MAPPING OF SINGLE NUCLEOTIDE POLYMORPHISM MARKERS FOR TOLERANCE TO CASSAVA BROWN STREAK DISEASE OF MKOMBOZI X TMS 4(2)1425 POPULATION

KASELE SALUM

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN CROP SCIENCE OF SOKOINE UNIVERSITY OF AGRICULTURE.

MOROGORO, TANZANIA.

2014

ABSTRACT

Cassava Brown Streak Disease (CBSD) is a viral disease that impacts cassava tuberous root quantity and quality by causing a brown, dry, corky necrosis within the starch bearing tissues. The efficient way to control this disease is the use of resistant cassava varieties. This study was carried out to genotype the mapping population of Mkombozi x TMS 4(2)1425 and develop a single nucleotide polymorphic (SNP) genetic linkage map. Before SNP genotyping was done, the study to determine the integrity of 325 F1 individuals using ten polymorphic simple sequence repeat (SSR) markers to eliminate unknown parental combinations was done. SSR data revealed 158 true crosses, 135 off-types or admixtures, 31 selfs and one genotype (MT 210) was rejected due to many missing data. SNP genotyping was carried out using a reduced representation genotyping-by-sequencing approach at the University of Berkeley. Genetic linkage analysis was performed using Join Map software, version 4.1, Linkage analysis resulted in a total of 1136 SNP markers spanning 3207.9 cM assigned into 18 linkage groups (LGs) of the integrated linkage map. The average marker spacing of 3.4 cM that ranged from 1.3 cM to 10.1 cM was observed from this population. The total of 802 SNP markers were distributed on 19 LGs of female parent (Mkombozi) spanning 2848 cM with an average marker spacing of 3.7 cM ranging from 1.7 to 6.2 cM. On the male parent (TMS 4(2)1425), 593 SNP markers were distributed on 18 linkage groups spanning 2482.8 cM with average marker spacing of 5.2 cM ranging from 1.9 cM to 12.6 cM. The genetic linkage map developed in this study will be used for the identification of molecular markers and quantitative trait loci (QTLs) associated with CBSD tolerance in Mkombozi. The molecular markers and QTLs identified provide useful materials to breed and select CBSD resistance cassava varieties through marker-assisted selection (MAS).

DECLARATION

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

Kasele Salum — 9/7/2014

(MSc. Candidate) — Date

The above declaration is confirmed

Prof. Paul Kusolwa — 09/07/2014

(MSc. Supervisor) — Date

COPYRIGHT

ACKNOWLEDGEMENTS

To Almighty God, the Creator and my Redeemer, thank you, as always, for keeping me strong and healthy throughout my study. It is all because of you I have been able to complete my studies.

I wish to sincerely thank my supervisor Professor Paul Kusolwa of the Faculty of Agriculture, Department of Crop Science and Production for his patience, diligent guidance, constructive criticism and supervision during the course of research and dissertation writing which enabled me to accomplish this work. I am also thankful to Dr. Morag Ferguson for her supervisory role during the laboratory work of the study and for guiding me on data analysis and final dissertation write-up. I also wish to extend my gratitude to my wife Mamy Hassan Kozi for her encouragement during my studies. My gratitude goes to my sponsor. Thanks to the Bill and Melinda Gates Foundation for sponsoring my study through the grant which supported a project titled "Biotechnology Applications to Combat Cassava Brown Streak Disease" of which my MSc work formed a component, Jessica Lyons for doing genotyping-by-sequencing and Jessen Bredesen for SNP data capture under the Bill and Melinda Gates Foundation project "Cassava Genomics: bridging the gap between sequence and breeding applications". I thank BeCANet and BeCA-Hub for supporting my laboratory work at BeCA-ILRI biotechnology facility.

My gratitude to the Permanent Secretary, Ministry of Agriculture, Food Security and Cooperatives, Director for Research and Development and Zonal Director Research and Development Lake Zone for granting me permission to pursue this study. I thank the staff of ARI Maruku and Ukiriguru, who assisted me during development of crossing blocks

and collecting seeds and ARI Makutupora for providing me area for multiplication of my mapping population. I am also indebted to Dr. Geoffrey Mkamilo, the Root and Tuber Crops Coordinator, of ARI Naliendele for managing and disbursements of my funds. I wish to convey my special thank to Dr. Heneriko Kulembeka for his hard work and constant help during the write-up of my proposal and dissertation.

Last but not least, to my numerous friends and colleagues, thank you for your time and support during my study.

DEDICATION

This work is dedicated:

To my beloved mother Hawa Shabani Makenya and my father Salum Rashid Feruzi for bringing me up to who I am today.

To my lovely wife Mamy Hassan Kozi, my sons Salmin Kasele Feruzi and Haruna Kasele Feruzi for their support, patience and for hard times they went through during my study period. I will always love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND SYMBOLS

°C	Degrees Celsius
%	Percentage
ABI	Applied Biosystems
ACMV	African cassava mosaic virus
AFLP	Amplified fragment length polymorphism
AICC	Arusha International Conference Centre
ARI	Agricultural Research Institute
ATP	Adenosine triphosphate
BecA	Biosciences Eastern and Central Africa
BecANet	Biosciences Eastern and Central Africa Network
BMC	Bio Med Central
bp	Base pairs
cal	Calories
CBB	Cassava bacterial blight
CBSD	Cassava brown streak disease
CBSV	Cassava brown streak virus
CGM	Cassava green mite
CIAT	Centro International de Agricultural Tropical/International Center for Tropical Agriculture
cm	Centimetre
cM	centiMorgan
CMB	Cassava mealy bug
CMD	Cassava mosaic disease

DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DRC	Democratic Republic of the Congo
EACMV	East African cassava mosaic virus
EARRNET	Eastern Africa Root Crops Research Network
EDTA	Ethylenediaminetetraacetic acid
EST	Expressed sequence tag
FAO	Food and Agriculture Organization
g	Gram
GBS	Genotyping by sequencing
h	Hour
ha	Hectare
HCl	Hydrochloric acid
Hi-Di	Highly deionised
IFAD	International Fund for Agricultural Development
IITA	International Institute of Tropical Agriculture
ILRI	International Livestock Research Institute
IPGRI	International Plant Genetic Resources Institute
LG	Linkage group
LOD	Logarithm of odds
m	metre
M	Molar
MAP	Months after planting
MAS	Marker-assisted selection
masl	Metre above sea level

MgCl2	Magnesium chloride
min	Minute
ml	Millilitre
mM	Millimole
MSc	Master of Science
NaCl	Sodium chloride
ng	Nanogram
PCR	Polymerase Chain Reaction
pH	Hydrogen Potential
PhD	Doctor of Philosophy
PVP	Polyvinylpyrrolidone
QTL	Quantitative trait loci
RAPD	Randomly amplified polymorphism DNA
RFLP	Restriction fragment length polymorphism
Rnase	Ribonuclease
rpm	Revolutions per minute
SDL	Segregation distortion loci
SDS	Sodium dodecyl sulphate
Sec	Seconds
SNP	Single nucleotide polymorphism
SRAP	Sequence related amplified polymorphism
SRI	Sugar Research Institute
SSR	Simple sequence repeats
t	Tonne
Taq	*Thermus aquaticus*

TBE	Tris Borate EDTA
TE	Tris -HCL EDTA
TMS	Tropical *Manihot* species, a prefix for all cassava lines developed by IITA
Tris-HCL	Tris (Hydroxymethyl) aminomethane hydrochloride
TRTCP	Tanzania Root and Tuber Crops Program
U	Unit
UCBSV	Uganda cassava brown streak virus
UgV	Uganda variant
UK	United Kingdom
US$	United States dollars
USA	United States of America
UV	Ultraviolet
V	Volt
v/v	Volume per volume
w/v	Weight per volume
µl	Microlitre

CHAPTER ONE

1.0 INTRODUCTION

Cassava, (*Manihot esculenta* Crantz) (2n = 36) which originates from Latin America is an important food security crop for many tropical and subtropical countries. It is a source of calories for 800 million people in tropical and subtropical Africa, Asia and Latin America with an average consumption of approximately 500 cal/day (Iglesias *et al.*, 1997; Ceballos *et al.*, 2010). The high starch content (20-40%) makes cassava a desirable energy source both for human consumption and industrial biofuel applications (Balat and Balat, 2009; FAO, 2008; Schmitz and Kavallari, 2009). Some of the most important industrial uses of cassava include its use as a raw material in the feed industry, a source of starch and starch derived products such as high fructose-glucose syrup and for the production of ethanol (Ceballos *et al.*, 2007; Kunkeaw *et al.*, 2010a). Although roots are poor in nutrition, consisting largely of carbohydrates, leaves are rich in proteins, vitamins and minerals and are an important source of vegetables in the Democratic Republic of the Congo (DRC), Tanzania, Kenya, Madagascar, Sierra Leone, Uganda and Zambia (Nweke *et al.*, 2002). The worldwide production of cassava is 233 million tons, of which more than 50% occurs in Africa (FAO, 2010). Asia and the Americas contribute about 33% and 15% of the world production, respectively.

Cassava is an important staple crop in more than half of Tanzania and a subsistence crop, especially in the semi-arid areas. Eighty four percent of the total cassava production in the country is utilized as human food while the remaining 16% is for other uses like starch production, livestock feed and export (FAO and IFAD, 2001). Both roots and leaves of cassava are of major nutritional importance in the country. Major cassava producing areas in Tanzania include Mwanza, Tanga, Coast, Mtwara, Mara, Ruvuma, Lindi, Shinyanga, and

Kigoma (Kapinga *et al.*, 2001). The average cassava yield for Tanzania is 5.5 t/ha (FAO, 2010). Compared to yields achieved in some south Asian countries such as India (34.8t/ha), Indonesia (20.2 t/ha), Thailand (18.86 t/ha) and China (16.8t/ha) (FAO, 2010) yield in Tanzania is very low. The average yield in Tanzania is also well below the average yield of 12 t/ha of Africa's (and the world's) largest cassava producer - Nigeria (FAO, 2010). This low yield is caused by many factors, including susceptibility of commonly grown varieties to major diseases and pests such as cassava brown streak disease (CBSD), cassava mosaic diseases (CMD), cassava bacterial blight (CBB), cassava green mite (CGM), cassava mealy bug (CMB) and nematodes. Biotic stress constitutes the principal production constraint in Africa and Latin America. Whiteflies in particular are considered one of cassava's major pests due to its role as a vector for viruses that cause major diseases in cassava as well as causing direct damage. The species *Bemisia tabaci* is the vector of the most important production constraints (CBSD and CMD) in Africa (Maruthi *et al.*, 2005). CBSD is caused by two viruses Cassava Brown Streak Virus (CBSV) and Uganda Cassava Brown Streak Virus (UCBSV) (Genus: *Ipomovirus*; Family: *Potyviridae*). CBSD causes 20 - 100% yield losses in cassava, affecting peoples' livelihoods all over eastern Africa (IITA, 2007).

In conventional breeding for CBSD resistant cassava varieties, the number of years required for the breeding and evaluation of promising clones under field conditions, approximately 10 years. A quicker means of breeding for CBSD resistance is clearly required. Use of molecular markers offers an opportunity to efficiently accelerate the breeding in cassava through marker-assisted selection (MAS). MAS allows for selection to take place at the seedling stage, it allows for the reduction of population sizes and in the time taken to deliver a new variety. Molecular markers have been useful in introgression of useful traits from wild *Manihot* relatives into cassava through MAS.

Markers for root quality traits like dry matter content, protein and delayed post-harvest deterioration as well as disease resistance have been identified and used successfully in cassava selection at Centro International de Agricultura Tropical (CIAT) (Akano *et al.*, 2002; Fregene *et al.*, 2006; Egesi *et al.*, 2008).

It has been advocated that the development and use of resistant varieties could potentially form the basis for controlling the cassava diseases with an understanding of genetics and inheritance of their resistance (De Vries and Toenniessen, 2001). In cassava, using phenotypic information and molecular genetic maps, the genetics of resistance of three important diseases, CBB, CMD and CBSD have been studied and markers associated with the resistance trait identified (Mba *et al.*, 2001; Akano *et al.*, 2002; Kulembeka, 2010). Kulembeka (2010) identified one significant quantitative trait locus (QTL) in one of the best source of resistance to CBSD available, a variety named 'Namikonga' in Tanzania and 'Kaleso' in Kenya (Kanju *et al.*, 2010). Basing on the analysis of data from diallel and selfing experiments, this resistance is likely to be conferred by a few genes (Kulembeka, 2010). Attempt in understanding the inheritance and gene action of CBSD was firstly done at Amani in Tanzania and breeding for resistance to CBSD and CMD started in 1940s (Nichols, 1947; Jennings and Iglesias, 2002). This was done using *M. glaziovii* and *M. melanobasis*, as the sources of resistance to CMD and CBSD in interspecific hybridization with cassava varieties to produce F1 hybrids with improved resistance (Jennings, 1957; 1960a, b). It was reported that resistance to CBSD was multigenic and recessive in inheritance (Jennings, 1960a; 1978; Jennings and Iglesias, 2002). Kanju *et al.* (2010) reported a zigzag stem trait controlled by recessive genes associated with tolerance to the disease whereby homozygous recessive and heterozygous "zigzag" genotypes were tolerant to CBSD. Some works to produce cassava varieties

with resistance to CBSD through genetic transformation have been undertaken at the International Institute of Tropical Agriculture (IITA) (Ingelbrecht *et al.*, 2005).

Identification of molecular markers tightly linked to CBSD resistance gene/s would be useful in breeding CBSD resistant cassava varieties. This can efficiently accelerate the generation of elite cassava varieties resistant to the disease. A pre-requisite for identifying markers associated with traits of interest is a framework genetic linkage map on which variation in the phenotype can be mapped. Linkage maps have been constructed in a number of economically important species such as soybean (Hwang *et al.*, 2009), brassica rapa (Li *et al.*, 2009), jatropha (Wang *et al.*, 2011) and grass carp (*Ctenopharyngodon idella*) (Xia *et al.*, 2010). In cassava, linkage maps have been constructed using different markers such as Restriction Fragment Length Polymorphism (RFLP), Randomly Amplified Polymorphic DNA (RAPD), SSR and isozyme (Fregene *et al.*, 1997) and SNP (Rabbi *et al.*, 2012). A large variety of genetic molecular markers exists, each with its own advantages and drawbacks. SSRs or Microsatellites used to be the markers of choice for generating genetic linkage maps. However recently SNP markers have attracted significant attention in creating dense genetic linkage maps and genome-wide association studies (Wang *et al.*, 2005). SNPs are the most abundant class of polymorphisms in genomes and can be genotyped cost-effectively (Rafalski, 2002).

A number of sources of tolerance to CBSD have been observed in germplasm from Tanzania such as Mkombozi which is one among varieties developed by the East Africa Root Crops Research Network (EARRNET), and is a half sib of 92/0099S2 (SM). Through disease and diagnostic survey conducted by IITA and Tanzania Root and Tuber Crops Program (TRTCP) in 2009, this variety was observed to be tolerant to CBSD in the lake zone of Tanzania (Kasele *et al.*, 2009). Therefore there is a need to identify the molecular markers

associated with new sources of CBSD resistance in this variety and become utilized in cassava breeding programs.

In this study there were two objectives:

i. To determine the integrity of a mapping population using Simple Sequence Repeat (SSR) markers to eliminate unknown parental combinations.

ii. Genotype the mapping population for single nucleotide polymorphism (SNP) markers and generate a genetic linkage map.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Cassava: Origin and Distribution

2.1.1 Origin of cassava

Cassava is a member of the *Euphorbiaceae*, subfamily Crotonoideae, the tribe *Manihotae* and the genus *Manihot.* The *Manihot* genus is reported to have about 100 species, among which the only commercially cultivated one is *Manihot esculenta* Crantz (Alves, 2002). The species of the genus range from trees to shrubs and perennial herbaceous plants with a woody rootstock known for the production of latex and cyanogenic glucosides (Fregene *et al.,* 2006). This genus *Manihot* has been formally studied since 1886. Previous studies on Deoxyribonucleic acid (DNA) sequence and simple sequence repeats (SSR) markers used to analyze populations of *M. esculenta* and wild relatives in the genus *Manihot* revealed that haplotypes and microsatellite alleles of cassava were subsets of those found in *M. esculenta* ssp *flabellifolia* (Olsen and Schaal, 1999, 2001). This confirmed that cassava is derived from *M. esculenta* ssp *flabellifolia* (Schaal *et al.*, 2006). The inclusion of the *M. pruinosa* population in the analysis showed no evidence of its hybridization with *M. esculenta* ssp *flabellifolia.* The data sets provided evidence that cassava was domesticated in the southern region between the lower Amazon forest and the Cerrado region of Brazil from *M. esculenta* ssp *flabellifolia* (Schaal *et al.*, 2006).

2.1.2 Cassava introduction and distribution

Cassava was first introduced from Latin America into Africa by Portuguese traders in the 16th century, around 1550 (Jennings and Hershey, 1985; Carter *et al.*, 1992). Cassava was first cultivated in Africa for the sole purpose of saving slaves in ships until 1600. The crop was introduced in Western and East Coast of Africa by Portuguese sailors.

By the 17th century, cassava cultivation diffused to other parts of Africa through European explorers, French Navy colonialists and African farmers (Lebot, 2009). In Tanzania cassava reached Lake Tanganyika from West Africa by the Congolese farmers, from there it moved inland in Tanganyika through farmer to farmer diffusion (Carter *et al.*, 1992). Further spread was reinforced by the colonial administrators who encouraged farmers to grow cassava and cultivation increased in the late 19th and 20th centuries. Cassava continued to spread due to its ability to survive in harsh conditions and viability of the cuttings which facilitated the natural spread of the crop (Masumba, 2006). Currently, in Tanzania cassava is grown almost in every area below 2000masl.

2.2 Morphological Characteristics of Cassava

Cassava is a semi-woody perennial shrub mainly grown for its starchy roots. The morphological characteristics of cassava are highly variable. Cassava landraces are usually characterized on the basis of morphological and agronomic descriptors (Alves, 2002) but sometimes based on the crop utility. An International Plant Genetic Resources Institute (IPGRI) descriptor list has 75 cassava descriptors of which 54 are morphological and 21 are agronomic (IPGRI and CIAT, 2003) and an IITA descriptor list has 50 cassava descriptors (Fukuda *et al.*, 2010). Morphological descriptors (i.e., lobe shape, root pulp colour, stem external colour) have higher heritability than agronomic characters (such as root length, number of root per plant and root yield). Among morphological descriptors, the following are considered as the minimum or basic descriptors that should be taken in mind during identification of a cultivar. These are apical leaf colour, stem external colour, root peduncle presence, root external colour, root cortex colour, root pulp colour, root epidermis texture and flowering characteristics (Allem, 2002).

2.3 Characteristics of Mkombozi and TMS 4 (2) 1425

2.3.1 Mkombozi

Mkombozi is a half sib of 92/0099S2 (SM) and was developed by the East Africa Root Crops Research Network (EARRNET). It contains Light green Color of Apical Leaves, light green leaf color and reddish green petiole color. The growth habitat of stem is straight; the color for the outer root skin is Green-yellowish while the inner root skin color is Light green and possessed Cream root fresh color. It is a CMD resistant and CBSD tolerant cassava variety. A detail of its characteristics is shown in Appendix 1.

2.3.2 TMS 4 (2) 1425

This is among the tropical *Manihot* species developed by IITA. It is a pedigree of 58 308 x Oyarugba Funfun with Purplish green Color of Apical Leaves, green purple leaf color, silver stem colour and greenish red petiole color. The growth habitat of stem is zig- zag, the color for the outer root skin is cream while the inner root skin color is white and possessed white root fresh color. It is a CMD resistant cassava variety. A detail of its characteristics is shown in Appendix 2.

2.4 Reproduction in Cassava

Cassava can be propagated either by stem cuttings or sexual seeds, although the former is the most common practice by farmers for multiplication and planting. For plant breeding and under natural conditions, propagation by seeds is common. Farmers are known to occasionally use spontaneous seedlings for subsequent planting that is a starting point for generating useful genetic diversity (Alves, 2002). Cassava is monoecious with male flowers occurring near the tip of the inflorescence while female flowers occur close to the base and they open 10-14 days before the male flowers on the same branch (IITA, 1990; Ekanayake *et al.*, 1997). This protogyny phenomenon favors cross pollination in cassava,

but self-pollination can occur when male and female flowers on different branches on different plants of the same genotype open simultaneously (Jennings and Iglesias, 2002). Variation in flowering occurs between cultivars and in some, flowering is frequent and regular while in others it is rare or non-existent. Environmental factors such as temperature and photoperiod influence flowering. Synchronisation of flowering remains a challenge in cassava breeding (Ceballos *et al.*, 2004).

2.5 Seed Germination

After pollination and subsequent fertilization, the ovary develops into young fruits, which takes about 70-90 days to mature. The fruit contains three endocarp locules, each with one seed. When the fruit is dry the locules split to release the seeds. After maturity and harvesting, cassava seeds often have a physiological dormancy period of a few months that is common in *Manihot* species. Storing seeds at room temperature for two to three months in pest free and pathogen free conditions has been recommended by CIAT, (2004). Since at Ibadan, Nigeria, soil temperatures of 30-35°C and high soil moisture content are common, seeds are planted directly in the field at the International Institute for Tropical Agriculture (IITA) and these temperatures have been reported to be optimum for cassava germination (IITA, 1980).

2.6 Production Constraints

Cassava production is affected by both biotic and abiotic constraints. Other factors affecting cassava production include shortage of appropriate improved cultivars with high genetic potential, long growth cycle, inadequate availability of disease-free planting material, post-harvest physiological deterioration of roots, nutritional deficiency for people who solely depend on cassava, cyanide content, diseases and pests (IITA, 1990). The strongly out-crossing and monoecious nature of cassava (mediated by protogyny)

make it difficult to develop appropriate stocks for classical genetic studies (Fregene *et al.*, 1997). From a breeding point of view, this limits and makes it difficult to develop appropriate cultivars.

