

**Evaluation of Random Amplified Polymorphic  
DNA and Simple Sequence Repeat markers in  
*Moringa oleifera* (Lam.) to establish population  
diversity**

by

**Rynhard Smit**

**Student no: 27074112**

Submitted in partial fulfilment of the requirements for the degree

**MSc Biotechnology**

**In the Faculty of Natural and Agricultural Sciences**

**University of Pretoria**

**Supervisors: Dr. B.J. Vorster**

**Prof. E.S. duToit**

**October 2013**

## **Declaration**

The following document contains the written dissertation by Rynhard Smit, the undersigned, who hereby agrees that this document contains my own and independent work that is herewith submitted for the degree Master Scientia to the University of Pretoria.

**Signature:**

**Date:**

## Acknowledgements

I would like to convey my utmost gratitude to my supervisors, Dr Barend Juan Vorster and Prof. Elsa S. du Toit for their support and input with this project.

Special thanks to Dr. Barend Juan Vorster for his guidance during this project, help in interpreting results from RAPD's, continuous advice to achieve set out experimental goals, frequent follow up and critical reading of the manuscript including comments and suggestions on it.

I am grateful to Prof. Elsa S. du Toit for her continual follow up, concern for future prospects of my work, setting up the infrastructure for the study of Moringa as well as organizing the maintenance of the orchard and critical reading of the manuscript.

I extend my gratitude to both my supervisors for their initiation of this project and being able to contribute to the platform of research to be conducted on Moringa in South Africa. Moreover I am grateful for their financial support and material assistance.

I am very grateful to Miss. Kerry Reid for her assistance with interpretation of microsatellites, training in basic analyzing data scores and some software programmes and continuous suggestions, comments and willingness to assist via email.

I would like to thank the following organizations: University of Pretoria, Department of Plant Sciences and Forestry and Agricultural Biotechnology Institute, for their access to wonderful facilities including laboratories, administration offices, internet, library and computers. The National Research Foundation, for the funding opportunity to facilitate this study. I would also like to acknowledge the Plant Sciences department for their hospitality and support.

My gratitude also extends to my colleagues: Mr Stefan van Wyk, Mr Priyen Pillay, Ms Magdeleen du Plessis, Dr Berhanu Fenta, Dr Sonia Phillips, Ms Mischa Muller for their invaluable support and motivation throughout this project.

Lastly I would like to thank my parents, sister and family for their support, motivation and understanding in helping me exceed in my education.

# Table of contents

Acknowledgement.....	i
Table of contents.....	iii
List of Figures .....	vi
List of Tables.....	viii
List of abbreviations.....	ix
Abstract .....	xi
<b>1 Literature review .....</b>	<b>1</b>
1.1 Description of <i>Moringa oleifera</i> .....	1
1.2 The many uses of <i>Moringa oleifera</i> .....	4
1.3 Population structure and conservational genetics.....	6
1.4 Molecular markers .....	8
1.4.1. RAPD.....	9
1.4.2. Microsatellites.....	11
<b>2 Aims.....</b>	<b>13</b>
<b>3 Materials and methods .....</b>	<b>14</b>
3.1 Plant material.....	14
3.2 DNA extractions .....	14

3.3 DNA quality analysis .....	15
3.4 Selection of primers .....	16
3.4.1. RAPD primers .....	16
3.4.2. Microsatellite primer sets .....	16
3.5 PCR reactions .....	17
3.6 PCR amplification .....	18
3.7 PCR product analysis .....	21
3.8 Data analysis.....	21
3.9 Population variation.....	23
3.9.1. Marker system correlations.....	23
3.9.2. Genetic diversity .....	24
<b>4 Results .....</b>	<b>25</b>
4.1 Variation of RAPD loci .....	25
4.2 Variation of microsatellite loci .....	27
4.3 Combined marker analysis .....	31
4.3.1. Similarity coefficients of combined data .....	31
4.3.2. Principle Coordinate Analysis (PCoA).....	33
4.3.3 Correlation test.....	34
4.3.4 Genetic diversity analysis .....	35

<b>5 Discussion.....</b>	<b>37</b>
5.1 Primer efficiency .....	37
5.2 Marker system correlation .....	39
5.3 Population variation.....	39
<b>6 Conclusion .....</b>	<b>43</b>
<b>References .....</b>	<b>44</b>
<b>Annex.....</b>	<b>49</b>

## List of Figures

- Figure 1.1** - Distribution map of *Moringa oleifera* (Shaded) and different localities of sampling (Rings), including **a.)** Hawaii, **b.)** South Africa and **c.)** India.
- Figure 1.2** - **1.** Photographical illustration of *Moringa oleifera* with **2.** similar aged pods and their diverse phenotype observed within different varieties.
- Figure 4.1** - RAPD profile of 14 Moringa genotypes generated by RAPD primer 5. The lanes represent, lane L (100bp ladder); lanes Id1-Ii7 (Moringa genotypes).
- Figure 4.2** - The UPGMA clustering based on the Jaccard's similarity indices produced from the RAPD markers, clustering the 71 Moringa accessions included in this study, into three major clusters **RI**, **RII** and **RIII**. Note that 1 (similarity) = genetically identical.
- Figure 4.3** - Microsatellite-UPGMA dendrogram generated, based on the genetic similarity indices produced from dominant scores of microsatellite allele data for 71 Moringa accessions. Six major clusters, including **SI**, **SII**, **SIII**, **SIV**, **SV** and **SVI** were identified.
- Figure 4.4** - UPGMA clustering generated from the combination of both marker systems, RAPD and microsatellite, clustering foreign genotypes separately.
- Figure 4.5** - Scatter plot of principle coordinate analysis (PCoA) for 71 Moringa accessions based on data produced by both markers systems. Three grouping were formed, **PI**, **PII** and **PIII**, with the grouping pattern similar to the corresponding UPGMA clustering.



**Figure 4.6** - Genetic diversity values for each collection from their respective origins and cultivars estimated according to Shannon-Weaver ( $I$ ) and Nei's ( $h$ ) diversity calculations.

## List of Tables

- Table 3.1** - Sample identification, origins and cultivar description.
- Table 3.2** - Characteristics for the RAPD markers used in this study, including polymorphisms detected and their percentages.
- Table 3.3** - Properties of the twelve remaining microsatellite primer pairs after screening.
- Table 4.1** - Polymorphic Information Content (PIC) values estimated for the RAPD markers.
- Table 4.2** - Calculated PIC and  $H_{exp}$  values for the microsatellite markers used in this study.
- Table 4.3** - Comparison of the components used for evaluating each marker system within the analyzed data set.

## List of abbreviations

**RAPD** – Rapid Amplified Polymorphic DNA

**SSR** – Simple Sequence Repeats

**DNA** – Deoxyribonucleic acid

**PCR** – Polymerase chain reaction

**UPGMA** - Unweighted Pairwise Group Method with Arithmetic average

**PCoA** – Principle Co-ordinate Analysis

**PIC** – Polymorphic Information Content

**MI** – Marker Index

**S.A.** – South Africa

**AFLP** – Amplified Fragment Length Polymorphism

**ISSR** – Inter Simple Sequence Repeat

**RFLP** - Restriction Fragment Length Polymorphisms

**bp** – base pairs

**CTAB** – Cetyltrimethylammonium bromide

**EDTA** – Ethylenediaminetetraacetic acid

**NaCl** – Sodium chloride

**PVP** – Polyvinylpyrrolidone

**rpm** – Revolutions per minute

**min** – minutes

**nm** – Nanometers

**NCBI** – National Center for Biotechnology Information

**mM** – Millimolar

**ng** – Nanograms

**NH<sub>4</sub>** – Ammonium

**MgCl<sub>2</sub>** – Magnesium chloride

**dNTP** – Deoxyribonucleotide

**seq** – Sequence

**I.d.** – Identification

**TAE** – Tris-acetate-EDTA

**mA** – Milliamps

**EMR** – Effective Multiplex Ratio

**PAGE** – Polyacrylamide gel electrophoresis

## Abstract

*Moringa oleifera* is potentially an economically important tree species. It has gained interest globally for its multipurpose uses, in particular as a source of nutrition and oil, as well as for its various medicinal properties. Moringa is native to India, Malaysia and the Middle East, but have been introduced to many countries throughout Africa. There is however limited knowledge regarding the genetic variation of both native and introduced populations of Moringa, although phenotypic observations suggest the presence of significant genetic diversity. To do this we used Moringa collections from different locations including India, South Africa and Hawaii, with different cultivars present in the foreign samples. Molecular markers such as Random Amplified Polymorphic DNA (RAPD) and Simple Sequence Repeats (SSR) were evaluated for their efficiency as a marker system in Moringa, based on their success in other tropical tree population studies. The comparative analysis between the two revealed that both marker systems identified similar heterogeneity of 0.438 (RAPD) and 0.481 (SSR). However, the microsatellite marker was more effective for assessing genetic diversity, with a marker index (MI) value of 5.77 compared to 2.96 (RAPD), within our data set. Three major groups among the 71 accessions were clustered together in the dendrogram and Principle Co-ordinate Analysis (PCoA), based on the genetic similarity matrices generated by both markers. Both cultivars were grouped together irrespective of origins, suggesting a relationship of genetic identity rather than geographical origins. The markers investigated in this study were found to be effective for determining genetic diversity, verifying higher genetic variation among the S.A. accessions of Moringas and distinguishing them from the cultivars investigated. This information could possibly be exploited for cultivar development in tree improvement programmes.

# 1. Literature review

## 1.1 Description of *Moringa oleifera*

The taxonomic description was retrieved from the Integrated Taxonomic Information System on-line database (IT IS, 2010). The *Moringa oleifera* (Syn. *M. pterygosperma* Gaertn.), also known as the drumstick tree, is one of the thirteen species belonging to the monogeneric family Moringaceae. [41; 80]. The Moringaceae family are related to Brassicaceae and 15 other families that produce mustard oil [14].

**Kingdom** Plantae

**Subkingdom** Viridaeplantae

**Infrakingdom** Streptophyta

**Division** Tracheophyta

**Subdivision** Spermatophytina

**Infradivision** Angiospermae

**Class** Magnoliopsida – Dicotyledons

**Superorder** Rosanae

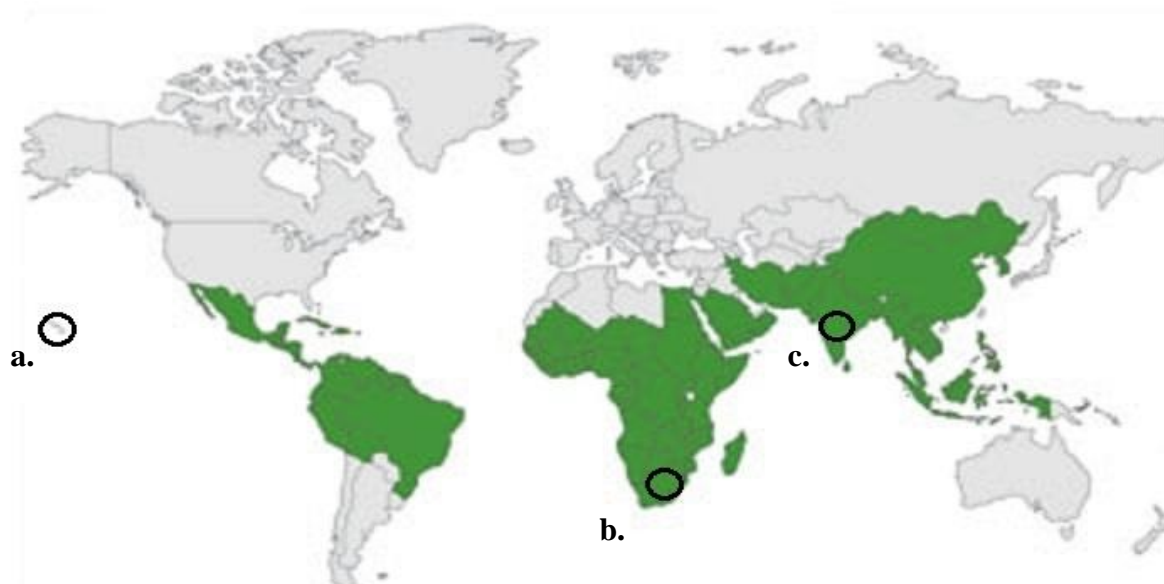
**Order** Brassicales

**Family** Moringaceae

**Genus** *Moringa*

**Species** *oleifera* Lam.

The Moringa tree is indigenous to Northwest India, western and sub-Himalayan tracts, Pakistan and Asia minor, where it has been naturalised and cultivated in Sri Lanka, Southeast Asia, tropical Africa, Central- and South America, Caribbean Islands, the greater of the tropical belt and many other countries (Figure 1.1) [41; 44; 59; 80; 82; 89]. It is capable of growing in a wide range of soil types, from light- to medium textured soils, as well as a wide range of conditions, including semi arid regions, and elevations from sea level to approximately 1400 meters above sea level [82]. Due to their broad distribution, both naturally and cultivated, the tree is well known under many vernacular names, i.e. English – Drumstick tree, Horseradish tree; Arabic – Rawag; Punjabi – Sanjna etc. [80; 82].



**Figure 1.1** – Distribution map of *Moringa oleifera* (Shaded); and different localities of sampling (Rings), including **a** Hawaii **b**. South Africa and **c**. India.

Derived from the vernacular Tamil name, *Moringa oleifera*, Moringa is a small, outcrossing diploid ( $2n = 28$ ) that reaches sexual maturity early at about 6 months to a year [43; 61; 80; 82]. The perennial shrub is evergreen-to-deciduous and reaches a height of up to 10 to 12 meters (Figure 1.2.1) with soft, corky and whitish bark and tuberous roots [18; 80; 82; 95]. It has bisexual, white-yellowish flowers, but reports of flowers with varying pink have been

noted [80; 82]. The bisexual nature of its flowers allow Moringa to be both adapted for selfing and outcrossing, using bees being their main pollinators with the latter producing more intense phenotypic variations than its counterpart mating system [43]. It produces dry triangular pods that split open at maturity releasing seeds with three papery wings for easy dispersal by the wind [25; 82]. Considerable phenotypic variations are observed fruit sizes, shapes, seed weight, amount of foliage, etc., with wild trees generally producing small fruits, as illustrated in Figure 1.2.2 [82; 92].

Cultivated varieties called, “Jaffna” and “Chavakacheri murunga”, produce fruits varying in length with the former producing lengths ranging from 60-90 cm and the latter from 90-120 cm respectively, are commonly grown in South India [80; 82]. Another variety grown, “Chemmurunga”, produces red tipped fruits and has been reported to flower all year [80; 82]. The Tamil Nadu district also has a variety with thick and bitter pulp known as “Palmurunga” [80; 82]. Some varieties are exploited in West Indies accordingly for their preferred phenotype, including high fruit yielding and reduced flowering varieties [80; 82]. The annual “PKM1” and “PKM2” cultivars are commonly grown commercial varieties developed in India and are superior to their perennial varieties [72].





**Figure 1.2** – 1.) Photographical illustration of *Moringa oleifera*, with 2.) similar aged pods and their diverse phenotype observed within different varieties.

## 1.2 The many uses of *Moringa oleifera*

The Moringa tree has great economical potential due to its i.) wide range of rich and novel products, ii.) its application in multiple industries such as cosmetics, biomedical and fuel, iii.) distribution throughout many under nourished regions and iv.) tolerance to many limiting factors such as environmental or stress factors [18; 25; 61; 82; 95]. Almost all plant parts are edible, including leaves, fruit, flowers and immature pods, which are used as a highly nutritive vegetable in the tropics as well as fodder [4; 5; 17; 41]. All edible organs are good sources for sustainable digestible protein, including amino acids, methionine and cysteine containing sulphur atoms, desirable fatty acids and other essential bioactive compounds such as vitamin C, B, A and iron, which is important in developing countries to counter under-nourishment [8; 25]. The leaves have been shown to be competitive compared to that of

soybean and rapeseed meals in ruminal nitrogen usage, due to its high protein content, sufficient essential amino acids and nitrogen availability [25; 88]. Both Abdulkarim et al. [1] and Oliveira et al. [68] reported higher protein content in Moringa than that of grain legumes, and lipid content than some soybean varieties, respectively [25].

The oil extracted from Moringa seeds has a higher oil content on kernel basis compared to that of either avocado or olive oil [8]. In addition they have a high oleic acid (18:1) content at ranges higher than 76.0% and, is reported to be highly resistant to oxidative rancidification [8; 25]. It is a clear, odourless, sweet tasting oil known as Ben oil, which is used as cooking oil in many countries and has several other industrial uses, including the production of cosmetics and “enfleurage” process for perfumes, lubricant for machinery and lamp illumination [8; 25]. Rashid et al. [81] evaluated Moringa oil as a potential alternative feedstock for biodiesel production and obtained one of the highest cetane numbers of approximately 67 for a biodiesel fuel, due to the high oleic acid content. The cetane number is used to describe the ignition delay time of a diesel fuel when injected into the combustion chamber of an engine [46]. There is a correlation between the higher cetane numbers and reduced NO<sub>2</sub> exhaust emissions, which is harmful to the environment [48]. This cetane number rating has become a fuel quality parameter in biodiesel standards, where *n*-hexadecane is assigned with a high cetane number of 100 and 2,2,4,4,6,8,8-heptamethylnonane a low cetane number of 15 [46]. This makes biodiesel made from Moringa an important alternative to consider for future biodiesel production and another avenue to explore to uplift rural communities. Other industrial applications can be attributed to other parts of the entire plant. The leaves can be used for biogas production, as a foliar nutrient source and as a food preservative [8; 25]. Bandana et al. [7] demonstrated that the application of leaf extracts to seeds or roots, increased root nodulation, nodule weight, and nitrogenase activity in *Rhizobium* treated mung bean [82]. The Moringa seed extract, both dry

and aqueous, have been successfully used to purify water, as it contains cationic polypeptides acting as a natural coagulant [28; 42; 61; 62; 64]. The removal of 99% of colloids from turbid water have been reported by Foidl et al. [27] and was found to be more efficient and less toxic than the chemical counterpart aluminium sulphate [2]. Antibacterial activity of seeds, capable of reducing 99% of bacterial suspensions, has been reported [42; 52; 90]. This process is an efficient, low cost solution to water treatment and adequate water supply to rural areas in developing countries and has potential for large scale application, with programmes already implemented in Brazil [25; 61].

The plant has a variety of therapeutic and pharmacological properties distributed amongst all parts from Moringa, determined by scientific research and documented use [25]. Therapeutic properties including, anti-ulcerative [69], hypocholesterolaemic [29], antihypertensive activity [19], cardioprotective potential [71] and many other uses have been reported. The roots and pods have been shown to be able to depress the central nervous system, generating an analgesic effect, and involved in liver detoxification [10; 34]. The seeds have multiple therapeutic properties, probably due to the high concentration of bioactive compounds. Some of these properties include, antimicrobial activity against bacteria [30; 78; 79] and fungi [23], antitumor, anti-inflammatory [31; 32] and diuretic [58]. These properties emphasise the importance of evaluating and exploiting certain high value, non-conventional oilseeds of their unique and well-rounded attributes.

### **1.3 Population structure and Conservational genetics**

Individuals of the same species found in a restricted geographical area and sharing a common gene pool with the potential to interbreed, define a population as a whole [9]. When a population has been established under a particular selective system, they will uphold genes with adaptive leverage over its environment and form local differentiated populations [94].

These recent patterns in genetic diversity can divulge historic clues that explain the current population structure [94].

There is an interplay between population structure affecting genetic variation, genetic diversity being influenced by biological forces, i.e. gene flow through population factors including seed and pollen dispersal, tree density, fragmentation, founder history and micro-environmental selection [47], and the genetic structure formed by natural processes such as mutation, genetic drift, selection, reproductive isolation and migration [13]. Thus, the importance of understanding the population structure is vital for logical decision making and planning of breeding programmes. Change in allelic frequencies within populations play an influential role in driving evolution, thus it is important to understand the genetic structure of populations with regards to allelic frequencies at different hierarchical levels, such as individuals, cultivars, natural ecological groupings, regions and various other groupings [6; 15; 61; 86].

