ASSESSMENT OF GENETIC DIVERSITY OF MAIZE LANDRACES IN TANZANIA USING RANDOM AMPLIFIED POLYMORHIC DNA MARKERS

THEODORE WORLANYO ASIGBEE

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN MOLECULAR BIOLOGY AND BIOTECHNOLOGY OF THE SOKOINE UNIVERSITY OF AGRICULTURE. MOROGORO, TANZANIA.

2016

ABSTRACT

The knowledge and comprehension of the genetic variation of maize (*Zea mays* L.) landraces is pivotal for the implementation of measures to address conservation and improvement. The purpose of this study was to assess the genetic diversity and relationship among selected maize genotypes in Republic of Tanzania by screening twenty Random Amplified Polymorphic DNA molecular markers.DNA was extracted from 160 maize genotypes and PCR was conducted using twelve informative primers. Amplification revealed 104 polymorphic bands with an average of 8.67 polymorphic fragments per primer. The number of amplified fragments ranged from 7 (OPP-02) to 10 (OPK-08), with the amplicon sizes ranging from 75 to 2000 base pairs. The polymorphic information content (PIC) ranged from 0.7487 to 0.954 with an average of 0.8647 and gene diversity value ranged from 0.7531 to 0.9577 with an average of 0.8698.The dendrogram drawn based on Neighbour- Joining method revealed the diversity and genetic relatedness among the landraces in the various clusters but did not reflect the geographical locations of the studied maize genotypes. This might be attributed to the high gene flow in the various study locations.The analysis of the RAPD molecular markers revealed a high genetic diversity among the maize landraces and proved to be a practical method for assessing polymorphism in maize cultivars. These findings will be useful to establish and improve the current germplasm collection of landraces and help maximize the utility of maize genetic resources.

DECLARATION

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

_____ _____

Theodore Worlanyo Asigbee Date

(Candidate: MSc. Molecular Biology and Biotechnology)

The declaration is hereby confirmed by:

_____ _____

Professor Joseph Ndunguru Date

(Supervisor)

_____ _____

Professor Paul Gwakisa Date

(Supervisor)

COPYRIGHT

ACKNOWLEDGEMENTS

Though only my name appears on the cover of this dissertation, a great deal of people have contributed in making the production of this thesis a reality.

I place on record my profoundest gratitude to Prof. Joseph Ndunguru and Prof.Paul Gwakisa of Mikocheni Agriculture Research Institute and Sokoine University of Agriculture respectively who took time off their busy schedules to help supervise and mentor me to finish my thesis.My experiences with the aforementioned men have been nothing short of amazing. Their words of advice and mentorship have enabled me to see the world from a different perspective.

Secondly, I am highly indebted to the Project Staff and members of Intra-ACP Academic Mobility Scheme for giving me the opportunity of a lifetime to pursue my postgraduate studies in Tanzania.

Furthermore, I thank MrutuBakari, Charles Kayuki,Dr. Mariam Mtungujaand all sundry of MARI who supported and lent a helping hand in making this a success. Without their help, production of the thesis would have been a challenge.

Finally, I thank my entire family, most importantly my parents who have been my backbone of support throughout my entire life. Although they hardly understand what I have researched on, my parents were willing to support any decision I made. For that, I am forever thankful.

DEDICATION

I dedicate this work to all strong and beautiful mothers of this great continent of Africa who fervently motivate and support their offspring regardless of their hardship and struggles.

May the good Lord uplift and strengthen their efforts. It is through their struggles,offspringshave attained great height and become a beacon of hope for younger generations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND SYMBOLS

µg	microgram(s)
µl	microliter(s)
AFLP	Amplified Fragment length Polymorphism
bp	base pairs
CGIAR	Consultative Group for International Agriculture Research
COSTECH	Commission for Science and Technology
DNA	deoxyribonucleic acid
dNTPs	dideoxynucleotide triphosphates
e.g.	example
FAO	Food and Agriculture Organization
Ft	feet
g	gram
GIS	Geographical Information System
ha	hectare(s)
IITA	International Institute of Tropical Agriculture
Kb	kilobase pairs
Kg	kilogram
M	molecular ladder
MAFAP	Monitoring African Food and Agricultural Policies
MAFSC	Ministry of Agriculture Food Security Cooperatives
MARI	Mikocheni Agriculture Research Institute
$MgCl_2$	magnesium chloride
min	minute(s)
ml	milliliter(s)
mm	millimeter(s)

mM	millimole(s)
mg	milligram
NC	negative control
NaCl	sodium chloride
Ng	nanogram
^{o}C	degrees Celsius
OECD	Organization for Economic and Co-operative Development
PCA	Principal Component Analysis
PCoA	Principal Co-ordinate Analysis
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information Content
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
rpm	revolution per minute(s)
S/No.	Sample Number
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
t	tonnes
U	unit
UK	United Kingdom
UPGMA	Unweighed Paired Group Method with Arithmetic Mean
USA	United States of America
UV	Ultraviolet
v	volume
w	weight

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background Information

Maize (*Zea mays* L.) is the third most important grain after wheat (*Triticumaestivum*L.) and rice (*Oryza sativa* L.) with high potential for production and productivity(Langade *et al*., 2013).Maize was first introduced in Tanzania in the Arusha region but later became popular in Mbeya and Rukwa in the 1950s and 1970s respectively(Coulson and Diyamett, 2012).Tanzanians depend on maize mainly for its consumption as part of a component of the nation's caloric intakeand for income generating purposes through its production(Nkonya *et al*., 1998). Maize is the most widely cultivated stable crop in Tanzania with the Southern highlands accounting for its highest production(Lyimo *et al.*, 2005).

Sub-Saharan Africa has a higher occurrence of pest, viruses, fungi, and predators than any part of the globe and this could be attributed to climate change and weather vagaries, trade liberization and agricultural intensifications(Nyambo,2009).Maize has been a controversial and risky crop because of its dependence on rain fed regime at specific times in the season. Feed the Future (2013), stated that in 2011,the Tanzanian government banned the transport of maize across borderspreventing the purchase of the crop because it was focal to their food security concerns. The export ban followed a period of drought, bad harvest and price peaks. The embargo on productionprevented maize from been diverted to Kenya and other environs where prices are significantly higher than in Tanzania. According to theUnited Nations (2004), it is projected that by the year 2050 the population of Africa will be about 1.8billion leading to a greater demand for food supply. Hence,there is the need to promote measures to increase productivity to combat the global

challenge of food by harnessing the wealth of genetic variations in the germplasm to be utilized for crop improvement.

One pivotal way to improve productivity is through genetic diversity assessment which provides the baseline for cultivar enhancement and development(Aremu, 2011a).Genetic diversity among maize germplasm can be achieved via physiological, morphological, and molecular differences existing among the population(Govindaraj *et al.,* 2015).Molecular markers are more efficient because marker interactions are not influenced by the environment and the plant can be assessed at any stage of development(Garcia *et al.*, 2004).The use of PCR-based markers for the genetic analysis and manipulation of plant traits has become a useful tool for plant breeding because it helps screen for polymorphism, productivity and yield stability early in the program(Konstantinov and Mladenovic, 2007).Random Amplified Polymorphic DNA (RAPD)is the first of PCR-based markers and requires little genetic material for the creating of multiple copies of DNA strands for analysis. RAPD are useful for assessing genetic diversity(Williams *et al.,* 1990) because its relatively cheaper, simple to use and faster than other types of markers(Abuali *et al*., 2014). Screening of RAPD markers provide a baseline to select markers that respond better during the pre-selection process. The variation found in the level of polymorphism by the pre-selected markers could be attributed to differences in the maize genotypes being used (Sun *et al*., 2001).

The current average yield per hectare is between 1.2 and 2.0 tons (FAOSTAT, 2012) as compared with expected potential yield of more than 8 t/ha and this change can be attributed to stress factors (FAOSTAT, 2012). Stress factors continue to cripple production and income of farmers, hence pose a threat to food security to the people of Tanzania and its environs.In order to improve the genetic resources available in Tanzania, there is a need

to know the currentgenetic pool of resources. However, there is limited information on the genetic diversity of maize grown in Tanzania.The present study seeks to use Random Amplified Polymorphic DNA, a molecular marker to assess the genetic diversity among maize genotypes in Tanzania.

1.2 Problem Statement and Justification

1.2.1 Problem statement

Maize (*Zea mays* L.) provides 61% of dietary calories and more than 50% of utilizable protein to the Tanzanian populace(Nkonya *et al*., 1998; Saidia *et al.*, 2010). Tanzania is endowed with more than 3.3 million hectares (ha) of land with suitable climate for production of specialty maize commanding high prices on the international market. In Tanzania, small scale farmers tend to recycle seeds from their best plants with preferred traits for many agricultural seasons (MAFAP, 2013). These local varieties grown have adapted to local conditions and farmers practices and represent a myriad of genetic resource and diversity offering new alleles for stress tolerance in maize (Magorokosho, 2006). Efforts by the Tanzanian Commission of Science and Technology (COSTECH) towards maize improvement has been carried out producing new varieties in Tanzania for farmers (Anonymous, 2014). To further improve the genetic diversity of local germplasm, it is crucial to know the extent of already existing variability.

1.2.2 Justification

The devastating impact of stress factors and the impact of global warming is an alarming concern on food security and economic development in sub-Sahara Africa; effective solutions to this are of utmost importance since maize is a major stable food crop in this country. Genetic diversity studies of maize in Tanzania will serve as antecedent for further research to be conducted, enabling the improvement of maize research.Research in Tanzania regarding maize relies heavily on conventional breeding methods that take many

years before suitable maize varieties are released. In order to improve on the maize germplasm, knowledge of genetic variability needs to be ascertained as a precursor for other developmental breeding programs that will boost crop yield and productivity to meet the ever-increasing population.

Furthermore, information obtained through the knowledge of genetic variability of maize germplasmin Tanzania will help in tracking down variety adoption through the process of DNA barcoding.This will help to accurately identify maize cultivars without using morphological maize descriptors and interviews from farmers, thus enhances the ability to identify mislabelled maize varieties on the market and redundancies in the germplasm.

1.3 Objectives

1.3.1 Overall objectives

The main objective of this study was to understand the genetic diversity and relationship among selected maize genotypes grown in Tanzania.

1.3.2 Specific objectives

i. To screen the RAPD markers against the maize genotypes in Tanzania
ii. To determine the level of polymorphism and diversity of maize revealed by the RAPD molecular markers.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 General Description

Maize is an annual, monoeciousplant characterized by a cane stalk and possessing tassels and ear shoot having staminateand pistillate flowers respectively. The plant usually grows between 3-13ft with broad leaves at every node, overlapping around the stalk. Maize is covered in rows of kernel encased in a husk and protected by maize silk. Maize is pollinated by wind either by self or cross with viable pollen lasting for 10-30 days or longer depending on favourable environmental conditions(OECD, 2008). Maize has the ability to adapt to diverse geographical and environmental conditions making it unparalleled to any other crop(Abuali *et al*., 2014).

2.2Taxonomy and Centre of Domestication (Origin) of maize

2.2.1Taxonomy of maize

Zea is a genus of the family Poaceae commonly referred to as the grass family. Crops such as sorghum, wheat, rice and barley are found in this family.The genus *Zea*consists of four species of which *Zea* mays is the only cultivated species and of economic importance. The other *Zea* species are termed teosintes and are mostly wild grasses. Species of *Zea*have a chromosome number of 20 except for *Zeaperennis* which has a total of40 (Doebley, 2004).Maize is classified as:

Table 1: Taxonomy and classification of maize

Classification	Taxonomy
Kingdom	Plantae
Division	Magnoliophyta
Class	Liliopsida
Order	Poales
Family	Poaceae
Subfamily	*Panicoideae*
Genus	*Zea*L.
Species	*Zeamays*L.

Source**:** Verheye (2010)**.**

2.2.2 Origin and centre of domestication

The centre of origin of maize is the Mesoamerican region, now called Mexico. Domestication of maize began about 6000 years ago and it started spreading around the world after the discovery of America(Anonymous, 2008).

According to Miracle (1965), maize was first introduced to Africa by the Portuguese in 1500s and later to Zanzibar and the mouth of River Ruvuma in the seventeenth century. Maize penetrated Africa mostly through the coastal regions and it was a major food for the slave ship during that phase(Miracle, 1965).

2.3 Types of Maize

Maize can be classified based on either grain colour or kernel and endosperm composition (FAO, 1992).Based on colour, maize can be grouped as white, yellow or red(plata).Some have a mixture of various shades of the named colours. In Sub-Saharan Africa the white maize is more predominant than the others(Kassie *et al.*, 2014).

On the basics of kernel and endosperm composition, Paliwal (2000) and Darrah *et al.*(2003)stated that maize can be grouped as:

1. Dent maize: the sides and base of kernel is presented with hard endosperm with high proportion of starch content with lightweight grain texture. The name dent is derived from the depression that occurs when the topmost part of the kernel contracts due to drying. This type of maize is mostly preferred for milling and livestock feed.
2. Floury maize: it is the oldest type of maize characterized by soft starchy endosperm that makes the grain easier to grind.
3. Sweet maize: The grains have a transparent appearance with a high constituent of sugar in the makeup. This is usually due to one or more recessive mutations that block the conversion of sugar to starch.
4. Pop corn: kernels possess the highest level of hard endosperm compared with other maize kernels. The kernel contains small amount of starch inside the pericarp. The application of heat results in the rupturing of the epidermis. Low levels of moisture in the kernel result in poor popping. Popcorn, a snack derived from pop maize is patronized and consumed globally.
5. Flint maize: they possess hard,vitreous endosperm with less soft starch. Flint maize has the ability to withstand freezing because of its low water content.

2.4 Nutritional Content and Allergens

One pivotal reason why the world regards maize as a staple crop is its par excellence in nutritional content.The composition of edible dry maize as described by Gopalan *et al*. (2007) is stated in Table 2 below:

Table 2: **Composition per 100g of edible portion of dry maize**

Nutrient	Composition	Nutrient	Composition
Protein	11.1g	Fibre	2.7g
Carbohydrates	66.2g	Magnesium	139mg
Minerals	1.5g	Sodium	15.9mg
Fat	3.6g	Sulphur	114mg
Iron	2.5mg	Amino acids	1.78mg
Riboflavin	0.10mg	Vitamin C	0.12mg
Copper	0.14mg	Calories	342
Potassium	286mg	Calcium	10mg
Phosphorus	348mg	Moisture	14.9g

Source: Gopalan *et al*.(2007)

However, major factors such as plant age, environmental conditions, geographical locations and genetic composition play and impact on the composition of the kernel between and within maize varieties (Nuss and Tanumihardjo, 2010).This implies that kernel composition analysis should be considered as an approximated value rather than a constant value.

