MOLECULAR CHARACTERIZATION OF CASSAVA MOSAIC DISEASE (CMD) RESISTANT VARIETIES AND WILD RELATIVES OF CASSAVA (*Manihot esculenta* Crantz) USING SSR AND SNP MARKERS

by

DHANYA O.G.

(2010-09-114)

THESIS

Submitted in partial fulfilment of the

requirement for the degree of

MASTER OF SCIENCE (INTEGRATED) IN BIOTECHNOLOGY

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Department of Plant Biotechnology

B.Sc - M.Sc. Integrated Biotechnology

COLLEGE OF AGRICULTURE

VELLAYANI, THIRUVANANTHAPURAM-695 522

KERALA, INDIA

2015

DECLARATION

I hereby declare that this thesis entitled **"MOLECULAR CHARACTERIZATION OF CASSAVA MOSAIC DISEASE (CMD) RESISTANT VARIETIES AND WILD RELATIVES OF CASSAVA (*Manihot esculenta* Crantz) USING SSR AND SNP MARKERS"** is a bonafide record of research work done by me during the course of research and that the thesis has not previously formed the basis for the award of any degree, diploma, associate ship, fellowship or other similar title, of any other university or society.

Place: Vellayani

Date: 11/01/2016

DHANYA O.G.

(2010-09-114)

भाकृअनुप-केन्द्रीय कन्द फसल अनुसंधान संस्थान
(भारतीय कृषि अनुसंधान परिषद)
श्रीकार्यम, तिरुवनन्तपुरम -६९५०१७, केरल,भारत
ICAR-CENTRAL TUBER CROPS RESEARCH INSTITUTE
(Indian Council of Agricultural Research)
Sreekariyam, Thiruvananthapuram-695 017, Kerala, India
ISO 9001:2008

Dr. C. Mohan
Principal Scientist

CERTIFICATE

Certified that this thesis entitled "**Molecular characterization of cassava mosaic disease (CMD) resistant varieties and wild relatives of cassava (*Manihot esculenta* Crantz) using SSR and SNP markers**" is a record of research work done independently by Dhanya O. G. (2010-09-114) under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to her.

Place : CTCRI, Trivandurm
Date : 11.01.2016

Dr. C. Mohan
(Chairman of the Advisory Committee)
Principal Scientist (Plant Breeding)

Phone : 91-0471-2598551 to 2598554
Director (Per.) : 91-0471-2598431
(Res.) : 91-0471 - 2441957
Admn.Officer : 91-0471-2598193
RIGHT TO INFORMATION
Fax : 91-0471-2590063
E-Mail: ctcritvm@yahoo.com
ctcritvm@gmail.com
Web : http://www.ctcri.org

CERTIFICATE

We, the undersigned members of the advisory committee of Ms. Dhanya O.G., a candidate for the degree of Master of Science (Integrated) in Biotechnology, agree that the thesis entitled **"Molecular characterization of cassava mosaic disease (CMD) resistant varieties and wild relatives of cassava (*Manihot esculenta* Crantz) using SSR and SNP markers"** may be submitted by Ms. Dhanya O.G., in partial fulfillment of the requirement for the degree

11/1/16

Dr. C. Mohan
(Chairman, Advisory Committee)
Principal Scientist
Division of Crop Improvement
ICAR- Central Tuber Crops Research Institute
Sreekariyam
Thiruvananthapuram-695 017

11/1/16

Dr. B. R. Reghunath
(Member, Advisory Committee)
Professor & Head
Department of Plant Biotechnology
College of Agriculture, Vellayani
Thiruvananthapuram-695 522

11/1/2016

Dr. K. I. Asha
(Member, Advisory Committee)
Principal Scientist
Division of Crop Improvement
Central Tuber Crops Research Institute
Sreekariyam
Thiruvananthapuram-695 017

10/1/16

Dr. J. Sreekumar
(Member, Advisory Committee)
Senior Scientist
Section of Social Science and Extension
Central Tuber Crops Research Institute,
Sreekariyam,
Thiruvananthapuram-695 017

11/1/2016

Dr. Deepa S. Nair
(Member, Advisory Committee)
Assistant Professor
Department of Plant Biotechnology
College of Agriculture, Vellayani
Thiruvananthapuram- 695 522

ACKNOWLEDGEMENT

I place my deep sense of reverence in the lotus feet of the God almighty who has Showered his boundless blessings on me enabling the successful completion of this research work.

I takes me immense pleasure to express my deep sense of gratitude to my chairman of advisory committee, Dr. C. Mohan, principal scientist, Division of division of crop improvement, Central Tuber Crops Research Institute Sreekaryam , Thiruvanandhapuram, for having given me his intimous guidance, constant encouragement, constructive criticism, incessant inspiration, sustained patience and untiring attention at each stage of the research work.

I am thankful to all the members of my advisory committee, Dr. B. R. Reghunath, Dr. K. I. Asha, Dr. J. Sreekumar and Dr. Deepa S. Nair for their valuable counseling and constructive suggestions that were much helpful throughout my research progress.

I place on record my sincere and whole hearted thanks to Dr. Swapna Alex our course director for her encouragement and help rendered during the course of the study.

I will always remember with gratitude our dearest, retired professor of plant biotechnology Dr. K. Rajmohan, former course director Dr. Lekha Sreekandan for their valuable suggestions and advice during this M.Sc. Programme.

This research work would not have been possible without the help of Ms. Vidhya, Mss Aswathy and Vishalakshi mam so i am very thankful to them for their help during our work.

I wish to express my deep gratitude to Ramadhasan sir, Prakashettan, Abhilashettan, Sreekumarettan members of CTCRI for their timely help.

I sincerely thank my dear friends Aishu, kuttan, Amru, Rini, Jaggu, Raju, Achu, Paru, Sudheer, Roshan, Ambu, and Edwin they always be stood with me in all my happiness, difficulties and sorrow.

I owe a great deal to my beloved parents and my ever loving brother Vichumon, loving sister Kechu and my relatives for their blessings, continuous encouragement, overwhelming interest and guidance which they showered on me throughout my study

I humbly dedicate my thesis at the feet of my parents and god whose grace and blessings, love, continuous support and encouragement made it possible for me.

DHANYA O.G.

CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS AND SYMBOLS USED

CMD	Cassava mosaic disease
SSR	Simple sequence repeats
SNP	Single nucleotide polymorphism
DNA	Deoxyribose nucleic acid
UPGMA	Unweighted Pair-Group method
PIC	Polymorphism information content
PCA	Principal component analysis
ssDNA	Single stranded DNA
ORSs	Open reading frames
CMV	Cassava mosaic virus
AFLP	Amplified fragment length polymorphism
ISSR	Inter simple sequence repeats
RFLP	Restriction fragment length polymorphism
%	Per cent
mM	millimolar
µl	Micro litre
@	At the rate of
°C	Degree Celsius
bp	Base pair
et al.	And other co workers

Cont

LIST OF ABBREVIATIONS AND SYMBOLS

Fig.	Figure
g	Gram
g-1	Per gram
mg	Milli gram
ml	Millilitre
sec	Seconds
min	Minutes
ng	Nanogram
µM	Micromolar
PCR	Polymerase chain reaction

LIST OF APPENDICES

Introduction

1. INTRODUCTION

Cassava (*Manihot esculenta*) is the most widely cultivated tuber crop in tropics and is grown across a broad range of agro-climatic conditions. It is grown for various end uses such as human food, animal feed and as industrial raw material. Much of the world cassava production is centered in Sub-Saharan Africa, with a production of 167 MT in 2014 and would account for some 57 per cent of the global total (FAO, 2014). Cassava was introduced into India by the Portuguese during the 16th century as a food crop. The importance of cassava as a food crop was well recognized in Kerala, south India during the 20th century, when famine struck India at the time of Second World War. Adaptability to different soils, ability to establish in high as well as low rainfall areas, and relative resistance to pests and diseases are the few factors that helped to anchor cassava in India.

