the thousands of known legume species, less than 20 are used extensively today and those that are commonly used include peanuts, soybean, peas, alfalfa (Lucerne), sweet clover and vetches (USA National Academy of Science, 1979, p. 1). The remaining species such as *Tylosema esculentum* (marama bean) are seldom used and many of them are almost unknown to science (van der Maesen, 2006).

Marama bean a perennial-creeper species belongs to the Fabaceae, sub family Caesalpinioideae and it is considered as the most promising of the wild species of the genus *Tylosema*. The primary agronomic potential is based upon the high nutritional value of the seeds. The protein content is slightly higher than that of soybean and the oil content is twice that of soybean and approaches that of peanuts. The plant has been described as a legume crop of considerable potential in arid and semi-arid land agriculture (Powell, 1987, p. 217). Marama bean could help in alleviating hunger and provide food security in Southern Africa. Thus there is an increasing interest on its cultivation, due to its potential as a cash crop and food source especially in the face of climate changes.

CHAPTER 2: LITERATURE REVIEW

2.1 Background Information

Understanding the molecular basis of the essential biological phenomena in plants is crucial for the effective conversation, management, and efficient utilization of plant genetic resources (Mondini, Noorani and Pagnotta, 2009, p. 20). In particular, an adequate knowledge of existing genetic diversity, where in plant population it is found and how to best utilize it, is of fundamental interest for basic science and applied aspects like plant domestication and the efficient management of crop genetic resources. The improvement of crop genetic resources and domestication is dependent on continuous infusions of wild relatives, traditional varieties and the use of modern breeding techniques (Mondini et al., 2009, p. 20).

The assessment of genetic diversity within and between populations is routinely performed at the molecular level using different laboratory-based techniques such as allozyme (Agarwal, Shrivastava and Padh, 2008, p. 620) or DNA analysis, which measures levels of variation directly. According to (Mondini et al., 2009, p. 20), genetic diversity may also be gauged using morphological, and biochemical characterization and evaluation:

i. **Morphological** characterization does not require expensive technology but large tracts of land are often required for these experiments, making it possibly more expensive than molecular assessment.

These traits are often susceptible to phenotypic plasticity; because it clouds the level of diversity as it's difficult to partition genetic and environmental effects.

ii. **Biochemical** analysis can be based on the separation of proteins into specific banding patterns. It is a fast method which requires only small amounts of biological material and it is subjected to differences caused by various tissue types, developmental variation and environmental effects. However, only a limited number of enzymes are available and thus, the resolution of diversity is limited.

iii. **Molecular** analyses comprises of a large variety of DNA molecular markers, (and RNA or expression markers) which can be employed for analysis of variation. Markers can be dominant, recessive or co-dominant and can be derived from different regions of the genome such as the anonymous or characterized loci and can contain expressed or non-expressed sequences.

2.2 Molecular assessments of genetic diversity

Analyses of genetic diversity are usually based on assessing the diversity of many individuals using either allozymes (i.e. variant forms of an enzyme that are coded for by different alleles at the same locus) or molecular markers, which tend to be selectively neutral (Agarwal et al., 2008, p. 622). Most of molecular markers, if not all of them, offer numerous advantages over conventional, phenotype-based alternatives as they are stable and detectable in all tissues regardless of growth, differentiation, development or defense status of the cell. Additionally, they are not confounded by environmental, pleiotropic and epistatic effects.

According to Primack (1993), genetic variability within a population can be assessed through:

(i) The number (and percentage) of polymorphic genes in the population.

(ii) The number of alleles for each polymorphic gene.

(iii) The proportion of heterozygous loci per individual.

Molecular methods, such as DNA analysis, directly measure genetic variation, giving a clear indication of the levels of genetic variation present in a species or population (Karp et al., 1996, p. 144) without direct interference from environmental factors. The different methods of molecular assessment differ from each other with respect to important features such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and cost (Mondini et al., 2009, p. 22).

2.3 Molecular markers

Characterization of diversity has previously mainly been based on morphological traits (Karp et al., 1996, p. 143). However, morphological variability is often restricted because characters may not be obvious at all stages of the plant development. Some traits may interfere with the expression of others and thus cannot be measured simultaneously and their appearance may be affected by the environment. The limitation of phenotype-based genetic markers led to the development of DNA-based markers, i.e. molecular markers.

Barcaccia, Alberini and Rosellini (2000, p. 529), defined a molecular marker as a genomic locus, detected through probe or specific starters (primer) which by virtue of its presence, distinguishes unequivocally the chromosomic trait which is represented as well as the flanking region at the 3' and 5' extremities.

Ford-Lloyd (2001, p. 59), classified DNA-based markers into three different categories:

1. **Non-PCR-based methods**: Restriction fragment length polymorphism analysis (RFLP) is an example of this category.
2. **Arbitrary (or semi-arbitrary) primed techniques**: These are PCR-based category that uses random primers during the PCR reaction. The most well known and widely used of these methods is the "Random Amplified Polymorphic DNA" (RAPD) and Amplified Fragment Length Polymorphism (AFLPs).
3. **Site-targeted PCR**: In this class, primers that amplify specific regions of the DNA are used during the PCR reaction. Examples are single nucleotide polymorphism (SNPs), microsatellite repeats (Variable Number Tandem Repeats), and many others.

Molecular markers work by highlighting differences (polymorphisms) within a nucleic sequence between different individuals. These differences include insertions, deletions, translocations, duplications and point mutations (Mondini et al., 2009, p. 22). They do not, however, necessarily encompass the activity of specific genes.

In addition to being relatively unaffected by environmental factors, molecular markers have the advantage of: (i) being applicable to any part of the genome (introns, exons and regulation regions); (ii) not possessing pleiotrophic or epistatic effects; (iii) being able to distinguish polymorphisms which do not produce phenotypic variation and finally, (iv) many are co-dominant (Agarwal et al., 2008, p. 619).

Different techniques can be employed to detect differences either by restriction-hybridization of nucleic acids or by the polymerase chain reaction (PCR) or both. In addition, the different techniques can assess either multi-locus or single-locus markers. According to Agarwal et al. (2008, p. 625), these types of markers are also defined as dominant since it is possible to observe the presence or the absence of a band for any locus, but it is not possible to distinguish between heterozygote conditions and homozygote for the same allele. By contrast, single-locus markers employ probes or primers specific to genomic loci, and are able to hybridize or amplify chromosome loci with well-known sequences. They are defined as co-dominant since they allow discrimination between homozygote and heterozygote loci.

2.4 Molecular markers based on amplification techniques (PCR-derived)

The use of this kind of marker has been exponential following the development of PCR by Mullis, Faloona, Scharf, Saiki, Horn and Erlich (1986, p. 264). This technique consists of the amplification of one or more discrete DNA products, deriving from regions of DNA which are flanked by regions of high homology with

primers. The use of random primers overcame the limitation of prior sequence knowledge for PCR analysis and being applicable to all organisms, facilitated the development of genetic markers for a variety of purposes (Williams, Kubelik, Livak, Rafalski and Tingey, 1991, p. 6533). PCR based marker techniques such as restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), simple sequence repeats (SSRs) and amplified fragment length polymorphism (AFLP) are now in common use for ecological, evolutionary, taxonomical, phylogenic and genetic studies in the plant sciences (Ayad, Hodking, Jaradat and Rao, 1995).

Table 1: Comparison of different characteristics of most frequently used molecular markers techniques.

Molecular Marker	**RFLP**	**RAPD**	**AFLP**	**SSR**	**SNP**
Degree of polymorphism	Medium	Medium	Medium	Medium	High
Locus specificity	Yes	No	No	No	Yes
Dominance / (D)Co-dominance (C)	Co-dominance	Dominance	Dominance	Co-dominance	Co-dominance
Ease of replication	High	Low	High	Medium	High
Abundance	High	High	High	Medium	High
PCR-based	No	Yes	Yes	Yes	Yes
Quantity of DNA required	High	Low	Medium	Low	Low
Automation	No	Yes	Yes	Yes	Yes
Costs per assay	High	Low	Medium	Low/Medium	Low
Technical requirement	High	Low	Medium	Low/Medium	Medium

Adapted from Mondini et al. (2009, p. 24).

2.4.1 Restriction Fragment Length Polymorphisms (RFLPs)

These are markers detected by treating DNA with restriction enzymes (enzymes that cut DNA at a specific sequence) (Dreher et al., 2000) and need more steps before they can be detected. For example, the *Eco*R1 restriction enzyme cuts DNA whenever the base sequence GAATTC is found. Differences in the lengths of DNA fragments will then be seen if, for example, the DNA of one individual contains that sequence at a specific part of the genome (e.g. tip of chromosome 3) whereas another individual has the sequence GAATTTC (which is not cut by *Eco*R1). RFLPs were the first molecular markers to be widely used, their use is, however, time-consuming and expensive and simpler marker systems have subsequently been developed (Dreher et al., 2000).

2.4.2 Random Amplified Polymorphic DNA markers (RAPDs)

RAPDs were the first PCR-based molecular markers to be employed in genetic variation analyses (Welsh and McClelland, 1990, p. 7214). RAPD markers are generated through the random amplification of genomic DNA using short primers (decamers), separation of the obtained fragments on agarose gel in the presence of ethidium bromide and finally, visualization under ultraviolet light. The use of short primers is necessary to increase the probability that, although the sequences are random, they are able to find homologous sequences suitable for annealing. DNA polymorphisms are then produced by “rearrangements or deletions at or between oligonucleotide primer binding sites in the genome” (Williams et al., 1991, p. 6532).

As this approach requires no prior knowledge of the genome analyzed, it can be employed across species using universal primers. The analysis for RAPD markers is quick and simple, but the major drawback of this method is that the profiling is dependent on reaction conditions which can vary between laboratories, even a difference of a degree in temperature is sufficient to produce different patterns (Williams et al., 1991, p. 6532). Additionally, as several discrete loci are amplified by each primer, profiles are not able to distinguish heterozygous from homozygous individuals (Bardakci, 2001, p. 187).

2.4.3 Amplified Fragment Length Polymorphism (AFLPs)

This method is based on the combination of the main analysis techniques: digestion of DNA through restriction endonuclease enzymes and PCR technology (Vos et al., 1995, p. 4411). It can be considered an intermediate between RFLPs and RAPDs methodologies as it combines the power of RFLP with the flexibility of PCR-based technology. The primer pairs used for AFLP usually produce 50–100 bands per assay. Althoff, Gitzendanner and Segraves (2007, p. 480), stated that the number of amplicons per AFLP assay is a function of the number selective nucleotides in the AFLP primer combination, the selective nucleotide motif, GC content, and physical genome size and complexity. AFLP generates fingerprints of any DNA regardless of its source, and without any prior knowledge of DNA sequence. Most AFLP fragments correspond to unique positions on the genome and hence can be exploited as landmarks in genetic and physical mapping.

The technique can be used to distinguish closely related individuals at the sub-species level (Althoff et al., 2007, p. 478) and can also map genes. According to Zabeau and Vos (1992), the origins of AFLP polymorphisms are multiple and can be due to: (i) mutations of the restriction site which create or delete a restriction site; (ii) mutations of sequences flanking the restriction site, and complementary to the extension of the selective primers, enabling possible primer annealing; (iii) insertions, duplications or deletions inside amplification fragments. These mutations can cause the appearance or disappearance of a fragment or the modification (increase or decrease) of an amplified-restricted fragment.

2.4.4 Single Nucleotide Polymorphisms (SNPs)

Single base changes in DNA sequence, have become an increasingly important class of molecular markers. Single nucleotide variations in genome sequence of individuals of a population are known as SNPs (Mondini et al., 2009, p. 27). The potential number of SNP markers is very high, meaning that it should be possible to find them in all parts of the genome. Micro-array procedures have been developed for automatically scoring hundreds of SNP loci simultaneously at a low cost per sample (Korzun, Malyshev, Voylokov and Borner, 2001, p. 712). SNPs are the most abundant molecular markers in the genome and they are widely dispersed throughout genomes with a variable distribution among species. The SNPs are usually more prevalent in the non-coding regions of the genome (Sobrino, Briona and Carracedoa, 2005, p. 182).

Within the coding regions, when SNP is present, it can generate either non-synonymous mutations that result in an amino acid sequence change (Sunyaev et al., 1999, p. 757), or synonymous mutations that don't alter the amino acid sequence. Synonymous changes can, however, modify mRNA splicing, resulting in phenotypic differences (Richard and Beckman, 1995, p. 259). Improvements in sequencing technology and an increase in the availability of the increasing number of Expressed Sequence Tags (EST) sequences have made analysis of genetic variation possible directly at the DNA level. The majority of SNP genotyping analyses are based on allele-specific hybridization, oligonucleotide ligation, primer extension or invasive cleavage (Sobrino et al., 2005, 190).

Genotyping methods, including DNA chips, allele-specific PCR and primer extension approaches based on SNPs, are particularly attractive for their high data throughput and for their suitability for automation. They are used for a wide range of purposes, including rapid identification of crop cultivars and construction of ultra high-density genetic maps (Sunyaev et al., 1999, p. 756).

2.4.5 Microsatellites-based marker technique (SSRs)

Microsatellites, alternatively known as simple sequence repeats (SSRs), short tandem repeat (STRs) or simple sequence length polymorphisms (SSLPs) are tandem repeats of sequence units generally between 2 and 6 bp in length found within eukaryotic genomes (Bruford and Wayne, 1993, p. 941).

Polymorphisms associated with a specific locus are due to the variation in length of the microsatellite, which in turn depends on the number of repetitions of the basic motif. Pure SSRs consist of uninterrupted repeats of a single motif, e.g., (AT)*n*, compound SSRs consist of two or more repeat types, e.g., (GT)*n*(AT)*m*; and interrupted SSRs contain an interruption in the repeat, e.g., (GT)*n*GG(GT)*m* (Weber, 1990, p. 168). One common example of microsatellite is a $(CA)_n$ repeat (Figure 1),where n is the variable between alleles.

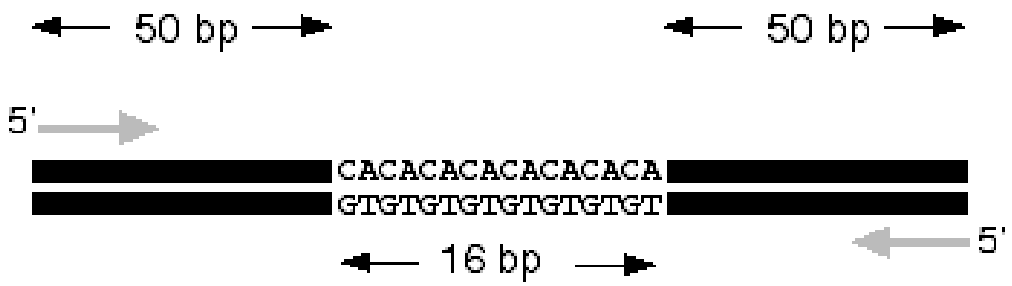

Figure 1: An illustration of a microsatellite (CA repeats) (Adapted from Weber, 1990).

The size of the repeat unit, the number of repeats, the presence of variant repeats and the frequency of transcription in the area of DNA repeat are the factors responsible for generating polymorphisms. Large changes in repeat number are thought to be the result of processes such as unequal crossing over (Strand, Prolla, Liskay and Petes, 1993, p. 274). Such differences are detected on either agarose or polyacrylamide gels (Figure 2), where amplified units migrate different distances according to their sizes (based on the number of repeating units).

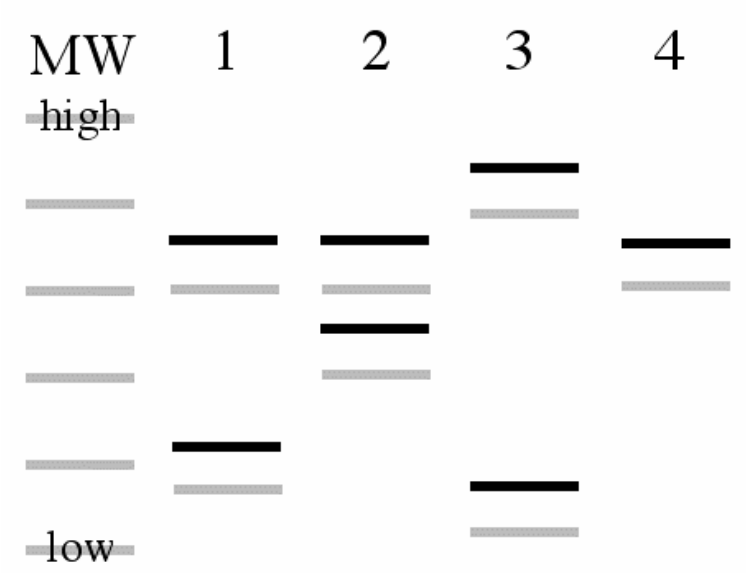

Figure 2: Polymorphism in microsatellite (Adapted from Strand et al., 1993)

Variations in the number of tandemly repeated units are mainly due to strand slippage during DNA replication where the repeats allow matching via excision or addition of repeats (Schlotterer and Tautz, 1992, p. 212). As slippage in replication is more likely than point mutations, microsatellite loci tend to be hypervariable. Microsatellite assays show extensive inter-individual length polymorphisms during PCR analysis of unique loci using discriminatory primer sets.

Microsatellites are highly popular genetic markers as they possess the following attributes: co-dominant inheritance, high abundance, enormous extent of allelic diversity, ease of assessing SSR size variation through PCR with pairs of flanking primers and high reproducibility (Schlotterer and Tautz, 1992, p. 211). However, the development of microsatellites requires extensive knowledge of DNA sequences, and sometimes they underestimate genetic structure measurements, hence they have been developed primarily for agricultural species, rather than wild species.

2.4.5.1 Application of microsatellites on plant genetic diversity studies

The applications of microsatellites spans over different areas ranging from ancient and forensic DNA studies, to population genetic conservation and management of biological resources (Jarne and Lagoda 1996, p. 426). Microsatellites have proven to be an extremely valuable tool for genome mapping in many organisms (Knapik, Goodman and Ekker, 1998, p. 338).

Plant microsatellite markers have been developed for germplasm conservation, cultivar identification, for assessing genetic diversity and have shown high levels of polymorphism in many important crops including rice (*Oryza sativa* L., Chen Temnykh, Cho and McCouch, 1997, p. 553), wheat (*Triticum aestivum* L., Roder, Plaschke and Konig, 1995, p. 329), barley (*Hordeum vulgare* L., Liu, Biyashev and Saghai-Maroof, 1996, p. 870), oat (*Avena sativa* L., Li, Rossnagel and Scoles, 2000), maize (*Zea mays* L., Senior, Murphy, Goodman and Stuber, 1998, p. 1088), sorghum (*Sorghum bicolor* (L.) Moench, Brown, Hopkins and Mitchell, 1996, p. 191), soybean (*Glycine max* (L.) Merr., Akkaya, Bhagwat and Cregan, 1992, p. 1132), beans (*Phaseolus* and *Vigna*, Yu, Park and Poysa, 1999, p. 28), *Brassica* species (Szewc-McFadden, Lamboy, Hokanson and McFerson, 1996, p. 619), alfalfa (*Medicago* spp., Diwan, Bhagwat, Bauchan and Cregan, 1997, p. 888), sun-flower (*Helianthus annuus* L.,Brunel, 1994, p. 397), White clover (*Trifolium repens* L., George, van Zijll de Jong, Cogan and Forester, 2006, p. 919) and in tomato (*Lycopersicon esculentum* Mill., Smulders, Bredemeijer, Rus-Kortekaas, Arens and Vosman, 1997, p. 264).

Dolanská and Čurn (2004, p. 99) evaluated and compared genetic variation among white clover cultivars (*Trifolium repens* L.) one of the most important and widely used legumes in temperate regions of the world using different DNA markers such as RAPD and SSRs. The study revealed that SSR markers did not show variability within one cultivar but between cultivars. The RAPD analysis in the above mentioned study, disclosed that the analysis was an unsatisfactory method due to high polymorphism within one population and RAPD profiling can characterize a genotype but not a population. Only by SSR analysis was it possible to identify different *T.repens* populations, even if this method expects to test a number of primer pairs and select one or several primer pairs that generate binding patterns representative for each population (Dolanska and Čurn (2004, p. 98).

Blair, Giraldo, Buendia, Towar, Dugue and Beebe (2006, p. 101) estimated allelic diversity of common bean (*Phoseolus vulgaris*. L), employing microsatellites markers, and revealed higher intra specific diversity (80%) than inter specific diversity within common bean. The study revealed higher polymorphic information content ranging from 0.927 to 0.937. Krishna et al. (2004, p. 686) used microsatellite markers to identified genetic diversity in Valencia type peanut (*Archis hypogaea* .L). Higher level of polymorphism ranging from 66.6% to 100% was observed among the genotypes screened.

