GENETIC DIVERSITY AND POPULATION STRUCTURE OF NAPIER GRASS (*Pennisetum purpureum* Schumach) IN WESTERN KENYA

BY

AWALLA BORNIFACE JUMA

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN BOTANY (GENETICS) OF UNIVERSITY OF ELDORET, KENYA.

SEPTEMBER, 2014

DECLARATION

Declaration by the Candidate

This thesis is my original work and has not been presented for a degree in any other University. No part of this thesis may be reproduced without the prior written permission of the author and/or University of Eldoret.

Awalla Borniface Juma

……. ……
Signature Date

SC/PGB/047/09

Declaration by Supervisors

This thesis has been submitted for examination with our approval as University Supervisors.

……..…… ……..……

Signature Date

Dr. Beatrice A. Were,

Department of Biological sciences,

University of Eldoret,

P.O Box 1125-3011 Eldoret, Kenya.

…… ……..……

Signature Date

Professor. O. G. Dangasuk,

Department of Biological sciences,

University of Eldoret,

P.O Box 1125-3011 Eldoret, Kenya.

DEDICATION

This work is dedicated to my late father, George Awalla Ododa who succumbed to prostate cancer, my dear Wife Olgah Anyango Otieno and my daughters, Faith Atieno Juma, Shalom Adhiambo Juma and Vicky Akoth Juma.

ABSTRACT

Napier grass (*Pennisetum purpureum* Schumach*)* is the dominant livestock fodder in all stall feeding systems in Kenya with other numerous applications. To date, no study in Kenya has been done on genetic diversity in napier grass that use both molecular and morphological markers. The objectives of this study was to determine genetic diversity and population structure of napier grass in western Kenya using simple sequence repeat (SSR) markers and morphological traits. A total of 96 samples out of 116 were studied using 25 selected SSR markers. The mean polymorphic loci index was 91.11% with the lowest being Alupe population at 85.19% while Busia and Butere at 94.44% each. Shannon information index ranged from 0.028 to 0.492 while Nei's diversity index ranged from 0.021 to 0.329. Nei genetic distance ranged from -0.002 to 0.047 while identity ranged from 0.954 to 0.985. Most (90%) of the molecular variation in napier grass populations exists among individuals within populations, with lesser amount (1%) between populations. There was no distinct population structuring in the five populations studied. Morphological characterization showed moderate diversity with two major clusters and one minor cluster. This corroborated with molecular clustering data. The most important morphological characteristics for diversity study were; plant height, stool diameter, growth form, leaf length and tiller numbers. This study recommends increasing the level of diversity in the Western Kenya napier grass germplasm through introductions of new napier clones and proper selection so as to increase the chances of getting resistant genes to Napier Stunting Disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

16Sr	16S Ribosomal gene sub-unit
AFLP	Amplified Fragment Length Polymorphism
Amp	Antigenic Membrane Protein
ASAL	Arid and Semi Arid Land
BeCA	Biosciences East and Central Africa
DM/ha	Dry Matter per hectare
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
dNTPs	Deoxy Nucleotide Triphosphate
EDTA	Ethylenediaminetetracetic acid
EST	Expressed Sequence Tags
GDP	Gross Domestic Product
ICIPE	International Center for Insect Physiology and Ecology
ILRI	International Livestock Research Institute
KARI	Kenya Agricultural Research Institute
LAMP	Loop Mediated Isothermal Amplification
MAS	Marker Assisted Selection

mDNA	Mitochondrial DNA
$MgCl_2$	Magnesium chloride salt
NaCl	Sodium Chloride salt
NCBI	National Centre for Biotechnology Information
nPCR	Nested Polymerase Chain Reaction
NSD	Napier Stunt Disease
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PIC	Polymorphic information content
PVP	Polyvinylpyrrolidine
RAPD	Random Amplified Polymorphic DNA
rDNA	Ribosomal DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
SDS	Sodium dodecyl sulphate
*Sec*A/Y/E	Secretion proteins A, Y & E

SSR	Simple Sequence Repeats
STR	Short Tandem Repeats
Taq	*Thermus aquaticus*
TE	Tris EDTA buffer
Tris-HCl	Trizma base and Hydrochloric acid

ACKNOWLEDGEMENTS

This research would be void without the support from different personalities along different levels of the study. I am extremely grateful to my supervisors Dr. Beatrice Were and Prof. Otto George Dangasuk who granted me immense support and especially for their in depth critique of the work and above all being understanding to the challenges that came along in my study life. Special thanks to the sponsors of this work, the World Bank through the East Africa Agricultural Productivity Project (EAAPP), Biosciences East and Central Africa (BeCA) for offering me the African Biosciences Challenge Fund fellowship for nine months based at ILRI Nairobi. For the fellowship period, Dr Rob Skilton was instrumental in the supervision of the overall work to ensure that the outputs were on target and achievable within the fellowship period. The BeCA technical team which included Mr. Moses Njahira, Ms. Martina Kyalo, Mr. Wilson Kimani, Mr. Nzuki Innosters, Mrs. Lucy Muthui among others were of immense help from DNA extraction to data analysis and interpretation. Analysis of data was made even easier by the assistance of Dr Mark Wamalwa and Joyce Njuguna.

I draw my attention to Bramwel Waswa Wanjala who always encouraged me and for his immense technical prowess in genetic diversity studies. My sincere gratitude goes to my late father, George Awalla Ododa who was always keen on my academic journey and who all along this study time had been battling with prostate cancer but unfortunately succumbed to the disease.

Many a time I found myself in low esteem and could hardly think of finishing this work on time. I found myself away from the people dear to my life for nearly two years. To my three lovely daughters, Faith Atieno Juma, Shalom Adhiambo Juma and Victoria Akoth Juma, sorry for leaving you with mom during my study time, but I thank you for always inquiring about my progress. To my dear wife Olgah Anyango Otieno, who persevered the loneliness and huge family responsibility, I salute you for your strength as a woman. Above all, this work was made possible through the wonderful power of almighty God through my savior Jesus Christ; he is merciful and powerful.

CHAPTER ONE

INTRODUCTION

1.1 Background information

Dairy farming is vital for the livelihoods and food security of millions of Kenyans. More than 80% of milk produced and sold in Kenya comes from smallholder farmers having one or two dairy cows on small plots of land (Orodho, 2006).

Kenya has over 3 million dairy cows, the largest in East Africa and 75% of which are kept by stall feeding or zero grazing (Orodho, 2006). Zero grazing or stall feeding is where cattle are confined in stalls and fed on fodder. Many of the feeding units are found among the small holder farmers in densely populated areas of Central, Western, North and Central Rift Valley and some parts of Coast region of Kenya. Stall feeding is also rising due to the inevitable spread of aridity due to climate change, and shrinking rangelands (Lekasi, 2000).

Napier grass (*Pennisetum purpureum* Schumach) also known as Elephant grass is a perennial C4 monocot originating from Africa that grows in bamboo-like clumps (Anderson et al., 2008) and may grow to a height of about 10 meters (Boonman., 1997). Napier grass is the livestock fodder of preference in all stall feeding systems in Kenya. Farmers use over 30% of their land for napier grass (Lekasi, 2000), and some are grown on the roadsides. Valued for its high biomass, perennial nature, high leaf nutritive value and moderate pest resistance (Bhandari et al., 2006), napier grass is cultivated for forage worldwide and widely used in South America and Africa (Hanna et al., 2004). Napier

grass can grow in a wide range of soil conditions, has good drought tolerance, high photosynthesis efficiency and good water use efficiency (Anderson et al., 2008). It also traps cereal stem borer insect pest. Many cereal farmers especially in western Kenya usually establish a hedge of Napier grass around their cereal plots to trap and control the pest using push- pull technology (Khan et al., 2001).

Napier grass survives aridity and its deep roots and rhizomes bind up the soil to prevent soil erosion. However, utilization of Napier grass in Kenya is threatened by Napier stunting disease which is caused by phytoplasma 16sr XI (Khan et al., 2001).

Napier grass stunt disease has been shown to have devastating effect on smallholder dairy farmers in Kenya and other East African Countries. The affected plants are severely stunted, yellowed, and dies in 6 months (Mulaa et al., 2004). Stunted plants with low biomass are unable to sustain feed requirements of dairy cows hence farmers are forced to reduce their herd size with related reduction in farm income in the absence of alternative feeds. Phytoplasma 16sr XI is closely related to those that cause rice yellow dwarf, sorghum grassy shoot, sugarcane white leaf and sugarcane yellow leaf (Jones et al., 2004, 2006) this presents a challenge of the disease spreading to infect sugarcane, rice and sorghum leading to food insecurity in Kenya (Farrell et al., 2001, Jones et al., 2004).

Farmers have benefited from the management strategies which include rouging and burning of affected clones, use of clean planting materials, not sharing farm tools, use of farm yard manure among others. These measures have been packaged by researchers and other stakeholders through extension offices at various levels of administration (Awalla

et al., 2009). However these management strategies are not sustainable hence the need to be supplemented with production of tolerant/resistant high yielding clones adaptable to these regions. The convenient approach is to develop resistant clones that can form a base population carrying the candidate genes for future selection (Awalla et al., 2009).

Characterization of genetic resource collections has been greatly facilitated by the availability of a number of morphological and molecular marker systems. Morphological markers were among the earliest to be used in germplasm management and were fully exploited in Mendelian era. However, morphological variation does not always reflect real genetic variation because of genotype by environment interaction and the largely unknown genetic control of polygenic morphological and agronomic traits. The limitations encountered with morphological traits include low polymorphism, low heritability, late expression, and vulnerability to environmental influences. On the other hand, molecular markers do not have such limitations and can be used to detect variation at the DNA level and have proven to be effective tools for distinguishing between closely related genotypes (Edwards et al., 1991).

Over the last years, many accessions of Napier grass or Napier grass inter specific hybrids with pearl millet (*P. glaucum)* have been produced or collected throughout the world (Azevedo et al., 2012). The nomenclature of napier grass vary from one region another region but with the same morphological and biochemical characteristics. In the East and Central Africa, there is lack of passport/ pedigree information of Napier grass for efficiency in breeding. This passport information is important especially when developing clones resistant to Napier stunting disease therefore; assessment of the genetic

relationships within the available germplasm is needed. Genetic diversity studies in plant species have been performed using isozymes, Amplified Fragment Length Polymorphism (AFLP) (Vos et al.,1995), Random Amplified Polymorphic DNA (RAPD) (Lynch et al., 1994), Sequence Related Amplified Polymorphism (SRAP), Simple Sequence Repeats (SSR) (Glaszmann et al., 2003), Restriction Fragment Length Polymorphism (RFLP), Diversity Array Technology (DArT) (Jaccoud et al., 2001) among the many available molecular markers. In Napier grass, isozymes were used to classify accessions from India (Bhandari et al., 2006), and RAPD have been used to examine napier grass germplasm accessions in the International Livestock Research Institute (Lowe et al., 2003).

1.2 Statement of the problem

Napier grass (*Pennisetum purpureum* Schumach) is the major feed for the cut-and-carry zero and semi-zero grazing dairy systems in the region where it constitutes between 40 to 80% of the forage for the smallholder dairy farms (Staal et al., 1997).

Despite this, napier stunt disease, caused by phytoplasma 16 sr XI has in the recent years caused up to 90 % reduction in forage yield (Lusweti et al., 2004), and is currently the biggest forage disease threat to the dairy industry in the region.

Stunting disease remains a challenge to scientists since identification of resistant clones has not been forthcoming. International Centre for Insect Physiology and Ecology (ICIPE) is in the process of identifying candidate accessions through diagnostic procedures of nested Polymerase Chain Reaction (PCR) and Loop Mediated Isothermal Amplification (LAMP) and through artificial challenge in the screen house where the promising clones are challenged with the vector *Maiestas banda* (Obura et al., 2011).

Management strategies for NSD are among many ways that can be used to curb the disease though not sustainable (Obura et al., 2010). These management strategies need to be supplemented with production of resistant/tolerant high yielding clones adaptable to this region for sustainability. The convenient and faster approach is to develop a wide and variable population for the candidate genes to be used for breeding resistant materials and subsequent supply to farmers in the affected areas. To initiate this, there is need to have information on the available germplasm in view of identification of clones that form the base population for breeding resistant clones. The initial step of breeding for resistance is to create variability within and without the germplasm so as to widen the base population of the candidate materials/clones to select from.

Variability can be intrinsic or due to different effects e.g. random genetic drift, mutation, natural selection, gene flow and transfer. This variability can be determined by assessing the level of polymorphism /relatedness between accessions and within each accession so as to know whether and where to create variability in the population to be used for clone development.

In order to determine variability of the available germplasm, there is need to do diversity study based on stable and informative molecular markers. Unfortunately, there is limited data to date on genetic diversity of napier grass in Western Kenya in relation to napier stunting disease candidate materials have been conducted using both morphological and SSR markers. There is need to have this information in order to know the relationship among the candidate clones and as pedigree information for breeding for resistant clones to napier stunting disease.

1.3 Justification

Napier grass is an open-pollinated species with low seed production. Its propagation is mostly done vegetatively using root splits and stem cuttings. Owing to its ability to undergo crossing, the genetic diversity within this species is high. Napier grass in Kenya is important especially to the dairy industry where up to eighty percent of farmers are small scale and depend almost entirely on this fodder due to unavailability of alternative fodder crops and expensive commercial feeds. Identification of the best clone(s) in terms of productivity and livestock preference has been a major problem to the farming community. This compounded by napier stunting disease (NSD) caused by phytoplasma pose a serious threat to the existence of napier grass and major food crops. Solutions to NSD is to breed resistant clones and make these clones available to farmers.

Germplasm exchanges of napier grass within and outside Kenya have been made extensively with no proper pedigree records. Germplasm collections must be characterized to maintain identity and purity for conservation and for use in breeding. Pedigree and collection information of many of napier grass accessions used in napier stunt disease breeding is unknown. Moreover, there is a tendency for a single cultivar to receive different names in different regions of collection.

Characterization increases breeding efficiency by providing important information of the breeding population being used. Through characterization of the available napier grass germplasm, this study envisaged to provide information on the level of diversity and population structure of napier grass collection used in the study of NSD and development of resistant clones. Information on characterization in this study will aid in identification

and introgression of resistant genes into high yielding susceptible clones through Marker Assisted Selection (MAS) and conventional breeding. This will then lead to increase in milk production by development of napier grass clones resistant to phytoplasma through provision of vital information on relatedness of the available napier clones.

1.4 Objective

1.4.1 General objective

To contribute to improved napier grass production in Kenya through clonal characterization.

1.4.2 Specific objectives

(i) To determine genetic diversity of napier grass collections in western Kenya using morphological and molecular markers.

(ii) To determine the population structure (deviation from Hardy-Weinberg proportions) of napier grass based on Simple Sequence Repeats molecular profile.

1.5 Hypotheses

1. There is no diversity within and between accessions of Western Kenya napier grass germplasm.
2. Napier grass germplasm in Western Kenya does not have a clear cut classification and can just form a few groups hence limited population structure.

CHAPTER TWO

LITERATURE REVIEW

2.1 Napier grass (*Pennisetum purpureum* Schumach)

Pennisetum is a genus in the grass family Poaceae, subfamily Panicoideae, and tribe Paniceae. Most members of the genus *Pennisetum* are widely adapted making the genus one of the most diverse in the *Poaceae* family (Jauhar et al., 1988). The genus *Pennisetum* is distributed throughout the tropical and subtropical regions of the world and consists of both tropical and subtropical species (Jauhar, 1981).

Pennisetum consists of more than 140 species, and it is considered one of the most important genera in the *Paniceae* tribe (Jauhar, 1981). Species in this genus are heterogeneous and characterized by genetic, physiological and reproductive variation (Augustin et al.,1993). The species and sub species in this genus have variation in ploidy levels ranging from diploids to octoploids, life cycles which range from annual to perennial (Chen et al., 2007). The modes of reproduction can either be sexual or apomictic while basic chromosome numbers range from five to nine (x = 5,7,8,9) (Shenoy et al.,1992). This genus is also characterized by the existence of three gene pools (Rao, 1929).

The genus *Pennisetum*, includes pearl millet (*P.glaucum)*, fountain grass (*P.setaceum)*, kikuyu grass (*P.clandestinum*) and napier grass (*P.purpureum*) (Boonman, 1997) among others. *P. glaucum* (L.) R. Br. was domesticated as the cereal pearl millet while *P. purpureum* Schumach became widely distributed as a tropical forage grass (Brunken ,1977).

The name elephant grass is derived from napier grass being a favorite food for elephants (Burton., 1989). On the other hand, napier grass was named after colonel Napier of Bulawayo in Zimbabwe who urged Rhodesia's (now Zimbabwe) department of agriculture to explore the possibility of using it for commercial livestock production (Orodho, 2006). Napier grass is an allotetraploid ($2n=4x=28$) and has a genome formula A'A'BB where A'A' is homologous to the AA genome of pearl millet ($2n=2x=14$) (*Pennisetum glaucum*) while BB is homologous to napier grass genome (Xie et al., 2005). Pearl millet and napier grass interbreed leading to a hybrid (2n =3x= 21) with genome formula AA'B. In this hybrid, the AA'B genome has greater similarity to the elephant grass type due to the larger genetic contribution (66.7% chromosomes) and dominance of the elephant grass B genome over the pearl millet A genome (Gonzalez and Hanna, 1984).

Napier grass was introduced into the United States by the United States Department of Agriculture in 1913 (Thompson, 1919). After its introduction, test plots and improvement programs were established. However, labor requirement to establish napier grass vegetatively and susceptibility to the eyespot disease (*Helminthosporium sacchari* Butler), caused farmers to loose interest in the crop (Burton, 1944). Overall, napier grass poses desirable characteristics such as insect resistance, rapid growth rate, and a high nutrient value (Rao et al., 2003). It is a hardy species that tolerates frost, fungi, high and low pH, monsoon, savanna, sewage sludge, virus, weeds and water lodging (Duke, 1978).

When Merkeron variety was registered by Burton in 1984, several specific genes of interest were noted (Burton., 1989). The two most valuable traits were; resistance to the eyespot disease and a dwarfing gene that resulted in the release of 'Mott' (Burton, 1989). Napier

grass was documented to produce as much as 100 tons of fresh weight per ha of biomass and has one of the fastest growth rates among higher plants (Renard et al., 2011). Karlsson and Vasil (1986) reported that napier grass cell cultures were the fastest growing of all the C_4 species tested. Napier grass can form hybrid in combination with pearl millet. The combination of napier grass' perennial growth habit and biotic stress tolerance with pearl millet's abiotic stress tolerance and integrated agronomic systems leads to development of large-seeded hybrids having superior perennial biomass production and forage potential.