Cassava is monoecious with 36 chromosomes, and is highly heterozygous due to its out-crossing nature. Along with these characters the diploid, cross pollinated, recombination and gene flow during traditional farming practices, and increased species variations on wild cassava collections also contribute to increased genetic variability on cassava. More emphasis is given for improvements in cassava to target CMD resistance, early maturity, high dry matter content, reducing antinutrient factors like cyanogens and increasing nutrient contents like Provitamin *A* and proteins. In cassava, disease resistant breeding for CMD resistance is one of the important targets in cassava improvement.

Cassava mosaic disease (CMD) is one of the most important constraints in cassava production which will cause an average yield loss of 50 per cent (Fauquet and Fargette, 1990). It is a viral disease caused by gemini viruses of the genus Begomovirus (Family Geminiviridae) transmitted by a white fly vector, *Bemisia tabaci*. Different approaches followed to control CMD, such as use of virus free

planting materials and deployment of resistant varieties. The disease is best kept under control by exploiting the available host plant resistance, which was introgressed from *M. glaziovii* (*CMD1*) to cassava was recessive with a heritability of about 60 per cent (Jennings, 1976). Akano *et al.*, (2002) reported that the resistance could be due to a major dominant gene (*CMD2*) using Nigerian landraces. Bulk sergeant analysis on F_1 population derived from a cross between TME7 × TMS30555 cross using SSR and AFLP markers introduce more resistance genes CMD3 and CMD4 (Okogbenin *et al.*, 2008). The complexity of the genetics behind the CMD resistance indicates the possibility of presence of several components responsible for CMD resistance.

Contributions of such disease resistance genes as well as other desired characters on cassava can be highly promoted by the exploitation of specific characters present in wild relatives of cassava. They are the progenitors of cassava and constitute valuable genetic reservoirs presenting genes that show new characters, so it become frequently used as an important source of genetic diversity and have been employed effectively in a variety of breeding programs.

One of the major challenges in cassava breeding is the quest to overcome the evolutionary capacity of the disease by identifying additional sources of resistance to the disease. So the need to identify other sources of CMD resistance is critical in building durable and stable resistance to CMD through genetic diversity studies. The overwhelming developments in molecular marker technology have generated various DNA marker systems. Among those systems number of recent studies on cassava research program using SSR and SNP markers reveals a better performance than all other markers and it is regarded as one of the most powerful tools for the analysis of plant genomes and the association of heritable traits with underlying genetic variations. To date there are nearly 1000 SSR markers and 10,546 SNPs from cassava draft genome sequence and ESTs from GenBank were discovered. The

reducing cost of DNA sequencing and increasing availability of large sequence data sets permits the mining of this data for large numbers of SSRs and SNPs. So it can contribute to wide applications in cassava research such as genetic linkage analysis, trait mapping, diversity analysis, association studies, marker-assisted selection *etc.*

With this view the present study is mainly formulated for the exploitation of the SSR and SNP markers to reveals the extent of genetic diversity associated with CMD resistance in cassava accessions to increase the levels of resistance within the gene pool by identifying additional sources of resistance with a wider genetic base. The study is undertaken with the following objectives

- To analyse the extent of genetic variations, and phylogenetic relationship among CMD resistant and wild relatives of cassava using SSR and SNP markers.

Review of literature

2. REVIEW OF LITERATURE

Origin

Cassava has its genetic, geographical and agricultural origin in Latin America. Its domestication began 5000 – 7000 years BC in the Amazon, Brazil (Allen, 2002) and it was distributed by Europeans to the rest of the world (Henry and Hershey, 2002). Cassava was taken from Brazil to the West coast of Africa by Portuguese navigators in the 16th century (Jones, 1959). Most traditional domestication hypotheses have envisioned the crop to be a "compilospecies" derived from one or more species complexes, either in Mexico and Central America (Rogers and Appan, 1973) or throughout the Neotropics (Sauer, 1993; Rogers, 1963; Ugent *et al.*, 1986). However, evidence provided from recent findings studying the glyceraldehyde 3-phosphate dehydrogenase (*G3pdh*) gene showed that the crop does not seem to be derived from several progenitor species, and cassava does not share haplotypes with *Manihot pruinosa*, a closely related, potentially hybridizing species (Olsen and Schaal, 1999). The *G3pdh* locus provides high levels of non-coding sequence variation in cassava and its wild relatives. The studies showed that cassava was likely domesticated from wild *M. esculenta* populations along the Southern border of the Amazon basin. Furthermore, studies using single nucleotide polymorphisms (SNPs) and simple sequence repeat (SSR) markers showed that cassava was likely domesticated from the wild *Manihot* species, *Manihot esculenta* ssp. *flabellifolia* (Olsen and Schaal, 2001; Olsen, 2004). These findings have provided the clearest insight to date on cassava's origin.

Economic importance

Cassava is one of the most important and widely grown root crops in Africa, Asia and South America with a total production of over 250 MT (FAO, 2013). It is a starchy root crop that supplies carbohydrate energy to millions of people in the tropics (Ceballos *et al.*, 2004) and being used increasingly as an industrial crop

(Jansson *et al.*, 2009). Though its remarkable ability to tolerate unfavourable conditions such as drought and poor soils makes it a food security crop in many parts of Sub-Saharan Africa (SSA), on-farm productivity of cassava has remained stagnant for many years due to several production constraints. The storage roots form the basic carbohydrate component of the diet and the leaves, which contain appreciable amounts of vitamins, minerals and proteins are consumed as a preferred green vegetable in many parts of Africa, providing protein, mineral and vitamins (Hahn, 1989; Lancaster and Brooks, 1983). It can be harvested anytime from 6 to 24 months after planting, and can be left in the ground as a food reserve for household food security in times of famine, drought and war (FAO, 2000; Best and Henry, 1992; Cock, 1985).

Wild cassava

Wild species of *Manihot* are progenitors of cassava. About 98 species belonging to the genus *Manihot* (Rogers and Appan, 1973), ranging from shrubs, to trees, are known. The majority of them produces latex and contains cyanogenic glucoside. Chromosome number in all species investigated is $2n = 36$. In spite of this high chromosome number, *Manihot* species behave meiotically as diploids. This seems to have preceded the emergence of the group as a whole and to have being responsible for their rapid speciation and their weak interspecific barriers, leading to interspecific hybridization. An extremely heterozygous gene pool is thus created, followed by differentiation; this begins a sequence of hybridization followed by speciation.

Exploitation of wild cassava

Cassava cultivars are lacking in many economically important characters such as resistance to insects, diseases, drought, and have low protein content (Nassar and Dorea, 1982; Nassar and Grattapaglia, 1986). This can be attributed to the mode

of evolution of the species and modifications of the allogamy system of the plant (Nassar and O'Hair, 1985). Lost genes can be restored to the gene pool of the cultigen by interspecific hybridization with wild relatives which possess these genes (Nassar *et al.,* 1986). Wild species of cultivated crops have been frequently used as an important source of genetic diversity and have been employed effectively in a variety of breeding programs. Controlled introgression of genes could alleviate stress problems in cassava in view of the availability of wild relatives which exhibit diversity in adaptation and attributes (Nassar, 1985).