The most important diseases affecting cassava production in Tanzania include CMD, CBSD, CBB and root rot (*Phytopthora* spp*)*. These diseases have been identified as the major biotic constraints to cassava production in Tanzania. CMD occurs in all cassava growing areas in Africa and is caused by the East African cassava mosaic virus (EACMV), African cassava mosaic virus (ACMV) and (Uganda variant) UgV (Legg, 1999; Otim-Nape *et al.*, 2000). Bacterial blight is common in the wet and humid areas of east, central, south and West African countries. It is also found in South America and Asia and this disease is caused by *Xanthomonas campestris pv manihoti.* Whilst cassava green mite (*Mononychellus tanajoa* Bondar) and cassava mealybug (*Phenaccocus manihoti*) are the major arthropod pests (Yaninek, 1994). Yield losses can be as high as 95% depending on the time of infection by any one or more of the diseases (Storey and Nichols, 1938; Brian and John, 1940; Legg, 1999; Hillocks and Thresh, 2000).

2.7 Cassava Brown Streak Disease

Cassava Brown Streak Disease (CBSD) is a viral disease that impacts on cassava tuberous root quantity and quality (Hillocks and Thresh, 2000). It causes a brown, dry, corky necrosis within the starch bearing tissues, sometimes accompanied by pitting and distortion that is visible externally and renders the root useless (Hillocks, 1997). It also causes foliar chlorosis and sometimes stems lesions. CBSD is considered the major threat to food security in worst affected areas of coastal, eastern and southern Africa.

2.7.1 Causal agents

CBSD is caused by Cassava Brown Streak Virus (CBSV) and Uganda Cassava Brown Streak Virus (UCBSV), Genus: *Ipomovirus*; family *Potyviridae* and it is spread both through propagation of infected cuttings and by a whitefly vector, *Bemisia tabaci* (Maruti *et al.*, 2005; Ntawuruhunga and Legg, 2007). The shape and size of the virion has recently been confirmed by Mbanzibwa *et al.* (2009a). The presence of pinwheels and slightly flexuous rod-shaped particles associated with CBSD symptoms was also detected in isolates from cassava samples collected from Kenya, an indication that a Potyvirus is associated with CBSD infection (Were *et al.*, 2004a, 2004b; 2007). Winter *et al.* (2010) characterized CBSV at the molecular level by comparing the complete coat protein encoding sequence of isolates from the coastal lowlands of Tanzania and Mozambique to those from Uganda, Kenya and Malawi, and detected two distinct virus species.

2.7.2 Virus transmission and spread

The Amani work done by Storey (1936) demonstrated that CBSV was perpetuated through vegetative propagation and transmissible through grafting. Work on pathological studies by Storey (1939) led to suspicions that whitefly (*Bemisia* spp) was the vector of the virus in the field. Vector transmission of CBSV by whitefly (*Bemisa tabaci* Gennadus) was confirmed by Maruthi *et al.* (2005) during their transmission studies although the transmission efficiency was low (22%). It was first reported and distinguished from CMD in Tanzania during the 1930s (Storey, 1936). Since CBSD was first reported at Amani in Tanzania (Storey, 1936), a number of surveys have been conducted that confirmed the presence of the disease at high incidences. A survey carried out by IITA and TRTCP in 2009 identified the Lake Zone of Tanzania as a severely affected area (Kasele *et al.*, 2009). A disease survey that was conducted in Tanga Region of coastal lowlands of Tanzania revealed crop losses of up to 74% (Muhana and

Mtunda, 2002) but in severely affected areas, entire fields were destroyed leading to 100% yield losses. In recent years it has spread to many areas in other countries such as Uganda, DRC, Burundi and Rwanda where it is threatening cassava production and food security (Alicai *et al.*, 2007; Ntawuruhunga and Legg, 2007).

2.7.3 Disease symptoms

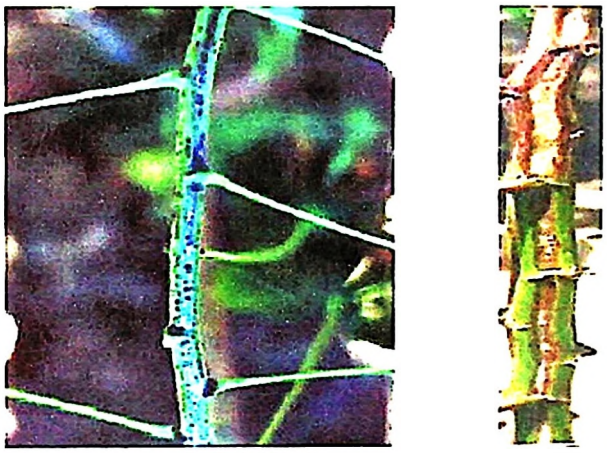
2.7.3.1 Leaf symptoms

Symptoms of CBSD on the leaf can take different forms depending on the cassava variety, the age of the plant and conditions under which it is grown (Nichols, 1950). Often, the most obvious symptom of CBSD is chlorosis (yellowing) on the lower older leaves between and along the secondary y veins - veinal chlorosis (Plate 1). During advanced stages much of the leaf lamina may be affected. The diseased leaves remain attached to the plant for a long time. Unlike CMD in which leaf shape is distorted, the leaf shape remains intact in CBSD. The distinguishing feature for CBSD infection is that the lower, older leaves are usually affected (Nichols, 1950).

Plate 1: Leaves infected by cassava brown streak virus showing symptoms on the lower and older leaves (Courtesy J.P. Legg)

2.7.3.2 Stem symptoms

On young, maturing stem tissue, purple brown lesions may be observed on the exterior surface of the stem, particularly on sensitive varieties. Necrotic lesions in the leaf scars on the stem appear after the plant has shed leaves due to normal senescence. In severe infections, lesions develop to kill dormant auxiliary buds. When auxiliary buds die, a general shrinkage of the node occurs causing the death of internodes, leading to die back (Plate 2 a) and (Plate 2 b).



(a) (b)

Plate 2: (a) Brown lesions on young stem of plant infected with cassava brown streak virus (Courtesy J.P. Legg) (b) advanced stage of disease infection causing die back (Courtesy G. Rwegasira)

2.7.3.3 Root symptoms

The appearance of root symptoms varies with variety. However, for many varieties the most damaging symptom that renders the roots economically unmarketable is the appearance of yellow or brown, dry or corky necrotic rot in the internal tissues of the storage roots. The corky necrosis sometimes appears together with blue or black streaks. The lesions seem to remain discrete, but in sensitive varieties almost the entire tissues of the starch storage root may be affected. In the advanced stages of infection, soft rot may

appear due to secondary infection by secondary organisms (Nichols, 1950; Hillocks and Thresh, 2000). Symptoms on the outside of the roots are variable; they may appear as constrictions across the root or pits and fissures along the roof surface. Sometimes the root appears healthy on the outside with no obvious constrictions or size reduction but upon cutting, they are found to be necrotic (Plate 3 a) and (Plate 3 b).

(a) (b)

Plate 3: (a) Roots infected with cassava brown streak virus showing root constrictions (b) Plant roots infected with cassava brown streak virus showing root necrosis (Courtesy H. Kulembeka)

2.7.4 Economic importance of cassava brown streak disease

The main economically significant effect of CBSD is root necrosis. Yield loss due to CBSD is more pronounced and obvious due to loss of root quality rather than root weight. Comprehensive studies on the effect of CBSD on yield and root quality of cassava was done in Tanzania by Hillocks *et al.* (2001). Susceptible cultivars have been reported to sustain root yield loss of up to 70% per plant (Hillocks *et al.*, 2001). Using these estimates and production data of FAO (2006), it was estimated that about 35 000 tonnes of cassava is lost annually in the coastal regions due to CBSD. In economic terms it is estimated that US$ 16.5 million are lost per year due to CBSD (Kanju *et al.*, 2007).

In Malawi, yield losses of up to 60% were reported and the range of loss at an average disease incidence of 40% translated to 137 000-172 000 tonnes per year (Gondwe *et al.*, 2003). This loss in monetary terms is estimated at US$ 6-7 million per year. All studies indicated that losses were largely due to root necrosis, and in some cases, constriction of roots. In cases where roots were severely necrotic, the quality of roots was significantly affected and roots rendered unmarketable, leading to total losses. In Tanzania and Malawi, farmers indicated that roots with mild root necrosis were sometimes eaten but extensively damaged roots, whether processed or unprocessed, had a bad taste and were usually thrown away (Gondwe *et al.*, 2003; Mkamilo and Hillocks, 2006). Flours from heavily necrotic roots are discoloured and have a bitter taste that makes it unpalatable (Gondwe *et al.*, 2003).

2.7.5 Disease management

It has been advocated that the development and use of resistant varieties could potentially form the basis of a sustainable management strategy for cassava diseases (Asiedu *et al.*, 1994; Mahungu *et al.*, 1994; De Vries and Toenniessen, 2001).

2.8 Integrity of the Mapping Population

Genomics tools have been used to verify crosses (Supawadee and Sompong, 2009; Terzic *et al.*, 2006; Gomez *et al.*, 2008). Tools utilized are mainly molecular markers capable of detecting false progenies in a hybrid or selfed line. The most desirable marker tools are co-dominant and capable of showing heterozygous polymorphic alleles in parents. Of the major DNA marker types, restriction fragment length polymorphism (RFLP) and simple sequence repeat (SSR) markers are co-dominant (Gomez *et al.*, 2008). SSR markers are however more convenient as they require smaller quantities of genomic DNA (Powell *et al.*, 1996). SSR markers have been used for the verification of inter-specific crosses

(Terzic *et al.*, 2006) as well as verification of purity of intra-specific crosses in crops like cotton (Dongre and Parkhi, 2005; Asif *et al.*, 2009), maize (Salgado *et al.*, 2006) and rice (Tamilkumar *et al.*, 2009). They can be used to identify selfed progenies in crosses (Gomez *et al.*, 2008) as well as false hybrids in populations by observation of banding patterns not consistent with Mendelian segregation laws (Salgado *et al.*, 2006). Polymorphism in SSR markers in diploid crops such as cassava can be characterized by two, three or four alleles (multi-allelic), with a maximum of two in each parent; thus SSR markers are good for checking the integrity of the mapping population.

2.9 Molecular Markers and Construction of Linkage Map

Molecular markers are useful and powerful tools for genetic improvement of crops and can be applied in different aspects of plant breeding and germplasm management (Thottapilly *et al.*, 2000; Kulembeka, 2010). They can be used to assess genetic diversity of individuals in a population as well as to elucidate genetic organization in different crops (Mace *et al.*, 2006). In the genetic improvement of crops, MAS is an efficient strategy and is based on linkage relationships between markers and traits. A genetic linkage map is a common intermediary for determining the linkage between markers and traits. Genetic linkage maps have been used for genetic studies in many crops like wheat (*Triticum aestivum*) (Mustafa *et al.*, 2004), Sorghum (*Sorghum bicolor*) (Emma *et al.*, 2009) and Rice (*Oryza sativa*) (Akkareddy, 2010).

In cassava different types of molecular markers such as RFLP, RAPD, Amplified Fragment Length Polymorphism (AFLP), Microsatellites or SSR, isoenzymes, Expressed Sequence Tag derived SSR (EST-SSR), Sequence Related Amplified Polymorphism (SRAP) and SNP have been deployed to construct the cassava molecular genetic linkage map. The first genetic linkage map of cassava contained 132 RFLPs, 30 RAPDs, 3

microsatellites and 3 isoenzymes and spanned 931.6 cM with 20 linkage groups (Fregene *et al.*, 1997). It provided initial tools for genetic analysis of several important traits of cassava. The second genetic linkage map of cassava was developed by Mba *et al.* (2001). The map consisted of only 36 additional markers that were placed on the cassava RFLP framework map in order to increase density. The map possessed 18 linkage groups. RFLPs could not be transferred readily to national programs of the developing countries since it was expensive and laborious; again facilities for the radioactive procedures were not available in many laboratories. The SSRs were preferred to overcome the problems of RFLP, since they are highly informative due to the number and frequency of alleles detected and can distinguish closely related individuals, easy to implement in most laboratories and amenable to high throughput marker genotyping (Okogbenin *et al.*, 2006). Okogbenin *et al.* (2006) developed the first SSR based linkage map of cassava; the map contained 122 SSR markers spanning 1236.7 cM on 20 linkage groups with an average marker density of 17.92 cM.

After the construction of the first SSR-based linkage map in cassava by Okogbenin *et al.* (2006), other SSR based maps were developed. In 2010, a genetic linkage map of cassava was generated with a total of 355 molecular markers where 231 FLPs, 41 SSRs, 48 SRAPs and 35 EST – SSRs were included (Chen *et al.*, 2010). This map contained 18 linkage groups and spanned 1707.9 cM with an average marker interval of 4.81 cM. Other cassava SSR linkage maps were published in 2011, by Kunkeaw *et al.* (2010b) and Sraphet *et al.* (2011). The map by Kunkeaw *et al.* (2010b) showed 20 linkage groups consisting of 211 markers (56 EST-SSR and 155 SSR) spanning 1179 cM, with an average distance between markers of 5.6 cM and about 11 markers per linkage group. Sraphet *et al.* (2011) published a map, which had 510 markers and was 1420. 3 cM in length distributed on 23 linkage groups. The map had a mean distance between markers

of 4.54 cM. Apart from cassava, SSR-based maps have been constructed in many plants such as wheat (Torado *et al.*, 2006), barley (Varshney *et al.*, 2007), tobacco (Bindler *et al.*, 2011), chickpea (Gaur *et al.*, 2011) and soyabean (Hwang *et al.*, 2009). SSRs were found to have limitations, since they were not sufficiently dense to provide the genomic coverage necessary for the dissection of complex traits. Moreover, automation in high-throughput genotyping of SSRs was not very common. Currently, SNPs are now proposed as the most useful markers, especially for the construction of highly saturated maps, which will have very wide and novel applications in plant genomics.

SNPs and small insertions and deletions (indels) represent the most frequent kind of naturally occurring genetic variation in population (Cho *et al.*, 1999; Ferguson *et al.*, 2011) and normally occur at a frequency of one SNP every 100–300 bp (Edwards *et al.*, 2010). Although SNPs are biallelic compared to SSRs that are multi-allelic, they are highly amenable to ultra-high throughput genotyping technologies compared to SSR (Appleby *et al.*, 2009). SNP genotyping can be performed using predefined SNP arrays or *de novo* using techniques such as 'genotyping- by-sequencing' (Elshire *et al.*, 2011).

SNP markers have been used in generating genetic linkage maps in several crop species including Jatropha (Wang *et al.*, 2011), apple (Antanaviciute *et al.*, 2012) and chickpea (Rashmi *et al.*, 2012). The jatropha SNP map contained 506 markers (216 microsatellites and 290 SNP from ESTs) while 2272 and 1063 SNP markers were found in the apple and chickpea genetic linkage maps respectively. Only one SNP-based genetic linkage map of cassava has been published (Rabbi *et al.*, 2012). This map contained 568 markers (434 SNPs and 134 SSRs) and spanned 1837 cM across 19 linkage groups with average distance between markers of 3.4 cM.

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Cassava Plant Materials

Parents with contrasting response to CBSD infection were selected to generate an F1 population for this study. Mkombozi was selected as the female parent and is field resistant to CBSD. TMS 4(2)1425 was selected as the male parent and is susceptible to CBSD. Stakes of the parents were obtained from Ukiriguru and Maruku Agricultural Research Institutes which were multiplied by farmer groups in their fields during the Great Lakes Cassava Initiative project (GLCI).

3.1.1 Genetic crosses

Crossing blocks consisting of the two parents were planted at Ukiriguru and Maruku experimental stations in Mwanza and Bukoba respectively in the 2010 growing season. Genetic crosses were performed by hand pollination according to Kawano (1980) and IITA (1990). Plants started to flower six months after planting (MAP) and pollination was performed by rubbing anthers with pollen on the stigma of female flowers. After pollination, flowers were bagged with muslin bags to prevent honey bees or other insects from pollinating opened female flowers. Pollinated flowers were labelled to indicate each of the cross combinations and mature seed were harvested and collected 70-90 days after pollination.

3.1.2 Seed germination and establishment of seedlings

In April 2011 about 800 seeds were germinated and grown in seed trays containing sterilized forest soil in the screen house at SRI-Kibaha. Since temperatures up to 35°C are required for seed germination of cassava (IITA, 1980; Ellis *et al.*, 1982). seed trays were

put in the screen house to ensure optimum temperature for seed germination. In June 2011, 34 days after sowing the seeds, seedlings were transported and transplanted in the field at Makutupora experimental station- Dodoma. The seedlings of about 20-25 cm long were planted at spacing of 1.0 m x 1.0 m, 50 rows with 10 plants each row with irrigation and NPK fertilizer to maximize the multiplication rate.

3.1.3 Plant sample collection

Young leaf samples from 350 individual plants were harvested from the field at Makutupora and placed in labeled aluminum foil packets and immediately kept in a cool box with ice blocks after which were transported to International Livestock Research Institute - Biosciences eastern and central Africa (ILRI-BecA) hub for DNA extraction. The leaves were then stored at -80 °C deep freezer until DNA was extracted.

3.1.4 Extraction of DNA

Extraction of DNA from each of the 325 leaf samples was done at ILRI-BecA hub Nairobi, Kenya. DNA was extracted using a modified mini preparation extraction protocol of Dellaporta (Dellaporta *et al.*, 1983).

3.1.5 Determination of DNA quantity and quality

Quantity and quality of DNA samples were determined using a spectrophotometer (NanoDrop. ND-1000 Spectrophotometer) and agarose gel electrophoresis. The quality and quantity of DNA was checked by running samples on a 0.8% (w/v) or 0.8g agarose gel in 100ml 1x Tris-boric EDTA (TBE). The agarose solution was brought to the boil in a microwave oven. The product was cooled slightly and 5µl of Gel Red solution added. This solution was then poured into a gel tank with comb and left for 20-30 min to polymerize. The gel was then placed in a running tank containing 1XTBE buffer.

Approximately 3µl DNA and 3µl of loading dye were mixed and then loaded into each well. The running tank was then connected to power at 120V for 30 minutes. Ultraviolet (UV) illumination photograph was used to check the sample DNA quality and quantity and after determination of DNA concentration and quality the DNA was stored at -20°C.

3.2 Screening for SSR Polymorphic Markers and Genotyping

To identify polymorphic markers, 38 SSR markers were screened against the two parental genotypes. These were selected from a subset of 48 SSR markers developed at CIAT (Chavariagga-Acquirc *et al.*, 1998; Mba *et al.*, 2001). Amplification of DNA samples was carried out on a GeneAmp®PCR System 9700 Base Module (Applied Biosystems Inc) PCR machine as described by Mba *et al.* (2001). PCR amplified product were visualized on a 2% (w/v) or 2g agarose gel (Invitrogen, UK) as described in section 3.1.5 above for checking the quality and quantity of DNA. Approximately 2µl PCR product, 1µl of distilled water and 2µl of loading dye were mixed and then loaded into each well at the extreme end wells.

3.2.1 Parental genotyping

Before PCR products were run on the ABI 3730 capillary sequencer for fragment analysis, marker panels comprising of SSRs with no overlapping allele sizes or being differently labeled were used to design co-loading primer sets for genotyping. Co-loading markers were selected and put into groups of four based on their dye label. About 2 µl of each fluorescently-labeled PCR product (i.e. 8 µl total products for four PCR products) were combined in one new plate. Products were briefly vortexed and centrifuged at 3500 rpm for 30 s. Then 900µl HI-DI form amide and 20µl LIZ was mixed and 9 µl formamide-standard mix was added into each well of a new, empty PCR plate. One µl of the PCR product mixture was added to 9 µl formamide-standard mix.

The pooled plate was vortexed and centrifuged at 3500 rpm for 60 s. Samples were denatured at 95°C for 3 min, placed on ice for 5 min, centrifuged, and then loaded into the ABI 3730. DNA fragments were analyzed for polymorphism (fragment size variation) between the parents using Gene Mapper software version 4.1.

3.2.2 SSR Genotyping of F1 progenies

Ten SSR markers that showed polymorphism amongst the parents were used to genotype the entire mapping population (Table 1) - SSRY38, SSRY51, SSRY52, SSRY63, SSRY102, SSRY119, SSRY147, SSRY155, SSRY169 and NS193. Amplification reactions were set using the optimized PCR conditions for each marker in 96 PCR plates. PCR amplifications were carried in an auto-Lid Dual 384-Well GeneAmp® PCR System 9700 (Applied Biosystems Inc). The thermo cycling as well as visualization of PCR amplification product was performed as described in section 3.1.5 above.

Amplified SSR fragments were separated by capillary electrophoresis using an ABI 3730 sequencer and allele sizes were scored using GeneMapper. Co-loading markers were selected and put into groups of three, two and one based on their dye label. Co-loading group 1 contained SSRY38 (FAM), NS193 (PET) and SSRY147 (VIC), co-loading group 2 consisted of SSRY119 (FAM) and SSRY51 (PET), co-loading group 3 contained SSRY119 (FAM) and SSRY63 (PET), co-loading group 4 consisted of SSRY102 (FAM) and SSRY52 (PET) and co-loading group 5 had only SSRY155 which was NED. About 2 µl of each fluorescently-labelled PCR product (i.e. 2-6 µl total products for one, two and three PCR products) were combined in one new plate and were prepared for capillary electrophoresis.