Dao Li, Fotokun, Ub, Singh and Scoles (2001, p. 190) demonstrated the application of cowpea microsatellites for the differentiation and estimation of genetic relationships on cowpea breeding lines. The study showed that microsatellites markers were highly polymorphic in cowpea.

The level of polymorphism ranged from 0.02 to 0.73, and they managed to distinguish the cowpea's breeding line. Based on these findings, SSRs are likely to be useful for the analysis of genetic diversity of marama bean and other legumes.

2.5 The Genus Tylosema and *Tylosema esculentum*

The genus *Tylosema* (Schweinf.) Torre & Hillc. Leguminosae,Caesalpinioideae) comprises of four taxonomically accepted species [*T.esculentum* (Burch.) A.Schreib., *T. fassoglense* (Schweinf.) Torre & Hillc., *T. argenteum* (Chiov.) Brenan and *T. humifusum* (Pic.Serm. & Roti Mich.) Brenan] occurring in eastern, central tropical Africa (Castro, Silveira, Pereira-Coutinho and Figueiredo, 2005, p. 99) and Southern Africa (Figure 3). According to Castro et al. (2005, p. 99), this genus is characterized by the presence of only two fertile stamens, the remaining seven or eight stamens being sterile, variously shaped and coloured; with the presence of a lobed non-spathaceous calyx-limb.

Marama bean is presently known only in the wild state and it is endemic to Southern Africa, occurring in the northern areas of Namibia, Botswana, western and north-western Transvaal and Northern Cape where it is well adapted to the arid zones of Southern Africa (Castro et al., 2005, p. 101). This species produces a nutritious and palatable food, the marama bean. The bean which compares well in protein and oil content with both soya and peanut (Ketshajwag, Holmback and Yeboah, 1998, p. 741) is collected and eaten by local people.

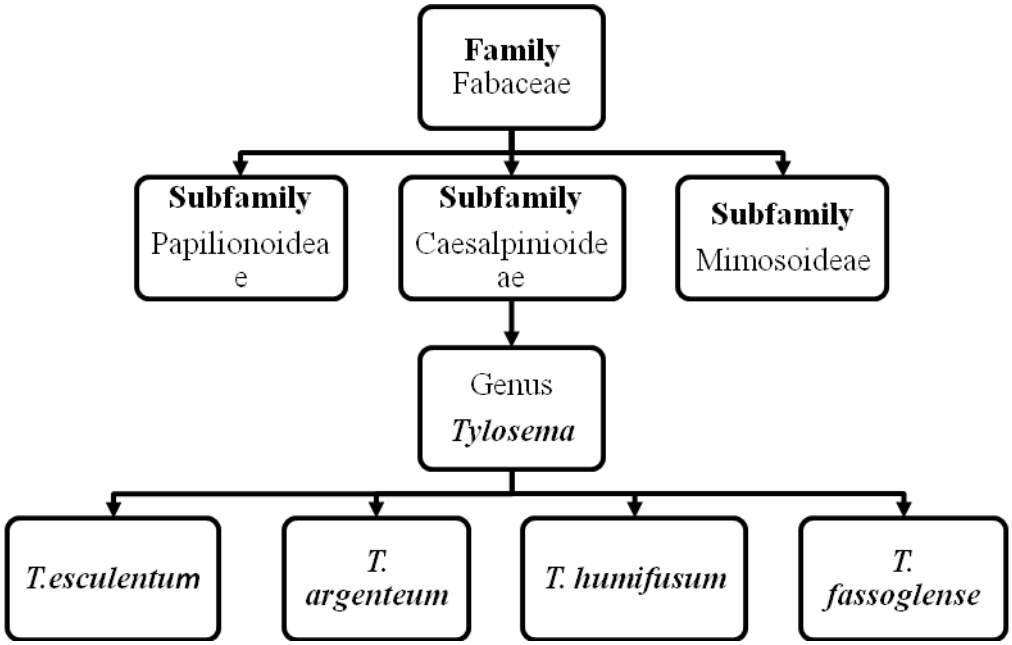

Figure 3: Taxonomic position of *T. esculentum* in the genus *Tylosema* (Adapted and modified from Hartley, 1997, p. 17).

2.5.1 Botanical characterization of Marama bean

The marama bean is an herbaceous perennial legume with creeping stems of up to six metres in length which spread from a woody tuberous root below the ground (National Academy of Sciences, 2006, p. 243). Tubers have a reddish-brown bark and usually taper to a thinner neck-like structure near the soil surface, from the annual branches growing during the rainy season. Axillary forked tendrils are found along the stems, but the plant is not really a creeper (vine) and the branches merely spread out along the ground surface in several directions. The vines carry double-lobed or two lobed leaves that are hairless with a firm texture. The leaves are soft and red brown when young, but turn leathery and gray-green with age.

Attractive bright yellow flowers are borne along the stems, each with erect petals and stamens, and are followed by fruits (Figure 4). Each fruit comprises of a broad woody pod containing between two and six large chest-nut brown seeds inside that weigh 2-3 grams (van der Maesen, 2006). The seed coat is about two millimetres thick and encloses the delicious white nut inside.

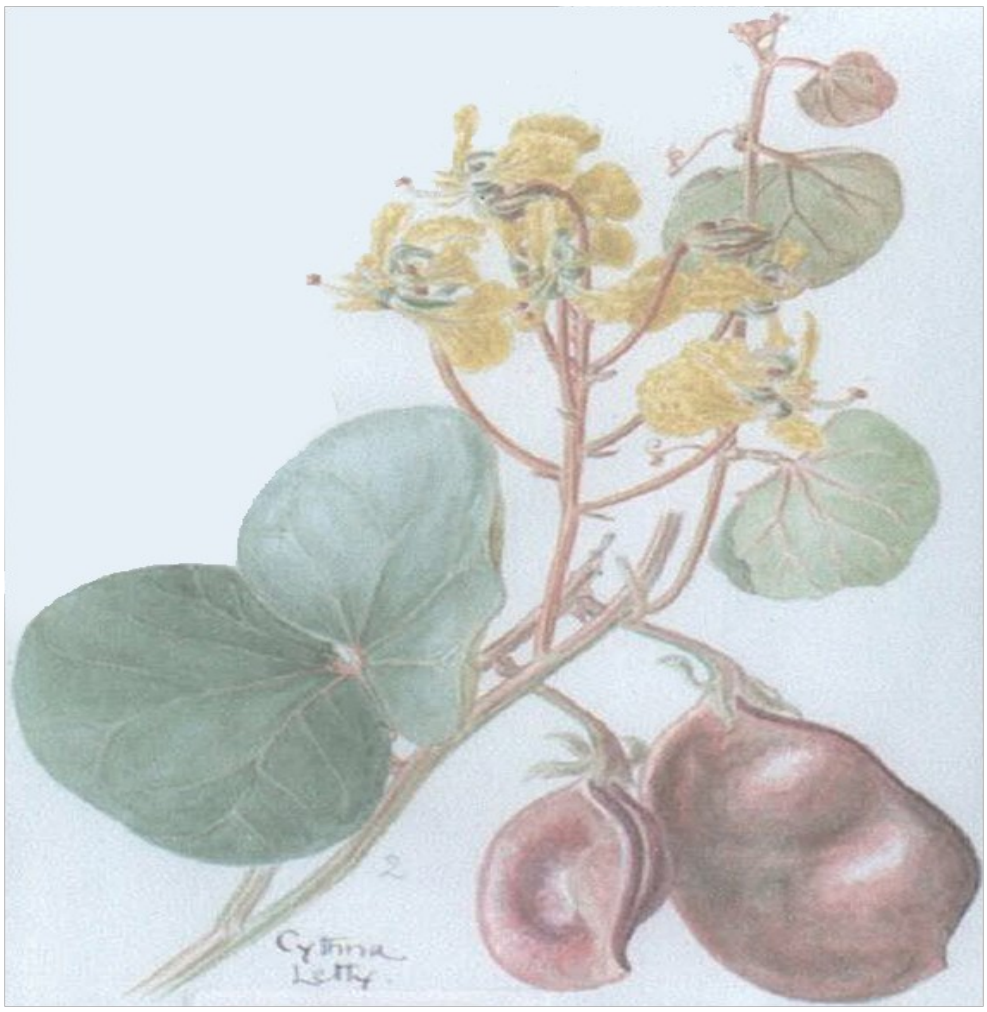

Figure 4: Bright yellow flowers, leaves and pods of Marama bean (Adapted from Flowering Plants of Africa: vol, 33, plate no 1131, Verdoon, 1959).

2.5.2 Agronomic characterization of Marama bean

The marama bean is a drought tolerant plant and survives in regions where few conventional crops survive, yet it appears adapted to a wide range of climatic conditions. Marama bean grows well in areas with low rainfall ranging from 50 to 500 millimetres annually, and long periods without precipitation (Powell, 1987, p. 217). Marama bean grows well in neutral to acidic soils, soils that are leached, infertile and low in phosphorus (Mmonatau, 2005, p. 8). It is adapted to poorly drained soils, with limited water-holding capacity and is frequently exposed to very high light and extreme temperature (Mitchell, Keys, Madgwick, Parry and Lawlor, 2005, p. 969). It has also been observed to grow well on deep sand, but also where there are outcrops of dolomite and neutral shaly soils (National Research Council, 2006, p. 244).

Soils which support the growth of marama bean are generally low in Nitrogen. The plant does not have nodules; therefore it is a non-nodulating legume. It does not obtain its Nitrogen from symbiotic fixation of atmospheric Nitrogen with soil rhizobia nor from the sources different from that utilized by associate plant species (Dakora, Lawlor and Sibiuga, 1999, p. 272). It efficiently obtains its Nitrogen from low concentrations and rapidly builds reserves in its large tuber and seeds to serve as a buffer from formation of protein rich organs. Therefore it depends on soil Nitrogen for its Nitrogen nutritional needs (Dakora et al., 1999, p. 272).

2.5.3 Domestication of Marama bean, an under-utilized species

The anthropogenic pressures in response to climate change have brought significant changes in fauna and flora. The most significant examples of human impact on the evolution of ecological niches come from domestication of animals and plants. Crop plant domestication began approximately 10,000 years ago at the dawn of agriculture (Harlan, 1992).

During the domestication process, early agriculturalists consciously or unconsciously selected among wild germplasm for material that was better adapted to human use and cultivation (Frary and Douanlar, 2003, p. 66). Domestication refers to the process of reciprocation, by which animal and plant species come to depend on humans for survival, while providing humans with numerous benefits in turn (Gupta, 2004, p. 55). Domesticated plants and animals are of prime importance for agriculture.

Without agriculture, the complex, technically innovative societies and large human populations that exist today could not have evolved. Agriculture allowed people to become sedentary (living for a prolonged period in one place), establish permanent villages and towns and develop classified societies that included specialized and dedicated segments such as farmers, artisans, soldiers, religious leaders, teachers and governors. Agriculture is thus defined as the cultivation of domesticated plants and animals for use by the human societies, as many domesticated plants and animals would not survive without human intervention (MacDonald, 2003).

An under-utilized species is, by definition, one that has not been fully exploited in terms of its potential use as a food and non food product (Azam-Ali, 1996). Although marama bean is presently known in the wild state, it is a prime candidate for improvement. Marama bean has the potential to address hunger and malnutrition in semi-arid areas and in other parts of the world. Despite the potential to address hunger and malnutrition, marama plant is still under-utilized and not yet domesticated due to lack of practical data prior to the plant breeding programs.

The natural populations of marama are under pressure both from human exploitation of its seeds, rapid human population growth that lead to habitat loss, land degradation and other land use practices (van der Maesen, 2006). In the frame of a breeding program, a detailed knowledge of the genetic diversity within a gene pool would facilitate a more efficient selection of parental genotypes. Scientific approach to diversity conservation, the exploration of plant genetic resources and the design of plant improvement programmes require a detailed knowledge of the amount and distribution of genetic diversity within species. Knowledge of genetic variation within a species can help plant breeders to collect and utilize different genetic resources and forecast potential gain in breeding programs. This research therefore provided vital knowledge on the genetic diversity of marama prior to its domestication.

2.5.4 Importance of Marama bean

The marama bean forms an important part of the diet of rural people in the Kalahari, especially the Khoisan people, where subsistence agriculture is marginal due to drought and low soil fertility and it is a popular delicacy of the Herero, Tswana and other Bantu-speaking people in Southern Africa (van der Maesen, 2006). Young pods are edible as a vegetable, but the main attraction in the large seeds. The beans are not eaten raw but are delicious when roasted. The seeds may be eaten as nuts and the taste is considered to be superior to that of ground nuts.

Marama beans are highly nutritious and are comparing favourably with many existing legume crops (Bower, Hartley, Oh and Storey, 1988, p. 535), the protein content ranges from 37- 39%, and that rivals the value of soybean. Like most legume protein, marama protein is rich in lysine (5%) and deficient in methionine (0.7%). An oil content of 36 – 43 % has been recorded in deshelled Marama seed, which is about twice that of soybean, and approaches that of peanut (van der Maesen, 2006). Essential amino acids composition is slightly better than that for soybeans and is comparable to that of casein (milk protein) (Powell, 1987, p. 217).

The beans are never eaten raw, but nearly always roasted, which gives them a rich, nutty flavour due to certain inhibitors that are heat inactivable. For many Kalahari Desert forages the marama bean has been the preferred staple food for centuries (Keegan and van Staden, 1981, p. 388) and even today it is an important dietary component for some in the region. Hence, marama bean has been described as a plant of considerable potential for semi-arid agriculture.

Despite these surprising qualities, little is known about the plant and nothing is understood about its cultivation. Dedicated research and development effort will lift this wild species out of obscurity and project it far enough to contribute importantly to the food supply is some of the most challenging of all agricultural locations.

2.5.5 Assessments of marama bean genetic diversity using AFLP and RAPD

A few studies have assessed the genetic diversity of marama using different molecular markers, namely, RAPD and AFLP. At present, there is no study that analysed genetic diversity within and among marama populations using SSRs. A study conducted by Monaghan and Halloran (1996, p. 287) on genetic variations within and between natural populations of marama in Southern Africa. The study has revealed a considerable amount of genetic variations exists in marama bean and that most of this variation occurs within populations (84.6%), rather than between populations (15.4%). Since marama bean is a widely distributed, perennial, predominantly out crossing species, their finding fits the general pattern of genetic variation, which Hamrick and Godt (1990, p. 98) suggest is typical of such species. A similar study was conducted by Naomab (2004, p. 12) to assess the genetic variation in the natural populations of marama bean using molecular markers such as AFLP and RAPD. This study also revealed a higher existence of genetic variability the within marama populations (85.2%), rather than between populations (14.8%).

2.6 Ribosomal DNA encoding genes in plants

Within their genomes, all organisms have DNA sequences that code for ribosomal RNA (rRNA), an essential component of cellular protein synthesis machinery (Kollipara, Singh and Hymowitz, 1997, p. 59). The basic organization of ribosomal DNA (rDNA), namely tandem arrayed repeats of the 18S, 5.8S and 28S coding regions separated by an intergenic spacer region have been maintained in most eukaryotic systems. One repeating unit consists of the 18S, 5.8S, and 25S rDNA-coding regions, the internal transcribed spacers (ITS) between these units, and an intergenic spacer (IGS). In contrast to the conserved coding regions, the ITS and IGS regions are more variable within a species than the RNA coding regions.

Generally the intergenic spacer of the rDNA cluster evolves quickly and is highly polymorphic, providing a useful tool for assessing genetic variability, taxonomic and phylogenetic studies (Singh, Deverumath, RamaRao, Singh and Raina, 2008, p. 70). The IGS frequently contains sufficient variation to allow for the examination of genetic relationships between closely related species or even populations or cultivated varieties (Nickrent and Patrick, 1998, p. 184). The variation is normally the result of length variation in the sub-repeats in the IGS (Polanco and Pérez de la Vega, 1997, p. 117).

The variation in the length of IGS has been used as a taxonomic tool in many species including the Triticeae (Dvorak and Appels, 1982, p. 350), maize (Zimmer, Juppe and Walbot, 1988, p. 1126), soybean (Doyle and Beachy, 1985, p. 371), mugbean

(Gerstner, Schiebel, von Waldberg and Hembleben, 1988, p. 724), and rice (Liu et al., 1996, p. 1110). The ITS region comprising ITS1, 5.8S rRNA gene and ITS2 is also an area of particular importance for discrimination at the species level.

2.6.1 Molecular markers based on selective genic (intergenic) amplification

A number of PCR-based DNA markers have been developed in recent years to evaluate genetic variation at the intra-specific and inter-specific levels (Wolfe and Liston, 1998). Molecular markers allow the selection of desired traits based on genotype and can therefore accelerate plant breeding programs (Dolanská and Čurn, 2004, p. 97). The rDNA genes have been shown to be particularly suitable for analyzing genetic variation and for phylogenetic analyses (Hamby and Zimmer, 1992).

Therefore, the structural organization of ribosomal RNA genes (18S-5.8S-28S) clusters originating from nuclear as well as organelle genomes genes had led to the design of many universal primers (Kirti, 2008) and development of techniques for rapid determination of the primary nucleotide sequence of rRNA molecules.

2.7 Statement of the research problem

Agriculture faces the problem of increasing demand from an expanding human population, coupled to threats of reduced area for production as a consequence of climate changes, for example through water deficit ,soil salinity or unpredictable weather at harvest (Callow et al., 1997, p. 14). In essence within the next 50 years agricultural productivity will need to double. There is no doubt that the major determinants of the success or failure of this endevour will depends on factors remote from molecular genetics and in the end it will be farmers that produce food and inevitably most will have limited resources.

Marama bean is such a wild nutritional plant that can be utilized to its full extent due to its higher nutritional quality and it can provide the immediate dietary needs for humans (Moholo and Amarteifio, 1998, p. 330). Despite a wide array of human benefits that can be derived from marama bean, no concerted effort on its breeding and domestication has been undertaken. More information concerning the genetic variability among marama ectypes is needed in order to move forward with a marama improvement program, which will be initiated by this study. Therefore, it is imperative to put efforts into the determination of the range of genetic and associated phenotypic variations present in extant Namibian Marama germplasm using microsatellite markers which have never been done before, to assist in finding the best marama traits for domestication.

This study focused on using microsatellites to determine the amount of genetic variation existing within and between wild populations of marama bean. Such understanding is a prerequisite for successful selection, identifying, conserving and domestication of marama bean, a considerable crop in Southern Africa in the face of climate changes.

2.8 Objectives of the study

The main objective of this study was to use the newly developed microsatellites for the determination of genetic diversity within and between populations of Marama bean in geographically separated populations of Marama bean in the Namibian germplasm.

The study has set specific objectives that will act as stepping stone to the achievement of the main objective.

2.8.1 Specific objectives of the study

1. To use microsatellites for the analysis of genetic variation among Marama bean populations.
2. To estimate the levels of genetic diversity within and among geographically separated populations of Marama bean in the Namibian marama bean germplasm.
3. To assess genetic diversity in marama based on ribosomal intergenic spacer length variation.

2.9 Hypotheses of the study

1. It is possible to use newly developed microsatellite of Marama bean to analyze genetic diversity among Marama bean populations because they are highly informative, polymorphic and have high variability which makes them very powerful genetic markers.

2. A considerable amount of genetic variation exists within and between geographically separated populations of Marama bean because it is generally an out-crossing species.

3. The intergenic spacer of the rDNA cluster evolves quickly, contains sufficient variation and is highly polymorphic allowing for the assessment of Marama bean genetic diversity.

CHAPTER 3: MATERIALS AND METHODS

3.1 Sampling sites and Sample collection

A total of 332 Marama bean individuals representing 11 populations described in Nepolo, Takundwa, Chimwamurombe, Cullis and Kunert (2009, p. 2091) were sampled in the Namibian germplasm (Table 2). The 11 populations were from six localities in the Omaheke region (Sandveld, Otjovanatje, Omipanda, Post 3, Harnas, and Okomombonde), four localities in Otjozondjupa region (Ozondema, Ombunjonjou, Osire, and Otjiwarongo) and one locality in Khomas region (Omitara) of Namibia, (Figure 5). The localities where the Marama bean was growing were recorded using a hand held Global Positioning System receiver (GPS).

The marama populations in these three regions (Otjozondjupa, Khomas and Omaheke regions) are separated by a distance of approximately 100 – 250 km but occur in a similar dry savannah biome. Fresh leaf material was collected from randomly selected marama plants at each location and placed in Ziploc bags well labelled with sample number, date of collection and location. Neighbouring marama plants were not sampled because the plants have long vines and vines from different individuals can intertwine and can appear as one individual.

Table 2: Location, number of accessions of marama individuals collected and region where the locations are within the Namibian germplasm used for genetic variation analysis.