Marker reveals on wild Cassava

Grattapaglia *et al.,* (1986) conducted a biosystematic analalysis of 19 wild *Manihot* species based on soluble seed protein patterns through which the several species were found to be highly similar, *viz., M. fruticolosa, M. pentaphyla, M. pilosa* and *M. corymbiflora.* Nassar *et al.,* (1996) studied 11 interspecific hybrids of cassava with wild relatives and detected high frequencies of dyads formation: 3.7% in a progeny of the hybrid of *M. glaziovii* with *M. esculenta.*

A study was conducted to identify the allele polymorphism, degree of relationship and cross species amplification of cassava microsatellites in 6 different species of *Manihot* genes. In this study, greater allelic diversity was noticed in wild species than in cassava. Among 124 alleles, 79 alleles including unique ones, were detected in the wild species but were not found in the crop (Roa *et al.,* 2000). The utility of SSR primed PCR markers for germplasm assessment was determined in a study by Carvalho and Schaal (2001). These markers as well as RAPD markers was used to characterize the Brazilian collection. In this study cassava and its wild relatives confirmed the close relationship of cassava, *M.esculenta sp.* esculenta to *M. esculenta* sp. *flabelifolia* as well as identified several other closely related wild species and PCR based markers indicated a strong grouping of varieties related to the region of cultivation in Brazil

Olsen (2004) made an analysis of SNPs and SSR variations using 20 cassava varieties from germplasm diversity and 212 wild collections to trace evolutionary and geographical origin of cassava. The resulting SNP and indel variation examined on *BglA* and *Hnl* nuclear genes when compared with the *G3pdh* nuclear gene revealed cassava was likely domesticated from a single wild *Manihot* species, *M. esculenta ssp. flabellifolia*, rather than from multiple hybridizing species, as traditionally believed. Recently Gedil *et al.*, (2012) carried out the isolation of resistance gene analogues (RGAs) from genomic DNA and cDNA in cassava, wild *Manihot* species, and castor bean (*Ricinus communis*) leading to the development of functional gene targeted markers that can be used in molecular resistance breeding aimed at combating cassava brown streak disease (CBSD) and CMD .

Cassava mosaic disease (CMD)

CMD, a severe viral disease is one of the most important biotic stresses that limit cassava production. The main symptoms are chlorotic areas intermixing with normal green tissue, which gives the mosaic patterns. In severe conditions, leaves will be with reduced size, twisted and distorted. It will also affects the plant growth resulting in significant reduction in plant height, stem girth, petiole length and leaf size and causes 25-80% yield reduction depending upon the varieties. The main spread of these viruses is through the use of infected planting material (Fargette *et al., 1988).*

Cassava viruses

CMD is caused by white fly-transmitted viruses of the genus Begomovirus (family Geminiviridae). Geminiviruses are small plant viruses with circular single stranded DNA (ssDNA) genomes encapsidated in twinned isometric particles. Virus replicated by a rolling circle mechanism in the plant cell nucleus. These geminiviruses possess two DNA molecules called DNA-A and B. DNA-A encodes all viral proteins necessary for replication and encapsidation of both components,

while the DNA-B component encodes for two proteins necessary for efficient systemic spread of the virus throughout the plant. The two DNAs share a common region which is approximately 200 bp in size and has 90- 100 per cent sequence homology between the A and B components.

Analysis of DNA-A and DNA-B sequences revealed the presence of six open reading frames (ORSs), four on DNA-A and two on DNA-B and a conserved intergenic common region of approximately 200 bp shared by the two DNA components. Genes on DNA-A were shown to code for a coat protein (AV1, CP), the replication associated protein (ACI, Rep) and proteins associated with movement (AV2), transactivation of AVI and BV1 (AC2, TrAP) and replication enhancement (AC3, REn). Genes on DNA-B were demonstrated to have important roles in nuclear transport (BV1, NSP) and cell to cell movement (BC1, MP) and these components are involved in virus spread throughout the plant, symptom production and host range (Legg and Fauquet, 2004).

Nine distinct cassava mosaic virus (CMV) have been characterized worldwide from CMD- affected cassava plants and seven of them are from SSA. These viruses are African cassava mosaic virus (ACMV), East African cassava mosaic virus (EACMV), East African cassava mosaic Cameroon virus (EACMCV), East African cassava mosaic Kenya virus (EACMKV), East African cassava mosaic Malawi virus (EACMMV), East African cassava mosaic Zanzibar virus (EACMZV) and South African cassava mosaic virus (SACMV). Two other viruses, Indian cassava mosaic virus (ICMV), and Sri Lankan cassava mosaic virus (SLCMV) were reported from the Indian sub-continent (Asare *et al.*, 2014).

CMD resistance in cassava

Among different methods for controlling CMD the most effective one is the deployment of resistant varieties (Thresh *et al.,* 1997). Quantitative and qualitative

resistance mechanisms are the main two resistance mechanism against CMD in cassava. Here the quantitative resistance derived from *M. glaziovii,* was conferred as single resistance gene. Genetic studies reveal that the polygenic resistance from *M. glaziovii* (*CMD1*) is recessive with a heritability of about 60% (Jennings, 1976), its effect is detected only in back cross progeny and not in the F_1 (Akano *et al.,* 2002) and the qualitative resistance is given by a single gene (*CMD2*) with dominant effect.

It is found that there was a high level of resistance to CMD found in several Nigerian cassava land races including TME3 (Dixon *et al.,* 2001; Akano *et al.,* 2002 and Fregene *et al.,* 2004). By using classical genetic techniques such as genetic mapping, resistance in several cassava cultivars was thought to be attributed to the presence of a major dominant resistance (R) gene, *viz., CMD2* (Akano *et al.,* 2002 and Fregene *et al.,* 2004). Furthermore, several molecular markers associated with *CMD2,* including SSRY28, NS158 and RME1 (Akano *et al.,* 2002) also identified.

Mohan *et al.* (2013) reported that, the variation for the resistance to CMD based on the damage scores across the 141 progenies and the segregating marker data, detected four markers viz., SSRY28, SSRY235, SSRY44 and NS136. SSRY28 and SSRY235 which were located on linkage group G and SSRY44 and NS136 on linkage group P (Fregene *et al.,* 1997). Among the four markers, three (SSRY235, SSRY44 and NS136) are new markers associated with CMD resistance. Fregene *et al.,* (1997) identified the markers *viz.,* SSRY235, SSRY28 and NS158 on linkage group G closer to each other. The strong association of SSRY28 and SSRY235 located on the same linkage group G with resistance to CMD established the possibility of having a major QTL for resistance to CMD in that region.

Quantitative Trait Loci (QTLs) associated with resistance to CMD were identified using F_1 progeny derived from a cross between the CMD resistant (TMS30572) and the susceptible landrace (TME117). Seven highly significant marker-associated (SSRY6, 7, 21a, 21b, 42, 77, 324) QTL effects, explaining

between 7.31 and 12.15 per cent of the total phenotypic variation, were detected by regression and Kruskal Wallis analysis. Significant marker trait associations were due to markers donated by both parents, which confirms the polygenic and recessive nature of this source of resistance (Lokko *et al.*, 2004).

The discovery of a new QTL (*CMD3*) associated with a marker NS198 for CMD resistance in TMS 97/2205 done by a F_1 population derived from a TMS 97/2205 × NR 8083 cross using 530 SSR markers offers new opportunities for pyramiding CMD genes for enhanced durability resistance in cassava (Okogbenin *et al.*, 2012).

A bulk sergeant analysis (BSA) on F_1 population derived from TME7 × TMS30555 cross using SSR and AFLP markers revealed that an SSR marker SSRY28, donated by the resistant parent was linked with resistance to CMD (Lokko *et al.*, 2005). An another BSA using 150 polymorphic SSR marker found out two markers viz., NS198 and NS199 were significantly associated to new source of CMD resistance (Okogbenin *et al.*, 2008). Here the CMD gene linked to NS198 is designated *CMD3* while that linked to NS119 is designated *CMD4*. The alleles for these markers in both families were from the male parent TMS2205.