Napier grass is a perennial C4 monocot grass native to Eastern and Central Africa but has been introduced to most tropical and sub- tropical countries (Boonman, 1993). Napier grass resembles sugarcane and mature plants normally grow in dense bamboo-like clumps with large flat leaves that may be 30-90 cm long and up to 3 cm broad and reach up to 4m in height up to 20 nodes (Hanna and Monson, 1988). The grass grows well at sea level and up to 2000 m and grows best in high-rainfall areas (in excess of 1500 mm/year), but its deep root system allows it to survive in drought times (Heng et al., 2012).

Napier grass is a poor breeding grass due to low viability of pollen grains and seed yields are usually very low, rarely more than 1-2 kg/ha pure germinating seed thus it is established vegetatively from stem cuttings or root splits.

Napier grass is highly heterozygous giving rise to a very heterogeneous population of seedlings, which are not morphologically uniform. Due to low stability and viability of pollen grains and seeds, little research has been done to enhance seed production as a planting material. Napier grass acts as a valuable source of genetic variation for pearl millet due to its high morphological variability and cross pollination though the hybrids

formed (2*n*=*3x*=21) are sterile. There are several varieties of napier grass developed through mass selection and hybridization techniques (Martel et al.,1997). Uganda hairless was developed in Uganda by A.S. Thomas (Tiley, 1959). Cameroon and Gold Coast varieties were developed in South Africa from the seed of West African origin (Kennan, 1952). Clone 13 was developed from French Cameroon (Wijk, 1977) while both Kakamega 1 and Kakamega 2 were developed from ILRI accessions. The above clones were developed through mass selection hence they also have the genome formula (2*n*=4*x*=28). Bajra napier hybrid is a cross between napier grass and bulrush millet while dwarf napier grass variety 'Mott' was bred at a research station in Gainesville, Florida (Hanna & Monson, 1988). These two have a genome formula (2*n*=3x=21).

2.2 Importance of Napier grass

Napier grass is a major livestock feed in cut-and-carry zero grazing systems in the Central and Western Kenya and has been shown to constitute between 40% to 80% of all forage for the smallholder dairy farms (Mwendia et al., 2006). In Kenya alone, about half a million smallholder dairy producers rely on napier grass as a major source livestock feed (Staal *et al.*, 1997). Research has shown that, one kilogram of fresh napier grass to contain 77.8g water, 1.0g protein, 0.5g fat, 17.6g total carbohydrate, 3.1g ash, 0.12% calcium and 0.07% phosphorus (Duke, 1983). Napier grass can yield 50-100 tones green matter per hectare if recommended agronomic practices are implemented. Among the fodder grasses, napier grass has been known to produce more dry matter per unit area when compared to other fodder crops such as guinea grass (*Panicum maximum*) and Rhodes grass (*Chloris gayana* Kunth)(Smith et al.,1989). Napier grass withstands repeated cutting and four to six cuts in a year can produce 50 to 150 tones fresh herbage

per ha. Its multiple uses include fire breaks, mulch, green manure, wind break, grazing, soil erosion control and as a constituent of fish ponds (Farrel et al.,2002). Napier rhizomes ramify soil particles and prevent runoff, its deep root system enables it to survive dry season(s).

It has been reported that napier grass is useful for phyto-remediation of petroleum-hydrocarbon-contaminated agricultural soils (Ayotamuno et al., 2006). In central Africa, the leaves are used as a source of a medicinal salt while in Spanish Guinea (Equitorial Guinea), the leaf and stalk infusion is used in anuria and oliguria (Watt and Breyer,, 1962). There are other species of *Pennisetum* which have also been identified to have medicinal values. These are; *setaceous*, *villoseum* and *divisum* (Sujatha et al., 1989).

In Kenya, napier grass is used in 'push-pull' technology for stem borer management (Khan et al., 2001). Valued for its high biomass napier grass is also perennial in nature, high leaf nutritive value, pest resistance and can equally be used for alcohol or methane production (Souza et al., 2005).

2.3 Napier stunting disease

Napier stunting disease has been present in Eastern Africa for about 30 years but it has become more noticeable in the last few years. Napier stunting disease is characterized by yellowing of leaves, thinning of leaves and stems, stunted growth, reduced biomass per unit area, necrosis and wilting of the entire plant (Mulaa et al., 2004). NSD is widely spread in Western Kenya resulting in economic loss within the small-scale dairy industry. NSD is highly expressed during re-growth upon cutting or grazing and all varieties of napier grass are susceptible to the disease where the infected foliage turn yellow green in

colour and become stunted eventually leading to death of the infected plants. In the year 2004, the disease was estimated to have affected over 23,298 km^2 of napier grass crop, an estimated 2 million households (about 30% of the population) in Western and Rift Valley provinces of Kenya (Mulaa et al.,2004). Napier cultivars such as Kakamega 1 and 2, French Cameroon and Bana are currently susceptible to NSD and the disease is spreading rapidly to other parts of Kenya (Mulaa et al., 2004).

Tiley, (1969) reported a stunting disease of napier grass in Uganda and the cause was suspected to be a virus transmitted by insects. Jones, (2004) identified the cause of napier stunt disease in Kenya and Uganda to be a phytoplasma. Analysis of rDNA sequences (GeneBank accession numbers AY377874-AY37787) identified these phytoplasmas as members of the 16SrXI (*Candidatus* Phytoplasma *oryzae*) or Rice yellow dwarf Phytoplasma. Two years later, 16S rDNA of Phytoplasma in Ethiopia was sequenced and its phylogeny determined by blast analysis (Jones *et al.* 2006). The nucleotide sequence deposited in GenBank (Acc. No. DQ305977) showed these phytoplasmas to be similar to Sugarcane yellow leaf phytoplasma (Accession No. AF056095), a member of the 16SrIII (*Ca.* P. *pruni*) (Jones et al., 2006).

In general, napier stunt phytoplasma of the groups16SrI, 16SrXI and 16SrIII have been observed to significantly hinder the development of napier grass (*Pennisetum purpureum*) farming in East Africa (Obura et al., 2011). Bermuda grass white leaf (BGWL) disease caused by the phytoplasma group (16SrXIV) and characterized by phyllody (leaf-like petals and sepals), whitening of the leaves and shortening of the stolons has been reported in the western regions of Kenya (Obura et al., 2010). There is a probability that the disease could spread to other crop species apart from the forages. Napier stunting disease

is spreading very fast in the Eastern and Central Africa countries especially Kenya, Uganda, Tanzania, and Rwanda due to various factors that include; lack of stringent regulations on country to country vegetative plant material transfer, borderline farmers easily transferring plant materials and farm tools from one farmer to another.

2.4 Genetic diversity in Napier grass

Techniques for assessing genetic diversity within plant species using DNA fingerprinting are now well established and can be applied quickly and easily (Weising et al., 1995). The use of molecular markers for diversity analysis can also serve as a tool to discriminate between closely related individuals from different breeding sources (Lombard et al., 2000) and may help to eliminate redundancy in phenotype base germplasm collections.

Some studies using molecular markers have been carried out in napier grass to improve on the understanding about diversity and kinship. Some molecular markers that have been used in napier grass diversity studies include; Isozymes (Bhandari et al., 2006), RFLP (restriction fragment length polymorphism; Smith et al., 1993), RAPD (random amplified polymorphic DNA; Smith et al., 1993; Daher et al., 2002; Lowe et al., 2003; Passos et al., 2005; Pereira et al., 2008; Babu et al., 2009), AFLP (amplified fragment length polymorphism; Harris et al., 2009), ISSR (inter-simple sequence repeats; Babu et al., 2009) and Sequence Related Amplified Polymorphism (SRAP) (Xie et al., 2009). Smith et al. (1989) verified the existence of polymorphism in different populations of elephant grass using RFLP. RFLP markers were also utilized to show genetic uniformity in plants derived from somatic embryos (Shenoy, 1992). Use of various probes has made

it possible to detect unwanted mixtures and off types in breeding stocks and advanced lines previously thought to be pure (Smith et al., 1993). RAPD have been used to characterize ILRI's napier grass germplasm that also included clones from East and Central Africa (Lowe et al. 2003).

Harris (2009) used AFLP to evaluate diversity of napier grass. In this study, the genetic relationships between the accessions in the Tifton napier grass nursery were evaluated using AFLPs and morphological traits. Wanjala (2013) used AFLP to study diversity of different collections of napier grass in Kenya and East Africa as compared to clones found in the ILRI germplasm and also clones from the USA. ISSR have been used to characterize napier grass from different parts of North America (Lima et al 2011). SRAP has also been used to assess genetic variability and relationship between MT-1 elephant grass and closely related cultivars in China (Xie Min et al., 2009). SSR markers have been used for cross species amplification using primers from *Pennisetum glaucum* (pearl millet), in the study, the investigators used clones from Brazil's EMBRAPA dairy cattle research center (Azevedo et al, 2012).

2.5 Molecular markers in diversity studies

A molecular marker is a piece of DNA with known position on the genome. It is also defined as fragment of DNA that is associated with a certain location within the genome. (Matsuoka et al.,2002). Molecular markers can be classified into type I and type II markers. Type I markers are associated with genes of known function, while type II markers are associated with anonymous genomic regions (O'Brien, 1991).Molecular markers are DNA sequence variants that can readily be detected and whose inheritance can be monitored (Newbury and Ford-Lloyd, 1999). Molecular marker technology can

facilitate the precise determination of the number, individual chromosomal location and interactive effects of genes that control traits (Peleman and van der Voort, 2003). Markers together with the genes they mark are within the proximity of each other or can be present on the same chromosome and tend to stay together generation after generation (Cheng et al., 1994). Molecular markers do not have any biological effect, and are transmitted by the standard laws of inheritance from one generation to the next.

Different molecular marker systems have been developed. These include; Restriction Fragment Length Polymorphisms (RFLPs)(Bolstein et al., 1980), Random Amplified Polymorphic DNAs (RAPDs)(Roy, 1992), Inter simple sequence repeats (ISSR)(Godwin et al., 1997, Zietkiewicz et al., 1994), Sequence Tagged Sites (STS)(Olson et al., (1989), Sequence Related Amplified Polymorphism (SRAP) (Li, 2001), Amplified Fragment Length Polymorphisms (AFLPs)(Vos et al., 1995), Simple Sequence Repeats (SSRs) (Litt et al., 1989; Jacob et al., 1991; Edwards et al., 1991). Diversity Array Technology (DART) (Jaccoud et al., 2001) and Single Nucleotide Polymorphisms (SNPs) (Chee et al., 1996) among others. Each marker system has its own merits and demerits according to its applicability, crop species and whether it is PCR based or hybridization based. The most widely used markers in major crops are simple sequence repeats (SSRs) or microsatellites (Gurta and Varshney, 2000).

A summary of some of the commonly used molecular markers and their features is shown in table 2.1.

Table 2.1 Characteristics of commonly used molecular markers

FEATURE	RFLPs	RAPDs	AFLPs	SSRs	SNPs
DNA required (µg)	10	0.02	0.5-1.0	0.05	0.05
DNA quality	High	High	Moderate	Moderate	High
PCR-based	No	Yes	Yes	Yes	Yes
Number of polymorph loci analyzed	1.0-3.0	1.5-50	200-100	1.0-3.0	1.0
Ease of use	Not easy	Easy	Easy	Easy	Easy
Automation	Low	Moderate	Moderate	High	High
Reproducibility	High	Low	High	High	High
Development cost	Low	Low	Moderate	High	High
Cost per analysis	High	Low	Moderate	Low	Low
Dominance	Co-dominant	Dominant	Dominant	Co-dominant	Co-dominant
DNA sequence	Not required	Not required	Not required	Required	Required
Radioactive detection	Yes	No	Yes	No	No
Gel System	Agarose	Agarose	Acrylamide	Acrylamide/Agarose	Sequencing Required

Adapted from: Semagn et al., 2006

2.5.1 Simple sequence repeats (SSR)

Simple Sequence Repeat (SSR) markers, also known as microsatellites, or short tandem repeats are tandemly repeated motifs of 1-6 nucleotides found in all prokaryotic and eukaryotic genomes (Zane et al., 2002). SSRs have been widely used by human geneticists for parentage testing, forensic identification, and medical diagnostics (Edwards et al., 1992; Alford et al., 1994). SSRs are present in both coding and non-coding regions (Hancock, 1995) and are usually characterized by a high degree of length polymorphism (Zane et al., 2002). The use of microsatellite loci are limited by inherent unstable mutation rates, a phenomenon that is reported to be caused by DNA polymerase slippage and/or unequal recombination (Li et al., 2002).

SSRs play a significant role as molecular markers for evolutionary and population genetic studies due to their high mutability (Yu et al., 2000), Microsatellites offer several advantages compared to other molecular markers; First ,they are highly reproducible, have high polymorphic information content (PIC), co dominant inheritance, locus specific, extensive genome coverage, simplicity in interpretation of results, PCR-based, readily portable within a species (Edwards et al., 1996) and requirement of only a small amount of template DNA (~5 ng/reaction). In a study comparing SSRs, RAPDs and AFLPs for the genetic analysis of yeast *(Saccharomyces cerevisiae)* strains, Gallego et al. (2005) reported that SSR analysis gave the highest level of information content. Microsatellites have also attracted scientific attention because they have been shown to be part of or linked to some genes of agronomic interest (Yu et al., 2000). In plants, SSRs have wide applications including construction of linkage maps in *Arabidopsis thaliana*

(Bell and Ecker, 1994), soybean (Akkaya et al., 1995), and maize (Senior et al., 1996) among other crops.

SSRs have also been successfully utilized to assess genetic diversity in barley (Saghai-Maroof et al., 1994), wheat (Plaschke et al., 1995), rice (Xiao et al., 1996) and sugar cane (Corderio et al., 1999, 2000). SSR markers are frequently used in DNA fingerprinting in many species including peach (Downey and Iezzoni, 2000; Testonlin et al., 2000), apricot (Hormaza, 2002), grapes (Sanchez-Eschribano et al., 1999). These studies showed that the genotyping of homozygous and heterozygous genotypes using SSRs is readily accomplished and that SSRs are inherited in a Mendelian fashion (Hormaza, 2002). Besides being excellent molecular markers for genetic mapping, microsatellite markers are very useful for population genetics, variety identification and protection, monitoring of seed purity and hybrid quality, gene tagging, germplasm evaluation and phylogenetic studies, studies of kinship, conservation genetics and forensics (Li et al., 2002). SSR markers are however costly and the discovery of the repeat motifs is a complex process that requires expensive equipment and expertise.

2.5.2 Morphological characterization of Napier grass

Morphology or phenotype is the observable trait of a given organism. Phenotypic identification of an organism is governed or affected by both the genotypic and environmental factors (Van de wouw et al., 1999). Classical plant breeding studies relied only on the morphological and agronomical traits which were extensively utilized in the identification of landraces, inbred lines and hybrids developed at that time. The use of morphological characters is cost-effective when compared to biochemical and molecular

markers for preliminary characterization of many individuals. Morphological traits can be used to identify similarity and differences among cultivars (Martinez, 2003).

Morphological characterization can be used to distinguish different clones and is also a powerful tool used by plant breeders to incorporate desired traits from one plant (donor) to another plant (recipient). Morphological characterization of napier grass is important for easy identification of clones of farmers' preference. Fourteen traits are commonly used to characterize napier germplasm. These were initially developed by Van de Wouw (Van de Wouw et al., 1999) but modifications are always made for different study objectives. These traits involve leaf characteristics, stem characteristics and growth form. Leaf morphology has an important role in identifying taxa in which variation in floral structures is uninformative or in which flower specimens are infrequent owing to a limited flowering season (Meade, 2003). Leaf traits in napier grass are the most important both to the farmer and the livestock due to handling and feeding preference of the livestock. Leaf traits used are, leaf length, leaf width, leaf seratteness, leaf hairiness adaxial and abaxial, length of hairs adaxial and abaxial, leaf sheath hairiness, length of sheath hairs and length of sheath edge hairiness. Stem characteristics used include stem thickness, internode length, node hairiness, length of node hairs. Growth habit includes growth form, rhizome number and rhizome length.

2.5.3 Molecular characterization of Napier grass

A number of different molecular marker systems are available for identifying genetic differences between and within napier grass populations. The choice of a marker system for any one specific use will depend upon the nature of the questions being addressed.

These marker systems differ in the way they sample within the genome and in the type of data that they generate. A number of markers have been used in napier grass and these include RAPD (Random Amplified Polymorphic DNA) (Lowe et al., 2003), SSR (Simple Sequence Repeats) (Azevedo et al., 2012), ISSR (Inter simple sequence repeats) (Babu et al., 2009), RFLP (Restriction Fragment Length Polymorphism)(Smith et al., 1989), AFLP (Amplified Fragment Length Polymorphism) (Wanjala et al., 2013) and SRAP (sequence related amplified polymorphism) (Xie et al., 2009).

CHAPTER THREE

MATERIALS AND METHODS

3.1. Study site

This study was located in Alupe, Busia county in Kenya. Busia is a county in the former western province of Kenya. It borders Kakamega to the east, Bungoma and Teso to the north, Uganda to the west, and Lake Victoria to the south. The average temperature is 22°C and the rainfall amount ranges between 750mm and 1,800mm per annum. The main economic activity is trade with neighboring Uganda, with Busia town - the county headquarters - being a cross-border centre. The county economy is heavily reliant on fishing and agriculture, with cassava, millet, sweet potatoes, beans, rice and maize being the principal crops and a small percentage of livestock products. A survey carried out in the year 2009 on NSD incidence in this area confirmed that this county has a high prevalence of NSD as compared to other counties (Mulaa et al., 2009). Clones were collected on farmers' fields and planted in KARI Alupe for evaluation and for quarantine purposes.

3.2 Plant Materials

Collection of samples was done using guidance from district livestock extension officers and collection points was on farmers' fields using GPS reading to identify the exact location of the farm and the collection point in each farm. This information were to be used later to trace back the exact location of the plants to generate a data base and pedigree information of the accessions. Collection of the samples involved clean materials not affected by NSD in the farm and within the stool. For each sample, root

splits were collected for use as planting material. The accessions were planted in 2 liters pots in the screen houses at ICIPE Mbita point. Each was labeled according to the collection district, division, location, sub location, farm and exact GPS reading for collection point in the farm. These materials were also planted at KARI Alupe Sub centre in a quarantine site in plots measuring 1m by 1m replicated thrice in Randomized Complete Block Design (RCBD).

3.3 Sample harvesting

The first three apical leaves were harvested using a pair of sterilized scissors for each plant 2 weeks after the first cut back and collected in medical envelopes containing silica gel and stored in the deep freezer at -20 ^{0}C for longer storage. In this study, materials from the same county were considered as a population the assumption being that the napier grass clones could freely interbreed within the county taking into consideration the viability of napier grass pollen.

3.4 DNA isolation and quantification

From each napier plant, the youngest leaf was harvested using a sterilized pair of scissor and stored in a medical envelope containing silica gel for genomic DNA isolation. Genomic DNA was extracted following the cetyl-trimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987) (Appendix 1).