Location	Number of accession	Region
Omitara	19	Khomas
Otjovanatje	20	Omaheke
Sandveld	21	Omaheke
Harnas	25	Omaheke
Ozondema	26	Otjozondjupa
Epukiro/Post 3	30	Omaheke
Omipanda	31	Omaheke
Osire	40	Otjozondjupa
Ombujondjou	40	Otjozondjupa
Otjiwarongo	40	Otjozondjupa
Okomumbonde	40	Omaheke
Total	**332**	

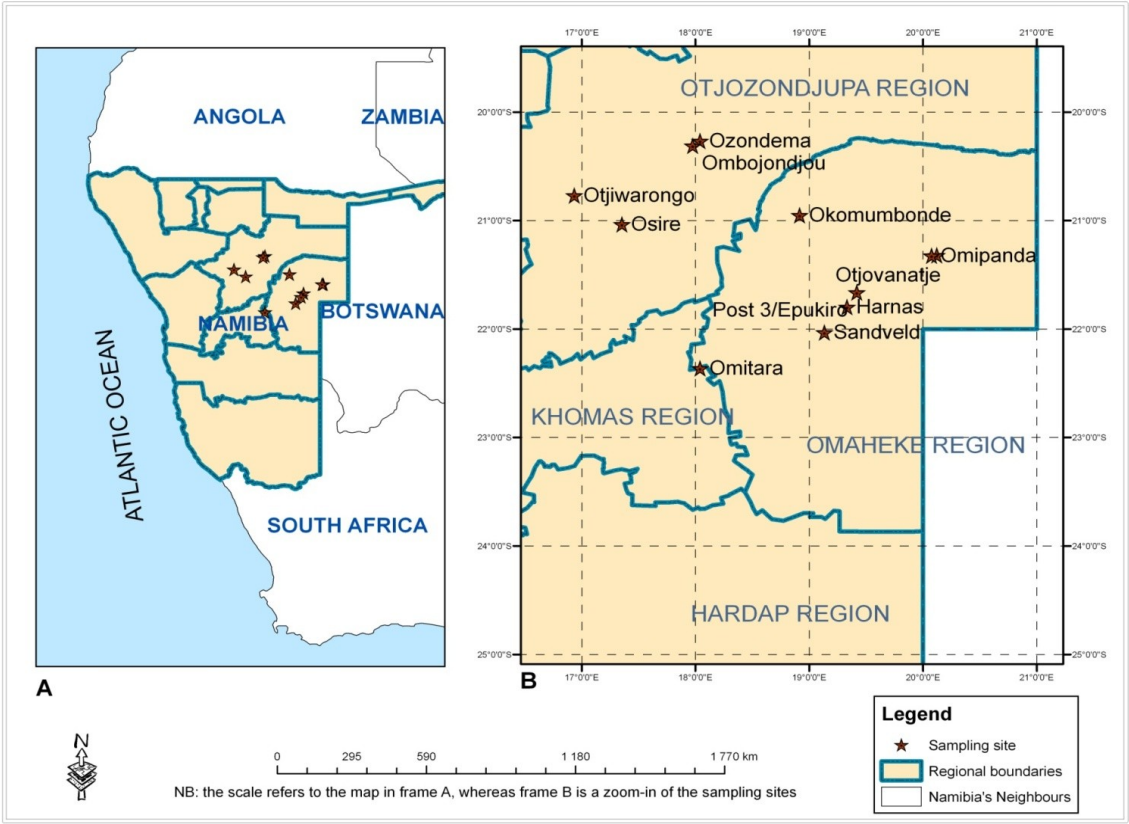

Figure 5: Locations and sampling sites of the Namibian marama bean germplasm.

3.2 Genomic DNA extraction

DNA was extracted from leaves of each of the plant samples collected from the 11 sampling sites using the manufacturer's protocol for the DNeasy (Qiagen, Valencia, CA, USA) mini-kit for purification of genomic DNA from plant tissue. Genomic DNA was extracted from approximately 100 mg of the frozen leaf samples. The leaf material was grounded to a fine powder under liquid nitrogen using a mortar and a pestle. After the grounded leaf material had been finely grounded, the tissue powder was placed into a 2 ml micro-centrifuge tube. After transferring the tissue powder into a micro-centrifuge tube, 400 µl buffer AP1 and 4 µl RNAase stock solutions (100 mg/ml) were added to a maximum of 100 mg of grounded tissue powder followed by vigorously vortex, which ensured that no tissue clumps were visible. The thoroughly mixed mixture was incubated for 10 minutes at 65 ˚C with an occasional inversion of the tube during incubation. The inversion of the tube during incubation lyses the cells. Following cell lysis, 130 µl Buffer AP2 was added to the lysate, mixed thoroughly and incubated for 5 minutes in ice. This step precipitates detergents, proteins and polysaccharides. Following incubation for 5 minutes in ice, the lysate was centrifuged for 5 minutes at 13,800 rpm to remove the majority of precipitates that may arise from lysis. If these precipitates are not removed, they can result in shearing of DNA. After centrifugation, the lysate was pipetted into a QIAshredder Mini spin column placed in a 2 ml collection tube and centrifuged for 2 minutes at 13,800 rpm. The QIAshredder Mini spin column removes most precipitates and cell debris, but a small amount will pass through and form a pellet in the collection tube.

During centrifugation a flow-through fraction is produced. The flow-through fraction produced was transferred into a new micro-centrifuge tube carefully without disturbing the cell-debris pellet. Approximately 450 µl of lysate was recovered and the volume was determined for the following step. Once the flow-through fraction was transferred, 1.5 volumes of Buffer AP3/E was added to the cleared lysate and mixed by pipetting. The amount of Buffer AP3/E to be added was determined by the volume of lysate. If less lysate is recovered, the amount of Buffer AP3/E added was reduced in accordance with the volume of lysate. After the addition of Buffer AP3/E to the cleared lysate, 650 µl of the mixture was pipetted including any precipitate that may have formed into a DNeasy Mini column placed in a 2 ml collection tube and centrifuged for 1 minute at 8000 rpm and the flow through was discarded. The collection tube from this step was reused with the remaining samples. Once all samples have been pipetted into the DNeasy Mini column and centrifuged, the DNeasy Mini column was placed into a new 2 ml collection tube, 500 µl Buffer AW, was added centrifuged for 1 minute at 8000 rpm, the flow-through was discarded and the collection tube was reused for the following step. The next step engaged the addition of 500 µl Buffer AW to the DNeasy Mini spin column and centrifuged further for 2 minutes at 13,800rpm to dry the membrane. It is important to dry the membrane of DNeasy Mini spin column since residual ethanol may interfere with subsequent reactions. This centrifugation step ensures that no residual ethanol was carried over during elution. Following the drying of the membrane, the DNeasy Mini spin column was transferred to a 1.5 and 100 µl Buffer AE pipetted directly onto the DNeasy membrane followed by incubation for 5 minutes at room temperature (15 – 25 ˚C) and then centrifuged for 1 minute at 8000 rpm to elute.

The elution step was repeated once more with a new micro-centrifuge tube to prevent the dilution of the first elute. The DNA concentration was determined on a 1% agarose gel stained with ethidium bromide using known molecular weight standards and stored in clearly labelled microcentrifuge tubes at -20^{o}C.

3.3 Microsatellites Amplification

Four primer pairs MARA001, MARA065, MARA068 and MARA077 were used to amplify regions of genomic DNA from 332 marama bean individuals under thermocycling conditions (Table 2). The four primer pairs used were selected out of 80 newly developed microsatellites primers for marama bean developed by a modified FIASCO technique (Takundwa, 2009). The primer pairs were dissolved in nuclease free water to make 100µM stock solutions. Primer pairs were diluted by the factor of 10 to make 10µM working solutions of each primer pair. DNA with a concentration of 25- 250µg/µl. DNA samples were diluted to 10 ng of DNA per µl for PCR amplification. PCR amplifications were performed in 25 µl reaction volumes, with a 2X PCR master mix from Fermentas. Each PCR reaction contained 1 µl template genomic DNA, 1 µl (1 µM) of SSR forward primer, 1 µl of SSR reverse primer, 12.5 µl of the 2X PCR master mix and 9.5 µl nuclease free water. The thermocycling reaction profile used involved an initial denaturation step of 95^{o}C for 4 minutes, followed by 35 cycles of denaturation at 95^{o}C for 30 seconds, an annealing at between 55^{o}C and 65^{o}C (primer sequence dependent) for 1 minute and an extension at 72^{o}C for 2 minutes, a final extension at 72^{o}C for 5 minutes and then held at 4^{o}C.

Amplification products were separated by electrophoresis on 2% agarose gels, and the 100bp DNA ladder (Fermentas) was used as a standard. Banding patterns were visualized by staining the gels in ethidium bromide solutions and viewed under UV radiation.

Table 3: List of selected primer pairs used to detect polymorphism and for the analysis of genetic variation in 11 marama populations of the Namibian germplasm.

Primer	Sequence 5'→ 3'	Repeat	Expected PCR product size (bp)
MARA001	L - GCACAACCAATTTCCTGCTT R - TCCCTCACTGGCCTATATCC	$(gag)_5$	137
MARA065	L - TGGTGGTAGGGTGGTGGTAT R - CCACTTTTCACAGGCAAACA	$(ttc)_6$	191
MARA068	L - GGAGGAGGAGGAGGATTTG R - GAGGATCCACTCCCTCACTG	$(gag)_5$	192
MARA077	L- CTCAGCACTCTCCAGCCTCT R - GGGTTGGTTGAAGAGGGAGT	$(aag)_5$	197

3.4 Expected SSR band sizes validation through amplicons sequences

To confirm whether the observed SSR amplicons patterns were caused by SSRs or by insertions that has nothing to do with SSRs, a validation analysis was done by means of gel extraction and sequencing of selected amplicons. Individuals used for this analysis were selected from Harnas and Ozondema populations.

3.4.1 Gel extraction of SSR amplicons

SSR amplicons to be extracted were electrophoresed through a 2% agarose gel in TBE buffer containing ethidium bromide to fractionate DNA fragments. Amplicons were visualized by UV trans-illuminator and the expected band sizes were excised from the gel using a sterile scalpel and placed into a 1.5 mL microtube. The DNA was extracted using E.Z.N.A.™ Gel Extraction Kit (Omega Bio-tek, Inc., USA) according to the manufacturer's specifications. The mass of the gel sliced was calculated and three volume of Binding Buffer (XP2) was added. The microtube was incubated at 55 °C – 60 °C for seven minutes until the gel was completely dissolved. For the recovery of 25 µg DNA, 700 µl of the DNA/agarose solution was added to the HiBind® DNA mini column placed in a 2 ml collection tube and centrifuged at 10,000 x g for 1 minute at room temperature. The supernatant was discarded and the HiBind® DNA mini column was placed back in the same collection tube.

The HiBind® DNA mini column was washed by adding 300 µl of Binding Buffer (XP2) and centrifuged at 10,000 x g for 1 minute. It was further washed by adding 700 µl of SPW Wash Buffer and centrifuged at 10,000 x g for 1 minute. The supernatant was discarded and the empty HiBind® DNA mini column was centrifuged for 2 minutes at maxi speed 13,000 x g to dry the column matrix. The HiBind® DNA mini column was placed into a clean 1.5 mL microtube and 30 – 50 µl of Elution Buffer (10 mM Tris-HCI, pH 8.5) was added directly onto the column matrix and incubated at room temperature for 1 minute.

It was than centrifuged for 1 minute at maximum speed of 13,000 rpm to elute the DNA. The yield and quality of DNA was determined by the absorbance of an appropriate dilution of the sample at 260 nm and then at 280 nm. The extracted DNA was stored at – 20 °C, until it was sent for sequencing.

3.4.2 Gel extracted DNA amplification

The extracted DNA from the gel was amplified using the specific primer pair which gave the rise to bands that were used in gel extraction to obtain single bands. PCR amplifications were performed in 25 µl reaction volumes, with a 2X PCR master mix from Fermentas. Each PCR reaction contained 1 µl template genomic DNA, 1 µl (1 µM) of SSR forward primer, 1 µl of SSR reverse primer, 12.5 µl of the 2X PCR master mix and 9.5 µl nuclease free water. The PCR reaction profile used involved an initial denaturation step of 95^{o}C for 4 minutes, followed by 35 cycles of denaturation at 95^{o}C for 30 sec, an annealing at between 55^{o}C and 65^{o}C (primer sequence dependent) for 60 sec and an extension at 72^{o}C for 2 minutes, a final extension at 72^{o}C for 5 minutes and then held at 4^{o}C. Agarose gel (2.5%) visualization of PCR products was then used to determine if single bands were obtained, which is an indication of a successful gel extraction.

3.4.3 SSR amplicon sequences and analysis

All amplicons that showed single band pattern were sequenced and analyzed with the Roche 454 GS-FLX platform at Inqaba Biotech (Pretoria, Gauteng, South Africa). Sample preparation and analytical processing such as base calling, were performed at Inqaba Biotech using the manufacturer's protocol as described previously (Santana, Coetzee and Steenkamp, 2009, p. 218).

The sequence information gathered was sorted into BioEdit Sequence Alignment Editor and ABI Chromatogram. The obtained sequences were blasted using Blast (nt) in NCBI to determine and read if the sequences are really SSR either in legumes or the other plant species. The sequences were also compared with the sequence information from where the SSRs were developed to see if they possess the repeat unit as that one present in the original SSRs sequences.

3.5 Data Analyses

3.5.1 Genetic diversity Analysis

The genetic diversity within each marama population was measured by the Shannon diversity index (King and Schaal, 1989, p. 1117) which is calculated as:

$$H = -\sum_{i=1}^{k} pi \log_e p_i$$

Where *k* is the number of bands and *pi* is the frequency of the *i*th band in a given population. *H* is the population diversity for each primer pair. *H* was averaged to give the average per-locus diversity within population (H_O). The average diversity for all populations (H_{POP}) was calculated at two levels for each primer pair as the average of *H* and over all primer pairs as the average of H_O.

SSR diversity for the species (H_{SP}) was calculated using pooled band frequencies for all individuals from the eleven populations. The proportion of within-population diversity relative to total diversity was calculated by:

$$\text{Genetic diversity within populations} = \frac{H_{POP}}{H_{SP}}$$

Genetic diversity between marama populations was calculated by:

$$\text{Genetic diversity between populations} = \frac{H_{SP} - H_{POP}}{H_{SP}}$$

3.5.2 Cluster Analysis

The products of SSR amplification were recorded with regard to the presence (1) or absence (0) of bands to obtain a similarity matrix. Bray-Curtis Similarity method estimates were used to construct phenetic trees by using the unweighted pair-group method with arithmetic averages (UPGMA) cluster analysis to infer the phenetic relationships. The dendogram were constructed with the Primer-E 5 for Windows (Plymouth Routines in Multivariate Ecological Research) software package (PRIMER-E Ltd. 2001).

3.6 Assessment of marama bean genetic diversity based on ribosomal intergenic spacer length variation

3.6.1 Collection of samples

Leaf samples from 31 marama bean individuals were used for ribosomal intergenic spacer length variation analysis. These samples were collected from the Namibian germplasm (Omipanda, in Omaheke region, Eastern Namibia) as described in section 3.1.

3.6.2 DNA Extraction

Genomic DNA was extracted from each of the plant leaf samples collected using the DNeasy mini kit (Qiagen) for purification of total DNA from plant tissues. DNA extraction from different individual plant samples took place at the Molecular Biology Laboratory at the University of Namibia. The DNA for each sample was stored at -20 °C.

3.6.3 PCR amplification of the IGS region

The pair of primers 18SL (5'-CTCAATGAGCCCGGTATTGT-3') and 28SR (5'-ACGAGAGGAACCGTTGATTC- 3') was used to amplify across the IGS region of the rDNA. A hot start PCR protocol with Takara SpeedStar Taq polymerase was used. The 50 µl reaction mixture containing 10X Fast Buffer I, 2.5mM of each dNTP, 5 units of SpeedStar *Taq* DNA polymerase, 1 µM each of primers 18SL and

28SR, with 20 ng genomic DNA template. The PCR conditions were: 98 °C for 45 seconds for initial genomic DNA denaturation; 30 cycles of 94 °C (DNA denaturation) for 5 seconds, 58 °C (primer annealing) for 20 seconds, 72 °C (DNA amplification) for 1 minute, and final extension at 72 °C for 5 minutes. The PCR products were separated on a 1.5 % agarose gel stained with ethidium bromide using 1X TBE buffer. The size of amplified fragments was determined using Bioline marker I.

3.6.4 Analysis of IGS region amplification products

The amplification products were scored as present (1) or absent (0) binary scores for all the primer pair combinations. Percent polymorphism was computed by the formula given below:

$$Per\ cent\ polymorphism = \frac{[Total\ number\ of\ polymorphic\ bands]}{Total\ number\ of\ \ bands} \times 100$$

Molecular data were used to determine Bray-Curtis similarity coefficients using PRIMER 5 software. These similarity coefficients were then used to construct dendogram depicting the genetic relationship employing the similar software.

3.7 Statistical Analysis

Data were tested for normality using a Kolomogorov-Smirnov test and the data that were normally distributed data were analyzed by ANOVA and means were compared using a post-hoc Scheffé multiple comparison test, using SPSS 16™ for Windows® at a confidence interval (CI) of = 95%, $\alpha = 0.05$. ANOVA was performed to test for the significant differences in mean genetic diversity within eleven marama bean populations.

CHAPTER 4: RESULTS

In this section results from the four microsatellite markers used to assess genetic diversity within and among Marama bean populations are presented. Results on the expected SSR band size validation by amplicon sequencing are also presented under this section. Additionally, results from the assessment of marama bean genetic diversity based on ribosomal intergenic spacer length variation is also made know under this section.

4.1 DNA extraction

Genomic DNA was extracted from fresh frozen young marama bean leaf samples. The extraction process employed a commercial extraction kit and is described in chapter 2 and, it was observed that a good yield of marama bean DNA was obtained (Figure 6). Total genomic DNA was successfully extracted from all 332 marama bean individual leaf samples from 11 populations in the Namibian germplasm. The resulting DNA with a concentration of 25- 250ng/µl was collected and electrophoresed through a 1% agarose gel stained with ethidium bromide to determine the DNA concentration. DNA samples were diluted accordingly to get equal concentrations of 10 ng/µl.

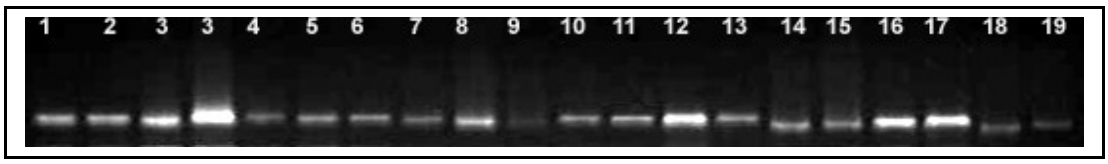

Figure 6: Electrophoresis results of genomic marama bean DNA extraction from leaf samples, using a Qiagen extraction kit.

4.2 Microsatellites (SSRs) Analysis

Polymorphic primer pairs MARA 001, MARA 065, MARA 068 and MARA 077 were used to determine genetic variation in each of the 11 ecotypes using the same reaction profile and conditions, with the difference being in the annealing temperature for each primer pair which was primer pair sequence dependent. The SSR amplicons products were electrophoresed through a 2.5% agarose gel to detect polymorphism in all 11 populations of the Namibian germplasm. A total number of 2465 bands were generated by the four microsatellites primers analysed, with an average of 616.3 bands per primer pair. The number alleles per locus ranged from two for MARA001 and MARA068, three for MARA065 to six for MARA077 among marama bean individuals.

4.3 Primer pair (SSRs) MARA001 and MARA068 amplicons analysis

Primer pair MARA001 produced 475 reproducible bands and MARA068 produced 551 bands. The number of alleles for these two loci ranged from one to two, and gave the characteristic diploid type profile. Therefore, genetic variation in each of the populations was calculated as heterozygosity for the 2 microsatellite loci MARA 001 and MARA 068 as these 2 loci gave the characteristic diploid type profile enabling calculation of heterozygosity (Table 4). For instance, out of 20 individuals from Otjovanatje amplified with primer pair MARA001, ten (50%) were heterozygous (Figure 7).

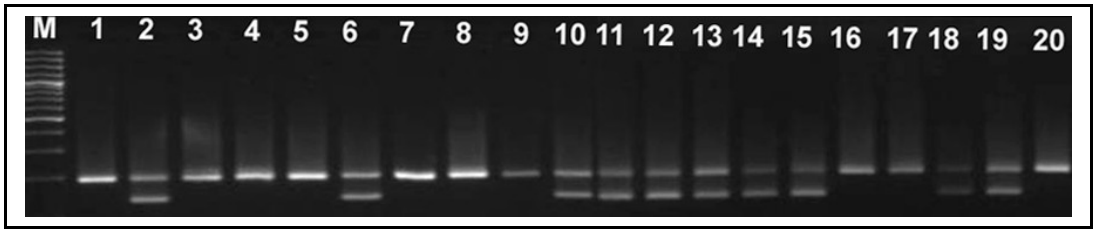

Figure 7: A 2.5% Agarose gel electrophoresis of primer pair MARA001 amplification products showing 20 marama bean individuals from Otjovanatje. M indicates the DNA molecular size marker (O'gene Ruler 100 bp).