A comparison of four makers NS198, SSRY28, NS158 and RME1 among four field resistant cassava genotypes revealed markers RME1 and NS198 were more reliable for the detection of the *CMD2* resistance gene (Asare *et al.*, 2014).

Molecular genetics in cassava

Genetic markers represents the genetic differences between individual organisms or species and are located in close proximity to genes may be referred to as gene tags. Such markers do not affect the trait of interest and are located only near or linked to genes controlling the trait.

There are three types of genetic markers, morphological markers, biochemical markers and DNA markers. Morphological markers are usually visually characterized phenotypic characters such as flower color, seed shape, growth habits or pigmentation (Sumarani *et al.*, 2004), whereas biochemical markers are mainly proteins that can be extracted and observed *viz.*, isozymes and storage proteins. The major disadvantages of these two markers are that they are limited in number and are influenced by environmental factors or the developmental stage of the plant (Winter and Kahl, 1995).

DNA markers are the most widely used type of marker predominantly due to their abundance. They arise from different classes of DNA mutations such as substitutions (point mutations), rearrangements (insertions or deletions) or errors in replication of randomly repeated DNA (Paterson, 1996). Keys to the emergence of genomics were advances in DNA marker technology resulting in a wealth of genetic markers including allozymes, mtDNA (mitochondrial DNA), RFLPs, RAPDs, AFLPs, microsatellites SNPs, and ESTs (Expressed sequence tag) with potentially widespread utility in cassava breeding programmes.

Isozymes

Isozyme diversity was done in a collection of 365 cassava cultivars and 109 accessions from wild relatives in Africa (Lefervre and Charrier, 1993). Seventeen polymorphic loci were found for the ten enzyme systems, with 59 alleles. All the markers showed disomic heredity and three linkage groups were identified. Studies at CIAT found out one of the most informative isozyme marker, α- β- esterase (Ocampo *et al.*, 1993). Most recently there are four different enzyme systems (acid phosphatase (PAC), esterase (EST), malate dehydrogenase (MDH) and phosphoglucose isomerase (PGI) gave successful enzymatic polymorphism for genetic diversity in cassava (Edmond *et al.*, 2015).

Randomly amplified polymorphic DNA (RAPD)

RAPD analysis, a PCR based molecular marker technique, was developed independently by Welsh and McClelland (1990) and Williams *et al.,* (1990). They generated by PCR amplification of random genomic DNA segments with single synthetic decamer primers of arbitrary sequence (Williams *et al.*, 1990; Jacobson and Hedrén, 2007).

RAPD markers developed for cassava were to characterize elite cassava germplasm held at CIAT (Bonierbale *et al.,* 1995; Gomez *et al.*, 1994). They were also used to estimate genetic distances in African cultivars (Marmey *et al.*, 1994; Mignouna and Dixon, 1997) and Brazilian cultivars (Carvalho *et al.,* 1993). Colombo *et al.*, (1998) investigate the genetic diversity of 31 Brazilian cassava clones using RAPD markers. The results were compared with the genetic diversity revealed by botanical descriptors. Both sets of varieties revealed identical relationships among the cultivars.

The 126 genotypes of cassava studied for genetic diversity studies using RAPD resulting in important implications for cassava germplasm collections and genetic breeding (Colombo *et al.,* 2000). Eighty eight polymorphic bands were analysed with weak genetic structure of the cassava. The PCA pattern revealed an overlapping of the São Paulo State genotype, the Brazilian group and the core collection accessions (CIAT).

In cassava, 30 accessions with yellow-orange root color characterized by RAPD marker exploited for the development of cassava varieties with higher β-carotene contents (Ferreira *et al.*, 2007).

Amplified fragment length polymorphism (AFLP)

The AFLP technique combines components of RFLP analysis with PCR technology (Vos *et al.,* 1995). Here total genomic DNA is digested with a pair of

restriction enzymes, normally a frequent and a rare cutter. Adaptors of known sequence are then ligated to the DNA fragments. Primers complementary to the adaptors are used to amplify the restriction fragments.

AFLP was used to assess the resistance and susceptibility to CMD in 20 land races and nine elite lines of cassava from Africa and eleven accessions from Latin America. AFLP data from all accessions was analyzed by both the unweighted pair group mean average (UPGMA) and multiple cluster analysis (MCA) methods. Results revealed a genetic divergence between African and Latin American accessions, although some overlapping was found between them. African land races resistant to CMD, were also found to be genetically differentiated from susceptible land races and from resistant elite lines (Fregene *et al.*, 2000).

The importance of interaction between human and ecological factors in the dynamics of genetic diversity in cassava was assessed for 31 varieties of cassava traditionally grown in Guyana and wild cassava accessions using AFLP markers. While clonality of the varieties was expected due to the vegetative propagation of cassava, 21 varieties presented intravarietal polymorphism. All wild forms of cassava clustered together and separately from the cultivated varieties in a neighbour joining dendrogram. These results are consistent with the hypothesis of a limited domestication event in a restricted area, followed by rapid diffusion of cultivated phenotypes and convergent evolution (Elias *et al.*, 2000). Raji *et al.*, (2009) performed a combinational marker study using both 20 AFLP and 50 SSR markers among a collection of African cassava landraces and elite cultivars.

Inter simple sequence repeat (ISSR)

ISSR is a marker technique, which involves the use of microsatellite sequences as primers in a PCR to generate multilocus markers. It is a simple and quick method that combines most of the advantages of SSRs and AFLP to the universality of RAPD.

Ten primers for ISSR were successful in generating reproducible and reliable amplicons for the four the imported cassava genotypes identified (Zayed *et al.*, 2012). It can be used to measure genetic distance and to generate molecular profile among selected accessions.

Simple sequence repeats (SSR)

SSR markers are stretches of DNA, consisting of tandomly repeating small nucleotide units (Powell *et al.*, 1996). Conserved regions flanking the repeats are suitable for designing PCR primer pairs to be used for amplifying the intervening repeat loci. These were first referred to as microsatellites by Litt and Luty (1989) and later as SSRs by Jacob *et al.*, 1991.

The hyper variability of SSRs among related organisms makes an excellent markers for a wide range of applications, including genetic mapping, molecular tagging of genes, genotype identification, analysis of genetic diversity, phenotype mapping and marker assisted selection (MAS) (Tautz, 1989 and Powell *et al.*, 1996). SSRs demonstrate a high degree of transferability between species, as PCR primers designed to an SSR within one species frequently amplify a corresponding locus in related species, enabling comparative genetic and genomic analysis.

SSRs are believed to be involved in gene expression, regulation and function (Kashi *et al.*, 1997 and Gupta *et al.*, 1994) and there are numerous lines of evidence suggesting that SSRs in non-coding regions may also be of functional significance (Mortimer *et al.*, 2005. At present more than 1000 SSR markers are available in cassava which helps to have genetic tags for various phenotypes in cassava.

Novel technologies for SSRs have been limited to new approaches to increase the multiplex ratio of the SSRs, to increase throughput and decrease costs. One such technology is the Multiplex-Ready™ Marker technology (MRT), developed at the University of Adelaide. This reduces marker deployment costs for fluorescent-based SSR analysis, and increases genotyping throughput (Nikki *et al.*, 2009).

SSR diversity in Cassava

Fourteen microsatellites containing GA repeats were isolated and characterized in cassava. Here microsatellite heterozygosity (*h*) was estimated in 48 accessions using (^{32}P)-end-labeled primers and in more than 500 accessions using fluorescence based genotyping. Heterozygosity values ranged from 0.00 to 0.88 and the number of alleles detected varied from 1 to 15 (Chavarriaga-Aguirre *et al.*, 1998).