Table 1: SSR primers, sequences and annealing temperature for each SSR marker used for assessing the integrity of mapping population

Marker	Left primer (Forward 5' - 3')	Right primer (Reverse 5' - 3')	Repeat motif	Expected allele size (bp)	Reference author	Annealing temperature (°C)
SSRY38	GGCTGTTCGTGATCCTTATTAAC	GTAGTTGAGAAAACTTTGCATGAG	(CA)17	122	Mba *et al.* (2001)	57
SSRY51	AGGTTGGATGCTTGAAGGAA	GGATGCAGGAGTGCTCAACT	(CT)11CG(CT)11(CA)18	298	Mba *et al.* (2001)	57
SSRY52	GCCAGCAAGGTTTGCTACAT	AACTGTCAAACCATTCTACTTGC	(GT)19	266	Mba *et al.* (2001)	57
SSRY63	TCAGAATCATCTACCTTGGCA	AAGACAATCATTTTGTGCTCCA	(GA)16	290	Mba *et al.* (2001)	55
SSRY102	TTGGCTGCTTTCACTAATGC	TTGAACACGTTGAACAACCA	(GT)11	179	Mba *et al.* (2001)	57
SSRY119	AACATAGGCATTAAAGTTTGGCA	GCAAATGTGTTTTCAATATAAGGC	(GA)8(G)3(GA)3(N)4(GA)3(N)32 (A)5(GT)2 (N)6(GT)3	155	Mba *et al.* (2001)	57
SSRY147	GTACATCACCACCAACGGGC	AGAGCGGTGGGGCGAAGAGC	n/a	113	Mba / Fregene (CIAT)	57
SSRY155	CGTTGATAAAGTGGAAAGAGCA	ACTCCACTCCCGATGCTCGC	n/a	158	Mba / Fregene (CIAT)	59
SSRY169	ACAGCTCTAAAAACTGCAGCC	AACGTAGGCCCTAACTAACCC	GA19A3GAA2	100	Mba *et al* (2001)	57
NS193	TTGGGGGCTTTAAGTTGTTG	AAAGCCCATCCCCTCTATGT	n/a	258	DNA-Angela Zarate/Martin Fregene	57

n/a = not available

3.2.3 Integrity of the mapping population

Sizing of fragments, allele calling and scoring of polymorphic markers were performed using GeneMapper Version 4.1 computer software (SSRY155 is one of the SSR marker used to score the allele size as shown in Fig. 1). The size of each marker was recorded.

True crosses, off-types (showing additional alleles, not present in parents), and selfs (showing an excess of homozygote) were identified. The true cross progenies were identified by examining the identity of each allele at each locus, and then considering the results of all loci together.

SSRY 155

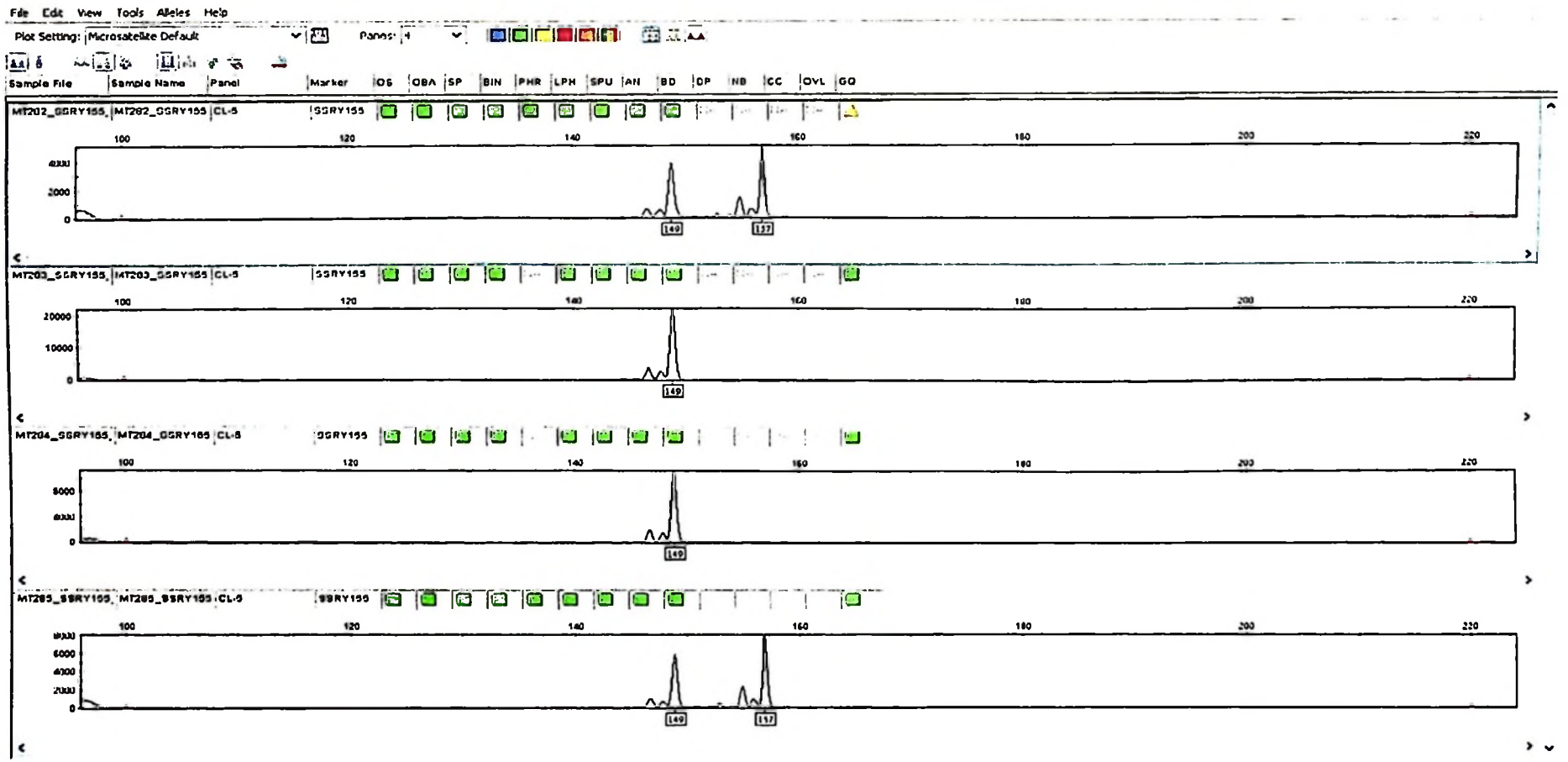


Figure 1: Image of allele (for MT 282, MT 283, MT 284 and MT 285 genotypes) scored using marker SSRY 155 for the evaluation of the integrity of the mapping population

3.3 SNP Genotyping

Genotyping of 157 F1 progenies from Mkombozi x TMS 4(2)1425 was carried out using a reduced representation genotyping-by-sequencing (GBS) approach at the University of Berkeley (USA). A reduced representation approach involved restriction enzyme (*Ape*KI) to digest small sample of DNA, this enzyme recognize and cut the motify (GCWGC). The figments with low reads enable to capture the SNPs.

3.3.1 Restriction digest and ligation

DNA was first digested for 2 h at 75 °C with *Ape*KI enzyme. The *Ape*KI is a methylation sensitive restriction endonuclease that recognizes a degenerate 5 bp sequence (GCWGC), where W is A or T. It creates a 5′ overhang sequence (GWC). Barcoded adapters were then ligated to the sticky ends of the template DNA left by the restriction digestion. This was done by adding 1.66x ligase buffer with Adenosine triphosphate (ATP) and T4 ligase to each well to ligate adapters. Each digested sample in a 96 wells had a different barcoded adapters facilitating multiplexing of all samples into a single sequencing lane. Unlike Elshire *et al.* (2011), Y- shaped adapters were used which ensure that each strand of DNA has a different sequence on each end.

3.3.2 Multiplexing and amplification (pooling samples)

Barcoded adapters were ligated on to the restriction fragments, once the DNAs have unique adapters; they were pooled and size-fractioned. Pooled, cleaned, and concentrated adapter-ligated DNA was run on an agarose gel. Fragment of sizes 400–800 bp were excised and well-removed from any adapter-mers and selecting fragments that will cluster well during sequencing. Well clustered fragments were then PCR amplified. PCR primers complementary to adapter sequences were used in a PCR reaction and thus enrich for properly-ligated DNA fragments which are subsequently sequenced on the HiSeq2000 (Illumina) sequencer.

3.4 Genetic Linkage Analysis

Markers segregating only in the first (female) parent were coded as *lm* x *ll*. Those segregating in the second (male) parent were coded as *nn* x *np*. Both classes of these markers were expected to segregate in the ratio of 1:1. The markers that were segregating in both parents were coded as *hk* x *hk* and were expected to show 1:2:1 segregation ratio. From 157 F1 progenies, twenty six had a large number of missing data (>10%) and were removed from the analysis, leaving 131 progeny. A total of 4351SNPs were obtained from the GBS procedure. Of these, 2783 were excluded from the analysis as they had a large proportion of missing data (>10%), leaving 1568 SNPs.

Genetic linkage analysis was performed using JoinMap software, version 4.1 (Van Ooijen, 2006). Chi-square results for each marker were used to test for significant deviations from Mendelian expected ratios (segregation distortion). Linkage groups were identified using LOD values (logarithm of odds) and the LOD threshold ranged from 3.0 to 7.0 were used to create linkage groups. Regression mapping was used to construct the map using a recombination frequency < 0.4. The recombination frequencies were converted to map distances (cM) using the Kosambi mapping function (Kosambi, 1994).

3.5 Distribution of Markers in Scaffolds

To identify scaffolds associated with markers, SNPs needed to be queried against the cassava genome assembly using BLAST. A cassava genome assembly from whole genome shotgun sequencing is available at Phytozome. Using the SNPs identification code which consisted of a scaffold name followed by the base pair position, the numbers of markers per scaffold used in this study were counted.

CHAPTER FOUR

4.0 RESULTS

4.1 Quantity and Quality of the DNA

Concentrations of DNA ranged from 62.1 to 2950.03ng/µl with an average of 364.04 ng/µl according to nanodrop readings. Extracted DNA was of good quality (Fig. 2) with A_{260}/A_{280} absorbance ratio values ranging from 1.2 to 2.0.

4.2 Screening Markers for Polymorphism

A total of 38 SSR markers were screened to obtain the polymorphic markers against parental genotypes. Of these, 10 primers were polymorphic and were used to provide the information of the integrity of the mapping population, 11 were monomorphic and not informative and 17 markers did not generate good PCR amplification products. Therefore 28 markers were excluded from genotyping.

4.3 Composition of the Mapping Population

From an initial mapping population of 325 individuals, SSR data revealed 158 progenies from true crosses between Mkombozi and TMS 4(2)1425, 135 off-types or admixtures with allele(s) that were not present in either of the parents, 31 were identified as selfs (Appendix 3) and one genotype (MT 210) was rejected due to many missing data.

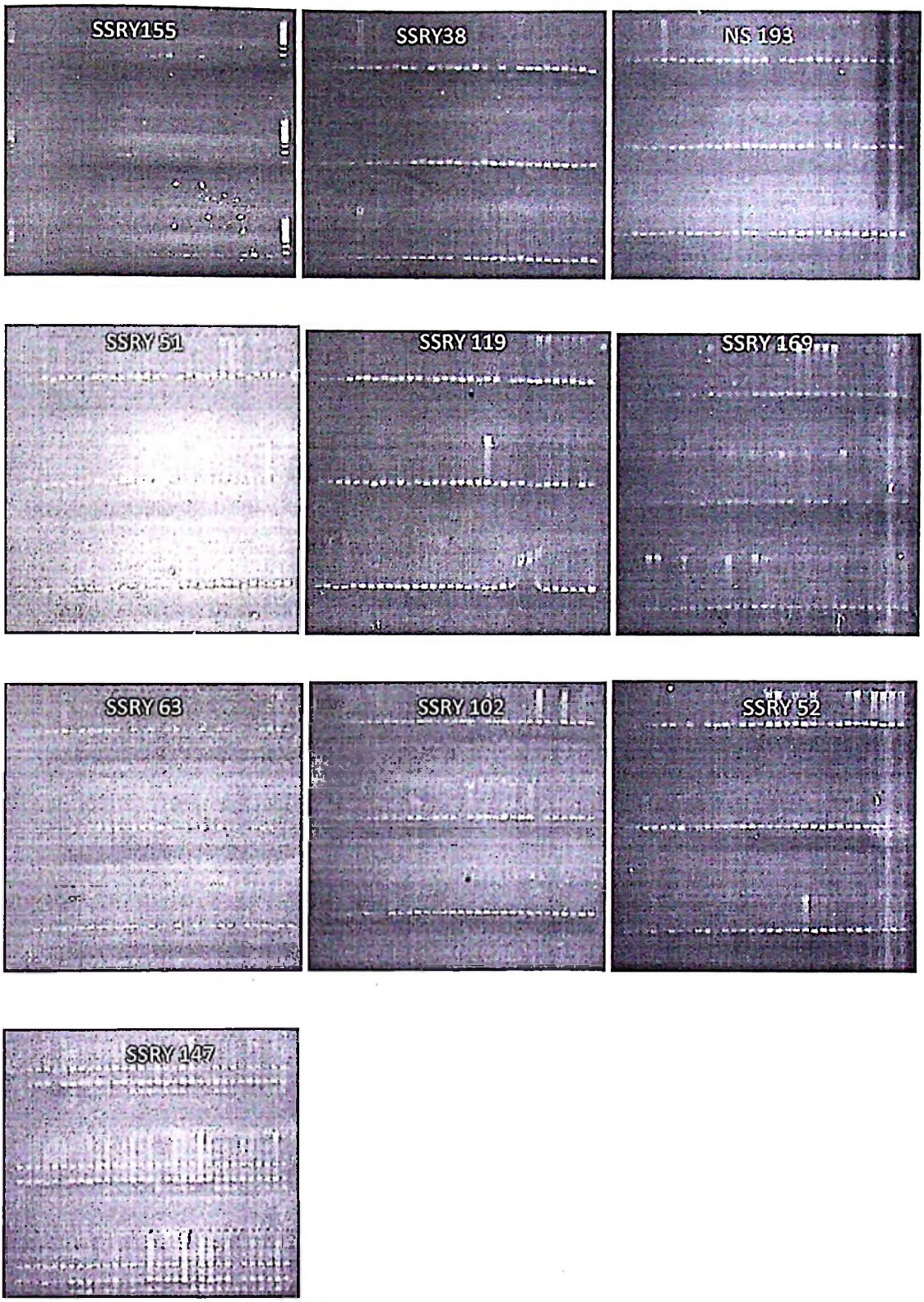


Figure 2: Agarose gel stained images of PCR product to check amplification of SSR primers used to screen the population

4.4 Integrity of the Mapping Population

Marker (NS193) was homozygous in Mkombozi and heterozygous in TMS 4(2)1425 (Table 2). Using this marker all heterozygous progenies can be described as true crosses because there is no possibility of having male selfs. However, the homozygote progenies (250bp/250bp) could be either selfs of Mkombozi or a true cross receiving one 250bp allele from each parent.

Using SSRY38, Mkombozi and TMS 4(2)1425 were heterozygous and homozygous respectively. All homozygous (106bp/106bp) F1 progenies could either be true cross progenies because they contain 106bp from each parent or could be male selfs. In this case, since seed was collected from the female parent, there is no possibility of a male self so all 106bp/106bp progenies were recorded as true crosses. However the heterozygous progenies could not be classified as either true crosses or products of selfing due to the alleles being found in the female parent.

Marker (SSRY63) was particularly informative as both parents were heterozygous and shared one allele. True crosses were those from all four expected combinations and selfed individuals are those possessing 289bp/289bp. SSRY102 was heterozygous in the female parent and homozygous in the male parent. Homozygous (179bp/179bp) progenies could either be true cross progenies or female self. The heterozygous progenies (179bp/181bp) could not be classified as either true crosses or product of selfing due to the alleles being found in female parent, this also required confirmation of other markers.

Using SSRY119, Mkombozi was heterozygous while TMS 4(2)1425 was homozygous. Homozygous progenies were true crosses since they contain the unique male allele while heterozygous progenies could not be easily described as either true crosses or product of

selfing due to the alleles being found in female parent, this also required confirmation from other markers. The individuals with alleles of the 169bp/169bp were selfs from Mkombozi.

The female parent was homozygous at SSRY155 while the male parent was heterozygous In this case the homozygous individuals (149bp/149bp) were either true cross progenies or female selfed progenies. The heterozygous individuals from expected alleles in this marker were regarded as true crosses.

Using SSRY169, the female parent had heterozygous allele and male parent had homozygous allele. The heterozygous individuals (89bp/99bp) were either true crosses receiving 89bp from Mkombozi and 99bp from TMS 4(2)1425 or female selfed progenies. Therefore the heterozygous progenies could not be confidently classified as either true crosses or product of selfing due to the alleles being found in female parent. This required confirmation from other markers. The homozygous putative progenies (99bp/99bp) could be true crosses.

Mkombozi was homozygous for SSRY147 and was heterozygous. All heterozygous progenies (111bp/101bp) were classified as true crosses. The homozygote progenies were not easily classified because they possess a single allele (111bp) common to both parents. They could be female selfs or true crosses receiving the same allele from both parents. At locus SSRY51, the female parent was homozygous for the allele 278bp while the male parent was heterozygous (Table 2). All progenies homozygous for 259bp were classified as female selfs, but the heterozygous progenies with expected combinations from the alleles of the two parents were confidently assigned as true cross progenies.

Both Mkombozi and TMS 4(2)1425 were heterozygous at locus SSRY52 (Table 2). All individuals which were heterozygous for the expected allele combinations were classified as true crosses. Those that were (246bp/263bp were classified as selfs). The homozygote individuals (246bp/246bp and 263bp/263bp) were described as purely female selfs.

Table 2: Size of parental alleles used to determine the integrity of the population

Marker	Allele size (bp) in Mkombozi	Allele size (bp) in TMS 4(2) 1425	Additional allele (bp)
NS193	250/250	250/258	256
SSRY38	106/108	106/106	104, 80, 120 and 104
SSRY63	285/289	285/287	Nil
SSRY102	179/181	179/179	Nil
SSRY119	157/169	157/157	Nil
SSRY155	149/149	149/157	Nil
SSRY169	89/99	99/99	97
SSRY147	111/111	101/111	83, 100 and 110
SSRY51	278/278	259/298	279 and 281
SSRY52	246/263	257/267	265

4.5 SNP Markers Segregation Types

Marker segregation type indicated the total number of alleles segregating in the parents of the F1 progenies. Of 1568 markers which were used for linkage mapping, 774 (49%) were segregating only in female parent (Mkombozi), 526 (34%) segregated only in TMS 4(2)1425 (male parent) and 268 (17%) were heterozygous in both parents.

4.6 Segregation Distortion

Markers showing significant deviation from expected segregation ratios at P = 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005 and 0.0001 levels were marked with *, **, ***, ****, *****,

****** and ******* respectively. Of the 1568 markers segregating in the mapping population, 1312 (84%) conformed to expected Mendelian segregation ratios and 18 markers (1%) were severely distorted with deviations of significance level marked with ******* (P = 0.0001) based on chi-square value and were excluded from the analysis. The remaining, 238 markers (15%) with moderate distortion ($0.0005 \leq P \leq 0.05$) were retained in the analysis and were observed not to significantly inflate the map distance or disturb marker order.

4.7 Distribution of SNP Markers on Female and Male Parents

804 markers were distributed on 19 linkage groups (LG1-LG19) of the female parent (Mkombozi). Distribution of markers in female parent spanned 2848 cM with an average marker spacing of 3.7 cM ranging from 1.7 to 6.2 cM (Table 3).

Table 3: Characteristics of marker distribution on linkage groups of the female parent (Mkombozi)

Linkage group	Length (cM)	Number of SNP markers	Unique loci	Number of common markers	cM/ marker	cM/ unique locus
LG1	184.8	64	49	23	2.9	3.8
LG2	153.1	88	56	7	1.7	2.7
LG3	243.1	39	27	15	6.2	9
LG4	254.2	47	37	12	5.4	6.9
LG5	133.4	25	23	4	5.3	5.8
LG6	132.5	45	35	-	2.9	3.8
LG7	55	14	12	7	3.9	4.6
LG8	209.1	66	49	17	3.2	4.3
LG9	178.5	48	35	30	3.7	5.1
LG10	175.2	34	32	12	5.2	5.5
LG11	192.3	53	42	15	3.6	4.6
LG12	186.1	45	37	7	4.1	5
LG13	94.8	30	21	9	3.2	4.5
LG14	137.7	47	32	24	2.9	4.3
LG15	132.6	46	39	6	2.9	3.4
LG16	167.2	30	24	10	5.6	7
LG17	15.3	8	7	2	1.9	2.2
LG18	76.3	37	30	11	2.1	2.5
LG19	126.8	36	32	2	3.5	4
Total/Mean	**2 848**	**804**	**619**	**213**	**3.7**	**4.7**

Key: - = missing data

The linkage group (LG4) was the longest (254.2 cM) and LG17 being the shortest (15.3 cM). The maximum number of markers occurred on LG2 with 88 markers and the minimum number of markers occurred on LG17 with eight markers. The 804 genetic markers were located at 619 discrete positions (unique loci) on the 19 linkage groups of the female parent; hence the average space of the discrete positions was 4.7 cM, ranging from 2.2 cM to 9.0 cM for LG17 and LG3 respectively. Marker distribution on the male parent (TMS 4(2)1425) had a total length of 2483 cM with average marker spacing of 5.2 cM ranging from 1.9 CM to 12.6 cM. The male parent contained a total of 593 markers in 18 linkage groups (Table 4).