Twelve individuals (46%) out of 26 individuals from Ozondema amplified using primer pair MARA068 were heterozygous (Figure 8). It was observed that these two primer pairs gave between one and two bands (alleles) with dominant upper bands in all 11 populations.

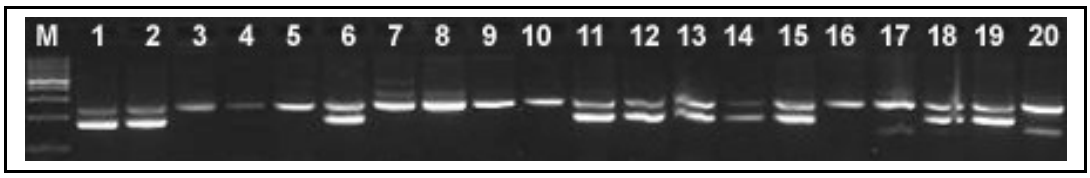

Figure 8: A 2.5% Agarose gel electrophoresis of primer pair MARA068 amplification products showing 20 marama bean individuals from Otjovanatje. M indicates the DNA molecular size marker (O'gene Ruler 100 bp).

The direct count of heterozygosities for locus MARA001 in 11 marama bean populations ranged from 0.25 to 0.61, with the average heterozygosity of 0.43 for all populations. The heterozygosities for locus MARA068 ranged from 0.33 to 0.93 with 0.51, the average heterozygosity for all 11 populations (Table 4). The average heterozygosity for the two loci (MARA001 and MARA068) ranged from 0.30 to 0.74, with a mean value of 0.51 for all loci.

Table 4: Heterozygosity and average heterozygosity at two loci MARA 001 and MARA 068 in 11 populations of marama bean in the Namibian germplasm as determined by electrophoresis.

Population	**Number of Individuals**			**Heterozygosity**		**Average heterozygosity at two loci MARA 001 and MARA 068**
	Heterozygotes MARA 001	***Heterozygotes MARA 068***	***Total***	***Locus MARA 001***	***Locus MARA 068***	
Omitara	8	11	19	0.42	0.58	0.50
Otjovanatje	10	11	20	0.50	0.55	0.53
Sandveld	9	7	21	0.43	0.33	0.38
Harnas	10	12	25	0.40	0.48	0.44
Ozondema	12	12	26	0.46	0.46	0.46
Epukiro/Post 3	13	16	30	0.43	0.53	0.48
Omipanda	19	27	31	0.61	0.87	0.74
Osire	14	20	40	0.35	0.50	0.43
Ombujondjou	17	37	40	0.43	0.93	0.68
Otjiwarongo	16	37	40	0.40	0.93	0.67
Okomumbonde	10	14	40	0.25	0.35	0.30
Average				**0.43**	**0.51**	**0.51**

4.4 Primer pair MARA065 amplicons analysis

A total number of 700 bands were produced by primer pair MARA065. It also gave one to three bands with minimum separation. This profile pattern was observed in Harnas as well as the other 10 populations amplified using the same primer pair.

Individual 1,2,10 and 19 from Harnas gave three distinctive bands, while individual 20, 22 and 25, gave two bands each. Other individuals such as 20, 22 and 24 gave two bands, with minimum separation (Figure 9).

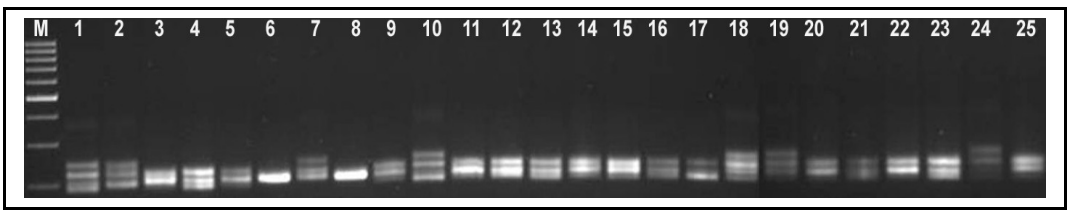

Figure 9: A 2.5% Agarose gel electrophoresis of primer pair MARA065 amplification products showing 25 marama bean individuals from Harnas. M indicates the DNA molecular size marker (O'gene Ruler 100 bp).

4.5 Primer pair MARA077 amplicons analysis

Amplification of all 11 populations using primer pair MARA077 produced a total of 739 bands. The number of alleles for this locus ranged from one to six bands among individuals of marama bean. The amplicons obtained with primer pair MARA 077 shows smears together with bands for certain individuals. For instance, in individuals 1, 4 6, 9, 15 and 21 shows smears - bands association was observed (Figure 10).

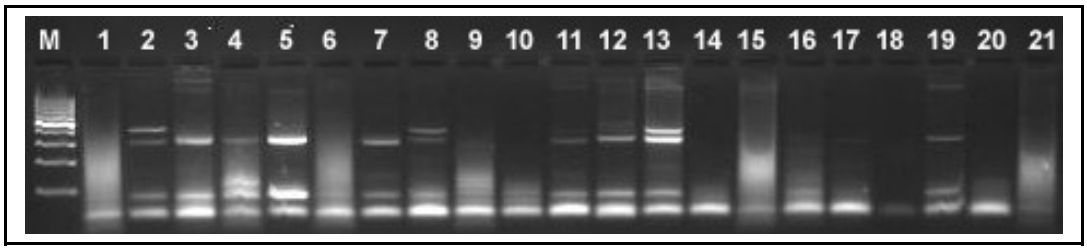

Figure 10: A 2.5% Agarose gel electrophoresis of primer pair MARA077 amplification products showing 21 marama bean individuals from Sandveld. M indicates the DNA molecular size marker (O'gene Ruler 100 bp).

4.6 Genetic variability within and between marama populations

The average within population diversity using Shannon information index for the eleven marama bean populations was 0.898 and ranged from 0.862 (± 1.065 SE) for population Otjovanatje to 0.931 (± 0.173 SE) for population Omitara (Table 5). This

suggests that Omitara is the most diverse marama population, but the ranking of populations in terms of highest to lowest genetic diversity changed with the primer pair used. The highest amount of genetic diversity for an individual population was 1.349 for Omitara and it was revealed by SSR primer pair MARA077.

SSR primer pair MARA001 revealed the lowest individual level diversity of 0.608 in the Omitara population, but the greatest estimate of average population diversity was revealed by primer pair MARA077 (1.217). Most of the diversity (66.5%) in the marama populations occurs within population rather than between (33.5%) populations. Genetic diversity within populations ranged from 48.7% to 86.3%, while genetic diversity between marama populations ranged from 13.7% to 51.3%, however the estimate and distribution of genetic variability within and between marama populations varies between primer pairs. For instance, primer pair MARA077 detected up to 51.3% of variation between populations, while primer pair MARA068 detected up to 13.7% of variation between marama populations. Consequently, MARA068 detected up to 86.3% of variation within populations, while MARA077 detected up to 48.7% variation within marama populations.

Table 5: Genetic diversity of eleven marama populations and the partitioning of the genetic diversity within and between populations (Shannon Information index) for four microsatellite primer pair analyzed.

	H														
Primer pair	**OMI**	**HAR**	**SAN**	**OZO**	**OTV**	**EPK**	**OMP**	**OTR**	**OSE**	**OKO**	**OMB**	H_{POP}	H_{SP}	(%) H_{POP}/H_{SP}	(%) $(H_{SP}-H_{POP})/H_P$
MARA001	0.608	0.69	0.664	0.637	0.637	0.613	0.673	0.679	0.632	0.632	0.614	**0.644**	0.823	78.3	21.7
MARA065	1.077	1.047	1.053	1.09	1.065	0.974	1.032	1.089	1.07	1.061	1.054	**1.056**	1.999	52.8	47.2
MARA068	0.688	0.666	0.673	0.624	0.65	0.668	0.693	0.693	0.688	0.68	0.693	**0.674**	0.781	86.3	13.7
MARA077	1.349	1.046	1.229	1.248	1.239	1.267	1.106	1.146	1.268	1.175	1.311	**1.217**	2.5	48.7	51.3
	***H*o**														
	0.931	**0.862**	**0.905**	**0.899**	**0.898**	**0.881**	**0.876**	**0.902**	**0.915**	**0.887**	**0.918**	**0.898**	**1.526**	**66.5**	**33.5**

diversity for all marama populations; $H_{SP,}$ genetic diversity of species; H_{POP}/H_{SP}, % of diversity within populations; $(H_{POP} - H_{SP})/ H_{SP,}$ % of diversity between populations.

Population from Omitara (OMI), population from Otjovanatje (OTV), population from Sandveld (SAN), population from Harnas (HAR), population from Ozondema (OZO), population from Epukiro (EPK), population from Omipanda (OMP), population from Otjiwarongo (OTR), population from Osire (OSE), population from Okomombonde (OKO), population from Ombunjonjou (OMB).

The lowest marama average population genetic diversity was 0.862 (± 1.065 SE) for population Harnas and the highest average population genetic diversity was 0.931 (± 0.173 SE) for population Omitara (Figure 11). However, one-way ANOVA analysis revealed that there was no significant difference ($p > 0.05$) in mean genetic diversity within the 11 marama bean populations.

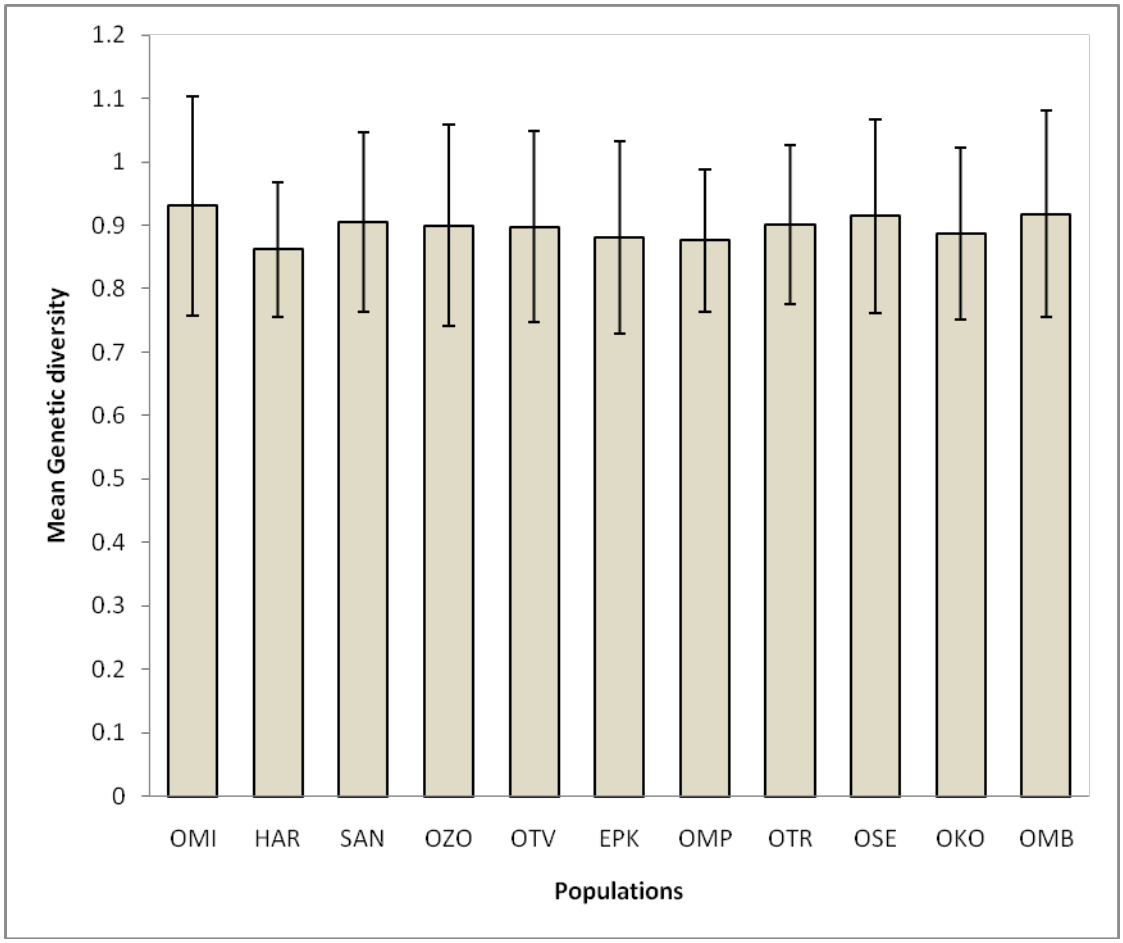

Figure 11: Mean per-locus genetic diversity and their Standard Errors (SE) among 11 Marama bean populations in the Namibian germplasm, (OMI = Omitara, OTVO = Otjovanatje, SAN = Sandveld, HAR = Harnas, OZO = Ozondema, EPK = Epukiro, OMP = Omipanda, OTR = Otjiwarongo, OSE = Osire, OKO = Okomombonde and OMB = Ombunjonjou).

The average within marama population genetic diversity (H*pop*) revealed by the four primer pairs used ranged from 0.644 (± 0.0086 SE) for primer pair MARA001 to 1.217 (± 0.027 SE) for primer pair MARA077 (Figure 12). The highest within population genetic diversity (1.217) was revealed by primer pair MARA77 and the lowest within population genetic diversity (0.644) was revealed by MARA001.The ANOVA analysis revealed a significant difference (p <0.05) in the mean marama bean genetic diversity revealed by the four primer pairs.

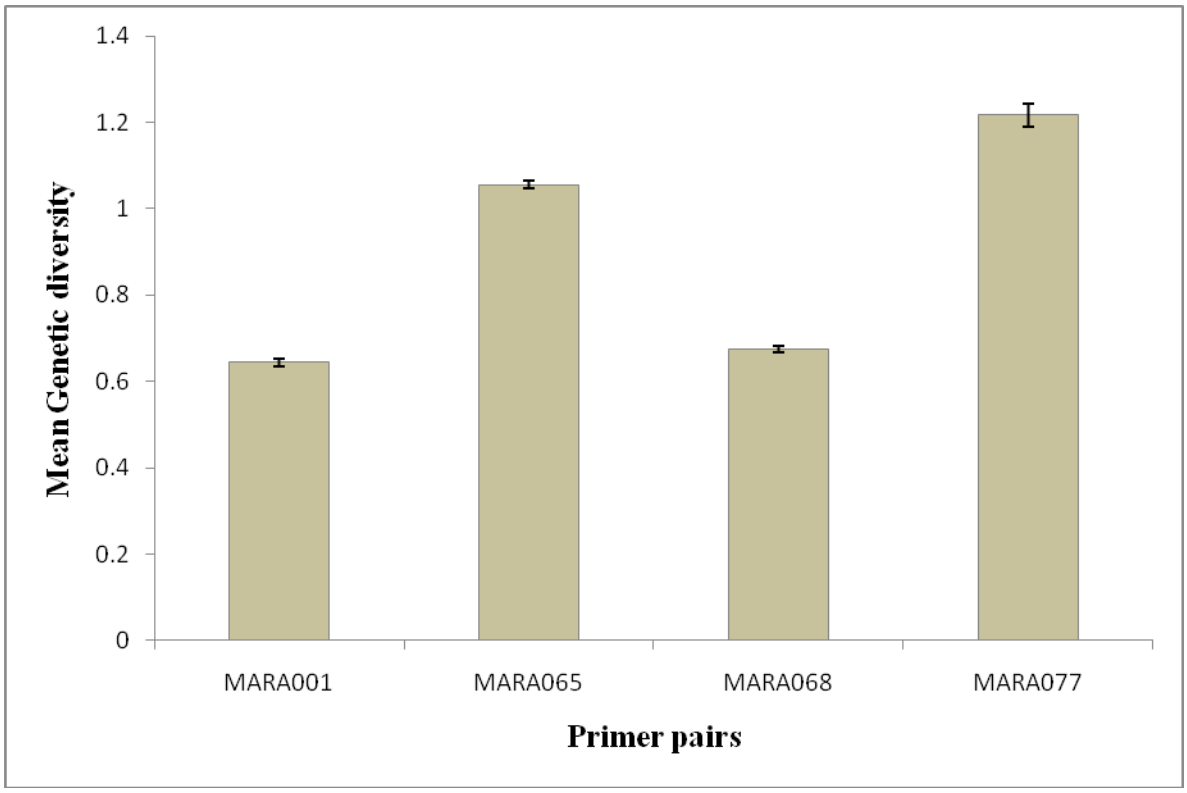

Figure 12: Mean genetic diversity of all 11 Marama bean populations and their SE for each SSR primer pair.

4.7 Expected SSR bands size validation by amplicons sequences

A total of 8 bands were sequenced to determine whether the banding patterns observed such as band sizes either more than 30 bp above or below the expected SSR band size were caused by SSRs or by insertions that have nothing to do with SSRs (Figure 13). Of the entire generated band sequences blasted only one was confirmed to be a microsatellite sequence of Coffee (*Coffea canephora*) with 90% similarity (Figure 14) However, other band sequences such 17 MARA001 and 20U2 MARA001 have the repeat unit of $(gag)_5$ as present in the sequence from where the used SSR primer pair was developed.

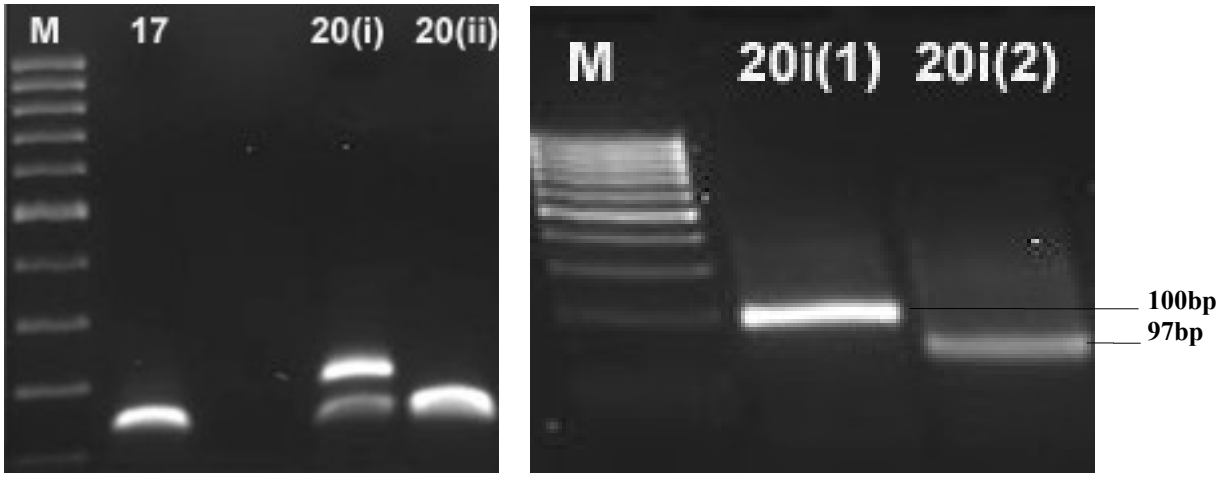

Figure 13: Amplicons of gel extraction produced by MARA001 and visualized on a 1.5% agarose gel. M indicate a molecular DNA size marker (O'gene Ruler 100 bp).

Identities = 29/32 (90%),

```
Query  116  CTTGTCCTCAAGACT-GAAAGCAGGAAATTGG  146
            ||||||||||||||| ||||  ||||||||||
Sbjct  253  CTTGTCCTCAAGACTGGAAATGAGGAAATTGG  222
```

(a)

SEQUENCE 20U1_MARA001_R

TCAGCATCGCTTTCATCATTGGCCCCCCAAAGCCCAACTTAATACAACCCAT
AACCCGCTAAGTGTCACTCATTTTCCTGGTCCATATCACTCACCGTCCCTTGA
GAGCCATC**CTTGTCCTCAAGACTGAAAGCAGGAAATTGG**TTGTGCAAACA
****

(b)

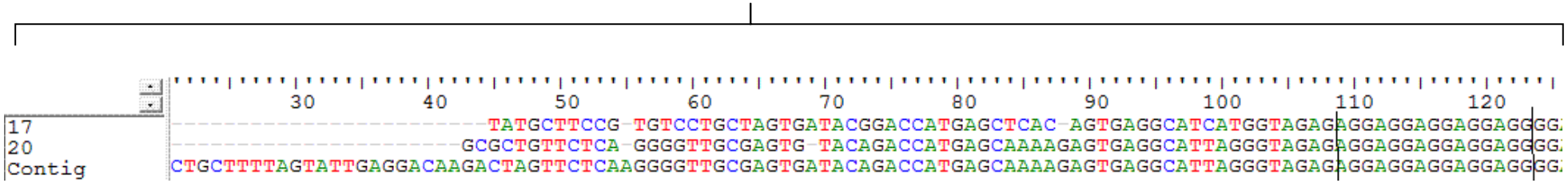

Figure 14: SSR validation sequences and alignments. (a) Alignment of the SSR band obtained sequences with the BLAST match of *Coffea canephora* with 90% similarity, (b) contig from where primer pair MARA001 was designed aligned with SSR band obtained sequences (17 and 20) showing the presence of the SSR MARA001 repeats $(GAG)_5$ in their sequences.