Genetic variation offers insurance against genetic erosion and maintenance of biodiversity. This threat of erosion of biodiversity has multiple facets, including habitat fragmentation increasing inter-population genetic divergence, misappropriate land reformation, and ecological and life history traits such as reproductive dynamics [3]. Genetic parameters such as inbreeding levels, gene flow, and genetic differentiation are important keys to ecologists and conservation biologists [14]. Conservation genetics focuses on understanding plant genetic responses to environmental changes and to develop informed strategies for the conservation of genetic resources [3]. Okun et al. [66] suggested knowledge on the genetic variation, as well as ecological and geographical differentiations, are important for breeding and sampling strategies. High levels of genetic variation can aid against co-evolving biotic factors [63]. It is thus necessary to elucidate the genetic diversity and relatedness of an

introduced population that has been established for future conservation and breeding strategies, as suggested by Wu et al. [95] on the study of Moringa trees.

There have been a lot of biochemical studies conducted on Moringa due to its therapeutic and nutritional properties; however few studies have focussed on elucidating genetic variation within Moringa until recently. In the study, conducted by Wu et al. [95], they isolated and characterised 20 polymorphic microsatellite loci but have not fully tested their capabilities as markers for investigating genetic variation. They screened twenty-four individuals with germplasms of India and Myanmar for polymorphisms of these microsatellite markers and found that the number of alleles per locus ranged from two to six. Only seven of these loci were found to be significant. Muluvi et al. [61] identified two sources of germplasm introductions within Kenyan natural populations of Moringa, using amplified fragment length polymorphism (AFLP) markers, and further suggested that understanding their provenance source is an important factor for conservation and exploitation. Further an investigation into the outcrossing rate in Moringa using AFLP markers, identified that the combination of their natural selfing capability and early sexual maturity, provides the opportunity for developing inbred cultivars. A multilocus outcrossing rate of 0.74 and a twenty-six percentile natural selfing rate that could potentially lead to an overestimation of the proportion of genetic variance, was observed for Moringa in the above mentioned. This emphasises that appropriate management and understanding of its genetic structure is necessary for protecting genetic variation as well as minimising selfing within seed orchards [60]. RAPD marker analysis of genetic diversity in Moringa has proven very successful lately and reaffirms the importance of elucidating genetic variation for management of Moringa [2; 18; 83]. A genetic diversity study conducted on 161 accessions of Moringa, 131 wild accessions collected from the wild in Pakistan and 30 accessions obtained from ECHO (Florida), has identified the usefulness of 19 microsatellite markers in determining great genetic diversity within wild

collections globally [87]. Very recently the genetic diversity of commercially grown Moringa cultivars from India was investigated in a comparative evaluation of three genetic marker systems, including RAPD's, ISSR's and cytochrome P<sub>450</sub>-based, and determined the cultivars to be highly diverse genetically based on all three markers [84]. High genetic diversity was found among the cultivars and out of the three marker systems used, the ISSR marker system was found to be most effective for genetic diversity evaluation of Moringa trees within that study.

In the present study DNA techniques involving molecular markers, which include twelve random amplified polymorphic DNA (RAPD) markers and twenty-one identified microsatellite markers were evaluated for their utility in genotyping Moringa, and characterising the levels of genetic diversity within a collection found at the Hatfield experimental farm at the University of Pretoria. Phenotypic variation has been observed, with limited knowledge on the genetic diversity within this collection. By defining the genetic structure, proper management programmes can be exploited in their resources with the assurance of scientifically informed selection that may serve as a valuable commodity.

#### **1.4 Molecular markers**

Although morphological and ecological characterisations have been effective in evaluating germplasms and biological forces, which remain essential, they have a lack of genetic informative support and low discriminating power, in comparison to that of molecular methods used for genetic diversity characterisation [49]. Certain molecular methods have been available for some time, i.e. in the form of allozymes and Restriction Fragment Length Polymorphisms (RFLP) assays, but advancements in DNA analysis techniques have made them redundant [14; 38].

DNA methods, especially Polymerase Chain Reaction (PCR) -based techniques, have made a variety of new approaches available to investigate genetic variability as they became more easily accessible, reproducible, more polymorphic informative, less time consuming, cost effective and more specific. They provide powerful and reliable tools for genetic analysis of variation within and amongst plant populations [38]. These genetic markers form a part of the observable differentiation within and amongst populations that are geographically isolated. Studying the genetic diversity of populations allows us to assess genetic parameters such as, levels of inbreeding, gene flow, genetic differentiation, and processes of micro evolutionary changes of populations [14].

Although molecular markers are found to be selectively neutral marker-trait associations or association mapping has become a powerful approach to identify correlations for genotype and phenotype within diverse germplasm collections [93]. Also these molecular markers are valuable tools for identifying genetic diversity of populations, recognising units of conservation including strategies for their management, and assisting breeding programmes [18]. Other applications of molecular markers have become widespread across many studies including genome mapping, mating patterns and phylogenetics. However, careful consideration is required when choosing such methods, as the repertoire of different types of molecular markers is vast and each differs in their principle, methodologies and applications [85]. The attractiveness, considering advantages and disadvantages, of a specific technique can vary depending on the researcher.

#### **1.4.1 RAPD**

According to the prior knowledge of sequence information for the type of primers used, two types of PCR-based techniques can be defined, arbitrary or semi-arbitrary primed PCR techniques and site targeted PCR techniques [85]. With no previous information of the

genome sequence available, two arbitrary molecular marker techniques are commonly used in diversity studies with high success rates, including the RAPD and AFLP marker technique. The AFLP marker technique combines the technology of restriction digestion of genomic DNA, using the combination of both a frequent and rare cutting restriction enzyme, and PCR amplification of these digested fragments, with the help of designed adaptors ligated to the cut fragments [85]. Many different applications of this technique have been successfully done including the above mentioned study involving the estimation of outcrossing rates in Moringa [60]. However the key drawbacks of the AFLP technique when compared to the RAPD technique, are the added cost of the specific restriction enzymes used as well as the designed double-stranded oligonucleotide adaptors including the time and effort spent on the added number of steps involved [85]. The RAPD technique focuses on sampling genomic regions without discrimination of coding and non-coding regions to identify random segments of polymorphic DNA [51]. Discrete segments of DNA are amplified with primers that anneal to unknown flanking sequences, when these priming sites are within amplification distance [91]. These primer binding sites are inverted repeats found ubiquitously throughout the genome [9]. The discrete bands that are generated by amplification are scored as gel phenotype through gel electrophoresis and collectively forms individual profiles for each. The polymorphisms generated are due to base substitutions or deletions in the binding sites, insertions stretching binding sites too far for amplification and insertion and deletions modifying fragment size [35]. Preconditions for primers include decamer primers consisting of 60%-80% GC content and no self-compatible ends [74]. The application of the RAPD technique with the use of statistical methods has been effective in analysing genetic diversity and similarity amongst populations of various species [18].

An important limitation to this technique is its inability to distinguish between homozygotes and heterozygotes determinants and can only identify dominant markers. Detection is



restricted to only a single allele i.e. amplified allele, and not the other, thus monoallelic detection system [9]. It still allows for the estimation of allelic frequencies required for population genetic structure characterisation, but when compared to other co-dominant techniques there is an underestimation of the genetic variability present [51]. In conjunction to dominant inheritance, the co-migration of bands with similar mobility; and reproducibility reduces the assurance of homology which is unfavoured by other studies with phylogenetic components [9]. In some cases the lack of resolution of co-migrating fragments can be reduced by polyacrylamide gels as suggested by a study done by Huff et al. [39].

However the lack of prior sequence information, reduced labour intensive work of generating specific primers, minimal use of limited DNA sources and ease of operation set up, makes this technique a valuable diagnostic tool for the efficient assessment and manipulation of a gene pool for breeding programmes, diversity evaluation, management of biodiversity etc. [9; 49; 85]. In addition this technique allows for the addition of large number of primers to identify useful markers which then have the potential to identify clones, genome mapping, study of sexual differentiation and microbiological relationships between human and the environment [9; 50].

#### **1.4.2 Microsatellites**

Simple Sequence Repeats are short repetitive DNA sequences of a few base pairs (1-6bp) in length found ubiquitously throughout the genome, variations in repeat length of the motif forms due to a natural process of polymerase slippage [11]. They contain unique flanking regions that allow specificity when evaluating loci. These repeats can be classified according to their motif structure, as perfect, imperfect, interrupted or composite and they can be defined depending on the repeat length, such as mono-, di-, trinucleotide etc. repeat motifs [67]. The sequences are unique and present high levels of allele polymorphism. There are

multiple advantages such as high variability or polymorphisms, abundance throughout the genome, specificity of loci evaluation, co-dominance markers, straightforward and easy assay using PCR setup and scoreability and high reproducibility [11; 77]. The initial costs involved for setting up microsatellite work is high due to proper screening for informative microsatellite's identified from the vast amounts of estimated numbers and labelled probes [14]. However, once it has been developed it remains an inexpensive procedure and is easily maintained.

Microsatellites were developed for the use in human studies, but have spread to many other organisms due to its versatility. They have become useful markers used in various studies, including genetic mapping, human forensic analysis, conservation biology phylogenetics and population studies [11; 33]. They have become of great interest for accurately studying genetic parameters and the analysis of population structures in both naturally occurring and experimentally derived populations [14]. The potential of microsatellite's lie in their level of allelic diversity, but loci with remarkably high number of alleles are rare as seen in a study by Maroof et al. [55]. It would be assumed that these variable alleles at certain loci are due to the area they are found in.

In comparison to other marker techniques this gives us the essential resolution to investigate and accurately estimate genetic parameters of populations [14]. As suggested by Wu et al. [95], these markers are important for elucidating founder components of introduced species and have been important in studying other aspects of ecological evolution patterns such as genetic distribution patterns of pollen within *Moringa* trees.

## 2. Aim

This study aims to evaluate the potential of DNA techniques involving two molecular marker systems for their utility in detecting genetic variation, genotyping Moringa, and characterising the levels of genetic diversity within a collection found at the Hatfield experimental farm at the University of Pretoria and between foreign sampled Moringa varieties. The two marker systems is comprised of twelve random amplified polymorphic DNA (RAPD) markers and twenty-one identified microsatellite markers.

### 2.1 Objectives

- Determine the resolution power of selected markers, compare them and estimate the level of genetic variation detected among collections of Moringa.
- Based on this, evaluate the genetic diversity of an uncharacterised SA population compared to known Moringa plants from different localities and cultivars.

### **3. Materials and methods**

#### **3.1 Plant material**

Moringa leaf samples were collected from an experimental orchard grown at the University of Pretoria's experimental farm and will be representing an initial South African source. The seeds are of unknown source or cultivar. This orchard was well established and grown according to the appropriate propagation methods reported by De Saint Sauveur and Broin [21] and Palada and Chang [70]. Seeds obtained from populations of different localities, including PKM1 and PKM2 from India and Hawaii were also used in this study (Figure 1.1. a., b. and c.). Foreign seeds were imbibed overnight in distilled water before being planted into individual cells (14 seeds per area) of 100 cell seedling trays (3-4 cm wide and deep). Samples were grown in potting soil, consisting of soil and compost in these seedling trays at 24 °C in 30%-50% shade and watered with 15ml of water every second day. Germinated seedlings of approximately 20cm in length were then transplanted to pots (3 per pot), containing commercial potting mix. Leaf samples were collected in liquid nitrogen and stored in a freezer at -80°C.

#### **3.2 DNA extractions**

A total of seventy-one samples were used, including fifty accessions from the South Africa (S.A.) accessions, six PKM1 and three PKM2 accessions from India and twelve accessions from Hawaii. The Hawaiian accessions had six from both PKM1 and PKM2 cultivars (Table 3.1). DNA extractions were conducted by using the DNeasy plant mini kit (Qiagen®, U.S.A) to produce uniform DNA samples and concentrations. Alternatively a modified version of a CTAB method, by Doyle and Doyle [24] and Cullings [16], was performed on all the samples for obtaining higher yields. The CTAB method was performed using a total of 200mg leaf

tissue, crushed in liquid nitrogen and submerged in a CTAB extraction buffer containing 2% CTAB, EDTA – 0.02M at pH 8.0, Tris-Cl – 0.1M pH 8.0, NaCl – 1.4M,  $\beta$ -mercaptoethanol and PVP, similarly used by da Silva et al. [18]. The mixture was incubated at 60°C for 1.5 hours, after which an equal volume of ice cold chloroform was added. The samples were centrifuged at 15 000 rpm for 15 minutes, where the top aqueous layer was carefully transferred to a new 1.5ml eppendorf tube not to disturb the lower layers. All centrifugation steps were conducted at room temperature. Ice cold Isopropanol was added to the solute and incubated at -20°C for 2 hours. The samples were centrifuged for 20 min at 15 000 rpm and the supernatant removed. Firstly a 96% ethanol and lastly a 70% ethanol wash were done. Samples were air dried to remove all traces of ethanol and finally eluted in TE buffer and kept overnight. RNA was digested by adding *RNase* (0.02 mg/ml), incubated for 30min at 37°C and followed by quality analysis.

**Table 3.1** – Sample identification, origins and cultivar description.

<sup>1</sup> Sample Id	Cultivar	Geographical origins
Mo - 1 to 50	Undefined seed source	South Africa
MoID - 1, 2, 6, 7, 8 and 9	PKM1	India
MoIi - 1, 5 and 7	PKM2	India
MoHI – 1, 2, 4, 7, 8 and 9	PKM1	Hawaii
MoH – 2, 4, 6, 7, 8 and 9	PKM2	Hawaii

<sup>1</sup>Cultivars were separated accordingly. Fifty accessions were from South Africa, 6 of PKM1 and 3 of PKM2 from India and lastly 6 of PKM1 and 6 of PKM2 from Hawaii.

### 3.3 DNA quality analysis

DNA was analysed using Nanodrop 3000 spectrophotometer (Thermo Scientific, USA) at 260 and 280 nm absorbance to estimate its purity and concentration. With a further quality check, based on visual observation of genomic DNA, mixed with loading dye containing gelred, ran on a 1% agarose gel. Working stocks of 50ng/ $\mu$ l were prepared and kept at -20°C.

### **3.4 Selection of primers**

#### **3.4.1 RAPD primers**

Twelve decamer RAPD primers (Inqaba biotec™, R.S.A.), were identified and selected from previous research papers to have polymorphic characteristics in plants and subsequently screened for suitability in the current research project (Table 3.2). Out of the twelve primers screened, RAPD 2, 3 and 8 were found to be monomorphic but reproducible. RAPD 3 was subsequently replaced with RAPD3b from another paper by da Silva et al. [18]. All other RAPD primers were found to be polymorphic and reproducible with replicates performed, producing enough bands to suit this study. The incorporation of RAPD markers are used to aid in better resolution of the diversity among these Moringa trees.

#### **3.4.2 Microsatellite primer sets both labelled and unlabelled**

Sequences of regions containing microsatellites were obtained from the NCBI database and twenty-one PCR primer pairs (Table 3.3) for microsatellites were designed with the aid of the program CLC bio Main Workbench (CLC bio, Denmark). Four primer sets produced non-specific bands and were thus discarded from the study, with a total of seventeen microsatellite primer sets remaining for further analysis. The amplified fragments were isolated from agarose gels with the QIAquick Gel Extraction Kit (Qiagen, U.S.A.) and then cloned into the pGEM-T Easy vector (Promega, U.S.A.), according to the manufacturer's instructions, for sequencing. Sequencing reactions were prepared to generate DNA sequencing templates with BigDye® Terminator v3.1 Cycle Sequencing Kit (Life technologies, U.S.A.) via PCR. Reactions contained 1µL of 5 x Sequencing buffer, 2 µL of BigDye®, 60ng of DNA, 3.2 pmol primer and was filled to a final volume of 10 µL. Cycling conditions included an initial denaturation step at 95°C for 5 minutes, followed by 40 cycles of 95°C denaturation step for 30 seconds, extension at 50°C for 10 seconds and a final

extension step at 60°C for 4 minutes. Sequencing samples were cleaned with NucleoSEQ (Macherey-Nagel, France) columns as per manufacturer's recommendations.

Suitable microsatellite primers were labelled with fluorescent dyes (FAM, NED, VIC and PET) according to their respective allelic size ranges for detection in a multiplex reaction, twelve primer sets were subsequently found to be viable candidates to be used for multiplex genotyping or fragment analysis. After allelic ranges were defined for this collection, two panels were designed to avoid pull-up from different overlapping or closely migrating primers within each multiplex reaction (Table 3.3.). Panel 1 contained primer pairs Mic7 ALT, Mic9 ALT, Mic23 ALT, Mic10 ALT, Mic15 ALT, Mic20 ALT, Mic6 ALT and Mic18 ALT. Whereas panel 2 contained primer pairs Mic8 ALT, Mic5a ALT, Mic16 ALT and Mic12 ALT.

### 3.5 PCR reactions

DNA amplification was performed for both markers using two different *Taq* systems. Each of the 10µl RAPD reaction mixtures, contained 25 ng of genomic DNA template, 1µl of the 10x NH<sub>4</sub> reaction buffer, 0.5µl of 50mM MgCl<sub>2</sub>, 1µl of 1mM dNTP mix (Fermentas International, Canada), 0.15µl of a 10µM solution of an individual RAPD primer, with 0.1µl of 'BIOTaq' *Taq* DNA polymerase (Bioline, U.K.) and was filled up to volume with nuclease-free water (Fermentas International, Canada).

Using designed microsatellite primers, a total volume of 10 µl was used for each multiplex reaction, containing 2 µl of the 5 x MyTaq reaction buffer. The reaction buffer supplied with MyTaq DNA Polymerase (Bioline, U.K.) contained premixed and appropriate concentrations of dNTP's and MgCl<sub>2</sub> for efficient PCRs. Moreover each reaction further consisted of 0.15µM of each primer pair in a complex mix of multiple primers for a multiplex reaction.

Zero point one microliters of MyTaq DNA Polymerase was added and a final volume was achieved with nuclease-free water (Fermentas International, Canada).

Master mixes were prepared for each individual primers and multiplex primer mixes, with concentrations adapted accordingly. This was done to circumvent measurement deviations when using small volumes for pipetting. The mixtures were mixed thoroughly for each individual reaction by pipetting up and down and all PCR reaction setups were conducted on ice before placing in a preheated thermocycler.

### **3.6 PCR amplification**

Reactions were carried out using Bio-Rad S1000 Thermocycler (Bio-Rad, U.S.A.). Each marker system required specific optimised conditions, in accordance to their respective DNA polymerase cycling conditions and programmed into the thermocycler to achieve effective amplification. The RAPD reactions were set to an initial melting temperature step of 95°C for 2 minutes, following 37 cycles of each at a denaturing temperature of 95°C for 30 seconds, annealing temperature of 35.5°C for 30 seconds and extension temperature of 72°C for 30 seconds. The final cycling step was an additional extension step at 72°C for 5 minutes to allow unfinished extensions to complete.

Using MyTaq allowed for better PCR cycling times due to the optimised buffer supplied with the DNA polymerase. Amplifications were carried out in the thermocycler programmed for an initial denaturing step at 94°C for 1 minute, following 36 cycles of each at a denaturing temperature of 94°C for 20 seconds, annealing temperature of 57°C for 20 seconds and extension temperature of 72°C for 15 seconds. A final extension step was included at 72°C for 2 minutes, for above mentioned reasons.