The quality and concentration of the DNA was determined using Nanodrop and run on 0.8% agarose gel at 100V for 30 minutes. Aliquots of 1 µl for each sample were loaded into the nanodrop (Applied Biosystems ND 1000) and OD260, OD280 read to determine purity. Samples with OD ratio beyond 2 were purified further by reprecipating in ethanol.

RNA contamination was detected using 0.8% agarose gel electrophoresis with gel red staining. DNA samples were treated with RNase A to remove contaminating RNA. DNA was then recovered by adding a tenth equivalent volume of 3M sodium acetate (pH 6.8) and 95% ethanol to the DNA containing solution. Ethanol was carefully discarded and another wash done with 70% ethanol and pellet dried at room temperature. The DNA pellet was then dissolved in low salt TE buffer and stored at -20 ^{0}C for further analysis.

3.4.1 Optimization of SSR conditions and primer selection

Optimization was done on the samples and on the PCR conditions. DNA was normalized to 20 ng/µl which was used in all subsequent reactions. The optimal concentrations of $MgCl_2$, dNTPs and Taq DNA polymerase were determined. Gradient PCR was used to determine primer annealing temperatures and the number of cycles needed in each step to sufficiently amplify the SSR loci. A total of eighty four primers were tested using the same reaction conditions but with varying annealing temperature. Among these twenty five were selected based on the level of amplification and ease in scoring as shown in table 3.1 below.

Table 3.1: Summary of primer information and annealing temperatures used in the study

Primer name	Forward sequence	Reverse sequence	Repeat motiff	Annealing temperature (^{O}C)
ctm_8	GCTGCATCGGAGATAGGGAA	CTCAGCAAGCACGCTGCTCT	(CT)8	53
ctm_10	GAGGCAAAAGTGGAAGACAG	TTGATTCCCGGTTCTATCGA	(CT)22	54
ctm_12	GTTGCAAGCAGGAGTAGATCGA	CGCTCTGTAGGTTGAACTCCTT	(CT)12	53
ctm_27	GTTGCAAGCAGGAGTAGATCGA	CGCTCTGTAGGTTGAACTCCTT	(CT)71	53
ctm_59	TCCTCGACATCCTCCA	GACACCTCGTAGCACTCC	(CT)11	54

Table 3.1 continuation

Primer name	Forward sequence	Reverse sequence	Repeat motiff	Annealing temperature (^{O}C)
pgird_5	CAACCCAACCCATTATACTTAG	GCAACTCTTGCCTTTCTTGG	(GA)7	56
pgird_46	GAACAATTGCTTCTGTAATATCT	GCCGACCAAGAACTTCATACA	(CTC)6	56
pgird_56	ATCACTCCTCGATCGGTCAC	ACCAGACACACGTGCCAGT	(TG)6	55
pgird_57	GGCCCCAAGTAACTTCCCTA	TCAAGCTAGGGCCAATGTT	(AG)7	53
pgird_13	CAGCAGCGAGAAGTTTAGCA	GCGTAGACGGCGTAGATT	(AGC)8	51
pgird_21	GCTATTGCCACTGCTTCACA	CCACCATGCAACAGCAATA	(ACC)8	53
pgird_25	CGGAGCTCCTATCATTCCAA	GCAAGCCACAAGCCTATCC	(GA)9	58
phi_227562	TGATAAAGCTCAGCCACAAGG	ATCTCGGCTACGGCCAGA	(ACC)4	56
psmp_2235	GCTTTTCTGCTTCTCCGTAGAC	CCCAACAATAGCCACCAATAAAGA	(TG)9	61
psmp_2248	TCTGTTTGTTTGGGTCAGGTCCC	CGAATACGTATGGAGAACTGCGCC	(TG)10	61
psmp_2255	CATCTAAACACAACCAATCGAAC	TGGCACTCTTAAATTGACGCAT	(TG)4	61
psmp_2266	CAAGGATGGCTGAAGGGCTATG	TTTCCAGCCCACACCAGTAATC	(GA)17	61
psmp_2267	GGAAGGCGTAGGGATCAATCTCAC	ATCCACCCGACGAAGGAAACGA	(GA)16	61
xcup_14	TACATCACAGCAGGGACAGG	CTGGAAAGCCGAGCAGTATG	(AG)10	54
xcup_53	GCAGGAGTATAGGCAGAGGC	CGACATGACAAGCTCAAACG	(TTTA)5	54
xcup_63	GTAAAGGGCAAGGCAACAAG	GCCCTACAAAATCTGCAAGC	(GGATGC)4	54
xipes_0093	GGGGAAGAGATAGGGTTGGT	CGTGGACCGATCAGAACAAC	(GAA)6	55
xipes_0191	GAAGAACCTCCAGCTTTCCC	TGATGATTAGACAGACCGAC	(GAC)7	54
xipes_0219	GGGGAAGAGATAGGGTTGGT	AGACGATCGGCATCGCTACG	(CGC)11	56
xtxp_278	GGGTTTCAACTCTAGCCTACCGAACTTCCT	ATGCCTCATCATGGTTCGTTTTGCTT	(TTG)12	50

3.4.2 Morphological characterization

Fourteen morphological characters were scored on the regrowth of each plot 8-9 weeks after cutting. Leaf and stem characteristics were observed on 10 plants per plot and all observations done on the leaf below the first completely unrolled leaf at the top of the plants. The characters were evaluated in each clone as indicated in table 3.2 below.

Table 3.2: Morphological characters assessed in 120 Napier grass clones at KARI Alupe in 2013

Character	**Definition**	**No. of plants observed**
Morphological characteristics		
Growth habit		
1. Growth form	Average angle of stem to the ground from 0^0 to 90^0	Full plot
2. Tiller number	Average number of tillers on the stool	Full plot
Leaf characteristics		
All observations were done on the third leaf below the first completely unrolled leaf at the top of the plants		
1. Leaf length	Length from ligule to tip of leaf (cm)	10 plants
2. Leaf width	Width of leaf at widest point (cm)	10 plants
3. Leaf serrateness	An estimate of the average number of teeth on 1cm of leaf edge at middle of the leaf; (1) <15, (2) 15-20, (3) >20	10 plants
4. Leaf hairiness – adaxial	An estimate of the average hairiness of the abaxial face of the leaf at the middle of the leaf; (1) none, (2) sparse, (3) dense	10 plants

Table 3.2 continuation

Character	Definition	No. of plants observed
Morphological characteristics		
5. Leaf roughness – abaxial	An estimate with the tip of the finger of the average roughness of the abaxial face of the leaf; (1) smooth (2) rough (3) very rough	10 plants
6. Leaf sheath hairiness	An estimate of the average hairiness of the leaf sheath (excluding the edge of the leaf sheath); (1) none, (2) sparse, (3) dense	10 plants
7. Length of the sheath hairs	An estimate of the average length of the hairs on the leaf sheath (excluding the edge of the leaf sheath); (1) <2mm, (2) 2-3 mm, (3) 3-4 mm, (4) >4mm	10 plants
8. Leaf sheath edge hairiness	An estimate of the average hairiness of the edge of the leaf sheath; (1) none, (2) sparse, (3) dense	10 plants
9. Leaf colour	Estimate of leaf colour; (1) dark green, (2) mid green, (3) pale green, (4) yellow, (5) pale yellow, (6) white/cream (7) red/orange	10 plants
Stem characteristics		10 plants
1. Stem thickness	Stem diameter above the lowest node (cm)	10 plants
2. Internode length	Length of the fifth internode from the lowest internode (cm)	10 plants
3. Node hairiness	An estimate of the hairiness of the lowest node; (1) none (2) sparse (3) dense	10 plants

3.4.3 Data scoring and analysis

Allele sizing and scoring was done using GeneMapper version 4.1 (Applied Biosystems Microsatellite mapping protocol 2005) for each marker used. This data was exported into excel and the efficiency of each marker based on Polymorphic Information Content (PIC) determined using Power Marker version 3.25. Markers with low PIC were not used in the final analysis. GenAlex Software version 6.5 (Peakall and Smouse, 2009) was used to calculate genetic diversity indices. Data from GenAlex were exported to DARwin version 5 (Perrier and Jacquemoud Collet, 2006) using CONVERT software for phylogenetic analysis and structuring of the population.

Morphological scores were divided into qualitative and quantitative traits and entered into XLSTAT 2014 for descriptive statistics and characterization.

Figure 3.1: A) Morphological characterization of napier grass at KARI Alupe and (B) DNA normalization at BeCA ILRI hub

(Source: Author, 2014)

CHAPTER FOUR

RESULTS

Microsatellite variability such as the observed and expected number of alleles at each locus for each population, mean number of alleles per locus, observed heterozygosity, expected heterozygosity, Phylogenetic analysis and multivariate analysis revealed substantial variation which was used to quantify the genetic variation within the population and among the populations.

4.1 Quantifying intrapopulation diversity

4.1.1 Test for deviation from Hardy – Weinberg equilibrium (HWE)

Each population's single locus deviation from HWE was determined. The assumption or H_o hypothesis was that clones or cultivars in each population were in HWE (observed = expected). In the Alupe population a total of 5 loci (which was equivalent to 20% of the total loci) out of 25 PGIRD 46, Phi 227562, PSMP 2248, PSMP 2266 and XCUP 63 were in HWE with X^2 values below the expected at 95% confidence level ($P<0.005$). In this case, 20 loci (80% of the total loci) deviated significantly from HWE.

In the Bungoma population, a total of 13 loci (52% of the total loci) out of 25 deviated significantly from HWE at confidence level ranging from 95% to 99.9 % ($P<0.005$) to ($P<0.001$). (CTM 8, CTM 10, CTM 12, CTM 27, PGIRD 46, PGIRD 56, Phi 227562, PSMP 2235, PSMP 2248, PSMP 2266, XCUP 14, XCUP 63 and XIPES 0219). Twelve loci deviated significantly from HWE.

The Busia population recorded a total of 16 loci (64% of the total loci) that significantly showed deviation from HWE at confidence level ranging from 95% to 99.9 %. (CTM 8, CTM 10, CTM 12, CTM 27, PGIRD 56, PGIRD13, Phi 227562, PSMP 2266, PSMP 2267, XCUP 14, XCUP 53, XCUP 63 and XIPES 0093).

The Mumias population recorded a total of 16 loci (64% of the total loci) that showed significant deviation from HWE. These were, CTM 8, CTM 10, CTM 12, CTM 27, PGIRD 46, PGIRD 57, PGIRD 13, PGIRD 25, Phi 227562, PSMP 2248, PSMP 2255, PSMP 2267, XCUP 14, XIPES 0093, XIPES 0191 and XIPES 0219.

The Butere population had a total of 10 loci (40% of the total loci) that significantly showed deviation from HWE at confidence level ranging from 95% to 99.9 %. (CTM 10, CTM 12, CTM 27, PGIRD 46, PGIRD 57, Phi 227562, PSMP 2248, PSMP 2266, PSMP 2267 and XCUP 14) Out of the 25 loci studied, 3 loci (12%) (PGIRD 46, Phi 227562, PSMP 2248) significantly deviated from HWE for all the populations. Three loci namely, CTM 59, PGIRD 21 and XTXP 278 also did not deviate significantly across all the sub populations (12%). The Alupe population had the lowest number of loci at 20% of the total loci studied that deviated significantly from the HWE while Mumias and Busia showed the highest number of loci that deviated significantly both at 64% of the total loci studied (Appendix 2).

4.1.2 Proportion of polymorphic loci (Polymorphic information content)

PIC is a measure of usefulness of a molecular marker. PIC is a ration between the polymorphic loci and the total number of loci and is useful in providing information on

the marker(s) that are most informative in discriminating among individuals in a population. PIC is calculated using allelic frequencies and depends on the number of alleles and allele frequency. A total of 25 markers each with an average of 4 alleles were evaluated for efficiency in Napier grass diversity studies (Table 4). PIC values ranged from 0.1697 to 0.6793 in XTXP 278 and CTM 10 respectively and this was correlated to the gene diversity indices of 0.7297 and 0.1753 respectively.

4.1.3. Mean Number of Alleles per locus

Mean number of alleles is a ration between the sum of all detected alleles in all loci and the total number of loci. A total of 97 alleles were observed in the entire population. The most polymorphic marker in the entire population was PGIRD 13 that had 5 alleles, while the least polymorphic marker was PGIRD 21 with two alleles. Allele frequencies across all loci ranged from 0% to 94.4% (Appendix 3). The mean number of alleles ranged from 2.00 (PGIRD 21) to 4.40 (PGIRD 13) with the grand mean of allele number being 3.26.

Table 4.1: Markers used in the study and their PIC values analyzed using Powermarker version 3.25

marker	Sample size	Gene Diversity	Heterozygosity	PIC
pgird46	96.0000	0.5870	0.9792	0.4981
ctm12	96.0000	0.5427	0.9101	0.4375
phi227562	96.0000	0.5698	0.9101	0.4814
ctm27	96.0000	0.5619	0.8737	0.4657
ctm8	96.0000	0.6232	0.8588	0.5476
xipes0093	96.0000	0.6186	0.8211	0.5441
xcup53	96.0000	0.6788	0.6707	0.6113
psmp2266	96.0000	0.6361	0.6588	0.5690
psmp2267	96.0000	0.6096	0.6508	0.5654
pgird25	96.0000	0.5603	0.6316	0.5042
ctm10	96.0000	0.7297	0.6042	0.6793
psmp2255	96.0000	0.4594	0.5833	0.4069
pgird13	96.0000	0.7058	0.5579	0.6598
xipes0191	96.0000	0.3806	0.5056	0.3123
pgird56	96.0000	0.5082	0.4835	0.4692
pgird5	96.0000	0.3877	0.4479	0.3239
ctm59	96.0000	0.3352	0.4043	0.2992
xipes0219	96.0000	0.5425	0.3934	0.4940
psmp2235	96.0000	0.4221	0.2813	0.3657
pgird21	96.0000	0.2679	0.2747	0.2320
psmp2248	96.0000	0.4958	0.2714	0.4548
xcup14	96.0000	0.5978	0.2394	0.5226
pgird57	96.0000	0.3249	0.2234	0.3078
xtxp278	96.0000	0.1753	0.1875	0.1697
xcup63	96.0000	0.2750	0.1310	0.2460

4.1.4 Private alleles

Private alleles are alleles that were unique in only one population. These were observed in all populations but at different loci. Busia and Alupe populations each had two private alleles namely XIPES 0093, PGIRD 5 and PGIRD 5, XCUP 63, respectively. The Bungoma population had a private allele at locus XCUP 63, Butere population at XIPES 0191 while Mumias population had at CTM 59 loci. The frequencies of the private alleles ranged from 0.019 to 0.077 with the lowest being allele 335 of CTM 59,in Mumias population while the highest was allele 324 of XCUP 63, in Busia population (Table 5). A total of six samples in the entire population had one or more private alleles as shown in table 4.2.

Table 4.2: Summary of Private Alleles by Population analyzed using GenAlex 6.5

Sample	Pop	Locus	Allele	Freq
12	**Alupe**	pgird_5	90	0.056
12	**Alupe**	xipes_0093	136	0.056
62	**Bungoma**	xcup_63	172	0.028
10	**Busia**	pgird_5	163	0.053
35	**Busia**	xcup_63	324	0.077
95	**Butere**	xipes_0191	267	0.026
94	**Mumias**	ctm_59	335	0.019

4.1.5 Effective number of alleles (Ne)

Effective number of alleles not only represents the number of alleles that can be present in a population but also determines the level of intrapopulation diversity. The higher the alleles are in a population, the higher the chances of having diversity in the alleles. On average, the Alupe population had the lowest effective number of alleles at 2.023, while the Mumias population had the highest value at 2.190 (Figure 2). Marker CTM 10 had the highest N_e value at 3.365 as compared to XTXP 278 with the lowest index of 1.206 as shown in figure 4.1 below.

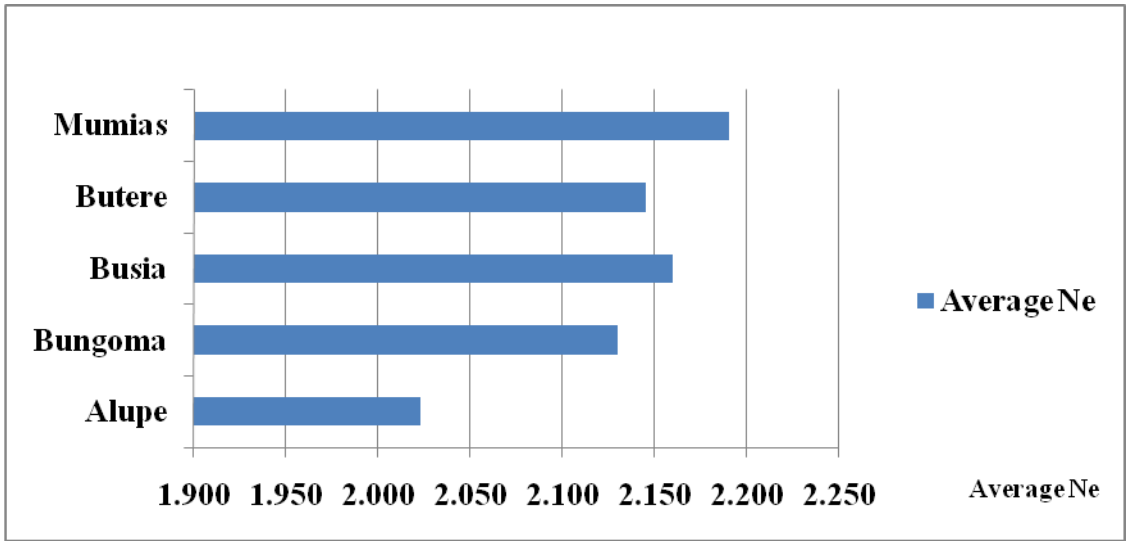

Figure 4.1: Average effective Number of alleles in five populations of napier grass from western Kenya.

4.1.6 Mean Expected Heterozygosity (He)

Mean expected heterozygosity is the probability that at a single locus, any two alleles, chosen at random from the population will be different from each other. The average H_e over all loci is an estimate of genetic variability in a population. A mean heterozygosity

deficit (difference between observed and expected) was observed in the Alupe population with the mean observed (H_o) at 0.509 as compared to the expected (H_e) at 0.511 resulting in a difference of 0.002 (Figure 4.2). The other populations had high heterozygosity ranging from 0.028 to 0.044 with the highest being in the Bungoma and Butere populations while the lowest was observed in the Mumias population. The grand mean heterozygosity observed was 0.539 as compared to an expected heterozygosity of 0.508 in all the populations observed. From these results, Alupe population had the lowest level of diversity (extent of genetic differentiation) as compared to the other populations. The total gene diversity in the 25 loci across the entire populations was also determined. Gene diversity varied from 0.170 to 0.724 in XTXP 278 and CTM 10 respectively which corresponded to the Mean expected heterozygosity for the loci with the highest being 0.693 for marker CTM 10 and 0.167 for XTXP 278.