4.8 Cluster analysis

Dendograms were constructed for the pooled data using Bray-Curtis Similarity to elucidate the genetic relationships among marama populations over four SSR primer pairs. All dendograms were found to explain between 23% and 87% of the genetic similarity coefficient.

4.8.1 UPGMA dendogram of 332 marama individuals based on SSR primer pair MARA001

The dendogram constructed for the pooled primer pair MARA001 data indicated the existence of two clusters (Figure 15). The first cluster (1) was formed by individual like SAN6,OZO2,0TV1,EPK9,OMP3,OTR14,OSE19,OKO1,OMB5,OMI4 and HAR2 at the genetic similarity level of 52%. The second cluster (2) had two sub-clusters at a genetic similarity level of 68%. Marama bean individuals HAR18, SAN20, OMP16, OTR11, OSE7 and OKO29 formed the first sub-cluster (A), whereas OKO26, OMB8, OMI5, HAR17, SAN3, OZO10, OTV2, EPK6, OMP4, OTR2 and OSE38 formed the second (B) sub-cluster. The detailed clustering of all individuals that formed different main clusters and sub-clusters are illustrated in Appendix A.

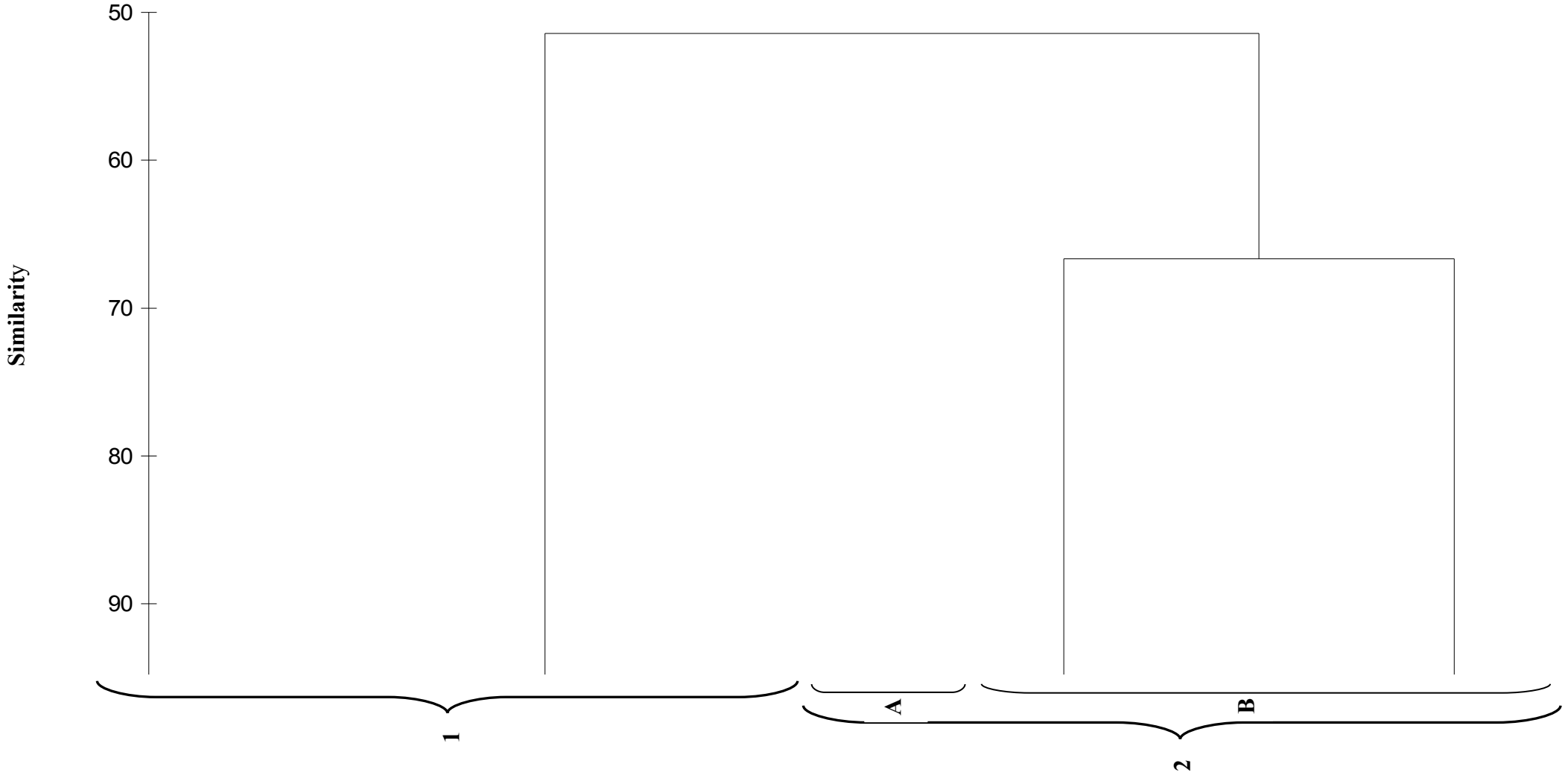

Figure 15: UPGMA dendogram constructed from MARA001 SSR profiles of 332 individuals from 11 populations of Marama bean. The scale represents Bray-Curtis similarity coefficients.

4.8.2 UPGMA dendogram of 332 marama individuals based on SSR primer pair MARA065

According the dendogram for the pooled data of primer pair MARA065, at the genetic similarity level of 23% the marama individuals were divided in the first cluster (1) containing OTR25, SAN13, OTV2, EPK4 and EPK18 respectively. The second cluster (2) was formed at genetic similarity level of 24% and it had six sub-clusters. The first sub-cluster (A) was only formed by OMI11 at a genetic similarity level of 24%, whereas OTV14, EPK7, OMP22, OSE37, OKO23, OMB14, OMI2, HAR3 and OZO10 formed the second sub-cluster (B) at a genetic similarity level of 44% (Figure 16). The third sub-cluster consisted of OZO26, OTV11, EPK19, OMP27, OTR11, OSE28, OKO16, OMB5, OMI14, HAR9 and SAN6 at a genetic similarity of 60%, whereas the fourth sub-cluster was formed by HAR1,SAN1, OZO25, OTV1, EPK10, OMP19, OTR40, OSE17, OKO34, OMB32, OMI10 and HAR13 formed at a genetic similarity of 63% respectively. The fifth sub-cluster was formed by OZO11 and OZO18, whereas the sixth sub-cluster consisted of OZO19, EPK6, OMP9, OTR5, OKO7, OMB35, OMI18, HAR10 and SAN5 at a genetic similarity level of 70%. Marama populations showed some similarities, even though none of them was clustered under a single sub-cluster. Populations like Omitara, Okomombonde, Osire, Harnas, Ozondema, Omipanda and Ombunjonjou were all clustered under the second main cluster (2) under different sub-clusters (Appendix A).

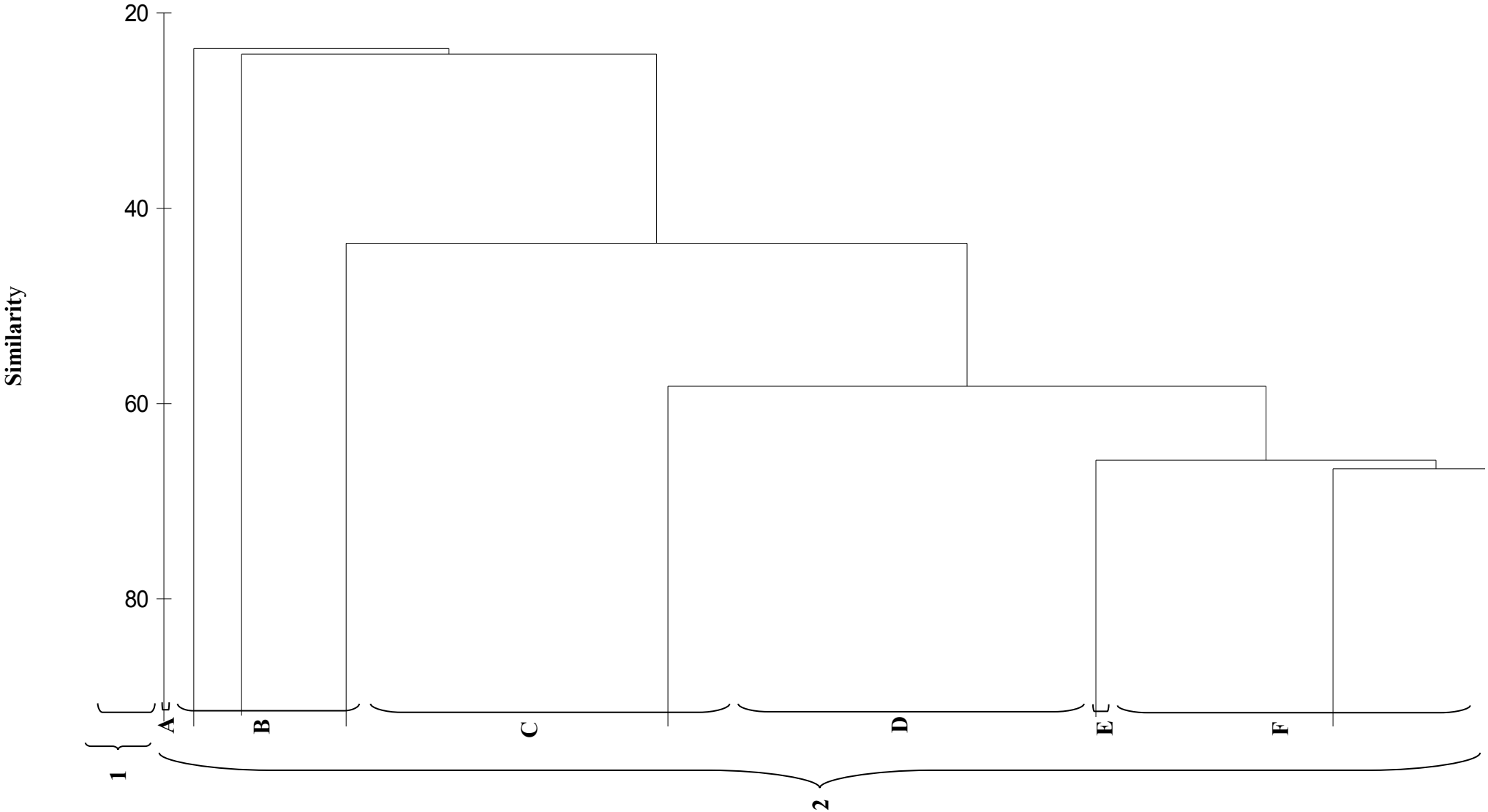

Figure 16: UPGMA dendogram constructed from MARA065 SSR profiles of 332 individuals from 11 populations of Marama bean. The scale represents Bray-Curtis similarity coefficients.

4.8.3 UPGMA dendogram of 332 marama individuals based on SSR primer pair MARA068

The dendogram constructed for the pooled data of primer pair MARA068 indicated the existence of two clusters (Figure 17). The first cluster (1) consisted of OMB40, OTR2, OMI15, HAR4, SAN21, OZO26, OTV5, EPK27, OKO37 and OSE27, at a genetic similarity of 62% respectively. The second cluster (2) had two sub-clusters at a genetic similarity of 65%. OMI19, OKO16, OKO28, OSE5, OMB28, OMI16 and OMI18 formed the first sub-cluster (A), whereas SAN3, OZO18, OTV11, EPK4, OMP12, OKO39, OSE40, OMB33, OTR24, OMI10, HAR13 and SAN1 formed the second sub-cluster respectively. The dendogram indicated that Ombunjonjou, Osire and Otjiwarongo are the most similar populations with most of their individuals being grouped in sub-cluster B of cluster 1 (Appendix A). The primer pair discriminated population Otjiwarongo and Ombunjonjou from the rest of the populations with all of its individuals being clustered under one sub-cluster with only one individual from Ombunjonjou being clustered under main cluster 2.

4.8.4 UPGMA dendogram of 332 marama individuals based on SSR primer pair MARA077

The dendogram constructed for the pooled primer pair MARA077 data indicated the existence of two clusters (Figure 18) at a genetic similarity level of 55%. The first cluster (1) had six sub-clusters. The first sub-cluster (A) was formed at a genetic similarity of 69% and contained marama individuals such as OTV5, EPK17, OMP7,OTR11,OSE9, OKO36, OMB37, HAR24, SAN10 and OZ015 respectively.

The second sub-cluster (B) was formed at a genetic similarity level of 74% and was formed by OMI16, OTV2, OTV6, EPK2, EPK21, OMP24, OSE24 and OSE39. The third sub-cluster (C) was formed by OMB29, OMI6, HAR12, SAN5, OZO17, OTV20, EPK4, OMP31, OTR8, OSE1 and OKO37, at a genetic similarity level of 81%, whereas OTR36, OMB16, OMB18,EPK24 and OTR34 formed the second fourth-cluster (E) at a genetic similarity level of 85%.

The fifth sub-cluster (F) was formed by OMP11, OTV18 and OMP3 only at a genetic similarity of 87%, whereas the sixth sub-cluster contained OKO3,OMB7, OMI4, HAR18, SAN13 ,OZO19, EPK29, OMP28, OTR2, OSE40 and OKO02 at the same genetic similarity value of 87%. The second cluster (2) had two sub-clusters formed at a genetic similarity value of 66%. The first sub-cluster (i) was formed by only two individuals OMB21 and OMB26, whereas the second sub-cluster (ii) was formed by OTV4, EPK7, OMP6, OTR13, OSE6, OKO1, OMB5, OMI1, HAR2 and SAN21 respectively. None of the sub-clusters managed to group individuals from the same population in one cluster. Individuals from different populations were scattered around the produced sub-clusters and neither population was grouped in the main cluster under sub-clusters (Appendix A) even though most of the different populations fell under the first main cluster.

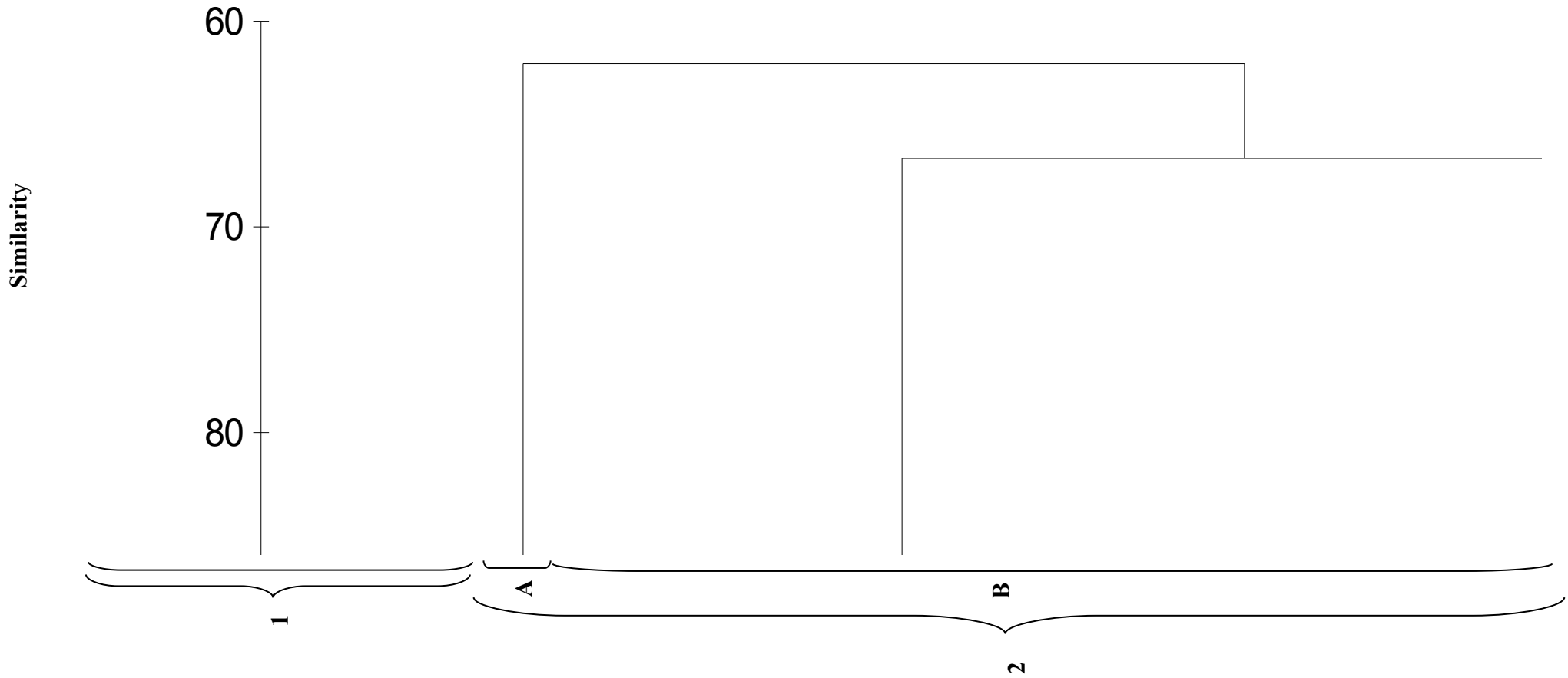

Figure 17: UPGMA dendogram constructed from MARA068 SSR profiles of 332 individuals from 11 populations of Marama bean. The scale represents Bray-Curtis similarity coefficients

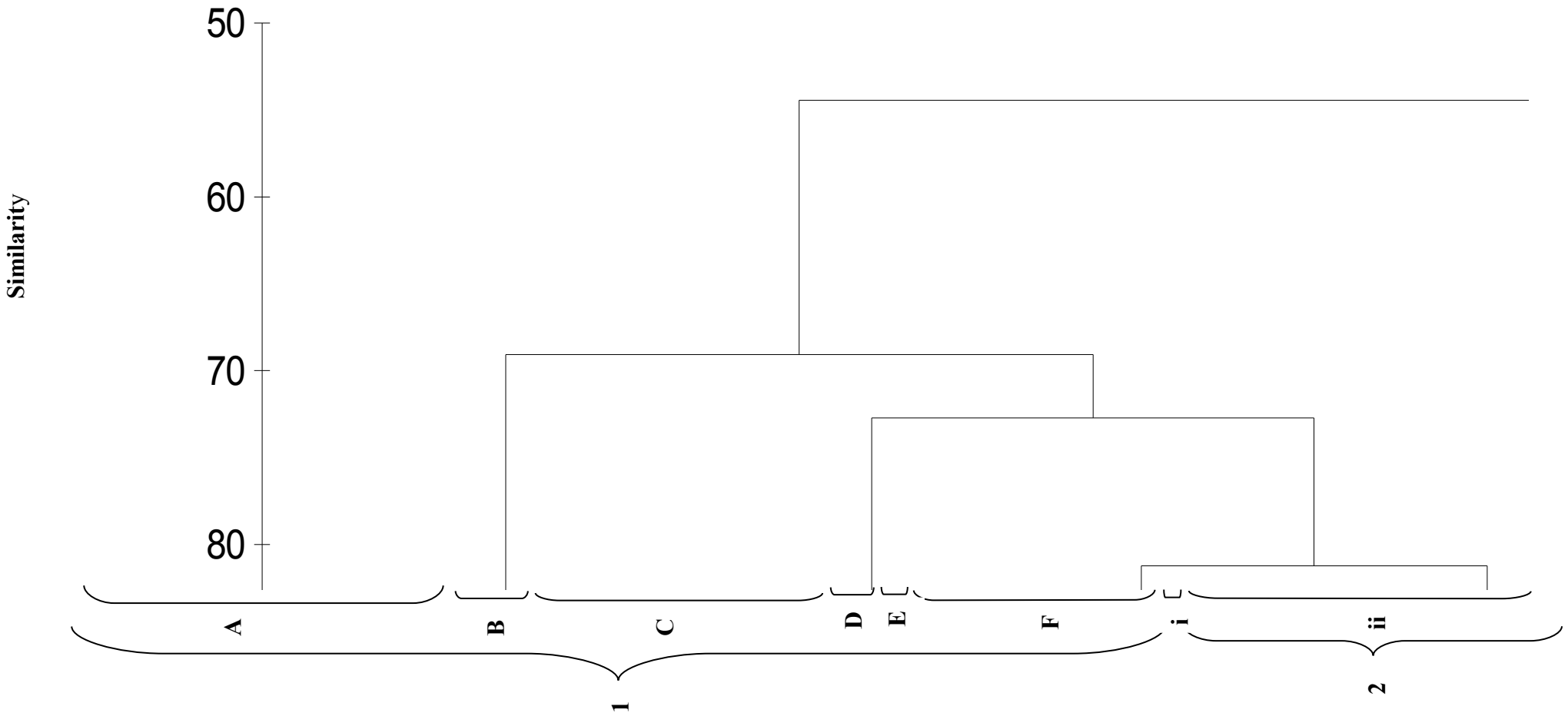

Figure 18: UPGMA dendogram constructed from MARA077 SSR profiles of 332 individuals from 11 populations of Marama bean. The scale represents Bray-Curtis similarity coefficients.