**Table 3.2** - Characteristics for the RAPD markers used in this study, including polymorphisms detected and their percentages

Primer Label	OP-system	RAPD seq 5'-3'	% GC	Polymorphic bands/total no. of bands	% Polymorphism
RAPD 1	OPA-02	TGCCGAGCTG	70	5/7	71.43
RAPD 2	OPA 04	AATCGGGCTG	60	0/6	0.00
RAPD 3	OPA-06	GGTCCCTGAC	80	0/3	0.00
RAPD 3b	IDT02	TGATCCCTGG	60	7/11	63.63
RAPD 4	OPA-10	GTGATCCCAG	60	3/5	60.00
RAPD 5	OPA-11	CAATCGCCGT	60	15/17	88.24
RAPD 6	OPA-12	TCGGCGATAG	60	6/6	100.00
RAPD 7	OPA-15	TTCCGAACCC	60	8/10	80.00
RAPD 8	OPA-16	AGCCAGCGAA	60	0/5	0.00
RAPD 9	OPAB19	ACACCGATGG	60	8/10	80.00
RAPD 10	OPAC19	AGTCCGCCTG	70	6/6	100.00
RAPD 11	OPAH02	GAGACCAGAC	60	7/10	70.00
<b>Total</b>				65/96	
<b>Average</b>				5.42/8.00	59.44%

**Table 3.3** – Properties of the twelve remaining microsatellite primer pairs after screening.

Primer labels	GenBank accession no.	Primer sequences (5'-3')	<sup>1</sup> Motif	No. of alleles	<sup>2</sup> Allele size (bp)	Fluorescent dye	Panels
Mic7 ALT	GQ853890.1	CACCTCAGTATCCCTCTCT CGGCTTATGTTCTCGTTT	(CT) <sub>17</sub> (CT) <sub>5</sub>	10	279-291	6FAM	
Mic9 ALT	GQ853892.1	CTTCCCTCCTTATGGCTC GCTCACTTCCATCTCCA	(CT) <sub>15</sub>	5	195-205	6FAM	
Mic23 ALT	GQ853906.1	CCACCGTCACCTACAGAA CAGTAATCAGTCCTCAGC	(GA) <sub>8</sub>	7	159-183	VIC	
Mic10 ALT	GQ853893.1	CCCCTCTATTTCCATTTTCCC CCTCTTGCTCCAATTTCTCTC	(TC) <sub>10</sub> CCT(TC) <sub>6</sub>	5	125-133	6FAM	
Mic15 ALT	GQ853898.1	AAGGGTTTCAACTGCTGG AGCTTGGTGGATTCTTCG	(GA) <sub>14</sub>	7	115-129	PET	1
Mic20 ALT	GQ853903.1	GCCTTTTCTCTCTCTTTTGT CAGGCCTTTGGTCGTTTAG	(TC) <sub>17</sub>	7	112-124	NED	
Mic6 ALT	GQ853889.1	GGGCCCTTATCAGTTGTGA GGGGTCTGTGTTCTTATTG	(CT) <sub>9</sub>	6	100-110	VIC	
Mic18 ALT	GQ853901.1	CCTATTAGCTATACTGCG ACGATCGAACCCTTCTG	(AG) <sub>11</sub>	7	81-91	6FAM	
-----							
Mic8 ALT	GQ853891.1	GCTACAGTGAATGCGATAA CAAAAGGAAGAACGCAAGAG	(CT) <sub>13</sub>	3	154-164	6FAM	
Mic5a ALT	GQ853888.1	GGTTAGTTAGGTCTGCGT CTCGCGGCTTATTTTACG	(AG) <sub>14</sub>	6	123-137	VIC	
Mic16 ALT	GQ853899.1	GAACCCAACAGAGGATAAAC CTTTCTGGGTACTTGTGTC	(TC) <sub>9</sub> C(CT) <sub>13</sub>	10	128-144	NED	2
Mic12 ALT	GQ853895.1	GGGCATTAGAAAACAGAGG CAACAGCACACTTTCCAATC	(GA) <sub>9</sub>	3	100-106	6FAM	
<b>Total</b>				76			
<b>Average</b>				6.3			

<sup>1</sup> Repeat motifs identified from sequencing. <sup>2</sup> Allelic ranges detected for individual loci that were considered for fragment analysis.

### 3.7 PCR product analysis

The RAPD PCR amplifications were electrophoresed in 2% (w/v) agarose gel prepared in TAE buffer (40 mM Tris acetate pH8 and 1mM EDTA), which was also used as the running buffer. The PCR samples were loaded into the wells mixed with 2  $\mu$ L of 6 x DNA loading dye from Fermentas (Fermentas International, Canada.) containing 1 x GelRed™ (Biotium, U.S.A.). The electrophoresis was done at 93 volts and 300 mA for 35 minutes and the gel phenotypes (DNA bands) were visualised using the BioRad Molecular Imager® Gel Doc™ XR system (Bio-Rad, U.S.A.). Each RAPD band was considered as independent characters or loci and assigned molecular weight designations according to their positions estimated against the DNA molecular ladder. This was done to avoid error when scoring.

Microsatellite amplifications were first visualised on 2% agarose gels to confirm amplification and reduce false runs on the fragment analyzer. Capillary electrophoresis was performed on the ABI 3130 genetic analyzer (Applied Biosystems, U.S.A.). Each reaction was prepared by diluting the samples to the appropriate concentrations, to accommodate for optimal fluorescent detection range, and added to a fresh 10  $\mu$ l mix of Formamide and LIZ500 size standard (Applied Biosystems, U.S.A.). A G5 dye filter set was used for the analysis to accommodate for the five different dyes utilized. GeneMapper software version 3.0 (Applied Biosystems, U.S.A.) takes fragment data and generates electropherograms of fragment sizes for downstream analysis.

### 3.8 Data analysis

The RAPD profile (bands) generated and visualised on the gels were scored as a binary model, either present (1) and absent (0), whereas the microsatellite alleles generated from fragment analysis were converted to the binary model for the following analyses. The intensity of the fluorescence of RAPD bands were not taken into consideration. The data

generated were used to calculate parameters, such as number of polymorphic loci, percentage polymorphisms, and polymorphic information content. The polymorphic information content (PIC), which is an estimate of the discriminatory power of molecular markers per primer, was calculated according to the formula described by Botstein et al. [12]:

$$\text{Polymorphic information content (PIC)} = 1 - \sum_{j=1}^n P_{ij}^2$$

Where  $P_{ij}$  is the frequency of the  $j^{\text{th}}$  allele for the  $i^{\text{th}}$  marker, and then summed over  $n$  alleles. The expected heterozygosity ( $H_{exp}$ ), also known as the gene diversity, of the polymorphic locus for a genetic marker was calculated as suggested by Nei [65]:

$$H_{exp} = 1 - \sum P_i^2,$$

Where  $P_i$  represents the allele frequency for the  $i^{\text{th}}$  allele. The arithmetic mean ( $H_{ave}$ ) was calculated as the expected heterozygosity per loci:

$$H_{ave} = \sum H_{exp}/n,$$

Where  $n$  is the number of markers (loci analysed). The Marker Index (MI) is a good estimate of the utility of a marker system. This was devised as the product of the expected mean heterozygosity ( $H_{ave}$ ) and the Effective Multiplex Ratio (EMR) which is expressed as follows:

$$\text{EMR} = n_j \times \beta,$$

Where  $n_j$  number of polymorphic markers and  $\beta$  the fraction of markers that were polymorphic. The loci that were found to be monomorphic in the data set, were excluded from this analysis but remained in the calculation of both the EMR and MI.

### 3.9 Population variation

The Genetic similarity coefficients and the principle coordinate analysis (PCoA) were conducted to estimate the genetic identity or similarity between accessions and to highlight the resolution power of the clustering, respectively. Genetic similarity matrices were generated using Jaccard's coefficients [40] with the module SIMQUAL within the software program called NTSYS-PC [26]. These similarity coefficients were calculated using the following equation as suggested by Panwar et al. [73]:

$$\text{Jaccard's coefficient} = N_{xy} / (N_{xy} + N_x + N_y),$$

Where  $N_{xy}$  is the number of bands shared by the samples,  $N_x$  represents the amplified fragments in sample X, and  $N_y$  represents the fragments in sample Y. Similarity matrices based on these indices were calculated. The similarity matrices were used to construct the Unweighted Pairwise Group Method with Arithmetic average (UPGMA) dendograms, for cluster analysis amongst genotypes. The robustness of the dendograms were determined by bootstrapping the similarity matrices with 2000 replications using the software program FreeTree [36]. The PCoA was conducted using principle component analysis programs such as DCENTER and EIGEN of the NTSYS-pc software [26] based on similarity matrices of Jaccard generating a scatter plot.

#### 3.9.1 Marker system correlations

The Mantel test [54] was employed to determine the coefficient of correlation ( $r$ ) between the similarity matrices obtained from the marker types RAPD, microsatellite and the combination of both markers. The correlation to measure the relatedness among the similarity matrices obtained and 'goodness of fit' of the UPGMA dendograms generated by each marker system (and combination of both), were calculated by means of the MxCOMP and COPH modules

from the NTSYS-pc software [26]. To analyse the ‘goodness of fit’ for the UPGMA dendograms produced by each marker, the cophenetic similarity matrices were generated from each UPGMA dendogram and then compared with the original Jaccard’s similarity matrices.

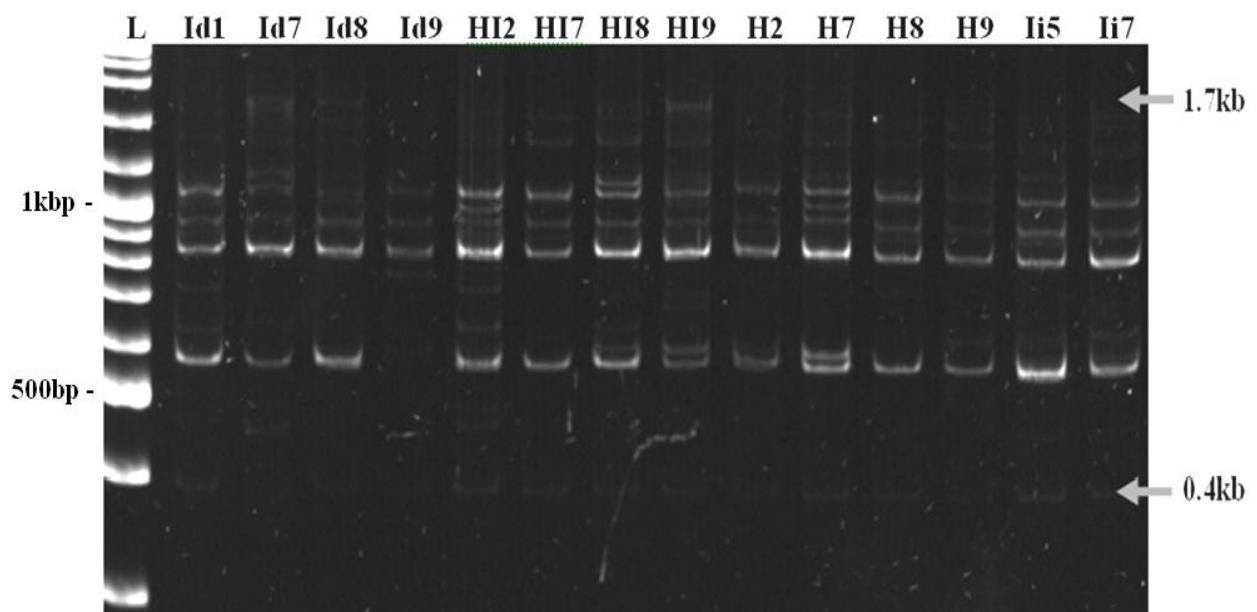
### **3.9.2 Genetic diversity**

Additionally the Shannon-Weaver diversity ( $I$ ) and the Nei genetic diversity ( $h$ ) indices set were determined using both the RAPD and microsatellite markers (microsatellite data were analyzed as dominant markers) with the software program Popgene, version 1.32 [96]. This was conducted on the entire data set for all groups.

## 4. Results

### 4.1 Variation of RAPD loci

Out of the twelve RAPD primers tested, only nine primers were selected for further use in this study based on their optimal and strong reproducible patterns. The assays were done in triplicate to assess the reproducibility of RAPD profiles and generated similar banding patterns. A total of 96 amplicons were generated, from 71 individuals investigated, where 65 bands were found to be polymorphic, producing an average polymorphism of 59.44%. The number of polymorphic bands per primer and the size range of bands generated varied from three (RAPD 4) to 15 (RAPD 5) and 0.2 to 2.5kb, respectively. The average number of amplicons produced per primer was 5.42 RAPD and no population specific band(s) were observed at all. A representative RAPD profile obtained from RAPD 5 (Figure 4.1) produced a total of seventeen bands (0.4-1.7kb) and fifteen bands were found to be polymorphic (88.24%). The RAPD primer 6 and 10 were the most efficient with 100% polymorphism, followed by RAPD primer 5, 7, 9 and 1.



**Figure 4.1** – RAPD profile of 14 *Moringa* genotypes generated by RAPD primer 5. The lanes represent, lane L (100bp ladder); lanes Id1-Ii7 (*Moringa* genotypes).

The PIC values, also a reflection of allele diversity and frequency among the varieties, ranged from the lowest 0.249 (RAPD 11) to the highest 0.375 (RAPD 10), with a mean value of 0.255 for dominant markers. The expected heterozygosity was determined to be significantly higher to the PIC values, ranging from the lowest 0.291 and the highest 0.499 for RAPD primer 11 and primer 10, respectively. The PIC and the  $H_{exp}$  values (Table 4.1) were then used to calculate the arithmetic mean heterozygosity ( $H_{ave}$ ) which was calculated to be 0.438. Determining the marker index requires the calculation of the effective multiple ratio, which was calculated as 6.75, thus producing a value of 2.959 for the marker index.

**Table 4.1** - Polymorphic Information Content (PIC) values estimated for the RAPD markers.

Primer Label	<sup>1</sup> PIC	Expected Heterozyg. ( $H_{exp}$ )
RAPD 1	0.343	0.44
RAPD 2	0	0
RAPD 3	0	0
RAPD 3b	0.373	0.497
RAPD 4	0.374	0.498
RAPD 5	0.364	0.478
RAPD 6	0.351	0.455
RAPD 7	0.331	0.418
RAPD 8	0	0
RAPD 9	0.3	0.368
RAPD 10	0.375	0.499
RAPD 11	0.249	0.291
Average	0.255	<sup>2</sup> 0.438

<sup>1</sup> Values for dominant markers. <sup>2</sup> Calculated per polymorphic primer assay.

The similarity coefficients based on 65 RAPD amplicons ranged from 0.292 to 0.889 and an average similarity index of 0.536. Genotype Mo17 and Mo47 (S.A. accessions) showed the lowest similarity index (0.292), while genotypes MoF1HI1 and MoF1HI2 (Hawaii PKM1 cultivars) showed the highest similarity index (0.889) with RAPD primers. Cluster analysis based on Jaccard's genetic similarity coefficient matrix generated from RAPD data, separated the genotypes into three (**RI**, **RII** and **RIII**) main clusters (Figure 4.2). The first cluster formed six sub clusters, with the majority of the South African (S.A.) accessions found

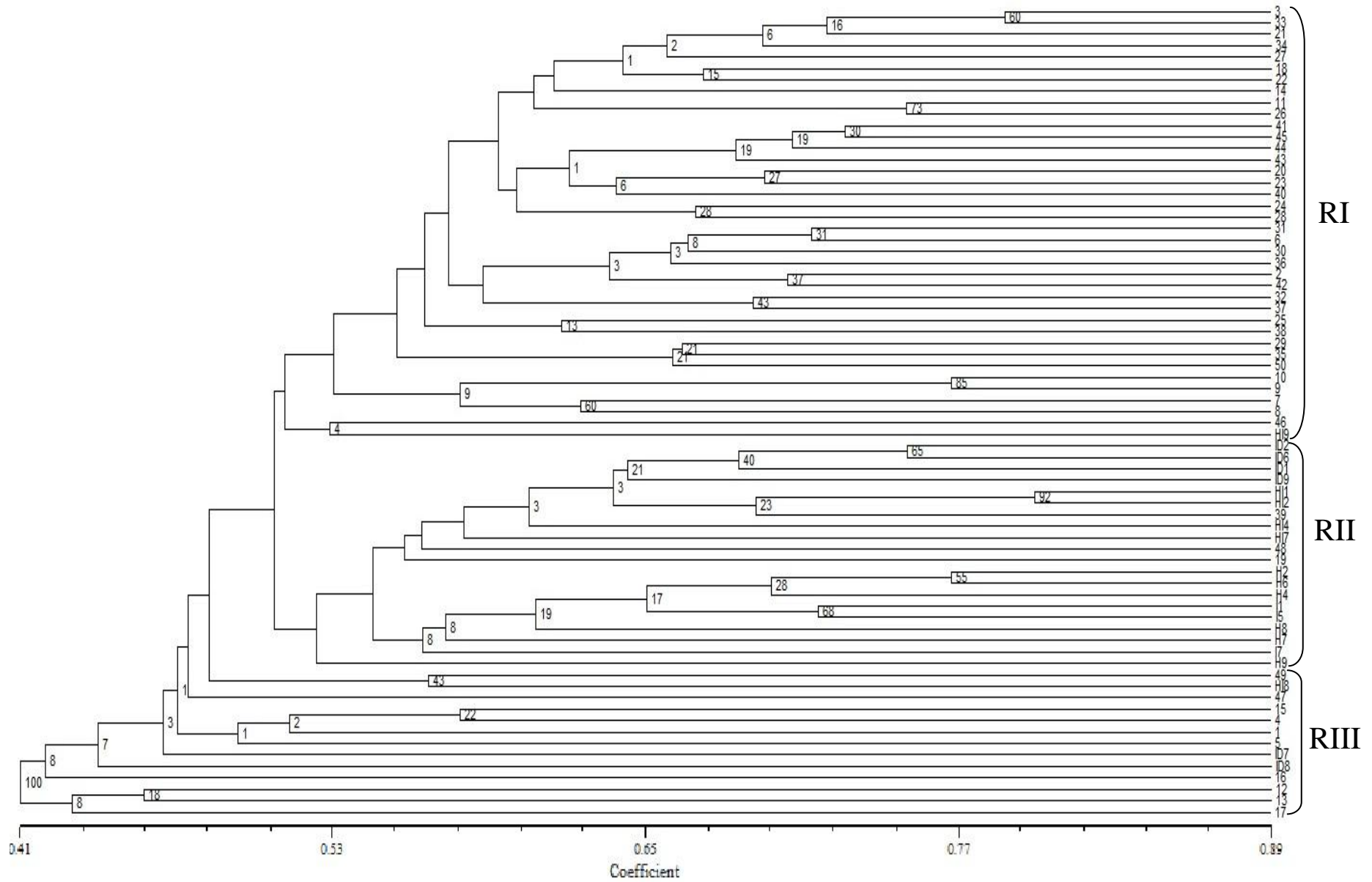


within it. The last sub cluster is an outlier containing Mo46 and MoHI9 (PKM1). The second cluster was split up into two sub clusters. The first contained many of the genotypes from the PKM1 cultivar, including MoID1, MoID2, MoID6, MoID9 and S.A. accessions, Mo19, Mo39 and Mo48. The second sub cluster formed all of the genotypes from the PKM2 cultivar, genotype MoH9 formed a separate branch. The third major cluster was divided into four sub clusters containing the remaining S.A. accessions and separate singular genotypes, MoID7 and MoID8.

#### **4.2 Variation of microsatellite loci**

Only twelve out of twenty-one microsatellite were found to be polymorphic and used for fragment analysis. A total of 76 alleles were detected in all samples with an average of 6.33 alleles per locus. The number of alleles per primer ranged from 3 (Mic8 ALT and Mic12 ALT) to 11 (Mic7 ALT) and allelic variation found spanned from 81 bp (Mic18 ALT) to 291 bp (Mic7 ALT).

The average PIC value for 12 primers was 0.598, with values ranging from lowest 0.278 (Mic12 ALT) to highest 0.812 (Mic7 ALT) for co-dominant markers (Table 4.2). There were three unique alleles generated with microsatellite primers Mic10 ALT, Mic20 ALT and Mic18 ALT identified in the S.A. accessions Mo17, Mo28 and Mo31, respectively. The expected heterozygosities ( $H_{exp}$ ) were calculated for all loci and ranged from lowest 0.268 (Mic18 ALT) to highest 0.732 (Mic20 ALT). The average heterozygosity ( $H_{ave}$ ) was determined to be 0.481 for all loci and thus produced a value of 5.776 for the marker index from an EMR of 12.