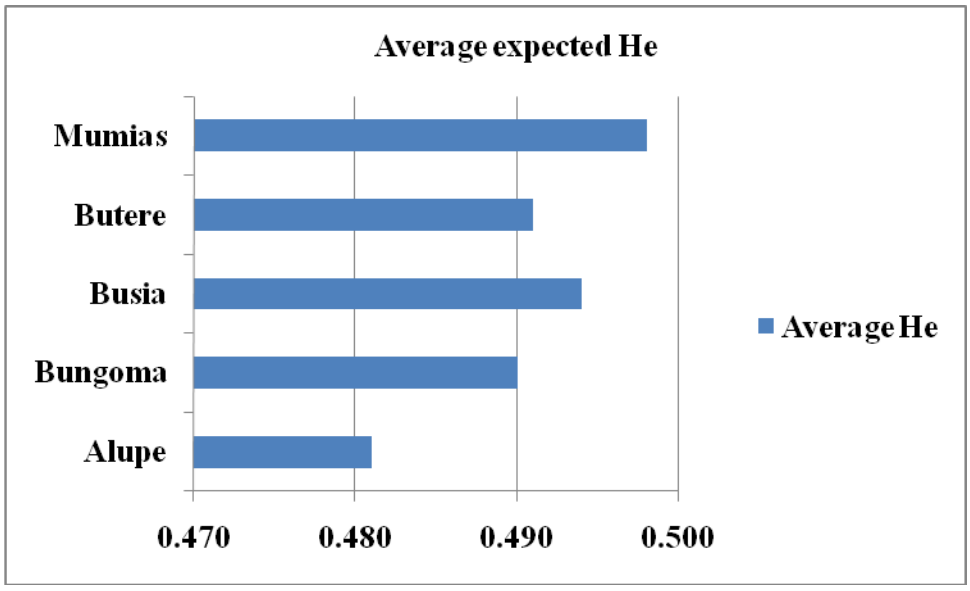

Figure 4.2: Average expected heterozygosity in 5 populations of napier grass from western Kenya

4.2 Distribution of genetic diversity among napier grass populations

4.2.1 Degree of gene differentiation among populations (F_{ST})

The highest F_{ST} value was observed at 0.065 at locus PGIRD 57 across all the populations (Table 4.3). This means that 6.5 % of the total genetic variation at this locus is distributed among the populations with 93.5% of the variation within populations. The lowest F_{ST} value was at 0.002 at locus Phi 227562. Mean F_{ST} value for all the populations was 0.026.

Table 4.3: F_{ST} values for the 25 loci used in Napier grass

Locus	Gene differentiation (F_{ST})	Gene differentiation (F_{IT})	Fixation index (F_{IS})
ctm_8	0.006	-0.388	-0.396
ctm_10	0.042	0.204	0.169
ctm_12	0.017	-0.666	-0.695
ctm_27	0.016	-0.574	-0.599
ctm_59	0.005	-0.206	-0.212
pgird_5	0.025	-0.158	-0.188
pgird_46	0.006	-0.686	-0.697
pgird_56	0.021	0.087	0.067
pgird_57	**0.065**	0.336	0.290
pgird_13	0.035	0.204	0.175
pgird_21	0.024	0.000	-0.025
pgird_25	0.013	-0.078	-0.091
phi_227562	**0.002**	-0.627	-0.631
psmp_2235	0.027	0.279	0.260
psmp_2248	0.025	0.470	0.456
psmp_2255	0.026	-0.299	-0.334
psmp_2266	0.033	-0.011	-0.045
psmp_2267	0.027	-0.068	-0.098
xcup_53	0.035	0.042	0.007
xcup_63	0.048	0.566	0.545
xipes_0093	0.018	-0.315	-0.338
xipes_0191	0.023	-0.338	-0.369
xipes_0219	0.048	0.269	0.232
xtxp_278	0.014	-0.068	-0.083
Mean	**0.026**	-0.057	-0.081
SE	**0.003**	0.075	0.074

4.2.2 Fixation Index (F_{IS})

The measure of deviation of genotypic frequencies from panmictic frequencies in terms of heterozygous deficiency or excess was also estimated using the formula;

$$F_{IS} = (H_{exp} - H_{obs}) / H_{exp}$$

Where; H_{exp} = expected heterozygosity while, H_{obs} = observed heterozygosity.

The PGIRD 57 locus in the Mumias population recorded the highest positive F_{IS} value at 0.764 while the highest negative F_{is} value observed was at -1 at CTM 27 locus in the Bungoma population (Table 4.4). The mean fixation values for Bungoma, Busia and Butere populations were the same (-0.110) while Alupe and Mumias populations had mean F_{is} values -0.052 and -0.048, respectively. The grand mean F_{is} value was -0.086.

Table 4.4: Fixation indices (FIS), highest positive and negative values of the 5 populations and their respective loci

Population	Highest positive F_{IS} Value	Locus	Highest negative F_{IS} value	Locus	F_{is} ± SE
Alupe	0.723	xcup 63	-0.820	pgird 46	-0.0592±0.092
Bungoma	0.644	xcup 63	-1.000	ctm 27	-0.110±0.082
Busia	0.677	xcup 14	-0.810	ctm 12	-0.110±0.078
Butere	0.588	xcup 14	-0.774	pgird 46	-0.110±0.073
Mumias	0.764	pgird 57	-0.598	pgird 46	-0.048±0.076

4.2.3 Genetic differentiation between napier grass populations

The differences among napier grass populations for two measures of population differentiation namely, the G_{ST} (Nei, 1987) and F_{ST} was also determined. G_{ST} is defined as the proportion of genetic diversity that resides among populations or a measure of population differentiation for multiple alleles. It is equivalent to Wright's (1951) F_{ST} when there are only two alleles at a locus, and, in the case of multiple alleles, G_{ST} is equivalent to the weighted average of F_{ST} for all alleles (Nei, 1973). F_{ST} can also be defined as the reduction in heterozygosity due to population structure or the variance in allele frequencies among populations. Values of G_{ST} and F_{ST} ranged from zero to one, with low values indicating that little variation is proportioned among populations while high values denote that a large amount of variation is found among populations. G_{ST} ranged from -0.022 for locus psmp 2248 to 0.033 for locus pgird 57. F_{ST} ranged from 0.002 for locus phi 227562 to 0.065 for locus pgird 57. The mean values were 0.026 for F_{ST} and 0.000 for G_{ST} (Table 4.5).

Table 4.5: Measure of population differentiation in Napier grass populations

Locus	G_{ST}	F_{ST}
ctm_8	-0.013	0.006
ctm_10	0.014	0.042
ctm_12	0.008	0.017
ctm_27	0.006	0.016
ctm_59	-0.015	0.005
pgird_5	0.005	0.025
pgird_46	-0.002	0.006
pgird_56	-0.006	0.021
pgird_57	**0.033**	**0.065**
pgird_13	0.006	0.035
pgird_21	-0.001	0.024
pgird_25	-0.010	0.014
phi_227562	-0.008	**0.002**
psmp_2235	-0.005	0.027
psmp_2248	**-0.022**	0.025
psmp_2255	0.010	0.026
psmp_2266	0.006	0.033
psmp_2267	-0.006	0.027
xcup_14	-0.010	0.040
xcup_53	0.007	0.035
xcup_63	0.005	0.048
xipes_0093	0.001	0.018
xipes_0191	0.007	0.023
xipes_0219	0.001	0.048
xtxp_278	-0.009	0.014
Mean	**0.000**	**0.026**
SE	**0.002**	**0.003**

Key note: Standard errors SE were estimated by jackknifing over loci. The bolded values show the minimum and maximum values for each of the G_{ST} and F_{ST}.

4.3 Measuring genetic relationship between the populations

4.3.1 Nei's genetic distance (D_A)

This measure of genetic distance assumes that genetic differences arise due to mutations and genetic drift but this distance measure is known to give more reliable population trees than other distances particularly for microsatellite data. In this study, it was more meaningful than D_S. Nei's D_A index of genetic distance was calculated for the populations, with the lowest D_A observed between Mumias and Butere populations (0.017) while the highest distance was between Bungoma and Alupe populations (0.048) as shown in table 4.6 below.

Table 4.6: Pair wise population matrix of Nei's genetic distance

	Alupe	**Bungoma**	**Busia**	**Butere**	**Mumias**
Alupe	0.000				
Bungoma	0.048	0.000			
Busia	0.046	0.019	0.000		
Butere	0.033	0.035	0.020	0.000	
Mumias	0.046	0.035	0.023	0.017	0.000

4.3.2 Phylogenetic relationship among populations

The phylogenetic relationships between the 5 Napier grass populations based on average Nei's genetic distance (D_A) and hierarchical clustering yielded two distinct clusters which did not reflect the geographical origins of the cultivars. There was much overlap among cultivars spread across different clusters. The dendogram shows two Major groups and one minor group with no clear distinction within the groups (figure 4.3).

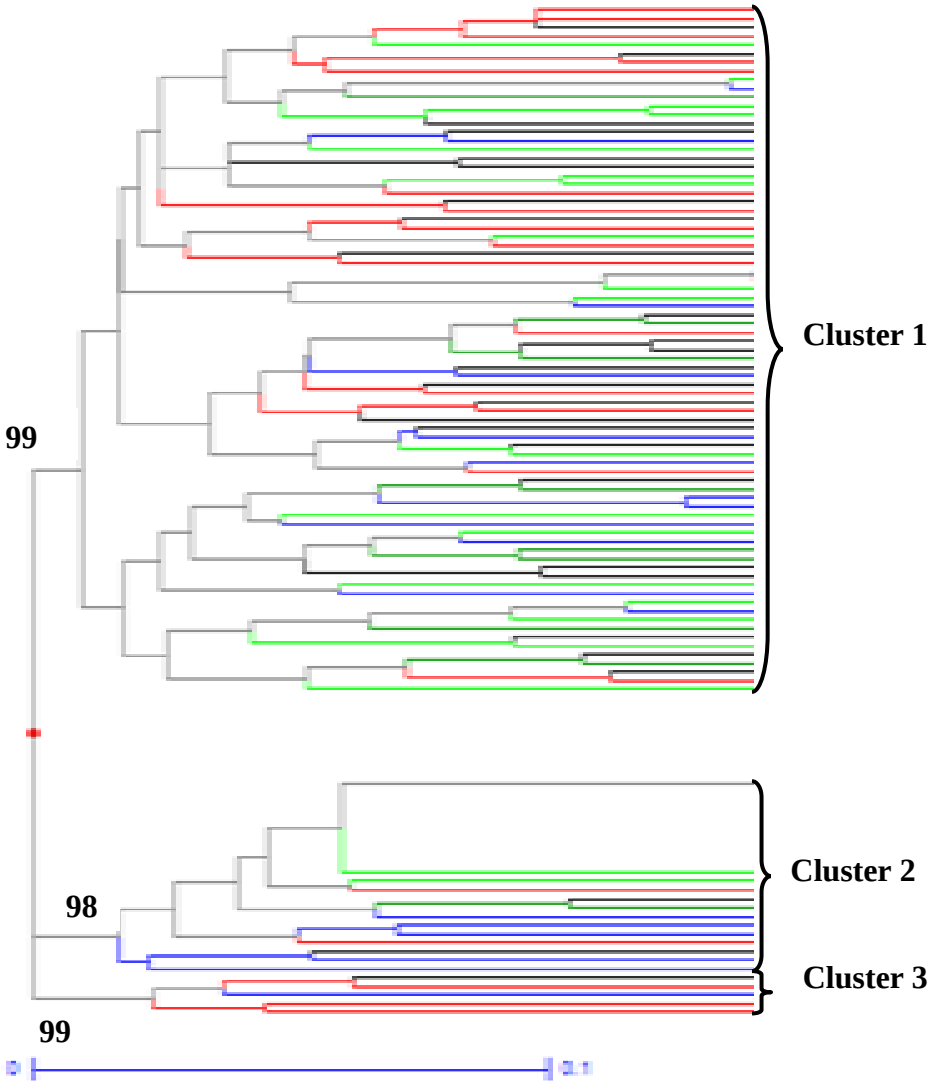

Figure 4.3: Dendogram from Neighbor joining (NJ) method showing genetic relationship among the 5 populations using Nei's genetic distance (D_A)

Key: Bungoma Butere Alupe Busia Mumias

4.3.3 Principal Coordinate Analysis (PCoA)

Principal Coordinate analysis for the allele frequency data for the 5 populations using 25 microsatellite markers was performed. The variation accounted for by PCo1, PCo 2 and PCo 3 were 34.75%, 16.36% and 14.01% respectively.

PCo1 separated samples 56 (Mumias), 51(Mumias), 13 (Mumias), 49 (Bungoma) and 8 (Bungoma) on extreme ends on one side and samples 58 (Busia), 28 (Busia), 10 (Busia) and 1 (Busia) on the other extreme end separating them from the two main groups (Figure 4.4).

PCo 2 separated samples 51 (Mumias), 8 (Bungoma), 13 (Mumias), 58 (Busia), 10 (Busia), 5 (Busia), 49 (Bungoma), 28 (Busia) and 1(Busia) on one extreme end and 56 (Mumias), 95 (Mumias) and 33 (Busia) on the other end separating them from the two main populations.

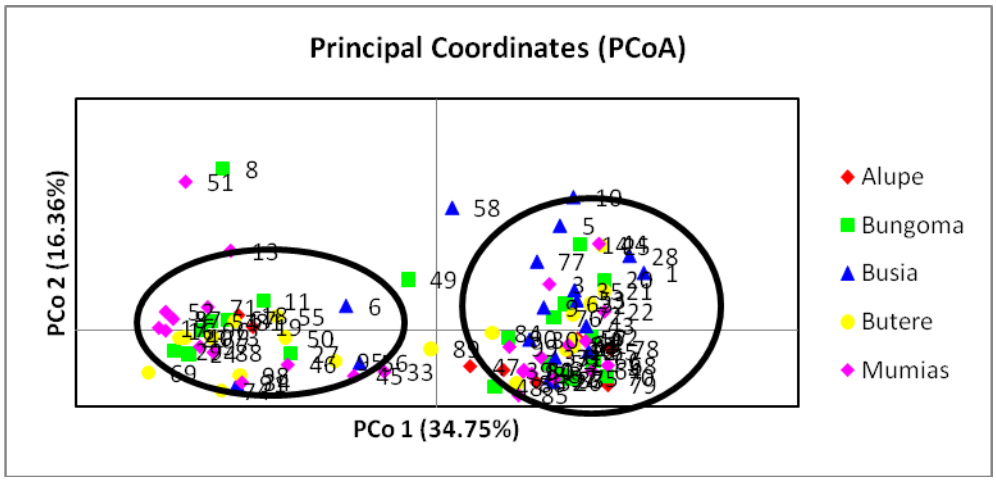

Figure 4.4: Principal Coordinate analysis showing variation of 5 napier populations explained by Coordinates 1 and 2

4.3.4 Analysis of Molecular variance (AMOVA)

In order to understand the partitioning levels of genetic diversity of the napier grass under study, an analysis of molecular variance (AMOVA) was conducted. The result of AMOVA revealed that 90% of the total genetic diversity existed within individuals, 9% among individuals and 1% among the populations as shown in table 4.7.

Table 4.7: Analysis of molecular variance (AMOVA) in 5 Napier grass populations using genotype data from 25 microsatellite loci

Source of variation	Degree of freedom	Sum of Squares	Variance components	Percentage variation
Among Populations	4	36.562	0.041	1%
Among Individuals	91	691.021	0.656	9%
Within Individuals	96	603.000	6.281	90%
Total	**191**	**1330.583**	**6.979**	**100%**

4.3.5 Morphological characterization of napier grass

Morphological traits were divided into quantitative and qualitative characteristics. Quantitative traits involved those traits that are controlled by many genes while qualitative involved those that are controlled by one or few genes. Morphological characters based on the leaf, stem and growth form were scored. A number of analyses were performed using XLSTAT statistical package in Excel.

4.3.6 Correlation among quantitative characters of napier clones

Pearson correlation analysis was carried out to determine how the quantitative characters were correlated with each other. There was a high positive correlation between leaf

length and leaf width (r =0.678), growth form and plant height (r = 0.637), internode length and plant height (r = 0.626), growth form and leaf length (0.539), growth form and internode length (r = 0.520).

On the contrary, there was a negative correlation between tiller number and stem thickness (r =-0.438), tiller number and leaf width (r = -0.368), tiller number and leaf length (r = -0.076), stem thickness and stool diameter (r = -0.062)(Table 4.8).

Table 4.8: Pearson correlation matrix for eight quantitative traits evaluated for five napier grass populations from western Kenya

Variables	LL	LW	SD	PH	ST	IL	GF	TN
LL	1							
LW	0.678	1						
SD	0.254	0.096	1					
PH	0.496	0.285	0.322	1				
ST	0.335	0.506	-0.062	0.062	1			
IL	0.368	0.105	0.215	0.626	0.074	1		
GF	0.539	0.226	0.148	0.637	0.078	0.520	1	
TN	0.076	-0.368	0.455	0.302	-0.438	0.242	0.159	1

Key: LL= Leaf length; LW= Leaf width; SD= Stool diameter; PH= Plant height; ST= Stool diameter; IL= Internode length; GF= Growth form; TN= Tiller number.

4.3.7 Principal component analysis of quantitative characters of napier clones

The first principal component accounted for 38.06% of the total variation while the second component accounted for 26.13% of the total variation; however there was no distinct grouping of the clones. This could not be correlated with the molecular partitioning (Fig 4.5).