Olsen and Schaal (2001) investigated the evolutionary and geographical origins of cassava (220) and the population structure of cassava's wild relatives (33) using five SSR markers. Seventy three alleles were observed across all loci and populations. These data indicate the following on cassava's origin: (1) genetic variation in the crop is a subset of that found in the wild *M. esculenta* subspecies, suggesting that cassava is derived from wild relative.(2) Phenotypic analyses group cassava with wild populations from the southern border of the Amazon basin, indicating this region as the likely site of domestication. (3) *Manihot pruinosa*, while closely related to *M. esculenta*, is probably not a progenitor of the crop. Genetic differentiation among the wild populations is moderately high. This differentiation has probably arisen primarily through random genetic drift following recent population divergence.

SSR marker variation was assessed at 67 loci in 283 accessions of cassava landraces from Africa and the Neotropics. Average gene diversity was high in all countries. Although the highest was found in Brazilian and Colombian accessions, genetic diversity in Neotropical and African materials is comparable. Forces shaping differences in allele frequency at SSR loci and possibly counter balance successive founder effects involve probably spontaneous recombination as assessed by parent offspring relationships and farmer selections for adaptations (Fregene *et al.*, 2003).

Elias *et al.*, (2004) have used SSR markers to assess the genetic structure of traditional landraces of sweet and bitter cassava collected from five South American

sites and also a sample of 38 accessions from a world collection of cultivated cassava. For a total of 10 loci examined, it was found that 15 alleles were not represented in this sample. The geographical structure of genetic variability was weak, but the genetic differentiation between bitter and sweet landraces was significant, suggesting that each form had evolved separately after domestication. The results showed that traditional landraces form an important source of genetic diversity and merit more attention from managers of crop genetic resources.

Lokko *et al.,* (2006) conducted a study to determine the extent of genetic diversity among African cassava accessions resistant to the CMD using SSR markers. The accessions include a breeding stock (Clone 58308), 5 improved lines, 62 CMD resistant and 10 CMD susceptible landraces. Average gene diversity, *He*, was high in all cluster groups, with an average heterozygosity of 0.591 ± 0.061. Gene diversity among all accessions was 51.4% and gene diversity within cluster groups was 46.6%, while 4.8% was due to diversity between the different cluster groups.

SSR markers were used to study the effect of CMD on the genetic diversity in five agroecologies in Uganda with high and low incidence of CMD (Kizito *et al.*, 2006). Surprisingly, high gene diversity was detected. Most of the diversity was found within populations, while the diversity was very small among agroecological zones and the high and low CMD incidence areas. The high genetic diversity suggests a mechanism by which diversity is maintained by the active involvement of the Ugandan farmer in continuously testing and adopting new genotypes that will serve their diverse needs. However, in spite of the high genetic diversity, there is a loss of rare alleles in areas with high CMD incidence.

One hundred and sixty local cassava varieties identified as sweet or bitter cassava by traditional farmers from Atlantic forest and Amazon, Brazil were studied for diversity using SSR markers (Peroni *et al.*, 2007). In the Atlantic forest sample the average genetic diversity (HS =0.654) was higher for the sweet varieties than for

bitter ones (0.582). The genetic differentiation coefficient (RST), used to estimate the diversity among groups, was 0.057 ($P < 0.001$), indicating that the divergence between the two groups is low.

SSR allele diversity and association mapping were performed by Raghu *et al.*, (2007) by using 29 morphological traits and 15 SSR primers in 58 cassava accessions. In this SSR analysis, a total of 71 alleles were generated and out of which 63 alleles were found to be polymorphic and the level of polymorphism was 86.56%.

Moyib *et al.*, (2007) have conducted diversity study among thirty one improved cultivars and five Nigerian landraces of cassava at genomic DNA level with 16 SSR primers. The similarity coefficients generated between improved cultivars and Nigerian landraces ranged from 0.42 to 0.84, and 12 distinct DNA cluster groups were identified at 0.70 coefficients. The study also revealed combination of 5 SSR primers could readily distinguish the 36 cassava genotypes at 0.93 similarity coefficients. These five primers clustered the 36 cassavas into 16 groups at 0.70 similarity coefficient.

Thirteen microsatellite markers were used to determine the genetic diversity of cassava, grown in home gardens in two Chibchan Amerindian reserves in Costa Rica (Oscar *et al.*, 2008). The level of genetic diversity was compared with the commercial varieties cultivated in Costa Rica and found high levels of genetic diversity among the cassava plants. 12 of the 13 loci examined were polymorphic in each Amerindian reserve (P=92.3). In the commercial varieties only nine loci were polymorphic (P=69.2).

A set of 124 new unique polymorphic EST-SSRs was developed and characterized Adebola *et al.*, (2009) which extends the repertoire of SSR markers for cultivated cassava and its wild relatives. The markers show high Polymorphism information content (PIC) values and therefore will be useful for cultivar identification, taxonomic studies, and genetic mapping. The study further showed

that mining ESTs was highly efficient strategy for polymorphism detection within the cultivated cassava gene pool.

Siqueira *et al.*, (2010) used nine microsatellite loci to investigate genetic structure and diversity in 83 Brazilian cassava accessions, including several landraces. All nine loci were polymorphic, averaging 6.00 alleles per locus. Total genetic variability was high (0.668) and most of this genetic variability was concentrated within municipalities (0.577). Cluster and structure analyses divided accessions into two major clusters or populations (K=2). A significant genetic versus geographic correlation was found in the study.

Molecular differentiation and diversity of cassava taken from 162 locations across Puerto Rico were assessed with 33 microsatellite markers (Montero *et al.*, 2011). The results of this study show that genetic diversity of unknown cassava samples is slightly higher than in the current Puerto Rico cassava collection. So it concluded that traditional practices, especially inter cropping and incorporation of volunteer seedlings, have led to the contribution of recombinant genotypes to cultivated stocks.

Twelve released varieties of cassava and 24 Central Kerala collections were assessed at the genomic DNA level with 36 SSR primers for genetic diversity study (Lekha *et al.*, 2011). The minimum number of SSR primers that could readily be used for identification of the 36 cassava genotypes was also determined. For the genetic diversity study, the similarity coefficients generated between released varieties and central Kerala varieties ranged from 40 to 95% and two separate DNA cluster groups were formed at 0.60 coefficients using "numerical taxonomy" and "multivariate analysis system software package". The similarity index for released varieties ranged from 60 to 93% and in the case of central Kerala varieties it ranged from 70 to 98%.

Genetic diversity and relationships among cassava genotypes identified in the African gene pool. A total of seventeen landraces around Africa, three elite lines from International Institute of Tropical Agriculture (IITA) and four inter-specific hybrid lines were analyzed. Ninety six SSR primers produced 31% polymorphic and 5% monomorphic markers, totaling 82. The cassava genotypes analyzed were genetically closely related, with a coefficient of genetic similarity ranging from 0.43 to 0.86, exhibiting similar SSR amplification profiles. This may imply low genetic variability among the genotypes studied (Kabeya *et al.,* 2012*).*

Genetic variation at SSR markers loci was used to evaluate 21 Tanzanian farmer- preferred cassava landraces collected and maintained at Mikocheni Agricultural Research Institute laboratory in Tanzania. Two West African cassava genotypes and one Kenyan cassava were included in the clustering analysis. Average gene diversity among the Tanzanian cassava was high (0.71) with an average heterozygozity of 0.46. Total number of alleles across all loci was 127 with mean number of alleles per locus being 6.35 and SSR markers had a mean polymorphic information content of 0.67 (Elibariki *et al.,* 2013).