Currently, maize is devoid of toxins but there are some allergic reactions that are caused by maize. Baker's asthma and hay fever are examples of allergic reactions caused by pollen and maize dust. Four allergens namely Zea m1(Broadwater *et al*., 1993), Zea m12 (Staiger *et al*., 1993), Zea m14 (Pastorello *et al*., 2002) and Zea m25(Weichel *et al*., 2006) are the main allergens of maize.

2.5 Maize Production and Utilization in Tanzania

2.5.1 Maize production

The total maize area under cultivation in Tanzania as stated by various authors has ranged between 1.7 and 2.0 million hectares (Nkonya *et al*., 1998; Pingali, 2001).According to the

FAOSTAT(2014),Tanzania has produced 5.1million to 6.7million tonnesfrom the 2012-2014 year. Only about 5% of the maize is produced by large scale farms,10% produced by medium scale farms whilst the remaining 85% is produced by small-scale subsistent farmers dependent on a rain-fed regime (Moshi *et al.,* 1998; Lyimo *et al*., 2014).Maize is produced almost throughout the entire country during the Masika main season and the Vuli second season. During the Masika season (March –June),41% of arable land is used whilst 47% of the cultivated land is used during the Vuli season occurring between October to November (MAFAP, 2013). Regions such as Morogoro, Mbeya,Coast, Kagera, Kigoma, Mwanza, Tanga, Arusha, Kilimanjaro and Mara have two major agricultural seasons. The remaining maize production is accounted for by the unimodal and bimodal Masika season thereby enhancing domestic production of maize all year round.Regions such as Dar es salaam, Lindi,and Singida are maize-deficient regions hence depend on other regions during the bounty harvest(Nkonya *et al*., 1998;MAFAP, 2013).

There are geographical variations in consumption of calories in the country. Maize accounts for 51% of total calories in the Southern highlands, 32% of the total calories is accounted for in the Lake Zone and Dar es Salaam accounting for 23% of calories(Cochrane and Souza, 2015).The Lake Zone has a lower caloric percentage because cassava is the other stable consumed whilst in Dar es Salaam diets are more diverse because of the higher average income and the greater access to markets.

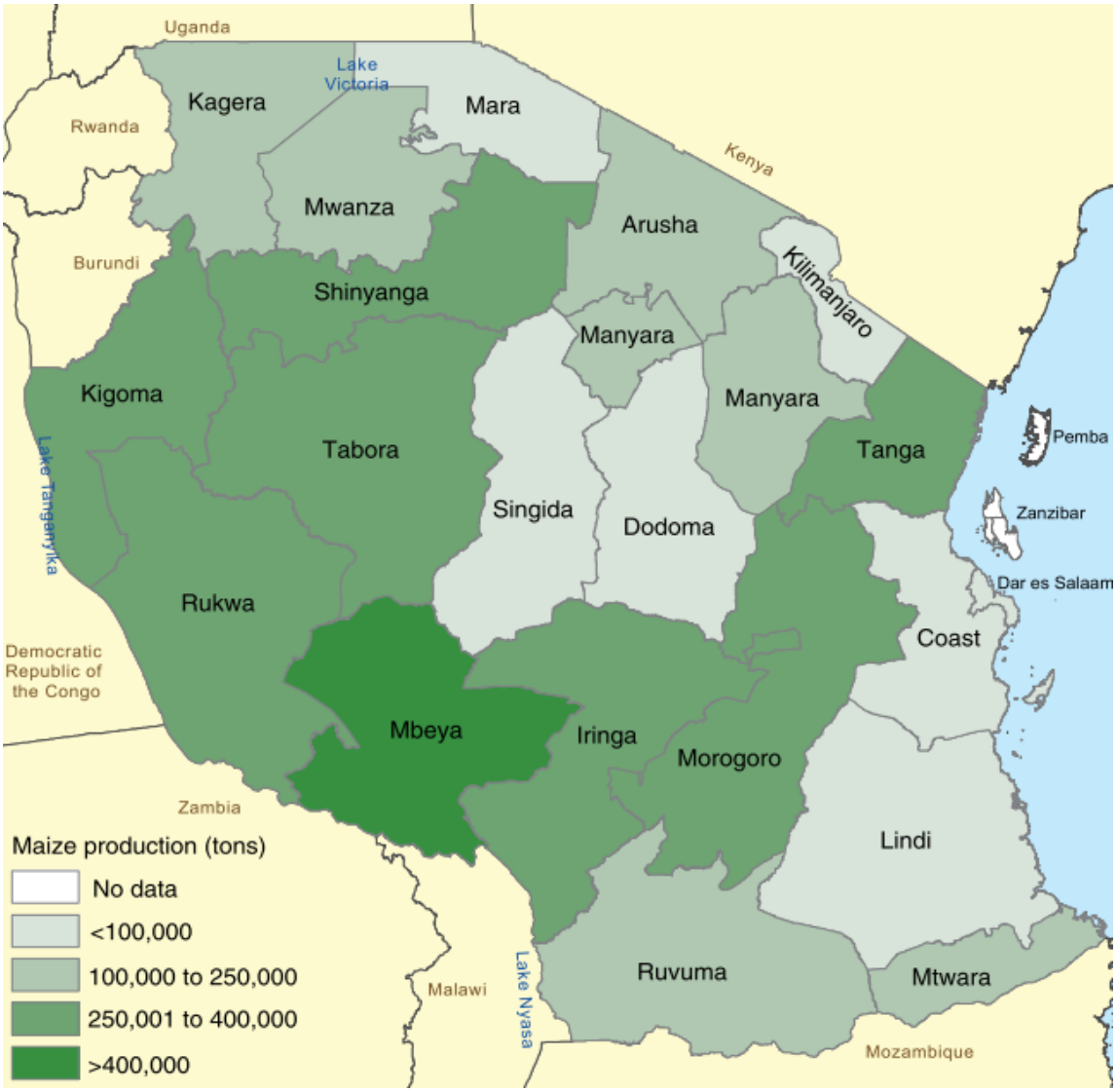


Figure 1:Maize production zones in Tanzania

Source: Cochrane and Souza (2015).

Note: Mbeya region in the Southern highlands is the leading producer of maize in Tanzania producing over 400 000 tons annually. Agro-ecological regions such as Rukwa, Iringa, Morogoro, Tabora, Kigoma, Tanga, and Shinyanga produce between 250 001 to 400 000 tonnes annually whilst the other regions vary in production of up to 250 000 tons.

2.5.2 Maize Utilization

Maize is mostly consumed by both the rural and urban dweller in United Republic of Tanzania.The maize produced in Tanzania is mostly consumed by the people with only about 5% been used for animal feed (Pingali, 2001).The inhabitants diet constitute about 61% of maize dependent meals making the per capita utilization of maize to be estimated over 114kg per year (Kaliba *et al*., 1998; Nkonya *et al*., 1998).

The maize seeds are usually milled into flour to prepare porridge and stiff porridge locally called Ugali made by mixing and continuously stirring ground maize flour with boiling water. Most school feeding programs across the country consume this meal, hence increasing the consumption of maize. The Tanzania schoolfeeding program as of 2013 has reached over 1064 primary schools across the country(MAFAP, 2013). Fresh maize on its cob is either boiled or roasted and eaten as a delicacy. Fermented maize is processed for the production of locally made beer. Popcorn is also obtained from maize, which is mostly served as a snack.The stover and cobs from the maize is a ration for ruminants and other animals whilst the dried cobs are usually used as firewood in small scale farms (Abate *et al.*, 1993). In developed countries, there is the production of ethanol from maize, which is further processed into biofuel for use as the next generation replacement for oil. The United States of America produces about 2.81billion gallons of ethanol yearly (Pimentel and Patzek, 2005). Large hectares of land are required to cultivate corn in order to generate biofuel for use.

2.6 Development of Improved Maize Cultivars

Maize variety development in Tanzania was initiated by the public sector and taken over by the private sector in 1993(Lyimo *et al*., 2014).The private sector released its first variety in 1994 and it has dominated since, even after the re-emergence of the public

sector in 2008 (Lyimo *et al*., 2014).By the year 2010, 91 varieties have been released of which 79% are hybrids. Two-thirds of the improved varieties were released by the private sector mostly by multinationals and regional companies.

In 2014, Tanzanian Commission for Science and Technology developed three drought tolerant maize varieties for small-scale farmers with the collaborative efforts of the private and public sector. The new varieties are WE2109,WE2112 and WE2113 and are developed with the understanding to combat climate change challenges(Anonymous, 2014). According toLyimo *et al*. (2014) most farmers do not use the improved varieties due to a myriad of reasons including the high price of the varieties, susceptibility to pest and diseases,and lack of information about the improved varieties. The farmers who are well informed about improved maize cultivars gain higher productivity than the uninformed farmers, thus impact of improved varieties on reduced food security concerns was well pronounced(Kassie *et al*., 2014).In West Africa, there have been new varieties that have been developed to combat constraints such as drought, low soil fertility, pest diseases and parasitic weeds through conventional plant breeding. Two Striga resistant and 4 hybrids with drought-tolerant varieties have been developed focusing on natural available plant traits in Nigeria (IITA, 2010). Every year International Institute of Tropical Agriculture (IITA), distributes improves open-pollinated varieties and hybrids to national partners and private sector within and outside of the region through regional trials. In 2010, IITA a member of CGIAR consortium released two vitamin A enriched maize containing high levels of beta-carotene (the vitamin A precursor) that helps against vitamin A deficiency (IITA, 2012).

2.7 Morphological Character for Diversity Analysis

Maize is mostly classified by taxonomist based on their race in which they belong. Anderson and Cutler (1942) defined race from the genetically point of view as group of plants possessing dominant alleles in common, with sub-races possessing larger number of

alleles in common than major races. Maize has been classified based on varied number of parameters to explain genetic variation in landraces. According to studies conducted byGalarreta and Alvarez (2001),leaf area, ear shape, tassel branches, kernel rows, plant height, cob weight and ear length are the most important criteria for taxonomical classification.Contrary to Galarreta and Alvarez findings, Sanchez *et al*. (1993) stated that taxonomic classification based on morphology is based on number of leaves per plant, number of branches, part length/tassel length, spike internode length, kernel width, pith diameter and ear/ kernel diameter. These findings are inconclusive and must be noted that characters must be less subjective to environmental changes and biases.Parameters that can be measured quickly on the farms and fields are mostly of great importance with regards to morphological parameters selected (Magorokosho, 2006).

2.8 Genetic Diversity Studies

2.8.1 Genetic diversity

Genetic diversity among germplasm can be defined as the process of studying variations among genotypes and populations and later analyzed by a specific method or group of methods (Mohammadi and Prasanna, 2003). Information obtained from analysis of genetic diversity facilitates classification of germplasm, the right choice of parents, and improves plant variety protection. Breeders are critical with the diversity of elite germplasm because information obtained are used to develop both farmers desirable traits (yield and large seeds etc.) and the breeders preferred traits such as pest and drought resistance (Govindaraj *et al.*, 2015). Genetic diversity in crops can be analyzed at any stage of growth of the plant and can be analyzed using markers and software packages. With regards to software choices, users prefer user-friendly and statistical purchases that offer best evolutionary models with NTSYS-pcRohlf (2000) emerging as one of the best software been used to analyze diverse datasets (Mohammadi and Prasanna,

2003).Assessment of genetic diversity can be ascertained at various plant stages using techniques viz. morphological, biochemical and molecular marker analysis.

Morphological markers are phenotypic traits that are visually accessible such as plant height and seed colour. These markers do not require expensive technology to use but large hectares of land is required for experimentations and trials to visualize these markers.The major disadvantage of morphological markers are its ability to be influenced by phenotypic plasticity hence are not able to assess genetic diversity(Noel *et al.*, 2007).

Biochemical markersalso termed as isozyme markers are variation in enzyme that are detected by electrophoresis and specific staining(Magorokosho, 2006). Isozyme markers are co-dominantin nature and require small amount of plant material to detect variation at the gene level(Govindaraj *et al*., 2015). Isozymes possess that major disadvantage of being influenced by the environment and at the developmental stage of the plant (Tanksley and McCouch, 1997).

Molecular markers are the most predominant markers used because of their abundance.They are selectively neutral because of their location in non-coding regions of DNA in a chromosome. They arise from different types of DNA mutations such as substitution, rearrangement, or errors in replication of tandemly repeated DNA(Matsuoka *et al.*, 2002). DNA markers can be divided into three main category based on their method of detection; (i) hybridization based e.g. RFLP (ii) PCR based e.g. AFLP, RAPD, SSR and (iii) DNA sequence based e.g. SNPs (Tanksley and McCouch, 1997). All these markers vary with regards to their principle, application, amount of polymorphism detected, cost and time requirement (Tanksley and McCouch, 1997).

Restriction Fragment Length Polymorphism is a hybridization-based marker based on variation in the length of DNA fragments produced by digestion of genomic DNA with specific restriction endonucleases (Dear, 2001). Variation in the pattern of RFLP can be caused by base pair deletions, mutations, inversions, translocations and transpositions resulting in polymorphism and differences in fragment length (Abdel-Mawgood, 2012).Despite RFLP detecting unlimited number of loci because of its co-dominance it is expensive, time consuming, labour-intensive, requires large amount of DNA and produces limited polymorphisms in closely related landraces and lines(Collard *et al.*, 2005). This led to the evolution of PCR based techniques for assessing diversity.

Amplified Fragment Length Polymorphisms is a technique for detecting polymorphism by the combination of both RFLP and PCR by digesting amplicon fragments using specific restriction enzymes that cut DNA at or near specific recognition site in nucleotide sequence (Mohan *et al*., 1997). Fingerprints are produced without prior knowledge of genomic content using a limited number of genetic markers (Vos *et al*., 1995).Kimani *et al.* (2012)used AFLP to characterize lablab beans in Kenya revealing low diversity among the bean accessions. The major disadvantages of AFLP are high operational cost, laborious methodology; alleles are not easily recognized and it produces medium reproducibility(Vos *et al*., 1995; Abdel-Mawgood, 2012).

Simple Sequence Repeat (SSR) also called microsatellite occur frequently in most eukaryotes, are highly variableand evenly distributed throughout the genome (Garcia *et al.*, 2004). Microsatellites are located in non-coding regions hence polymorphisms are determined by manufacturing forward and reverse primers that are complementary to target DNA. Microsatellites are effective markers because of their co-dominance, locus specificity and multi-allelic character(Magorokosho, 2006). Microsatellites have been

used for assessing the genetic diversity of crops. Shehata *et al.* (2009) used SSR to understand the genetic variability among and within maize inbred lines.