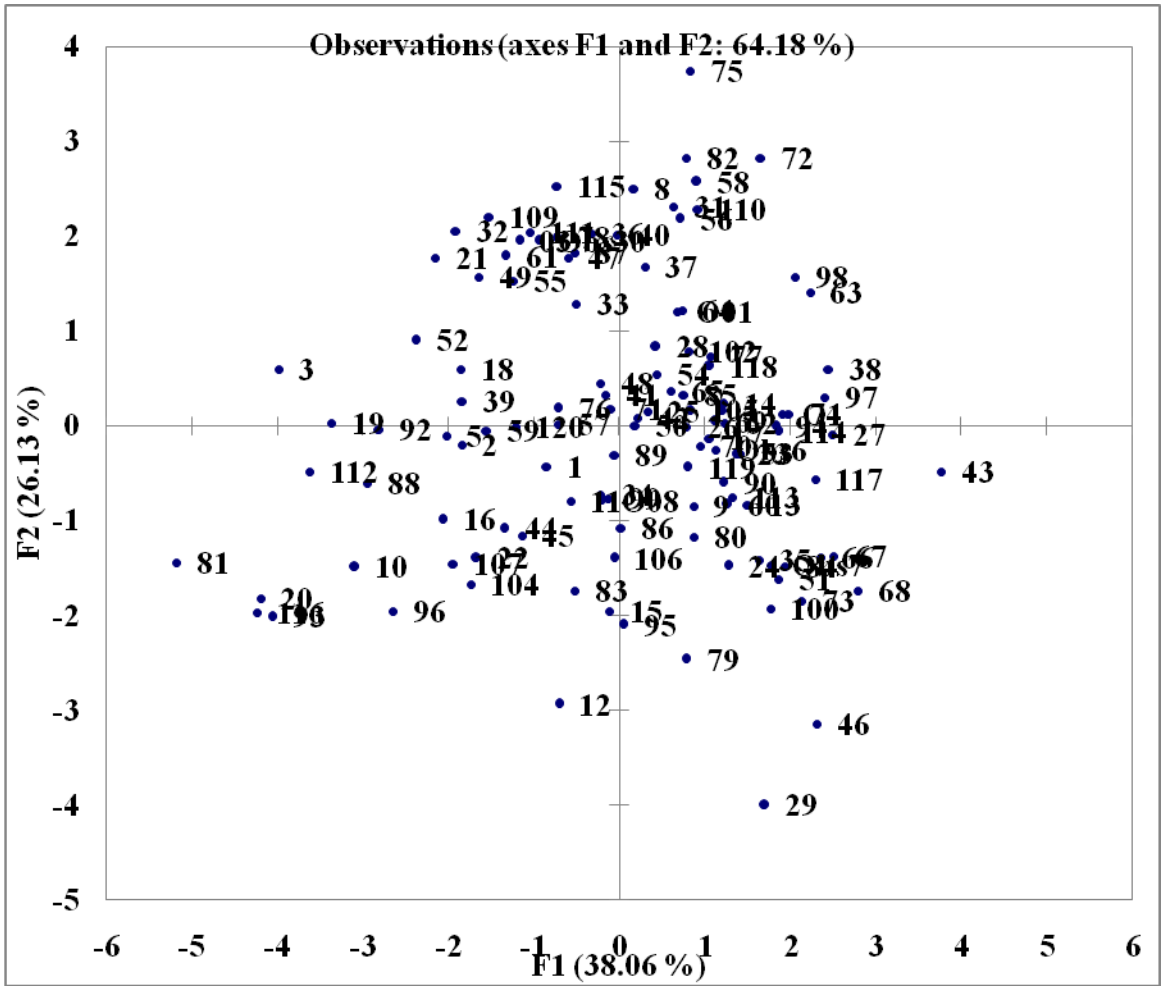

Figure 4.5: Two dimensional scatter plot showing the relationship between F1 and F2 using 8 Napier quantitative morphological characters

Key: Number represents a clone; F1= First principal component PC1, F2= Second principal component PC2

The first principal component accounted for 38.06% of the total variation while the second component accounted for 26.13% of the total variation. The first three quantitative characteristics or factors (leaf length, leaf width, stool diameter) contributed to 76% of the total variation. The other five factors had minimal contribution to the total variation (Fig 4.6).

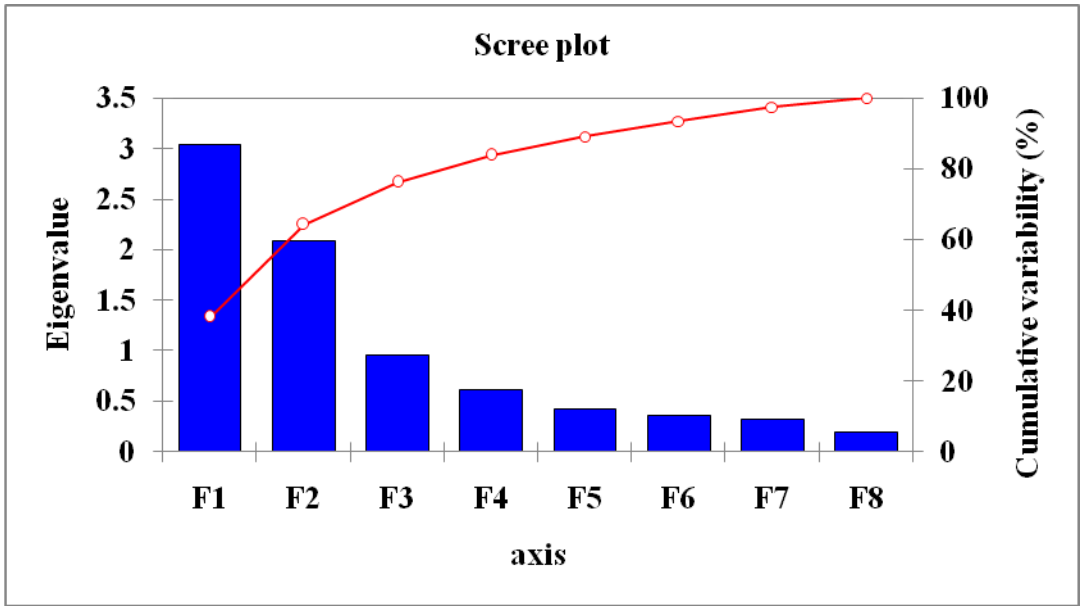

Figure 4.6: Scree plot showing variability contributed by each of the eight factors used

Key: F1= leaf length, F2= leaf width, F3= stool diameter, F4= plant height, F5=stem thickness, F6= internode length, F7= growth form, F8= tiller number

4.3.8 Correlation analysis between qualitative characters of napier clones

There was a significant positive correlation between length of sheath hairs and leaf sheath hairiness (r = 0.775, P≤ 0.05), abaxial and adaxial hairiness (r = 0.427), leaf sheath hairiness and abaxial (r = 0.332), leaf sheath hairiness and adaxial(r = 0.304), length of sheath hairs and leaf sheath edge hairiness (r = 0.283), length of sheath hairs and abaxial (r = 0.274), length of sheath hairs and adaxial (r = 0.265).

There was also a negative correlation between adaxial and leaf sheath edge hairiness (r = -0.282), leaf color and length of sheath hairs (r = -0.235), leaf color and leaf sheath hairiness (r = -0.211), leaf color and adaxial (r = -0.195) (Table 4.9)

Table 4.9: Correlation matrix for qualitative characters of napier clones

Variables	LC	NH	AD	AB	SR	LSEH	LSH	LESH
LC	1							
NH	0.046	1						
AD	0.195	-0.037	1					
AB	0.089	0.080	0.427	1				
SR	0.001	-0.001	0.147	0.164	1			
LSEH	0.026	0.076	0.282	0.165	0.019	1		
LSH	0.211	0.160	0.304	0.332	0.213	0.283	1	
LESH	0.235	0.189	0.265	0.274	0.146	0.202	0.775	1

Key: LC= Leaf color; NH= Node hairiness; AD= Adaxial hairiness; AB= Abaxial hairiness; SR= Seratteness; LSEH= Leaf sheath edge hairiness; LSH= Leaf sheath hairiness; LESH= Length of sheath hairs.

4.3.8.1 Principal component analysis of qualitative characters of napier clones

The first principal component accounted for 30.45% of the total variation seen while the second principal component accounted for 18.59% of the total variation seen. In total the variation observed using these two components was at 49.04% (Fig 4.7). The clones clustered into two major groups with a few exceptional clones (clones 8, 21, 10,112 and 32) which were in between these major groups.

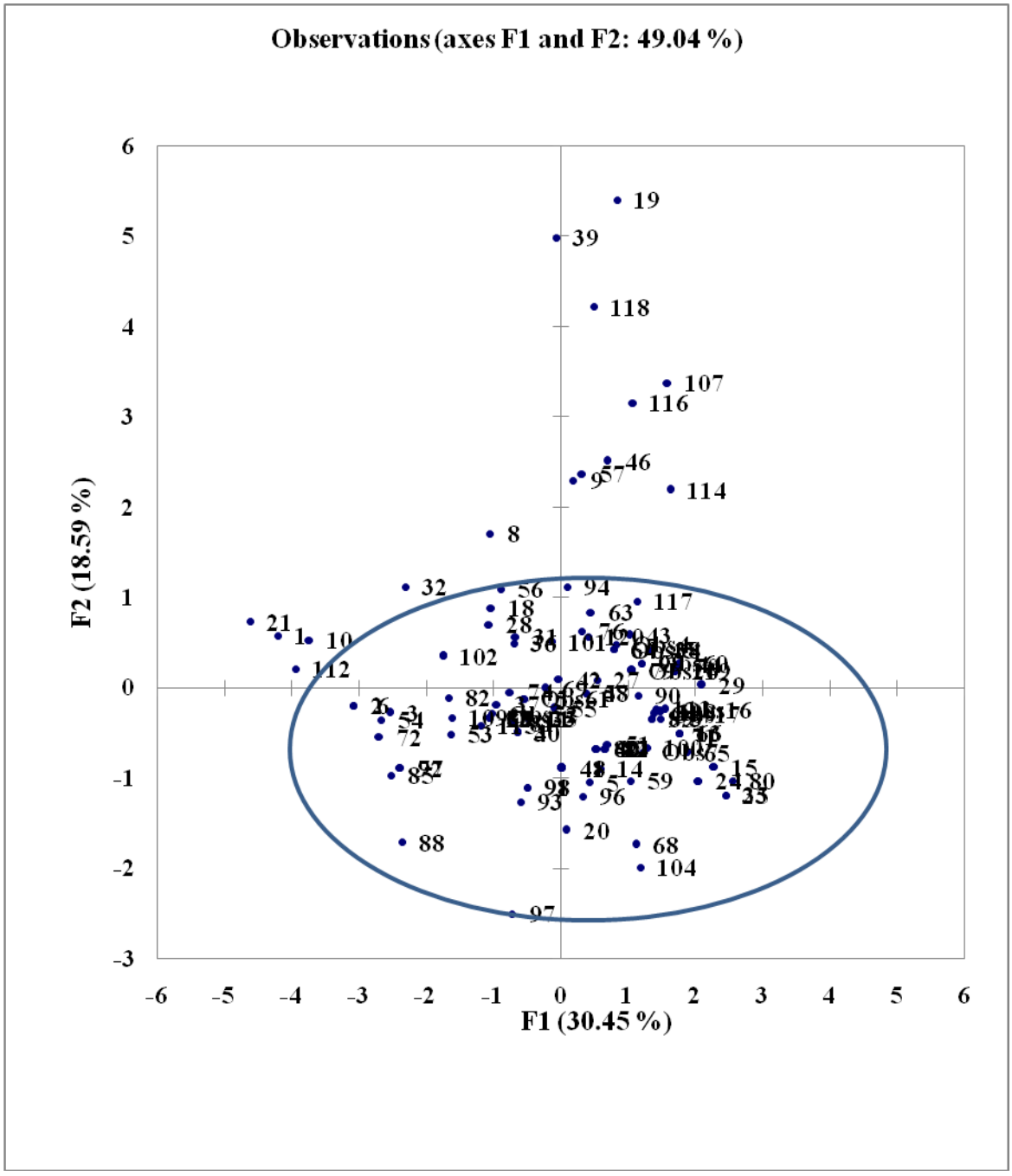

Figure 4.7: Principal component analysis of napier clones using qualitative characters

Key: F1 represents First principal component, F2 represents second principal component

The first two factors (leaf color and node hairiness) accounted for 49% of the total variability seen while the first five factors contributed to most variation seen in these qualitative traits at 83.9 % (Fig 4.8).

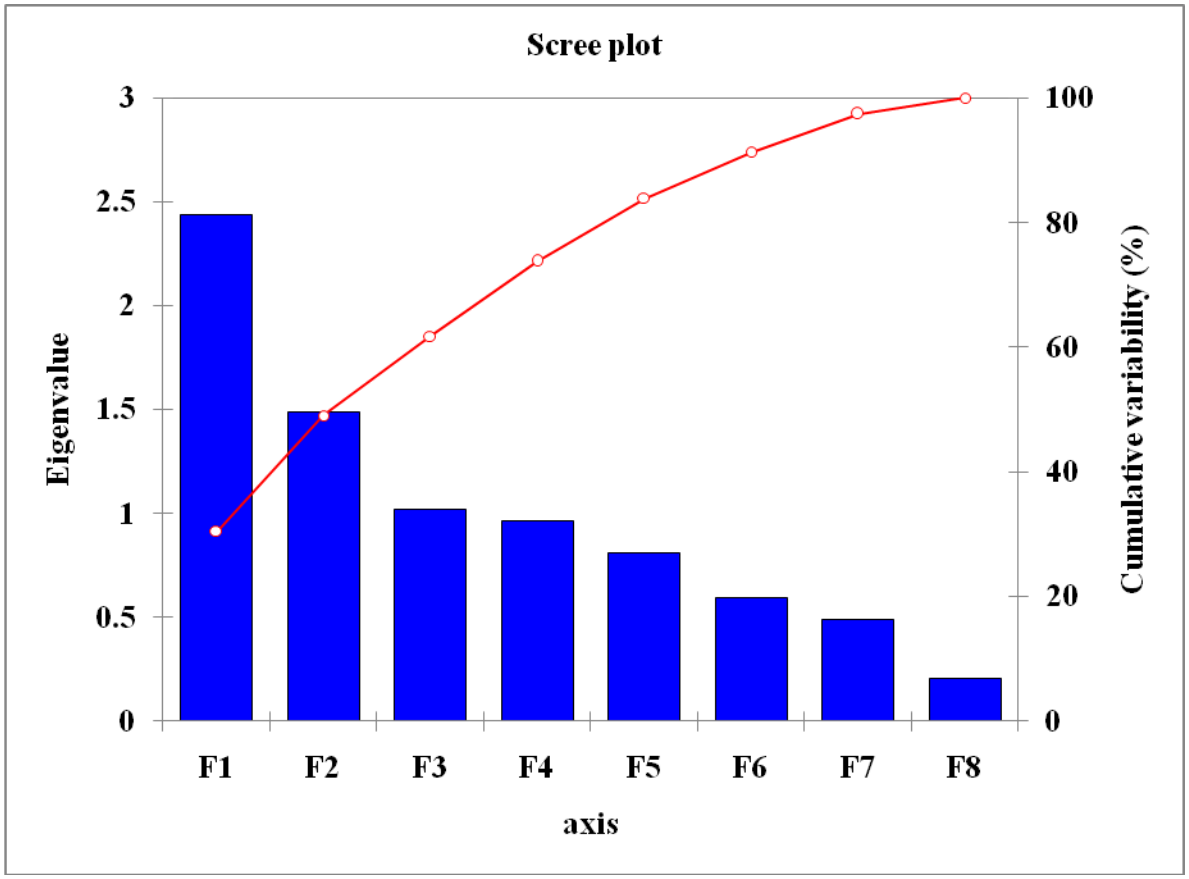

Figure 4.8: Graphical display of variance of each factor and the cumulative variability for qualitative characteristics

Key: F1= leaf color, F2= node hairiness, F3= adaxial, F4= abaxial, F5= seratteness, F6= leaf sheath edge hairiness, F7= leaf sheath hairiness, F8= length of sheath hairs

4.3.8.2 Combined morphological data of napier clones

A combined analysis for both qualitative and quantitative traits was carried out using XLSTAT to determine the most important morphological traits by measuring deviation

from the central tendency. This gave an indicator of the level of diversity in that, the more that a trait deviates from the mean, the more that it was expected to contribute to the total population diversity (Table 4.10).

Table 4.10: Summary statistics of the morphological traits showing central tendency

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
leaf_colour	120	1.000	4.000	2.042	0.438
node_hairness	120	0.000	2.800	1.179	0.459
adaxial	120	1.000	3.000	2.317	0.648
abaxial	120	1.000	3.000	1.900	0.541
serratness	120	1.000	3.000	2.392	0.539
lif_she_edge_hai	120	1.000	3.000	1.117	0.434
lif_she_hairiness	120	1.000	3.000	2.292	0.679
len_she_hairs	120	1.000	4.000	2.942	1.040
Leaf_length	120	66.200	111.800	91.408	9.556
Leaf_width	120	1.900	5.120	3.616	0.715
Stool_diameter	120	16.200	135.300	56.536	15.167
plant_height	120	34.200	213.900	139.133	30.996
stem_thickness	120	4.050	17.930	7.055	2.161
internode_length	120	0.000	19.310	6.968	2.371
growth_form	120	30.000	89.000	71.342	13.563
Tiller_no	120	1.400	35.800	13.155	7.143

Most of the variation seen in the morphological characterization was mainly contributed by plant height, stool diameter, growth form, leaf length and tiller number.

4.3.8.3 Cluster analysis of napier clones in the five populations studied

Cluster analysis partitioned the clones into two major clusters and four sub clusters. This was in line with the molecular cluster analysis of the same clones where two major clusters were identified. However, the clustering was not uniform in that clones could not

be traced based on the clusters formed for both molecular and morphological data (Figure 4.9).