Table 4: Characteristics of marker distribution on linkage groups of the male parent (TMS 4(2) 1425)

Linkage group	Length (cM)	Number of SNP markers	Unique loci	Number of common markers	cM/ marker	cM/ unique locus
LG1	277.6	144	98	23	1.9	2.8
LG2	152.5	22	17	7	6.9	9
LG3	173.7	43	34	15	4	5.1
LG4	187.4	28	24	12	6.7	7.8
LG5	37.9	10	9	4	3.8	4.2
LG6	-	-	-	-	-	-
LG7	339.5	27	21	7	12.6	16.2
LG8	110	25	20	17	4.4	5.5
LG9	184.7	59	50	30	3.1	3.7
LG10	165.9	55	41	12	3	4
LG11	122.3	35	28	15	3.5	4.4
LG12	28.3	17	12	7	1.7	2.4
LG13	192	34	33	9	5.6	5.8
LG14	177	33	28	24	5.4	6.3
LG15	37.8	8	8	6	4.7	4.7
LG16	120.5	19	17	10	6.3	7.1
LG17	66	6	5	2	11	13.2
LG18	49.4	17	12	11	2.9	4.1
LG19	60.3	11	11	2	5.5	5.5
Total/Mean	2 483	593	468	213	5.2	6.2

Key: - = missing data

The longest linkage group in this map was LG7 with 339.5 cM and LG12 with 28.3 cM was the shortest. The highest number of markers (144) was observed in LG1 and the lowest number of markers (6 markers) was found in LG17. All 593 markers were located in 468 discrete positions (unique loci) on the 18 linkage groups; therefore the average space of the discrete positions was 6.2 cM, ranging from 2.4 cM for LG12 to 16.2 cM for LG7.

4.8 Comparison of SNP Markers Distribution on Female and Male Parents

A comparison of the female and male parents indicated that there were some differences in terms of number of linkage groups and number of markers used to generate a linkage map. The female parent was having more markers than the male parent and it had 19 linkage groups compared to 18 linkage groups. Based on common markers found in both parents, there were 213 common markers in female and male parents as shown in Table 3 and 4. The number of common markers varied widely between linkage groups and the highest number of common markers (30) was observed in linkage group (LG9) while LGs 17 and 19 had the lowest number of common markers of two.

4.9 Construction of A Genetic Linkage Map

Of the 1568 markers, 1136 markers were mapped on 18 linkage groups of the integrated map shown in Fig. 3 (a-f) which covered a distance of 3 207.9 cM with an average marker spacing of 3.4 cM that ranged from 1.3 cM to 10.1 cM (Table 5). The largest linkage group (LG1) contained 185 markers while the smallest group (LG17) had only 12 markers. The shortest linkage group (LG17) was 68.4 cM while the longest (LG7) was 343.0 cM. There was a large shift of about 62.7 cM between adjacent markers (s11 216-_156 124 and s07 007_68 359) observed in linkage group 3. The 1136 markers occupied 900 discrete positions on the 18 linkage groups, whereby the average space of the discrete

positions was found to be 4.1 cM that ranged between 1.8 cM (LG12) and 12.3 cM (LG7). Interestingly, LG 6 only contained markers that segregated in the female parent, Mkombozi.

Table 5: Main characteristics of the integrated genetic linkage map

Linkage group	Length (cM)	Number of markers	Unique loci	cM/ marker	cM/unique locus
LG1	235.4	185	132	1.3	1.8
LG2	163.3	103	68	1.6	2.4
LG3	240.9	67	51	3.6	4.7
LG4	265.7	63	51	4.2	5.2
LG5	130.5	31	28	4.2	4.7
LG6	-	-	-	-	-
LG7	343	34	28	10.1	12.3
LG8	205.7	74	58	2.8	3.5
LG9	187.5	77	65	2.4	2.9
LG10	173.7	76	62	2.3	2.8
LG11	189.4	73	60	2.6	3.2
LG12	185.8	55	44	3.4	4.2
LG13	152.2	55	47	2.8	3.2
LG14	158.1	56	44	2.8	3.6
LG15	134.3	48	41	2.8	3.3
LG16	157.7	39	34	4	4.6
LG17	68.4	12	11	5.7	6.2
LG18	83.8	43	35	1.9	2.4
LG19	132.5	45	41	2.9	3.2
Total / Mean	3 207.9	1 136	900	3.4	4.1

Key: - = missing data

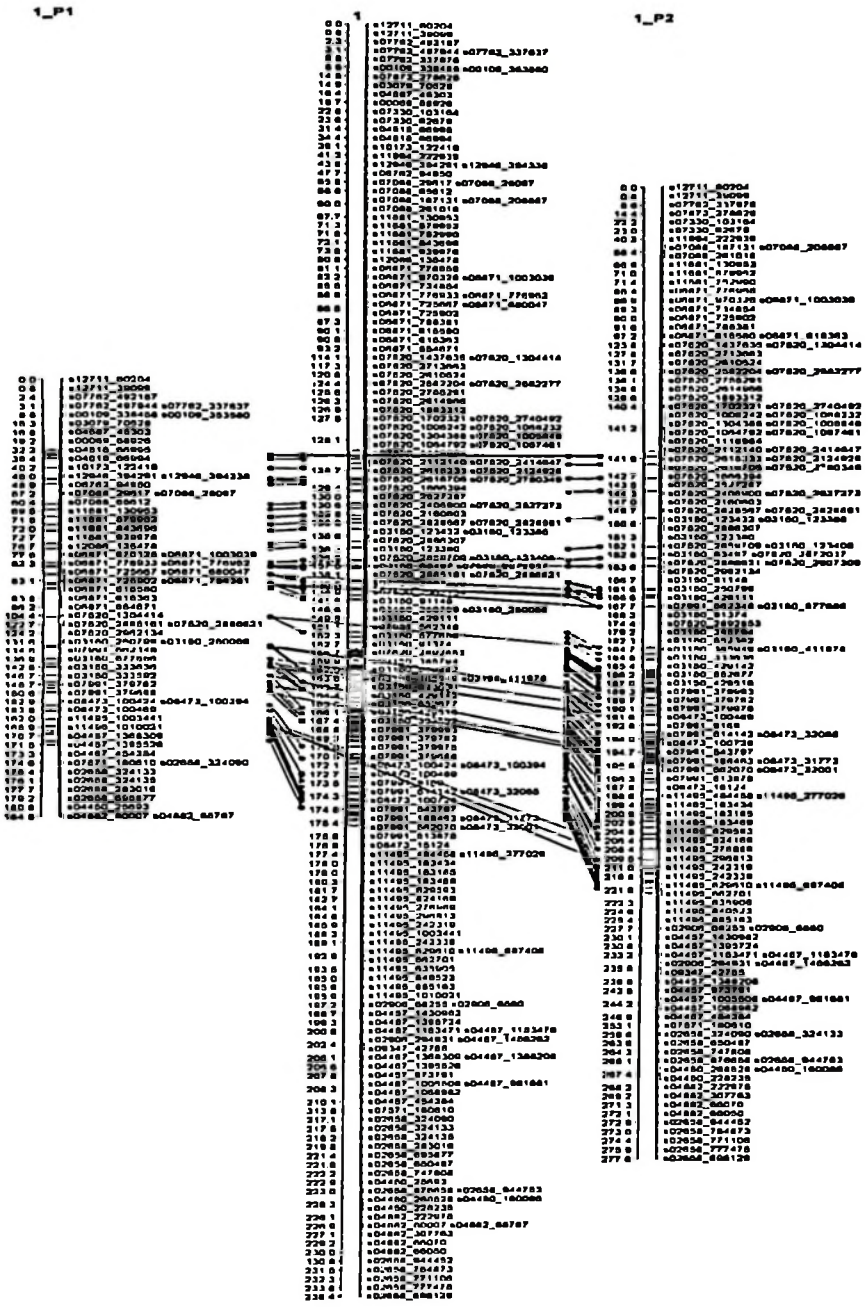


Figure 3 (a): LG1

Figure 3: (a - f): A genetic linkage map of cassava from Mkombozi x TMS 4 (2)1425. Female parental linkage groups on the left (P1), male parental linkage groups on the right (P2) and at the middle is the integrated map which shows the homologous parental linkage groups