Single Nucleotide Polymorphism (SNP) are the most widespread type of sequence variation in genomes and they occur in both coding and non –coding regions(Kwok *et al.*, 1996).SNPs are single base variation in a DNA sequence which results from mutation events(Wakeley *et al.*, 2001) and identified by using microarrays or denaturing high performance liquid chromatography (DHPLC) machines (Mammadov *et al.*, 2012). SNPs produce less levels of polymorphism as compared with SSR due to its biallelic nature (Foster *et al*., 2010)but is compensated for by the use of more markers and can identify haplotypes in the genome(Galeano *et al*., 2012).

Molecular markers have been used to ascertain the genetic diversity of maize. Thakur *et al.* (2016) used SSR markers to study the genetic diversity of maize accessions of North Western Himalayas. In Benin, a high genetic diversity was observed by Salami *et al.* (2016) using SSR markers to understand the variability of maize accessions. Molin*et al.*(2013) used a combination of RAPD, SSR and AFLP to understand the diversity of maize landraces in Brazil. It was observed that high variability was observed among the landraces.

2.8.2 RAPD

RAPD is a PCR-based technique since the 1990s for accessing intra-specific genetic variation at the nuclear level (Williams *et al.,* 1990).RAPD produces a quick and efficient screening to detect polymorphism without prior knowledge of specific target DNA sequences for amplification. RAPD markers are random or arbitrary markers consisting of a decamer of oligonucleotides that amplifies at a very large number of loci making it very

suitable for genetic mapping studies and DNA fingerprinting especially in population genetics (Govarthanan, 2011). It works on the principle of a single arbitrary primer in a PCR, resulting in the final synthesis of discrete DNA fragments (Hulse de SOuza *et al*., 2008). The RAPD amplicon is generated at the flanking region of the decamer priming site in the appropriate orientation and configuration (Abdel-Mawgood, 2012). Amplified products are usually visualized on agarose gels and stained with ethidium bromide (Abdel-Mawgood, 2012).

The main advantages of RAPD is its universality to be used across wide array of species, cost effectiveness, no prior knowledge of sequence of DNA, and requires less amount of DNA to obtain amplifications (Kumari and Thakur, 2014).

Although the RAPD marker is easy to use and cost effective the major concern associated with it has to do with reproducibility due to the sensitivity of the arbitrary primer amplifying regions of the genome (Kumari and Thakur, 2014;Govindaraj *et al*., 2015), but if care is taken reproducible results can be obtained when conditions are standardized (Lowe *et al.*, 1996).The collective pool of random markers gives the discriminatory power of the markers and enhances its effectiveness as compared with single primers (Kumari and Thakur, 2014).Information obtained from pooled markers provides more insight about genomes and genome expression. RAPD has been used to characterize a wide variety of species. In animals, Gwakisa *et al.*(1994), used RAPD to characterize indigenous African cattle in Tanzania. This study reported that primer (ILO 1127) amplified a RAPD fingerprint in 61% of the Tanzanian Zebu cattle animals but less than 6% in the other breeds. This study further revealed that RAPD could be used to reveal introgression. With regards to plant diversity studies, RAPD has been used to ascertain the genetic diversity among East African tall coconut. Masumbuko *et al.* (2014) reported that RAPD provided

the discriminatory power to determine the origin of the twelve coconut trees and help discriminate between the different provenances.

2.8.3 Grouping techniques in measuring genetic diversity

Genetic relationship between and within landraces can be categorized using multivariate grouping methods, which bases its identification on genetic distances among the landraces. The performance of genotypes across populations is assessed by using statistical algorithmsto categorize landraces into distinct groups, with members of a groups been closely related. The widely used techniques for grouping landraces are cluster analysis, Principal Component Analysis (PCA), Principal Co-ordinate Analysis (PCoA), Canonical Correlation and Multidimensional Scaling (Aremu, 2011a).Among the aforementioned techniques, cluster analysis and PCA are the two most commonly used techniques that are mostly used to explain variation and diversity among genotypes in genetic studies(Aremu, 2011b).

Cluster Analysis as described by Gordon (1999), states that clusters are based on the principle of internal cohesion-homogeneity and external isolation-separation. This definition has led to making the concept of homogeneity and separation mathematically concise. Cluster analysis present patterns of relationship between genotypes such that similar descriptors are mathematically assembled into similar groups. In generating patterns using cluster analysis, five methods are used namely Unweighed Paired Group Method using centroids, Single Linkage (SCLA), Complete Linkages(CLCA), and Median Linkage(MLCA). UPGMA provides more accurate grouping information on genotypes used than the others because it is consistent, effective and provide more accurate grouping information in grouping genotypes and also easier to group genotypes from different data sources and types(Aremu, 2011b). There are different types of clusters namely well

separated, prototype-based, graph based, density based, and shared property conceptual cluster(Everitt *et al*., 2011). These types of clusters are then usually represented as a tree or dendrogram to display all clusters(Mohammadi and Prasanna, 2003).In a study to assess the genetic diversity of maize inbred lines and hybrids in Southern highlands of Tanzania,Mrutu *et al*. (2014) performed cluster analysis on the hybrids and obtained two major groups.Afonso (2013)also used cluster analysis to evaluate the genetic diversity among maize accessions in Mozambique revealing no clear grouping of accessions belonging to the same agro-ecological zone.

Smith (2002) defined PCA as a way of identifying system patterns in data and expressing it in a new variable with fewer dimensions to highlight similarities and differences. Datasets can be compressed without losing informational data by the reduction in the number of dimensions producing new variables that are viewed as non-correlating groups(Aremu, 2011a).The geometrical distances among genotypes when analyzed by PCA reflects the genetic distances between them(Govindaraj *et al.*, 2015). One major shortcoming with regards to PCA is that it cannot analyze incomplete datasets hence it either discards the data or completes it using a variety of interpolation strategies(Chen, 2002).PCA is used in the application of data compression,image processing, exploratory data analysis, pattern recognition and time series prediction (Chen, 2002).Abuali *et al.* (2014) accessed the genetic diversity of maize inbred lines in Sudan using RAPD markers. It was observed that the PCA score obtained was 19.0 indicating the adaptability of the maize to the environment.

CHAPTER THREE

3.0 MATERIALS AND METHOD

3.1 Study Design and Study Population

The present study was carried out as a descriptive cross-sectional study, in which maize landraces were collected fromagro-ecological zones across Republic of Tanzania during the period of June to August 2015. One hundred and sixty genotypes were collected from various districts in Tanzania (Appendix1 and Figure 2).

3.2 Sample Collection and Procedure

Collected maize landraces were transported in well labelledbrown envelope of dimensions 229x324mm to the MikocheniAgriculture Research Institute situated in Mikocheni, Dar es Salaamwith latitude and longitude of -6.792354S and 39.208328E respectively.

3.3 Greenhouse Planting

The seeds of the maize genotypes were sown in plastic buckets and labelled with wooden tags. The labeled wooden tags were fixed into the buckets correspondingto the maize genotype planted.Watering was carried out twice during a period of 8 days. Further laboratory work was carried out at the 8 day after planting.

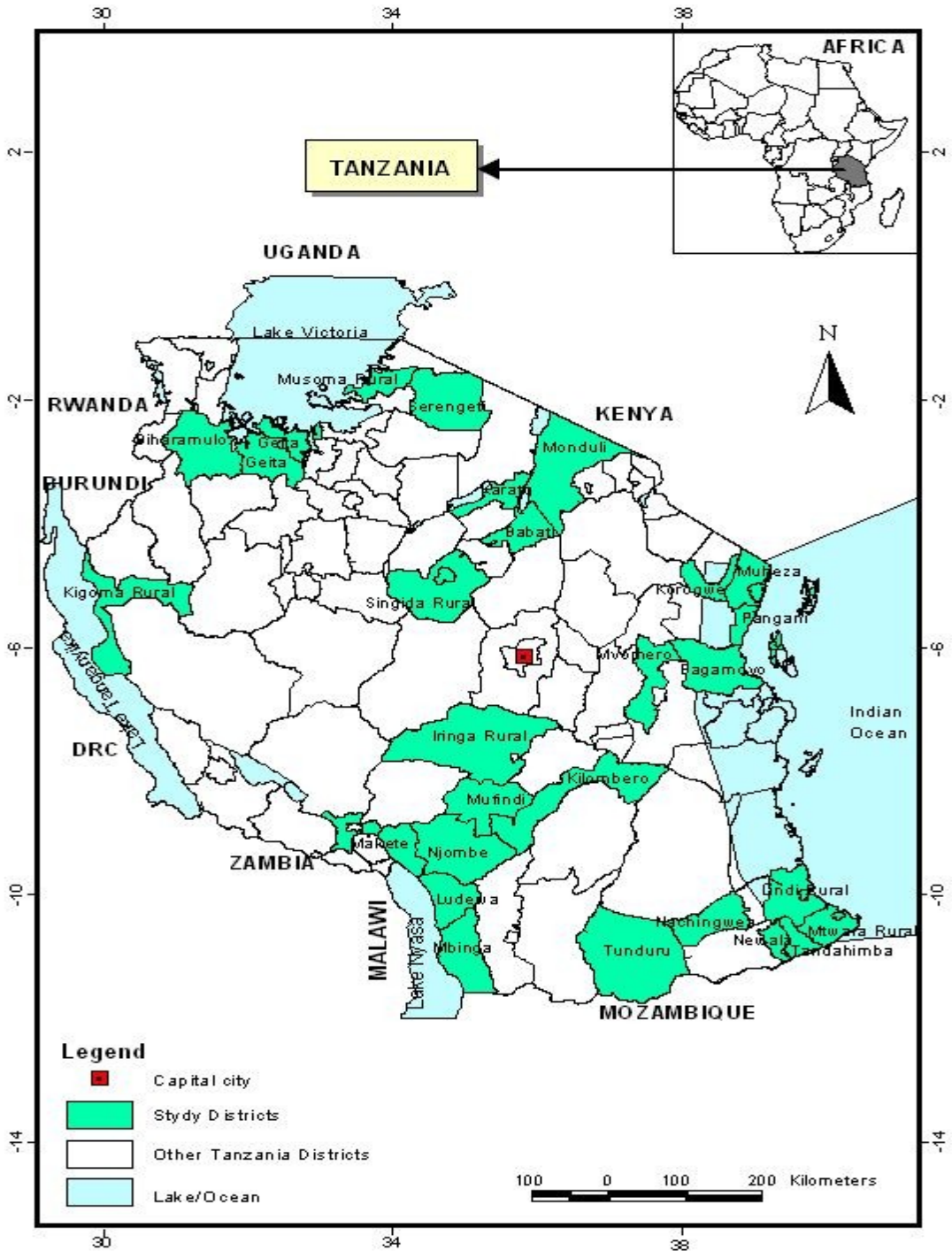


Figure 2: Map of Republic of Tanzania indicating the study districts.

Note: The map was constructed using Arc View GIS 3.2

Figure 3: The layout of maize genotypesin the MARI greenhouse. The maize samples shown were 8 day and 24 day old

3.4 Laboratory Analysis

3.4.1 Genomic DNA extraction

Fresh and tender maize leaf samples were excised and cleaned using 70% ethanol. Total genomic DNA was extracted using a protocol described byDellaporta *et al.* (1983),with some modifications (Appendix 2). Briefly, leaves were ground in 500µl of plant extraction buffer then transferred to 1.5mleppendorf tube and 40µl of 20% (w/v) sodium dodecyl sulphate (SDS) was added to the tube and mixed thoroughly by vortexing. Extracting buffer contained 100 mMTris (hydroxymethyl) aminomethane hydrochloride 88 (Tris-HCl), 50 mM ethylene-diaminetetraacetate (EDTA) and 500 mM sodium chloride (NaCl) at pH 8.0 plus one percent (w/v) polyvinylpyrrolidone (PVP) and 0.2 % (v/v) β-mecarptoethanol.The extracted DNA was stored at 20 ^{o}C until use.

3.4.2 Quantification of genomic DNA

The concentration and quality of DNA extracted was measured using Nano Drop 2000 UV Spectrophotometer(Thermo Scientific, Massachusetts, USA). For each sample, 1ul of

DNA was dispensed onto the Nano Drop sensor for quantification. The A260/A280 ratio was used to provide an estimate of DNA purity. A ratio between 1.8 and 1.9 was considered of acceptable quality. The obtained readings from the various samples were pooled together in aMicrosoft Excel 2016 spreadsheet.The quantified DNA samples were then normalized to a working concentration of 25ng/µl by using RNase-free water (Fisher Scientific, Geel, Belgium).

3.4.3 Primer screening and PCR

Conventional PCR was carried out in a total volume of 25 µl by mixing 10 mMTris-HCl (pH 8.6), 1xTaqpolymerase buffer, 50 mMKCl, 2 mM $MgCl_2$, 0.27mM of each dNTP, 0.8 mM primer, 1 U TaqDNA polymerase (Thermo Scientific),25 ng template DNA and the final volume made up with double distilled sterile water.Amplification was carried out in a Techne Prime Thermal cycler (Bibby Scientific, Staffordshire,UK).Twenty RAPD primers (Ascefran LLC, North Carolina, USA) were screened against the DNA samples to select the polymorphic primers.

Table 3: **List of twenty RAPD markers and their sequences**

S/No.	Name	Sequence 5'-3'	S/No.	Name	Sequence 5'-3'
1	0PP-02	GGCACGTAAG	11	OPO-05	AAGTCCGCTC
2	0PP-04	ACGTAGCGTC	12	OPP-05	CCCCGGTAAC
3	0PP-07	GTCCATGCCA	13	OPJ-07	CTCCATGGGG
4	OPK-05	CCGCCCAAAC	14	OPJ-05	CCACGGGAAG
5	OPK-04	CCAGCTTAGG	15	OPO-01	GTGCAACGTG
6	OPO-06	CCCAGTCACT	16	OPK-01	CATACCGTGG
7	OPK-10	GAACACTGCC	17	OPP-03	CTGATACGCC
8	OPK-08	AGCGAGCAAG	18	OPP-01	GTAAGCACTC
9	OPP-04	GTGTCTCAGC	19	0PP-03	CTGATACGCC
10	OPK-03	GTCTCCGCAA	20	OPJ-09	CATACCGTGG

Source: Mukharib *et al*. (2010)

3.4.4 Electrophoresis and visualization of PCR product

Eight µl of each PCR product was run on a 2 % agarose gel electrophoresis in 1.0x TAE buffer at 100 V for 60 min using Owl Separation System A6 (Thomas Scientific, Portsmouth, UK), along with a 1Kb plus DNA ladder (Thermo Scientific, Massachusetts, USA), after being mixed with 2 µl of 6x DNA gel loading dye (Thermo Scientific, Massachusetts, USA) and the gel stained with 2.5mg/ml ethidium bromide. The agarose gel was visualized under UV light using BioDoc-It Imaging System (VWR International, Cambridge, UK).