SSR diversity of cassava in South, East and Central Africa in relation to resistance to CBSD was identified (Pariyo *et al.*, 2013) using 13 SSR primer. They determined the frequency and distribution of alleles putatively associated with resistance to CBSD derived from the variety 'Namikonga'. In order to identify the genetic constitution of cassava accessions from different regions of Kenya, seven pairs of SSR primers previously developed from cassava were used to detect polymorphic 21 alleles in a sample of 69 accessions. The cluster analysis of similarity matrix obtained at 68% with SSR data showed that the 69 accessions were grouped into five marker based groups. This study proved that SSRs could be used to identify cassava accessions as well as in the assessment of level of genetic relatedness among accessions (John *et al.*, 2014).

Single nucleotide polymorphism (SNP)

SNP represents a single nucleotide difference between two individuals at a defined location. There are three different forms of SNPs: transitions (C/T or G/A), transversions (C/G, A/T, C/A, or T/G) or small insertions/deletions (Edwards *et al.*, 2007). SNPs are direct markers as the sequence information provides the exact nature of the allelic variants.

SNPs are evolutionarily stable, not changing significantly from generation to generation. The low mutation rate of SNPs makes them excellent markers for studying complex genetic traits and as a tool for understanding genome evolution (Syvanen, 2001). The high density of SNPs makes them valuable for genome mapping, and in particular, allow the generation of ultrahigh- density genetic maps and haplotyping systems for genes or regions of interest and map-based positional cloning. SNPs are used routinely in crop breeding programs, for genetic diversity analysis, cultivar identification, phylogenetic analysis, characterization of genetic resources and association with agronomic traits (Edwards *et al.*, 2007). One of the main challenges of SNP was discovery of SNPs on complex genome due to repetitive sequences. But genome complexity reduction techniques along with New Generation Sequencing (NGS) technologies can successfully solve this problem.

SNPs on cassava

SNPs are known to occur at frequencies of one per, 100–500 bp in plant genomes, depending on the species, eg. 1 SNP/121 bp in cassava (Kawuki *et al.*, 2009), 1SNP/204 bp in maize (Gore *et al.*, 2009) and 1 SNP/500 bp in *Arabidopsis* (Atwell *et al.*, 2010). Rapid advancement in sequencing capability together with the reduction in sequencing cost allow for effective genome-wide discovery of SNPs in cassava.

Despite the continual reduction in sequencing costs driven by next-generation sequencing, whole-genome sequencing for multiple individuals remains costly for species with large genomes. One approach that has successfully been applied to

obtain reduced representations of the genome is targeted enrichment sequencing (Turner *et al.,* 2009). This technique, together with the RNA-seq based SNP discovery, allows the predefined regions encompassing EST-SNPs to be selectively captured and sequenced in multiple individuals (Gnirke *et al.,* 2009).

Other available reduced representation sequencing strategies, such as restriction site-associated DNA sequencing (RAD-seq) and genotyping-by-sequencing (GBS), involve the use of restriction enzymes to secure similar genomic fractions from different individuals to screen for polymorphic markers (Elshire *et al.,* 2011). The advantages of these restriction enzyme-based techniques are their ability to simultaneously discover and genotype genomic derived SNPs and their low cost of operation.

An early study of sequence variation in cassava identified 136 SNPs from EST sequences and 50 SNPs from bacterial artificial chromosome (BAC) end sequences (Lopez *et al.,* 2005). Kawuki *et al.,* (2009) studied sequence diversity in nine genes and identified 26 SNPs. An SNP which generates yellow color in cassava root was identified Welsch *et al.,* (2010). An allelic variant of one cassava Phytoene synthase (PSY) gene defines cassava tuber color.

Identification and validation of 1,190 SNP markers in cassava has been reported from a total of 2,954 putative EST-derived SNPs (Ferguson *et al.,* 2011) to genotype 53 cassava varieties from the Americas, Asia, West Africa and Africa. This study shows a similar structure in the germplasm as that observed by Fregene *et al.,* (2003) using SSR markers. Jorge *et al.,* (2012) evaluated 358 cassava genotypes by using 390 SNPs for different traits such as shoot weight, fresh root yield, starch fraction amylase content, dry matter content and starch yield.

An EST-derived SNP and SSR genetic linkage map approach (Rabbi *et al.,* 2012) made a great contribution to dissect the genetic architecture of cassava

resistance and other economically important traits. Most recently a high resolution genetic mapping of the CMD resistance locus was carried out using genotype by sequencing (GBS), here a total of 6756 SNP markers were mapped across 19 linkage groups, this is the densest map developed for cassava so far

A study in cassava to test the effect of restriction enzyme choice on library quality was done Martha and Rabby (2014). And their results suggested that 6-cutter enzymes might be most appropriate for genotyping a modest number of markers at a high multiplexing level. Most recently, a high density genetic map using 2,141 SNPs with integrated genetic and physical localization of newly annotated immunity related genes in cassava was carried out by Soto *et al., 2015*.

Table 1. Molecular markers used for Cassava diversity studies

Accessions used	References
Isozymes	
365 accessions, 109 wild, 10 isozymes	Lefervre and Charrier, (1993)
α- β- esterase isozyme	Ocampo *et al.*, (1993)
200 accessions, 8 isoenzymes	Carbal *et al.*, (2002)
28 accessions, 4 isozymes	Montarroyos *et al.*, (2003)
77 accessions, 9 isozymes, Costa Rica	Zaldivar *et a*l., (2004)
32 accessions - 4 isozymes	Efisue A. A (2013)
4 enzyme systems PAC, EST, MDH and PGI	Koffi *et al.*, (2015)
Randomly amplified polymorphic DNA	
31 accessions – Brazil	Colombo *et al.*, (1998)
Elite cassava germplasm, CIAT	Bonierbale *et al.*, (1993); Gomez *et al.*, (1996)
126 accessions - South America	Colombo *et al.*, (2000)
Brazil and CIAT collections	Carvalho and Schaal (2001)
30 cassava accessions - Latin America	Ferreira *et al.*, (2007)
50 accessions – Ghana	Asante and Offei (2003)
RFLP	
Cassava and wild species	Beeching *et al.*, (1998)
45 accessions and wild species	Fregene *et al.*, (2000)
AFLP	
Cassava wild species	Roa *et al.*, (1997)
521 accessions	Chavarriaga - Aguirre *et al.*, (1999)
8 accessions	Wong *et al.*, (1999)
69 accessions - cassava wild species	Elias *et al.*, (2000)
20 landraces 9 accessions	Fregene *et al.*, (2000)
African cassava landraces and elite cultivars, 20 AFLP	Raji *et al.*, (2009)
48 manihot species - 12 primer pairs	Whankara *et al.*, (2012)
ISSR	
Cassava 4 cultivar	Zayed *et al.*, (2013)
SSR	
Heterozygosity in 48 accessions	Chavarriaga-Aguirre *et al.*, (1998)
220 accessions, 33 wild, 5 SSR	Olsen and Schaal (2001)
283 accessions, 67 SSR markers	Fregene *et al.*, (2003)

Cont.

Table 1. Molecular markers used for Cassava diversity studies

Accessions used	References
SSR	
38 accessions, 10 SSR	Elias *et al.*, (2004)
138 CMD resistant accessions – Nigeria	Lokko *et al.*, 2005
5 improved lines, 62 CMD resistant and 10 CMD susceptible landraces.	Lokko *et al.*, (2006)
245 accessions, 35 SSR markers - Uganda	Kizito *et al.*, 2006
160 cassava accessions	Peroni *et al.*, 2007)
58 cassava accessions, 15 SSR primers	Raghu *et al.*, (2007)
36 genotypes, 16 SSR markers - Uganda	Moyib *et al.*, (2007)
Cassava accessions in Costa Rica, 13 SSR	Oscar *et al.*, 2008
Cassava cultivars and its wild relatives, 124 EST-SSRs	Adebola *et al.*, (2009)
18 cultivars,3 SSR primers	Bi *et al.*, (2010)
83 Brazilian cassava accessions, including several landraces	Siqueira *et al.*, (2010)
43 Ghana in farmer preferred accessions, 20 SSR primers	Asare *et al.*, (2011)
93 cassava accessions, 14 microsatelite primers	Ribeiro *et al.*, (2011)
Cassava from 162 locations across Puerto Rico, 33 SSR markers	Montero *et al.*,(2011)
12 released varieties of cassava and 24 Central Kerala collections, 36 SSR primers	Lekha *et al.*, (2011)
Cassava genotypes in African gene pool, 96 SSR marker	Kabeya *et al.*, (2012)
21 Tanzanian farmer preferred cassava landraces, 20 SSR markers	Elibariki *et al.*, (2013)
Diversity of cassava in South, East and Central Africa in relation to resistance to CBSD, 13 SSR primer	Pariyo *et al.*, (2013)
43 accessions, 20 SSR markers	Tetteh *et al.*,(2013)
69 accessions from Kenyan germplasm, 7 SSR primer	Ndung'u *et al.*, (2014)

Cont.