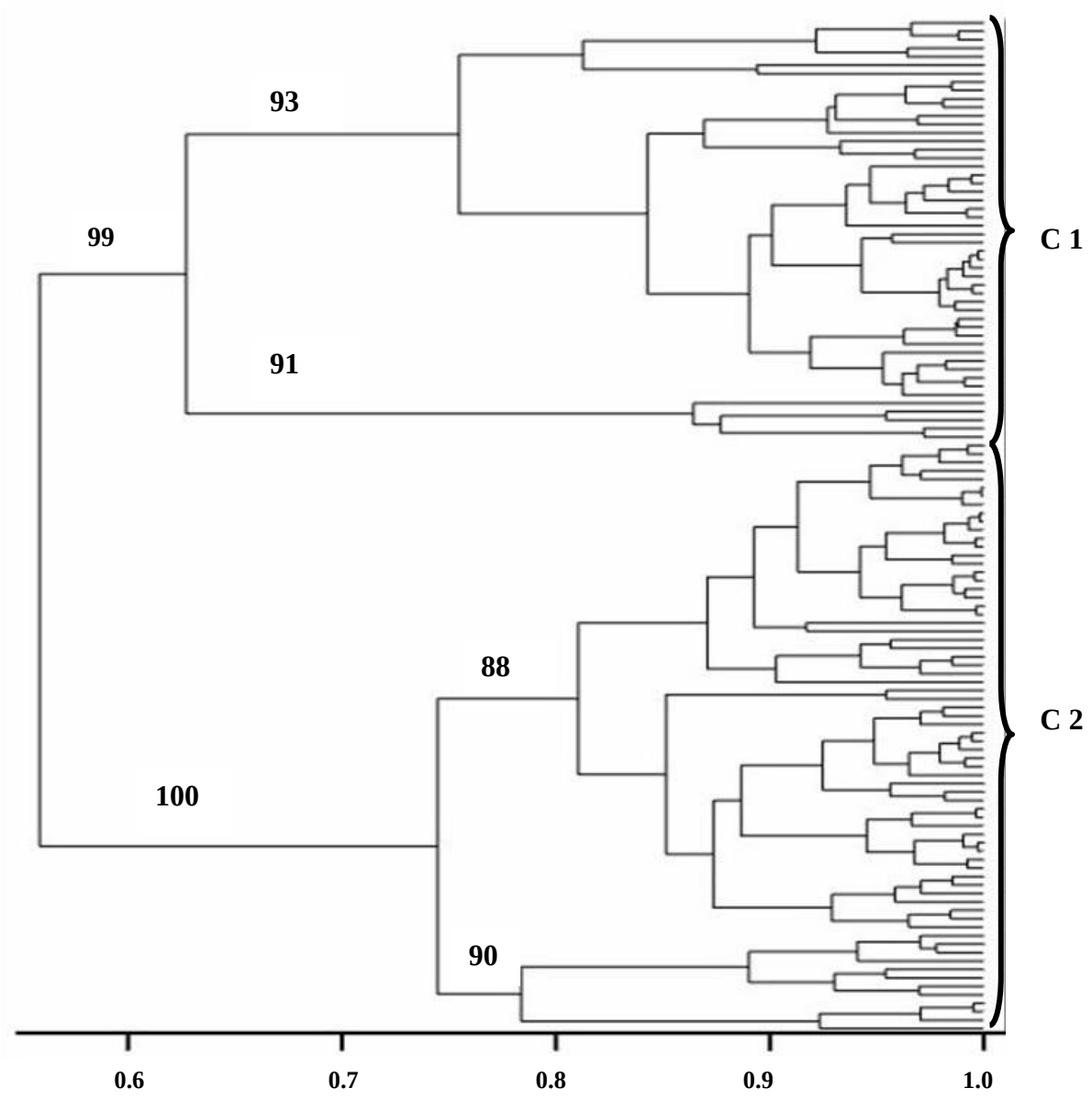

Figure 4.9: Unrooted dendogram showing clustering of napier clones using dissimilarity matrix
Key: C 1 = Major cluster 1; C 2= Major cluster 2

CHAPTER FIVE

DISCUSSION

This is the first time that Kenyan napier grass used for napier stunt disease resistance breeding has been characterized using SSR markers. These results provide important information for planning of breeding strategies for resistance against Phytoplasma 16SRXI that causes napier stunting disease. Genetic characterization of breeding plant materials is necessary before the materials are used for any breeding objective to ascertain how variable the gene pool is in a particular population.

5.1 Within population genetic diversity

5.1.1 Deviation from Hardy Weinberg Equilibrium

HWE principle predicts how gene frequencies will be inherited from generation to generation with assumptions that there is no mutation, gene migration, selection or genetic drift (Cedric et al., 2006). In Busia and Mumias clones, there was a significant deviation from the principle. These clones were collected in the farmer's fields where it was expected that these were farmers' preferred clones. Due to selection in terms of farmers' desired characteristics, there tend to be biasness in terms of existence of certain clones. This is directed or artificial selection which is one of the contributors of deviation from HWE (Azevedo et al., 2012). Selection leads to gene depletion within a specified gene pool while on the other hand increasing or accumulating certain genes in another gene pool which then leads to reduction and increase of gene frequencies respectively (Budak et al., 2003). Another factor that could have led to significant deviation observed

in Mumias and Busia populations could be the fact that clones are moved from one farmer to another due to various reasons or characters preferred by the recipient farmer. This could be explained by the fact that the emergence of NSD in these regions, and lack of resistant or tolerant variety has prompted the farmers to acquire clones from other farmers in their locality or even across the district borders in search of tolerant and high yielding clones which hence leads to gene migration from one gene pool to another. Gene migration or gene flow leads to changes in gene frequencies which then lead to significant deviation from HWE in the said populations. This study also does not rule out the fact that mutation over years could have contributed to deviation from HWE even though spontaneous mutations take a long time and that the occurrence of positive mutants is far minimal (Mariac et al., 2006). Another factor could be that, the collection sites are separated by barriers. In a diversity study using foxtail millet (*Setaria italica*) from Taiwan, Heng sheng et al., 2012 obtained a high HWE significant deviation (87.5%) using SSR markers. The study inferred that, the main contributor to this deviation could be the fact that the collection sites are geographically separated by distance. Other contributing barriers could be mountains, valleys and rivers. The same was inferred by Upadhyaya et al., 2008 where he suggested that the main contributors to significant deviation of HWE in a collection of chick pea could be due to man-made selection, isolation by geographical barriers and natural selections.

5.1.2 Mean number of alleles

The total number of alleles sampled in a population across all loci ranged between 71 in Alupe to 86 in Mumias. The mean number of alleles per locus across all the populations was 3.264 ±0.069. Alupe had the least value at 2.840±0.0138 while Busia, Butere and

Mumias populations had the highest number of alleles at 3.440±0.153. The mean number of alleles corroborates other studies using Pearl millet (*Pennisetum glaucum*) which ranged from 2.76 to 4.68 in diversity studies using SSR markers (Baskaran et al., 2009 and Sathyavath et al., 2009). Allele number is a factor of the plant population; the higher the population, the higher the number of the available alleles in the gene pool. In this case, Mumias, Busia and Butere populations had relatively higher number of clones sampled which translated to a higher mean number of alleles. This higher mean number of alleles in Busia, Butere and Mumias populations showed that there is more allelic diversity which implies that there exists cross pollination between clones in these regions even if the viability of the napier pollen is short lived. The higher number of alleles detected could also mean that the markers with the high mean allelic number could be used more universally than the ones with low mean allelic number (Otoo et al., 2009). Toro et al .(2008) proposed that the higher the number of alleles in a population under study, the higher the diversity because the maximum diversity occurs when alleles are at equal frequencies. The assumption here is that all alleles in a particular population have the identical frequency.

The high allelic diversity could not be related to admixture since most of the napier clones in this region are clonally propagated through either root splits or stem cuttings (Corderio et al., 2000).

5.1.3 Private alleles

Private alleles are alleles that were unique in only one population. These were observed in all populations but at different loci. Busia and Alupe populations had each two private alleles namely XIPES 0093, PGIRD 5 and PGIRD 5, XCUP 63, respectively and a total

of seven private alleles in the entire population. The total number of private alleles in this study was low as compared to a study using Sorghum which is a close relative to napier grass but not as close as Pearl millet. Mutegi et al (2010), studying the diversity of farmers' sorghum varieties and comparing with the wild types, had 15 private alleles for the cultivated and 13 for the wild with an average of 2.6 private alleles per locus.

The higher the number of private alleles, the higher the genetic diversity. This study shows moderate diversity which can be exploited in increasing the genetic base of napier grass breeding germplasm. The unique alleles imply that these clones could be reservoirs of important traits or novel genes such as pests and disease resistance. The idea of using the existence of private alleles has been well documented in sorghum breeding particularly for disease and pest resistance such as sorghum shoot fly (Kamala et al., 2009), sorghum midge (Sharma and Franzmann 2001), green bug (Duncan et al., 1991), downy mildew (Kamala et al., 2002) and ergot (Reed et al., 2002).

The existence of private alleles is also an indicator of gene flow which is brought about by migratory forces. The distance between Busia and Alupe towns is approximately 20 Kilometers. This close proximity could have enhanced gene flow between clones in these regions through farmers' exchange of clonal materials or through other means like field days, shows and demonstrations. The higher frequencies of the private alleles in Alupe and Busia materials could also be explained by the fact that these two towns border Uganda and porosity of the border could easily lead to free exchange of clonal materials hence gene flow (Wanjala et al., 2013).

5.1.4 Effective Number of alleles (Ne)

This is the number of alleles that can be present in a freely interbreeding population. In this study, Alupe population had the lowest value which could be attributed to the population size sampled for the study. In 2006, Cedric in a study comparing the diversity between the cultivated and wild Pearl millet, concluded that the number of alleles in a population is directly related to the population size which hence also affects the number of effective alleles. The number of clones in Alupe population could have reduced the chances of having more genotypes which also means that chances of having diverse alleles were low hence low intrapopulation gene diversity. On the other hand, Mumias had the highest value but with a higher population size. Higher population size on the other hand could mean high number of genotypes in the Mumias population and that the probabilities of having diverse alleles were high which could translate to high intrapopulation diversity in the Mumias population.

5.1.5 Expected Heterozygosity (Nei's genetic distance)

Heterozygosity is a measure of intrapopulation gene diversity or proportion of individuals heterozygous at a locus. The mean observed heterozygosity H_O and expected heterozygosity H_E was 0.539±0.023 and 0.491±0.013 respectively across all loci and all population. The observed heterozygosity across all loci ranged between 0.509±0.053 in Alupe population to 0.551 ± in Busia population. The highest mean expected heterozygosity (gene diversity) across all loci was 0.498±0.033 in Mumias population and the lowest was 0.481± 0.024 in Alupe population. The difference of both values

among the napier populations was statistically significant ($P \leq 0.05$) since the observed heterozygosity values were higher compared to the expected heterozygosity.

The level of genetic diversity in the entire napier grass population was moderate as compared to the standard range of diversity estimation using heterozygosity indices. This corroborates the findings of Wanjala et al. (2013) using AFLP, Xie et al. (2009) using SRAP and Lowe et al (2003) using RAPD. The low diversity in terms of heterozygosity in Alupe population could be attributed to the small population size. The higher heterozygosity seen in other populations can be attributed to larger populations sampled and high gene flow due to frequent exchange of plant materials. Another probable reason could be that, napier grass undergoes protogyny where the stigmas are excited prior to anther excition which facilitates out crossing. This hence reduces inbreeding depression and increases heterozygosity in napier grass (Passos et al., 2005).

AMOVA results showed that most of the diversity for the population studied resides within the population rather than between the populations which is also corroborated with heterozygosity indices. This could be attributed to napier grass breeding systems and the fact that napier grass is a highly heterozygous tetraploid species, it's an out crossing but a clonaly propagated species, low gene flow and probably mutation. Variations in napier grass cultivars were expected to be high due to its rich gene pool and wide parental diversity (Azevedo et al., 2012). Other probable reasons could be due to artificial selection for agronomic traits by the farmers who target specific or common target qualities which leads to uniformity among clones.

5.2 Between population genetic diversity

Wright's fixation indices (F statistics) commonly used for describing genetic differentiations among populations was calculated. The mean F_{IS} value was –0.081± 0.074. This means that in general, the populations showed excess heterozygosity across the 25 loci studied. This shows that there is low inbreeding of napier clones studied and that the population is not at a risk of inbreeding. Moderate genetic differentiation was observed in napier grass by Hartl and Clark et al. (1989) which is corroborates the findings frm this study. Other related out-crossing species such as perennial rye grass (*Lolium perenne*), meadow fescue (*Festuca pratensis*), orchard grass (*Dactylis glomerata*), and rhodes grass (*Chloris gayana* Kunth) (Huff, 1997; Kölliker et al., 1998; Ubi et al., 2003) also show moderate genetic differentiation using F statistics.

F_{ST} is the coefficient of gene differentiation and is used to measure the degree of genetic sub division among populations. The average F_{ST} value across all the loci in the entire population was low at 0.026± 0.003. This means that the genotype frequencies among populations are not randomly distributed. Both F_{ST} and G_{ST} indicate that the level of gene differentiation between the populations was low. Based on F_{ST} value, only 2.6% of total variation was contributed by genetic sub division among populations while 97.4% of total variation accounted for genetic variations among individuals within populations. This was in line with an earlier report on napier diversity studies where nested analysis partitioned by country among the countries studied showed that 97% of the variation observed in the Kenyan population was within individuals while 3% resided between the populations (Wanjala et al., 2013). In the same study, four populations were analyzed which included Kenya, Uganda, Tanzania and ILRI-FG populations which also showed

that 91% of the total variation accounted for genetic variations among individuals within populations while 9% was between the populations.

In the present study, the genetic variation among the populations was quantified by analysis of molecular variance (AMOVA), and the results showed that 90% of genetic variation accounted for variation within individuals in the population across all the loci, 9% among individuals and 1% among populations.

F_{ST} and AMOVA values obtained in this study between the 5 populations show a moderate genetic differentiation of the Napier clones. This could be attributed to the fact that napier grass is a clonal plant with low seed setting and germination and is spread by asexual stem and root reproduction. Due to this, gene flow is low and most genetic variation resides within rather than between individual clones because vegetative propagation rarely leads to creation of variability between populations (Kölliker et al., 1998).

Another probable reason for high intra population variation may be that napier grass is a highly heterozygous tetraploid species with a rich gene pool and wide parental diversity (Azevedo et al. 2012). Wide parental diversity could be attributed to the fact that napier grass has a good combining ability and can form hybrids with various *Pennisetum* species which fall in different gene pools and with the B genome being dominant over other genomes in such a hybrid combination (Ubi et al., 2003).

5.3 Genetic relationship between napier populations

The phylogenetic relationship between the napier populations (where a population was defined as clones from the same place of collection hence five populations namely;

Alupe, Busia, Bungoma, Mumias and Butere) was determined using Nei's genetic distance. Phylogenetic reconstruction was performed using Neighbor Joining method (Nei, 1987) to summarize the genetic distances between the napier cultivars. Neighbor joining method classified the napier populations into two major phylogenetic clusters and one minor cluster. Different napier populations from geographically distinct regions would be expected to emerge clearly distinct from each other. However, the contrary was true in this study. The clusters were not distinct with genotype overlap and redundancy. This would be attributed to the fact the materials under circulation share ancestry especially since napier grass is clonally propagated, and the farmers freely exchange planting materials. These finding supports the study by Wanjala et al (2013) where the napier clones under that study also clustered into two major clusters and one minor cluster. On the other hand, these findings contradict previous studies that demonstrated clustering of napier grass accessions based on geographical location (Lowe et al., 2003; Harris et al., 2009). However the marker used in the study by Lowe et al (2003) was RAPD which lacks reproducibility and is hence unreliable.

5.4 Morphological characterization of napier grass in the five populations

In this study, morphological characterization using fourteen characteristics produced two major clusters with an additional minor cluster which is supported by the molecular data. Principal component analysis using qualitative characteristics produced two major clusters with a few exceptional clones. However in both qualitative and combined morphological data methods, the clustering was not region specific. This is also supported by molecular analysis. It is not easy to separate the clones based on the fourteen morphological descriptors because both the qualitative and quantitative

characters are affected by the environment. Azevedo et al (2012) characterized 25 napier grass introductions using 15 descriptors and concluded that it is extremely difficult to separate the introductions by morphological characterization only. In contrast to morphological traits, which can be influenced by temperature, soil type, nutrients, insects, etc., the use of molecular markers can provide new insights to better understand the genetic variation within the germplasm collection (Harris et al., 2009). This indicates that both techniques can provide moderately consistent information for diversity analyses in accessions of napier grass from different origins, with different or similar edapho climatic adaptations.

Morphological analysis revealed highly significant differences among clones for all of the traits suggesting that there was a high degree of phenotypic diversity among the clones. Plant height, stool diameter, growth form, leaf length and tiller numbers showed wide variation which means that they could be the most important useful morphological characters for grouping the clones.

CHAPTER SIX

CONCLUSSION AND RECOMMENDATIONS

6.1 Conclusion

This work was unique in that both morphological and microsatellite profiles were utilized. SSR methodology used in this study was able to discriminate among the napier grass accessions and could be useful in future for clonal analysis especially to asses napier grass introductions and as a tool for fingerprinting. From this study, it can be concluded that;

1. The average genetic diversity obtained in napier grass from western Kenya is moderate. This is clearly demonstrated by clonal overlap and redundancy in both molecular and morphological data.

2. The napier grass populations studied are not in equilibrium as shown by the varied allele frequencies among the populations and variation in Hardy Weinberg indices.

6.2 Recommendations

1.There is need to increase diversity of the western Kenya gene pool through introductions of wild types and other clonal materials from other regions of Kenya and other countries. This will increase chances of breeding resistant clones to napier stunting disease and other traits of interest by having a diverse gene pool.

2. There is need to address the disequilibrium seen in the populations in terms of the allele frequencies by;

a) Increasing the level of gene flow through more free exchange of materials among these five populations and especially into areas with low allele frequencies.

b) Improving on methods of clonal selection through use of the most important morphological traits.

REFERENCES

Alford R.L, Hammond H.A., Coto I. and Caskey C.T. (1994). Rapid efficient resolution of parentage by amplification of short tandem repeats. *American Journal of HumanGenetics* 55: 190-195.

Anderson W.F., Dien B.S., Brandon S.K. and Peterson J.D. (2008). Assessment of Bermuda grass and bunch grasses as feedstock for conversion to ethanol.*Appl.Biochem.Biotechnol.* **145**: 13-21.

Augustin E., and Tcacenco F.A. (1993). Isozymatic characterization of elephant grass (*Pennisetum purpureum* Schum). Rev. *Brasil.Genet*. **16**:685–696

Awalla B.J., Mulaa M., Lusweti C.M. and Kute C. (2009).Morphological characterization of napier grass (*Pennisetum purpureum* schumach) collections in Kenya. Proceedings of the 12th KARI Biennial Scientific Conference.Transforming Agriculture for improved livelihoods through Agricultural Product Value Chains . 8^{th}–12^{th} Nov 2010.

Ayotamuno J.M., Kogbara R.B. and Egwuenum, P.N. (2006). Comparison of corn and elephant grass in the phyto remediation of a petroleum-hydrocarbon contaminated agricultural soil in Port Harcourt, Nigeria. *Int. J. Food, Agric.Environ.*, **4**: 218-222.

Azevedo A.L.S., Costa P.P., Machado J.C., Machando M.A., Pereira A.V. and Ledo F.J.S. (2012). Cross species amplification of Pennisetum glaucum microsatellite markers in Pennisetum purpureum and genetic diversity of Napier grass accessions. *Crop Science* **4**: 1776–1785.

Babu C., Sundaramoorthi J., Vijayakumar G. and Ram S.G. (2009). Analysis of genetic diversity in napier grass (Pennisetum purpureum Schum) as detected by RAPD and ISSR markers. *J. Plant Biochem. Biotechn.* **18**: 123-133.

Baskaran K., Hash C.T., Senthilvel S., Raj A.G.B. and Muthiah A.R. (2009). SSR allele frequency changes in response to recurrent selection for pearl millet grain yield and other agronomic traits, *SAT eJournal.*

Bell C.J. and Ecker J.R. (1994). Assignment of 30 microsatellite loci to the linkage map of Arabidopsis. *Genomics* **19** : 137-144

Bhandari A.P., Sukanya D.H. and Ramesh C.R. (2006). Application of isozymes data in fingerprinting Napier grass (*Pennisetum purpureum* Schum) for germplasm management. *Genetic Resources and Crop Evolution* **53**: 253–264

Boonman J.G. (1997). Farmers' success with tropical grasses: Crop/pastures rotation in mixed farming in East Africa. Ministry of Foreign Affairs, The Hague, Netherlands, p. 95.