3.4.5 Data Management and Analysis

The bands obtained from the gelswere scored in Microsoft Excel 2016 spreadsheet.

The RAPD product was assumed to represent a single locus and data were scored as the presence (1) or absence (0) of a DNA band. Individual primers were tabulated separately and the data were pooled together to obtain a combined matrix for the genotypes. The polymorphic percentage of the obtained band was calculated by using the formulae:

Polymorphism % = (No. Of Polymorphic band / Total bands) X 100

The Powermarker V3.25 (Liu and Muse, 2005)was used to compute the binary matrix and evaluate the genetic associations between the genotypes. The major allele frequency, allele number, gene diversity and Polymorphic Information Content (PIC) were computed. A phylogenetic analysis based on the relatedness of the maize genotypes was drawn using Neighbour Joining method.

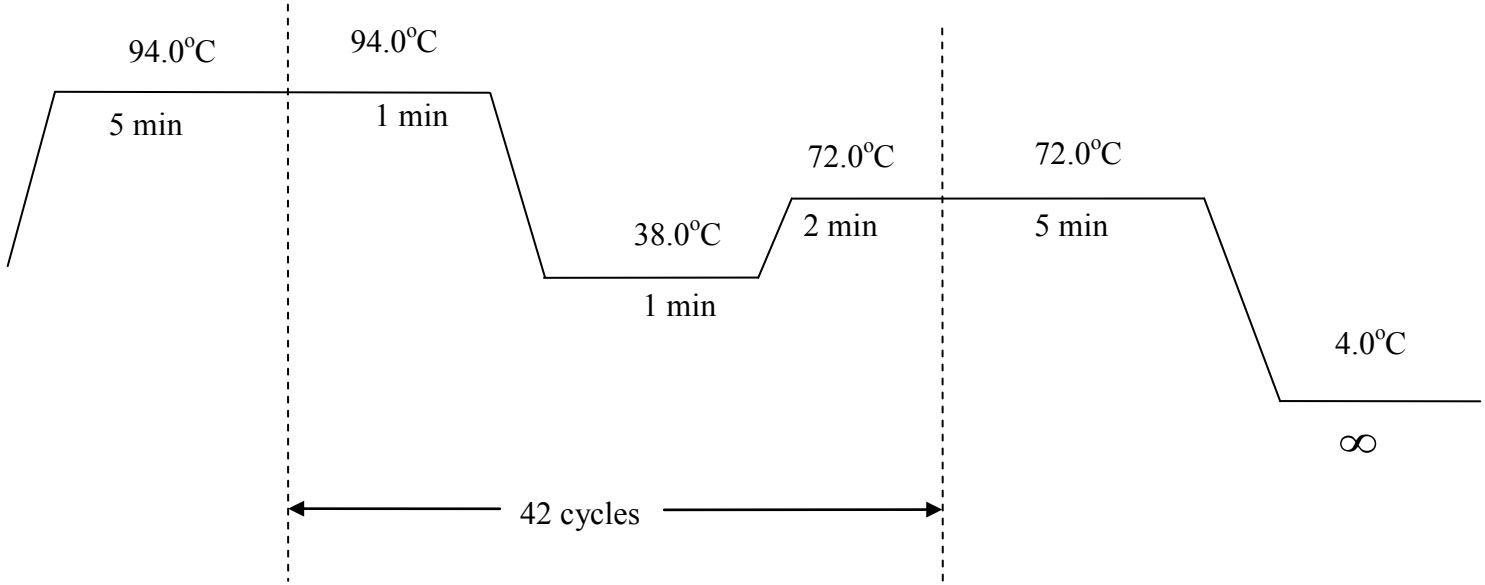


Figure 4: Conventional PCR cycling conditions used in the present study.

Note: The reactions consisted of 42 cycles, each comprising a denaturation at 94°C for 1min, annealing at 38°C for 1min and extension at 72 °C for 2 min. This was preceded by an initial denaturing step at 94°C for 5 min and followed by a final extension step of 5 min at 72 °C.

CHAPTER FOUR

4.0 RESULTS

4.1 Quality and Quantity of Maize Genomic DNA

In the present study, DNA was extracted from maize leaves using a modified protocol previously described by Dellaporta *et al*. (1983). The quality and quantity of DNA was ascertained using spectrophotometry. Good quality and quantity of extracted DNA was obtained(Appendix 3).DNA concentration between 200.7 and 4123 (ng/µl) were obtained (Appendix 3), and the quality of DNA ranged between 1.8 and 1.9 (Appendix 3).The quantified DNA was standardized to a working concentration of 25 ng/µl for PCRamplification.

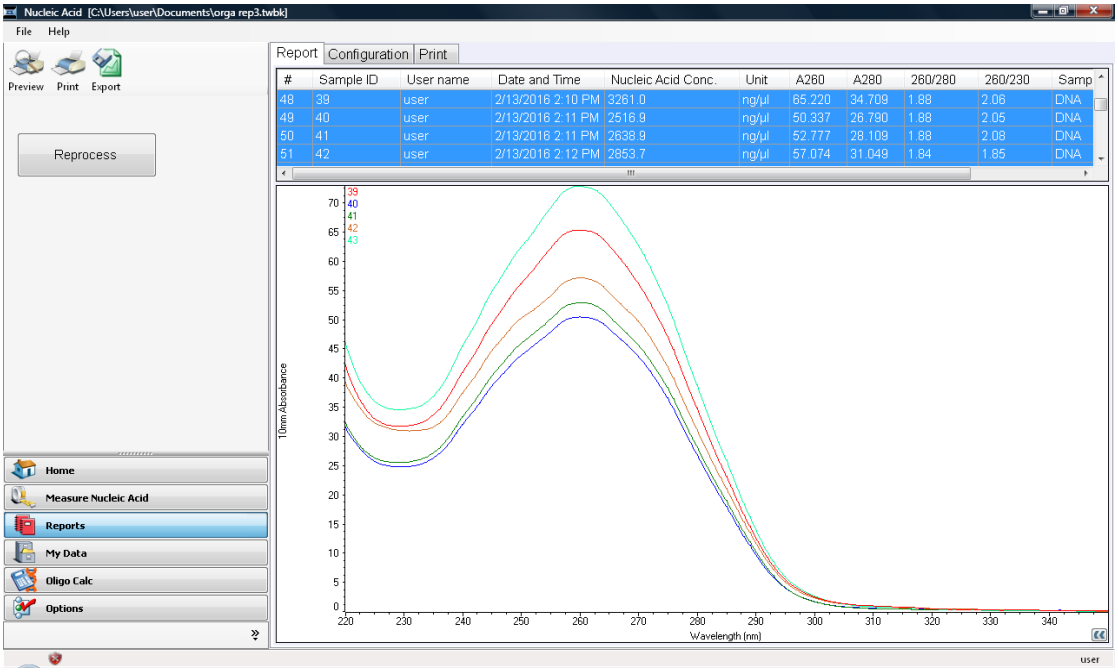
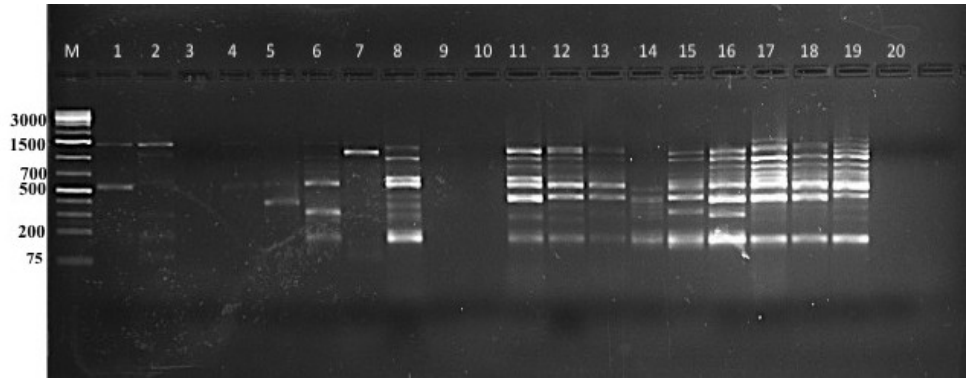
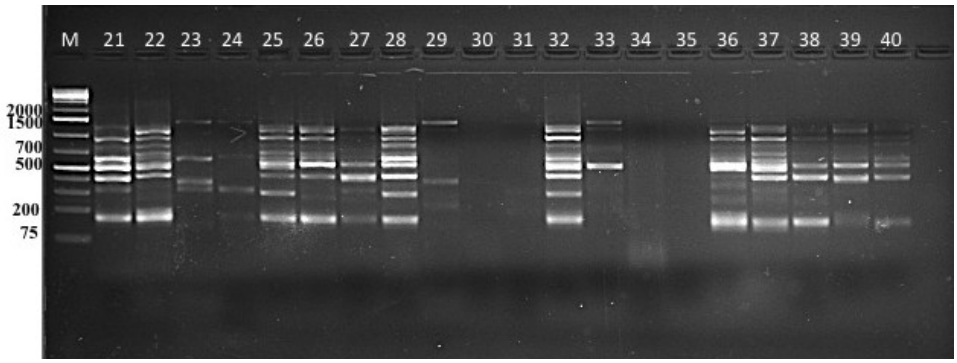


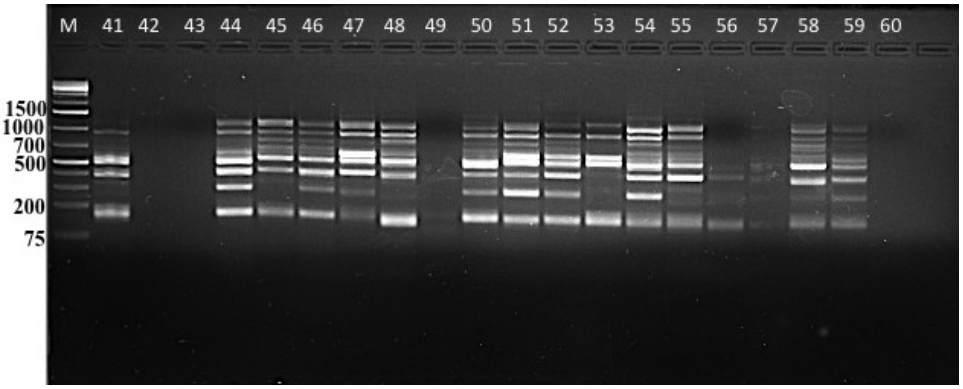
Figure 5:DNA concentration of selected maize samples.

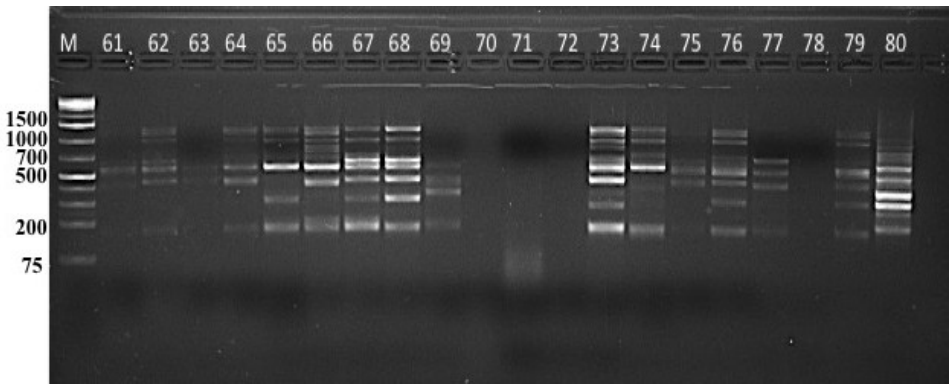
Note: Nano Drop reading of samples 39-42 showing the nucleic acid concentrationand absorbance ranging between 2516.9- 3261.0 ng/µl and 1.84-1.88 respectively.

4.2 RAPD Analysis









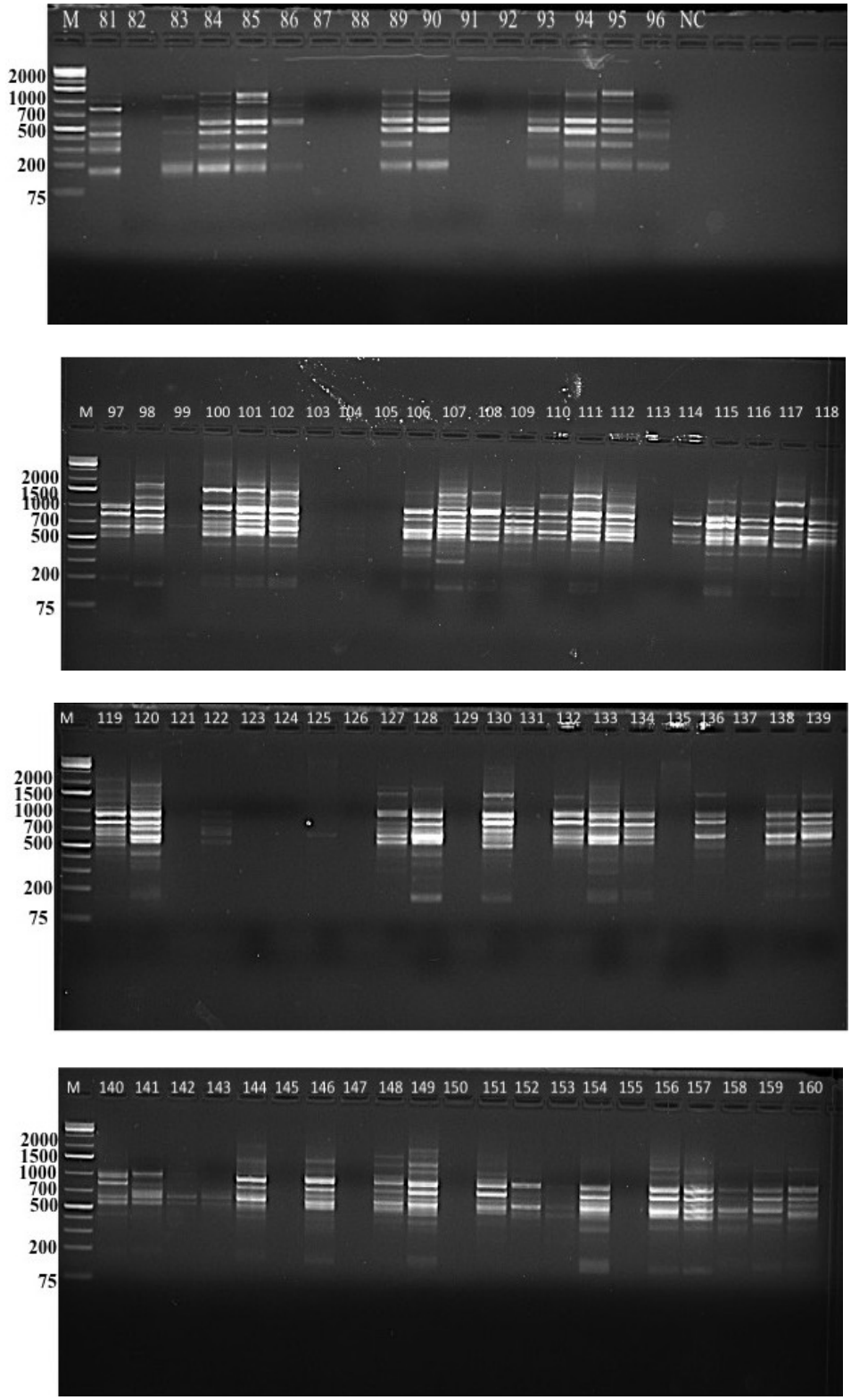


Figure 6: **Amplification profiles of 160 maize genotypes of primer OPO-06.**

Note: M= 1kb plus ladder, NC= Negative Control

RAPDs are useful to determine variability and polymorphism due to its universality across a wide range of species including maize requires less or no prior information about the sequence of DNA and less amount of DNA is needed for amplifications. The major limitation of RAPD has to do with its tendency to reproduce given results due to the sensitivity of the arbitrary primers. The RAPD approach can be improved for genetic diversity studies when conditions are standardized and the same equipment are reused.