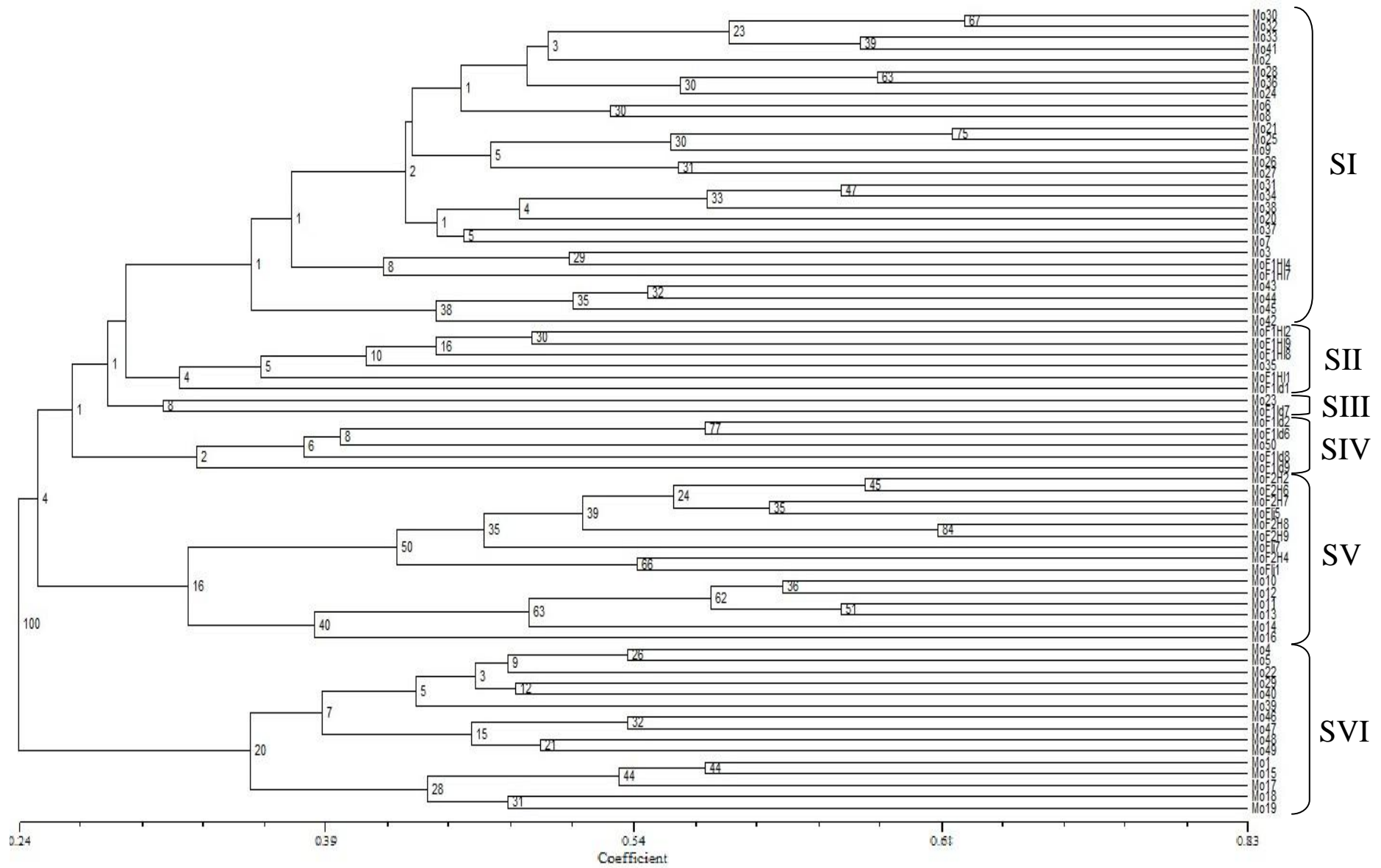
**Figure 4.2** – The UPGMA clustering based on the Jaccard's similarity indices produced from the RAPD markers, clustering the 71 *Moringa* accessions included in this study into three major clusters **RI**, **RII** and **RIII**. Note that 1 (similarity) = genetically identical.

**Table 4.2** - Calculated PIC and  $H_{exp}$  values for the microsatellite markers used in this study.

Primer Label	PIC	<sup>1</sup> Exp. Heterozyg. ( $H_{exp}$ )
Mic7 ALT	0.812	0.592
Mic9 ALT	0.576	0.535
Mic23 ALT	0.605	0.521
Mic10 ALT	0.587	0.282
Mic15 ALT	0.665	0.463
Mic20 ALT	0.722	0.732
Mic6 ALT	0.634	0.578
Mic18 ALT	0.53	0.268
Mic8 ALT	0.457	0.366
Mic5a ALT	0.581	0.507
Mic16 ALT	0.789	0.662
Mic12 ALT	0.278	0.271
<b>Average</b>	<b>0.598</b>	<b>0.481</b>

<sup>1</sup>Calculated for co-dominant loci.

Similarity indices were determined from the 76 microsatellite amplicons produced, ranging between 0.030 and 0.826, with an average of 0.308. The lowest index of 0.030 was determined for the genotypes Mo19 and MoF1Id1 and the highest, 0.826, for genotypes Mo30 and Mo32 when analysed with microsatellite primers. Cluster analysis of the microsatellite marker system separated the genotypes into six major clusters (**SI, SII SIII, SIV, SV and SVI**) (Figure 4.3). The first cluster contains four sub clusters, where sub clusters one, two, and four having S.A. accessions and sub cluster three Mo3, MoHI4 and MoHI7. Cluster two contains Mo35 and PKM1 genotypes (MoHI1, MoHI2, MoHI8, MoHI9 and MoID1), with the third cluster grouping Mo23 and MoID7 together. Accession Mo50 and remaining PKM1 genotypes were grouped into the fourth cluster. In cluster five all of the PKM2 genotypes and Mo10, Mo11, Mo12, Mo13, Mo14 and Mo16 were grouped. The last cluster grouped only S.A. accessions.



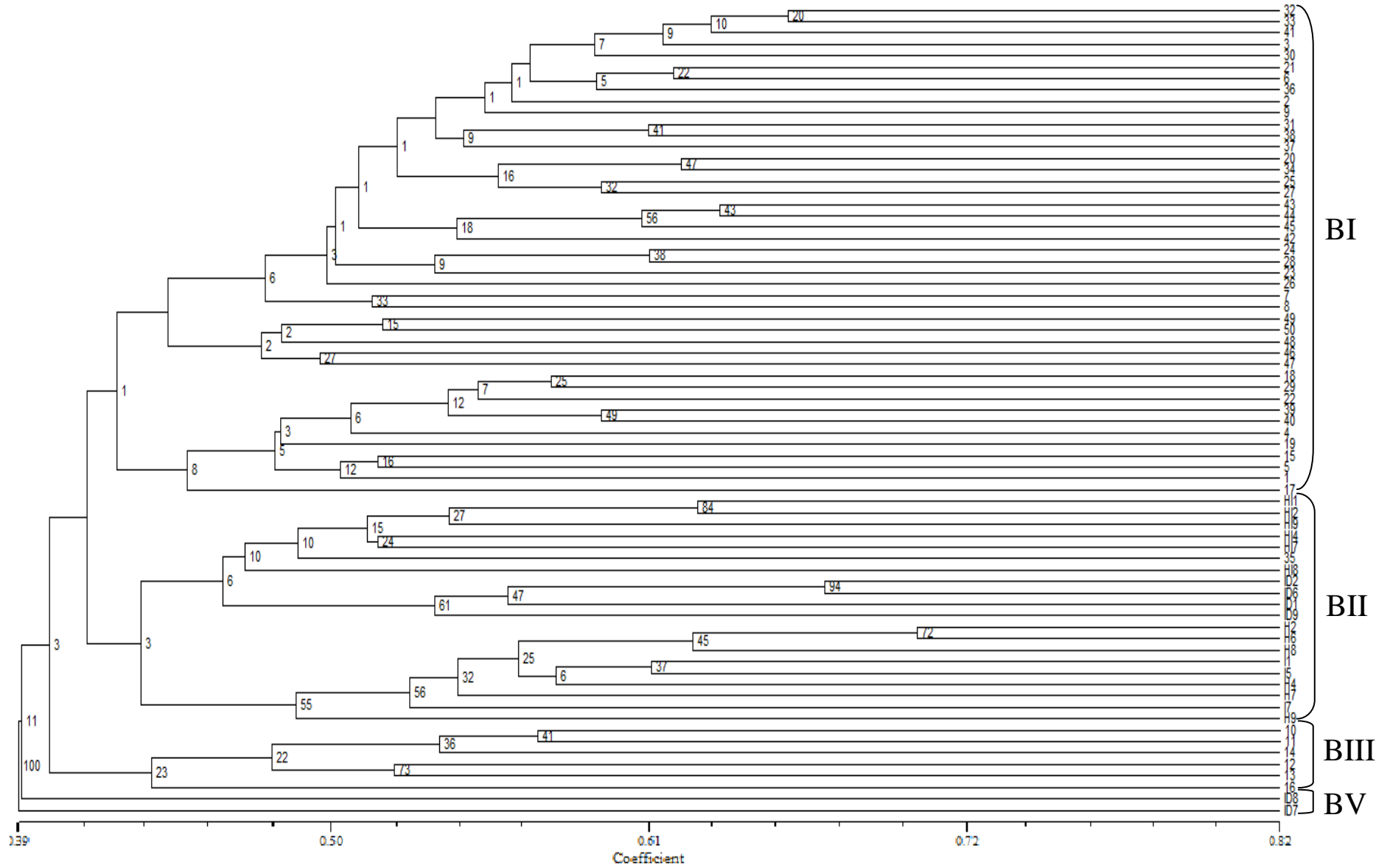
**Figure 4.3** – Microsatellite-UPGMA dendrogram generated, based on the genetic similarity indices produced from binary scores of microsatellite allele data for 71 Moringa accessions. Six major clusters, including **SI**, **SII**, **SIII**, **SIV**, **SV** and **SVI** were identified.

### 4.3 Combined marker analysis

#### 4.3.1 Similarity coefficients of combined data

The similarity coefficient of the combined dominant data sets (RAPD and microsatellite), generated the lowest similarity index of 0.220 for genotypes MoID1 and Mo16 and the highest index of 0.825 for genotypes MoH2 and MoH6. The average similarity coefficient index was determined at 0.452. The UPGMA cluster analysis (Figure 4.4) based on both RAPD and microsatellite marker data showing the relationship and diversity among the accessions, generated groupings into four clusters (**BI**, **BII**, **BIII**, and **BIV**).

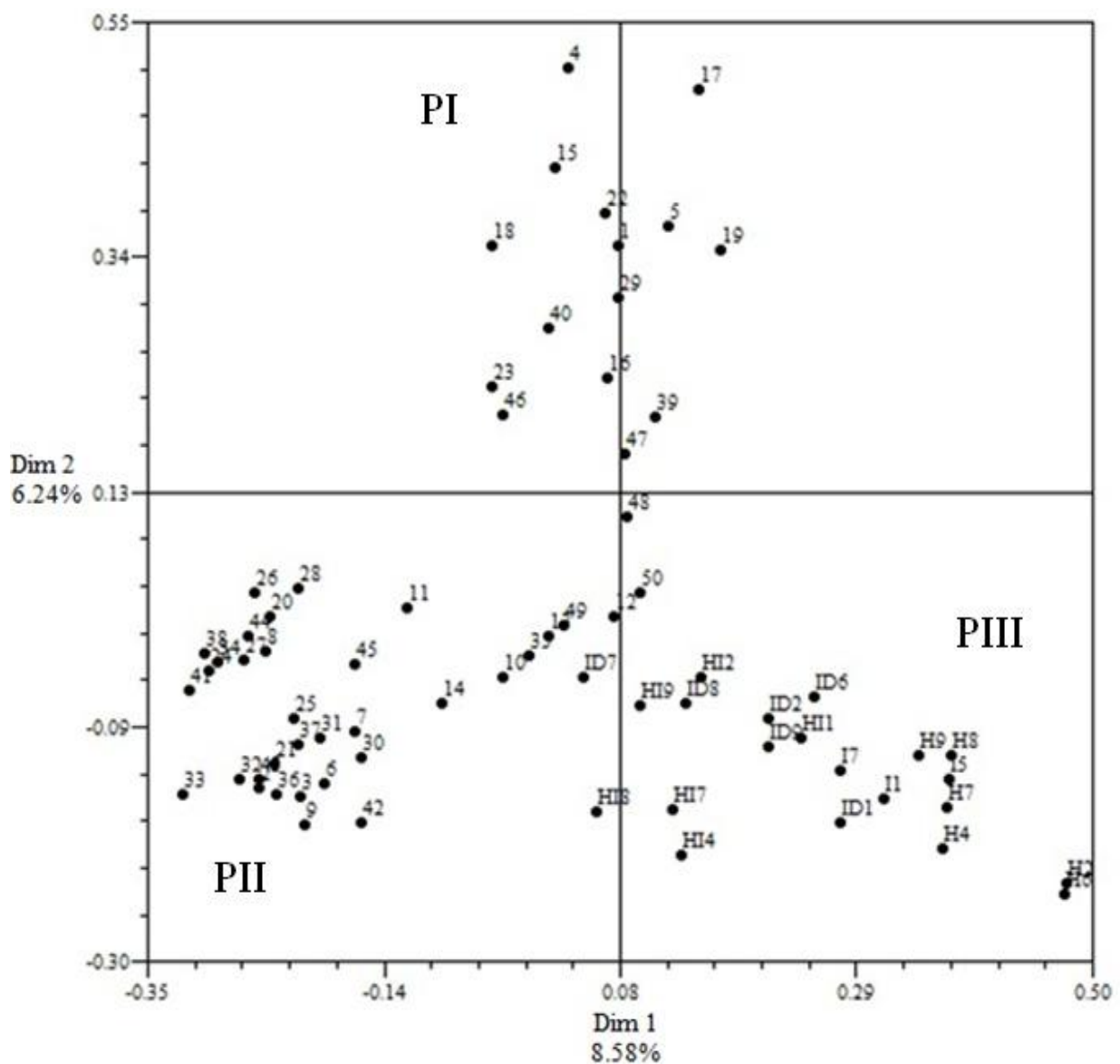
The first cluster formed two sub clusters both containing the majority of the S.A. accessions. The second cluster is divided up into two sub clusters, sub cluster one contains PKM1 genotypes (except MoID8 and MoID7). The second sub cluster contains all of PKM2 genotypes. Two sub clusters were generated within the third cluster, with genotypes Mo10, Mo11, Mo12, Mo13 and Mo14 within the first sub cluster and Mo16 as the outlying sub cluster. The last cluster contains the divided remaining genotypes MoID7 and MoID8. The pattern of sub-grouping of the clusters, which included genotypes from different cultivars and origins, varied between the two marker systems. Bootstrap analysis from the binary data was used to determine the robustness of the dendograms and confidence values as percentages at each node. The percentage values were not significant and resulted in low confidence levels for branches at certain nodes.



**Figure 4.4** – UPGMA clustering generated from the combination of both marker systems, RAPD and microsatellite, clustering foreign genotypes separately.

### 4.3.2 Principle Coordinate Analysis (PCoA)

The Principle Coordinate Analysis (PCoA) carried out using both RAPD and microsatellite marker data, generated a scatter plot of the first two principle ordinates (Figure 4.5). The first and second principle ordinates explained 8.58% and 6.24%, respectively of variance in the molecular data. The scatter plot revealed that the genotypes were grouped together into three major clusters (**PI**, **PII** and **PIII**) and this grouping is equivalent to the UPGMA clustering as shown by the dendrogram for both markers combined.



**Figure 4.5** – Scatter plot of principle coordinate analysis (PCoA) for 71 *Moringa* accessions based on data produced by both markers systems. Three grouping were formed, **PI**, **PII** and **PIII**, with the grouping pattern similar to the corresponding UPGMA clustering.

According to the PCoA, the first two axes or components, being the highest amongst other axis, represent 14.81% of the variation found. The first two groups (**PI** and **PII**), grouped only S.A. accessions. The first grouping (**PI**) corresponded to the second sub cluster within the first main cluster (**BI**) in the dendrogram (Figure 4.4) produced by the combination of both marker systems. The second grouping (**PII**), identified a grouping similar to that of the first sub cluster within cluster **BI** from the dendrogram containing both markers. The remaining genotypes found in the first cluster from both markers (**BI**) grouped close together in both cluster analysis and PCoA. The last grouping (**PIII**) contained all of the foreign genotypes and further grouped them into PKM1 and PKM2 collections. This confirms three distinct associations of genotypes into distinct subgroups. Some differences were found comparing the two methods. According to the UPGMA analysis, genotypes Mo11, 13, and Mo14 were more related to the foreign grouping than found in the PCoA (where they are grouped within cluster **PII**). Similarly, instead of observing genotypes Mo49, Mo50 and Mo35 grouped within cluster **PI**, as in the UPGMA analysis, they were grouped together with the foreign samples.

### 4.3.3 Correlation test

The cophenetic similarity matrices were used to analyse the UPGMA dendrograms generated, for their goodness of fit by means of the Mantel test. It revealed correlation coefficient values higher than 0.7 for the indices produced by all the markers, confirming their reliability. The generated similarity matrices comparison to the cophenetic matrices correlated to values of 0.767, 0.790 and 0.728, for each marker system RAPD, microsatellite and the combination of both, respectively. The correlation coefficient between the RAPD and microsatellite markers, based on their similarity matrices, was calculated as 0.300. This shows a significant difference in resolution power between the two markers. The correlation coefficients between the RAPD marker and microsatellite marker to the combination of markers were determined



as 0.853 and 0.748, respectively. The molecular markers, RAPD and microsatellite were compared on the basis of different criteria (Table 4.3). The high level of polymorphism compared between the two marker system i.e. RAPD and microsatellite in detecting alleles revealed that the microsatellite marker was very effective in polymorphic detection. By defining the level of polymorphism, the information content of the marker, and the extent to which the assay can identify multiple polymorphisms can give a clear estimate of the marker's utility. The marker indices (MI) were calculated and distinguished the microsatellite marker system from the other, due to the level of polymorphisms detected from the samples. However, the use of PAGE gels increased the EMR index and the extent to which multiple polymorphisms can be detected.

**Table 4.3** - Comparison of the components used for evaluating each marker system within the analysed data set.

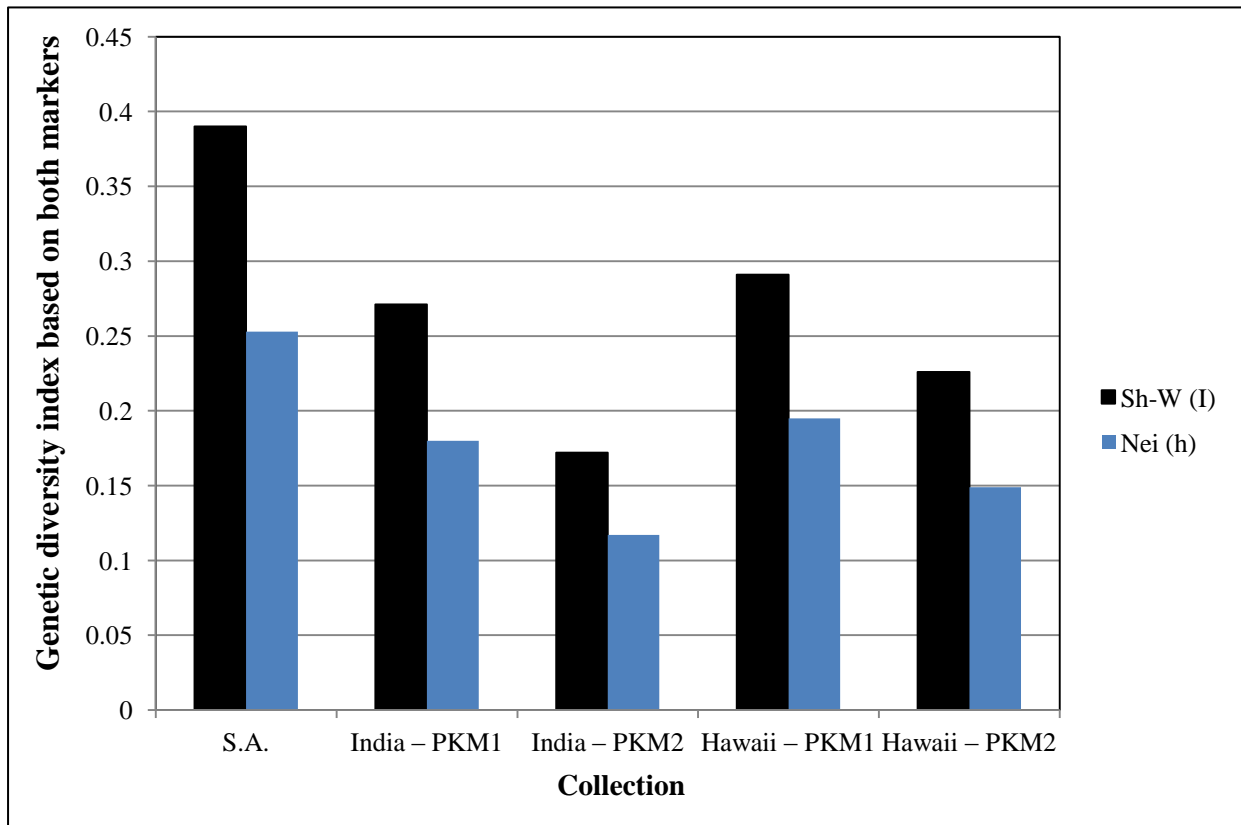
Component	RAPD	Microsatellite
No. accessions	71	71
No. assays/primer combination	12	12
No. polymorphic markers	9	12
<sup>1</sup> Mean No. of alleles per marker	5.42	6.5
<sup>2</sup> Average heterozygosity ( $H_{ave}$ )	0.438	0.481
<sup>3</sup> Fraction of polymorphic marker ( $\beta$ )	0.75	1
Effective Multiplex Ratio (EMR)	6.75	12
<sup>4</sup> Marker Index (MI)	2.959	5.776

<sup>1</sup>Average number of alleles that were observed from both fragment analysis and gel visualisation. <sup>2</sup> Calculated  $\sum(H_{exp})$  was divided by the number of polymorphic primer combinations ( $n$ ). <sup>3</sup> The number of polymorphic markers ( $n$ ) divided by the total number of markers. <sup>4</sup> Marker index determined for both markers using the Effective Multiplex Ratio (EMR) and Fraction of polymorphic markers ( $\beta$ ).