4.9 rDNA polymorphism and genetic relationship

The intra-population diversity using the Shannon diversity index for an individual population assessed was H′ = 1.700 as revealed by the primer pair used. PCR amplicons of *T.esculentum* obtained from the primer pair produced 79 scorable bands of which 7 alleles were polymorphic across the 31 individuals of marama bean (Figure 19). The primer pair used showed 8.9% polymorphism.

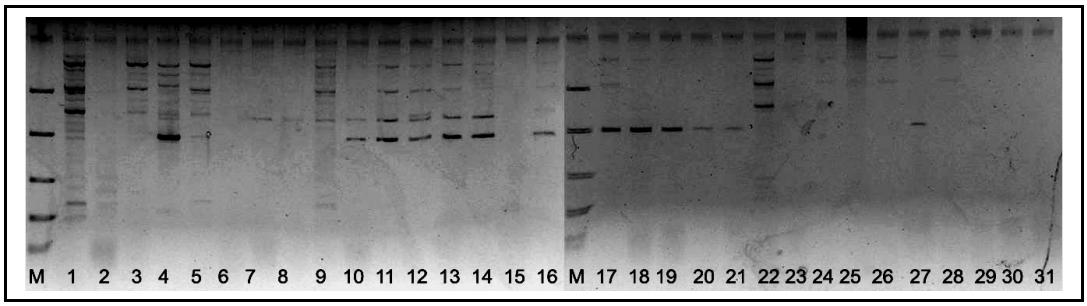

Figure 19: A 1.5 % Agarose gel electrophoresis of the IGS-PCR products for the large marama bean ribosomal RNA genes, M indicates the DNA molecular size marker (Bioline Marker I).

The dendrogram (Figure 20) constructed for the pooled data indicated the existence of four clusters. The first (1) was formed by OMP15, OMP29, OMP06, OMP02, OMP31 and OMP30 at a genetic similarity value of 0.1 with all individuals exhibiting 100% similarity (this is since no bands were seen and therefore they may be considered as a null set). The second (2) has two sub-clusters at a genetic similarity value of 0.3. OMP21 and OMP27 formed the first sub-cluster, whereas OMP19 formed the second sub-cluster. The third (3) cluster formed by OMP14, OMP17, OMP16, OMP01, OMP22, OMP24 and OMP26 at a genetic similarity of 0.63, consisting of seven sub-clusters.

OMP14, OMP09, OMP05 and OMP04 formed the first sub-cluster with the highest molecular variation of 100%. The second sub-cluster was formed OMP17, while OMP16, OMP13, OMP12 and OMP11 formed the third sub-clusters. OMP1 formed the fourth sub-cluster and OMP22 and OMP03 formed the fifth sub-cluster. The sixth sub-cluster was formed by OMP24 and OMP26, whereas OMP25 formed the seventh sub-cluster exhibiting a molecular variation of 100%. The fourth cluster (4) has three sub-clusters at a genetic similarity value of 0.45. Two individuals, OMP08 and OMP07 formed the first sub-cluster, whereas OMP18 and OMP10 formed the third sub-cluster. Only one individual (OMP23) formed the second sub-cluster.

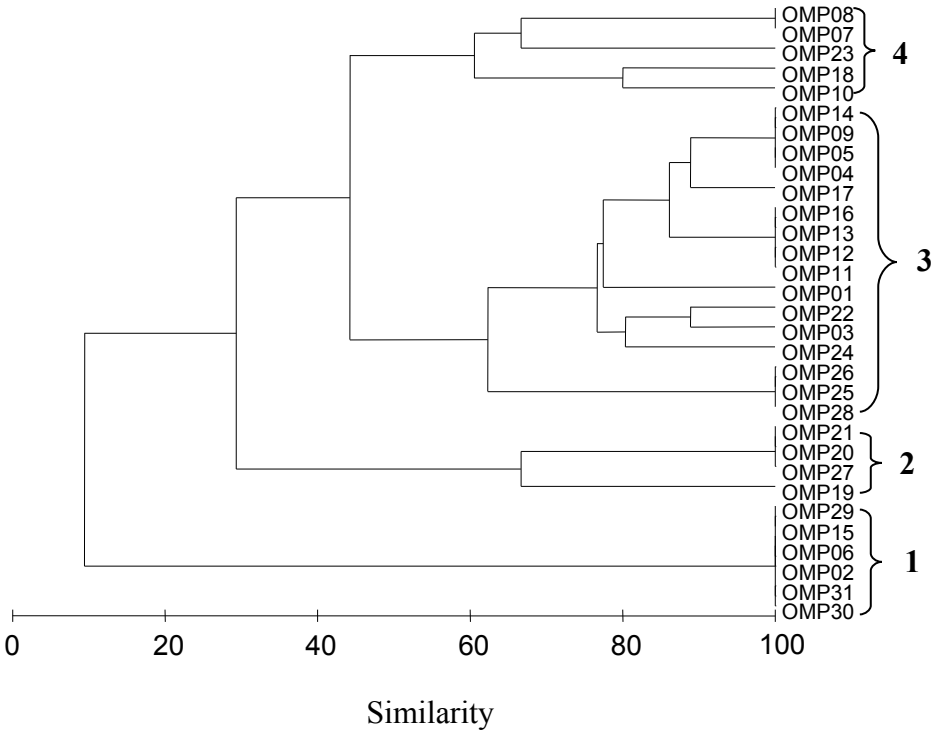

Figure 20: Dendrogram constructed from IGS polymorphism profile from the large ribosomal RNA genes of 31 individuals of Marama bean. The scale at the bottom represents Bray-Curtis similarity coefficients.

CHAPTER 5: DISCUSSION

A comparatively high level of genetic variation was detected in 11 *T. esculentum* populations in the Namibian germplasm using four marama SSR markers, compared with the results AFLP and RAPD (Monaghan and Halloran, 1996, p. 287) obtained from studies of the marama populations. This indicates that marama SSR assay provides a useful tool in estimating genetic variations of wild marama populations. The SSR analysis of the Marama bean populations in Namibia demonstrates an overall genetic variation pattern that is in agreement with the assumption that most of the genetic diversity of Marama bean resided within populations rather than between populations (Monaghan and Halloran, 1996, p. 287 and Naomab, 2005, p. 54).

Microsatellite markers have been used to assess genetic diversity of a large number of plants especially in cultivars such as in rice (Yang, Saghai Maroof, Xu, Zhang and Biyashev, 1995, p. 187), Soybean (Rongwen, Akkaya, Bhagwat, Lavi and Cregan, 1995, p. 44), wheat (Plaschke, Ganal and Roder, 1995, p. 1005), maize (Senior et al., 1998, p. 1088) and cowpea (Dao-Li et al., 2001, p. 193). The number of alleles amplified per primer pair was from 3 to 25 for rice, 11 to 26 for soybean, 3 to 16 for wheat, 2 to 23 for maize, and 2 to 7 for cowpeas. In the present study two to six alleles per primer pair were amplified from the 332 marama individuals among 11 populations. Thus the level of microsatellite polymorphism in marama, although relatively high, it's much lower than other legumes like soybean and cowpeas.

Monaghan and Halloran (1996, p. 289), reported the total number of bands scored per primer in marama ranging from 4 to 12, however, this finding was based on RAPD amplification. It has been suggested that the three mechanisms for creating a new allele at SSR loci are replication slippage (Tachida and Iizuka, 1992, p. 474), unequal crossing-over (Harding, Boyce and Clegg, 1992, p. 849) and genetic recombination. Replication slippage is considered to be a major factor affecting the repeat number for short tandem repeat sequences, whereas unequal crossing-over is thought to result in a very large number of alleles for long tandem repeat arrays. All microsatellites used in the present study were tri-nucleotide repeats (Table 3) hence replication slippage probably plays a major role in creating new alleles at these SSR loci.

Huang, Borner, Roder and Ganal (2002, p. 699) reported that in wheat microsatellites which have short repeat units have lower number of alleles, whereas compound microsatellites produced a very large number of alleles, because they possess very long repeat units. Therefore, the length of tandem repeat units could also be the possible reason for the low level of microsatellite polymorphism in comparison to other plants species especially legumes.

The distribution of alleles within populations was not as expected in random outcrossing populations that started out with equal numbers of the two forms of the alleles at each locus. The founding population in such a case could have been skewed towards the presence of a particular allele that could have been carried in the source population. Hence, the allele frequency of the founder is likely to differ slightly from that in the source population.

This could have led to the founder effect or to these observed dominant type bands (alleles) with high frequencies which might be selected and kept for a species adaptation reasons in the environment it is occurring. Although the PCR conditions were optimized in this study, it was observed that of the four pairs of primers tested, two pairs (MARA065 and MARA077) did not amplify clearly interpretable patterns. For, instance, MARA065 (Figure 9) gave many bands with minimum separation. This could be ascribed to the type of electrophoresis used; of which in this study it was agarose gel electrophoresis. Agarose gels cannot detect alleles in SSRs, because sequences differing by 2 bp could not be resolved using agarose gels (Senior et al., 1998, p. 1086).

Rosetto, Melauchan and Harris (1999, p. 1094) reported that agarose gel and acrylamide gels gave comparable electrophoresis during screening, but it was on a polyacrylamide gel where the banding patterns and resolution were more accurate when the interpretation involved small fragments were being analysed. However, because dinucleotides and tri-nucleotide repeats are more abundant in SSRs and more variable, they also generate stutter bands (*i.e.* smeared bands) during amplification (Schlotterer, 1998). In this study, stutter bands were produced by primer pair MARA077 (Figure 10), regardless of optimizing the PCR conditions. The stutter bands are caused by polymerase slippage during PCR amplification, which results in secondary products containing one or more repeat units less than the primary allelic band (Wang et al., 2009, p. 2434).

This problem is exacerbated in microsatellite loci that have large microsatellites, as the level of stuttering is generally higher in microsatellites with large repeated arrays (O'Reilly and Wright, 1995, p. 30). The stuttering at dinucleotide and tri-nucleotide microsatellite loci can often occlude adjacent alleles. Stutter bands can sometimes equal the intensity of the primary band, making it difficult to accurately characterize genotypes, particularly in population studies where the accurate scoring is essential for parentage determination and family reconstruction from wild populations (Liu and Cordes, 2004, p. 7).

Heterozygosity (H) is a measure of genetic variation in a population, and enlightens how many individuals on average in the population are heterozygotes (Ayala, 1982, p. 99). The amount of genetic variation present in the 11 populations of marama bean was analysed as heterozygosity present in each of the populations and then heterozygosity between populations was determined. The direct count of heterozygosities for individual loci ranged from 0.25 to 0.61 for MARA001 and from 0.33 to 0.93 for MARA068 for the eleven loci in this study, with a mean value for all loci of 0.51 showing a substantial variation within populations (Table 4). This value falls within the range of values being reported for plant SSR studies. For instance Li et al. (2009, p. 2106) reported 0.49 values in their study of 12 falcata (*Medicago sativa* spp. *falcata*) populations. The heterozygosity of a population is the measure of genetic variation preferred by most population geneticists.

It is a good measure of variation because it estimates the probability that two alleles taken at random from the population are different (Mason, Stevens and Jellen, 2005, p. 1617). Study of genetic diversity is the process by which variation among individuals or group of individuals or populations is analysed by a specific method or a combination of methods (Mohammandi and Prasanna, 2003, p. 1236). In this study, the pattern of genetic diversity within and among the Namibian populations of *T.esculentum* was assessed using SSR marker. Genetic diversity is the basis for genetic improvement and knowledge of the genetic variation of these populations is a pre-requisite for successful selection, identifying, conserving and domestication of the marama bean. Knowledge of germplasm diversity has a significant impact on the genetic improvement of any plant species.

SSRs consist of tandemly repeated units of short nucleotide motifs, 1–6-bp long. These regions occur frequently and randomly throughout the genomes of plants and animals, and typically show extensive variation within and among individuals, populations and species (Jarne and Lagoda 1996, p. 437). As multi-allelic single-locus co-dominant genetic markers, both homozygotes and heterozygotes are detected, improving the ease of scoring and the information content per locus (Maquire, Edwards, Saenger and Henry, 2001, p. 281).

This study confirmed the usefulness of SSR in providing a quick and reliable method of assessing genetic variation in plant populations. SSR analysis of genetic diversity within and between natural populations of marama bean illustrated that a higher level of genetic variation is present in marama.

Most of this variation occurs within (66.5%) populations, rather than between (33.5%) marama populations. This finding does fit the general pattern of genetic variation, which Hamrick and Godt (1990, p. 45), Monaghan and Halloran (1996, p. 287), and Naomab (2005, p. 54) reported, even though their remarks were based on isozyme, ALFP and RAPD analysis. Marama is predominantly outcrossing species and is pollinated by a number of insect species and that could be the reason of higher genetic diversity within marama populations. According to Hamrick, Godt and Sherman-Broyles (1992, p. 117), out-crossed species tend to have more within populations genetic diversity than among populations. Hamrick et al. (1992, p. 120) stated that genetic diversity within populations is also influenced by the geographic range of a species. Regionally distributed and widespread species have higher proportions of polymorphic loci, and higher genetic diversity, than geographically limited species.

Nepolo et al. (2009, p. 2092) revealed a widespread distribution of marama bean in Namibia and Watts and Breyer Brandwijk (1962), also revealed that marama bean is widespread, but restricted to clumps distributed in Southern Africa (Botswana, Namibia and South Africa).This also could be a possible reason for a higher within population genetic diversity in marama bean. Furthermore, more widespread species such as marama bean may have historically consisted of large continuous populations that were less susceptible to the random loss of genetic variation, thus leading to a more genetic diversity within populations in marama bean.

Genetic diversity in the present study was determined using Shannon diversity index. The Shannon diversity index (H') measures the degree of uncertainty in finding an electrophoretic pattern given for any individual of a population. In this, indices values close to zero indicate low genetic diversity and values close to one indicate high genetic diversity (King & Schaal, 1989, p. 1118). The mean Shannon index for eleven marama populations assessed in this study ranged from 0.862 to 0.931 with mean total diversity ($H' = 1.526$). It is closer and compares with other studies which have used the Shannon diversity index as a measure of genetic diversity.

For instance, Geleta, Labuschagne and Viljoen (2005, p. 3252) analysed the genetic diversity in sorghum germplasm using SSR markers and found $H' = 0.539$ and Song, Xu, Wang, Chen and Lu (2003, p. 1492) studied the genetic diversity of the wild rice (Oryza *rufipogon*) and found a H' value of 0.919. All of these Shannon diversity indices are lower than the H' estimated in this study. However, Jump and Penuelas (2007, p. 933), found a H' value of 1.593 in beeches (*Fagus sylvativa*), whereas Bao, Chen, Zhang, Cao and Yamamoto (2007, p. 968) found $H' = 2.226$ when they assessed the genetic diversity of pear (*Pyrus. L*), cultivars using SSR markers.

The H' value obtained by Jump and Penuelas (2007, p. 933), however higher it is closer and compare more favourably with the H' value computed in the present study for marama bean. The genetic diversity determined by H' in the present study is an indication that a considerable higher genetic diversity exists in marama bean. This higher genetic variability of marama could have been caused by various factors through the species evolutionary history.

The outcrossing nature of marama, fitness and relevant mutations could have resulted in higher genetic diversity in marama. Marama bean have a single locus outcrossing estimate rate of t =0.77 (Monaghan, 1995), which supports the proposition that marama is predominantly outcrossing, which could be the result of high migration rates between populations. Therefore, outcrossing species populations tend to have higher genetic diversity and fecundity.

Fitness may also bring evolutional changes to a genetic structure of a population. This relative success with which different genotypes transmit their genes to the next generation changes with environmental conditions (Ayala, 1982, p. 95). These changes depend on the fitness in other population of species and on the migration rate of seeds and pollen from neighbouring areas. It is these changes in the genes that bring variations in species. Therefore, fitness could be another possible reason for higher genetic diversity in marama bean.

Mutation even though it is a very slow process, that by itself, changes the genetic constitution of a population at a very slow rate, it's the ultimate source of all genetic variability (Ayala, 1982, p. 74). It arises from changes in the nucleotide sequence of DNA or from deletions, insertions of rearrangements of DNA sequences in the genome. The genome of most eukaryotes like marama bean, contain a very large number of locations of tandem repeats of a number of short nucleotides sequences called simple tandem repeat polymorphism (STRP) (Hartl and Jones, 1998, p. 388).

STRP are found in most organisms and the number of the repeat differs from one chromosome to the next and populations are usually highly polymorphic for the number of repeating units. Moreover, STRP are highly polymorphic and susceptible to errors in replication or recombination that changes the number of repeats or the length of the run (Hartl and Jones, 1998, p. 388). Therefore, any short sequence repeated in tandem a number of times may gain or lose a few copies because of these types of errors. This insertions or deletions of DNA sequence in a species genome give rise to variation species and this could be another possible reason to the observed higher genetic diversity in marama bean.

Gene flow, which is the exchange of genes (in one or both directions) at a low rate between two populations, due to the dispersal of gametes of individuals from one population to another (Ayala, 1982, p. 93) could also lead to high genetic variation. In species with unlimited gene exchange, new mutations have higher probability of being incorporated into additional populations this would lead to more polymorphic loci, more alleles per locus and ultimately higher genetic diversity in the species.

The primer pair with many alleles exhibited higher genetic diversity. The one way ANOVA test performed revealed that there was a significant difference ($p < 0.05$) in the level of genetic diversity of marama populations assessed with different primer pairs. In the present study primer pair MARA077 had the most number of alleles (four alleles) followed by primer pair MARA065 with three alleles. The value of genetic diversity increases with the number of alleles at a given locus (Huang et al., 2002, p. 702).

Significant observation made in this study does fit the statement by Huang et al. (2002, p.703), since higher marama genetic diversity values were evaluated with MARA077 and MARA065. Although, genetic diversity was higher within marama populations, one-way ANOVA analysis revealed that there was no significant difference ($p > 0.05$) in mean genetic diversity between eleven marama bean populations. This implies that marama genetic diversity assessed among populations was homogenous and not hereterogeneous as one would have anticipated.

Despite the geographic separation of marama populations in the study, there has been very little differentiation between the marama populations. This is not unexpected, given the uniformity of the dry semi-arid savannah environments in which they occur. Therefore, the effective population size in marama bean may be sufficiently large to reduce the influence of random genetic drift which is the changes in gene frequencies due to sampling variation from generation to generation (Ayala, 1982, p. 90).

It should be noted however that terms of sampling for analysis of genetic diversity, the law of diminishing marginal returns hold true. The cost of sampling new individuals particularly by means of molecular markers like SSRs is directly proportional to the size of the sample, the probability of detecting an additional allele with each added individual sample decreases rapidly with increasing sample size (Frankel, Brown and Burdon, 1995).

Therefore, in studies aimed at analysis of genetic variation in taxonomic units (populations or species), it is necessary to balance the need to collect as large a sample size as possible against the need to screen as many populations as possible and the need to get allele frequencies from as many loci as possible.

Gene flow does not only result in these two activities i.e. seed movements, either by grazing animals or people moving from region to region with harvested seed it may also occur by other mechanisms as well (Monaghan and Halloran, 1996, p. 287). Gene flow does not change allele frequency of the whole species, but may change them locally when the allele frequencies in the migrants are different from those in resident individuals. The fact that marama bean occur only in Southern Africa may have limited the gene flow of the species resulting in later differentiation.

The SSR band size verification analysis revealed that the obtained band sizes by SSR primer pairs were caused by SSRs, although in some cases observed bands were found to be either more than 30 bp above or below the expected SSR band size. This was confirmed by some of the sequencing data that revealed that some of these banding patterns were microsatellite sequences found in other plant species such as coffee (*Coffea canephora*). Even though, some sequencing data failed to reveal if they were microsatellites in other organisms especially in plants when blasted, comparisons of their sequences with the sequences from where these primer pairs were designed showed the existence of similar repeats that are present in the SSR. This is an indication that, the observed banding patterns were caused by SSRs and not by insertions that have nothing to do with SSRs.