Table 1. Molecular markers used for Cassava diversity studies

Accessions used	References
SNP	
136 SNPs from EST sequences and 50 SNPs from BAC	Lopez *et al.*, (2005)
26 SNPs	Kawuki *et al.*, (2009)
SNP in yellow color in tuber	Welsch *et al.*, (2010)
2,954 putative EST-derived SNPs, 53 cassava varieties	Ferguson *et al.*, (2011)
358 cassava genotypes, 390 SNPs	Jorge *et al.*, (2012)
EST-derived SNP and SSR genetic linkage map	Rabbi *et al.*, (2012)
6756 SNP by GBS	Rabbi *et al.*, (2013)
Study using restriction enzyme choice	Hamblin *et al.*, (2014)
2,141 SNPs	Soto *et al.*, (2015)

Materials and methods

3. MATERIALS AND METHODS

The field experiments of the present study were carried out at Central Tuber Crops Research Institute (CTCRI), Sreekaryam Thiruvanandhapuram during 2014-2015 and the laboratory experiments were conducted at Division Crop improvement. The materials used and the methods adopted in this study are described below.

3.1 Field Experiment

Stem cuttings of selected cassava genotypes were planted on Central Tuber Crop Research Institute (CTCRI). It consists of 40 CMD resistant, 20 susceptible cassava cultivars. After two to three months of planting the cultivars were analysed for the CMD disease incidence by visual screening. Based on this screening, 25 resistant and 16 susceptible varieties were selected. Along with these cultivars 7 wild relatives of cassava from CTCRI, also added for the study (Table 2)

Table 2. Cassava accessions selected for the study

Sl. No	Cultivars	Code	Sl. No	Cultivars	Code
1	TME3	R1	25	CMR60	S8
2	TME4	R2	26	H165	S9
3	Sree Padmanabha (MNga-1)	R3	27	H266	S10
4	30572	R4	28	Sree Vishakam	S11
5	96/1089A	R5	29	Sree Sahya	S12
6	0160	R6	30	Sree Prakash	S13
7	0304	R7	31	Sree Harsha	S14
8	Albert	R8	32	Sree Jaya	S15
9	CMR100	R9	33	Sree Vijaya	S16
10	CMR401	R10	34	Sree Rekha	S17
11	CMR402	R11	35	Sree Praba	S18
12	CMR403	R12	36	Sree Athulya	S19
13	CMR404	R13	37	Sree Apoorva	S20
14	CMR405	R14	38	CO2	S21
15	CMR406	R15	39	CO3	S22
16	CMR407	R16	40	Ambakadan	S23
17	CMR408	R17	41	BR 105	S24
18	CMR409	S1	42	*M. glaziovii*	*W1*
19	CMR410	S2	43	*M. caerulescence*	W2
20	CMR411	S3	44	*M. pseudo glaziovii*	W3
21	CMR412	S4	45	*M. caerulescence* (BC3)	W4
22	CMR413	S5	46	*M. peruviana*	W5
23	CMR414	S6	47	*M. flabellifolia*	W6
24	CMR415	S7	48	*M. tristis*	W7

3.2 Genomic DNA extraction

DNA was isolated from leaf samples by using the method described by Dellaporta *et al.*, (1983) with slight modifications. Samples were collected from fresh young leaves of cassava samples weighing about 1g was grounded it in to fine powder in sterile pestle and mortar using liquid nitrogen and 2% of PVP. The powder was then transferred in to sterile eppendorf tubes. 15 ml of extraction buffer (Appendix I) was added in to these tubes, mixed well and incubated at 4°C for 30 min. After the incubation 1ml of SDS (20%) was added, mixed well and incubated in water bath (Memmert) at 65°C for 30 min. After incubation 5M potassium acetate (5ml) was added, mixed well and incubated on ice at 4°C for 20 min. These samples were then centrifuged (Sigma laborzentrifge) at 12000 rpm at 4°C for 20 min. After centrifugation the supernatant was collected and chilled isopropanol was added at 2/3 volume of it. The tubes were inverted gently to precipitate the DNA and it was incubated at -20°C for 30 min. The precipitated DNA was again centrifuged at 12000 rpm for 15 min at 4°C. The supernatant was discarded and the pellets were resuspended in 500µl TE buffer (Appendix II). Then 5µl of Rnase (10mg/ml) was added and incubated at 37°C for 1 hr on water bath. Equal volume of chloroform: isoamylalcohol (24:1) was added in to the samples and mixed well. These tubes were centrifuged at 12000 rpm for 15 min at 4°C. Again the supernatant was collected and 10µl sodium acetate was added along with 200µl ice cold absolute ethanol to precipitate DNA. It was mixed well and incubated for 2hr at -20°C. Tubes were taken and it was centrifuged at 1000 rpm for 15 min at 4°C. Supernatent was discarded and to the DNA pellet 500µl 70% Ethanol was added and again it was centrifuged. By discarding the supernatant he DNA pellet was dried completely and resuspended in 100µl TE and stored on 4°C / -20°C.

3.3 Determination of quality of DNA

Agarose gel electrophoresis

Gel was prepared by weighing 0.8g of agar powder (Sigma) in 100ml of 1X TBE buffer (Appendix III), agarose were dissolved by keeping it in microwave oven for 2 min. After cooling to 50- 60°C ethidium bromide (0.5µg/ml) was added and mixed well. Pour the solution in to gel casting tray were the comb was placed. When the gel get solidified comb was removed and it is transferred it to the gel tank which was filled with 1X TBE running buffer. The samples (1µl DNA+2 µl gel loading dye) as well as 100bp ladder (Appendix IV) (2µl) was loaded into the well. The gels were run at 80 volts for 30 min. Bands were visualized and it was documented in gel documentation system (Alpha Imager, USA).

3.4 Quantification of DNA

UV spectrophotometry

The concentration and purity DNA samples were determined by taking absorbance at 260/280 nm. The spectrophotometer was calibrated (Systronics) at 260nm as well as 280nm with 1 ml TE buffer. Absorbance of each samples were taken by mixing 10µl of DNA sample with 900µl TE buffer and TE buffer were kept as blank. OD260 and OD280 were measured and absorbance ratios were calculated. By using the formula:

$$\text{DNA concentration (µg/ml)} = \frac{\text{OD260} \times 100\ (\text{dilution factor}) \times 50\ \text{µg/ml}}{1000}$$

According to the reading obtained from the quantification, genomic DNA were diluted to a concentration of 50ng/µl and stored at 4°C and the stock DNA was stored in -20°C.