#### 4.3.4 Genetic diversity analysis

The Shannon-Weaver ( $I$ ) diversity indices were evaluated among the five collections and ranged from lowest 0.172 (India with PKM2 cultivars) to highest (S.A. accessions) 0.390 and

an average of 0.270. The Nei's genetic diversity indices ( $h$ ) were calculated from the lowest for the Indian collection with PKM2 cultivars only to the highest for the S.A. accessions, 0.117 to 0.253 respectively. The average genetic diversity ( $h$ ) was determined at 0.179 among the collections (Figure 4.6).



**Figure 4.6** – Genetic diversity values for each collection from their respective origins and cultivars estimated according to Shannon-Weaver ( $I$ ) and Nei's ( $h$ ) diversity calculations.

## 5. Discussion

### 5.1 Primer efficiency

The genome of *Moringa oleifera* has not been sequenced yet, but advances in genomic research have provided new tools, such as molecular markers, that have become essential for crop improvements [18]. The value of a molecular marker technique is dependent on the efficiency of the markers to detect the amount of polymorphisms among the set of genotypes. The different marker systems used in this study are both based on DNA amplification by PCR and due to the nature of the primers and reaction used. It was expected that they would differ in the specific sequences targeted and in the number of amplified sequences. The comparison of these techniques, with respect to the eliciting of unique groups and level of polymorphism generated in each case would give a clear indication of the value of these markers within this data set (Table 4.3). There are many parameters to consider for the utility of markers in each crop species before large-scale characterisation and assessment of germplasms can be attempted. Therefore, this study focused on the comparative evaluation of different PCR-based molecular marker systems (RAPD and microsatellite) and their potential use for fingerprinting and assessing the genetic diversity among accessions of Moringa and cultivars PKM1 and PKM2. Comparative studies in major oilseed crops involving both RAPD and microsatellite markers have been successfully used by various researchers [45; 57] and has recently spread to the genetic diversity study of Moringa [84]. Genetic diversity studies using the individual PCR based marker systems have found to be more common in diversity studies of Moringa [2; 18; 60; 61; 83; 87; 95].

In this study the RAPD primers generated higher number of polymorphic bands compared to other studies conducted on Moringa [2; 18; 84]. This higher number of polymorphic detection could be due to the efficient separation of the bands on a poly-acrylamide gel and agrees with the study conducted by [83], where 3% MetaPhor agarose gels were used to resolve the

bands. Average percentage of polymorphism (59.44) was found to be similar to earlier studies of da Silva et al. [18], even though the better gel separation produced higher resolution for RAPD markers. Slight issues were observed for the reproducibility of a few bands, thus ambiguous bands were not counted and remained useful as well as informative. However, microsatellites generated a higher number of polymorphisms compared to RAPDs, which was expected as the level of polymorphic detection for microsatellites is mostly due to the sensitivity of detection from the capillary analysis for each assay, the ability to detect co-dominant and the high level of polymorphisms within the sequence repeat. Thus, the fragment analysis would produce a higher percentage of polymorphism. Microsatellites do require a considerable amount of effort to determine DNA sequences, defining the microsatellite loci and the laborious process of screening markers for polymorphism within the population, before they can be exploited. Once both these approaches were optimised, a high degree of polymorphism was observed for the DNA profiles however low levels of polymorphism remained among the varieties PKM1 and PKM2, showing that considerable inbreeding and selection have taken place.

The marker index, as well as the effective multiplex ratio, for each marker system was measured as an overall estimate of marker efficiency in polymorphism detection. The higher marker index produced by the microsatellite marker system (5.776) was due to the effective multiplex ratio of 12, because similar efficiencies of detecting heterogeneity, 0.438 and 0.481, for RAPD and microsatellite respectively were estimated. As suggested by Powell et al. [76], the marker index may be used to predict the relative utilities of different marker types for unknown germplasms within the same species. This is due to the fact that as an overall measure of utility it can be applied to different experimental situation where the average expected heterozygosity ( $H_{ave}$ ) and EMR is calculated, even though those parameters are calculated according to the germplasm and experimental conditions of the individual

study under investigation, respectively. The marker utility values for RAPD markers are consistent with other studies in Moringa [84]. Estimates of microsatellite marker index values were consistent with other studies, such as soybean [76] and Finger Millet [73], but no estimates for microsatellites in Moringa have previously been done. The identification and distinction of common genotype profiles from both marker systems within this set of data, facilitates the identification of a unified marker system that could be used for varietal fingerprinting of Moringa crops as suggested by Datta et al. [20].

## **5.2 Marker system correlations**

The scoring of RAPD markers may be prone to errors, such as bands that are absent or present, scored as being similar due to their similar fragment mobility on gels, compared to the scoring of microsatellites via fragment analysis and their co-dominant detection of polymorphisms. This would lead to an overestimation of relatedness between distantly related individuals [77]. With the low correlation observed between the two marker systems in this, scoring could potentially have played a role in the analysis of relatedness using RAPD's. Another consideration is the regions which are targeted along the genome that may be affected by the evolutionary mechanisms involved as suggested by Powell et al. [77].

## **5.3 Population variation**

High levels of genetic variation were observed from the clustering analysis of each marker among the S.A. accessions based on Jaccard's similarity coefficients. The clustering generated varied between the two marker systems for the 71 accessions but allowed for the successful distinguishing between cultivars PKM1 and PKM2, based on the similarity coefficient and cluster analysis. None of the PKM2 genotypes fell under the same grouping with the PKM1 genotypes showing a relationship of genetic identity and not geographical origins. The use of both marker systems in combination generated better resolution for the

clustering of genetic variation, separating them from the South African collection and establishing three distinct groups.

The PCoA is a two dimensional representation of the scalar products of similarity matrices and the relative relationships between individuals. This supported the major groupings of the UPGMA dendrogram generated by RAPD and microsatellite markers in combination. The S.A. accessions found in the two clusters from the PCoA (PI and PII) could potentially be used in breeding efforts and conserving germplasms as they are found to be genetically distinct. This clustering of genetic relationships observed suggests different sources of germplasm sourced to the S.A. collection. In Kenya, AFLP studies on Moringa revealed two sources of introduction from Indian origins [61]. The Jaccard's similarity coefficients based on both RAPD and microsatellite markers, 0.220 and 0.825, were found to be similar to that generated by Saini et al. [84]. They investigated the genetic diversity of commercially grown Moringa cultivars from India using RAPD, ISSR and cytochrome P<sub>450</sub>-based markers and observed high genetic diversity among the cultivars with low distinction between geographical origins.

The Shannon-weaver information and Nei's diversity indices generated higher values, 0.390 and 0.253 respectively, in comparison to foreign collections, which supports the assumption that higher levels of genetic diversity are observed in established out-crossing species. As seen in genetic variation between populations, approximately 10-20% have been found to be attributed to out-crossing nature of certain plants, maintaining most of their variation within, when compared to in-breeding plants [37]. The observed variation may be potentially attributed to founder effects or genetic drift. The lower diversity observed for foreign genotypes, could be due to the low sample size and can be seen by the clustering of these genotypes from different origins. These genetic diversity values are in congruence with genetic diversities estimated for Moringa in a study done by Mgendi et al. [56]. They

investigated the genetic diversity between and within cultivated and non-cultivated provenances of *Moringa* by using RAPD markers and they observed higher genetic diversity within the wild genotypes. In a study done by Manoko et al. [53], it was shown that the cultivated genotypes were exposed to optimum environments which lead to reduced competition and natural selection, and ultimately leading to lower genetic changes. The variation observed for the genetic diversity indicators analysed suggests that accessions found within the South African collection of *Moringa* are not strictly inbred and are appropriate candidates for future improvement strategies.

The RAPD primers 3b, 4, 5 and 10 would be good candidates to be used further in genetic studies due to their high detection of quality polymorphisms for those markers. Another measure for their informativeness was observed in the estimated heterozygosities. The RAPD marker system also has the advantage of being a quick and simple genetic analysis tool for initial screening of a collection and should perhaps be considered for that purpose compared to microsatellite analysis. These markers separated the underlying sub-groups which show its capability at detecting variation within groups. The microsatellite markers had mixed polymorphism informativeness, which is understood as mechanisms involved in DNA variation on different components that are being detected by certain microsatellite markers Powell et al. [75]. Thus marker Mic7 ALT, Mic23 ALT, Mic15 ALT, Mic20 ALT, Mic6 ALT and Mic16 ALT could be used to continue the genetic analysis of *Moringa*. It has also been suggested by Powell et al. [76] that with too large values of similarity, the efficiency of microsatellite markers decrease. In addition, a drawback for microsatellite fragment analysis is its labour intensive pre-screening and panel development. However in this study these markers have been optimised and a platform developed for future screening genetic relatedness in *Moringa*. By selecting a smaller group of microsatellite markers with high informativeness it would be sufficient in estimating further analysis. They can be utilised in

the design of plant breeding programs as well as multiple genomic studies on Moringa, including genetic mapping of agronomical traits and management of genetic resources.

On recommendations of further genotyping i) more samples from the S.A. collection as well as from different ecological regions to look at genetic variability adapted to those settings; ii) more of the foreign samples irrespective of origins and including more cultivars; and iii) including different species to verify if some of these markers are species specific probes, will essentially improve the outcome of the study of genetic relationships among Moringa. With these microsatellite markers, pollination biology could be further investigated due to their co-dominant nature and good utility shown as seen by its marker index which can aid in planning breeding programs. Another prospective avenue to target using these markers would be the identification of links between agronomic important traits and these molecular markers for marker-assisted indirect selection in the development of new lines as the South African accessions showed distinct separate grouping from the already established cultivars.



## 6. Conclusion

The characterisation of genetic variability among seed orchards will aid in conservation to uphold genetic diversity, for plant breeding evaluating varieties and cultivar development [18; 22]. Furthermore, due to Moringa's mixed mating system it is beneficial to clarify genotypes or cultivars and genetic markers as an efficient way of evaluating the genetic variability among genotypes, especially when no genome sequence is available. Key criteria for screening successful markers are their informativeness and the ease of their assays. In a comparative analysis of RAPD and microsatellite, the microsatellite marker was found to be the better marker as it was highly informative due to the higher information content with an average PIC value of 0.598, average heterogeneity and Marker Index (5.776). The RAPD marker system was found to be less efficient, with a lower utility and information content. These findings are consistent with that of previous studies on Moringa and *Eleusine coracana* [73; 84]. In the dendrogram using both markers, cultivars from foreign origins were distinguished from each other as well as two groupings within the S.A. collection that generated three distinct clusters. Grouping of accessions revealed relationships according to distance instead of geographical origins. This was supported by the similar pattern obtained from the PCoA. The variability observed among the South African accessions could potentially be exploited in plant breeding programs.

## Reference

- [1] Abdulkarim, S., Long, K., Lai, O., Muhammad, S., Ghazali, H., 2005. Some physico-chemical properties of *Moringa oleifera* seed oil extracted using solvent and aqueous enzymatic methods. *Food Chemistry* 93, 253-263.
- [2] Abubakar, B., Wusirika, R., MuAzu, S., Khan, A., Adamu, A., 2011. Detection of Genetic Variability using Random Amplified Polymorphic DNA Markers in Some Accessions of *Moringa oleifera* Lam. from Northern Nigeria. *International Journal of Botany* 7.
- [3] Aguilar, R., Quesada, M., Ashworth, L., HERRERIAS-DIEGO, Y., Lobo, J., 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology* 17, 5177-5188.
- [4] Anwar, F., Ashraf, M., Bhangar, M.I., 2005. Interprovenance variation in the composition of *Moringa oleifera* oilseeds from Pakistan. *Journal of the American Oil Chemists' Society* 82, 45-51.
- [5] Anwar, F., Bhangar, M., 2003. Analytical characterization of *Moringa oleifera* seed oil grown in temperate regions of Pakistan. *Journal of Agricultural and Food Chemistry* 51, 6558-6563.
- [6] Ayana, A., Bekele, E., Bryngelsson, T., 2000. Genetic variation in wild sorghum (*Sorghum bicolor* ssp. *verticilliflorum* (L.) Moench) germplasm from Ethiopia assessed by random amplified polymorphic DNA (RAPD). *Hereditas* 132, 249-254.
- [7] Bandana, B., Srivastava, R., Mathur, S., 1987. Nodulation and nitrogenase activity in *Vigna mungo* in response to seed-soaking and root-dressing treatments of *Moringa* leaf extracts. *Indian J. Plant Physiology* 30, 362-367.
- [8] Banerji, R., Bajpai, A., Verma, S., 2009. Oil and fatty acid diversity in genetically variable clones of *Moringa oleifera* from India. *Journal of Oleo Science* 58, 9-16.
- [9] Beyene, D., 2005. Genetic variation in *Moringa stenopetala* germplasm of Ethiopia by using RAPD as genetic marker. Addis Ababa University.
- [10] Bharali, R., Tabassum, J., Azad, M.R.H., 2003. Chemomodulatory effect of *Moringa oleifera*, Lam, on hepatic carcinogen metabolising enzymes, antioxidant parameters and skin papillomagenesis in mice. *Asian Pacific Journal of Cancer Prevention* 4, 131-140.
- [11] Bhargava, A., Fuentes, F., 2010. Mutational dynamics of microsatellites. *Molecular biotechnology* 44, 250-266.
- [12] Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics* 32, 314.
- [13] Buiteveld, J., Vendramin, G., Leonardi, S., Kamer, K., Geburek, T., 2007. Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. *Forest ecology and management* 247, 98-106.
- [14] Chase, M., Kesseli, R., Bawa, K., 1996. Microsatellite markers for population and conservation genetics of tropical trees. *American Journal of Botany*, 51-57.
- [15] Crandall, K.A., Bininda-Emonds, O.R., Mace, G.M., Wayne, R.K., 2000. Considering evolutionary processes in conservation biology. *Trends in ecology & evolution* 15, 290-295.
- [16] Cullings, K., 1992. Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology* 1, 233-240.
- [17] D'souza, J., Kulkarni, A., 1990. Comparative studies on nutritive values of tender foliage of seedlings and mature plants of *Moringa oleifera* (Lamk.).

- [18] da Silva, A.V.C., dos Santos, A.R.F., Lédo, A.d.S., Feitosa, R.B., Almeida, C.S., da Silva, G.M., Rangel, M.S.A., 2011. Moringa genetic diversity from germplasm bank using RAPD markers. *Tropical and Subtropical Agroecosystems* 15.
- [19] Dangi, S., Jolly, C., Narayanan, S., 2002. Antihypertensive activity of the total alkaloids from the leaves of *Moringa oleifera*. *Pharmaceutical biology* 40, 144-148.
- [20] Datta, J., Lal, N., Kaashyap, M., Gupta, P.P., 2010. Efficiency of three PCR based marker systems for detecting DNA polymorphism in *Cicer arietinum* L and *Cajanus cajan* L Millspaugh. *Genetic Engineering and Biotechnology Journal* 2010.
- [21] De Saint Sauveur, A., Broin, M., 2010. Moringanews, Moringa Association of Ghana Produire et transformer les feuilles de Moringa. Editions CTA, CDE; Horizon Géméno.
- [22] Demir, K., Bakir, M., Sarikamiş, G., Acunalp, S., 2010. Genetic diversity of eggplant (*Solanum melongena*) germplasm from Turkey assessed by SSR and RAPD markers. *Genetics and Molecular Research* 9, 1568-1576.
- [23] Donli, P.O., Dauda, H., 2003. Evaluation of aqueous Moringa seed extract as a seed treatment biofungicide for groundnuts. *Pest management science* 59, 1060-1062.
- [24] Doyle, J., Doyle, J., 1987. Genomic plant DNA preparation from fresh tissue-CTAB method. *Phytochemical Bulletin* 19.
- [25] Ferreira, P.M.P., Farias, D.F., Oliveira, J.T.d.A., Carvalho, A.d.F.U., 2008. Moringa oleifera: Bioactive compounds and nutritional potential. *Revista de Nutrição* 21, 431-437.
- [26] FJ, R., 2008. NTSYSpc: Numerical taxonomy system, ver. 2.20. Exeter Publishing, Setauket, NY.
- [27] Foidl, N., Makkar, H., Becker, K., 2001. The potential of *Moringa oleifera* for agricultural and industrial uses. *The Miracle Tree: The Multiple Attributes of Moringa*, 45-76.
- [28] Gassenschmidt, U., Jany, K.D., Bernhard, T., Niebergall, H., 1995. Isolation and characterization of a flocculating protein from *Moringa oleifera* Lam. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1243, 477-481.
- [29] Ghasi, S., Nwobodo, E., Ofili, J., 2000. Hypocholesterolemic effects of crude extract of leaf of *Moringa oleifera* Lam in high-fat diet fed wistar rats. *Journal of Ethnopharmacology* 69, 21-25.
- [30] Govardhan Singh, R., Negi, P.S., Radha, C., 2013. Phenolic composition, antioxidant and antimicrobial activities of free and bound phenolic extracts of *Moringa oleifera* seed flour. *Journal of Functional Foods*.
- [31] Guevara, A., Vargas, C., Uy, M., 1996. Anti-inflammatory and antitumor activities of seed extracts of malunggay, *Moringa oleifera* L.(Moringaceae). *Philippine Journal of Science* 125.
- [32] Guevara, A.P., Vargas, C., Sakurai, H., Fujiwara, Y., Hashimoto, K., Maoka, T., Kozuka, M., Ito, Y., Tokuda, H., Nishino, H., 1999. An antitumor promoter from *Moringa oleifera* Lam. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* 440, 181-188.
- [33] Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources* 11, 591-611.
- [34] Gupta, M., Chakrabarti, S., 1999. CNS activities of methanolic extract of *Moringa oleifera* root in mice. *Fitoterapia* 70, 244-250.
- [35] Hadrys, H., Balick, M., Schierwater, B., 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Molecular Ecology* 1, 55-63.
- [36] Hampl, V., Pavlíček, A., Flegr, J., 2001. Construction and bootstrap analysis of DNA fingerprinting-based phylogenetic trees with the freeware program FreeTree: Application to trichomonad parasites. *International Journal of Systematic and Evolutionary Microbiology* 51, 731-735.