Bolstein D., White R.L., Skolmick M. and Davis R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Amer. J of Hum Genet* **32**: 314–332.

Brunken J.N. (1977). A systematic study of Pennisetum sect. Pennisetum (Gramineae). *American Journal of Botany* **64**:161-167.

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

Burton G.W. (1944). Hybrids between napier grass and cattail millet. *J. Hered.* **35**:227-232.

Burton G.W. (1989). Registration of 'Merkeron' napiergrass. *Crop Sci*. **29**:1327.

Cedric M., Viviane L., Issoufou K., Fabrice S. and Yves V. (2006). Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. *Theor Appl Genet* **114**:49–58

Chee M., Yang R., Hubbell E., Berno A., Huang X. C., Stern D., Winkler J., Lockhart D. J., Morris M. S. and Fodor S. P. A. (1996). Accessing genetic information with high-density DNA arrays. *Science* **274**:610-614.

Cheng H.H. and Crittenden L.B. (1994). Microsatellite markers for genetic mapping in the chicken. *Poult Sci* **73**: 539-546

Chen P., Zhu J., Liu P. and Shi X. L. (2007). Analysis of genetic relationships in new lines of hybrid Pennisetum by RAPD markers. *Chin. J. Grassl.* **29**: 34–38.

Corderio G.M., Maguire T.L., Edwards K.J. and Henry R.J. (1999). Optimisation of a microsatellite enrichment technique in *Saccharum* spp. *Plant Molecular Biology Reporter* **17**: 225-29.

Corderio G.M., Taylor G.O. and Henry R.J. (2000). Characterisation of microsatellite markers from sugarcane (*Saccharum sp.*), a highly polyploid species. *Plant Science* **155**: 161-68.

Daher R.F., Pereira M.G., Pereira A.V. and Amaral A.T. (2002). Genetic divergence among Elephantgrass cultivars accessed by RAPD markers in composit samples. *Sci. Agric*. **59**: 623-627.

Downey L.D. and Iezzoni A.F. (2000). Polymorphic DNA markers in black cherry are identified using sequences from sweet cherry, peach and sour cherry. *Journal of The American Society For Horticultural Science* **125**: 76-80.

Doyle J.J. and Doyle J.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochemical Bulletin* **19**:11–15.

Duncan R.R., Bramel-Cox P.J. and Miller F.R. (1991). Contributions of introduced sorghum germplasm to hybrids development in the USA. In: Shands HL, and Wiesner LE (eds) Use of Plant Introductions in the Cultivar Development. *Crop Science Society of America,* Madison, USA, pp 69 - 101

Edwards A., Civitello A., Hammond H. A. and Caskey C. T. (1991). DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am. J. Hum. Genet.* **49**: 746-756.

Edwards A., Hammond H. A., Jin L., Caskey C.T. and Chakraborty R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**: 241-253

Edwards K.J., Barker J.H.A., Daly A., Jones C. and Karp A. (1996). Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques* **20**: 758-759.

Farrel G.S.A. and Hillocks R.J. (2001). Aspects of the biology of Ustilago Kamerunensis, a smut pathogen of Napier grass (*Pennisetum purpureum*). *J. Phytopathology* **149**: 739-744.

Gallego F.J., Pérez M.A., Núñez Y. and Hidalgo P. (2005). Comparison of RAPDs, AFLPs and SSR markers for the genetic analysis of yeast strains of *Saccharomyces cerevisiae. Food Microbiol.* **22**: 561-568.

Gitu w. kang'ethe. (2004). Agricultural Development and Food Security in Kenya. Food and Agriculture Organization (FAO).

Godwin I.D., Aitken E.A.B. and Smith L.W. (1997). Application of inter simple sequence repeat (ISSR) markers to plant genetics Electrophoresis.**18**:1524-1528.

Gonzalez B. and Hanna W.W. (1984). Morphological and fertility responses in isogenic triploid and hexaploid pearl millet x napier grass hybrids. *J. Hered.*. **75(4):** 317-318.

Gurta P. K. and Varshney R. K. (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica **113**:163–185.

Hanna W.W., Chaparro C.J., Mathews B.W., Burns J.C., Sollenberger L.E. and Carpenter J.R. (2004). Perennial *Pennisetums.* In: Moser LE, Burson BL and Sollenberger LE (eds)Warm season (C4) Grasses. American Society of Agronomy Monograph SeriesNo.45.Madison: ASA/CSSA/SSSA, pp. 503–535.

Hanna W.W. and Monson W. (1988). Registration of dwarf Napier grass germplasm. *Crop Sci.* **28:**870- 871.

Harris K. F. (1979). Leafhoppers and aphids as biological vectors: In K. Maramorosch and K. F. Harris (ed.), Leafhopper Vectors and Plant Disease Agents. Academic Press, New York, N.Y. Vector-virus relationships, p. 217-308.

Hartl D.L. and Clark A.G. (1989). Principles of Population Genetics, 3rd edn. Sinauer Associates, Inc, Sunderland, MA.

Heng S. L., Chih Y.C., Song B.C., Gwo I.L. and Chang S. K. (2012). Genetic diversity in the foxtail millet (*Setaria italica*) germplasm as determined by agronomic traits and microsatellite markers. *AJCS* **6(2):**342-349 (2012)**.**

Hormaza J.I. (2002). Molecular characterization and similarity relationships among apricot genotypes using simple sequence repeats. *Theoretical and Applied Genetics* **104**: 321-28.

Huff D.R. (1997). RAPD characterization of heterogeneous perennial ryegrass cultivars. *Crop Science.* **37**: 557-594.

Jaccoud D., Peng K., Feinstein D. and Kilian .A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* **29:** e25.

Jacob H. J., Lindpaintne K., Lincoln S. E., Kusumi K., Bunker R. K., Mao. Ganten D. Dzau V. J. and Lander E. S. (1991). Genetic mapping of a gene causing hypertensive rat. *Cell* **67:** 213-224.

Jauhar P.P. (1981). Cytogenetics and breeding of pearl millet and related species. A.R. Liss, New York.

Jauhar P.P. and Hanna W.W. (1998). Cytogenetics and genetics of pearl millet. *Adv. Agron.* **64**:2-21.

Jones P., Arocha T., Zerfy J., Proud J., Abebe G. and Hanson J. (2006). A stunting syndrome of Napier grass in Ethiopia associated with 16SrIII Group phytoplasma.New Disease Reports, http://www.bspp.org.uk/ndr/july2006/2006-19.asp

Jones P., Devonshire B.J., Holman T.J. and Ajanga, S. (2004). Napier grass stunt: a new disease associated with a 16SrXI group phytoplasma in Kenya. *Plant Pathology* **53**:519.

Kamala V., Sharma H.C., Manohar R.D., Varaprasad K.S.and Bramel P.J. (2009). Wild relatives of sorghum as sources of resistance to sorghum shoot fly, *Atherigona soccata. Plant Breeding* **128**:137-14

Kamala V., Singh S.D., Bramel P.J. and Rao D.M. (2002). Sources of resistance to downy mildew in wild and weedy sorghums. *Crop Sci* **42**:1357-1360

Karlsson S.B. and Vasil I.K. (1986). Morphology and ultra structure of embryogenic cell suspension cultures of *Panicum maximum* (guinea grass) and *Pennisetum purpureum* (napier grass). *Am. J. Bot.* **73**:894-901.

Kennan T.C.D. (1952). The Improved West African Napier fodders Rhod. *Agric J.* **49**:221-228.

Khan Z.R., Picket J.A., Wadhams L.J. and Muyekho F. (2001). Habitat management strategies for the control of cereal stem borers and striga in maize in Kenya. *Insect Science and its Application* **21**: 375-80.

Kolliker R., Stadelmann F.J., Reidy B. and Nosberger J. (1998). Genetic variability of forage grass cultivars: A comparison of *Festuca pratensis* Huds, *Lolium perenne* L.and *Dactylis glomerata* L. *Euphytica*. **106**:261–270.

Li G. and Quiros C. F. (2001). Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica* .*Theor Appl Genet* **103**:455–461

Lima R.S., Daher R.F, Goncalves L.S, Rossi D.A, Perreira M.G. and Ledo F.J. (2011). RAPD and ISSR markers in the evaluation of genetic divergence among accessions of elephant grass. *Genetic and molecular research* **10(3):**1304-1313

Litt M. and Luty J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**: 397-401.

Li Y.C., Korol A.B., Fahima T., Beiles A. and Nevo E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms:a review. *Mol. Ecol.* **11**: 2453-2465.

Lombard V., Baril C. P., Dubreuil P., Blouet F. and Zhang D. (2000). Genetic relationships and fingerprinting of rapeseed cultivars by AFLP: consequences for varietal registration. *Crop Sci.* **40:** 1417- 1425

Lowe A.J., Thorpe W., Teake A. and Hanson J. (2003). Characterization of germplasm accession of Napier grass (*Pennisetum purpureum* and *P. glaucum* hybrids) and comparison with farm clones using RAPD. Genetic Resources and *Crop Evolution* **50:** 121–132.

Lusweti C.M., Nandasaba J., Ong'injo E. and Asena D (2004). Preliminary results of disease survey on Napier grass in selected sites of Western Kenya. Pasture Research Annual Report 2004. National Agricultural Research Centre, Kitale Kenya. P 6-7.

Lynch M. and Milligan B.G. (1994). Analysis of population genetic structure using RAPD markers. Mol. Ecol., **3**: 91-99.

Mariac C., Luong V., Kapran I., Mamadou A. Sagnard F., Deu M., Chantereau J.,Gerard B., Ndjeunga J., Bezançon G., Pham J.L. and Y. Vigouroux. (2006).

Diversity of wild and cultivated pearl millet accessions [*Pennisetum glaucum* (L.) R. Br.] in Niger assessed by microsatellite markers. *Theor.Appl.Genet.***114**:49–58.

Martel E., De Nay D., Siljak-Yakovlev S., Brown S. and Sarr A. (1997). Genome size variation and basic chromosome number in pearl millet and fourteen related*Pennisetum* species. *J. Hered*. **88**:139-143.

Martinez L., Cavagnaro P., Masuelli R. and Rodriguez J. (2003). Evaluation of diversity among Argentine gapevine (*Vitis vinifera* L.) varieties using morphological data and AFLP markers. *J. Biotechnol*. **6** :242-250.

Matsuoka Y., Mitchell S.E., Kresovich S., Goodman M. and Doebley J. (2002). Microsatellites in Zea - variability, patterns of mutations, and use for evolutionary studies. Theor. Appl. Genet. 104:436- 450.

Meade C. and Parnell J. (2003). Multivariate analysis of leaf shape patterns in Asian species of the Uvaria group (*Annonaceae*). Bot. J. Linn. Soc. **143**: 231-242.

Mulaa M., Ajanga S. and Wilson M. (2004). A survey to collect and identify potential vectors of Napier grass stunting disease associated with phytoplasma in Western Kenya. Pasture Research Annual Report. National Agricultural Research centre, Kitale. p 8-13

Mwendia S. W., Wanyoike M., Wahome R. G. and Mwangi D. M. (2006). Farmers' perceptions on importance and constraints facing Napier grass production in Central Kenya. *Livestock Research for Rural Development* **18** :11

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

Nei M. (1987). Genetic distance between populations. *American Naturalist* **106**:283–292.

Newbury J. and Ford-Lloyd B. (1999). The production of user-friendly molecular marker for studying plants. *Biotechnology News* **40:** 5-6

O'Brien S.J. (1991). Molecular genome mapping: lessons and prospects. *Current Opinion in Genetic Development.* **1(1):** 105-111.

Obura E., Masiga D., Midega C.A.O., Wachira F., Pickett J.A., Deng A.L. and Khan Z.R. (2010). First report of a phytoplasma associated with Bermuda grass white leaf diseasein Kenya. *New Disease Reports.* **21**, 23.

Obura E., Masiga D., Midega C.A.O., Otim M., Wachira F., Pickett J. and Khan Z.R. (2011). Hyparrhenia grass white leaf disease, associated with a 16SrXI phytoplasma,newly reported in Kenya. *New Disease Reports.* **24**:17.

Obura E., Masiga D., Wachira F., Gurja B. and Khan Z.R. (2011). Detection of phytoplasma by loop-mediated isothermal amplification of DNA (LAMP). *Journal of Microbial Methods.* **84(2)**: 312-316.

Olson M., Hood L., Cantor C. and Bostein D. (1989). A common language for physical mapping of the human genome. *Science* **245**:1434-1435

Orodho A.B. (2006). Dissemination and utilization of research technology on forages and agricultural by-products in Kenya. In: Utilization of research results on forage and agricultural by-product materials as animal feed resources in Africa. Proceedings of the first joint workshop held in Lilongwe, Malawi 5-9 December 1988. PANESA/ARNAB, Addis Ababa, Ethiopia, pp 833.

Otoo E.,Akkromah R., Kolesnikova A.M. and Asiedu R. (2009). Ethno botany and morphological characterization of the yam pona complex in Ghana. In. *Proceedings of the African Crop Science* **9**:407-414.

Passos L.P., Machado M.A., Vidigal M.C. and Campos A.L. (2005). Molecular characterization of elephant grass accessions through RAPD markers. *Agro.***29**: 568-574

Peakall R. and Smouse P.E. (2009). GENALEX 6: Genetic Analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**:288–295.

Peleman J.D. and Van der Voort J.R. (2003). Breeding by Design. *Trends Plant Sci.* **8**:330-334.

Perrier X. and Jacquemoud-Collet J.P. (2006). DARwin software. *http:// darwin.cirad.fr/* (16 July 2014).

Plaschke J., Ganal M.W. and Roder M.S. (1995). Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theoretical and Applied Genetics* **91:** 1001-7.

Rao N.K., Reddy L.J. and Bramel P.J. (2003). Potential of wild species for genetic enhancement of some semi-arid food crops. *Genet. Resour. Crop Evol.* **50**:707 –721.

Rao N.S. (1929). On the chromosome numbers of some cultivated plants of south India. *Indian J Bot Sci* **8**:126-128.

Reed J.D., Ramundo B.A., Claflin L.F and Tuinstra M.R. (2002). Analysis of resistance to ergot sorghum and potential alternate hosts. *Crop Sci* **42**:1135-1138.

Roy A., Frascaria N., Mackay J. and Bonsquet J. (1992). Segregating random amplified polymorphic DNA (RAPDs) in *Betula alleghaniensis*. *Theor. Appl. Genet.* **85**:173-180.

Saghai-maroof M.A., Bishayev R.M., Yang G.P., Zhang Q. and Allard R.W. (1994). Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations and population dynamics. *Proceedings of the National Academy of Science* **91**: 466-70.

Sanchez-eschribano E.M., Martin J.R., Carreno J. and Cenis J.L. (1999). Use of sequence tagged microsatellite site markers for characterizing table grape cultivars. *Genome* **42:** 87-93.

Sathyavathi C.T., Begum S., Singh B.B., Unnikrishnan K.V. and Bharadwaj C. (2009). Analysis of diversity among cytoplasmic male sterile sources and their utilization in developing F1 hybrids in pearl millet [*Pennisetum glaucum* (L.) R. Br.], *Indian J. Genet.*, **69**: 352-360.

Semagn K., Bjørnstad A. and Ndjiondjop M.N. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology* **5 (25):** 2540-2568

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

Sharma H.C. and Franzmann B.A. (2001). Host plant preference and oviposition responses of the sorghum midge, *Stenodiplosis sorghicola* (Coquillett) (Dipt., *Cecidomyiidae*) towards wild relatives of Sorghum. *J.Appl Ent* **125**: 109-114

Shenoy V.B. and Vasil I. K. (1992). Biochemical and molecular analysis of plants derived from embryogenic tissue cultures of napier grass (*Pennisetum purpureum*. Schum). *Theoretical and Applied Genetics.* **83**: 947-955

Smith. R. L., Chowdhury M. K. U. and Schank S. C. (1989). Use of restriction fragment-length polymorphism (RFLP) markers in genetics and breeding of napier grass. Proceedings of Soil and Crop Science Society of Florida Boca Raton.**48**: 13-19.

Smith R.L., Schweder M.E., Chowdhury M.K.U. and Seib J.C. (1993). Development and application of RFLP and RAPD DNA markers in genetic improvement of *Pennisetum* for biomass and forage production. *Biomass Bioenergy* **5**: 51-62.

Souza S.F., Pereira A.V. and Ledo F.J.S. (2005). Agronomic evaluation of interspecific hybrids of elephant grass and pearl millet. *Brazilian J. Agric. Res.* **40(9):** 873-880.

Staal S., Chege L., Kenyanjui M., Kimari A., Lukuyu B., Njumbi D., Owango M.,Tanner J. C., Thorpe W. and Wambugu M. (1997). Characterisation of dairy systems supplying the Nairobi milk market. International Livestock Research Institute, Nairobi.

Sujatha D.M., Manga V., Subba M.V. and Murty J.S.R. (1989). Meiotic studies in some species of *Pennisetum* (L.) Rich. (*Poaceae*). *Cytologia*. **54(4):** 641-652.

Testonlin R., Marrazo T., Cipriani G., Quarta R., Verde I., Dettori T., Pancaldi M. and Sansavini S. (2000). Microsatellite DNA in peach (*Prunus persica* L Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. *Genome*. **43**: 512-20.

Thompson J.B. (1919). Napier and merker grasses, two new forage crops for Florida. Bull. *Univ. Fla. Agric. Exp. Stn*. 317-153.

Tiley G.E.D. (1959). Elephant grass in Uganda. Kawanda report.

Ubi G.N., Kenworthy W.J., Costa J.M., Cregan P.B. and Alvernaz J. (2003). Genetic diversity of soybean cultivars from China, Japan, North America, and North American ancestral lines determined by amplified fragment length polymorphism. *Crop Science.* **43**:1858-1867.

Upadhyaya H.D., Dwivedi S.L., Baum M., Varshney R.K., Udupa S.M. and Gowda C.L.L.(2008). Genetic structure, diversity, and allelic richness in composite collection and reference set in chickpea (*Cicer arietinum* L.) *BMC Plant Biol* **8**: 106.

Van de Wouw M., Hanson J. and Luethi S. (1999). Morphological and agronomic characterization of 9 accessions of Napier grass. *Tropical grasslands.***33**: 150-1

Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M., Frijters A., Pot J., Peleman J. and Kuiper M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407–4414.70

Wanjala B.W., Obonyo M., Wachira F.N., Muchugi A., Mulaa M., Harvey J., Skilton R.A., Proud J. and Hanson J. (2013). Genetic diversity in Napier grass (*Pennisetum purpureum*) cultivars: implications for breeding and conservation. *AoB PLANTS plt022; doi:10.1093/ aobpla/plt022*

Watt J.M. and Breyer B.M.G. (1962). The medicinal and poisonous plants of Southern and Eastern Africa. 2nd ed, E.&S. Livingstone Ltd.,Edinburg and London.