Twelve previously selected RAPD primers were used, and only those showing high ability to produce polymorphic bands were selected for this study.There were variations of band intensity when visualized under UV producing both strong and weak bands. Figure 5 depicts an example of the electrophoretic pattern of RAPD fragments, amplified by primer OPO-06 resulting in 9 polymorphic bands.

In total, 12 primers amplified 104 bands in the maize genotypes. All the 104 bands showed polymorphism (100%) and on average 9 bands per primer were observed, with maximum of 10 (OPK-08) bands and minimum of 7 bands (OPP-02). Primers OPP-02, OPK-08, and OPP-07 produced maximum band size (2000bp) whilst primer OPO-05 produced the lowest band size (75bp) (Table 4).

Table 4: List of selected decamer oligonucleotide primers revealing RAPD analysis

Primer Number	Primer Name	GC Content (%)	Molecular weight range (bp) Larger	Smaller	Number of bands	Number of polymorphic loci	Percentage Polymorphism (%)
1	OPP-02	60	2000	150	7	7	100
2	OPP-04	60	1500	200	9	9	100
3	OPP-07	60	2000	75	9	9	100
4	OPK-05	70	1500	300	8	8	100
5	OPK-04	60	1500	200	9	9	100
6	OPO-06	60	1500	150	9	9	100
7	OPK-10	60	1500	200	9	9	100
8	OPK-08	60	2000	140	10	10	100
9	OPP-04	60	1500	200	9	9	100
10	OPO-05	60	1500	75	9	9	100
11	OPJ-05	70	1500	150	8	8	100
12	OPO-01	60	1250	100	8	8	100
Total					**104**	**104**	
Total Average					**8.67**	**8.67**	**100**

Table 5:**Details of polymorphisms and genetic analysis of 12 RAPD markers across 160 maize genotypes**

Primers	Major Allele Frequency	Allele Number	Gene Diversity	PIC
OPP-02	0.1563	39.0000	0.9207	0.9158
OPP-04	0.2875	67.0000	0.9009	0.8979
OPP-07	0.1500	67.0000	0.9577	0.9564
OPK-05	0.1938	48.0000	0.9397	0.9375
OPK-04	0.3750	58.0000	0.8493	0.8464
OPO-06	0.3188	38.0000	0.8373	0.8232
OPK-10	0.2375	44.0000	0.9009	0.8948
OPK-08	0.2063	52.0000	0.9206	0.9165
OPP-04	0.2750	49.0000	0.8951	0.8900
OPO-05	0.4625	48.0000	0.7739	0.7686
OPJ-05	0.4875	48.0000	0.7531	0.7487
OPO-01	0.4375	47.0000	0.7885	0.7806
Mean	0.2990	50.4167	0.8698	0.8647

The PIC valueof the RAPD markers, which is a measure of the allele diversity at a locus ranged from 0.7487 to 0.9564 with a mean value of 0.8647.All the primers exhibited PIC values higher than 0.740 indicating that they were potentially informative in detecting differences among the maize genotypes. The gene diversity values were in the range of 0.7531 to 0.9577 with a mean gene diversity value of 0.8698.The allele number ranged between 38.0 to 67.0 with an average allelenumber of 50.4167, with primers OPP-04 and OPP-07 revealing the highest allele number.The mean allele frequency value is 0.2990 with the maximum and minimum values being recorded by primersOPJ-05 (0.4875) and OPP-07 (0.1500) (Table 5).

The Powermarker software was used to construct a dendrogram using a Neighbor Joining method. The results obtained when the data was bootstrapped 1000 times to verify the

consistency and reliability of the dendrogram.This indicates the consistency and reliability of the results obtained from the study (Figure 7). It was observed that the genotypes were grouped into various major and minor clusters.The red dot in the dendrogram indicates where the branching of the tree begins. The biggest cluster possessed numerous sub-clusters detailing the relatedness of the genotypes. These sub clusters possessed bootstrap values ranging from 55 to 96%. The biggest cluster which also contains a sub-cluster of sample identity 1760, 26, and 2840 displayed a bootstrap value of 96% signifying that the three genotypes are genetically similar. Genotype 1760 (MN 50) locally called Segosera was sampled from Mbeya, genotype 26 (LNA 274) locally called Matundila was sampled from Lindi and genotype 2840 (MS170) locally called Kitwele was sampled from Tanga region. All the three genotypes were clustered within one sub-group indicating their genetic similarity even though they were sampled from different agro-ecological zones. There was variation in the bootstrap values of other clusters that ranged from 52 – 72%.

The top most minor cluster with a bootstrap value of 55% consisted of maize genotypes with sample T11, 1766, 3951, 2375, 4070, 3873, and 1730 with local names Mehhi, Katumbili, Kibanchano, Serena, Gembe, Malombe, and Kienyeji respectively.All these maize genotypes were obtained from various villages such as Kilimatembo (T11), Nmima (2375), Ibambila (4070), Nachingwe (3873), Ludewa (1730), Karatu (1766), Kienyeji (1730). With the exception of genotypes T11 and 1766, which were sampled from Karatu district in the Arusha region but from different villages, all the other genotypes in this cluster belonged to different districts and regions. Genotypes from different villages and district clustered together based on their genetic relatedness. Genotypes that were clustered together were said to be genetically identical and were genetically different from other genotypes that formed other clusters. The genetically identical genotypes were observed not to cluster based on their geographical locationsbut rather their genetic

relatedness. Within the major cluster with bootstrap values ranging from 55-96%, it was observed that genotypes formed sub clusters among themselves. The genotypes in this major cluster are genetically similar but not as similar as the sub-clusters within the major cluster. Genotypes 1751(MN 42) and 707 (MNC 237) are genetically more similar than genotype 687 (MNC 217) of bootstrap value 72% because they are grouped within a sub-cluster of the major cluster whereas genotype 687 with reference code MNC 217 is grouped within a different sub-cluster within the major cluster. Genotypes 687 and 707 were found to have come from the same region of Mara but different districts. These two genotypes are all locally called Nyamula indicating the clustering among the same major cluster.

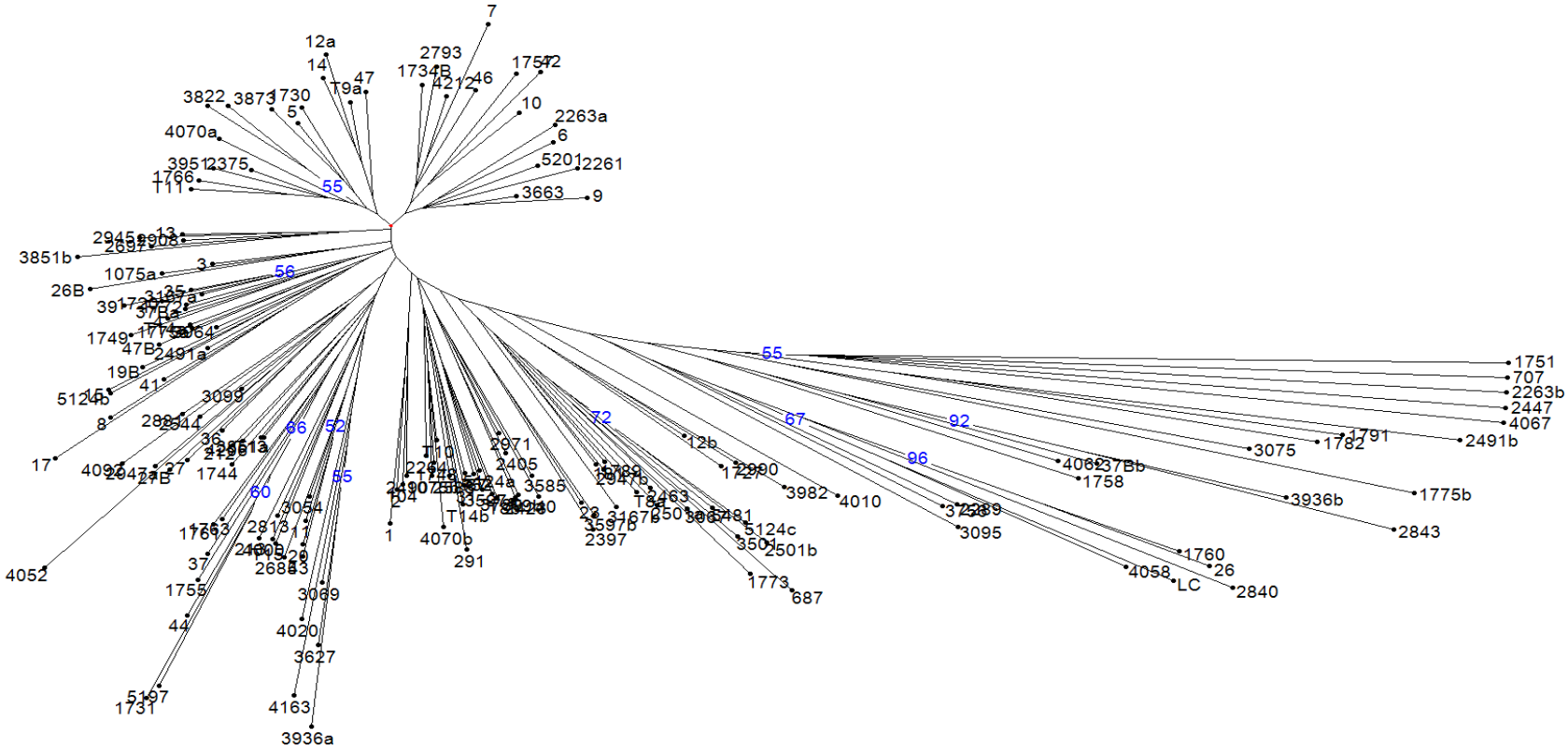


Figure 7: Phylogenetic analysis of the maize genotypes based on the Neighbour Joining method

Note: The numbers represent the Sample ID for the various maize genotypes used for the study (Appendix 1, gives more detailed information about the genotypes)

CHAPTER FIVE

5.0 DISCUSSION

Random Amplified Polymorphic DNA has exhibited to be an efficient and effective tool in studying genetic diversity of maize genotypes and other species. The level of polymorphism (100%) showed RAPD ability to detect differences among the genotypes. This diversity was the same as reported byBalazova *et al*. (2015) and Walunjkar *et al*. (2015) that assessed the genetic diversity of maize and pigeon peas genotypes respectively using RAPD.Other authors reported ability of RAPD to detect lower levels of diversity than those obtained in this study. Maize genetic diversity were reported at 72.2%, 89.33% and 97.03% byCarvalho *et al*. (2004),Abuali *et al.* (2011) and Handi *et al.* (2013) respectively. The level of polymorphism in this study can be attributed to the vigorous pre-selection of the primers contributing to the increase in the level of polymorphism. Additionally, variations found in the level of polymorphism could be the result of the distinct locations of the maize genome that were assessed by the pre-selected primers and/or the difference in the genotype among the maize used (Sun *et al.*, 2001).Mrutu *et al*. (2014) assessed maize inbred lines in Tanzania using the same RAPD markers and observed differences in number of alleles obtained and the level of polymorphism. This present study obtained 10 alleles and 100% polymorphism from primer OPK-08 whilst Mrutu's study obtained 6 alleles and 67% polymorphism.There are variations between the various primers used and this can be attributed to diversity among the maize varieties. Bruel *et al*. (2007) highlighted that divergence among genotypes under study can be a contributing factor to the number of alleles obtained and thus affecting the level of polymorphism. The high level of polymorphism revealed through the RAPD technique may be caused by a number of events, such as a deletion eliminating the primer binding site, an insertion making a fragment too large for polymerization, nucleotide substitutions

in the primer annealing site(s) leading to failure of polymerization, or a small addition or deletion leading to larger or smaller fragments respectively.

The RAPD primers pre- selected for this studyexhibited a high primer efficiency leading to its amplification. The range of primer efficiency showed the ability of the RAPD primers to give large proportion of polymorphic bands based on the total number of amplified bands, so the efficient primer is not the primer that gives highest number of amplified bands but the ability to differentiate and show the relationships between the maize landraces(Al-badeiry *et al*., 2013). There are variations in primer efficiency based on the DNA template of the species under study(Newton and Graham, 1997).

The Polymorphic Information Content gives detailed insight about the informativeness of the RAPD markers used for genetic diversity studies for all loci. The present study observed a mean PIC value of 0.8647, which was of almost the same magnitude (PIC = 0.88) as reported byAzad *et al*. (2014) that used RAPD to ascertain the diversity among Phoenix *spp.*Garcia *et al*. (2004)and Cholastova *et al*.(2011)reported lower PIC values of 0.75 and 0.61 among maize tropical inbred lines and hybrids respectively using RAPD. The high PIC value further highlights the high level of genetic diversity revealed among the maize genotypes by the markers.