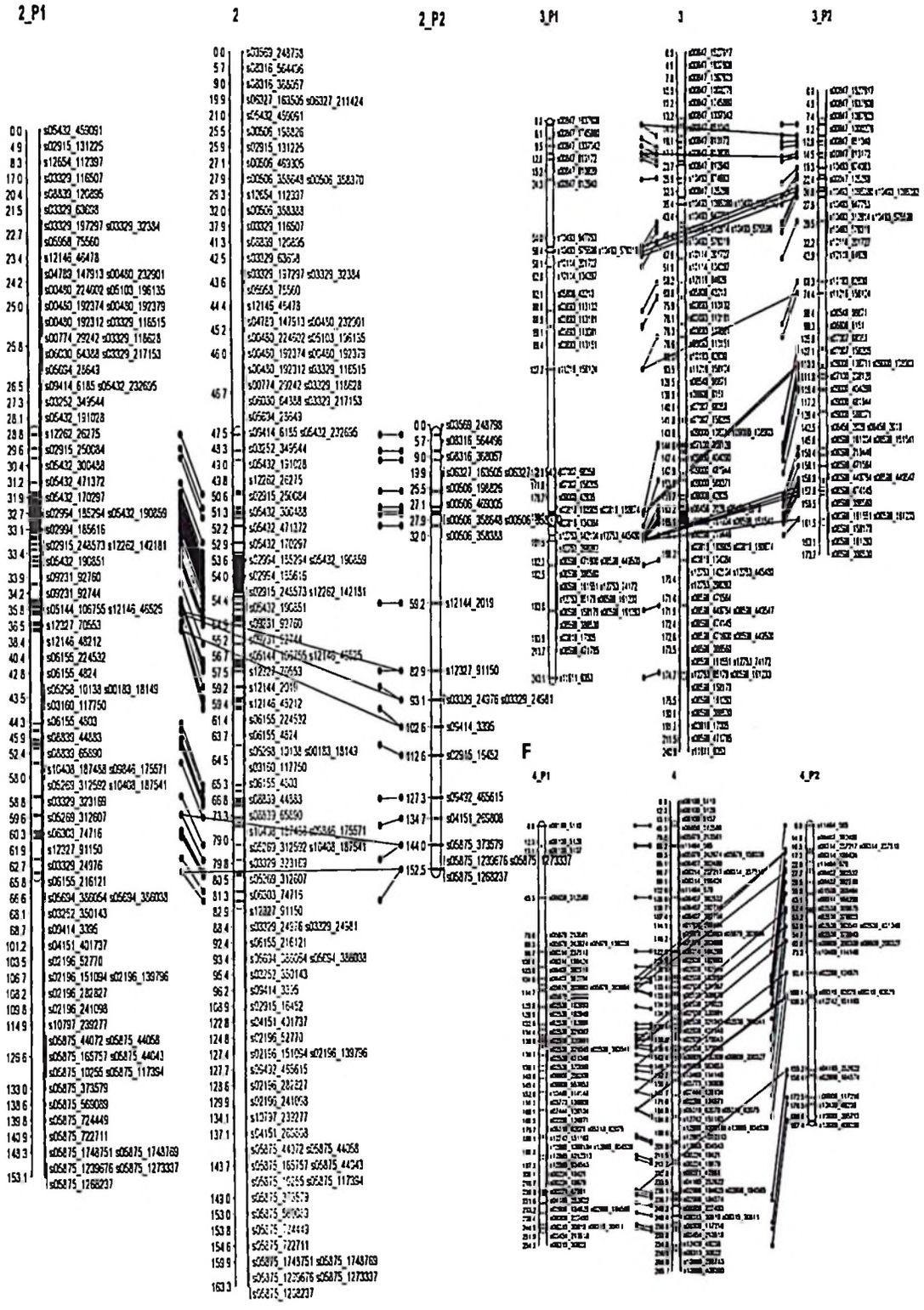


Figure 3 (b): LGs 2, 3 and 4

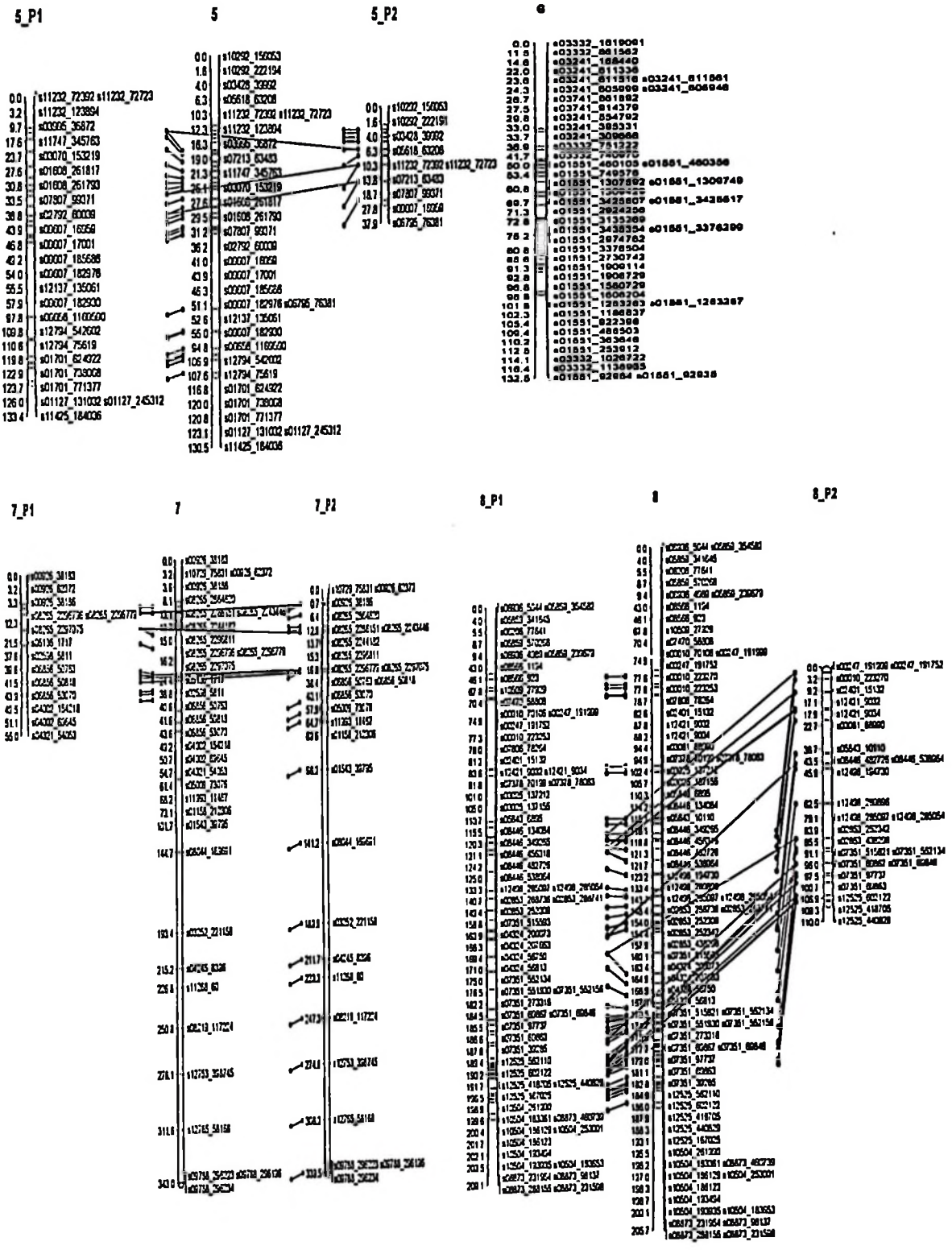


Figure 3 (c): LGs 5, 6, 7 and 8

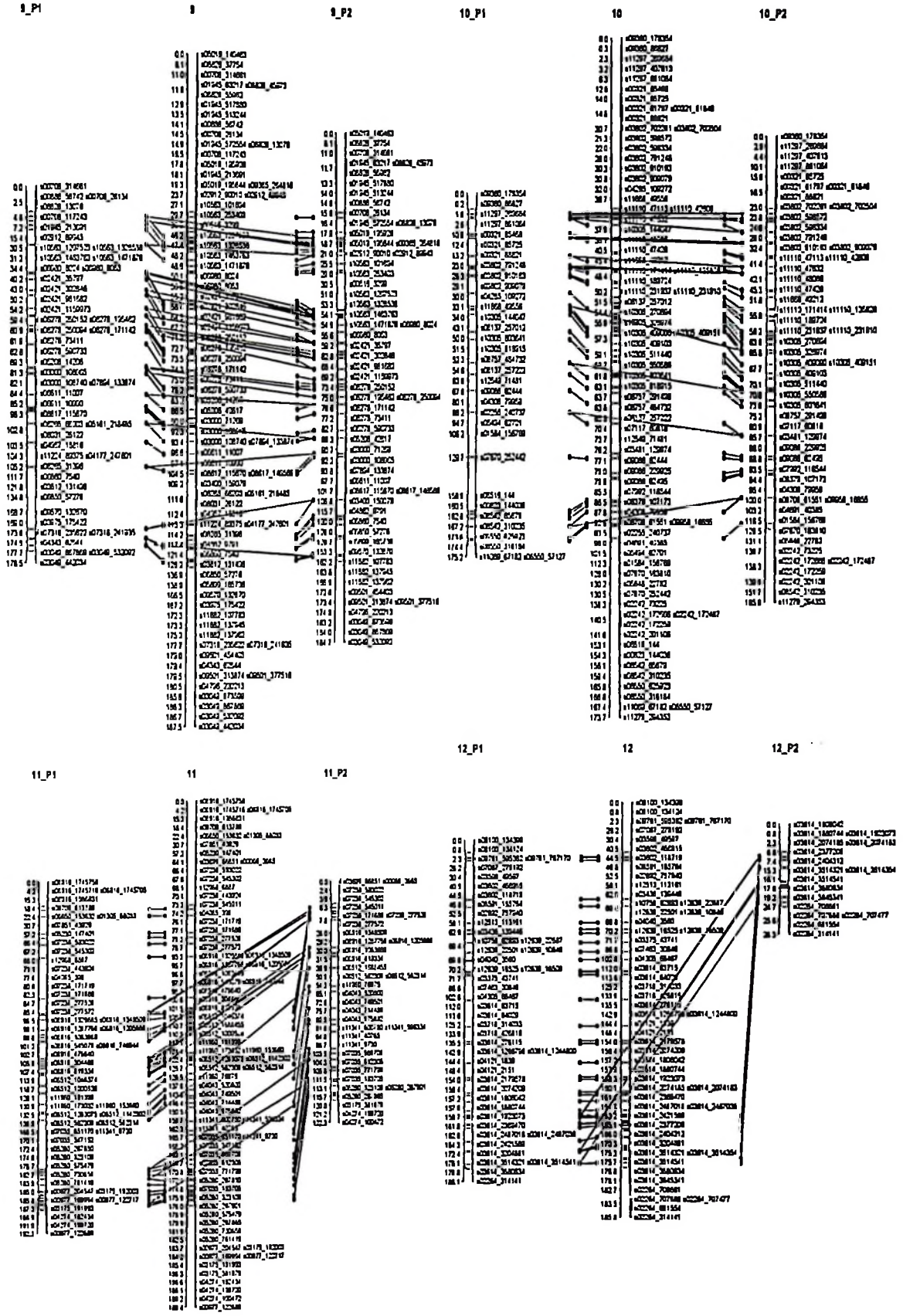


Figure 3 (d): LGs 9, 10, 11 and 12

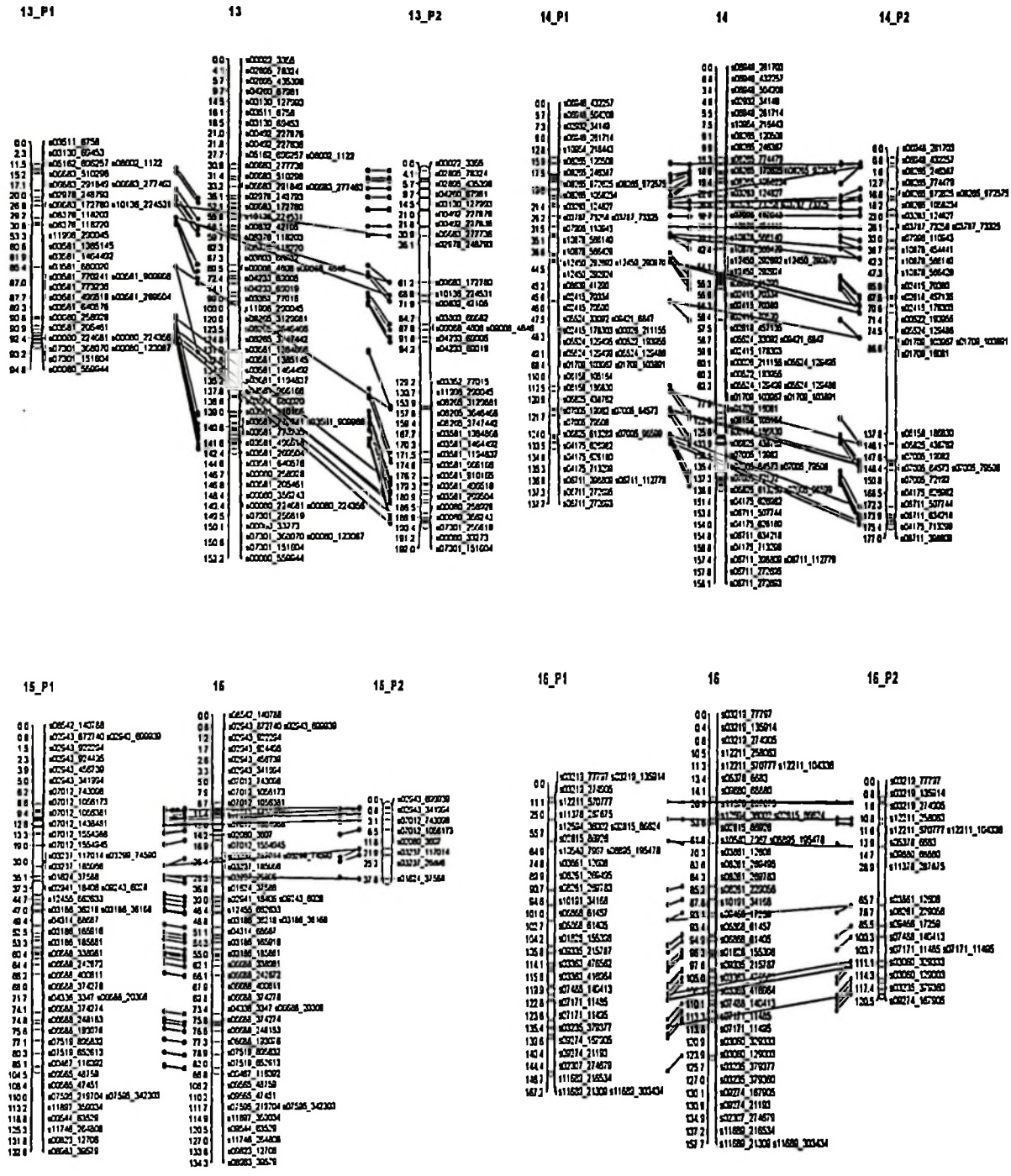


Figure 3 (e): LGs 13, 14, 15 and 16

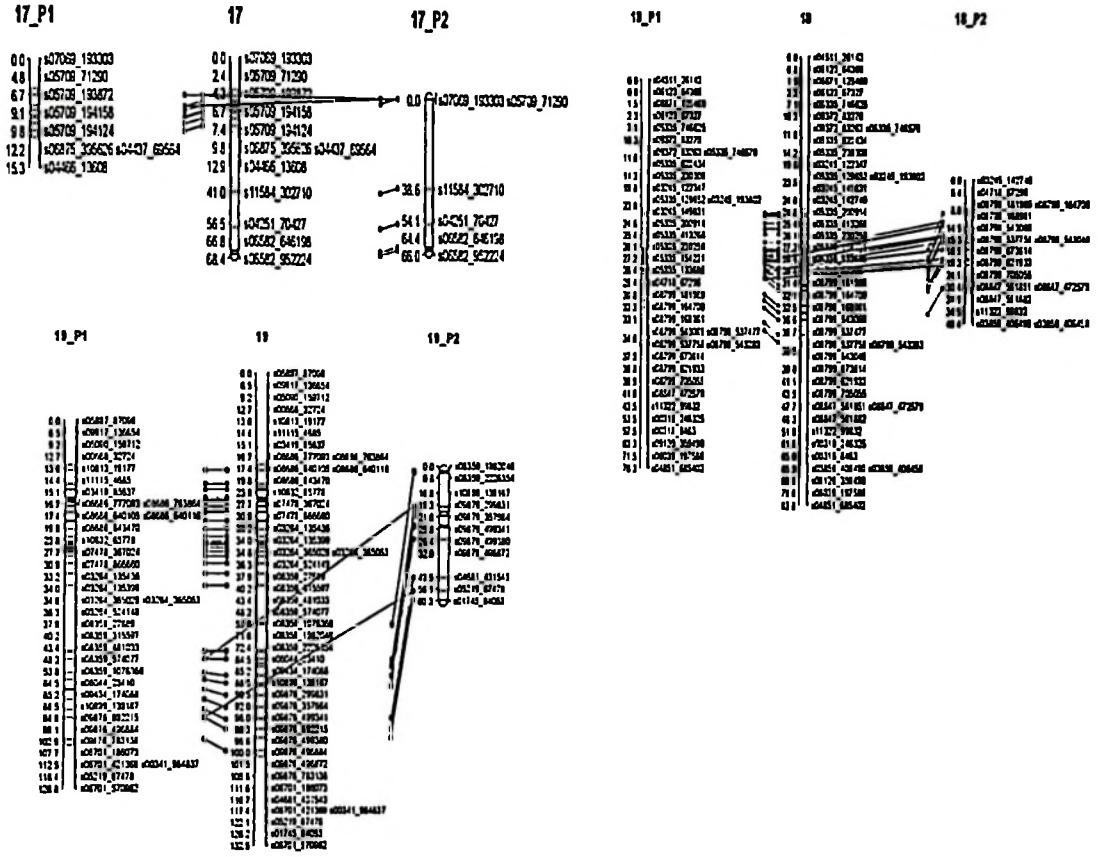


Figure 3 (f): LGs 17, 18 and 19

4.10 Distribution of markers on scaffolds

As shown in Fig. 3 (a-f), the integrated linkage map contained 1136 markers distributed on 415 unique scaffolds. Out of the 415 scaffolds, 211 were found to possess only one marker and 204 scaffolds had 2-39 markers. Scaffold 07 520 in linkage group one, had the highest number of 39 markers. Alternative way of locating SNP in the scaffolds is through BLAST the sequence against the draft cassava genome sequence available on phytozome.

4.11 Comparison with other SNP Map

By comparing the integrated map of this study and the integrated SNP map which was published by Rabbi *et al.* (2012). This map was found to be longer with an average map distance of 178.2 cM compared to 90 cM per group and had more SNP markers (1136 compared to 348). The average distance between adjacent markers of 3.4 cM in both maps.

CHAPTER FIVE

5.0 DISCUSSION

5.1 Integrity of the Mapping Population

In NS193, some F1 progenies possessed an allele 256bp which was not present in either parent. Since 258bp came from TMS 4(2)1425, 256bp allele is likely to have come from Mkombozi which indicates that there is genetic variation in female parent. Using SSRY 38, there were individuals with other alleles apart from the expected ones; these are 104bp, 80bp, 120bp and 104bp which suggesting there was variation in both parents. Possibly 104bp and 80bp came from both parents since there were individuals with 104bp/104bp and 80bp/80bp. 120bp might come in TMS 4(2)1425 if 104bp found in Mkombozi. This reflects variation in both parents-planted in the crossing block.

There was occurrence of the allele 97bp which was not found in either parent in SSRY169, which seemed to come in both parents since there was 97/99 and 89/97 suggesting that 97 came from both female and male parents, proposing variation in both parents. At locus SSRY147, some individuals were found to have alleles that were not present in either parent. These were 83bp, 100bp and 110bp. Another observation was that the 83bp/83bp alleles at SSRY147 always occurred with allele 80bp/80bp in SSRY38. SSRY 51 reveal heterozygous individuals with alleles not found in either parents (279bp and 281) suggesting that there was variation in both parents. There were many heterozygous individuals with an allele that was not found in the parents (265bp/267bp and 257bp/265bp) detected by SSRY52. Since 267bp and 257bp were from the male parent, 267bp must have been derived from variation in the female parent (Ferguson, M.E. personal communication, 2013).

The observed variation can be explained by the out crossing nature of cassava and past history of 'clones'. Mkombozi and TMS 4(2)1425 were brought to the lake zone of Tanzania in 1990s by Eastern Africa Root Crops Research Network (EARRNET) and International Institute of Tropical Agriculture (IITA) respectively. The varieties were grown in the field and adopted by farmers who grew them together with other cassava varieties. Some out crossing is likely to have occurred with other varieties. Volunteer seedlings are often allowed to develop resulting in a 'true' or 'typical' parental variety (Mkombozi or TMS 4(2)1425) and plants that look similar to the parental variety but are genetically slightly different (off-types) (Ferguson, M.E. personal communication, 2013). During the selection of plants to be grown in the crossing block, either 'true' or 'typical' parents can be selected or off-type. This variation within the parents causes the presence of unexpected alleles.

5.2 Genetic Markers

Several kinds of genetic markers have been used in the construction of genetic linkage maps in different crops (Okogbenin *et al.*, 2006; Xia *et al.*, 2007; Li *et al.*, 2009; Kunkeaw *et al.*, 2010b; Wang *et al.*, 2011; Rabbi *et al.*, 2012). Recently Single Nucleotide polymorphism (SNP) markers have attracted significant attention in creating dense genetic linkage maps and genome-wide association studies (Wang *et al.*, 2005) because SNPs are the most abundant class of polymorphisms in genomes, and can be genotyped cost-effectively (Rafalski, 2002). Since the parents segregates in response to CBSD infection, markers and QTLs associated with resistance to CBSD can be identified from mapping population of Mkombozi x TMS 4(2) 1425, through QTL analysis of phenotypic information and genetic results which have been obtained in this study.

5.3 Segregation Distortion

Segregation distortion is commonly encountered in actual genetic mapping populations which skew the frequency of genotypes from the expected Mendelian ratio (Lu *et al.*, 2002). The distortions are caused either by differential representation of segregation distortion loci (SDL) genotypes in gametes before fertilization or by viability differences of SDL genotypes after fertilization (Wang *et al.*, 2005). In out-crossing species like cassava, high proportions of segregation distortions are frequent. Cassava, being an out-crossing species has high genetic load, suffers from severe inbreeding depression upon selfing (Okogbenin *et al.*, 2006) and segregation distortions should therefore be expected. Segregation distortion is increasingly recognized as a potentially powerful evolutionary force and would affect the construction of genetic linkage maps. Segregation distortion detected in this study was 16% which is smaller than that reported by Okogbenin *et al.* (2006). A further explanation for segregation distortion could be the presence of mutations within the primer binding sites which results in preferential amplification of one of the alleles (Rabbi *et al.*, 2012). Distorted segregation has been reported in several crop species (Li *et al.*, 2010; Xu, 2008; Tai *et al.*, 2000). The segregation distortion detected in this study did not affect the constructing of the linkage map and hence could not cause difficulties on targeting markers linked to CBSD.

5.4 Linkage Map

After constructing the map, the female parent was having more markers as compared to the male. The integrated linkage map spanned 3207.9 cM. The map was longer than the SNP map of cassava developed by Rabbi *et al.* (2012) which had 568 markers and spanned 1837cM and longer than the cassava linkage maps developed by Kunkeaw *et al.* (2010b) which had 211 markers and spanned 1178 cM ; Sraphet *et al.* (2011) which consisted 510 markers and spanned 1420.3 cM; Chen *et al.* (2010) which contained 355

markers and spanned 1707.9 cM; Okogbenin *et al.* (2006) of 100 markers and spanned 1236.7 cM and Fregene *et al.* (1997) that had 168 markers spanned 931.6 cM. Marker distribution along the linkage groups within the map varied and was not uniform, low density of markers in some of the linkage groups might be due to a higher rate of recombination observed in cassava due to markers being far apart (Okogbenin *et al.*, 2006). The map generated in this study was expected to be longer as it consisted of many more markers compared to previously developed cassava linkage map.

The greatest gap between two adjacent markers was observed in the female parent (62.7 cM) of LG3 which was larger than that found in the map developed by Rabbi *et al.* (2012) of 30.9 cM. The large gap explains why there are 19 linkage groups in female parent instead of 18 linkage groups (haploid number of chromosome) in cassava i.e. n = 18. Cassava is diploid with 36 chromosomes and heterozygous, therefore it is expected to have 18 linkage groups as haploid number of chromosome. The number of linkage groups may be larger than the haploid chromosome number if there is recombination frequency is close to 0.5 as observed in the linkage map. There was also a gap > 10 cM in many other linkage groups which may suggest that there were 'hotspots' of recombination or insufficient marker coverage. This case had been reported in soybean (Hwang *et al.*, 2009; Rabbi *et al.*, 2012).

5.5 Comparison of SNP Markers Distribution on Female and Male Parents

The presence of heterozygous markers in both parents (*hk* x *hk*) favors the identification of homologous linkage groups as well as the integration of the female and male parents (Fig 3 a-i). Homologies were found in 18 linkage groups with a total of 213 common markers. A higher number of markers for the female parent compared to the male were observed. The difference in the total number of markers might be indicating that the

female parent (Mkombozi) had larger heterozygosity than the male parent (TMS 4(2) 1425). This larger heterozygosity might be resulting into genetically CBSD tolerant in Mkombozi.

5.6 Comparison with other SNP Maps

In terms of average map distance between markers and number of markers per linkage group, this map was longer than map published by Rabbi *et al.* (2012). The longer the total size of this map might have been contributed by the larger number of markers used.

CHAPTER SIX

6.0 CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

In the view of the above results it is concluded that:

Five markers (SSRY38, SSRY169, SSRY147, SSRY51 and SSRY52) revealed variation in both of the parents and NS193 revealed variation only in Mkombozi. This variation brought difficulties on identifying the true cross individuals, the self product from female parent and off type cross hybrids. The variation is likely to be derived from parental plants which were not typically Mkombozi or TMS 4(2)1425. The variation was likely to have come from outcrossing events with other clones, as the parental genotypes had been grown with a range of other genotypes for more than ten years.

A total of 1136 SNP markers were mapped on 18 LG of the integrated linkage map which will be used in QTL analysis to detect QTL linked with CBSD tolerance. QTL mapping is carried out by looking for associations between genotypes at individual markers and phenotypic traits of interest. It is expected that QTL linked with CBSD tolerance in Mkombozi will be identified after confirming the associations between the molecular markers data with CBSD phenotyping data. The identified QTL linked with this trait will be used in MAS. MAS is an efficient strategy to greatly improve selection efficiency for breeding of CBSD resistance in which selection for resistant genotypes (plants carrying CBSD resistance) can be done at seedling stage and allows for the reduction of population sizes and in the time taken to deliver a new variety.

6.2 Recommendations

According to the findings obtained in this study the following recommendations are made:

i. Breeding involves crossing between different species of the Manihot genus (inter-specific cross) or within the Manihot esculenta species (intra-specific crosses). Naturally, cassava is subject to both cross-pollination (outcrossing) and self-pollination (selfing) depending on genotype, planting design and type of pollinating insect present. Being outcrossing, cassava is subjected to variations within a variety that is maintained through continuous vegetative propagation. Genetic parental screen to be done before mapping populations are developed to ensure the uniformity of parents.

ii. To phenotype the mapping population to get the phenotypic data which will be correlated with molecular data in order to carry out QTL analysis for detecting QTL tightly linked to markers associated with CBSD tolerance gene/s. Identification of this QTL will be useful in breeding CBSD resistant / tolerant cassava varieties

iii. Molecular markers linked with CBSD tolerance from Mkombozi obtained from this study to be used for introgression of this trait into a range of farmer-preferred cassava varieties through MAS. This will quickly and effectively enable the selection of resistant plants that carrying this trait at early stage of breeding and hence can efficiently facilitate the generation of elite cassava varieties resistant / tolerant to the disease.

REFERENCES

Akano, A. O., Dixon, A. G. O., Mba, C., Barrera, E. and Fregene, M. (2002). Genetic mapping of dominant gene conferring resistance to cassava mosaic disease. *Theoretical and Applied Genetics* 105: 521 – 525.

Akkareddy, S., Lakshminarayana, R., Vermireddy, H., Jayaprad, M., Sridhar, S., Ramanarao, P. V., Anuradha, G. and Siddiq, E. A. (2010). Identification and mapping of landraces derived QTL associated with yield and its components in rice under different nitrogen levels and environments. *International Journal of Plant Breeding and Genetics* 4: 210 – 227.

Allem, A. C. (2002). The Origin and Taxonomy of Cassava. In: *Cassava Biology, Production and Utilisation.* (Edited by Hillocks, R. J., Thresh, M. J. and Bellotti, A.C.), International CABI Publishing, Wallingford, UK. 16pp.

Alicai, T., Omongo, C. A., Maruthi, M. N., Hillocks, R. J., Baguma, Y., Kawuki, R., Bua, A., Otim-Nape, G. W. and Colvin, J. (2007). Re-emergence of cassava brown streak disease in Uganda. *Plant Disease* 91: 24 – 29.

Alves, A. C. (2002). Cassava Botany and Physiology. In: *Cassava Biology, Production and Utilization* (Edited by Hillocks, J. R., Thresh, J. M and Belloti, A. C), Commonwealth of Agriculture Bureau International, Wallingford Oxon, UK. 89pp.

Antanaviciute, L., Fernandez- Fernandez, F., Jansen, J., Banch, E., Evans, K. M., Viola, R., Velasco, R., Dunwell, J. M., Troggio, M. and Sargent, D. J. (2012). Development of dense SNP-based linkage map of an apple rootstock progeny using the *Malus infinium* whole genome genotyping array. *BMC Genomics* 13(203): 1471 – 2164.

Appleby, N., Edwards, D. and Batley, J. (2009). Methods in Molecular Biology. In: *New Technologies for Ultra- high Throughput Genotyping in Plants*: (Edited by Gustafson, J. P., Lang-ridge, P. and Somers, D. J.), Humana Press, New York. 39pp.

Asiedu, R., Hahn, S. K., Vijaya, B. K. and Dixon, A. G. O. (1994). Interspecific hybridisation in the genus *Manihot*- progress and prospects. *Acta Horticulture* 380: 110 – 113.

Asif, M., Mehboob-Ur-Rahman., Mirza, J. I. and Zafar, Y. (2009). Percentage confirmation of cotton hybrids using molecular markers. *Pakistan Journal of Botany* 41(2): 695 – 701.

Balat, M. and Balat, H. (2009). Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy* 86: 2273 – 2282.

Bindler, G., Plieske, J., Bakaher, N., Gunduz, I., Ivanov, N., Van der Hoeven, R., Ganal, M and Donini, P. (2011). A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. *Theoretical and Applied Genetics* 122: 219 – 230.

Brian, A. K. and John, R. (1940). Cassava investigations in Zanzibar. *East African Agricultural Journal* 2: 404 – 406.

Carter, S. E., Fresco, L. O., Jones, P. G. and Fairbairn, J. N. (1992). *An Atlas of Cassava in Africa, Historical, Agro-ecological and Demographic Aspects of a Crop Distribution.* Centro Internacionale de Agricultura Tropical, Cali, Colombia. 85pp.

Ceballos, H., Iglesias, C. A., Perez, J. C. and Dixon, A. G. O. (2004). Cassava breeding: opportunities and challenges. *Plant Molecular Biology* 56: 503 – 516.

Ceballos, H., Okogbenin, E., Perez, J. C., Becerra Lopez-Valle, L. A. and Debouck, D. (2010). Cassava. Root and tuber crops, handbook. *Plant Breeding* 7: 53 – 96.

Ceballos, H., Perez, J. C., Morante, N., Sanchez, T., Chavez, A. L., Calle, F. and Fregene, M. (2007). Strategies to develop and identify high-value cassava clones. In: *Proceedings of the Second General Meeting on Biotechnology, Breeding and Seed Systems for African Crops Rockefeller Foundation*: *Research and Products Development that Reaches Farmers.* (Edited by Van Houten, H., Tom, K. and Tom-Wielgoz, V.), 26 – 29 March 2007, Maputo, Mozambique. pp. 42 – 43 .

Chavarriaga-Aguirre, P., Maya, M. M., Bonierbale, M. W., Kresovich, S., Fregene, M. A., Tohme, J. and Kochert, G. (1998). Microsatellites in cassava *(Manihot esculenta* Crantz): discovery, inheritance and variability. *Theoretical and Applied Genetics* 97: 493 – 501.

Chen, X., Zhiqiang, X., Yuhua, F., Cheng, L. and Wang, W. (2010). Constructing a genetic linkage map using an F1 population of non-inbred parents in cassava (*Manihot esculenta* Crantz). *Plant Molecular Biology* 28: 676 – 683.

Cho, R. J., Mindrinos, M., Richards, D. R., Sapolsky, R. J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T. L., Stammers, M., Federspiel, N., Theologis, A., Yang, W. H., Hubbell, E., Au, M., Chung, E. Y., Lashkari, D., Lemieux, B., Dean, C., Lipshutz, R. J., Ausubel, F. M., Davis, R. W. and Oefner, P. J. (1999). Genome-wide mapping with biallelic markers in *Arabidopsis thaliana. Nature Genetics* 23: 203 – 207.

CIAT (2004). *Improved Cassava for Developing World.* Annual Report No. 3. International Center for Tropical Agriculture, Cali, Colombia. 26pp.

Dellaporta, S. L., Wood, J. and Hicks, J. R. (1983). A plant Deoxyribonucleic acid minipreparation: version II. *Plant Molecular Biology Reports* 1: 19 – 21.

DeVries, J. and Toenniessen, G. (2001). *Securing the Harves, Biotechnology, Breeding and Seed Systems for African Crops.* Commonwealth of Agriculture Bureau International, Wallingford Oxon, UK. 155pp.

Dongre, A. and Pakri, V. (2005). Identification of cotton hybrids through the combination of PCR based RAPD, ISSR and microsatellite markers. *Journal of Plant Biochemistry and Biotechnology* 14: 53 – 55.

Edwards, D., Forster, J. W., Cogan, N. O. I., Batley, J. and Chagne, D. (2010). Single nucleotide polymorphism discovery. In: *Association Mapping in Plants*: (Edited by Oraguzie, N. C., Rikkerink, E. H. A., Gardiner, S. E., De Silva, H. N.), Springer, New York. 76pp.

Egesi, C., Cuambe, C., Rosero, C., Sanchez, A., Morante, T., Ceballos, H. and Fregene, M. (2008). Controlling delayed post-harvest physiological deterioration in cassava. In: *Cassava Meeting the Challenges of the new Millennium, Book of Abstracts, First Scientific Meeting of the Global Cassava Partnership, Institute of Plant Biotechnology for Developing Countries University of Ghent,* 21 – 25 July 2008, Ghent Belgium. 42pp.

Ekanayake, I. J., Osiru, D. S. O. and Porto, M. C. M. (1997). *Morphology of Cassava.* International Institute of Tropical Agriculture, Ibadan, Nigeria. 33pp.

Ellis, R. H., Hong, T. D. and Roberts, E. H. (1982). An investigation of the influence of constant and alternating temperature on the germination of cassava seed using a two-dimensional temperature gradient plate. *Annals of Botany* 49: 241 – 246.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S. and Motchell, S. E. (2011). A robust simple genotyping-by-sequencing approach for high diversity species. *PLoS ONE* 6(5): 94 – 99.

Emma, S. M., Rami, J. F., Bouchet, S., Klein, P. E., Rlein, R., Kilian, A., Wenzl, P., Xia, L., Halloran, K. and Jordan, D. R. (2009). A consensus genetic map of sorghum that integrates multiple component maps and high throughput Diversity Array Technology markers. *Plant Biology Journal* 1(6): 45 – 48.