- [37] Hamrick, J., Godt, M., 1989. Allozyme diversity in plant species. In 'Plant population genetics, breeding and genetic resources'. (Eds AHD Brown, MT Clegg, AL Kahler, BS Weir) pp. 43–63. Sinauer Associates: Sunderland, MA.
- [38] Heubl, G., 2010. New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques. *Planta medica* 76, 1963-1974.
- [39] Huff, D., Peakall, R., Smouse, P., 1993. RAPD variation within and among natural populations of outcrossing buffalograss [*Buchloe dactyloides* (Nutt.) Engelm.]. *Theoretical and Applied Genetics* 86, 927-934.
- [40] Jaccard, P., 1908. Nouvelles recherches sur la distribution florale.
- [41] Jahn, S., Musnad, H.A., Burgstaller, H., 1986. Tree that purifies water: Cultivating multipurpose Moringaceae in the Sudan. *Unasylva* 38, 23-28.
- [42] Jahn, S.A.A., 1984. Effectiveness of traditional flocculants as primary coagulants and coagulant aids for the treatment of tropical raw water with more than a thousand-fold fluctuation in turbidity.
- [43] Jyothi, P., Atluri, J., Reddi, C.S., 1990. Pollination ecology of *Moringa oleifera* (Moringaceae). *Proceedings: Plant Sciences* 100, 33-42.
- [44] Kantharajah, A., Dodd, W., 1991. Rapid clonal propagation of *Moringa oleifera* Lam., using tissue culture. *South Indian Horticulture* 39, 224-228.
- [45] Khurana-Kaul, V., Kachhwaha, S., Kothari, S., 2012. Characterization of genetic diversity in *Jatropha curcas* L. germplasm using RAPD and ISSR markers. *Indian Journal of Biotechnology* 11, 54-61.
- [46] Knothe, G., Matheaus, A.C., Ryan III, T.W., 2003. Cetane numbers of branched and straight-chain fatty esters determined in an ignition quality tester <sup>☆</sup>. *Fuel* 82, 971-975.
- [47] Kyndt, T., Assogbadjo, A.E., Hardy, O.J., Kakai, R.G., Sinsin, B., Van Damme, P., Gheysen, G., 2009. Spatial genetic structuring of baobab (*Adansonia digitata*, Malvaceae) in the traditional agroforestry systems of West Africa. *American Journal of Botany* 96, 950-957.
- [48] Ladommatos, N., Parsi, M., Knowles, A., 1996. The effect of fuel cetane improver on diesel pollutant emissions. *Fuel* 75, 8-14.
- [49] Lee, M., 1998. Genome projects and gene pools: New germplasm for plant breeding? *Proceedings of the National Academy of Sciences* 95, 2001-2004.
- [50] Liu, P., Yang, Y., Hao, C., Guo, W., 2007. Ecological risk assessment using RAPD and distribution pattern of a rare and endangered species. *Chemosphere* 68, 1497-1505.
- [51] Lynch, M., Milligan, B., 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3, 91-99.
- [52] Madsen, M., Schlundt, J., Omer, E., 1987. Effect of water coagulation by seeds of *Moringa oleifera* on bacterial concentrations. *The Journal of tropical medicine and hygiene* 90, 101.
- [53] Manoko, M.L., van den Berg, R.G., Feron, R.M., van der Weerden, G.M., Mariani, C., 2008. Genetic diversity of the African hexaploid species *Solanum scabrum* Mill. and *Solanum nigrum* L. (Solanaceae). *Genetic Resources and Crop Evolution* 55, 409-418.
- [54] Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer research* 27, 209-220.
- [55] Maroof, M.S., Biyashev, R., Yang, G., Zhang, Q., Allard, R., 1994. Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. *Proceedings of the National Academy of Sciences* 91, 5466-5470.
- [56] Mgendi, M., Manoko, M.K., Nyomora, A.M., 2010. Genetic diversity between cultivated and non-cultivated *Moringa oleifera* Lam. provenances assessed by RAPD markers. *J Cell Mol Biol* 8, 95-102.

- [57] Moghaddam, M., Mohammadi, S.A., Mohebalipour, N., Toorchi, M., Aharizad, S., Javidfar, F., 2009. Assessment of genetic diversity in rapeseed cultivars as revealed by RAPD and microsatellite markers. *Afr. J. Biotechnol* 8, 3160-3167.
- [58] Morton, J.F., 1991. The horseradish tree, *Moringa pterygosperma* (Moringaceae)—a boon to arid lands? *Economic Botany* 45, 318-333.
- [59] Mughal, M., Ali, G., Srivastava, P., Iqbal, M., 1999. Improvement of drumstick (*Moringa pterygosperma* Gaertn.)—a unique source of food and medicine through tissue culture. *Hamdard Med* 42, 37-42.
- [60] Muluvi, G., Sprent, J., Odee, D., Powell, W., 2004. Estimates of outcrossing rates in *Moringa oleifera* using Amplified fragment length polymorphism (AFLP). *African Journal of Biotechnology* 3, 145-151.
- [61] Muluvi, G., Sprent, J., Soranzo, N., Provan, J., Odee, D., Folkard, G., McNicol, J., Powell, W., 1999. Amplified fragment length polymorphism (AFLP) analysis of genetic variation in *Moringa oleifera* Lam. *Molecular Ecology* 8, 463-470.
- [62] Muyibi, S.A., Evison, L.M., 1995. Optimizing physical parameters affecting coagulation of turbid water with *Moringa oleifera* seeds. *Water Research* 29, 2689-2695.
- [63] Namkoong, G., 1998. The place of adaptation and production in genetic improvement programmes for sustainable development, Proc. 6th World Congr. Genet. Applied to Livest. Prod, pp. 69-79.
- [64] Ndabigengesere, A., Narasiah, K.S., Talbot, B.G., 1995. Active agents and mechanism of coagulation of turbid waters using *Moringa oleifera*. *Water Research* 29, 703-710.
- [65] Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583-590.
- [66] Okun, D., Kenya, E., Oballa, P., Odee, D., Muluvi, G., 2008. Analysis of genetic diversity in *Eucalyptus grandis* (Hill ex Maiden) seed sources using inter simple sequence repeats (ISSR) molecular markers. *African Journal of Biotechnology* 7.
- [67] Oliveira, E.J., Pádua, J.G., Zucchi, M.I., Vencovsky, R., Vieira, M.C., 2006. Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology* 29, 294-307.
- [68] Oliveira, J.T.A., Silveira, S.B., Vasconcelos, I.M., Cavada, B.S., Moreira, R.A., 1999. Compositional and nutritional attributes of seeds from the multiple purpose tree *Moringa oleifera* Lamarck. *Journal of the Science of Food and Agriculture* 79, 815-820.
- [69] Pal, S.K., Mukherjee, P.K., Saha, B., 1995. Studies on the antiulcer activity of *Moringa oleifera* leaf extract on gastric ulcer models in rats. *Phytotherapy Research* 9, 463-465.
- [70] Palada, M., Chang, L., 2003. Suggested cultivation practices for *Moringa*. AVRDC Publication# 03-545.
- [71] Panda, S., Kar, A., Sharma, P., Sharma, A., 2013. Cardioprotective potential of N,  $\alpha$ -l-rhamnopyranosyl vincosamide, an indole alkaloid, isolated from the leaves of *Moringa oleifera* in isoproterenol induced cardiotoxic rats: In vivo and in vitro studies. *Bioorganic & medicinal chemistry letters* 23, 959-962.
- [72] Pandey, A., Pradheep, K., Gupta, R., Nayar, E.R., Bhandari, D.C., 2011. 'Drumstick tree' (*Moringa oleifera* Lam.): a multipurpose potential species in India. *Genetic Resources and Crop Evolution* 58, 453-460.
- [73] Panwar, P., Nath, M., Yadav, V.K., Kumar, A., 2010. Comparative evaluation of genetic diversity using RAPD, SSR and cytochrome P450 gene based markers with respect to calcium content in finger millet (*Eleusine coracana* L. Gaertn.). *Journal of genetics* 89, 121-133.

- [74] Perez, T., Albornoz, J., Dominguez, A., 1998. An evaluation of RAPD fragment reproducibility and nature. *Molecular Ecology* 7, 1347-1357.
- [75] Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., Rafalski, A., 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular breeding* 2, 225-238.
- [76] Powell, W., Morgante, M., McDevitt, R., Vendramin, G., Rafalski, J., 1995. Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proceedings of the National Academy of Sciences* 92, 7759-7763.
- [77] Powell, W., Orozco-Castillo, C., Chalmers, K.J., Provan, J., Waugh, R., 1995. Polymerase chain reaction-based assays for the characterisation of plant genetic resources. *Electrophoresis* 16, 1726-1730.
- [78] Priadarshini, A., Pankaj, P.P., Varma, M., Kumar, K., 2013. Evaluation of the antibacterial potential of *Moringa oleifera* and *Azadirachta indica* against some pathogenic microbes: A comparative study. *Int. J. Drug Dev. & Res* 5, 214-218.
- [79] Rahman, M.S., Zerin, L., Anwar, M., 2013. Antibacterial and antifungal activity of *Moringa oleifera* stem bark. *Chittagong University Journal of Biological Sciences* 3, 109-117.
- [80] Ramachandran, C., Peter, K., Gopalakrishnan, P., 1980. Drumstick (*Moringa oleifera*): a multipurpose Indian vegetable. *Economic Botany* 34, 276-283.
- [81] Rashid, U., Anwar, F., Moser, B.R., Knothe, G., 2008. *Moringa oleifera* oil: A possible source of biodiesel. *Bioresource Technology* 99, 8175-8179.
- [82] Roloff, A., Weisgerber, H., Lang, U., Stimm, B., 2009. *Moringa oleifera* LAM., 1785. *Sea* 10, 10.
- [83] Rufai, S., Hanafi, M., Rafii, M., Ahmad, S., Arolu, I., Jannatul, F., 2013. Genetic Dissection of New Genotypes of Drumstick Tree (*Moringa oleifera* Lam.) Using Random Amplified Polymorphic DNA Marker. *BioMed Research International* 2013.
- [84] Saini, R., Saad, K., Ravishankar, G., Giridhar, P., Shetty, N., Genetic diversity of commercially grown *Moringa oleifera* Lam. cultivars from India by RAPD, ISSR and cytochrome P450-based markers. *Plant Systematics and Evolution*, 1-9.
- [85] Semagn, K., Bjørnstad, Å., Ndjiondjop, M., 2006. An overview of molecular marker methods for plants. *African Journal of Biotechnology* 5.
- [86] Semagn, K., Bjørnstad, A., Stedje, B., Bekele, E., 2000. Comparison of multivariate methods for the analysis of genetic resources and adaptation in *Phytolacca dodecandra* using RAPD. *Theoretical and Applied Genetics* 101, 1145-1154.
- [87] Shahzad, U., Khan, M.A., Jaskani, M.J., Khan, I.A., Korban, S.S., Genetic diversity and population structure of *Moringa oleifera*. *Conservation Genetics*, 1-12.
- [88] Soliva, C., Kreuzer, M., Foidl, N., Foidl, G., Machmüller, A., Hess, H., 2005. Feeding value of whole and extracted *Moringa oleifera* leaves for ruminants and their effects on ruminal fermentation in vitro. *Animal feed science and technology* 118, 47-62.
- [89] Somali, M., Bajneid, M., Al-Fhaimani, S., 1984. Chemical composition and characteristics of *Moringa peregrina* seeds and seeds oil. *Journal of the American Oil Chemists' Society* 61, 85-86.
- [90] Sutherland, J., Folkard, G., Grant, W., 1989. Seeds of *Moringa* species as naturally occurring flocculants for water treatment. *Science, Technology and Development* 7.
- [91] Tingey, S.V., del Tufo, J.P., 1993. Genetic analysis with random amplified polymorphic DNA markers. *Plant physiology* 101, 349.
- [92] Varalakshmi, B., 2006. Genetic Diversity in Drumstick (*Moringa oleifera* Lam) Germplasm, I International Conference on Indigenous Vegetables and Legumes. *Prospectus for Fighting Poverty, Hunger and Malnutrition* 752, pp. 411-412.
- [93] Wang, M., Sukumaran, S., Barkley, N., Chen, Z., Chen, C., Guo, B., Pittman, R., Stalker, H.T., Holbrook, C.C., Pederson, G., Yu, J., 2011. Population structure and marker-

trait association analysis of the US peanut (*Arachis hypogaea* L.) mini-core collection. *Theoretical and Applied Genetics* 123, 1307-1317.

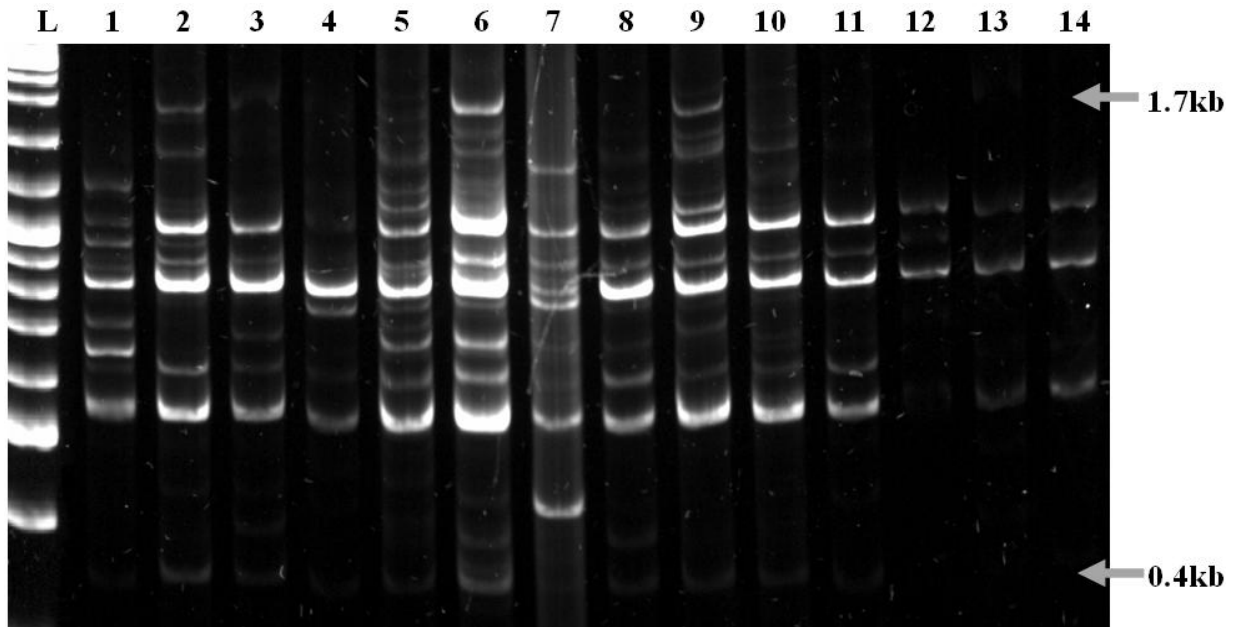
[94] Watson, R.T., Heywood, V.H., Baste, I., Dias, B., Gamez, R., Janetos, T., Reid, W., Ruark, G., 1995. Global biodiversity assessment.

[95] Wu, J.-C., Yang, J., Gu, Z.-J., Zhang, Y.-P., 2010. Isolation and characterization of twenty polymorphic microsatellite loci for *Moringa oleifera* (Moringaceae). *HortScience* 45, 690-692.

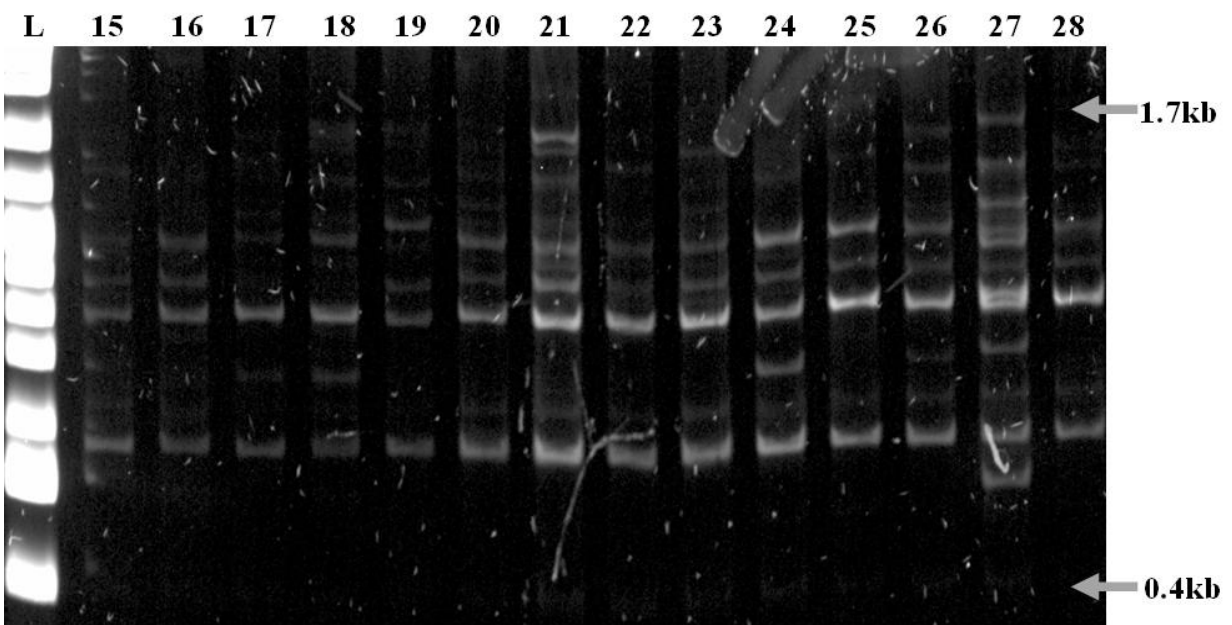
[96] Yeh, F.C., Yang, R.-c., Boyle, T.B., Ye, Z., Mao, J.X., 1997. POPGENE, the user-friendly shareware for population genetic analysis. *Molecular biology and biotechnology* centre, University of Alberta, Canada 10.

## Annex

**A1.1** – Gel electrophoresis of RAPD primer 5 on Moringa samples 1-14 (S.A.).

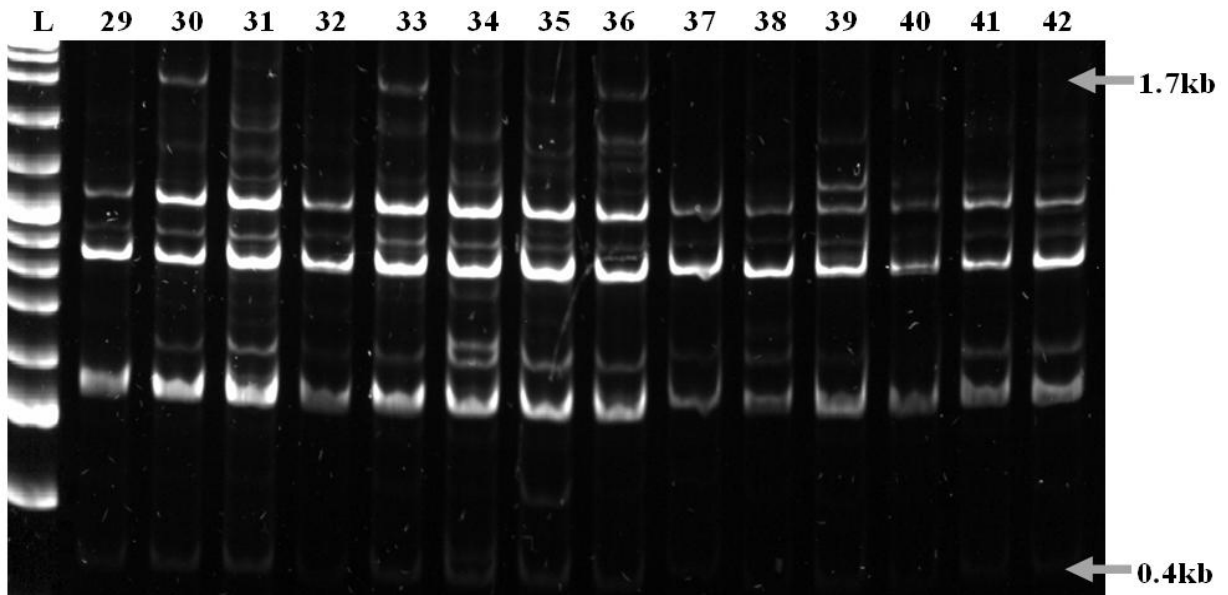


**A1.2** – Gel electrophoresis of RAPD primer 5 on Moringa samples 15-28 (S.A.).

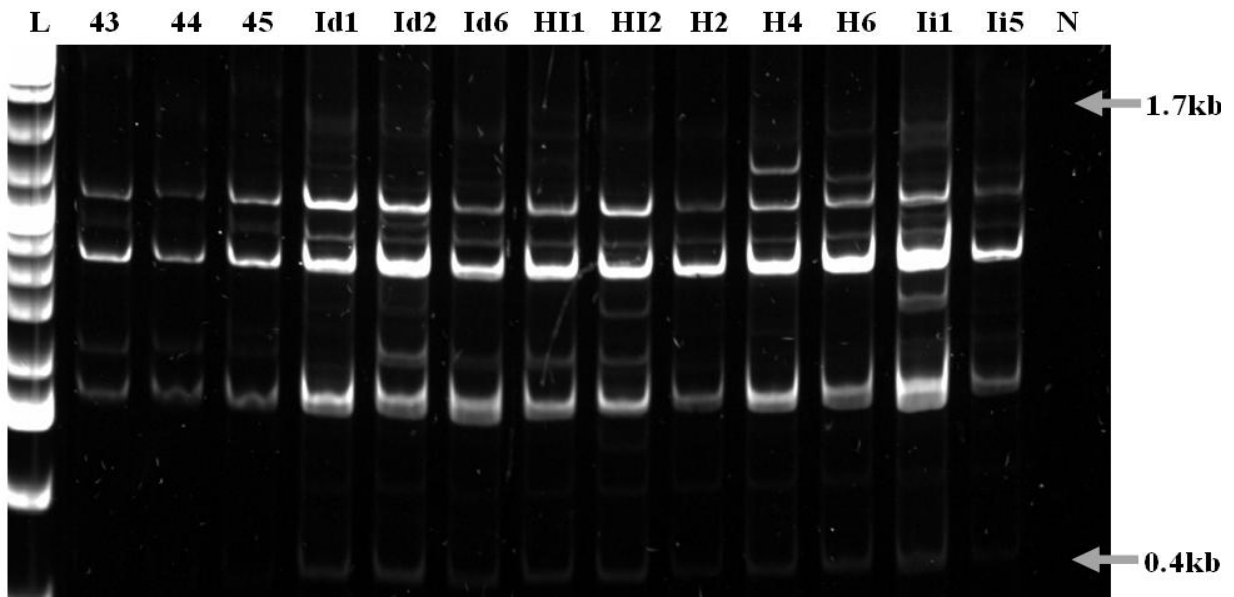




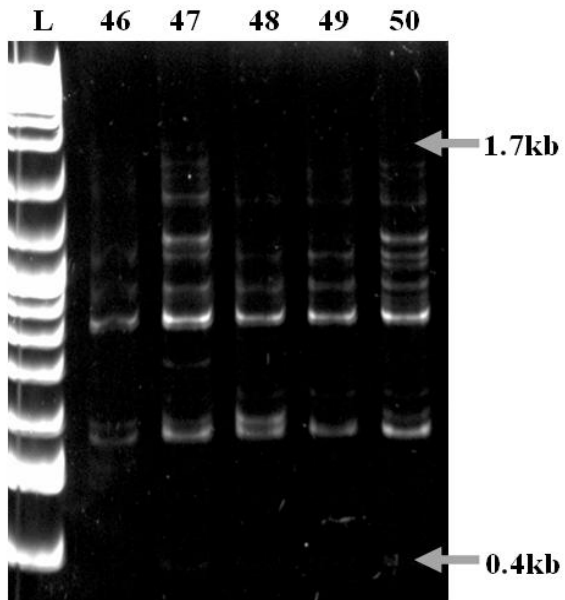
**A1.3** – Gel electrophoresis of RAPD primer 5 on Moringa samples 29-42 (S.A.).