RAPD marker pattern profiles have shown to be associated with geographical locations in barley (Fernández *et al.,* 2002), maize (Al-badeiry *et al*., 2013) and coconut (Masumbuko *et al*., 2014). Contrary to the above findings, patterns of geographical distribution of genetic diversity based on molecular data were not observed among the maize landraces. The dendrogram showed that clustering was rather based on genetic relatedness.The cluster with bootstrap value of 96% consisted of 3 landraces viz. 1760 (MN50), 26

(LNA274) and 2840 (MS170) with geographical locations of Segosera, Nachingwa, and Muheza districts respectively.The genotypes that clustered within the various major and minor clusters had varying agro-ecological locations.This observation relates with the entire landraces clustered within the dendrogram. The grouping of clusters in the dendrogram might possibly reflects the genetic history, agro-ecological relationship among the various landraces. This observation is in agreement with studies conducted in soybean (Brown-Guedira *et al*., 2000) and common bean (Gómez *et al.,* 2004) where RAPD had no correlation with geographical locations. A plausible explanation may be due to the high levels of gene flow within the various geographical locations and thus highlights the high level of diversity observed in the study.

Genotypes that were not found in the same cluster were genetically different and diverse from each other.In order to develop improved variety, genotypes belonging to different clusters should be targeted since the genetic relatedness varies. Closely related genotypes when crossed among themselves will not be vary genetically and will result in low genetic diversity hampering the improvement of the breeding stock. This information will help breeding programs to develop maize varieties that are more diverse to enable combat the global challenges surrounding food security. Challenges of developing varieties that will have lower improved qualities will be averted since the genetic resources of the new variety will possess traits that are more diverse and genetically different from parent's genotype.

CHAPTER SIX

6.0 CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The present study has shown that RAPD markers are able molecular markers for the assessment of genetic diversity and relationship among maize landraces in Tanzania. The RAPD molecular markers used recorded a 100 percent polymorphism signifying its ability to still analysethe genetic diversity of the studied maize landraces. This information helps to identify diverse sources in maize germplasm collection. RAPD revealed genetic diversity among the landraces but did not reflect the geographical locations of the studied genotypes.

Genotypes that were grouped in the same cluster were genetically similar, whilst those grouped in different set of clusters were genetically different from other group of clusters. The information obtained from this study will help plant breeders to further identify suitable landraces for introgression into breeding stocks promoting improvement and conservation of maize genetic resource in Tanzania.

6.2 Recommendations

Based on the findings of this study, it is recommended that plant breeders select landraces that are genetically diverse for further breeding programs to enhance maize improvement. The implication is that selecting landraces that are genetically related for breeding programswill lead to the development of new maize varieties that are not genetically diverse but rather identical to the selected landraces. The more diverse the parents are for breeding programs will result in highly improved new maize varieties.

Improvement can be achieved by selecting genetically diverse landraces that are not related (clustered) to characterize new variety. Genotypes from the topmost cluster with bootstrap value of 55% can be crossed with genotypes clustered with a bootstrap value of 92%. For example, local landrace Mehhi (T11) can be crossed with genotype Segosera (1760). These genetically diverse genotypes when crossed can improve and enhance the breeding stocks because the genotypes possess genetically different traits.

Also, selection process for maize conservation programshould be geared towards landraces that are not genetically identical since genetically related landraces may possess similar traits and characteristics thus increasing redundancies during maize conservation.Genotypes that are clustered together are more genetically identically hence with regards to conservation; not all genotypes within a cluster should be considered for conservation but rather selected genotypes to represent the entire group of genotypes in that clster. For example, landrace Kitwele with reference code MS173 from Tanga region with a bootstrap value of 92% can be conserved without conserving landrace Mwarabu (3936) or vice versa due to their genetic relatedness.

Furthermore, sequencing of the different maize genotypes and the application of SSR should be undertaken to provide supplementary data to support this study. The information obtained from this study will serve as the building block for more in depth research to be carried on the selected maize genotypes, hence the need to use more robust and improved molecular marker tools.

REFERENCES

Abate, A., Dzowela, B. H. and Kategile, J. A. (1993). Future of livestock industries in East and Southern Africa: *Proceedings of the Workshop Held at Kadoma Ranch Hotel,* Internatinal livestock Research Institute, Zimbabwe. 20-23 July 1992.227pp.

Abdel-Mawgood, A. L. (2012). *DNA Based Techniques For Studying Genetic Diversity*. Genetic Diversity in Microorganisms, Croatia. 122pp.

Abuali, A. I., Abdelmula, A. A. and Khalafalla, M. M. (2011). Assessment of genetic diversity in Sudanese maize (*Zea mays* L.) genotypes using random amplified polymorphic DNA (RAPD) markers. *African Journal of Biotechnology*10(42): 8245–8250.

Abuali, A. I., Abdelmula, A. A., Khalafalla, M. M., Hamza, N. B., Abdalla, A. H. and Idris, A. E. (2014). Assessment of Yield Stability and Adaptability of Parental Inbred Lines and F 1 - Hybrids of Grain Maize (*Zea mays* L.) Using AMMI Analysis. *British Biotechnology Journal*4(4): 339–349.

Afonso, A. V. (2013). *Genetic Diversity of Local Maize (Zea Mays* L. *) Germplasm From Eight Agro-Ecological Zones.* Swedish University of Agriculture Sciences, Sweden.37pp.

Al-badeiry, N. A. H., Merza, T. K. and Al-saadi, A. H. (2013). Assessment of genetic diversity and relationships among maize. *Journal of Life Sciences*7(12): 1260–1271.

Anderson E. and Cutler H. (1942). Races of *Zea mays*: Their recognition and classification. *Annuals of the Missouri Botanical Garden*21: 69–88.

Anonymous (2008). The biology of *Zea mays* L. ssp mays (maize or corn). Office of the gene technology regulator. [http://www.ogtr.gov.au] site visited on 29/4/2016.

Anonymous (2014). Drought tolerant maize varieties ready for Tanzanian farmers COSTECH. [http://www.costech.or.tz/?page_id=3410] site visited on 29/4/2016.

Aremu, C. (2011a). Genetic diversity: A review for need and measurements for intraspecie crop improvement. *Journal of Microbiology Biotech. Research*1(2): 1–4.

Aremu, C. (2011b). *Exploring Statistical Tools in Measuring Genetic Diversity for Crop Improvement.* Department of Crop Production and Soil Science Lakoke Akintola Univerity of Technnology, Omu-Aram, Nigeria. 348pp.

Azad, A. K., Mizanur, M., Arefin, H., Sima, S. N. and Hossian, M. A. (2014). Genetic Diversity and relationship of arabian and bangladeshi date palms (*Phoenix* spp.) revealed by RAPD fingerprints. *Journal of Nature Science And Sustainable Technology*8(4): 575–588.

Balazova, Z., Vivodk, M. and Galova, Z. (2015). Evaluation of molecular diversity of central European maize cultivars. *Emirates Journal of Food and Agriculture*28(2): 93–98.

Broadwater, A. H., Rubinstein, A. L., Chay, C. H., Klapper, D. G. and Bedinger, P. A. (1993). Zea mI, the maize homolog of the allergen-encoding gene of rye grass. *Gene*131(2): 227–30.

Brown-Guedira, G. L., Thompson, J. A., Nelson, R. L. and Warburton, M. L. (2000). Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Science*40(3): 815–823.

Bruel, D. C., Carpentieri-pípolo, V., Ruas, C. D. F., Gerage, A. C. and Graciele, S. (2007). Assessment of genetic diversity in maize inbred lines using RAPD markers. *Crop Breeding and Applied Biotechnology*7(2): 173–178.

Carvalho, V. P., Ruas, C. F., Ferreira, J. M., Moreira, R. M. P. and Ruas, P. M. (2004). Genetic diversity among maize (*Zea mays* L.) landraces assessed by RAPD markers. *Genetics and Molecular Biology*27(2): 228–236.

Chen, H. (2002). Principal component analysis with missing data and outliers. Rutgers University, New Jersey. [http://venus.unive.it/romanaz/datamin/articoli/chen.pdf\npapers3://publication/uuid]site visited on 10/4/2016.

Cochrane, N. and Souza, A. D. (2015). *Measuring Access to Food in Tanzania*. A Food Basket Approach, United States Deparments of Agriculture, Washington DC.135pp.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B. and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*142(2): 169–196.

Coulson, A. and Diyamett, B. (2012). Improving the contribution of agricultural research to economic growth: Policy implications of a scoping study in Tanzania. *Presernted at Research On Poverty Alleviation, 17th Annual Research Workshop Held at the Whitesands Hotel*, Dar es Salaam, Tanzania. 28 – 29 March, 2012. pp. 1 – 44.

Darrah, L. L., McCullen, D., Zubber, M. S. (2003). *Corn: Chemistry and Technology.* (2nd Edition), American Association of Cereal Chemist Inc., Minnesota, USA. 60pp.

Dear, P. H. (2001). Restriction fragment length polymorphism.London. [http://dx.doi.org/ 101038/npg.els005358] site visited on 19/6/2016.

Dellaporta, S. L., Wood, J. and Hicks, J. B. (1983). Plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter*1(4): 19–21.

Doebley, J. (2004). The genetics of maize evolution. *Annual Review of Genetics*38: 37–59.

Everitt, B., Landau, S. and Lesse, M. (2011). *Cluster Analysis.* (5th Edition),John Wiley and Sons,London, UK.pp15-17.

Food and Agriculture Organization (1992). Maize in human nutrition Rome,Italy. [http://www.fao.org/docrep/t0395e/T0395E01.htm#Acknowledgments] site visited on 15/4/2016.

Food and Agriculture Organization (2012). Yield per hectare of maize in Tanzania. [http://faostat3.fao.org] site visited on 18/4/2016.

Food and Agriculure Organization (2014). Annual maize production in Tanzania [http://faostat3.fao.org] site visited on 18/4/2016.

Feed the Future (2013). Policy reforms in Tanzania improve opportunities for maize farmers.[feedthefuture.gov/article/policy-reforms-tanzania-improve-opportunities -maize-farmers]site visited on 15/4/2016.

Fernández, M., Figueiras, A. and Benito, C. (2002). The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. *Theoretical and Applied Genetics*104(5): 845–851.

Foster, J. T., Allan, G. J., Chan, A. P., Rabinowicz, P. D., Ravel, J., Jackson, P. J. and Keim, P. (2010). Single nucleotide polymorphisms for assessing genetic diversity in castor bean (*Ricinus communis*). *BMC Plant Biology*10: 1 – 13.

Galarreta, R. de and Alvarez, A. (2001). Morphological classification of maize landraces from northern Spain. *Genetic Resources and Crop Evolution*48(4): 391–400.

Galeano, C. H., Cortés, A. J., Fernández, A. C., Soler, Á., Franco-Herrera, N., Makunde, G. and Blair, M. W. (2012). Gene-based single nucleotide polymorphism markers for genetic and association mapping in common bean. *BMC Genetics*13: 1 – 48.

Garcia, A. F., Benchimol, L. L., Barbosa, A. M. M., Geraldi, I. O., Souza, C. L. and de Souza, A. P. (2004). Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology*27(4): 579–588.

Gómez, O. J., Blair, M. W., Frankow-Lindberg, B. E.and Gullberg, U. (2004). Molecular and phenotypic diversity of common bean landraces from Nicaragua. *Crop Science*44(4): 1412–1418.

Gopalan, C. and Rama, S.B.V. (2007). *Nutritive Value of Indian Food Hyderabad*. National Institute of Nutrition, Indian Council of Medical Research, India. 54pp.

Gordon, K. (1999). *Classification.* (2nd Edition),Chapman and Hall/ CRC, Florida. 27pp.

Govarthanan, M. (2011). Genetic variability among Coleus sp. studied by RAPD banding pattern analysis. *International Journal for Biotechnology and Molecular Biology Research*2(12): 202–208.

Govindaraj, M., Vetriventhan, M. and Srinivasan, M. (2015). *Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives.*Hindawi Publishing corporation, India. 14pp.

Gwakisa, P. S., Kemp, S. J. and Teale, A. J. (1994). Characterization of Zebu cattle breeds in Tanzania using random amplified polymorphic DNA markers. *Animal Genetics*25(2): 89–94.

Hulse de SOuza, S. G., Carpentieri-Pipolo, V., Ruas, C. D. F., Carvalho, V. D. P., Ruas, P. M. and Gerage, A. C. (2008). Comparative Analysis of Genetic Diversity Among the Maize Inbred Lines (*Zea mays* L .) Obtained by RAPD and SSR Markers. *Brazilian Archives of Biology and Technology*51(1):183–192.

Internationa Institute of Tropical Agriculture (2012). Vitamin A-enriched maize released in Nigeria, seen to benefit health of millions /CGIAR /News. [http://www.cgiar. org/conso rtium-news/vitamin-a-maize-released-nigeria] site visited on 26/5/2016.

Kaliba, A., Verkuijl, H., Mwangi, W., Mwilawa, A. J. T., Anandajayasekeram, P. and Moshi, A. J. (1998). *Maize Production Technologies in Central Tanzania Adoption of Maize*. International Maize and Wheat Improvement Center, Mexico. 5pp.

Kassie, M., Jaleta, M. and Mattei, A. (2014). *Improved Maize Varieties and Household Food Security: Achieving Impact in Tanzania*.Policy Brief No. 6. Australian Centre for International Agricultural Research. Australia. 3pp.

Katinila, N., Verkuijl, H., Mwangi, W., Anandajayasekeram, P. and Moshi, A. J. (1998). *Maize Production Southern Tanzania Adoption of Maize*. InternationalMaize and Wheat Improvemtn Center, Mexico. 5pp.

Kimani, E. N., Wachira, F. N. and Kinyua, M. G. (2012). Molecular diversity of Kenyan lablab bean (*Lablab purpureus*L.) Sweet accessions using amplified fragment length polymorphism markers. *American Journal of Plant Sciences*3: 313–321.

Konstantinov, K. and Mladenovic, S. (2007). Molecular Genetics -step by step implementation in maize breeding. *Genetika*39(2):139–154.

Kumari, N. and Thakur, S. K. (2014). Randomly Amplified Polymorphic DNA - a brief review. *American Journal of Animal and Veterinary Science*9(1): 6–13.

Kwok, P.-Y., Deng, Q., Zakeri, H., Taylor, S. L. and Nickerson, D. A. (1996). Increasing the information content of sts-based genome maps: identifying polymorphisms in mapped STSs. *Genomics*31(1): 123–126.

Langade, D. M., Shahi, J. P., Srivatava, K., Singh, A., Agarwal, V. K. and Sharma, A. (2013). Appraisal of Genetic Variability and Seasonal Interaction for Yield and Quality Traits in Maize (*Zea mays* L.). *Plant Gene and Trait* 4(18): 95 – 103.