FAO (2006). Production Year Book for 2005. Rome, Italy. [http://faostat.fao.org] site visited on 22/4/2012.

FAO (2008). Cassava for food and energy security Newsroom. [http://www.fao.org] site visited on 16/7/2012.

FAO (2010). Statistical database of the food and agricultural organization of the United Nations. [http://faostat.fao.org] site visited on 24/7/2012.

FAO and IFAD (2001). *The World Cassava Economy.* Facts and Outlook, Rome, Italy. 31pp.

Ferguson, M. E., Hearne, S. J., Close, T. J., Wanamaker, S., Moskal, W. A., Town, C. D., de Young, J., Marri, P. R., Rabbi, I. Y. and de Villiers, E. P. (2011). Identification, validation and high-throughput genotyping of transcribed gene SNPs in cassava. *Theoretical and Applied Genetics* 122: 1739 – 1749.

Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavariaga, P., Roca, W., Tohme, J. and Bonierbale, M. (1997). A molecular genetic map of cassava. *Theoretical and Applied Genetics* 95: 431 – 441.

Fregene, M., Morante, N., Sánchez, T., Marin, J., Ospina, C., Barrera, E., Gutierrez, J., Guerrero, J., Bellotti, A., Santos, L., Alzate, A., Moreno, S. and Ceballos, H. (2006). Molecular markers for introgression of useful traits from wild Manihot relatives of cassava, marker assisted selection of disease and root quality traits. *Journal of Root Crops* 32(1): 1 – 31.

Fukudi, W. M. G., Guevara, C. L., Kawuki, R. and Ferguson, M. E. (2010). *Selected Morphological and Agronomic Descriptors for the Characterization of Cassava.* International Institute of Tropical Agriculture, Ibadan, Nigeria. 19pp.

Gaur, R., Sethy, N. K., Choundhary, S., Shokeen, B., Gupta, V. and Bhatia, S. (2011). Advancing the STMS genomic resources for defining new locations on the inraspecific genetic linkage map of chickpea (*Cicer arietinum* L). *BMC Genomics* 12(117): 1471 – 2164.

Gomez, S. M., Denwar, N. N., Ramasubramanian, T., Simpson, C. E., Burow, G., Burke, J. J. and Puppala, N. (2008). Identification of peanut hybrids using microsatellite markers and horizontal polyacrylamide gel electrophoresis. *Peanut Science* 35: 123 – 129.

Gondwe, F. M. T., Mahungu, N. M., Hillocks, R. J., Raya, M. D., Moyo, C. C., Soko, M. M., Chippings, F.P. and Benesi, I. (2003). Economic losses experienced by small-scale farmers in Malawi due to cassava brown streak virus disease. In: *Cassava Brown Streak Disease Past, Present and Future, Proceedings of the International Workshop*, (Edited by Legg, J. P. and Hillocks, R. J.), 27 - 30 October 200, Mombasa, Kenya. pp. 29 – 39.

Hillocks, R. J. (1997). Cassava virus disease and their control with special reference to southern Tanzania. *Integrated Pest Management Reviews* 2: 125 – 138.

Hillocks, R. J. and Thresh, J. M. (2000). Cassava mosaic and cassava brown streak virus diseases in Africa: A comparative guide to symptoms and aetiology. *Roots* 7: 3 – 27.

Hillocks, R. J., Raya, M. D., Mtunda, K. M. and Kiozya, H. (2001). Effects of brown streak virus disease on yield and quality of cassava in Tanzania. *Journal of Phytopathology* 140: 389 – 394.

Hwang, T. Y., Sayama, T., Takahashi, M., Takada, Y., Nakamoto, Y., Funatsuki, H., Hisano, H., Sasamoto, S., Sato, S., Tabata, S., Kono, I., Hoshi, M., Hanawa, M., Yano, C., Xia, Z., Harada, K., Kitamura, K. and Ishimoto, M. (2009). High-density integrated linkage map based on SSR markers in Soybean. *DNA Research* 16: 213 – 225.

Iglesias, C., Mayer, J., Chavez, L. and Calle, F. (1997). Genetic potential and stability of carotene in cassava roots. *Euphytica* 94: 367 – 373.

IITA (1980). *Annual report of the International Institute of Tropical Agriculture.* Ibadan, Nigeria. 185pp.

IITA (1990). *Cassava in Tropical Africa A Reference Manual.* Balding and Marsell International, Wisbech, United Kingdom. 98pp.

IITA (2007). Solution found for cassava root-rot devastation in Africa. [http://www.iita.org/cms/detail/news] site visited on 15/4/2011.

Ingelbrecht, I., Lokko, Y., Raji, A., Herron, C., Dixon, A. G. O., Mahungu, N. and Winter, S. (2005). Towards Molecular Enhancement of Cassava for Resistance to Cassava Brown Streak Disease. In: *Scientific Meeting on Biotechnology, Breeding and Seed Systems for African Crops*: Research and Product Development that Reaches the Farmers, 24 – 27 January 2005, Nairobi, Kenya. pp. 16 – 39.

IPGRI and CIAT (2003). *Descriptors for Cassava (Manihot esculenta)*. International plant Genetic Resources Institute, Rome, Italy and Centro International para la Agricultura Tropical, Cali, Colombia. 28pp.

Jennings, D. L. (1957). Further studies in breeding cassava for virus resistance. *East African Agricultural Journal* 22: 213 – 219.

Jennings, D. L. (1960a). Observations on virus diseases of cassava in resistant and susceptible varieties. I. Mosaic disease. *Empire Journal of Experimental Agriculture* 28: 23 – 34.

Jennings, D. L. (1960b). Observations on virus diseases of cassava in resistant and susceptible varieties. I. Brown streak. *Empire Journal of Experimental Agriculture* 28: 261 – 270.

Jennings, D. L. (1978). Inheritance of linked resistances to African cassava mosaic and bacterial blight diseases. In: *Proceedings of cassava protection workshop* (Edited by Brekelbaum, T., Bellotti, A. and Lozano, J. C.), Centro International para la Agricultura Tropical, Cali, Colombia. pp. 45 – 49.

Jennings, D. L. and Hershey, C. H. (1985). Cassava Breeding, A Decade of Progress from International Programmes. In: *Progress in plant breeding.* (Edited by Russell, G. E.), Butterworths, London, UK. 116pp.

Jennings, D. L. and Iglesias, C. H. (2002). Breeding for Crop Improvement. In: *Cassava biology, production and utilization.* (Edited by Hillocks, R. J., Thresh, J. M. and Bellotti, C. A.), CABI Publishing, Wallingford, UK. 166pp.

Kanju, E., Mkamilo, G., Mgoo, V. and Ferguson, M. (2010). Statistical evidence linking the zigzag stems habit with tolerance to cassava brown streak disease. *Roots* 12: 4 – 6.

Kanju, E., Masumba, E., Masawe, M., Tollano, S., Muli, B., Zacarias, A., Mahungu, N., Khizzah, B., Whyte, J. and Dixon, A. (2007). Breeding cassava for brown streak resistance: regional cassava variety development strategy based on farmers and consumer preferences at AICC. In: *Proceedings of the Thirteenth Triennial Symposium of the International Society for Tropical Root Crops: Tropical Root and Tuber Crops: Opportunities for Poverty Alleviation and Sustainable Livelihoods in Developing Countries.* (Edited by Kapinga, R., Kingamkono, R., Msabaha, M., Ndunguru, J., Lenmaga, B. and Tusiime, G.), 10 - 14 November 2007, Arusha, Tanzania. pp. 95 – 101.

Kapinga, R. J., Mafuru, J., Jeremiah, S., Rwiza, E., Kamala, R., Mashamba, F. and Mlingi, N. (2001). Status of cassava in Tanzania: Implication for future research and development. [http://www.globalcassavastrategy.net/Africa/Tanzania/t0000e00.htm] site visited on 15/02/2011.

Kasele, S., Merumba, S., Rweywmamu, P., Ndyetabula I. L., Chirimi, B., Kagimbo, F., Chuwa, Q. and Bagambisa. (2009). *A Report on Cassava Pests and Diseases Diagnostic Survey.* The Lake and Western Zones of Tanzania. 71pp.

Kawano, K. (1980). Cassava. In: *Hybridization of Crop Plants.* (Edited by Fehr, W. R. and Hadley, H. H.), American Society of Agronomy and Crop Science Society of America, Madison, Wisconsin, USA. 233pp.

Kosambi, D. D. (1994). The estimation of map distances from recombination values. *Ann Eugen* 12: 172 – 175.

Kulembeka, H. F. (2010). Genetic linkage mapping of field resistance to cassava brown streak disease in cassava (*Manihot esculenta crantz*) landraces from Tanzania. Thesis for Award of PhD Degree at University of the Free State, Bloemfontein, South Africa, 258pp.

Kunkeaw, S., Tangphatsornruang, S., Smith, D. R. and Triwitayakorn, K. (2010a). Genetic linkage map of cassava (*Manihot esculenta* Crantz) based on AFLP and SSR markers. *Plant Breeding* 129: 112 – 115.

Kunkeaw, S., Yoocha, T., Sraphet, S., Boonchanawiwat, A., Boonseng, O., Lightfoot, D. A., Triwitayakorn, K. and Tangphatsornruang, S. (2010b). Construction of a genetic linkage map using simple sequence repeat markers from expressed sequence tags for cassava (*Manihot esculenta* Crantz). *Molecular Breeding* 27: 67 – 75.

Lebot, V. (2009). *Tropical Root and Tuber Crops, Cassava, Sweet Potato, Yams and Aroids.* Crop production science in horticulture series. CAB International, UK. 85pp.

Legg, J. P. (1999). Emergency, spread and strategies for controlling the pandemic of cassava mosaic virus disease in East and central Africa. *Crop Protection* 18: 627 – 637.

Li, H. B., Kilian, A., Zhou, M. X., Wenzl, P., Huttner, E., Mendham, E., McIntyre, L. and Vaillancourt, R. E. (2010). Construction of a high-density composite map and comparative mapping of segregation distortion regions in barley. *Molecular Genetics: Genomics* 284: 319 – 331.

Li, F., Kitashiba, H., Inaba, K. and Nishio, T. (2009). A *Brassica rapa* linkage map of EST based SNP markers for identification of candidate genes controlling flowering time and leaf morphological traits. *DNA Research* 16: 311 – 332.

Lu, H., Romero-Severson, J. and Bernando, R. (2002). Chromosomal regions associated with segregation in maize. *Theoretical and Applied Genetics* 105: 622 – 628.

Macc, E. S., Phong, D. T., Upadhyaya, H. D., Chandra, S. and Crouch, J. H. (2006). SSR analysis of cultivated groundnut *(Arachis hypogaea* L.*)* germplasm resistant to rust and late leaf spot diseases. *Euphytica* 152: 317 – 330.

Mahungu, N. H., Dixon, A. G. O. and Mkumbira, J. M. (1994). Breeding cassava for multiple pest resistance in Africa. *African Crop Science Journal* 2: 539 – 552.

Maruthi, M. N., Hillocks, R. J., Mtunda, K., Raya, M. D., Muhanna, M., Kiozia, H., Rekha, A. R., Colvin, J. and Thresh, J. M. (2005). Transmission of cassava brown streak virus by *Bemisia tabaci* (Gennadius). *Journal of Phytopathology* 153: 307 – 312.

Masumba, E. A. (2006). Genetic diversity and field performance of cassava (*Manihot esculenta* Crutz) landraces commonly grown in eastern southern and Lake Zones. Dissertation for Award of MSc Degree at Sokoine University of Agriculture, Morogoro, Tanzania, 89pp.

Mba, R. E. C., Stephenson, P., Edward, K., Melzer, S., Mkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, J. and Fregene, M. (2001). Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theoretical and Applied Genetics* 102: 21 – 31.

Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., Kullaya, A., Jari, P and Valkonen, T. (2009a). Genetically distinct strains of Cassava brown streak virus in the Lake Victoria basin and the Indian Ocean coastal area of cast Africa. *Archives of Virology* 154: 353 – 359.

Mkamilo, G. S. and Hillocks, R. J. (2006). Impact of cassava brown streak disease on processing and eating quality of cassava. *Roots* 10: 13 – 16.

Muhana, M. and Mtunda, K. J. (2002). *Report on the Cassava Root Rot Problem* in Muheza District. Tanga, Tanzania. 13pp.

Mustafa, E., Sandhu, D., Sidhu, D., Dilbirligi, M., Baenzigerl, P. S. and Gill, K.S. (2004). Demarcating the gene-rich regions of the wheat genome. Oxford Journals. *Nucleic Acid Research* 32: 3546 – 3565.

Nichols, R. F. J. (1947). Breeding cassava for virus resistance. *The East African Agricultural Journal* 12: 184 – 194.

Nichols, R. F. W. (1950). The brown streak disease of cassava: distribution climatic effects and diagnostic symptoms. *East African Agricultural Journal* 15: 154 – 160.

Ntawuruhunga, P. and Legg, J. (2007). *New Spread of Cassava Brown Streak Virus Disease and its Implications for the Movement of Cassava Germplasm in the East and Central African Region.* Crop crisis control project report. International Institute of Tropical Agriculture, Catham Maritime, United Kingdom. 6pp.

Nweke, F., Spencer, D. and Lyam, J. (2002). *The Cassava Transformation, African Best Kept Secret.* Michigan State University Press, East Lansing, USA. 273pp.

Okogbenin, E., Marin, J. and Fregene, M. (2006). An SSR-based molecular genetic linkage map of cassava. *Euphytica* 147: 433 – 440.

Olsen, K. M. and Schaal, B. A. (1999). Evidence on the origin of cassava: Phylogeography of *Manihot esculenta. National Academy of Science* 96: 5586 – 5591.

Olsen, K. M and Sachaal, B. A. (2001). Microsatellite variation in cassava *Manihot esculentus*, *Euphorbiaceae* and its wild relatives: Evidence for the southern Amazon origin of domestication. *American Journal of Botany* 4: 131 – 142.

Otim-Nape, G. W., Bua, A., Thresh, J. M., Baguma, Y., Goral, S., Ssemakula, G. N., Alcoa, G., Byabakama, B., Colvin, J., Cooter, R. J. and Martin, A. (2000). *The Current Pandemic of Cassava Mosaic Virus Disease in East Africa and Its Control.* National Agricultural Research Organization Uganda, National Research Institute University of Greenwich, UK. 100pp.

Powell, W., Machray, G. and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends Plant Science* 1(7): 215 – 222.

Rabbi, I. Y., Kulembeka, H. P., Masumba, E., Marri, P. R. and Ferguson, M. (2012). An EST-derived SNP and SSR genetic linkage map of cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics* 122: 1836 – 1840.

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. Curr *Opinion Plant Biology* 5: 94 – 100.

Rashmi, G., Sarwar, A., Ganga, J., Khan, A.W., Shalu, C., Mukesh, J., Gitanjari, Y., Akhilesh, K.T., Debasis, C. and Sabhyata, B. (2012). High-Throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (*Cicer arietinum* L.). *DNA Research* 18: 1 – 17.

Salgado, K. C. P. C., Viera, M. G. G. C., Pinho, E. V. R., Guimara˜es, C. T., Pinho, R. G., Souza, L. V. (2006). Genetic purity certificate in seeds of hybrid maize using molecular markers. *Revista Brasileira de Sementes* 28(1): 169 – 175.

Schaal, B. A., Olsen, K. M. and Carvalho, L. J. C. B. (2006). Evolution, domestication and agrobiodiversity in tropical crop cassava. In: *Darwin's Harvest New Approaches to the Origins, Evolution and Conservation of Crops*. (Edited by Motley, T. J., Zerega, and N. and Cross, H.), Colombian University Press, New York. pp. 269 – 283.

Schmitz, P. M. and Kavallari, A. (2009). Crop plants versus energy plants - on the international food crisis. *Bioorganic and Medicinal Chemistry* 17: 4020 – 4021.

Sraphet, S., Boonchanawiwat, A., Thanyasiriwat, T., Boonseng, O., Tabata, S., Sasamoto, S., Shirasawa, K., Isobe, S., Lightfoot, D. A., Tangphatsornruang, S. and Triwitayakorn, K. (2011). SSR and EST-SSR-based genetic linkage map of cassava (*Manihot esculenta* Crantz). *Theoretical Applied Genetics* 122: 1161 – 1170.

Storey, H. H. (1936). Virus diseases of East African plants: A progress report on studies of the diseases of cassava. *East African Agricultural Journal* 2: 34 – 39.

Storey, H. H. (1939). *Report of the Plant Pathology*. East African Agricultural Research Station, Nairobi, Kenya. 9pp.

Storey, H. H. and Nichols, R. F. W. (1938). Studies of the mosaic diseases of cassava. *Annals of Applied Biology* 25: 790 – 792.

Supawadee, T. and Sompong, T. (2009). Application of molecular markers in the hybrid verification and assessment of somaclonal variation and assessment of somaclonal variation from oil palm propagated in vitro. *Science Asia* 35: 142 – 149.

Tai, G. C. C., Seabrook, J. E. A. and Aziz, A. N. (2000). Linkage analysis of anther derived monoploids showing distorted segregation of molecular markers. *Theoretical and Applied Genetics* 101: 126 – 130.

Tamilkumar, P., Jerlin, R., Ganesan, K., Senthil, N., Jeevan, R and Raveendran, M. (2009). Fingerprinting of rice hybrids and their parental lines using microsatellite markers and their utilization in genetic purity assessment of hybrid rice. *Research Journal of Seed Science* 2: 40 – 47.

Terzic, S., Atlagic, J. and Rankovic, D. (2006). Characterization of F1 interspecific hybrids between wild *Helianthus annus* L. population and cultivated sunflower. *Genetika,* 38(2): 159 – 168.

Thottapilly, G., Mignoun, H. D. and Omitongun, O. G. (2000). The use of DNA markers for rapid improvement of crops in Africa. *African Crop Science Journal* 8: 99 – 108.

Torado, A., Koike, M., Mochida, K. and Ogihara, Y. (2006). SSR-based linkage map with new markers using an intraspecific population of common wheat, *Theoretical and Applied Genetics* 1(12): 1042 – 1051.

Van Ooijen, J. W. (2006). Join Map4. Software for the calculation of linkage maps in experimental populations. Kyazama, B.V., Wageningen, Netherlands.

Varshney, R. K., Marcel, T. C., Ramsay, L., Russell, J., Roder, M. S., Stein, N., Waugh, R., Langridge, P., Nicks, R. E. and Graner, A. (2007). A high density barley microsatellite consensus map with 775 SSR loci. *Theoretical ad Applied Genetics* 114(6): 1091 – 1093.

Wang, C. M., Cheng, S. Z., Hu, Q. Z. and Jian, M. W. (2005). Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (*Oryza sativa* L.). *Genetic* Research 86: 97 – 106.

Wang, C. M., Liu, P., Yi, C., Gu, K., Sun, F., Li, L., Chueng, L. L., Liu, X., Feng, F., Lin, G., Cao, S., Hong, Y., Yin, Z. and Yue, G. H. (2011). A first generation microsatellite- and SNP-based kinkage map of Jatropha. *PLoS ONE* 6(8): 5 – 8.

Were, H. K., Winter, S. and Maiss, E. (2004a). Occurrence and distribution of cassava begomoviruses in Kenya. *Annals of Applied Biology* 145: 175 – 184.

Were, H. K., Winter, S. and Maiss, E. (2004b). Viruses infecting cassava in Kenya. *Plant Disease* 88: 17 – 22.

Were, H. K., Winter, S. and Maiss, E. (2007). Characterisation and distribution of cassava viruses in Kenya. *African Crop Science* 8: 909 – 912.

Winter, S., Koerbler, M., Stein, B., Pietruszka, A., Paape, M. and Butgereit, A. (2010). The analysis of cassava brown streak viruses reveals the presence of a two distinct virus species causing cassava brown streak disease in East Africa. *Journal of General Virology* 91: 1365 – 1372.

Yaninek, J. S. (1994). Cassava plant protection in Africa In, Root Crops for Food Security in Africa. Akoroda, M. O. (Ed.), *Proceedings of the Fifth Trieennial Symposium of the International Society for Tropical Root Crops Africa Branch*, Kampala, Uganda, 22 – 28 November 1992. pp. 26 – 34.

Xia, Z., Tsubokura, Y., Hoshi, M., Hanawa, M., Yano, C., Okamura, K., Ahmed, T. A., Anai, T., Watanabe, S., Hayashi, M., Kawai, T. and Hossain, K. G. (2007). An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using a single F-2 population. *DNA Research* 14: 257 – 269.

Xia, J. H., Liu, F., Zhu, Z. Y., Fu, J., Feng, J., Li, J. and Yue, G. H. (2010). A consensus linkage map of the grass carp (*Ctenopharyngodon idella*) based on microsatellites and SNPs. *BMC Genomics* 11: 356.

Xu, S. (2008). Quantitative trait locus mapping can benefit from segregation distortion. *Genetics* 180: 2201 – 2208.