The assessment of length variation in marama bean has disclosed a high intraspecific variation in the nuclear ribosomal DNA repeat units in marama bean. Genetic diversity which was determined using Shannon diversity index was H'= 1.700, is one indication of a relatively high level of intra- specific genetic diversity in Marama populations. This finding supports the proposition from other studies that marama is a predominantly out-crossing species (Monaghan and Halloran, 1996, p. 289) and higher genetic diversity is expected within a population. Genetic diversity from the binary scores of primer pairs analysed as percentage polymorphism revealed a low (0.089%) percentage of polymorphism within marama rDNA unit length.

Since, the length of the coding region is highly conserved in plants, variation in rDNA repeat is attributed to variation in the intergenic spacer region, this finding fits that of Pillay and Myers (1999, p. 1884), that revealed low level of genetic diversity in the rDNA gene of cotton (*Gossypium* spp). The relatively low variation of the rDNA in marama may be ascribed to concerted evolution. Concerted evolution tends to reduce variability between gene copies by mechanisms such as gene conversion and unequal crossing over (Ohta and Dover, 1984, p. 515).

Some studies have invoked mechanisms such as natural selection (Rocheford, 1994, p. 544), genetic drift accompanied by limited gene flow, small population size, and geographical isolation (Jellen, Philips and Rines, 1994, p. 24) as factors associated with reduced variability of the rDNA spacer. Nothing is known about the cause of the marama bean rDNA variation. Generally, it is in the IGS where most variation exists when comparing different isolates, species or genera of organisms.

Variation assessed in this study may be due to the presence of insertions or deletions as reported elsewhere (Kohn, Petsche, Bailey, Novak and Anderson, 1988, p. 1049). The variation in the IGS discriminated marama individuals, making the assessment of ribosomal intergenic length variation a promising marker system in marama bean. Since, these individuals were collected from the same geographic area the lack of differentiation between the marama bean individuals is not unexpected, given the uniformity of the semi-arid savannah environment in which they grow (Monaghan and Halloran, 1996, p. 290). The variability in the number of these length units from one IGS to another has been attributed to unequal crossing over between sub-repetitive units either on the same chromosome or its homology, giving rise to differentiation.

This study is the first to document genetic variation present in marama ecotypes in the Namibian germplasm using SSR markers. Therefore, the evaluation of genetic variation for marama is important for the process of the best strains identification, for the selection of the genetically diverge parent plants for marama genetic mapping and domestication. The study of genetic diversity based on SSRs is a suitable method for distinguishing marama ecotypes.

CHAPTER 7: CONCLUSION

The patterns of genetic diversity within and among the Namibian populations of Marama bean were assessed using SSR marker. The present genetic diversity study revealed that a considerable amount of genetic diversity exists in Namibian marama germplasm and most of this genetic diversity exists within populations rather than between populations. The assessed genetic diversity level varied among populations, which could be due to different environmental conditions in which they are growing. The assessment of spacer length variation and rDNA polymorphisms in the RNA genes in marama bean also revealed that this is a useful region for genetic variability studies in marama bean. Higher similarity was observed in marama populations, but low similarity was also observed among the populations which could indicate that there is a high level of genetic diversity among marama populations. The observed banding patterns in this study were caused by SSRs, as it was confirmed by the study conducted to verify this observation. The findings of this study revealed that microsatellites could be good tools for studying marama bean genetic diversity. While, there is no other comparative studies of SSR genetic diversity at the individual and populations level within and among natural populations of the Marama bean, this findings are matched by other comparative studies of this genetic markers in agricultural systems. This suggests that this finding may well be general. The data reported in this study are the first molecular studies that begin to provide the necessary data that can be used to understand the extent and distribution of Marama bean genetic variations.

CHAPTER 6: RECOMMENDATIONS

Marama bean is regarded as having a considerable potential as an arable crop for arid and semi-arid regions. Marama beans have notable prospects mainly on the basis of its nutritional composition which is comparable with soybean and peanut, the two mostly consumed legumes. Despite marama being a plant with prospects, it is still a wild plant with vast uncertainties to answer before it can be cultivated on any scale and need to be attacked on several knowledge fronts. Focused research and development efforts are needed if this wild plant species can be raised from obscurity and improved sufficiently to contribute to the food supply in Namibia and Sub-Saharan Africa in general. This comprehensive lack of knowledge is the major limitation to its cultivation with information needed on its adaptability to cultivation, genetic variation and on all aspects of its agronomy.

Trials are needed for learning how to manage marama as an arable food crop for Africa. Marama should be tested in projects aimed at alleviating rural poverty and malnutrition in the drought semi-arid areas of Southern Africa. Additionally, marama need to be protected since its wild existence offers a wealth of different germplasm that are being exterminated in many areas. Marama seeds are being relentlessly harvested for village consumption and for sale. Overgrazing poses another threat on its survival as livestock and wild animals eagerly devour the plant's leaves and runners.

Therefore, documented and approved germplasm collections need to be made immediately and desirable strains need to be selected. Research on the productivity and yield measure or estimate need to be done as this will help on the selection of the best strains. Vigorous strains that produce large numbers of pods, bigger seeds, or more seeds per pod and higher protein content could change the whole picture of marama bean future. Furthermore, strains with higher yield especially under adverse conditions should be sought in the harshest sites.

The wisdom or indigenous knowledge of the Kalahari people on the use of marama needs to be compiled in an easily and accessible way. As traditional activities are abandoned in favour of agricultural, pastoral and industrial activities, their intimate knowledge of the bean and its use is being lost. Such information needs to be made accessible to plant scientists and others especially those people who can help foster the progress of marama bean. Genetic improvement of marama bean is of paramount importance because at present the plant produces too few seeds per pod for the fraction of the field they occupy.

APPENDICES

AAPENDIX A

One way ANOVA test for the mean genetic diversity differences among 11 marama populations and four primer pairs

ANOVA

VAR00002					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.242	10	.024	.275	.983
Within Groups	2.909	33	.088		
Total	3.151	43			

Test of Homogeneity of Variances

VAR00002

Levene Statistic	df1	df2	Sig.
1.271	3	40	.297

ANOVA

VAR00002					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.364	3	.788	39.270	.000
Within Groups	.803	40	.020		
Total	3.166	43			

APPENDIX B

Similarities dendograms for all 332 marama individuals produced by four primer pairs

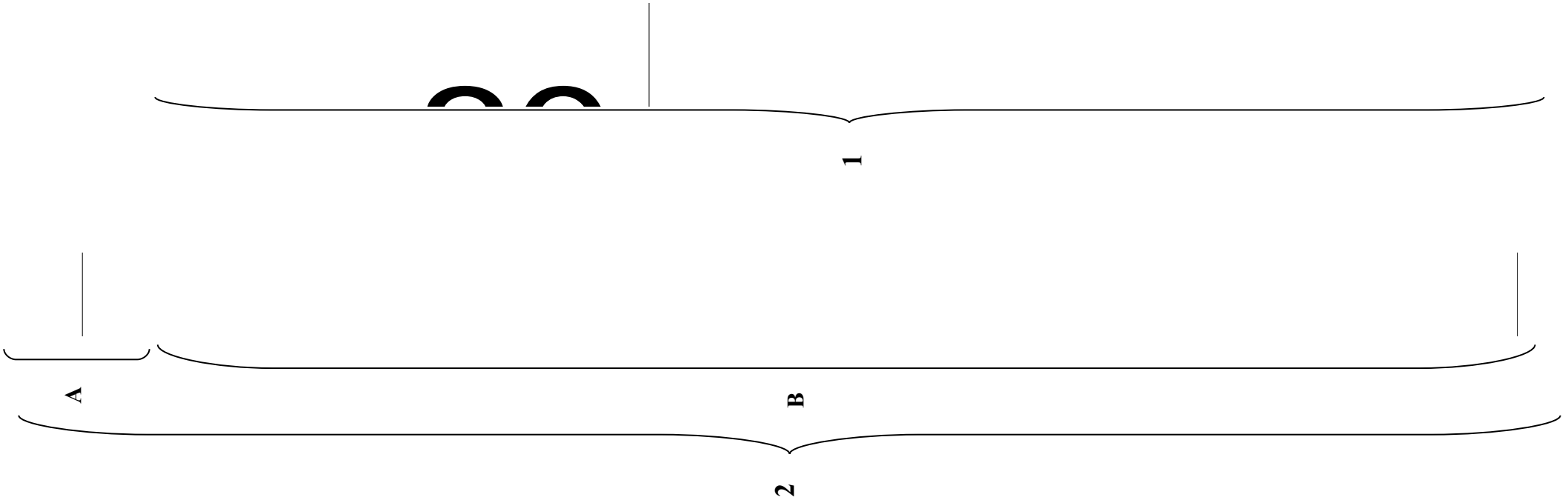

Figure 1: UPGMA dendogram constructed from MARA001 SSR profiles of all 332 individuals from 11 populations of *T. esculentum*.

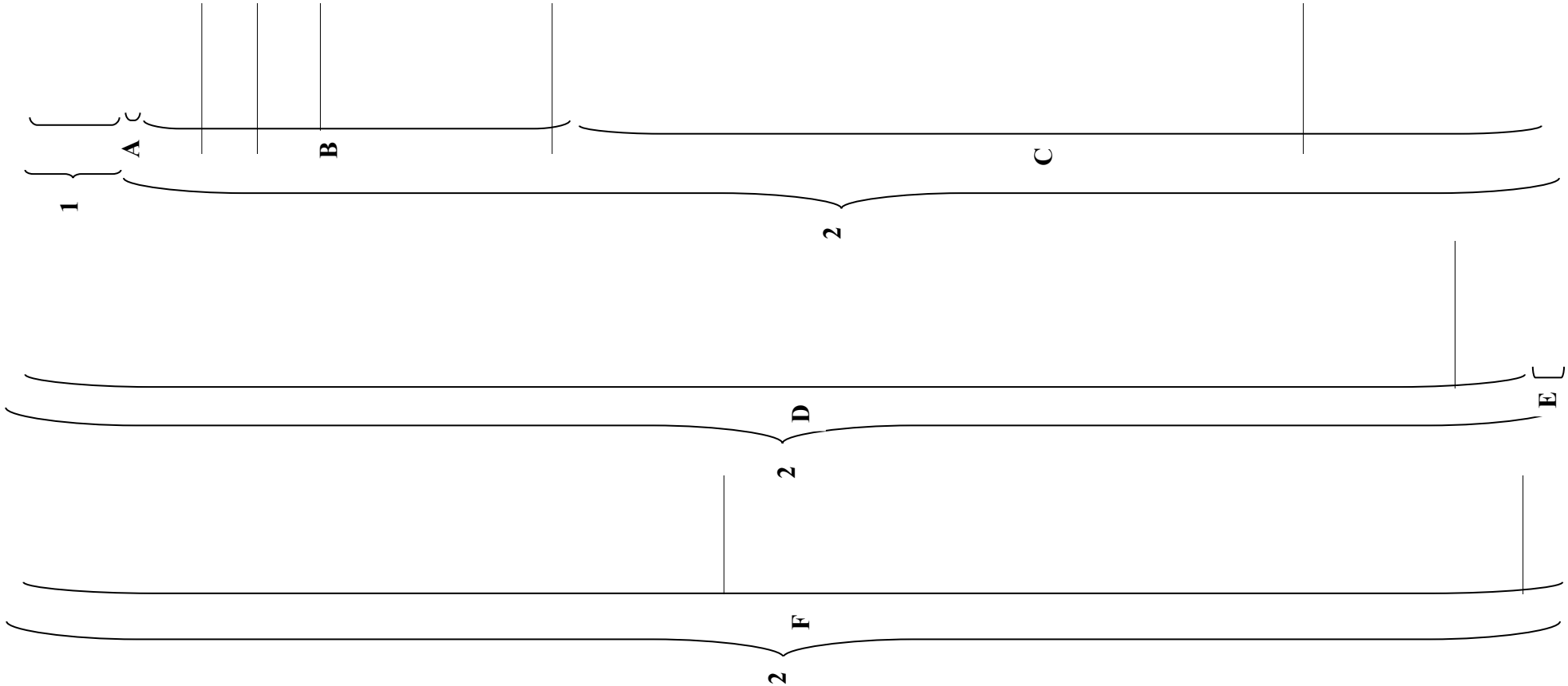

Figure 2: UPGMA dendogram constructed from MARA065 SSR profiles of all 332 individuals from 11 populations of Marama bean.

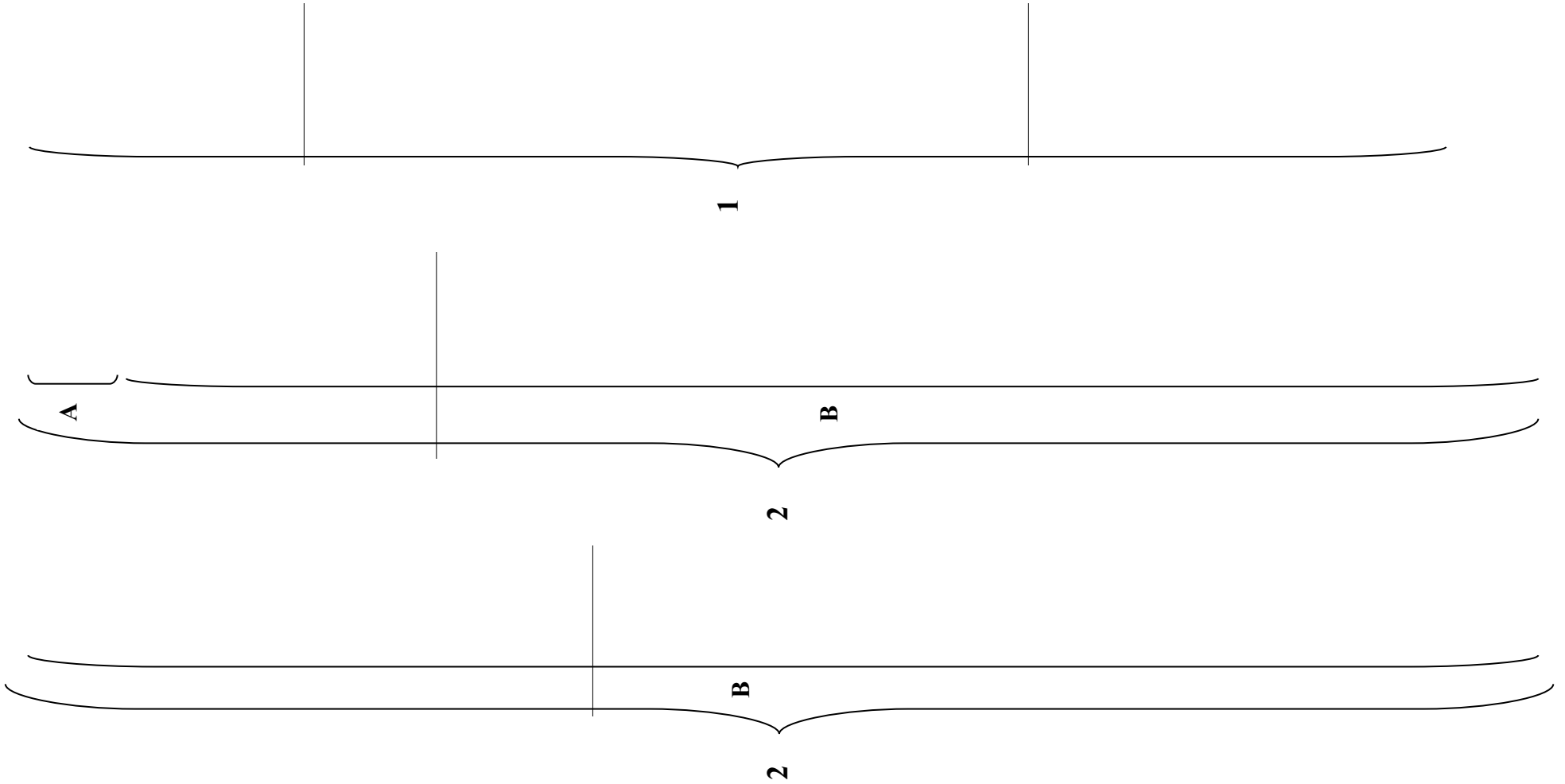

Figure 3: UPGMA dendogram constructed from MARA068 SSR profiles of all 332 individuals from 11 populations of Marama bean.

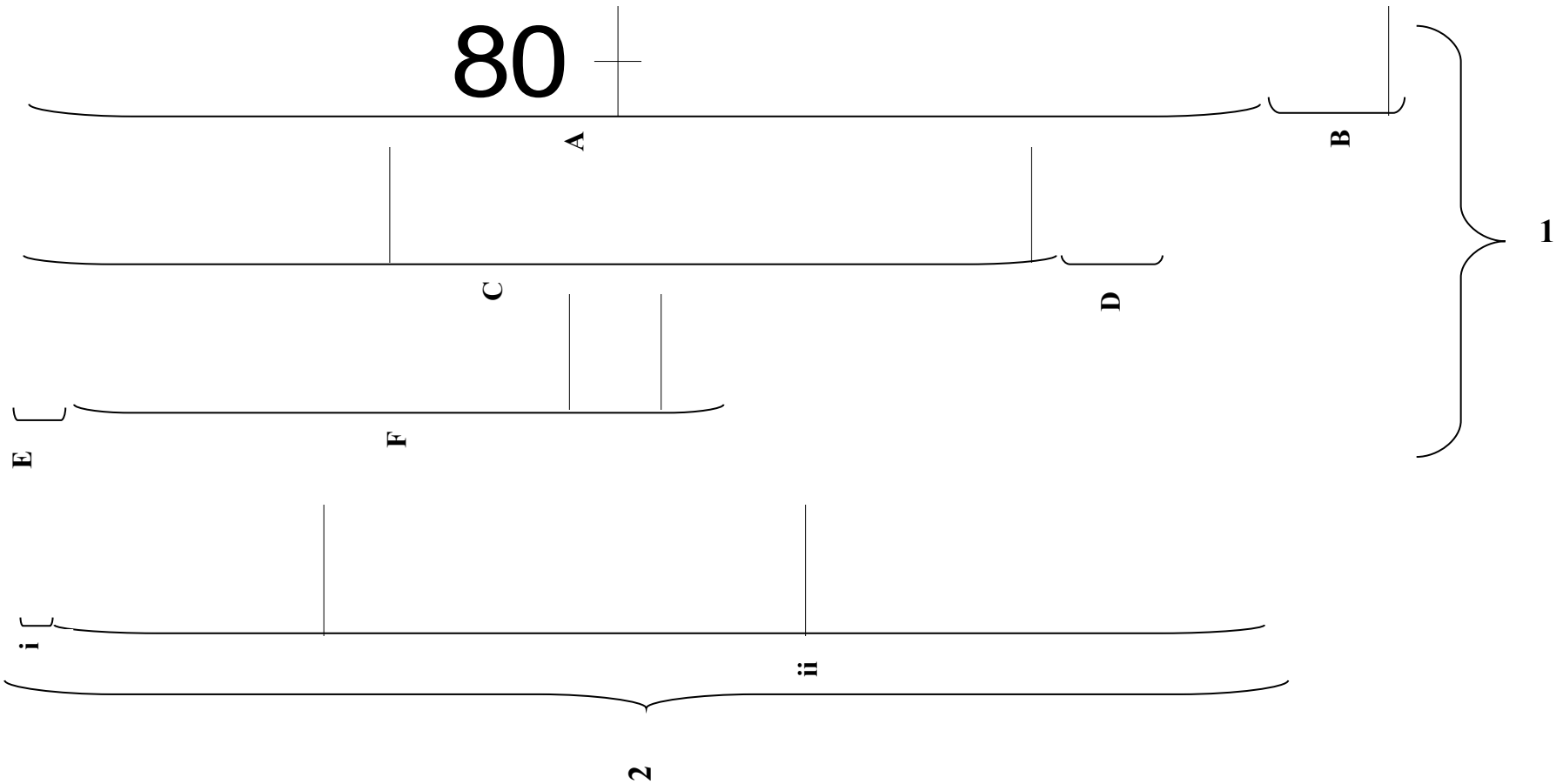

Figure 4: UPGMA dendogram constructed from MARA077 SSR profiles of all 332 individuals from 11 populations of Marama bean.

REFERENCES

Agarwal, M.; Shrivastava, N. and Padh, H. (2008). Advances in molecular marker techniques and their application in plant sciences. *Plant Cell Rep.*, **27**, 617 631.

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

Althoff, D.M.; Gitzendanner, M.A. and Segraves, K.A. (2007). The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. *Syst. Biol.* **56**, 477-484.

Arnold, C., Rosetto, M., McNally, J. and Henry, R.J. (2002). The application of SSR's characterized for Grape (*Vitis vinifera*) to conservation studies in *Viticea*. *American Journal of Botany*, **89**, 22-28.