Liu, K. and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*21(9): 2128–2129.

Lowe, A. J., Hanotte, O. and Guarino, L. (1996). *Standardization of Molecular Genetic Techniques for the Characterization of Germplasm Collections: The Case of Random Amplified Polymorphic DNA.*Plant Genetic Resources Newsletter, 54pp.

Lyimo Stephen, Zubeda, M. and Hugo, D. G. (2014). The use of improved maize varieties in Tanzania. *African Journal of Agricultural Research*9(7): 643–657.

Lyimo, N. G. (2005). *Improving Farmers' Access to and Management of Disease Resistant Cultivars in the Southern Highlands of Tanzania*. Uyole Agricultural Research Institute, Mbeya, Tanzania. 96pp.

Monitoring African Food and Agricultural Policies (2013). *Analysis of Incentives and Disincentives for Maize in the United Republic of Tanzania*.Food and Agriculture Organizaton, Rome, Italy. 21pp.

Magorokosho, C. (2006). Genetic Diversity and Performance of Maize Varieties From Zimbabwe, Zambia and Malawi.Thesis for Award of PhD Degree at A and M University, Texas, 179pp.

Mammadov, J., Aggarwal, R., Buyyarapu, R. and Kumpatla, S. (2012).*SNP Markers and Their Impact on Plant Breeding.*Hindawi Publishing Corportion, India. 11pp.

Masumbuko, L. I., Sinje, S. and Kullaya, A. (2014). Genetic diversity and structure of East african tall coconuts in Tanzania Using RAPD Markers. *Open Journal of Genetics*4:175–181.

Matsuoka, Y., Mitchell, E. S., Kresovich, S., Goodman, M. and Doebley, J. (2002). Microsatellites in Zea -variability, patterns of mutations, and use for evolutionary studies. *Theoretical and Applied Genetics*104(2): 436–450.

Miracle, M. (1965). The Introduction and Spread of Maize in Africa on JSTOR. *Journal of African History*6(1):39–55.

Mohammadi, S. A. and Prasanna, B. M. (2003). Review and interpretation analysis of genetic diversity in crop plants salient statistical tools. *Crop Science*43: 1235–1248.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R. and Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding*3(2):87–103.

Molin, D., Coelho, C. J., Máximo, D. S., Ferreira,F.S., Gardingo, J. R. and Matiello, R. R. (2013). Genetic diversity in the germplasm of tropical maize landraces determined using molecular markers.*Genetics and Molecular Research* 12 (1): 99-114

Mrutu, B. A. and Tileye, F. and Ndunguru, J. (2014). Assessment of genetic diversity of maize inbred lines and hybrids in Southern Highlands of Tanzania by using random amplified polymorphism markers. *American Journal of Research Communication*2(4): 84–98.

Mukharib, D. S., Patil, V. C., Biradar, D. P., Salimath, P. M. and Chimmad, V. P. (2010). Assessment of molecular diversity in selected maize inbreds. *Karnataka Journal of Agricultural Science*23(3): 409–412.

Newton, C. R. and Graham, A. (1997). Polymerase chain reaction. Oxford, UK.[http://catalog.hathitrust.org/api/volumes/oclc/36458142.html] site visited on 20/4/2016.

Nkonya, E., Xavery, P., Akonaay, H., Mwangi, W., Anandajayasekeram, P., Verkuijl, H. and Moshi, A. (1998). *Adoption of Maize Production Technologies in Northern Tanzania*. *Production*.International Maize and Wheat Improvement Center, Mexico. 29pp.

Noel, F., Machon, N. and Porcher, E. (2007). No genetic diversity at molecular markers and strong phenotypic plasticity in populations of Ranunculus nodiflorus, an endangered plant species in France. *Annuals of Botany* 99(6): 1203–1212.

Nuss, E. T. and Tanumihardjo, S. A. (2010). Maize: A paramount staple crop in the context of global nutrition. *Comprehensive Reviews in Food Science and Food Safety*9(4): 417–436.

Nyambo, B. (2009). *Agricultural Sector Development Program.Environmental Management.*United Republic of Tanzania*,* Dar es Salaam,Tanzania.24pp.

OECD (2008). Novel Food and Feed Safety SET 1: Safety Assessment of Transgenic Organisms Consensus Documents.[http://www.oecd.org/biotrack] site visited on 22/4/2016.

Paliwal R. (2000). *Maize Types. In: RL Paliwal, G Granados, HR Lafitte, AD Vlo Llc, eds*. *Tropical Maize: Improvement and Production.* Food and Agriculture Organization,Rome Italy. 43pp.

Pastorello, E. A., Pravettoni, V., Calamari, A. M., Banfi, E. and Robino, A. M. (2002). New plant-origin food allergens. *Allergy* 57(72): 106–110.

Pimentel, D. and Patzek, T. (2005). Ethanol production using corn, switchgrass, and u'ood; Biodiesel production using soybean and sunflower. *Natural Resources Research*14(1): 65–76.

Pingali, P. L. (2001).*CIMMYT 1999/2000 World Maize Facts and Trends. Meeting World Maize Needs.*Technological Opportunities and Priorities for the Public Sector,Mexico. 14pp.

Rohlf, F. J. (2000). NTSYS-pc: numerical taxonomy and multivariate analysis system, version 2.1.

Saidia, P., Chilagane, D., Wostry, A. and Maro, J. (2010). Evaluation of EM technology on maize (*Zea mays* L.) growth, development and yield in Morogoro, Tanzania. *Bustani Ya Tushikamane vol issue* 60pp.

Salami,H. A., Kamirou, C. S., Padonou, W., Djima, A., Yallou, C., Adjanohoun, A., Kotchoni, S., Baba-Moussa, L. (2016). Genetic Diversity of Maize Accessions (*Zea mays* L.) Cultivated from Benin Using Microsatellites Markers. *American Journal of Molecular Biology*6: 12-24.

Sanchez J. J., Goodman, M. M. and J. O. R. (1993). Appropriate characters for racial classification in maize. *Economic Botany*47: 44–59.

Shehata, A. I., Al-Ghethar, H. A. and Al-Homaidan, A. A. (2009). Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. *Saudi Journal of Biological Sciences*16(2): 57–62.

Smith, L. (2002). Tutorial on Principal Component Analysis. pp1-19.

Staiger, C.J., Goodbody, K.C., Hussey, P. J., Valenta, R. and Lloyd, C. W. (1993). The profilin multigene family of maize:differential expression of three isoforms. *Journal of Plant*4: 631–641.

Sun, G. L., William, M., Liu, J., Kasha, K. J. and Pauls, K. P. (2001). Microsatellite and RAPD polymorphisms in Ontario corn hybrids are related to the commercial sources and maturity ratings. *Molecular Breeding*7(1):13–24.

Suresh, H., Sasidharan N., Sudeshna, C., Sneha, M., Ruchi, T., Bhupendra, S. P. and Vala, A. G. (2013). Genetic diversity among maize varieties revealed by phenotypic descriptors and rapd profiles. *Journal of Agricultural Sciences*8(2): 1 - 91.

Tanksley, S. D. and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277(5329): 1063–6.

Thakur, N., Prakash, J., Sharma, J. K., Kumari, R., Rana, M., Lata, S. (2016).Genetic Diversity and Structure of Maize Accessions of North Western Himalayas Based on Morphological and Molecular Markers.Proceedings of the National Academy of Sciences, India Section B: Biological Sciences. doi:10.1007/s40011-016-0716-0, pp 1-14

United Nations (2004). *World Population to 2300.* Department of Economics and Social Affairs, New York. 415pp.

Verheye, W. (2010). Growth and production of maize: traditional low-input cultivation. *Encyclopedia of Life Support Systems*. National Science Foundation Flanders and Geography Department, University of Ghent, Belgium. 14pp.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V., Hornes, M. and Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*23(21): 4407–4414.

Wakeley, J., Nielsen, R., Liu-Cordero, S. N. and Ardlie, K. (2001). The Discovery of Single-Nucleotide Polymorphisms and Inferences about Human Demographic History. *American Journal of Human Genetics*69(6): 1332–1347.

Walunjkar, B., Parihar, A. and Singh, N. P. L. (2015). Genetic diversity of wild and cultivated genotypes of pigeonpea through RAPD and SSR markers. *Journal of Environmental Biology*36: 59–63.

Weichel, M., Glaser, A. G., Ballmer-Weber, B. K., Schmid-Grendelmeier, P. and Crameri, R. (2006). Wheat and maize thioredoxins: a novel cross-reactive cereal allergen family related to baker's asthma. *The Journal of Allergy and Clinical Immunology*117(3): 676–81.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*18(22): 6531–5653.

APPENDICES

Appendix 1: Locations of Maize Genotypes in Tanzania

Sample Id	Ref. Code	Local Name	Village	District	Region
4163	CCM 122	Epipo	Rushangwa	Biharamulo	Kagera
3069	MAN 209	_	Mkungugu	Iringa rural	Iringa
4092	CCM 51	Mahindi ya	Kazungute	Sengerema	Mwanza
3627	LNA44	Ndundungu	Nkutimang	Mtwara	Mtwara
3936	NAM 28	Malombe	Borenga	Serengeti	Mara
1744	MN 34	Ya kienyeji	Makete Iwa	Makete	Njombe
1763	MN 53	Ya kienyeji	Ruanda twa	Mbeya rural	Mbeya
1761	MN51	-	Isangati	Mbeya rural	Mbeya
4020	NAM113	Malingwa	Muluseni	Ukerewe	Mwanza
3851	LNA270	Pota	Namapwia	Nachingwe	Lindi
3951	NAM44	Mwarabu	Kibanchano	Serengeti	Serengeti
T14	TD-KR4	Mehhi	Slahhmo	Karatu	Arusha
2263	NJ32	Serena	Mahumbik	Lindi	Lindi
T13	AD-KR3	Mehhi	Upper Kitete	Karatu	Arusha
213	WMK160	Visaka	Kapele	Tunduma	Mbeya
4070	CCM 26	Gembe	Ibambila	Geita	Geita
2375	NJ143	Serena	Nnima	Mtwara	Mtwara
2945	MAN85	Katumbili	Kilosa	Mbinga	Ruvuma
T11	DM-KRI	Mehhi	Kilimatembo	Karatu	Arusha
3756	LNA174	Katumbili	Mkwiti Juu	Tandahimba	Mtwara
4206	CCM 165	Mapo	Sangabuye	Ilemela	Mwanza
1075	MN13	Ya kienyeji	Itunduma	Njombe	Iringa
T15	JL-KR5	Mehhi	Kambi ya Simba	Karatu	Arusha
2813	MS142	Mkonyoli	Boza	Pangani	Tanga
2261	NJ30	Yanga	Mahumbik	Lindi	Lindi
1772	NMA8	Mehe	Kansay	Karatu	Arusha
4010	NAM 103	Kagire	Wegero	Musoma	Mara
212	WMK159	Visaka	Kapele	Tunduma	Mbeya
2544	NJ315	Katumbili	Kichele	Nyamagana	Mwanza
4052	CCM7	Mzurunje	Busekeseke	Sengerema	Mwanza
43	MNC119	Kemphukwa	Senene Mf	Singida	Singida
4	MN7	Mbegu ya	Nzivi	Mufindi	Iringa
4000	NAM 93	Nchanana	Hamkoko	Ukerewe	Mwanza
26	LNA274	Matundila	Likongowe	Nachingwea	Lindi
23	LNA264	Pota	Namapwia	Newala	Mtwara
T9	DQ-BB5	Erikwatoo	Long	Babati	Manyara
14	NAM103	Kagire	Wegero	Musoma	Mara
12	NMA27	-	Alararashi	Monduli	Arusha
2	NMA6	Katumani	Kambi ya Simba	Karatu	Arusha
1	CCM156	Gembe	Lwanima	Nyamagana	Mwanza
5201	ML18	Majani me	Vuga	Mkinga	Tanga
36	LNA31	Malombe	Mkutimang	Mtwara	Mtwara
1760	MN50	Segosera	-	Mbeya	Mbeya

Sample Id	Ref. Code	Local Name	Village	District	Region
4212	CCM 171	Gembe	Kasamiko	Nyamagana	Mwanza
46	MNC138	Ipukele	Kisana	Iramba	Singida
6	MN14	Ya kienyeji	Mji mwema	Njombe	Iringa
47	NMA7	Katumani	Kansay	Karatu	Arusha
2908	MAN 48	Mapalapi	Mbatamila	Tunduru	Ruvuma
27	NH70	Ilonga	Maranzala	Tanga	Tanga
9	CCM 89	Gembe	Kabugozo	Geita	Geita
22	MAN73	Mawila	Likwiru Kilo	Mbinga	Ruvuma
38	NH76	Maricheni	Horohoro	Nkinga	Tabora
1766	NMA2	Katumbili	Kambi ya S	Karatu	Arusha
3663	LNA81	Mahindi	Kitaya	Mtwara	Mtwara
1730	MN20	Kienyeji	Nyamapino	Ludewa	Njombe
2685	MS12	Mampempe	Lusanga	Mvomero	Morogoro
3099	HPL15	-	Nyanganga	Kigoma	Kigoma
5	MN8	Kibena	Mninga	Mufindi	Iringa
3873	LNA292	Malombe	Ngunichile	Nachingwea	Lindi
LC	MN44	Ya kienyeji	Galijembe	Mbeya rural	Mbeya
8	MN29	Ya kienyeji	Mang'oto	Makete	Njombe
1734B	MN24	Mbegu ya	Malangali	Ludewa	Njombe
41	CCM165	Mapo	Sangabuye	Ilemela	Mwanza
39	NH77	Katumani	Horohoro	Nkinga	Tabora
2793	MS120	Mkonyoli	Katurukila	Kilombero	Morogoro
35	ML55	Kipekele	Matipwili	Bagamoyo	Pwani
2491	NJ262	Yanga	Mnima	Mtwara	Mtwara
13	MAN201	Katumbili	Lufunu	Mufindi	Iringa
37	LNA44	Ngundungu	-	Mtwara	Mtwara
2881	MS211	Kitwele	Potwe-Mpi	Muheza	Tanga
1755	MN45	Ya kienyeji	Galijembe	Mbeya rural	Mbeya
11	MN37	Ya kienyeji	Changarawa	Makete	Njombe
42	WMK160	Visaka	Kapele	Tunduma	Mbeya
1757	MN47	Ya kienyeji	Galijembe	Mbeya rural	Mbeya
44	CCM20	Ya kienyeji	Busekeseke	Sengerema	Mwanza
10	NAM134	Amalingwa	Igallu	Ukerewe	Mwanza
3054	MAN194	Katumbili	Igomaa	Mufindi	Iringa
20	LNA159	Gundugundu	Mkwiti Juu	Tandahimba	Mtwara
37B	LNA53	Malundila	Mkutimang	Mtwara	Mtwara
19B	ML3	Kinguu	Kuzekibago	Mkinga	Tanga
26B	ML17	Majani me	Mhinduro	Mkinga	Tanga
27B	NH67	Mahindi	Maranzala	Tanga	Tanga
47B	WMK9	Kalimwa	Myula	Sumbawanga	Rukwa
3964	NAM57	Amakuria	Kibancheba	Serengeti	Mara
3	MN5	Mbegu ya	Ibati	Mufindi	Iringa
1749	MN39	Filombe fre	Kidope Ithi	Makete	Njombe
4058	CCM13	Gembe	Busekeseke	Sengerema	Mwanza
1731	MN21	Mbegu ya	Mlangali	Ludewa	Njombe
7	MN31	Ya kienyeji	Tandala	Makete	Njombe