**A1.4** – Gel electrophoresis of RAPD primer 5 on Moringa samples 43 to 45 (S.A.) and foreign samples.



**A1.5** – Gel electrophoresis of RAPD primer 5 on Moringa samples 46-50 (S.A.).



## A2.1 All RAPD primer's scoring table for Moringa samples 1-50.

Moringa samples	R1-1250bp	R1-880bp	R1-800bp	R1-680bp	R1-580bp	R3B-1700bp	R3B-1200bp	R3B-900bp	R3B-800bp	R3B-760bp	R3B-480bp	R3B-400bp	R4-1100bp	R4-900bp	R4-620bp	R5-1700bp	R5-1500bp
1	1	1	1	1	0	0	0	1	0	1	0	0	1	1	0	0	0
2	0	1	1	1	0	0	1	0	1	1	0	0	0	1	1	1	1
3	0	1	1	1	0	0	1	0	1	1	0	0	0	0	1	1	0
4	1	1	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0
5	1	1	0	0	0	1	1	0	1	1	0	0	1	0	0	0	1
6	0	1	1	1	1	0	1	1	1	1	0	0	1	0	0	1	1
7	0	1	1	1	1	0	1	1	0	1	0	0	0	0	0	0	1
8	0	1	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0
9	0	1	1	1	1	0	1	0	0	1	0	0	1	0	0	1	1
10	0	1	1	1	1	0	1	0	0	1	0	0	1	0	0	0	1
11	0	1	1	1	1	0	1	1	0	1	1	0	0	1	0	0	0
12	1	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0
13	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0
14	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
15	1	1	1	1	0	1	1	0	0	0	0	0	1	0	0	0	0
16	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
17	1	1	1	1	0	1	1	1	1	0	0	0	1	1	0	0	0
18	1	1	1	1	0	0	1	0	1	0	0	0	1	1	0	0	0
19	1	0	1	1	0	1	1	0	0	1	0	0	1	0	1	0	0
20	1	0	1	1	1	0	1	0	0	1	0	0	1	1	0	0	0
21	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1
22	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1
23	1	1	1	1	1	0	1	0	0	1	0	0	1	1	0	0	0
24	0	1	1	1	0	0	1	0	0	1	0	0	1	1	0	0	0
25	0	1	1	1	1	0	1	1	0	1	0	0	0	1	0	0	0
26	0	1	1	1	0	0	1	1	0	1	1	0	0	1	0	0	0
27	1	1	1	1	1	0	1	1	1	1	0	0	0	1	1	0	0
28	0	0	1	1	0	0	1	0	0	1	0	0	1	1	0	0	0
29	1	0	1	0	0	0	1	1	1	1	0	0	1	1	0	0	0
30	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	1	0
31	0	1	1	1	0	1	1	1	0	1	0	0	1	0	0	0	1
32	0	1	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0
33	0	1	1	1	0	1	1	0	1	1	0	0	0	0	0	1	0
34	1	0	1	1	1	1	1	0	1	1	0	0	0	1	0	0	0
35	1	1	1	0	0	1	1	1	1	1	0	0	1	1	0	0	0
36	0	0	1	1	0	1	1	1	0	1	0	0	1	1	0	1	0
37	0	0	1	1	0	1	1	0	0	1	0	0	0	1	0	0	0
38	0	1	1	0	0	0	1	1	0	1	1	0	0	1	0	0	0
39	1	1	1	1	0	1	1	0	0	1	0	0	1	1	0	0	1
40	1	1	1	1	0	0	1	0	0	1	0	0	1	1	0	1	0
41	0	1	1	1	0	0	1	0	0	1	0	0	0	1	0	0	0
42	0	1	1	1	0	0	1	0	0	1	0	0	0	1	0	0	0
43	0	1	1	1	0	0	1	1	1	1	0	0	1	0	0	1	0
44	0	1	1	1	0	0	1	0	1	0	0	0	1	1	0	0	0
45	0	1	1	1	0	0	1	0	0	1	0	0	1	1	1	0	0
46	0	1	0	1	0	0	1	1	0	1	0	0	1	1	0	1	1
47	1	1	1	0	1	0	0	0	1	1	0	1	1	1	1	0	1
48	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1
49	0	1	1	1	1	0	1	0	0	1	1	0	0	1	1	1	1
50	0	1	1	1	0	0	1	1	1	1	0	0	1	1	0	0	1

## A2.2 All RAPD primer's scoring table for Moringa samples 1-50.

Moringa samples	R5-1200bp	R5-1100bp	R5-1000bp	R5-900bp	R5-800bp	R5-750bp	R5-710bp	R5-680bp	R5-640bp	R5-600bp	R5-550bp	R5-470bp	R5-400bp	R6-420bp	R6-400bp	R6-300bp	R6-260bp
1	1	0	1	0	0	1	0	1	1	1	1	0	0	0	1	0	1
2	0	1	1	1	0	0	0	0	1	0	1	0	0	1	0	1	0
3	0	1	0	1	0	0	1	0	1	0	1	0	0	1	0	0	0
4	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
5	1	1	0	1	1	1	1	0	1	0	1	0	0	0	1	0	0
6	1	1	0	1	1	0	1	0	1	0	1	0	0	1	1	1	0
7	0	1	0	1	1	0	1	0	1	0	1	0	1	1	0	0	1
8	0	1	0	1	1	0	1	0	1	0	1	0	0	0	0	0	0
9	0	1	0	1	1	1	0	0	1	0	1	0	0	1	0	0	0
10	0	1	0	1	1	0	1	1	1	0	1	0	0	1	0	0	0
11	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0
12	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	1
13	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
14	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
15	1	1	1	1	1	0	1	0	1	0	1	1	0	0	0	1	0
16	0	1	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0
17	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
18	0	1	0	1	0	0	1	0	1	0	1	0	0	0	1	1	1
19	0	1	0	1	0	1	0	0	1	0	1	0	0	0	0	1	0
20	0	1	0	1	0	0	0	1	1	0	1	0	0	1	0	1	0
21	0	1	0	1	1	0	1	0	1	0	1	0	0	1	0	1	0
22	0	1	0	1	0	0	1	0	1	0	1	0	0	0	0	1	0
23	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	1	1
24	0	1	0	1	1	1	0	0	1	0	1	0	0	1	0	1	0
25	0	1	0	1	0	0	1	1	1	1	1	0	0	0	0	0	0
26	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0
27	0	1	0	1	0	0	1	1	1	0	1	1	0	1	0	1	0
28	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1	0
29	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1
30	1	1	0	1	1	1	0	0	1	0	1	0	0	1	1	0	0
31	1	1	0	1	1	1	1	0	1	0	1	0	0	0	1	1	0
32	0	1	0	1	0	1	0	0	1	0	1	0	0	1	0	0	0
33	0	1	0	1	1	0	0	0	1	0	1	0	0	1	0	0	0
34	0	1	0	1	0	0	0	1	1	0	1	0	0	1	0	1	0
35	1	1	1	1	1	1	0	0	1	0	1	0	1	0	0	1	1
36	1	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1	0
37	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0
38	0	1	0	1	0	0	1	0	1	0	1	0	0	0	0	1	0
39	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0
40	0	1	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0
41	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0
42	1	1	1	1	0	0	0	0	1	0	1	0	0	1	1	0	0
43	0	1	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0
44	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0
45	0	1	0	1	0	0	0	1	1	0	1	0	0	1	0	0	0
46	1	1	0	1	0	0	1	0	0	1	1	0	1	1	1	0	0
47	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0
48	0	1	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0
49	1	1	1	1	0	0	1	0	1	1	1	0	1	1	0	0	1
50	0	1	1	1	0	0	0	0	1	0	1	0	1	0	1	0	1

### A2.3 All RAPD primer's scoring table for Moringa samples 1-50.

Moringa samples	R6-250bp	R6-200bp	R7-1600bp	R7-1500bp	R7-1100bp	R7-1000bp	R7-950bp	R7-850bp	R7-750bp	R7-700bp	R9-2500bp	R9-1800bp	R9-1300bp	R9-1200bp	R9-900bp	R9-800bp	R9-640bp
1	0	0	0	1	1	1	1	1	0	1	1	1	0	1	0	1	1
2	1	1	1	1	1	1	1	0	0	1	1	1	0	1	0	0	1
3	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1
4	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	1	1
5	0	1	0	1	1	1	0	0	0	1	1	1	0	0	0	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
7	0	1	0	1	0	1	1	1	0	1	1	1	1	1	0	0	0
8	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0	1	0
9	1	0	1	1	0	1	1	0	1	1	1	1	1	1	0	1	0
10	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	0
11	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1
12	0	1	0	1	1	0	0	0	0	0	1	1	0	1	0	1	0
13	0	1	0	0	0	0	0	1	0	0	1	1	1	1	1	1	0
14	1	0	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1
15	1	0	0	1	1	1	0	0	0	1	1	1	0	1	1	1	1
16	0	1	1	1	1	1	0	1	0	0	1	1	0	0	0	0	0
17	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	1	0
18	1	0	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1
19	0	0	0	1	1	0	0	0	1	1	1	1	0	1	0	1	1
20	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1
21	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1
22	0	0	0	1	1	1	1	0	0	1	1	1	1	1	0	1	1
23	1	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1
24	1	0	1	1	0	1	1	1	0	1	1	1	1	1	0	1	0
25	0	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	0
26	1	0	0	1	1	1	1	1	0	1	1	1	1	0	0	1	0
27	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1
28	0	0	1	1	0	0	0	0	0	1	1	1	1	1	0	1	1
29	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1
30	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1
31	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	1	1
32	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0
33	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1
34	1	0	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1
35	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	1
36	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1
37	1	0	1	1	0	1	1	1	0	1	1	1	1	1	0	0	0
38	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1
39	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1
40	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
41	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1
42	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	1
43	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0
44	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1
45	0	0	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1
46	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	1
47	1	0	1	1	0	1	1	1	0	1	1	1	1	1	0	0	0
48	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1
49	0	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1
50	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1

## A2.4 All RAPD primer's scoring table for Moringa samples 1-50.

Moringa samples	R9-600bp	R10-1800bp	R10-1400bp	R10-1300bp	R10-720bp	R10-680bp	R10-650bp	R11-2000bp	R11-1600bp	R11-1550bp	R11-1190bp	R11-1170bp	R11-1150bp	R11-880bp
1	1	1	1	1	1	0	0	1	1	1	1	1	0	1
2	1	1	0	0	0	0	1	1	1	1	1	1	1	1
3	1	1	0	0	0	1	0	1	1	1	1	1	1	1
4	1	0	1	1	1	1	0	1	1	1	1	0	0	0
5	1	0	1	0	1	1	0	1	1	0	1	1	1	1
6	1	0	1	0	0	0	1	1	1	1	1	1	1	1
7	1	0	0	1	0	1	0	1	0	0	1	1	1	0
8	1	0	0	0	0	1	0	0	0	0	1	1	0	0
9	1	0	0	0	0	1	0	1	1	1	1	1	1	1
10	1	0	0	0	0	1	0	1	1	1	1	1	1	1
11	1	0	0	0	0	1	0	1	1	1	1	1	0	1
12	0	0	0	1	0	1	0	1	1	1	1	0	0	1
13	0	1	1	1	1	1	1	1	1	1	1	1	0	1
14	1	1	0	0	0	1	0	1	1	1	1	1	1	1
15	1	0	1	0	1	1	1	1	1	1	1	1	0	0
16	1	0	1	0	0	1	0	1	1	1	1	1	1	0
17	1	0	1	1	0	1	0	1	1	0	1	1	0	1
18	1	0	0	1	0	1	0	1	1	1	1	1	1	1
19	0	0	1	1	0	1	0	1	1	1	1	1	1	0
20	1	1	1	0	0	1	0	1	1	1	1	1	0	0
21	0	1	1	0	0	1	0	1	1	1	1	1	1	1
22	1	1	1	0	0	1	0	1	1	1	1	1	0	1
23	1	1	1	0	0	1	0	1	1	1	1	1	0	0
24	1	1	1	0	0	1	0	1	1	1	1	0	0	1
25	0	1	1	1	1	1	1	1	1	1	1	1	0	1
26	1	0	0	0	0	0	0	1	1	1	1	1	0	0
27	0	1	0	1	0	1	0	1	1	1	1	1	0	1
28	0	1	1	0	0	1	0	1	1	1	1	0	0	0
29	0	1	0	1	1	1	1	1	1	1	1	1	1	1
30	1	1	0	1	0	0	1	1	1	1	1	1	1	0
31	1	1	0	1	0	0	1	1	1	1	1	1	1	1
32	0	0	0	0	1	0	0	1	1	1	1	1	1	1
33	1	1	0	0	0	1	0	1	1	1	1	1	1	1
34	1	1	0	0	0	1	0	1	1	1	1	1	1	1
35	0	1	0	1	0	0	1	1	1	1	1	0	1	1
36	0	0	0	0	0	0	1	1	1	1	1	0	1	1
37	0	1	1	0	1	0	1	1	1	0	1	1	1	1
38	0	1	1	1	0	0	1	1	1	1	1	1	1	1
39	1	1	0	1	0	0	1	1	1	1	1	1	1	1
40	1	1	0	1	1	1	1	1	1	1	1	1	0	0
41	1	1	1	1	0	1	0	1	1	1	1	1	1	1
42	1	1	0	1	0	0	1	1	1	1	1	1	1	1
43	0	1	0	1	0	1	0	1	1	1	1	1	1	0
44	1	1	0	1	0	1	0	1	1	1	1	1	1	0
45	1	1	1	1	0	1	0	1	1	1	1	0	1	0
46	0	0	0	0	0	1	1	1	1	1	1	1	0	1
47	0	1	1	0	1	0	1	1	0	1	1	1	0	0
48	0	1	1	1	1	0	1	1	1	1	1	1	1	1
49	1	1	0	0	0	0	1	1	1	1	0	1	1	0
50	0	1	0	1	1	1	1	1	1	1	1	1	1	1

### A2.5 All RAPD primer's scoring table for Moringa samples foreign samples.

Moringa samples	R1-1250bp	R1-880bp	R1-800bp	R1-680bp	R1-580bp	R3B-1700bp	R3B-1200bp	R3B-900bp	R3B-800bp	R3B-760bp	R3B-480bp	R3B-400bp	R4-1100bp	R4-900bp	R4-620bp	R5-1700bp	R5-1500bp
ID1	1	1	1	1	0	1	1	0	1	1	0	1	0	0	0	0	1
ID2	1	1	1	1	0	1	1	0	1	1	0	0	0	1	1	0	1
ID6	1	1	1	1	0	1	1	0	1	1	0	0	1	0	0	0	1
ID7	0	1	1	1	1	1	1	0	1	1	0	0	1	1	0	0	0
ID8	1	1	0	1	1	1	0	0	1	0	0	1	0	0	0	0	0
ID9	0	1	1	1	0	0	1	1	1	1	0	0	0	1	1	0	0
HI1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1	0	1
HI2	1	1	1	1	0	1	1	1	0	1	0	0	0	0	1	0	1
HI4	0	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1
HI7	1	0	1	1	0	1	0	0	1	1	0	0	0	0	1	1	1
HI8	1	0	1	1	0	0	1	1	0	1	1	0	0	0	0	1	1
HI9	1	1	1	0	0	1	1	1	0	0	0	0	1	0	1	1	1
H2	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	0	1
H4	1	1	1	1	0	1	1	1	0	1	0	1	1	0	0	0	1
H6	1	0	0	1	0	1	1	0	1	1	0	1	1	0	1	0	1
H7	1	1	1	1	0	0	1	0	1	0	0	1	1	0	0	1	1
H8	1	0	1	0	1	1	1	0	1	1	0	1	1	0	1	0	1
H9	0	0	0	1	0	1	1	1	0	1	0	0	0	1	1	0	1
I1	1	1	1	1	0	1	1	1	0	1	0	1	1	1	1	0	1
I5	1	1	1	1	0	1	1	0	1	1	0	1	1	1	0	0	1
I7	0	0	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1

### A2.6 All RAPD primer's scoring table for Moringa samples foreign samples.

Moringa samples	R5-1200bp	R5-1100bp	R5-1000bp	R5-900bp	R5-800bp	R5-750bp	R5-710bp	R5-680bp	R5-640bp	R5-600bp	R5-550bp	R5-470bp	R5-400bp	R6-420bp	R6-400bp	R6-300bp	R6-260bp
ID1	1	1	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0
ID2	0	1	1	1	0	1	0	0	1	0	1	0	1	0	0	0	0
ID6	1	1	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0
ID7	1	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
ID8	0	1	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0
ID9	0	1	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0
HI1	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0	1	0
HI2	0	1	0	1	0	1	0	0	1	0	1	1	1	0	0	1	0
HI4	0	1	0	1	0	1	1	0	1	0	1	0	1	0	0	1	0
HI7	0	1	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0
HI8	1	1	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0
HI9	0	1	0	0	0	0	1	1	0	1	1	0	1	0	1	1	0
H2	0	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0
H4	1	1	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0
H6	0	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0
H7	0	1	1	0	0	0	0	0	0	1	1	0	1	0	1	0	1
H8	0	1	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0
H9	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0
I1	1	1	1	1	0	1	0	0	1	0	1	0	1	0	0	0	0
I5	0	1	0	1	0	0	1	0	1	0	1	0	1	0	0	0	0
I7	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1