APPENDICES

Appendix 1: Characteristics of Mkombozi

Number	Parameter	Characterization
1	Pedigree	Half sib of 92/0099S2 (SM).
2	Color of Apical Leaves	Light green
3	Pubescence on Apical Leaves	Absent
4	Leaf Retention	Average leaf retention
5	Shape of Central Leaflets	Elliptic-lanceolate
6	Petiole Color	Reddish green
7	Leaf Color	Light green
8	Number of Leaf Lobes	Five lobes
9	Length of Leaf Lobes	12.9
10	Width of Leaf Lobe	4
11	Lobe Margin	Winding
12	Petiole Length	12.5
13	Color of Leaf Vein	Green
14	Orientation of Petiole	Horizontal
15	Flowering	Present
16	Pollen	Present
17	Prominence of Foliar Scars	Semi-prominent
18	Color of stem Cortex	Light green
19	Color of Stem epidermis	Cream
20	Color of Stem Exterior	Greeny-yellowish
21	Distance Between Leaf Scars	Short
22	Growth Habit of Stem	Straight
23	Color of End Branches of Adult Plant	Green
24	Length of Stipules	Short
25	Stipule Margin	Entire
26	Fruit	Present
27	Seed	Present
28	Plant Height	204
29	Height of First Branching	37
30	Levels of Branching	3
31	Branching Habit	Tetrachotomous
32	Angle of Branching	75-90
33	Shape of Plant	Open
34	Number of Storage roots per plant	7
35	Number of Commercial Roots Per Plant	7
36	Extent of Root Peduncle	Mixed
37	Root Constrictions	Some
38	Root Shape	Conical-cylindrical
39	External Color of Storage Root	Light brown
40	Color of Root Pulp (Parenchyma)	White
41	Color of Root Cortex	White
42	Cortex Easy of Peeling	Easy
43	Texture of Root Epidermis	Rough
44	Root Taste	Sweet
45	Cortex Thickness	Thick

Descriptors for cassava morphological characterization (Fukuda *et al.*, 2010)

Appendix 2: Characteristics of TMS 4(2)1425

Number	Parameter	Characterization
1	Pedigree	58308 x Oyarugba Funfun
2	Color of Apical Leaves	Purplish green
3	Pubescence on Apical Leaves	Absent
4	Leaf Retention	Outstanding leaf retention
5	Shape of Central Leaflets	Lanceolate
6	Petiole Color	Greenish red
7	Leaf Color	Dark green
8	Number of Leaf Lobes	Five lobes
9	Length of Leaf Lobes	17 cm
10	Width of Leaf Lobe	5.0 cm
11	Lobe Margin	Smooth
12	Petiole Length	19.4 cm
13	Color of Leaf Vein	Green
14	Orientation of Petiole	Horizontal
15	Flowering	Absent
16	Pollen	Present
17	Prominence of Foliar Scars	Prominent
18	Color of stem Cortex	Dark green
19	Color of Stem epidermis	Light brown
20	Color of Stem Exterior	Silver
21	Distance Between Leaf Scars	Short (≤ 8 cm)
22	Growth Habit of Stem	Zig-zag
23	Color of End Branches of Adult Plant	Green
24	Length of Stipules	Long
25	Stipule Margin	Entire
26	Fruit	Present
27	Seed	Present
28	Plant Height	86 cm
28	Plant Height	86 cm
29	Height of First Branching	45
30	Levels of Branching	3
31	Branching Habit	Tetrachotomous
32	Angle of Branching	75 – 90
33	Shape of Plant	Compact
34	Number of Storage roots per plant	6
35	Number of Commercial Roots Per Plant	3
36	Extent of Root Peduncle	Sessile
37	Root Constrictions	Few to one
38	Root Shape	Cylindrical
39	External Color of Storage Root	Yellow
40	Color of Root Pulp (Parenchyma)	White
41	Color of Root Cortex	White
42	Cortex Easy of Peeling	Easy
43	Texture of Root Epidermis	Rough
44	Root Taste	Sweet
45	Cortex Thickness	Thick

Descriptors for cassava morphological characterization (Fukuda *et al.*, 2010)

Appendix 3: Mkombozi x TMS 4(2)1425 F1 mapping population integrity using SSR markers

Sample ID	NS 193	SSRY 38	SSRY 63	SSRY 102	SSRY 119	SSRY 155	SSRY 169	SSRY 147	SSRY 51	SSRY 52	Conclusions
MT100	250/258	106/106	285/289	179/181	157/169	149/149	89/99	101/111	259/278	257/263	True cross
MT101	256/258	104/120	285/287	179/179	169/169	149/157	97/99	100/100	259/278	257/265	Off-type
MT102	-9/-9	-9/-9	285/289	179/181	157/169	149/149	99/99	132/132	278/298	263/267	Off-type
MT103	256/258	106/121	285/287	179/181	169/169	149/149	99/99	111/111	-9/-9	257/265	Off-type
MT104	250/258	106/108	287/289	179/179	169/169	149/149	99/99	101/111	278/298	263/267	True cross
MT105	250/256	104/120	285/285	179/179	157/157	149/149	97/99	100/111	259/279	267/267	Off-type
MT106	250/256	104/104	285/285	179/181	157/169	149/157	97/99	100/100	259/279	265/267	Off-type
MT108	250/256	106/106	285/289	179/179	169/169	149/149	89/89	111/111	278/278	263/263	Self
MT109	250/258	106/108	285/287	179/179	157/157	149/149	89/99	101/111	278/298	246/257	True cross
MT10	250/256	106/106	285/287	179/179	157/169	149/157	99/99	101/111	278/279	257/265	Off-type
MT110	250/250	106/108	285/289	179/181	157/169	149/149	99/99	111/111	278/298	246/257	True cross
MT111	250/256	106/106	285/287	179/179	157/169	149/157	97/99	101/101	278/298	267/267	Off-type
MT113	250/256	106/106	285/287	179/181	169/169	149/149	89/99	111/111	278/298	263/267	Off-type
MT114	250/256	106/121	285/287	179/181	157/169	149/149	97/99	101/101	259/278	257/265	Off-type
MT115	250/258	106/106	285/285	179/179	157/169	149/149	99/99	111/111	259/278	257/263	True cross
MT116	250/258	106/108	285/285	179/181	169/169	149/149	99/99	111/111	278/298	246/257	True cross
MT118	250/250	106/106	285/285	179/179	157/169	149/157	99/99	101/101	259/279	257/267	Off-type
MT119	-9/-9	80/80	285/285	179/181	157/169	149/149	99/99	-9/-9	259/278	257/263	Off-type
MT11	250/250	106/108	285/287	179/179	157/169	149/149	89/99	101/111	278/298	246/267	True cross
MT120	250/250	106/108	285/289	179/179	169/169	149/149	99/99	111/111	278/278	246/263	Self
MT121	250/250	106/108	287/289	179/181	169/169	149/149	99/99	101/101	259/278	263/267	True cross
MT122	250/250	106/121	285/285	179/181	157/157	149/149	97/97	101/101	278/278	265/267	Off-type
MT123	250/258	106/108	285/287	179/181	157/157	149/149	99/99	111/111	259/278	246/257	True cross
MT124	250/250	106/108	285/289	181/181	157/157	149/149	89/99	111/111	278/278	246/246	Self
MT125	250/250	104/120	285/285	179/181	169/169	149/157	99/99	100/100	259/278	257/265	Off-type
MT126	250/258	106/121	285/289	179/181	157/169	149/149	99/99	111/111	259/278	257/265	Off-type
MT129	250/250	106/106	285/287	179/181	157/169	149/157	99/99	111/111	279/298	257/267	Off-type
MT12	250/258	106/106	285/289	179/179	157/157	149/149	99/99	111/111	259/278	257/263	True cross
MT131	250/250	106/108	-9/-9	179/181	157/157	149/149	89/99	111/111	278/278	246/246	Self
MT132	-9/-9	80/80	287/289	179/179	169/169	149/149	89/99	83/83	259/278	246/257	Off-type
MT133	250/256	106/106	285/285	179/181	157/157	149/157	97/99	101/111	278/279	265/267	Off-type
MT134	250/258	106/108	285/287	179/179	157/169	149/149	89/99	101/111	259/278	263/267	True cross
MT135	250/258	106/106	285/287	179/179	157/169	149/149	89/99	101/111	278/298	246/257	True cross
MT136	250/250	106/108	289/289	179/181	157/169	-9/-9	89/89	111/111	278/278	246/263	Self
MT137	250/258	104/106	285/285	179/181	157/169	149/149	99/99	100/111	259/278	246/267	Off-type
MT138	250/256	106/106	285/285	179/181	157/157	149/149	97/99	101/111	298/278	257/265	Off-type
MT139	250/258	106/106	287/289	179/179	157/169	149/149	89/99	111/111	278/298	246/267	True cross
MT13	250/250	106/121	285/289	179/181	169/169	149/157	89/97	111/111	278/279	246/265	Off-type
MT140	250/258	106/108	-9/-9	179/181	157/157	149/149	99/99	111/111	278/298	257/263	True cross
MT141	250/256	106/106	285/287	179/179	157/169	149/149	97/99	101/111	259/279	257/265	Off-type
MT142	250/250	106/108	285/289	179/181	169/169	149/149	99/99	101/111	278/298	263/267	True cross
MT144	256/250	106/106	-9/-9	179/181	157/157	157/157	97/99	101/111	279/278	265/267	Off-type
MT145	250/258	106/106	287/289	179/179	157/157	149/149	99/99	111/111	278/298	246/257	True cross

Sample ID	NS 193	SSRY 38	SSRY 63	SSRY 102	SSRY 119	SSRY 155	SSRY 169	SSRY 147	SSRY 51	SSRY 52	Conc lusions
MT146	250/258	104/104	285/285	179/179	157/157	149/157	97/99	100/100	278/298	265/267	Off-type
MT147	250/258	106/106	285/287	179/179	157/157	149/149	99/99	111/111	278/279	267/267	Off-type
MT148	250/258	106/106	285/285	179/179	157/169	149/149	99/99	111/111	278/298	257/263	True cross
MT14	250/250	104/120	285/287	179/181	157/169	149/157	97/99	100/111	278/279	267/267	Off-type
MT150	250/250	106/108	285/289	179/181	157/169	149/149	99/99	111/111	278/278	263/263	Self
MT151	250/250	106/108	285/289	179/181	157/169	149/149	99/99	101/111	278/298	246/257	True cross
MT152	250/256	104/104	285/285	179/181	157/169	149/149	97/99	100/111	259/278	267/267	Off-type
MT154	250/250	106/106	285/285	179/179	157/157	149/149	99/99	111/111	278/278	263/265	Off-type
MT157	250/250	106/106	285/287	179/179	157/169	149/149	99/99	101/111	278/298	246/257	True cross
MT159	250/258	106/108	285/289	179/181	169/169	149/149	89/99	111/111	278/298	246/267	True cross
MT15	250/258	106/108	287/289	179/181	157/157	149/149	99/99	111/111	278/298	257/263	True cross
MT160	250/258	106/108	-9/-9	179/181	157/169	149/149	99/99	111/111	259/278	246/257	True cross
MT161	250/258	106/106	285/285	179/179	157/169	149/149	89/99	101/111	259/278	246/267	True cross
MT162	-9/-9	-9/-9	285/289	179/181	169/169	149/149	99/99	-9/-9	278/298	246/257	True cross
MT163	250/250	104/104	285/285	179/181	157/157	149/149	89/99	100/111	278/298	263/267	Off-type
MT164	250/258	106/106	285/289	179/179	157/157	149/149	89/99	111/111	259/278	246/267	True cross
MT165	250/250	106/106	285/285	179/181	157/169	149/149	99/99	101/111	278/298	263/267	True cross
MT166	250/256	106/121	285/285	179/181	169/169	149/149	97/99	101/111	278/298	265/267	Off-type
MT167	250/250	106/108	287/289	179/179	157/169	149/149	99/99	111/111	278/298	263/267	True cross
MT168	250/258	106/106	287/289	179/179	169/169	149/149	99/99	101/111	278/298	246/257	True cross
MT169	250/256	120/120	285/285	179/181	157/157	149/157	97/97	100/111	279/279	265/267	Off-type
MT16	250/258	106/106	285/285	179/181	157/157	149/149	89/99	111/111	278/298	257/263	True cross
MT170	250/256	106/121	285/285	179/181	157/157	149/149	97/99	111/111	278/279	265/267	Off-type
MT171	250/258	106/121	-9/-9	179/179	169/169	149/157	99/99	101/111	279/298	257/265	Off-type
MT172	250/258	106/108	285/289	179/179	157/169	149/149	89/99	101/111	278/298	246/267	True cross
MT173	250/250	106/106	285/287	179/181	157/157	149/149	89/99	101/111	259/278	263/267	True cross
MT174	250/258	106/108	285/289	179/179	157/169	149/149	89/99	101/111	278/298	257/263	True cross
MT175	250/250	106/108	285/285	181/181	157/169	149/149	89/89	111/111	278/278	263/263	Self
MT176	250/258	106/106	285/287	179/181	157/169	149/149	89/99	111/111	278/298	246/267	True cross
MT177	250/258	106/106	285/287	179/181	157/157	149/149	89/99	101/111	278/298	246/267	True cross
MT178	250/250	106/106	285/285	179/181	157/169	149/157	99/99	101/111	279/298	267/267	Off-type
MT179	250/258	106/106	285/287	179/181	157/169	149/149	89/99	101/111	259/278	246/257	True cross
MT17	256/258	106/121	285/285	179/179	157/157	149/149	99/99	101/111	278/279	257/267	Off-type
MT180	256/258	106/106	285/285	179/179	157/169	149/157	97/99	101/111	278/279	267/267	Off-type
MT181	250/258	106/108	287/289	179/181	157/169	149/149	89/99	101/111	259/278	263/267	True cross
MT182	250/258	106/106	287/289	179/181	157/169	149/149	89/99	101/111	259/278	246/257	True cross
MT183	-9/-9	80/80	-9/-9	179/181	157/169	149/149	99/99	83/83	298/298	257/267	Off-type
MT184	256/250	106/106	-9/-9	179/179	157/157	149/157	97/97	111/111	278/278	265/267	Off-type
MT186	250/250	106/108	285/287	179/179	157/169	149/149	97/99	111/111	278/298	263/267	True cross
MT187	250/258	106/108	-9/-9	179/181	169/169	149/149	89/99	101/111	278/298	246/267	True cross
MT188	250/250	104/106	287/289	179/181	169/169	149/149	89/99	100/100	278/298	246/267	Off-type
MT189	250/250	104/120	285/285	179/181	169/169	149/149	97/99	100/100	279/298	265/267	Off-type
MT18	250/250	106/108	285/289	179/181	157/169	149/149	99/99	111/111	278/278	263/263	Self
MT190	250/250	106/121	285/285	179/179	157/169	149/157	99/99	101/111	279/298	265/267	Off-type
MT191	250/250	106/108	-9/-9	179/179	169/169	149/149	89/99	101/111	259/278	257/263	True cross

MT192 Sample ID	250/258 NS 193	106/106 SSRY 38	287/289 SSRY 63	179/179 SSRY 102	157/157 SSRY 119	149/149 SSRY 155	89/99 SSRY 169	101/111 SSRY 147	278/298 SSRY 51	263/267 SSRY 52	True cross Conc lusions
MT193	250/258	106/108	285/289	179/179	169/169	149/149	99/99	111/111	278/298	246/267	True cross
MT194	250/258	106/106	-9/-9	179/179	157/169	149/149	89/99	111/111	278/298	246/257	True cross
MT195	250/250	106/106	285/287	179/179	169/169	149/149	99/99	111/111	259/278	263/267	True cross
MT196	250/250	106/108	285/289	179/181	157/169	149/149	89/99	101/111	278/298	257/263	True cross
MT197	250/250	106/106	287/289	179/181	157/157	149/149	99/99	111/111	278/298	263/267	True cross
MT198	250/258	104/120	285/287	179/181	157/169	149/157	99/99	100/111	259/278	257/265	Off-type
MT199	250/250	106/106	285/287	179/179	157/169	149/149	89/99	111/111	278/298	246/257	True cross
MT19	250/250	106/106	285/285	179/181	157/169	149/149	89/99	111/111	278/298	246/267	True cross
MT1	250/250	106/108	-9/-9	179/181	157/157	149/149	89/99	101/111	259/278	263/267	True cross
MT200	250/250	108/108	285/285	179/181	169/169	149/149	89/89	111/111	278/278	246/263	Self
MT202	250/250	104/104	285/289	179/181	157/169	149/149	99/99	100/111	259/278	246/267	Off-type
MT203	250/250	106/106	285/287	179/181	157/157	149/149	89/99	111/111	278/298	263/267	True cross
MT204	250/250	106/108	287/289	179/179	157/157	149/149	99/99	111/111	259/278	257/263	True cross
MT205	250/250	106/108	285/285	179/179	157/169	149/149	89/99	111/111	278/278	246/263	Self
MT206	250/258	106/108	285/285	179/181	169/169	149/149	99/99	101/111	259/278	246/267	True cross
MT207	250/250	106/108	285/285	179/181	157/169	149/149	89/89	111/111	278/278	263/263	Self
MT208	250/258	106/121	285/287	179/181	157/169	149/149	97/99	101/111	278/298	257/267	Off-type
MT209	250/258	106/108	287/287	179/181	157/169	149/149	89/99	101/111	259/278	246/263	True cross
MT20	250/258	106/108	285/289	179/181	157/169	149/149	89/99	101/111	259/278	246/257	True cross
MT210	-9/-9	106/106	-9/-9	-9/-9	157/157	149/149	-9/-9	111/111	-9/-9	-9/-9	Delete
MT211	250/256	106/106	285/287	179/181	157/169	149/157	99/99	101/111	278/298	257/267	True cross
MT212	250/250	106/121	285/287	179/179	157/169	149/157	97/99	101/111	279/298	267/267	Off-type
MT213	250/250	106/108	289/289	181/181	157/157	149/149	89/99	111/111	278/278	246/263	Self
MT214	250/258	106/108	285/287	179/181	157/169	-9/-9	89/99	111/111	278/298	257/263	True cross
MT215	-9/-9	80/80	285/287	179/181	157/169	149/157	97/99	83/83	278/298	257/265	Off-type
MT216	250/250	106/106	285/285	179/181	157/169	149/149	99/99	100/111	278/298	246/257	Off-type
MT217	250/250	106/106	289/289	179/181	157/169	149/157	89/99	101/111	259/278	246/263	True cross
MT218	-9/-9	80/80	-9/-9	179/179	157/169	149/149	89/99	83/83	259/278	246/257	Off-type
MT219	250/258	106/106	285/287	179/181	157/169	149/157	99/99	111/111	279/298	257/267	Off-type
MT21	250/250	106/121	285/287	179/179	157/169	149/149	99/99	111/111	278/279	267/267	Off-type
MT220	250/258	104/106	287/289	179/179	157/169	149/149	99/99	100/110	278/298	257/263	Off-type
MT221	250/250	106/108	285/289	179/179	157/169	149/149	99/99	111/111	278/298	246/257	True cross
MT222	250/258	106/106	-9/-9	179/181	157/169	149/149	99/99	101/111	259/259	267/267	True cross
MT223	-9/-9	80/80	-9/-9	179/179	169/169	149/149	99/99	83/83	278/278	246/263	Off-type
MT224	250/258	106/108	285/287	179/179	169/169	149/149	89/99	111/111	278/298	246/267	True cross
MT225	250/250	106/121	285/287	179/181	169/169	149/157	97/99	101/111	279/298	257/267	Off-type
MT226	250/258	106/108	285/287	179/181	157/169	149/149	99/99	101/111	259/278	257/263	True cross
MT227	250/250	106/106	285/287	179/179	157/157	149/149	99/99	101/111	259/278	257/263	True cross
MT228	250/250	106/108	285/287	179/179	169/169	149/149	89/99	111/111	259/278	263/267	True cross
MT229	250/256	108/108	285/289	179/181	157/169	149/157	97/99	101/111	278/278	263/267	Off-type
MT22	250/250	106/108	287/289	179/181	157/169	149/149	99/99	111/111	278/298	246/267	True cross
MT230	250/258	106/106	285/289	179/179	157/169	149/149	89/99	111/111	259/278	263/267	True cross
MT231	250/250	106/108	285/289	179/181	169/169	149/149	99/99	111/111	278/278	246/263	Self
MT232	250/258	106/108	287/289	179/181	-9/-9	149/149	99/99	111/111	-9/-9	246/257	True cross
MT233	250/250	106/108	-9/-9	181/181	169/169	149/149	89/89	111/111	278/278	263/263	Self