Weising K., Nybom H., Wolff K. and Meyer W. (1995). DNA Fingerprinting in Plants and Fungi. CRC Press, London.

Wijk A.J.P. (1977). The breeding of Napier grass (*Pennisetum purpureum* Schumach.) Kitale report.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenetics* **15**: 323–354.

Xiao J., Li J., Yuan L., Mccouch S.R. and Tingey S.V. (1996). Genetic Diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theoretical and Applied Genetics.* **92**: 637-643.

Xie X. M. and Lu X. L. (2005). Analysis of genetic relationships of cultivars in *Pennisetum* by RAPD markers. *Acta Prataculturae Sinica* **14**: 52–56.

Xie X.M., Zhou F., Zhang X.Q. and Zhang J.M. (2009). Genetic variability and relationships between MT-1 elephant grass and closely related cultivars assessed by SRAP markers. *Journal of Genetics* **88**: 281–290

Yu K., Park S.J., Poysa V. and Gepts P. (2000). Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). *J. Hered.* **91**: 429-434.

Zane L., Bargelloni L. and Patarnello T. (2002). Strategies for microsatellite isolation: a review. *Mol. Ecol.* **11**: 1-16.

Zietkiewicz E., Rafalski A. and Labuda D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics.* **20**: 176-183

APPENDIX

Appendix I: DNA extraction protocol (Modified CTAB)

Materials /Reagents

CTAB buffer

Microfuge tubes

Micropestles

Liquid Nitrogen

70 % Ethanol (ice cold)

Isopropanol (ice cold)

Chloroform : Iso Amyl Alcohol (24:1)

Water (sterile)

Water bath

CTAB buffer 100ml

2.0 g	CTAB (Hexadecyl trimethyl-ammonium bromide)
10.0 ml	1 M Tris pH 8.0
4.0 ml	0.5 M EDTA pH 8.0 (EthylenediaminetetraAcetic acid Di-sodium salt)
28.0 ml	5 M NaCl
40.0 ml	H2O
200µl	2-mercaptoethanol

Procedure

1. 300mg of Napier leaf samples was transfred into a labeled sterile 1.5ml microfuge tube.
2. The leaf was ground to a fine paste in liquid nitrogen using a micropestle.
3. 600µl of hot (65˚C) CTAB buffer was added and vortexed gently to mix.
4. The CTAB/plant extract mixture was incubated for 1 hour at 65˚ C in a water bath.
5. The mixture was left to cool at room temperature, equal volume of chloroform:isoamyl alcohol (24:1), was added and mixed by inverting the solution.

6. The plant extract mixture was spinned at 4000rpm for 10 min to spin down cell debris.
7. The upper aqueous phase was transferred (containing the DNA) to a clean microfuge tube.
8. 800µl ice cold iso propanol was added to the aqueous layer and the tubes inverted several times to precipitate the DNA. The tubes were placed for 2hrs at -20˚C to precipitate the DNA.
9. After precipitation, the tubes were centrifuged at 14,000rpm for 30mins and supernatant poured out.
10. The DNA pellet was washed by adding two changes of ice cold 70 % ethanol**.** Ethanol was poured out carefully.
11. The pellet was dried at room temperature for 1 hour, with caution not to over dry the DNA. 12. DNA was resuspended in 50 µl low salt TE.
13. A quick ascertain of DNA integrity was done by running in 0.8% agarose gel.
14. 2µls of RNAse A enzyme was added and incubated in a water bath set at 37°C for 1 hour and storing followed at -20°C.

Appendix II:Summary of chi square tests for HWE

Pop	**Locus**	**DF**	**ChiSq**	**Prob**	**Signif**
Alupe	ctm_8	3	4.238	0.237	ns
Alupe	ctm_10	6	10.756	0.096	ns
Alupe	ctm_12	1	2.880	0.090	ns
Alupe	ctm_27	1	3.645	0.056	ns
Alupe	ctm_59	3	0.735	0.865	ns
Alupe	pgird_5	3	0.735	0.865	ns
Alupe	pgird_46	3	9.000	0.029	*
Alupe	pgird_56	6	10.440	0.107	ns
Alupe	pgird_57	3	4.160	0.245	ns

Appendix II continuation

Pop	Locus	DF	ChiSq	Prob	Signif
Alupe	pgird_13	3	1.332	0.722	ns
Alupe	pgird_21	1	0.258	0.612	ns
Alupe	pgird_25	3	3.181	0.365	ns
Alupe	phi_227562	3	8.000	0.046	*
Alupe	psmp_2235	1	0.735	0.391	ns
Alupe	psmp_2248	1	4.022	0.045	*
Alupe	psmp_2255	3	3.645	0.302	ns
Alupe	psmp_2266	3	8.222	0.042	*
Alupe	psmp_2267	6	4.407	0.622	ns
Alupe	xcup_14	3	4.810	0.186	ns
Alupe	xcup_53	3	1.007	0.799	ns
Alupe	xcup_63	1	4.706	0.030	*
Alupe	xipes_0093	6	6.107	0.411	ns
Alupe	xipes_0191	1	2.250	0.134	ns
Alupe	xipes_0219	3	6.421	0.093	ns
Alupe	xtxp_278	1	0.031	0.860	ns
Bungoma	ctm_8	6	40.619	0.000	***
Bungoma	ctm_10	6	13.343	0.038	*
Bungoma	ctm_12	3	16.411	0.001	***
Bungoma	ctm_27	1	21.000	0.000	***
Bungoma	ctm_59	3	0.623	0.891	ns
Bungoma	pgird_5	1	0.107	0.743	ns
Bungoma	pgird_46	3	18.025	0.000	***

Appendix II continuation

Pop	Locus	DF	ChiSq	Prob	Signif
Bungoma	pgird_56	6	22.592	0.001	***
Bungoma	pgird_57	3	1.562	0.668	ns
Bungoma	pgird_13	10	8.762	0.555	ns
Bungoma	pgird_21	1	0.247	0.619	ns
Bungoma	pgird_25	3	3.064	0.382	ns
Bungoma	phi_227562	6	28.271	0.000	***
Bungoma	psmp_2235	3	25.573	0.000	***
Bungoma	psmp_2248	3	15.850	0.001	**
Bungoma	psmp_2255	3	1.562	0.668	ns
Bungoma	psmp_2266	6	21.501	0.001	**
Bungoma	psmp_2267	3	4.320	0.229	ns
Bungoma	xcup_14	6	19.846	0.003	**
Bungoma	xcup_53	6	5.157	0.524	ns
Bungoma	xcup_63	3	18.017	0.000	***
Bungoma	xipes_0093	3	5.230	0.156	ns
Bungoma	xipes_0191	1	2.222	0.136	ns
Bungoma	xipes_0219	6	22.500	0.001	***
Bungoma	xtxp_278	3	0.583	0.900	ns
Busia	ctm_8	3	14.122	0.003	**
Busia	ctm_10	6	18.457	0.005	**
Busia	ctm_12	1	12.451	0.000	***
Busia	ctm_27	3	30.494	0.000	***
Busia	ctm_59	3	1.469	0.689	ns

Appendix II continuation

Pop	Locus	DF	ChiSq	Prob	Signif
Busia	pgird_5	3	22.678	0.000	***
Busia	pgird_46	3	13.526	0.004	**
Busia	pgird_56	6	19.704	0.003	**
Busia	pgird_57	3	0.436	0.933	ns
Busia	pgird_13	10	24.900	0.006	**
Busia	pgird_21	1	0.720	0.396	ns
Busia	pgird_25	6	8.995	0.174	ns
Busia	phi_227562	6	31.222	0.000	***
Busia	psmp_2235	3	4.114	0.249	ns
Busia	psmp_2248	6	25.722	0.000	***
Busia	psmp_2255	6	3.154	0.789	ns
Busia	psmp_2266	6	17.942	0.006	**
Busia	psmp_2267	6	25.356	0.000	***
Busia	xcup_14	3	12.178	0.007	**
Busia	xcup_53	6	18.560	0.005	**
Busia	xcup_63	3	13.107	0.004	**
Busia	xipes_0093	3	12.477	0.006	**
Busia	xipes_0191	1	1.049	0.306	ns
Busia	xipes_0219	6	7.518	0.276	ns
Busia	xtxp_278	6	0.263	1.000	ns
Butere	ctm_8	6	12.330	0.055	ns
Butere	ctm_10	6	14.410	0.025	*
Butere	ctm_12	3	34.000	0.000	***

Appendix II continuation

Pop	Locus	DF	ChiSq	Prob	Signif
Butere	ctm_27	6	34.211	0.000	***
Butere	ctm_59	3	1.250	0.741	ns
Butere	pgird_5	1	1.250	0.264	ns
Butere	pgird_46	3	20.000	0.000	***
Butere	pgird_56	6	12.124	0.059	ns
Butere	pgird_57	3	19.157	0.000	***
Butere	pgird_13	10	12.790	0.236	ns
Butere	pgird_21	1	0.969	0.325	ns
Butere	pgird_25	6	9.862	0.131	ns
Butere	phi_227562	6	53.107	0.000	***
Butere	psmp_2235	6	5.045	0.538	ns
Butere	psmp_2248	6	13.925	0.030	*
Butere	psmp_2255	3	2.878	0.411	ns
Butere	psmp_2266	6	22.580	0.001	***
Butere	psmp_2267	6	13.750	0.033	*
Butere	xcup_14	3	13.295	0.004	**
Butere	xcup_53	3	0.254	0.968	ns
Butere	xcup_63	1	6.027	0.014	*
Butere	xipes_0093	3	10.945	0.012	*
Butere	xipes_0191	3	0.969	0.809	ns
Butere	xipes_0219	3	3.941	0.268	ns
Butere	xtxp_278	6	0.131	1.000	ns
Mumias	ctm_8	6	43.041	0.000	***

Appendix II continuation

Pop	Locus	DF	ChiSq	Prob	Signif
Mumias	ctm_10	6	31.499	0.000	***
Mumias	ctm_12	6	50.000	0.000	***
Mumias	ctm_27	3	49.688	0.000	***
Mumias	ctm_59	6	2.204	0.900	ns
Mumias	pgird_5	1	1.395	0.238	ns
Mumias	pgird_46	3	17.323	0.001	***
Mumias	pgird_56	6	12.123	0.059	ns
Mumias	pgird_57	6	27.056	0.000	***
Mumias	pgird_13	6	26.029	0.000	***
Mumias	pgird_21	1	0.718	0.397	ns
Mumias	pgird_25	6	12.960	0.044	*
Mumias	phi_227562	6	37.870	0.000	***
Mumias	psmp_2235	6	4.954	0.550	ns
Mumias	psmp_2248	6	21.469	0.002	**
Mumias	psmp_2255	6	18.509	0.005	**
Mumias	psmp_2266	3	4.096	0.251	ns
Mumias	psmp_2267	6	14.770	0.022	*
Mumias	xcup_14	3	16.964	0.001	***
Mumias	xcup_53	6	5.337	0.501	ns
Mumias	xcup_63	1	1.469	0.225	ns
Mumias	xipes_0093	3	11.449	0.010	**
Mumias	xipes_0191	1	5.385	0.020	*
Mumias	xipes_0219	6	15.728	0.015	*

Mumias	xtxp_278	3	0.173	0.982	ns

Key: ns=not significant, * P<0.05, ** P<0.01, *** P<0.001

Appendix III: Allele Frequencies and Sample Size by Populations

Locus	Allele/n	Alupe	Bungoma	Busia	Butere	Mumias
ctm_8	**N**	7	19	16	18	25
	234	0.429	0.395	0.438	0.389	0.440
	254	0.071	0.079	0.125	0.139	0.100
	259	0.500	0.421	0.438	0.444	0.420
	265	0.000	0.105	0.000	0.028	0.040
ctm_10	**N**	9	21	19	20	27
	55	0.167	0.262	0.263	0.200	0.259
	71	0.111	0.167	0.184	0.100	0.130
	89	0.111	0.238	0.368	0.300	0.315
	165	0.611	0.333	0.184	0.400	0.296
ctm_12	**N**	8	20	19	17	25
	158	0.000	0.000	0.000	0.059	0.080
	292	0.625	0.500	0.447	0.471	0.440
	297	0.375	0.475	0.553	0.471	0.460
	314	0.000	0.025	0.000	0.000	0.020

Appendix III continuation

Locus	Allele/n	Alupe	Bungoma	Busia	Butere	Mumias
ctm_27	**N**	9	21	19	19	27
	160	0.000	0.000	0.000	0.053	0.148
	292	0.389	0.500	0.474	0.421	0.407
	296	0.611	0.500	0.474	0.500	0.444
	314	0.000	0.000	0.053	0.026	0.000
ctm_59	**N**	9	20	18	20	27
	100	0.056	0.025	0.028	0.050	0.019
	173	0.778	0.850	0.778	0.800	0.778
	291	0.167	0.125	0.194	0.150	0.185
	335	0.000	0.000	0.000	0.000	0.019
pgird_5	**N**	9	21	19	20	27
	90	0.056	0.000	0.000	0.000	0.000
	153	0.167	0.333	0.289	0.200	0.185
	158	0.778	0.667	0.658	0.800	0.815
	163	0.000	0.000	0.053	0.000	0.000
pgird_46	**N**	9	21	19	20	27
	54	0.056	0.095	0.184	0.075	0.093
	88	0.500	0.476	0.447	0.500	0.444
	91	0.444	0.429	0.368	0.425	0.463
pgird_56	**N**	9	20	18	20	24
	100	0.222	0.150	0.056	0.075	0.063
	132	0.111	0.125	0.194	0.225	0.167
	141	0.556	0.625	0.722	0.650	0.729

Appendix III continuation

Locus	Allele/n	Alupe	Bungoma	Busia	Butere	Mumias
	156	0.111	0.100	0.028	0.050	0.042
pgird_57	**N**	8	21	19	19	27
	92	0.000	0.000	0.000	0.053	0.148
	104	0.313	0.143	0.053	0.079	0.019
	126	0.625	0.786	0.868	0.868	0.815
	296	0.063	0.071	0.079	0.000	0.019
pgird_13	**N**	9	21	18	20	27
	128	0.000	0.024	0.028	0.050	0.000
	160	0.167	0.167	0.278	0.150	0.148
	231	0.389	0.524	0.417	0.400	0.426
	236	0.444	0.238	0.222	0.250	0.148
	262	0.000	0.048	0.056	0.150	0.278
pgird_21	**N**	9	20	18	19	25
	158	0.278	0.100	0.167	0.184	0.140
	160	0.722	0.900	0.833	0.816	0.860
pgird_25	**N**	8	21	19	20	27
	58	0.000	0.000	0.026	0.025	0.037
	65	0.188	0.119	0.105	0.150	0.185
	154	0.125	0.238	0.211	0.275	0.222
	161	0.688	0.643	0.658	0.550	0.556
phi_227562	**N**	8	18	18	20	25
	166	0.063	0.056	0.056	0.025	0.080
	179	0.000	0.028	0.028	0.025	0.040

Appendix III continuation

Locus	Allele/n	Alupe	Bungoma	Busia	Butere	Mumias
	311	0.500	0.528	0.500	0.525	0.520
	330	0.438	0.389	0.417	0.425	0.360
psmp_2235	**N**	9	21	19	20	27
	76	0.000	0.000	0.000	0.025	0.037
	154	0.000	0.048	0.026	0.050	0.019
	161	0.778	0.595	0.711	0.750	0.796
	192	0.222	0.357	0.263	0.175	0.148
psmp_2248	**N**	8	15	18	13	16
	58	0.313	0.167	0.111	0.115	0.250
	83	0.000	0.000	0.056	0.115	0.031
	100	0.000	0.133	0.111	0.154	0.063
	159	0.688	0.700	0.722	0.615	0.656
psmp_2255	**N**	9	21	19	20	27
	221	0.611	0.786	0.711	0.725	0.630
	234	0.333	0.095	0.237	0.250	0.259
	267	0.000	0.000	0.026	0.025	0.056
	289	0.056	0.119	0.026	0.000	0.056
psmp_2266	**N**	8	18	16	19	24
	121	0.125	0.056	0.063	0.053	0.000
	165	0.375	0.472	0.406	0.237	0.271
	171	0.500	0.417	0.438	0.579	0.458
	175	0.000	0.056	0.094	0.132	0.271
psmp_2267	**N**	7	12	16	11	17

Appendix III continuation

Locus	Allele/n	Alupe	Bungoma	Busia	Butere	Mumias
	98	0.071	0.000	0.063	0.182	0.206
	174	0.214	0.125	0.188	0.182	0.206
	194	0.643	0.625	0.531	0.545	0.559
	217	0.071	0.250	0.219	0.091	0.029
xcup_14	**N**	9	16	10	16	20
	100	0.056	0.031	0.000	0.219	0.075
	122	0.278	0.438	0.300	0.250	0.450
	151	0.000	0.063	0.200	0.000	0.000
	161	0.667	0.469	0.500	0.531	0.475
xcup_53	**N**	9	17	16	18	22
	75	0.333	0.559	0.313	0.278	0.227
	88	0.278	0.176	0.313	0.250	0.364
	166	0.389	0.235	0.313	0.472	0.386
	282	0.000	0.029	0.063	0.000	0.023
xcup_63	**N**	9	18	13	19	25
	99	0.278	0.056	0.077	0.237	0.120
	122	0.722	0.917	0.846	0.763	0.880
	172	0.000	0.028	0.000	0.000	0.000
	324	0.000	0.000	0.077	0.000	0.000
xipes_0093	**N**	9	20	19	20	27
	116	0.389	0.500	0.500	0.475	0.519
	123	0.444	0.275	0.395	0.400	0.259
	130	0.111	0.225	0.105	0.125	0.222

Appendix III continuation

Locus	Allele/n	Alupe	Bungoma	Busia	Butere	Mumias
	136	0.056	0.000	0.000	0.000	0.000
xipes_0191	**N**	9	20	18	19	23
	68	0.667	0.750	0.806	0.816	0.674
	119	0.333	0.250	0.194	0.158	0.326
	267	0.000	0.000	0.000	0.026	0.000
xipes_0219	**N**	7	14	17	9	14
	130	0.000	0.286	0.235	0.222	0.214
	134	0.786	0.464	0.647	0.722	0.643
	139	0.071	0.071	0.029	0.056	0.036
	194	0.143	0.179	0.088	0.000	0.107
xtxp_278	**N**	9	21	19	20	27
	115	0.056	0.048	0.053	0.025	0.056
	157	0.944	0.857	0.895	0.925	0.926
	231	0.000	0.000	0.026	0.025	0.019
	246	0.000	0.095	0.026	0.025	0.000