### A2.7 All RAPD primer's scoring table for Moringa samples foreign samples.

	R6-250bp	R6-200bp	R7-1600bp	R7-1500bp	R7-1100bp	R7-1000bp	R7-950bp	R7-850bp	R7-750bp	R7-700bp	R9-2500bp	R9-1800bp	R9-1300bp	R9-1200bp	R9-900bp	R9-800bp	R9-640bp
ID1	1	0	0	1	1	0	0	0	1	1	1	1	0	1	1	1	1
ID2	1	0	0	1	1	0	1	0	1	1	1	1	1	1	0	1	1
ID6	1	0	0	1	0	0	1	0	1	1	1	1	1	1	0	1	1
ID7	0	0	0	1	1	0	1	0	0	1	1	1	1	1	0	0	1
ID8	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	1	0
ID9	1	0	0	1	0	0	0	0	1	1	1	1	1	1	1	1	1
HI1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HI2	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HI4	1	0	0	1	0	0	1	1	1	1	1	1	0	1	0	1	1
HI7	0	0	0	1	0	1	1	1	1	0	1	1	0	1	0	1	1
HI8	0	1	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1
HI9	1	0	1	1	1	1	1	1	0	1	0	1	1	1	0	1	1
H2	1	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1
H4	0	0	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1
H6	1	0	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1
H7	1	0	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
H8	1	0	1	1	1	1	1	1	0	1	1	1	0	1	0	1	0
H9	0	1	1	1	0	0	0	1	1	0	1	1	0	1	0	1	1
I1	1	0	1	1	0	0	1	1	1	1	1	0	0	1	0	1	1
I5	1	0	1	1	0	0	1	1	1	1	0	0	0	1	0	1	1
I7	1	0	0	1	1	0	1	0	1	1	1	0	1	1	0	1	1

### A2.8 All RAPD primer's scoring table for Moringa samples foreign samples.

	R9-600bp	R10-1800bp	R10-1400bp	R10-1300bp	R10-720bp	R10-680bp	R10-650bp	R11-2000bp	R11-1600bp	R11-1550bp	R11-1190bp	R11-1170bp	R11-1150bp	R11-880bp
ID1	1	1	0	1	0	0	1	1	0	1	1	1	1	0
ID2	1	1	0	1	0	0	1	1	1	1	1	1	0	0
ID6	1	1	0	1	0	0	1	1	1	1	1	1	0	0
ID7	0	1	0	1	0	0	1	1	0	1	1	1	1	0
ID8	1	1	0	1	0	0	1	1	0	0	1	0	0	1
ID9	1	1	0	1	1	0	1	1	1	1	1	1	1	0
HI1	1	1	0	1	1	1	1	1	1	1	1	1	1	0
HI2	1	1	0	1	0	0	1	1	1	1	1	1	1	0
HI4	1	1	0	1	1	0	0	1	0	1	1	1	1	0
HI7	0	1	0	1	0	1	1	1	1	1	1	1	1	0
HI8	1	1	0	1	0	0	0	1	1	1	0	0	1	0
HI9	1	1	0	1	1	0	1	1	1	1	1	1	1	0
H2	1	1	0	1	1	1	1	1	0	1	1	0	1	1
H4	1	1	0	1	0	1	0	1	1	1	1	0	1	1
H6	1	1	0	1	0	1	0	1	1	0	1	0	1	1
H7	1	1	0	1	0	1	0	1	1	1	1	0	1	1
H8	1	0	1	0	1	1	1	1	0	1	1	0	1	1
H9	0	1	0	1	0	0	0	1	1	1	1	1	1	0
I1	1	1	1	1	1	0	0	1	1	0	1	1	1	1
I5	1	1	1	1	1	1	1	1	1	0	1	1	1	1
I7	1	1	1	1	0	1	1	1	1	1	1	0	0	1





#### A4.1 SSR allele data set.

	SSR7	SSR7	SSR9	SSR9	SSR23	SSR23	SSR10	SSR10	SSR15	SSR15	SSR20	SSR20
1	287	291	197	197	159	171	133	133	127	127	124	124
2	279	279	205	205	159	159	133	133	121	129	112	120
3	283	283	197	197	161	171	133	133	121	129	112	112
4	283	289	197	197	159	159	131	131	127	127	118	118
5	283	283	197	197	159	171	131	131	127	127	118	124
6	279	279	195	205	161	171	133	133	127	129	120	120
7	289	283	195	197	169	171	131	131	121	129	112	120
8	279	289	195	205	159	171	131	133	129	129	118	120
9	285	281	205	195	171	171	133	133	129	121	112	112
10	285	285	195	197	169	169	131	143	129	129	118	120
11	285	285	197	195	169	169	131	143	129	129	112	112
12	285	285	195	197	169	169	131	143	129	129	112	120
13	285	285	197	195	169	169	131	143	125	129	112	120
14	287	287	195	197	165	171	131	143	125	129	112	120
15	291	287	197	197	159	183	133	133	127	127	118	124
16	285	285	195	197	165	171	131	143	127	129	118	122
17	285	291	197	197	159	159	133	121	127	127	118	124
18	285	291	197	197	159	159	131	131	115	127	112	124
19	285	291	197	197	159	159	131	131	119	125	124	118
20	281	285	195	205	159	171	131	133	123	125	112	114
21	281	285	195	197	159	171	133	133	121	129	112	120
22	283	289	197	197	159	159	131	131	127	127	118	124
23	295	291	197	197	159	159	133	133	121	121	112	118
24	289	289	197	197	159	171	133	133	121	129	112	112
25	285	281	195	205	159	171	133	133	121	129	112	120
26	281	289	197	197	171	171	133	133	121	121	112	118
27	281	281	195	205	159	171	133	133	121	121	112	112
28	279	289	197	205	159	171	131	133	121	121	110	112
29	289	283	197	197	159	159	131	131	129	129	124	124
30	279	285	195	197	169	171	131	133	121	129	118	120
31	283	289	195	205	159	171	131	133	121	129	112	120
32	279	279	197	195	169	171	131	133	121	129	112	120
33	279	279	197	205	171	171	131	133	121	129	112	120
34	279	283	195	205	159	171	131	133	121	125	112	112
35	287	285	197	205	181	171	133	133	121	121	112	112
36	279	279	205	197	171	159	131	133	121	129	112	120
37	283	287	197	197	159	171	131	133	129	121	112	120
38	283	291	205	195	159	171	133	133	121	121	112	112
39	283	289	197	197	159	159	131	131	129	129	118	124
40	283	289	197	197	159	159	131	131	129	129	124	124
41	279	289	197	205	171	171	133	133	121	121	112	120
42	279	279	195	195	171	171	133	133	129	121	112	112
43	279	279	197	205	171	171	133	133	121	121	112	120
44	289	179	205	195	159	159	133	133	121	121	112	112
45	283	289	197	197	159	171	133	133	129	129	112	120
46	289	289	197	197	159	159	131	131	121	121	122	124
47	289	283	197	197	159	159	131	131	129	121	124	118
48	283	283	197	195	171	171	131	131	121	121	122	124
49	289	289	197	197	171	171	131	131	121	121	118	124
50	289	293	195	197	171	171	131	131	129	129	118	118
ID1	283	287	195	205	171	161	125	125	121	121	118	114
ID2	283	283	195	195	169	171	131	131	121	121	118	114
ID6	283	283	195	195	157	171	131	131	121	121	118	114
ID7	287	281	203	203	171	171	133	133	0	0	112	118
ID8	285	285	207	195	157	157	131	131	0	0	112	112
ID9	285	283	207	207	157	171	131	131	0	0	120	114
I1	287	285	203	197	159	171	133	133	129	129	114	120
I5	285	285	197	197	171	171	131	131	129	129	120	120
I7	285	285	197	205	161	159	125	125	0	0	112	120
HI1	287	287	205	197	171	171	125	125	129	121	120	118
HI2	283	287	205	203	171	171	125	125	125	121	118	112
HI4	283	283	197	205	171	161	133	125	121	129	118	112
HI7	285	283	205	205	171	161	133	133	129	121	112	118
HI8	287	283	205	205	171	161	125	125	125	121	120	112
HI9	283	287	197	205	169	171	133	125	121	129	112	112
H2	285	285	197	197	171	171	125	125	129	129	120	118
H4	285	287	195	197	171	171	125	125	129	115	114	120
H6	285	289	197	197	161	171	125	125	129	129	120	118
H7	285	285	197	205	171	171	125	125	129	129	114	120
H8	285	295	197	197	171	159	131	125	129	115	118	120
H9	285	295	197	197	171	159	131	131	129	129	118	120

#### A4.2 SSR allele data set.

	SSR6	SSR6	SSR18	SSR18	SSR8	SSR8	SSR5a	SSR5a	SSR16	SSR16	SSR12	SSR12
1	100	100	83	83	164	164	133	137	142	142	100	102
2	100	108	83	83	162	164	123	145	128	130	100	102
3	100	108	91	83	162	162	123	133	128	142	0	0
4	100	108	83	85	162	164	137	137	142	142	102	102
5	100	100	83	83	164	162	133	137	138	142	100	102
6	100	108	83	83	162	162	123	123	128	144	102	102
7	100	108	83	83	162	162	133	123	130	144	100	102
8	100	108	83	83	162	164	123	123	128	136	102	102
9	100	108	83	83	164	164	123	123	128	136	102	102
10	104	104	83	91	162	162	123	123	128	142	102	102
11	104	104	83	91	164	164	123	123	130	142	102	102
12	104	106	83	91	162	164	123	133	128	128	102	102
13	104	104	83	91	162	164	123	123	130	144	102	102
14	102	104	83	91	162	164	123	145	128	142	102	102
15	100	108	83	83	164	164	123	137	128	142	100	102
16	104	104	83	91	154	164	135	137	142	142	100	102
17	100	100	83	83	154	164	135	137	142	142	100	102
18	100	108	83	83	154	164	123	137	142	142	102	102
19	108	108	83	83	164	164	133	137	142	142	102	102
20	100	108	83	83	162	162	123	133	130	144	102	102
21	100	108	83	83	162	162	123	133	128	144	102	102
22	100	100	83	83	154	164	123	133	140	140	100	102
23	104	104	83	83	164	164	133	133	128	144	102	102
24	108	108	83	83	162	162	123	133	128	130	102	102
25	100	108	83	87	162	162	123	123	128	144	102	102
26	100	108	83	83	164	164	123	123	130	144	102	102
27	100	108	83	83	164	164	145	123	130	144	102	102
28	104	108	83	83	164	162	123	133	128	128	102	102
29	100	100	83	85	164	154	137	137	128	142	102	102
30	100	108	83	83	164	162	123	123	128	142	102	102
31	100	108	83	95	162	162	123	133	130	142	102	102
32	100	108	83	83	162	164	133	123	128	142	102	102
33	100	108	83	83	162	164	123	123	130	142	102	102
34	100	108	83	83	162	162	123	133	130	142	102	102
35	100	100	83	83	162	162	123	123	134	134	100	102
36	104	108	83	83	162	162	133	123	128	128	102	102
37	100	108	85	85	162	164	123	123	130	136	102	102
38	108	100	83	83	164	162	123	123	142	130	102	102
39	108	108	85	85	162	164	133	133	132	140	102	102
40	100	108	83	83	162	164	137	137	140	140	100	100
41	100	108	83	83	162	164	123	123	128	142	102	102
42	102	102	91	83	164	164	123	133	144	144	102	102
43	102	110	83	83	164	164	123	123	130	130	102	102
44	102	110	83	83	164	164	123	123	130	144	102	102
45	102	110	83	83	164	164	123	123	130	144	102	102
46	100	108	83	87	164	164	133	133	142	142	100	100
47	100	108	83	87	164	154	133	133	142	142	106	102
48	100	100	83	87	164	154	123	123	142	128	100	102
49	100	108	83	87	164	164	123	133	144	128	100	102
50	100	104	83	83	164	164	123	133	128	128	102	102
ID1	100	100	81	81	162	162	123	131	134	128	106	106
ID2	100	100	89	89	164	164	123	133	128	128	102	102
ID6	100	100	81	81	164	164	133	131	128	128	102	102
ID7	100	100	89	89	164	162	133	135	128	136	102	102
ID8	100	100	83	83	164	164	133	133	128	128	102	102
ID9	100	100	89	89	154	154	123	123	128	134	102	106
I1	104	104	91	91	154	164	123	133	128	144	106	102
I5	104	104	91	91	164	164	133	133	144	134	102	102
I7	100	104	91	91	164	164	123	133	134	128	102	102
HI1	100	104	87	87	162	162	133	137	134	146	102	102
HI2	100	100	83	83	164	164	137	123	134	146	100	102
HI4	100	108	91	89	164	164	135	123	128	128	102	102
HI7	100	100	83	83	162	162	135	135	146	144	102	102
HI8	100	100	83	83	162	162	123	123	146	136	102	106
HI9	100	100	83	89	164	164	123	133	146	146	100	102
H2	104	104	91	91	164	164	133	133	134	134	102	102
H4	104	100	91	91	154	164	133	123	128	128	102	102
H6	104	106	91	91	164	164	133	133	128	134	102	102
H7	104	104	91	91	164	164	133	133	134	144	100	102
H8	102	104	91	91	164	164	133	133	128	134	102	102
H9	104	102	91	89	164	164	133	133	128	134	100	102

## A5.1 SSR converted binary data set.

Moringa samples	7-279	7-281	7-283	7-285	7-287	7-289	7-291	7-293	7-295	9-195	9-197	9-203	9-205	9-207	23-157	23-159	23-161	23-165	23-169	23-171	23-181	23-183
1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
7	0	0	1	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0
8	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
9	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0
10	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
11	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
12	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
13	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0
15	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
16	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0
17	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
18	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
19	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
20	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0
21	0	1	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0
22	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
23	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0
24	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
25	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	1	0
26	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
27	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0
28	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	1	0	0
29	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
30	1	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0
31	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0
32	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0
33	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
34	1	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0
35	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	1	0
36	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0
37	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0
38	0	0	1	0	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0	1	0	0
39	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
40	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
41	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0
42	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
43	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0
44	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0
45	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0
46	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
47	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
48	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0
49	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
50	0	0	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	1	0	0
ID1	0	0	1	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0
ID2	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0
ID6	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0
ID7	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
ID8	0	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0
ID9	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0
HI1	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0
HI2	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
HI4	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0
HI7	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0
HI8	0	0	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0
HI9	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0
H2	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0
H4	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0
H6	0	0	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	1	0	0
H7	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
H8	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0
H9	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0
I1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0
I5	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0
I7	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0

## A5.2 SSR converted binary data set.

Moringa samples	10-121	10-125	10-131	10-133	10-143	15-115	15-119	15-121	15-123	15-125	15-127	15-129	20-110	20-112	20-114	20-118	20-120	20-122	20-124	6-100	6-102	6-104	6-106
1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
4	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
8	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
9	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
10	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
11	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
12	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
13	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	0	1	0
14	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	1	1	0
15	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0
16	0	0	1	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0	1	0
17	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0
18	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0
19	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
20	0	0	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0	0
21	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
22	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0
23	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0
24	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
25	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0
26	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0
27	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
28	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0
29	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0
30	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	1	0	0	0
31	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0
32	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0
33	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0
34	0	0	1	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0
35	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
36	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0	0	1	0	0
37	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0
38	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
39	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0
40	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0
41	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
42	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0
43	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
44	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
45	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	1	0	0
46	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
47	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0
48	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
49	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0
50	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0
ID1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0
ID2	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0
ID6	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0
ID7	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0
ID8	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
ID9	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0
HI1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	1	0
HI2	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0
HI4	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0
HI7	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	1	0	1	0
HI8	0	1	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0
HI9	0	1	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
H2	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0
H4	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0	1	0	0	0
H6	0	1	0	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
H7	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	0
H8	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	1	0
H9	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	1
I1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	1	0
I5	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	1	0
I7	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	1	0

### A5.3 SSR converted binary data set.

Morninga samples	6-108	6-110	18-81	18-83	18-85	18-87	18-89	18-91	18-95	8-154	8-162	8-164	5a-123	5a-131	5a-133	5a-135	5a-137	5a-145	16-128	16-130	16-132	16-134	16-136
1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
2	1	0	0	1	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0
3	1	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0
4	1	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
6	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
7	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
8	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	1
9	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	1
10	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0
11	0	0	0	1	0	0	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0
12	0	0	0	1	0	0	0	1	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0
13	0	0	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
14	0	0	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	1	1	0	0	0	0
15	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0	0
16	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0
17	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0
18	1	0	0	1	0	0	0	0	0	1	0	1	1	0	0	0	1	0	0	0	0	0	0
19	1	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0
20	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0
21	1	0	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0
22	0	0	0	1	0	0	0	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0
23	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
24	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	1	0	0	0
25	1	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0
26	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
27	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	1	0	0	0
28	1	0	0	1	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0
29	0	0	0	1	1	0	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0
30	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
31	1	0	0	1	0	0	0	0	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0
32	1	0	0	1	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0
33	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
34	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0
35	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0
36	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0
37	1	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0	1
38	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1
39	1	0	0	0	1	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0	0	0
40	1	0	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0
41	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0
42	0	0	0	1	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0
43	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
44	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
45	0	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0
46	1	0	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
47	1	0	0	1	0	1	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0
48	0	0	0	1	0	1	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0
49	1	0	0	1	0	1	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0
50	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0
ID1	0	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	0	0	1	0	0
ID2	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	1	0	0	0	0
ID6	0	0	1	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	0	0	0	0
ID7	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	1	0	0	1	0	0	0	1
ID8	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0
ID9	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0
HI1	0	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	0	1	0	0	0	0
HI2	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0
HI4	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	0	1	0
HI7	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0
HI8	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	1	0
HI9	1	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	0	0	1	0	0	0	0
H2	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
H4	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1
H6	0	0	0	1	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0
H7	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0
H8	0	0	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	0	1	0	0	0	0
H9	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0
I1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0
15	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0
17	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	1	0	0	1	0

### A5.4 SSR converted binary data set.

Mornings samples	16-138	16-140	16-142	16-144	16-146	12-100	12-102	12-106
1	0	0	1	0	0	1	1	0
2	0	0	0	0	0	1	1	0
3	0	0	1	0	0	0	1	0
4	0	0	1	0	0	0	1	0
5	1	0	1	0	0	1	1	0
6	0	0	0	1	0	0	1	0
7	0	0	0	1	0	1	1	0
8	0	0	0	0	0	0	1	0
9	0	0	0	0	0	0	1	0
10	0	0	1	0	0	0	1	0
11	0	0	1	0	0	0	1	0
12	0	0	0	0	0	0	1	0
13	0	0	0	1	0	0	1	0
14	0	0	1	0	0	0	1	0
15	0	0	1	0	0	1	1	0
16	0	0	1	0	0	1	1	0
17	0	0	1	0	0	1	1	0
18	0	0	1	0	0	0	1	0
19	0	0	1	0	0	0	1	0
20	0	0	0	1	0	0	1	0
21	0	0	0	1	0	0	1	0
22	0	1	0	0	0	1	1	0
23	0	0	0	1	0	0	1	0
24	0	0	0	0	0	0	1	0
25	0	0	0	1	0	0	1	0
26	0	0	0	1	0	0	1	0
27	0	0	0	1	0	0	1	0
28	0	0	0	0	0	0	1	0
29	0	0	1	0	0	0	1	0
30	0	0	1	0	0	0	1	0
31	0	0	1	0	0	0	1	0
32	0	0	1	0	0	0	1	0
33	0	0	1	0	0	0	1	0
34	0	0	1	0	0	0	1	0
35	0	0	0	0	0	1	1	0
36	0	0	0	0	0	0	1	0
37	0	0	0	0	0	0	1	0
38	0	0	1	0	0	0	1	0
39	0	1	0	0	0	0	1	0
40	0	1	0	0	0	1	0	0
41	0	0	1	0	0	0	1	0
42	0	0	0	1	0	0	1	0
43	0	0	0	0	0	0	1	0
44	0	0	0	1	0	0	1	0
45	0	0	0	1	0	0	1	0
46	0	0	1	0	0	1	0	0
47	0	0	1	0	0	0	1	1
48	0	0	1	0	0	1	1	0
49	0	0	0	1	0	1	1	0
50	0	0	0	0	0	0	1	0
ID1	0	0	0	0	0	0	0	1
ID2	0	0	0	0	0	0	1	0
ID6	0	0	0	0	0	0	1	0
ID7	0	0	0	0	0	0	1	0
ID8	0	0	0	0	0	0	1	0
ID9	0	0	0	0	0	0	1	1
HI1	0	0	0	1	0	0	1	1
HI2	0	0	0	1	0	0	1	0
HI4	0	0	0	0	0	0	1	0
HI7	0	0	0	0	1	0	1	0
HI8	0	0	0	0	1	1	1	0
HI9	0	0	0	0	0	0	1	0
H2	0	0	0	1	1	0	1	0
H4	0	0	0	0	1	0	1	1
H6	0	0	0	0	1	1	1	0
H7	0	0	0	0	0	0	1	0
H8	0	0	0	0	0	0	1	0
H9	0	0	0	0	0	0	1	0
I1	0	0	0	1	0	1	1	0
I5	0	0	0	0	0	0	1	0
I7	0	0	0	0	0	1	1	0