Ayad, W.G.; Hodking, A.; Jaradat, A. and Rao, V.R. (1995). Molecular genetic techniques for plant genetic resources. IPGRI Workshop: Rome, Italy, October 9-11, 1995.

Ayala, F.J. (1982). Population and evolutionary genetics*: A primer*. Benjamin Cummings, California.

Azam-Ali, S.N. (1996). Bambara groundnut: How did we get here? Proceedings of the International Bambara Groundnut Symposium, University of Nottigham,UK.

Barcaccia, G.; Alberini, E. and Rossellini, D. (2000). Inheritance and mapping of 2n egg production in diploid alfalfa. *Genome,* **43**, 528-537.

Bao, L., Chen, K., Zhang,D., Cao, Y. and Yamamoto, T.Y. (2007). Genetic diversity and similarity of pear (*Pyrus* L.) cultivars native to East Asia revealed by SSR (Simple Sequence repeat) markers. *Genetic Resources Crop Evolutions*, **54**, 959- 971.

Blair, M. W., Giraldo, M.C., Buendia, H.F., Towar, E., Dugue, M.C. and Beebe, S.E. (2006). Microsatellite markers in common bean (*Phaseolus vulgaris* L.). *Theoretical Applied genetics*, **113**, 100-109.

Bardakci, F. (2001). Random amplified polymorphic DNA (RAPD) markers. *Turkish Journal of Biology,* **25**, 185-196.

Bower, N., Hartley, K., Oh, J., and Storey, R. (1988). Nutritional evaluation of Marama Bean (*Tylosema esculentum, Fabaceae*): Analysis of the seed. *Economic Botany*, **42**, 533-540.

Brown, S.M., Hopkins, M.S. and Mitchell, S.E. (1996). Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics,* **93**, 190 198.

Bruford, M.V. and Wayne, R. K. (1993). Microsatellite and their application to population genetics studies. *Current Opinions in Genetics and Development*, **3**, 939-943.

Brunel, D. (1994). A microsatellite marker in *Helianthus annus* L. *Plant Molecular Biology*, **24**, 397-400.

Callow, J.A., Ford-Lloyd B.V. and Newbury, H.J. (1997). Biotechnology and plant genetic Resources: conservation and use. CAB International, Wallingford.

Carvalho, M.A. (2004). Germplasm Characterization of *Arachis pintoi* Karp.and Greg. (*Leguminosae*). A DISSERTATION presented to the Graduate School of the University of Florida in partial fulfilment of the requirements for the degree of Doctor of Philosophy, University of Florida.

Castro, S., Silveira, P., Pereira-Coutinho, A. and Figueiredo E. (2005). Systematic studies in *Tylosema* (Leguminosae). *Botanical Journal of Linnaeus Society*, **147**, 99-115.

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

Dakora, F.D., Lawlor, D.W.and Sibiuga, K.P. (1999). Assessment of symbiotic Nitrogen Nutrition in Marama Bean (*Tylosema esculentum* L.) a tuber producing underutilized African Grain Legume. *Symbiosis*, 27(**4**), 269-277.

Dao Li, C., Fotokun, C.A., Ub, B., Singh, B.B. and Scoles, J.G. (2001). Determining genetic similarities and relationships among Cowpea breeding lines and cultivars by microsatellite markers. *Crop Science*, **41**, 189-197.

Diwan, N., Bhagwat, A.A., Bauchan, G.R. and Cregan, P.B. (1997). Simple sequences repeat (SSR) DNA markers in alfalfa and perennial and annual Medicago species. *Genome,* **40**, 887-895.

Dolanska, L. and Čurn, V. (2004). Identification of white clove (*Tri folium ripens* L.) cultivars using molecular markers. *Plant, Soil and Environment*, **50**, 95-100.

Doyle, J.J. and Beachy, R.N. (1985). Ribosomal gene variation in the soybean (*Glycine*) and its relatives. *Theoretical and Applied Genetics,* **70**, 369– 376.

Dreher, K., Morris, M., Khairallah, M., Ribaut, J.M., Pandey, S. & Srinivasan, G. (2000). *Is marker-assisted selection cost-effective compared to conventional plant breeding methods?* **The case of quality protein maize.** Paper presented at the 4th ICABR Conference on Economics of Agricultural Biotechnology, Ravello, Italy, 24-28 August 2000.Retrieved June 12, 2008, from http://www.abneta.org/

Dvorak, J. and Appels, R. (1982). Chromosome and nucleotide sequence differentiation in genomes of polyploid *Triticum* species. *Theoretical and Applied Genetics*, **63**, 349–360.

Ford-Lloyd, B.V. (2001). Genotyping in plant genetic resources. *In* R.J. Henry(ed.). Plant genotyping: the DNA fingerprinting of plants. CABI Publishing: Lismore, Australia.

Frankel, O.H., Brown, A.H.D. and Burdon. J.J. (1995). The conservation of plant biodiversity. Cambridge University Press, Cambridge, England.

Frary, A. and Douanlar, J. (2003). Comparative genetics of crop plant domestication and evolution. *Turkish Journal of Agriculture*, **27**, 59- 69.

Geleta, N., Labuschagne, M.T. and Viljoen, C.D. (2006). Genetic diversity analysis in Sorghum germplasm as estimated by AFLP, SSR and morpho-agronomic markers. *Biodiversity and Conservation*, **43**, 3251- 3265.

George, J., van Zijll de Jong, E., Cogan, N.O.I. and Forester, J.W. (2006). Assessment of genetic diversity in cultivars of white clover (*Trifolium repens* L.) detected by SSR polymorphisms. *Genome,* **49**, 919-930.

Gerstner, J. K., Schiebel, von Waldberg, G. and Hembleben, V. (1988). Complex organization of the length heterogeneous external spacer of mug bean (*Vigna radiata*) ribosomal DNA. *Genome*; **30**, 723–733.

Gupta, A. K. (2004). Origin of Agriculture and domestication of plants and animals linked to Holocene climate amelioration. *Current Science*, **87** (1), 54-59.

Hamby, R.K and Zimmer E.A. (1992). Ribosomal RNA as phylogenetic tools in plant systematic. Chapman and Hall Publisher, London.

Hamrick, J.L., & Godt, M.J. (1990). Allozyme diversity in plant species. In A. Brown, M.C.Clegg, A. Kahler and B. Weir (Eds.), Plant population genetics, breeding and genetic resources (pp. 43 -63). Sinaurer Associates Inc., Mass.

Hamrick, J.L., Godt, J.OW. and Sherman-Broyles, S.L. (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests*, **6**, 95- 124.

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

Harlan, J.R. (1992). Crop and Man. *American Society of Agronomy and Crop Science society of America*, Madison, USA.

Hartl, D.L. and Jones, E.W. (2001). Genetic: analysis of genes and genomes.(5^{th}) Jones and Bartlett publisher. USA.

Hartley, L.M. (1997). Breeding systems and generic variation in *Tylosema esculentum* (PHD) thesis. Falmer. Sussex University.

Hartley, L.M., Tshamakeng, E. and Thomas, S.M. (2002). Functional heterostyly in *Tylosema esculentum* (Caesalpinioideae). *Annal Botany*, **89**, 67- 76.

Huang, X.O., Borner, A., Roder, m.S., and Ganal, W. M. (2002). Assessing genetic diversity of wheat (*Triticum aestivum* L.) using microsatellite markers. *Theoretical Applied Genetics*, **105**, 699-707

Jarne, P., & Lagoda, P.J.L. (1996). Microsatellites from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424-429.

Jellen, E.N., Philips, R.L. and Rines, H.W. (1994). Chromosomal localization and polymorphisms of ribosomal DNA in oat (*Avena* spp). *Genome*, **37**, 23–32.

Jump, A.S. and Penuelas, J. (2007). Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind-pollinated tree, *Fagus sylvativa*. *Molecular Ecology*, **16**, 925- 935.

Karp, A., Seberg, O. and Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. *Annals of Botany,* **78**, 143-149.

Keegan, A.B., & van Staden, J. (1981). Marama bean *Tylosema esculentum,* a plant worthy of cultivation. *South African Journal of Botany,* **77**, 387-388.

Ketshajwag, K, K., Holmback, J, Yeboah, S.O. (1998). Quality and compositional studies of some edible leguminosae seed oils in Botswana. *Journal of the American Oil Chemist's Society*, **75,** 741 -743.

King, L.M. and Schaal, B.A. (1989). Ribosomal DNA variation and distribution in *Rudbeckia missourensis*. *Evolution*, **43**, 1117-1119.

Kirti, P.B. (2008). Handbook of new technologies for genetic improvements of legumes. Technology and Engineering. CRC Press, United Kingdom.

Knapik, E.W., Goodman, A., & Ekker,M. (1998). A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genetics,* **18**, 338 -343.

Kohn, L.M., Petsche, D.M., Bailey, S.R., Novak, L.A and Anderson, J.B. (1988). Restriction fragment length polymorphisms in nuclear and mitochondrial DNA of *Sclerotinia* species. *Phytopathology*; **78**, 1047–1051.

Kollipara, K.P., Singh, R. J. and Hymowitz, T. (1997). Phylogenetic and genomic relationships in the genus *Glycine* Willd, based on sequences from the ITS region of nuclear rDNA. *Genome*, **49**, 57-68.

Korzun, V., Malyshev, S., Voylokov, A.V. and Borner, A. (2001). A genetic map of rye (*Secale cereale* L.) combining RFLP, isozyme, microsatellite and gene loci. *Theoretical and Applied Genetics,* **102**, 709-717.

Krishna. G.K., Zhang. J., Burow,M., Pittman, R., Delikostadinov, S.G., Lu, Y. and Puppala,N. (2004). Genetic diversity analysis in valencia peanut (*Arachis hypogaea* L.) using microsatellite markers. *Cellular and Molecular Biology Letters*, **9**, 685- 697

Li, P., Wang, Y., Sun, X. and Han, J. (2009). Using microsatellite (SSR) and morphological markers to assess the genetic diversity of 12 falcata (*Medicago sativa* spp. falcata) populations from Eurasia. *African Journal of Biotechnology,* **8** (10), 2102- 2108.

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

Liu, Z. J. and Cordes, J. F., (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, **238**, 1–37.

Liu, Z.W., Biyashev, R.M. and Saghai-Maroof, M.A. (1996). Development of simple sequence repeat markers and their integration into a barley linkage map. *Theoretical Applied Genetics*, **93**, 869-876.

Liu, K.D., Yang, G.P., Zhu, S.H., Zhang, Q., Wang, X.M. and Saghai-Maroof, M.A. (1996). Extraordinarily polymorphic ribosomal DNA in wild and cultivated rice. *Genome*, **39**, 1109–1116.

Maguire, T.L., Edwards, K.J., Saenger, P. and Henry, R. (2001). Characterization and analysis of microsatellite loci in mangrove species, *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae). *Theoretical Applied Genetics*, **101**, 279-285.

MacDonald, G.M. (2003). Biogeography: Space, time and life. John Wiley, New York.

Mason, S.L., Stevens, M.R. and Jellen, E.N. (2005). Development and use of microsatellite markers for germplasm characterization in Quinoa (*Chenopodium quinoa* Willd.). *Crop Science*, **45**, 1618-1630.

Mitchell, R.A.C., Keys, A.J., Madgwick, P.J., Parry, M.A.J. and Lawlor, D. W. (2005). Adaptation of photosynthesis in marama beans *Tylosema esculentum* (Burchell A. Schreiber) to a high temperature, high radiation, drought-prone environment. *Plant physiology and Biochemistry,* **43**, 969-976.

Mmnonatau, Y (2005). Flour from the Marama bean: composition and sensory properties in a Botswana perspective, University of Stellenbosch, South Africa.

Mohammandi, S.A. and Prassana, B.M. (2003). Analysis of genetic diversity in crop plants- Salient Statistical tools and Considerations. *Crop Science*, **43**, 1235 1248.

Moholo, D.B.G. and Amarteifio, J.O. (1998). The chemical composition of four legumes consumed in Botswana. *Journal of food composition analysis,* **11**, 329-332.

Monaghan, B.G., & Halloran, G.M. (1996). RAPD variation within and between natural populations of morama [*Tylosema esculentum* (Burchell) Schreiber] in southern Africa. *South African Journal of Botany,* **62**, 287-291.

Monaghan, B.G. (1995). Genetic variation in morama bean (*Tylosema esculentum*) M.Sc (Agri) thesis, University of Melbourne, Melbourne.

Mondini, L., Noorani, A and Paguotta, M.A. (2009). Assessing plant genetic diversity by molecular tools. *Diversity*, **1**, 19- 35.

Mullis, K.B.; Faloona, F.A.; Scharf, S.J.; Saiki, S.K.; Horn, G.T. and Erlich, H.A. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor. Symp. Quant. Biol.* **51**, 263-273.

Takundwa. M.M. (2009). Development of microsatellite (SSR) and DNA barcoding (MatK) primers for marama bean [*Tylosema esculentum* (Burchell) Schreiber] for germplasm characterization of the Namibian genotypes (in press) M.Sc thesis submitted to the University of Namibia for a M.Sc degree.

Naomab, E. (2004). Assessment of genetic variation in natural populations of Marama bean (*Tylosema esculentum*) using molecular markers. M.Sc thesis. University of Namibia. Retrieved March 15, 2008, from University of Namibia Special Collection Library Section.

National Academy of Science (1979). Tropical legumes; resources for the future. National Research Council, Washington, DC.

National Research Council (2006). Lost Crops of Africa: Volume II, Vegetables. National Academies Press, Washington, DC, USA.

Nepolo, E., Takundwa. M., Chimwamurombe. M., Cullis. C. A. and Kunert. K. (2009). A review of geographical distribution of marama bean [*Tylosema esculentum* (Burchell) Schreiber] and genetic diversity in the Namibian germplasm. *African Journal of Biotechnology*, Vol. 8 (**10**), 2088-2093.

Nickrent, D.L and Patrick, J.A. (1998). The nuclear ribosomal DNA intergenic spacers of wild and cultivated soybean have low variation and cryptic subrepeats. *Genome*, **41**: 183–191.

Ohta, T and Dover, G.A. (1984). The cohesive population genetics of molecular drive. *Genetics,* **108**, 501–521.

E.Z.N.A Plant DNA kit (2008). Omega Bio-tek. USA.

O' Reilly, P. and Wright, J.M. (1995). The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *Journal of Fish Biology*, **l47**, 29-55.

Pillay, M. and Myers, G.O. (1999). Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. *Crop Science*, **39**, 1881-1886.

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

Polanco, C. and Pérez de la Vega, M. (1997). Intergenic ribosomal spacer variability in hexaploid oat cultivars and landraces. *Heredity*, **78**, 115–123.

Powell, M.A. (1987). Marama Bean (*Tylosema esculentum,Fabaceae*).Seed crop in Texas. *Economic Botany,* **41**, 216 -220.

Primack, R.B. (1993). Essentials of Conservation Biology. Sinauer Associates: Sunderland, MA, USA.

Primack, R.B. (2004). A primer of Conservation Biology. 3rd edition. Sinauer Associates: Sunderland, MA, USA.

PRIMER 5 (Plymouth Routines In Multivariate Ecological Research) for Windows (2001) Version 5.2.0. Primer-E Ltd, United Kingdom.

Qiagen. (2006). DNeasy Plant Handbook. Qiagen, New York, pp. 24-27.

Richard, I. and Beckman, J.S. (1995). How neutral are synonymous codon mutations? *Nature Genet*ics, ***10*****,** 259- 265.

Rocheford, T.R. (1994). Changes in ribosomal DNA intergenic spacer length composition in maize recurrent selection populations. I. Analysis of BS13, and BSCB1. *Theoretical and Applied. Genetics*, **88**, 541–547.

Roder, M.S., Plaschke, J. and Konig, S.U. (1995). Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular and General Genetics*, **246**, 327-333.

Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U. and Cregan, P.B. (1995). The use of microsatellite DNA markers for soybean genotype identification. *Theoretical Applied Genetics*, **90**, 43-48.

Rosseto, M., Melauchan, A. and Harris, F.C.L. (1999). Abundance and polymorphism of microsatellite markers in the tea tree (*Melaleuca alternifolia*, Myrtaceae). *Theoretical and Applied Genetics*, **98**, 1091-1098.

Santana, Q.C., Coetzee, M.P.A. and Steenkamp, E.T. (2009). Microsatellite discovery by deep sequencing of enriched genomic libraries. *Biotechniques*, **46**, 217 -223.

Schlotterer, C. (1998). Microsatellites. In: Hoezel, A.R. (ed) Molecular genetic analysis of populations: A practical approach. Oxford University Press, New York, USA.

Schlötterer, C. and Tautz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, **20**, 211–215.

Senior, M.L., Murphy, J.P., Goodman, M.M. and Stuber, C.W. (1998). Utility of SSRs for determining genetic similarities and relationships in maize using an agarose system. *Crop Science* **38**, 1088-1098.

Stephen, B. B. (2002). Genes in the field: On-farm conservation of crop diversity. Lewis Publishers, USA.

Singh, A., Deverumath, R.M., RamaRao, S., Singh, V.P., and Raina, S.N. (2008). Assessment of genetic diversity, and phylogenetic relationships based on ribosomal DNA repeat unit length variation and Internal Transcribed Spacer (ITS) sequences in chickpea (*Cicer arietinum*) cultivars and its wild species. *Genetic Resources and Crop Evolution* , **55**, 65–79.

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

Sobrino, B.; Briona, M. and Carracedoa, A. (2005). SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Sci. Int*, **154**, 181-194.

Song, P. Z., Xu, X., Wang, B., Chen, J.K. and Lu, B.R. (2003). Genetic diversity in the northernmost *Oryza rufipogon* populations estimated by SSR markers. *Theoretical Applied Genetics*, **107**, 1492-1499.

Strand, M., Prolla, T.A., Liskay, R.M. and Petes, T.D. (1993). Destabilisation of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*, **365**, 274- 276.

Sunyaev, S.; Hanke, J.; Aydin, A.; Wirkner, U.; Zastrow, I.; Reich, J. and Bork, P. (1999). Prediction of non-synonymous single nucleotide polymorphisms in human disease-associated genes. *Journal of Molecular Medicine*, **77**, 754 760.

Szewc-McFadden, A.K., Lamboy, W.F., Hokanson, S.C and McFerson, J.R. (1996). Utilization of identified simple sequence repeats (SSRs) in *Malus domestica* (apple) for germplasm characterization. *Horticulture Science*, **31**, 619- 626.

Tachida, W. and Iizuka, M. (1992). Persistence of repeated sequences that evolve by replication slippage. *Genetics,* **131**, 471–478.

Van der Maesen, L.J.G. (2006). *Tylosema esculentum* (Burch.)A. Schreib. In Brink, M. and Belay, G. (Editors).PROTA 1: Cereals and pulses/ Cereales et legumes secs. Retrieved April 10, 2008, from PROTA, [CD-Rom]. Wageningen, Netherlands.

Verdoorn. I.C. (1959). *Bauhinia esculenta*. The flowering plants of Africa. 33: sub t. 1311.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M.(1995). AFLP: a new techniques for DNA fingerprinting. *Nucleic Acid Research,* **23**, 4407 – 4414.

Wang, X., Rinehart,T.A., Wad, P.A., Spiers, J.M., Hadziabdic, D., Windham, M.T and Trigiano, R.N. (2009). A new electrophoresis technique to separate microsatellite alleles. *African Journal of Biotechnology*, 8 (**11**), 2432-2436.

Watt, J. M. and Breyer-Brandwijk, M.G. (1962). The medicinal and poisonous plants of southern and western Africa, 2nd edition. E .and S. Livingstone, Edinburgh, London.

Weber, J. L. (1990). Human DNA polymorphisms and methods of analysis. *Curr. Opinions in Biotechnologies,* **1**,166 –171.

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

Westman, A.L. and Kresovich, S. (1997). Use of molecular marker techniques. *In* Callow, J.A., Ford-Lloyd, B.V and , H.J. (ed.).Biotechnology and plant genetic Resources: conservation and use. CAB International, Wallingford.

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

Wolfe, A and Liston, A. (1998). Contribution of PCR based methods to plant systematic and evolutionary biology. In: Soltis D.E., Soltis P.S., Doyle J.J. (eds): Molecular biology of plants II DNA sequencing. Kluwer Acad. Publ., Boston, Dordrecht, London.

Yang, G.P., Saghai Maroof, M.A., Xu, C.G., Zhang, Q.F and Biyashev, R.M. (1995). Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. *Molecular and General Genetics*, **245**, 187-194.

Yu, K., Park, S.J. and Poysa, V. (1999). Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome*, **42**, 27-34.

Zabeau, M. and Vos, P. (1992). Selective Restriction Fragment Amplification: General Method for DNA Fingerprinting. European Patent EP0534858, September 24, 1992.

Zimmer, E.A., Juppe, E.R. and Walbot, V. (1988). Ribosomal gene structure, variation and inheritance in maize and its ancestors. *Genetics,* **120**, 125–1136.