Sample Id	Ref. Code	Local Name	Village	District	Region
1720	MN10	Mbegu ya	Mninga	Mufindi	Iringa
5197	ML14	Majani me	Mhinduru	Mkinga	Tanga
2840	MS170	Kitwele	Potwe-Mpi	Muheza	Tanga
1775	NMA11	Yanga	Usale	Mbulu	Manyara
3167	HPL83	Urubinga	Nyakasanda	Kigoma	Kigoma
15	CCM122	Epipo	Rushangwa	Biharamulo	Kagera
17	CCM124	Ibahakazi	Luganzo	Biharamulo	Kagera
1727	MN17	Mbegu ya	Uwemba	Njombe	Njombe
T9	DQ-BB5	Erikwatoo	Long	Babati	Manyara
2491	NJ262	-	Mnima	Mtwara	Mtwara
2501	NJ272	Ipukele	Kinyangiri	Iramba	Singida
5124	NH29	Utakuna	Kwemhosi	Muheza	Tanga
3597	LNA13	Malombe	Chemchem	Mtwara	Mtwara
2947	MAN87	Lusewa	Likwiru Kilo	Mbinga	Ruvuma
3597	LNA13	Malombe	Chemchem	Mtwara	Mtwara
3095	HPL11	Fumandoli	Ilagala	Kigoma	Kigoma
T10	JM-ARI	-	Olorien	Arumeru	Arusha
4070	CCM 26	Gembe	Ibambila	Geita	Geita
1075	MN13	Kibena	Itunduma	Njombe	Iringa
5124	NH29	Utakuna	Kwemhosi	Muheza	Tanga
3501	AHM24	Oloman	Kibogwa	Morogoro	Morogoro
1748	MN38	Ya kienyeji	Kidoni	Makete	Njombe
3067	MAN207	Katumbili	Mkungugu	Iringa rural	Iringa
707	MNC237	Nyamula	Kimakorere	Tarime	Mara
291	WMK159	Visaka	Kapele	Tunduma	Mbeya
T14	TD-KR4	Slahhmo	Mehhi	Karatu	Arusha
T8	MA-BB4	Ikweto	Endaw	Babati	Manyara
2588	NJ358	Mampempe	Lusanga	Mvomero	Morogoro
1789	NMA25	-	Orarashi	Monduli	Arusha
2264	NJ33	Serena	Mahumbik	Lindi	Lindi
104	WMK57	Makonde	Liapona	Sumbawanga	Rukwa
2397	NJ165	Malingwa	Igallu	Ukerewe	Mwanza
1791	NMA27	Yanga	Alararashi	Monduli	Arusha
2447	NJ215	Segosera	Nachunyu	Lindi	Lindi
3936	NAM 28	Mwarabu	Borenga A	Serengeti	Serengeti
4062	CCM 18	Chalana	Busekeseke	Sengerema	Mwanza
1775	NMA11	-	Usale	Mbulu	Manyara
1773	NMA9	-	Usale	Mbulu	Manyara
2697	MS24	Mkonyoli	Katurukila	Kilombero	Morogoro
2843	MS173	Kitwele	Potwe-Mpi	Muheza	Tanga
2490	NJ261	Malundila	Mnima	Mtwara	Mtwara
3851	LNA270	Pota	Namapwia	Nachingwe	Lindi
2947	MAN87	Lusewa	Likwela-Ny	Mbinga	Ruvuma
3585	LNA1	Katumbili	Mtwara	Mtwara	Mtwara
181	WMK128	Amangagu	Igamba	Vwawa	Mbeya
4067	CCM23	Gembe	Kazungute	Sengerema	Mwanza

Sample Id	Ref. Code	Local Name	Village	District	Region
687	MNC217	Nyamula	Ochuna	Tarime	Mara
3075	MAN215	Fumandoli	Mkungugu	Iringa rural	Iringa
2463	NJ193	Mahindi	Kitaya	Mtwara	Mtwara
1752	MN42	Filombe fre	Misiwa	Makete	Njombe
3167	HPL83	Urubinga	Nyakasanda	Kigoma	Kigoma
2423	NJ193	Dungudung	Nkutimang	Mtwara	Mtwara
2990	MAN130	Kienyeji	Yakobi	Njombe	Njombe
1782	NMA18	Yanga	Gidimari	Babati	Manyara
2971	MAN111	Mapalapi	Yakobi	Njombe	Njombe
2289	NJ73	Mehe	Mahumbik	Lindi	Lindi
2405	NJ175	Serena	Nachunyu	Lindi	Lindi
1758	MN48	Mbatagwa	Maganzu	Mbeya rural	Mbeya
3789	LNA208	Ngundungu	Mapili	Newala	Mtwara
67	WMK9	Makonde	Muimwa	Namanyere	Rukwa
1751	MN 42	Filombe fre	Misiwa	Makete	Njombe
31	NH12	Ngoni	Lwengera	Korogwe	Tanga
5481	TLMZ19	Mwihindi	Matemwe	Kaskazini	Zanzibar
3982	NAM75	Amalingwa	-	Bugwema	Mara
40	CCM171	Gembe	Kasamiko	Nyamagana	Mwanza
2263	NJ32	-	Mahumbik	Lindi	Lindi
12	ML1	Wahi	Kuzekibago	Mkinga	Tanga
T8	MA-BB4	Ikweto	Endamanang	Babati	Manyara
5124	NH29	Utakuna	Kwemhosi	Muheza	Tanga
2501	NJ272	Serena	Nachunyu	Lindi	Lindi
37B	MN43	Ya kienyeji	Kimondo	Mbeya rural	Mbeya

Appendix 2: Extraction protocol of nucleic acid (DNA)

A. Extraction of DNA from plant tissue as described by Dellaporta *et al.* (1983)with some modifications.

Note: Before working with your samples makes sure you sterilize your equipment, autoclave tips, eppendorf tubes, double distilled water, mortar and pestles. Sterilize your working bench with 70% ethanol. The aim is to avoid nuclease enzymes, inhibitors and contaminants**.**

Protocol:

1. Grind 200 mg of plant leaf tissue in 700 µl of Dellaporta extraction buffer, pre-warmed to 65 °C (to dissolve any salts) in a mortar using pestles until a uniform mixture.

2. Transfer the mixture to 1.5 ml eppendorf tube and add 40 µl of 20 % Sodium dodecyl sulphate (SDS) to the tube.

3. Cap the appendorf tube and vortex or mix briefly and incubate the mixture in water bath at 65 °C for 30 min.

4. Add 200 µl of 3M sodium acetate (p.H 5.2) to each tube and mix thoroughly by briefly vortexing. Incubate the tubes in ice for 20 min.
5. Centrifuge the tube at 13,000 rpm for 10 min at room temperature in a microfuge.
6. Transfer 500 µl of the supernatant to new appendorf tube (labelled appropriately) avoiding leaf debris.

7. Add 10 µl of 10 mg/ml RNase A. Mix gently by inverting the tube 5 X and incubate in water bath at 37 °C for 30 min. This step digests any contaminating RNA.

8. Briefly spin the centrifuge to bring down all droplets.

9. In a fume hood, add equal volume of 500 µl of chloroform: isoamyl (24:1). Mix gently by inverting the tube 10 X.

10. Spin at 13,000 rpm for 10min at room temperature in a microfuge and transfer the upper aqueous layer (approximately 400-450 µl) into a new 1.5 ml eppendorf tube.

11. Add an equal volume of cold isopropanol (pre-chilled in -20 °C freezer) mix thoroughly by gently inverting the tube 10 X. Incubate at -20 °C in a freezer for 30 min.

12. Spin at 13,000 rpm for 10 min in a microfuge at room temperature. Carefully pouring off the liquid phase leaving the DNA pellet at the bottom of the eppendorf tube.

13. Wash the pellet by adding 500 ml of 70 % ethanol (at room temperature) to the appendorf tube containing the DNA pellet. Centrifuge at 14,000 rpm for 5 min at room temperature. Carefully pour off the ethanol.

14. Air dry the DNA pellet for 20-30 min by leaving the open and inverted the appendorf tube on tissue paper on the bench or in a fume hood.

15. Resuspend/dissolve the DNA pellet in 150-300 µl of sterile double distilled water depending on the size of the pellet.

16. Store at -20 ^{o}C for further use.

Appendix 3: DNA concentrations and purity of extracted maize samples

LAB. ID.	SAMPLE ID	DNA Concentration (ng/µl)	A260/280
1	4163	207.6	1.81
2	3069	932.0	1.84
3	4092	3376.0	1.89
4	3627	322.2	1.86
5	3936	361.3	1.82
6	1744	571.2	1.89
7	1763	461.3	1.86
8	1761	313.5	1.89
9	4020	961.9	1.82
10	3851	209.7	1.82
11	3951	486.1	1.85
12	T14	661.8	1.88
13	2263	552.1	1.81
14	T13	541.2	1.83
15	213	621.9	1.81
16	4070	418.0	1.85
17	2375	754.2	1.87
18	2945	205.8	1.83
19	T11	682.4	1.81
20	3756	717.1	1.89
21	4206	518.0	1.84
22	1075	361.0	1.85
23	T15	302.9	1.87
24	2813	501.0	1.89
25	2261	219.6	1.84
26	1772	639.4	1.82
27	4010	200.7	1.81
28	212	843.1	1.86
29	2544	135.8	1.84
30	4052	215.0	1.89
31	43	332.8	1.88
32	4	677.8	1.83
33	4000	851.1	1.81
34	26	642.8	1.86
35	23	520.7	1.84
36	T9	532.3	1.87
37	14	669.6	1.89
38	12	670.2	1.82
39	2	3261.0	1.88
40	1	2516.9	1.88
41	5201	2638.9	1.88
42	36	2853.7	1.84
43	1760	325.6	1.81

LAB. ID.	SAMPLE ID	DNA Concentration (ng/µl)	A260/280
44	4212	212.9	1.86
45	46	750.8	1.82
46	6	333.5	1.85
47	47	561.3	1.87
48	2908	203.2	1.84
49	27	572.9	1.89
50	9	767.1	1.81
51	22	537.8	1.87
52	38	568.8	1.82
53	1766	849.1	1.89
54	3663	390.8	1.87
55	1730	299.3	1.89
56	2685	669.5	1.82
57	3099	203.1	1.84
58	5	764.2	1.81
59	3873	486.8	1.86
60	LC	732.1	1.85
61	8	241.1	1.82
62	1734B	431.9	1.88
63	41	297.4	1.89
64	39	221.5	1.81
65	2793	208.5	1.86
66	35	643.2	1.82
67	2491	523.4	1.86
68	13	352.8	1.89
69	37	561.1	1.84
70	2881	468.8	1.89
71	1755	328.6	1.84
72	11	399.4	1.87
73	42	854.9	1.85
74	1757	535.1	1.88
75	44	416.7	1.81
76	10	310.5	1.89
77	3054	384.2	1.84
78	20	392.0	1.87
79	37B	238.1	1.87
80	19B	249.0	1.85
81	26B	367.4	1.83
82	27B	334.7	1.85
83	47B	362.0	1.86
84	3964	1604.0	1.89
85	3	1764.0	1.86
86	1749	1488.0	1.88
87	4058	851.4	1.81
88	1731	1665.4	1.85

LAB. ID.	SAMPLE ID	DNA Concentration (ng/µl)	A260/280
89	7	363.4	1.83
90	1720	1272.3	1.89
91	5197	1449.4	1.87
92	2840	902.1	1.83
93	1775	824.3	1.87
94	3167	242.0	1.82
95	15	943.0	1.89
96	17	2056.6	1.88
97	17	239.0	1.85
98	1727	1349.6	1.87
99	T9	1918.9	1.88
100	2491	483.1	1.84
101	2501	2997.9	1.87
102	5124	205.3	1.89
103	3597	2015.5	1.82
104	2947	205.3	1.86
105	3597	2352.6	1.83
106	3095	214.4	1.87
107	T10	313.6	1.82
108	4070	753.5	1.89
109	1075	220.8	1.84
110	5124	418.9	1.87
111	3501	1460.6	1.81
112	1748	988.8	1.88
113	3067	1977.1	1.83
114	707	3374.3	1.86
115	291	397.6	1.89
116	T14	742.2	1.81
117	T8	830.0	1.86
118	2588	1532.9	1.83
119	1789	2403.8	1.86
120	2264	2274.4	1.84
121	104	1814.5	1.81
122	2397	1804.0	1.86
123	1791	681.8	1.83
124	2447	402.9	1.88
125	3936	660.9	1.85
126	4062	3143.2	1.89
127	1775	2088.4	1.84
128	1773	2275.7	1.88
129	2697	1901.1	1.87
130	2843	1585.8	1.81
131	2490	234.5	1.86
132	3851	228.6	1.89
133	2947	4123.0	1.85

LAB. ID.	SAMPLE ID	DNA Concentration (ng/µl)	A260/280
134	3585	472.8	1.81
135	181	511.0	1.89
136	4067	738.9	1.87
137	687	1524.2	1.82
138	3075	1989.5	1.85
139	2463	2183.8	1.81
140	1752	1151.6	1.87
141	3167	1989.5	1.82
142	2423	2183.8	1.81
143	2990	1151.6	1.86
144	1782	1947.2	1.84
145	2971	894.9	1.89
146	2289	841.5	1.86
147	2405	1910.0	1.81
148	1758	3478.0	1.87
149	3789	654.3	1.84
150	67	395.5	1.81
151	1751	1129.6	1.89
152	31	482.8	1.87
153	5481	246.4	1.83
154	3982	980.6	1.87
155	40	820.0	1.85
156	2263	220.3	1.89
157	12	681.0	1.87
158	T8	241.2	1.85
159	5124	235.4	1.87
160	2501	229.5	1.82

A=Absorbance