MT234 Sample ID	-9/-9 NS 193	80/80 SSRY 38	285/287 SSRY 63	179/181 SSRY 102	157/169 SSRY 119	157/157 SSRY 155	99/99 SSRY 169	-9/-9 SSRY 147	279/281 SSRY 51	265/267 SSRY 52	Off-type Conc lusions
MT236	250/250	106/108	287/289	179/179	169/169	149/149	99/99	101/111	278/298	246/257	True cross
MT237	250/250	106/108	285/285	179/179	157/169	149/149	99/99	111/111	278/298	246/267	True cross
MT238	250/256	106/121	285/285	179/181	157/169	149/157	97/99	101/111	259/278	257/267	Off-type
MT239	250/250	106/108	287/289	179/179	157/169	149/149	89/99	101/111	278/298	246/267	True cross
MT23	250/258	106/106	285/287	179/179	169/169	149/149	89/99	101/111	259/278	246/267	True cross
MT240	256/258	106/106	285/287	179/181	169/169	149/149	99/99	101/111	279/298	265/267	Off-type
MT241	250/250	108/121	-9/-9	181/181	169/169	149/149	97/99	111/111	-9/-9	263/267	Off-type
MT242	250/250	106/108	285/289	179/181	157/169	149/149	99/99	111/111	278/298	246/267	True cross
MT243	250/256	106/106	285/285	179/179	157/169	149/149	97/99	101/111	278/298	267/267	Off-type
MT244	250/250	106/106	285/287	179/179	-9/-9	149/149	99/99	101/111	-9/-9	246/267	True cross
MT245	250/250	106/108	285/285	179/181	157/169	149/149	99/99	101/111	278/298	257/263	True cross
MT246	250/258	106/121	-9/-9	179/179	157/169	149/149	99/99	111/111	278/298	257/267	Off-type
MT247	250/258	106/106	287/289	179/179	157/157	149/149	89/99	111/111	278/298	257/263	True cross
MT248	250/256	104/120	285/285	179/181	157/169	149/157	97/99	100/100	279/298	267/267	Off-type
MT249	250/256	108/108	285/285	179/181	157/169	149/157	89/99	101/111	278/278	246/267	Off-type
MT24	-9/-9	80/80	285/287	179/179	157/169	149/149	99/99	83/83	278/298	265/267	Off-type
MT250	250/258	106/108	285/287	179/179	157/169	149/149	89/99	101/111	259/278	257/263	True cross
MT251	250/250	106/108	289/289	179/181	157/157	149/149	99/99	111/111	278/278	246/263	Self
MT252	250/258	106/108	285/289	179/179	157/169	149/149	99/99	101/111	-9/-9	246/267	True cross
MT253	250/250	106/106	285/287	179/179	157/169	149/149	97/99	101/111	279/298	265/267	Off-type
MT254	250/250	106/108	287/289	179/181	157/169	149/149	99/99	101/111	278/298	246/257	True cross
MT255	250/250	-9/-9	285/287	179/179	169/169	149/149	89/99	100/100	259/278	263/267	Off-type
MT256	250/258	106/108	285/285	179/179	157/157	149/149	99/99	111/111	259/278	246/257	True cross
MT257	250/250	106/108	285/287	179/181	169/169	149/149	99/99	101/111	278/298	246/267	True cross
MT258	250/258	104/120	285/287	179/181	157/169	149/157	97/99	100/100	298/298	257/267	Off-type
MT259	250/250	106/108	287/289	179/179	157/157	149/149	89/99	111/111	259/278	246/257	True cross
MT260	250/256	121/121	-9/-9	179/181	169/169	149/149	97/99	101/111	-9/-9	-9/-9	Off-type
MT261	250/258	106/106	285/285	179/179	157/157	149/149	99/99	101/111	259/278	267/267	True cross
MT262	250/250	108/108	285/289	181/181	157/157	149/149	89/99	111/111	278/278	263/263	Self
MT263	250/250	106/106	285/285	-9/-9	157/169	149/149	89/99	111/111	278/278	-9/-9	Self
MT264	250/250	106/108	285/289	179/179	157/169	149/149	89/99	111/111	259/278	257/263	True cross
MT265	250/250	106/108	289/289	179/181	157/169	149/149	89/99	111/111	278/278	263/263	Self
MT268	250/258	106/121	285/285	179/179	169/169	149/149	99/99	101/111	259/278	265/267	Off-type
MT26	250/258	106/108	285/287	179/179	169/169	149/149	99/99	111/111	278/298	246/257	True cross
MT270	250/250	104/120	285/287	179/179	157/169	149/149	97/99	100/110	259/278	257/265	Off-type
MT271	250/258	106/121	285/287	179/179	157/157	149/149	99/99	111/111	279/298	257/267	Off-type
MT272	250/250	106/108	287/289	179/181	157/169	149/149	99/99	101/111	278/298	246/267	True cross
MT273	250/250	106/106	285/285	179/181	157/169	149/149	99/99	101/111	259/278	257/263	True cross
MT274	250/250	106/106	285/287	179/181	157/169	149/149	89/99	101/111	259/278	246/257	True cross
MT275	-9/-9	-9/-9	285/285	179/179	157/169	149/149	99/99	-9/-9	259/278	263/267	True cross
MT276	250/250	106/108	287/289	179/181	169/169	149/149	99/99	111/111	278/298	246/267	True cross
MT277	250/258	-9/-9	285/287	179/179	157/169	149/149	99/99	-9/-9	278/298	246/257	True cross
MT278	250/258	106/106	285/287	179/179	157/157	149/157	97/99	111/111	259/279	265/267	Off-type
MT279	250/250	108/108	285/289	179/181	157/169	149/149	99/99	111/111	278/278	246/263	Self
MT27	250/250	108/108	285/285	179/181	169/169	149/149	89/89	111/111	278/278	246/263	Self

MT280 Sample ID	250/250 NS 193	106/108 SSRY 38	285/287 SSRY 63	179/181 SSRY 102	157/169 SSRY 119	149/149 SSRY 155	89/99 SSRY 169	101/111 SSRY 147	259/278 SSRY 51	246/267 SSRY 52	True cross Conc lusions
MT281	250/258	106/106	285/289	179/181	157/169	149/149	99/99	111/111	278/298	246/267	True cross
MT282	250/250	106/121	285/287	179/179	169/169	149/157	97/99	101/101	279/298	257/267	Off-type
MT283	250/258	106/108	285/287	179/181	157/169	149/149	99/99	101/111	259/278	246/257	True cross
MT284	250/258	106/106	287/289	179/179	157/157	149/149	89/99	111/111	259/278	246/257	True cross
MT285	250/256	121/121	285/285	179/181	157/169	149/157	99/99	101/111	279/279	265/267	Off-type
MT286	250/258	106/106	285/287	179/179	157/157	149/157	97/99	101/111	279/298	267/267	Off-type
MT287	250/250	106/106	285/285	179/181	157/157	149/149	89/99	111/111	278/278	246/246	Self
MT288	250/250	106/108	285/289	179/179	157/169	149/149	99/99	101/111	278/298	246/267	True cross
MT289	250/250	106/108	285/289	179/179	157/157	149/149	89/99	111/111	278/298	246/267	True cross
MT28	250/250	106/108	285/287	179/181	169/169	149/149	99/99	111/111	259/278	257/263	True cross
MT290	250/258	106/121	285/287	179/179	169/169	149/149	99/99	101/111	279/298	257/265	Off-type
MT291	250/256	106/106	285/287	179/179	157/157	149/149	99/99	101/111	259/278	257/265	Off-type
MT292	250/258	106/121	285/287	179/181	157/169	149/149	97/99	111/111	259/278	257/265	Off-type
MT293	250/250	106/106	287/289	179/179	157/169	149/149	89/99	111/111	259/278	263/267	True cross
MT294	250/250	104/106	285/285	179/181	157/169	149/149	89/99	100/110	278/298	246/257	Off-type
MT296	250/250	106/108	285/289	179/181	157/157	149/149	99/99	111/111	259/278	246/267	True cross
MT297	250/250	106/121	285/285	179/181	157/157	149/157	99/99	101/111	259/279	257/265	Off-type
MT298	250/250	106/108	285/285	179/179	157/157	149/149	89/99	101/111	-9/-9	246/257	True cross
MT299	250/258	106/108	287/289	179/181	157/157	149/149	89/99	101/111	259/278	257/263	True cross
MT29	250/256	106/106	285/287	179/179	157/157	149/149	89/99	111/111	279/298	257/265	Off-type
MT2	250/258	106/106	285/289	179/181	157/157	149/149	89/99	111/111	278/298	246/267	True cross
MT300	250/256	104/120	285/287	179/181	157/169	149/157	99/99	100/100	259/279	257/265	Off-type
MT301	250/250	106/106	287/289	179/179	157/169	149/149	99/99	111/111	259/278	257/263	True cross
MT302	-9/-9	-9/-9	285/289	179/181	157/169	149/149	89/99	-9/-9	278/278	246/263	Self
MT303	250/250	106/121	285/287	179/179	157/169	149/157	97/99	-9/-9	278/279	257/265	Off-type
MT304	250/250	106/106	285/285	179/179	157/169	149/149	89/99	-9/-9	259/278	246/257	True cross
MT305	250/258	106/121	285/287	179/179	157/169	149/149	97/99	-9/-9	278/298	267/267	Off-type
MT306	250/250	106/106	285/287	179/179	157/157	149/149	89/99	101/111	259/278	246/257	True cross
MT307	250/258	106/108	287/289	179/179	157/169	149/149	89/99	-9/-9	278/298	257/263	True cross
MT308	250/258	106/108	285/287	179/179	169/169	149/149	99/99	101/111	259/278	257/263	True cross
MT309	-9/-9	-9/-9	285/285	179/181	157/157	149/149	97/99	-9/-9	278/298	257/267	Off-type
MT30	250/258	106/108	287/289	179/179	157/169	149/149	99/99	101/111	278/298	246/257	True cross
MT310	250/250	106/106	285/285	179/181	157/169	149/149	89/99	101/111	259/278	263/267	True cross
MT311	-9/-9	80/80	285/287	179/179	157/169	149/149	97/99	83/83	278/279	257/267	Off-type
MT312	250/250	106/108	285/285	179/181	157/169	149/149	89/99	101/111	278/298	263/267	True cross
MT313	250/250	106/106	287/289	179/179	157/169	149/149	89/99	111/111	259/278	246/267	True cross
MT314	-9/-9	80/80	287/289	179/181	157/157	149/149	89/99	83/83	278/298	246/267	Off-type
MT315	250/258	106/106	285/285	179/181	157/169	149/149	97/99	101/111	279/298	257/265	Off-type
MT316	250/256	106/106	285/285	179/179	157/157	149/157	97/99	101/111	259/279	257/267	Off-type
MT317	250/258	104/104	285/287	179/179	157/169	149/157	99/99	100/100	259/278	257/267	Off-type
MT318	250/250	106/121	-9/-9	-9/-9	157/169	-9/-9	99/99	111/111	-9/-9	-9/-9	Off-type
MT31	250/258	106/106	285/289	179/181	157/157	149/149	89/99	111/111	278/298	257/263	True cross
MT320	250/250	106/106	287/289	179/179	157/169	149/149	99/99	101/111	278/298	246/257	True cross
MT321	250/250	106/108	285/289	179/181	157/169	149/149	99/99	111/111	259/278	263/267	True cross
MT322	250/258	106/106	287/289	179/179	157/169	149/149	99/99	111/111	278/298	246/267	True cross

Sample ID	NS 193	SSRY 38	SSRY 63	SSRY 102	SSRY 119	SSRY 155	SSRY 169	SSRY 147	SSRY 51	SSRY 52	Conc lusions
MT323	250/250	106/108	285/287	179/181	157/169	149/149	99/99	111/111	278/298	263/267	True cross
MT324	250/258	106/106	285/285	179/181	169/169	149/149	89/99	101/111	259/278	257/263	True cross
MT325	250/258	106/106	285/287	179/181	157/157	149/157	97/99	101/111	278/281	263/267	Off-type
MT326	250/258	106/121	285/287	179/179	157/169	149/149	99/99	101/111	278/298	257/267	Off-type
MT327	250/258	106/108	285/287	179/179	157/169	149/149	99/99	101/111	259/278	246/257	True cross
MT329	250/258	106/108	287/289	179/181	157/157	149/149	89/99	111/111	278/298	263/267	True cross
MT32	250/258	106/121	285/285	179/179	157/169	149/149	99/99	101/111	298/298	257/267	Off-type
MT330	250/258	106/106	285/287	179/179	157/169	149/149	97/99	101/111	278/298	265/267	Off-type
MT331	250/258	106/108	285/287	179/181	157/169	149/149	99/99	101/111	278/298	257/263	True cross
MT332	250/250	106/106	285/289	179/181	169/169	149/149	89/99	111/111	278/298	246/267	True cross
MT333	250/250	106/121	285/287	179/181	157/169	149/157	97/99	101/101	279/298	257/267	Off-type
MT334	250/258	106/121	285/287	179/179	157/169	149/157	99/99	101/111	279/298	267/267	Off-type
MT335	250/258	106/108	285/287	179/181	169/169	149/149	99/99	101/111	259/278	257/263	True cross
MT336	250/250	106/106	287/289	179/179	157/169	149/149	99/99	101/111	278/298	246/257	True cross
MT337	250/258	106/108	285/285	179/179	157/157	149/149	99/99	101/111	278/298	263/267	True cross
MT338	250/256	106/106	285/287	179/181	157/157	149/157	97/99	101/111	259/278	257/265	Off-type
MT339	250/250	104/104	285/285	179/181	157/169	149/149	99/99	100/100	278/298	257/267	Off-type
MT33	250/250	106/106	285/289	179/181	157/157	149/149	99/99	101/111	259/278	246/267	True cross
MT340	250/258	106/108	285/287	179/181	157/157	149/149	89/99	111/111	278/298	263/267	True cross
MT341	250/250	106/108	289/289	179/181	157/157	149/149	89/99	111/111	278/278	246/246	Self
MT342	250/258	106/121	285/285	179/179	157/157	149/149	99/99	111/111	279/298	257/265	Off-type
MT343	250/256	106/106	285/285	179/179	157/169	149/157	97/99	101/111	279/298	257/265	Off-type
MT344	250/250	106/108	285/289	179/179	157/169	149/149	99/99	111/111	278/278	246/263	Self
MT345	250/258	106/108	-9/-9	179/179	157/169	149/149	99/99	111/111	259/278	246/267	True cross
MT346	250/256	106/106	285/287	179/179	169/169	149/149	97/99	101/111	-9/-9	257/265	Off-type
MT347	250/258	104/120	285/287	179/179	169/169	149/157	97/99	100/100	279/298	265/267	Off-type
MT348	250/250	106/108	285/285	179/179	157/169	149/149	89/99	101/111	278/298	263/267	True cross
MT349	-9/-9	80/80	285/285	179/179	157/169	149/149	99/99	83/83	259/278	263/267	Off-type
MT350	250/250	106/121	285/287	179/179	157/169	149/149	97/99	101/111	278/298	257/265	Off-type
MT36	250/258	-9/-9	285/285	179/179	157/157	149/149	99/99	111/111	278/298	246/257	True cross
MT38	250/250	-9/-9	285/285	179/181	157/169	149/149	89/99	111/111	278/298	246/267	True cross
MT39	250/258	106/121	285/285	179/181	157/157	149/157	97/99	111/111	278/298	265/267	Off-type
MT3	250/258	106/108	287/289	179/179	169/169	149/149	89/99	111/111	259/278	257/263	True cross
MT40	250/250	106/106	285/289	179/181	157/157	149/149	89/99	111/111	278/278	246/263	Self
MT41	250/258	106/106	285/285	179/179	169/169	149/149	99/99	101/111	278/298	246/257	True cross
MT42	250/250	106/121	285/285	179/179	169/169	149/157	97/99	101/101	278/298	267/267	Off-type
MT43	250/256	106/106	285/287	179/179	157/157	149/149	99/99	111/111	-9/-9	257/265	Off-type
MT44	250/250	106/106	285/287	179/181	157/169	149/149	89/99	111/111	278/298	263/267	True cross
MT45	250/258	106/106	285/287	179/179	157/169	149/149	89/99	111/111	259/278	246/267	True cross
MT46	250/258	106/106	285/287	179/181	157/157	149/149	89/99	101/111	259/278	257/263	True cross
MT47	250/258	106/106	285/287	179/179	157/157	149/149	-9/-9	111/111	259/278	246/257	True cross
MT48	250/256	104/120	285/285	179/179	157/169	149/157	97/99	100/100	259/279	267/267	Off-type
MT49	250/258	106/106	-9/-9	181/181	169/169	149/157	89/89	101/101	281/298	263/267	Off-type
MT4	256/258	104/104	285/287	179/179	157/169	149/157	99/99	100/100	278/298	265/267	Off-type
MT50	250/250	106/108	285/287	179/179	169/169	149/149	99/99	111/111	259/278	246/257	True cross
MT51	250/250	106/108	285/289	179/179	169/169	149/149	89/99	101/111	259/278	246/267	True cross

MT52 Sample ID	250/250 NS 193	106/106 SSRY 38	285/289 SSRY 63	179/181 SSRY 102	157/169 SSRY 119	149/149 SSRY 155	99/99 SSRY 169	101/111 SSRY 147	278/298 SSRY 51	246/267 SSRY 52	True cross Conc lusions
MT53	250/250	106/106	-9/-9	179/181	157/157	149/149	89/99	111/111	278/298	257/263	True cross
MT54	250/258	106/106	-9/-9	179/181	169/169	149/149	99/99	101/111	259/278	263/267	True cross
MT55	250/258	106/108	285/285	179/181	157/169	149/149	89/99	101/111	259/278	246/267	True cross
MT56	256/258	106/106	285/285	179/179	157/169	149/149	97/99	111/111	259/279	257/267	Off-type
MT57	250/258	106/106	-9/-9	179/181	157/157	149/157	99/99	101/111	278/279	263/263	Off-type
MT58	250/258	106/106	285/285	179/181	157/169	149/149	99/99	111/111	278/279	257/265	Off-type
MT59	250/250	106/106	285/285	179/181	157/157	149/149	89/99	111/111	278/278	246/263	Self
MT5	250/258	104/104	285/287	179/179	157/157	149/149	97/99	100/100	279/298	257/267	Off-type
MT60	250/256	104/120	285/285	179/181	157/169	149/149	97/99	100/100	279/298	257/265	Off-type
MT61	250/250	106/106	287/289	179/181	157/157	149/149	99/99	111/111	-9/-9	-9/-9	True cross
MT62	250/250	106/108	285/289	179/179	169/169	149/149	99/99	111/111	278/278	246/263	Self
MT63	250/250	106/108	287/289	179/179	157/169	149/149	89/99	111/111	278/298	263/267	True cross
MT64	250/250	106/108	285/285	179/179	169/169	149/149	89/99	111/111	278/298	257/263	True cross
MT65	250/258	106/106	285/287	179/181	157/157	149/157	89/99	101/111	278/279	246/265	Off-type
MT66	250/250	106/108	287/289	179/179	157/169	149/149	99/99	101/111	278/298	257/263	True cross
MT67	250/250	106/106	285/289	179/181	157/157	149/149	89/99	111/111	278/298	246/267	True cross
MT68	250/258	106/108	285/285	179/181	157/157	149/149	99/99	101/111	278/298	263/267	True cross
MT69	250/258	106/108	285/285	179/179	157/169	149/149	89/99	101/111	259/278	263/267	True cross
MT6	250/250	106/108	285/287	179/179	157/169	149/149	89/99	101/111	278/298	246/267	True cross
MT70	256/258	106/106	-9/-9	179/181	157/169	149/157	97/99	101/111	278/298	267/267	Off-type
MT71	250/258	106/106	285/285	179/179	157/169	149/149	99/99	101/111	278/298	246/267	True cross
MT72	250/250	106/106	289/289	179/181	157/157	149/149	99/99	111/111	278/278	246/246	Self
MT73	250/258	108/108	285/287	179/181	157/157	149/157	89/99	101/111	278/279	246/263	Off-type
MT74	250/250	106/106	285/289	179/181	157/157	149/149	99/99	101/111	278/298	246/263	True cross
MT75	250/256	106/121	285/285	179/179	169/169	149/149	99/99	111/111	279/298	267/267	Off-type
MT76	250/250	106/106	287/289	179/179	157/169	149/149	99/99	101/111	278/298	263/267	True cross
MT77	250/258	106/106	285/285	179/179	157/157	149/157	99/99	111/111	259/278	257/267	True cross
MT78	250/250	104/106	285/285	179/181	157/169	149/149	99/99	100/110	259/278	246/257	Off-type
MT79	256/258	104/104	285/287	179/181	157/157	149/157	97/99	100/110	278/298	257/267	Off-type
MT7	256/258	106/106	285/287	179/179	157/157	149/157	97/99	101/101	278/298	257/265	Off-type
MT80	250/258	106/106	285/289	179/181	-9/-9	149/149	99/99	101/111	278/298	246/267	True cross
MT81	250/258	106/106	285/287	179/179	169/169	149/149	89/99	111/111	259/278	263/267	True cross
MT82	250/258	106/121	285/287	179/181	157/169	149/149	97/99	101/111	259/279	257/265	Off-type
MT83	250/258	104/104	287/289	179/181	157/157	149/149	99/99	100/110	259/278	263/267	Off-type
MT84	-9/-9	-9/-9	285/285	179/179	157/157	149/157	99/99	-9/-9	259/279	257/267	Off-type
MT85	256/258	106/106	285/285	179/179	169/169	149/157	97/99	111/111	259/279	267/267	Off-type
MT86	250/258	106/106	285/285	179/179	157/169	149/149	97/99	111/111	259/279	257/265	Off-type
MT87	250/258	104/106	285/287	179/181	157/169	149/149	89/99	100/110	278/298	263/267	Off-type
MT88	250/250	106/106	285/289	179/179	157/169	149/149	99/99	111/111	278/278	263/263	Self
MT89	250/250	106/108	285/289	181/181	157/157	149/149	89/89	111/111	278/278	246/263	Self
MT8	250/258	106/106	285/287	179/179	157/169	149/149	-9/-9	111/111	259/279	257/265	Off-type
MT90	250/250	106/106	285/287	179/179	157/169	149/149	97/99	101/111	278/279	257/267	Off-type
MT91	250/258	104/106	285/287	179/179	169/169	149/149	89/99	100/110	278/298	257/263	Off-type
MT92	-9/-9	-9/-9	285/285	179/181	157/157	149/157	97/99	-9/-9	259/279	257/265	Off-type
MT93	250/258	104/106	285/287	179/181	157/169	149/149	89/99	100/110	278/298	246/267	Off-type

Sample ID	NS 193	SSRY 38	SSRY 63	SSRY 102	SSRY 119	SSRY 155	SSRY 169	SSRY 147	SSRY 51	SSRY 52	Conclusions
MT94	250/258	106/106	285/285	179/181	169/169	149/149	89/99	101/111	259/278	246/257	True cross
MT95	250/258	106/121	285/285	179/181	157/169	149/149	97/99	111/111	279/298	257/265	Off-type
MT96	250/258	106/106	285/289	179/179	157/169	149/149	89/99	111/111	278/298	246/267	True cross
MT97	250/258	106/106	285/287	179/181	169/169	149/149	99/99	101/111	278/298	263/267	True cross
MT98	250/256	106/121	285/287	179/179	157/157	149/157	97/99	101/111	298/298	265/267	Off-type
MT99	250/250	106/108	285/285	181/181	157/169	149/149	89/99	111/111	278/278	246/263	Self
MT9	250/250	106/104	285/289	179/181	157/169	149/149	89/99	100/110	259/278	246/257	Off-type
MK	250/250	106/108	285/289	179/181	157/169	149/149	89/99	111/111	278/278	246/263	
TMS	250/258	106/106	285/287	179/179	157/157	149/157	99/99	101/111	259/298	257/267	

Key

Off-type

Self

Exactly as female

Exactly as male or male self

True crosses

-9/-9 Missing data