3.5 PCR Amplification for SSR primers

A total of 14 SSR primers (Table 3) were used for molecular characterization of samples. The PCR amplifications were performed in a volume of 20µl reaction containing 5ng genomic DNA, 0.2µM of each forward and reverse primer, .4µM of dntp, 1X buffer (10 mM Tris-Hcl (pH 8.3), 50 mM KCl, 1.5 mM MgCl), .03U Taq DNA polymerase and autoclaved nano pure water (ELGA). The cocktail for PCR was prepared according to the Appendix V. Amplifications were proceeded in a BioRad C1000^{TM} thermal Cycler which was programmed for initial denaturation of 5 min at 95°C then 30 cycles of, 1 min at 95°C, 1.30 min at 58°C, 2 min at 72°C, and a final extension of 5 min at 72°C. After amplification, a volume of 7µl of loading dye were added to each of the amplified product, and the products were run in 2% agarose gel which is mixed with ethidium bromide and it was visualized in a gel documentation system. The sizes of the amplified products were determined using an appropriate molecular ladder.

3.6 PCR Amplification for SNP primers

The PCR amplifications for SNP primers (Table 3) were performed in a volume of 20µl reaction containing 5ng genomic DNA, 0.2µM of each forward and reverse primer, 0.4µM of dntp, 1X buffers(10 mM Tris-Hcl (pH 8.3), 50 mM KCl, 1.5 mM MgCl), 0.03U Taq DNA polymerase and autoclaved ultra pure water. Amplifications (Bio-Rad : C1000^{TM}) programmed for initial denaturation of 2 min at 94°C then 30 cycles of 1 min at 94°C, 1 min at 53°C, 170sec at 72°C, and a final extension of 30 min at 72°C. The amplification of PCR products were analyzed by 2% agarose gel electrophoresis.

Table 3. SSR and SNP primers used for molecular characterization of cassava

Locus	Forward primer (5'-3') And Reverse primer	Types of repeat	Product size (bp)	Source
SSRY21	F:- CCTGCCACAATATTGAAATGG R:- CAACAATTGGCTAAGCAGCA	(GA)26	192	Mba *et al.*, (2001)
SSRY28	F:- TTGACATGAGTGATATTTTCTTGAG R:- GCTGCGTGCAAAACTAAAAT	(CT)26(AT)3AC(AT)2	180	Mba *et al.*, (2001)
SSRY32	F:- CAAATTTGCAACAATAGAGAACA R:- TCCACAAAGTCGTCCATTACA	(CA)11	292	Mba *et al.*, (2001)
SSRY36 *	F:- CAACTGTTTCAACCAACAGACA R:- ATTCTCGTGAACTGCTTGGC	(GT)3GC(GT)11(GA)19	134	Mba *et al.*, (2001)
SSRY39 *	F:- TCAATGCATAGGATTTTGAAAGTA R:- AATGAAATGTCAGCTCATGCT	(CT)24AT(CT)3(AT)3	293	Mba *et al.*, (2001)
SSRY40	F:- TGCATCATGGTCCACTCACT R:- CATTCTTTTCGGCATTCCAT	(GA)16	231	Mba *et al.*, (2001)
SSRY44	F:- GGTTCAAGCATTCACCTTGC R:- GACTATTTGTGATGAAGGCTTGC	(GA)28	194	Mba *et al.*, (2001)
SSRY106	F:- GGAAACTGCTTGCACAAAGA R:- CAGCAAGACCATCACCAGTTT	(CT)24	270-280	Mba *et al.*, (2001)
SSRY235	F:- CAGCTTTGCCATCCAATTTT R:- CAGCAAAATGACATGAGTGTATCTC	(GT)20	216	Fregene *et al.*, (1997)
NS158 *	F:- GTGCGAAATGGAAATCAATG R:- TGAAATAGTGATACATGCAAAAGGA	-	166	Akano *et al.*, (2002)
NS169 *	F:- GTGCGAAATGGAAAATCAATG R:- GCCTTCTCAGCATATGGAGC	-	319	Asare *et al.*, (2014)
NS198 *	F:- TGCAGCATATCAGGCATTTC R:- TGGAAGCATGCATCAAATGT	-	170-210	Asare *et al.*, (2014)
SSRY324	F:- CGCTTACAACACCACCTTCA R:- GCTTGATCTCAGCCATGTCA	-	206	Lokko *et al.*, (2004)
RME-1 *	F:- AGAAGAGGGTAGGAGTTATGT R:- ATGTTAATGTAATGAAAGAGC	-	650	Akano *et al.*, (2002)
SNP ERF *	F:- GAAAGCAATGGAAAATCTCTATCA R:- CAACCCAAGTCCAGTGTCTCA	-	520-530	Kawuki *et al.*, (2009)
SNP APX3 *	F:- GGTGCCACAAGGAACGCTCTG R:- CTTCCGACCATCATCATTCAAC	-	680-692	Kawuki *et al.*, (2009)

* - Fluorescent labelled primers

3.7 Capillary electrophoresis for SSR and SNP primers using Genetic analyzer

Among available SSR and SNP primer pairs, 5 SSR and 2 SNP were chosen based on their ability to give clear PCR products as visualised on an agarose gel. They were labelled to enable detection on the 3500 capillary DNA Genetic Analyzer (Applied Biosystem). Freshly prepared PCR products were diluted with 90 µl of sterile nanopure water (1:10). Above samples was vortexed for few seconds for fine mixing. Samples (0.5µl) were loaded on to the plates sequentially and 9.5 µl of master mix (9 µl of HIDI formamide and 0.5 µl Liz600 size standard for each sample) was then added . Loading pattern of each sample on the plate was noted. The plate was covered using sterile septa and allow for proper mixing using shaker. Then load the samples on the genetic analyzer and run for the fragment analysis.

3.8 Polyacrylamide gel electrophoresis (PAGE)

3.8.1 Larger plate cleaning

Glass plate was cleaned using deionized water to remove all the contaminants. It was then wiped with Kimwipe soaked in absolute ethanol. Air dried the plate and laboline was applied evenly using kim wipes.

3.8.2 Small plate cleaning

Glass plates were thoroughly cleaned with water and laboline and rinsed the glass plate with deionized water to remove detergent residues and wiped with kimwipes soaked in absolute ethanol. The glass plates were air dried and bind silane and were gently and evenly applied on the inner surface of small plate.

3.8.3 Gel preparation and casting

The larger glass plate was kept on a flat bench and the spacers were placed on its side's then smaller plate were placed on it (coated side should be towards spacer). Aligned the edges and the unit was assembled with side clamps and bottom caster assembly and the unit was locked. 6% of polyacrylamide solution containing 7M urea (Appendix VII) was prepared for gel casting. 60µl TEMED and 600µl APS was

added to it at the time of gel casting. The above mixed solution was taken into a barrel of 120ml of syringe and inverts the syringe to expel any trapped air that has entered the syringe. Introduce the nozzle of the syringe into the notched region on the caster base where both the glass plates are aligned. Expel the mixed solution from the syringe, filling the space almost to the top. Once the solution is filled, the comb is inserted in to the gel at the edge of the plate. Keep the apparatus in an appropriate position and the unit was left undisturbed for ½ an hour for polymerization.

3.8.4 Gel running

Assembled apparatus with casted gel was then dislodged from the precision caster base and fitted vertically into the universal base using a stabilizer bar. The temperature indicator was adhered to the surface of the outer plate to monitor the temperature during the run. The upper and the lower buffer chambers were filed with the required volume of 1X TBE buffer. The gel was prerun for 20 minutes at 100 W. Following completion of the prerun, the power supply was stopped. The wells were thoroughly rinsed using a pipette to remove any deposited urea. Samples were prepared with denatured PCR samples by denaturation of all PCR amplified DNA samples along with gel loading dye at 95°C for 5 min in a thermal cycler. 3-4 µl of each denatured samples were loaded along with 100bp ladder and each of the empty wells are loaded with empty well dye (Appendix IV). The samples were electrophoresed at 100 W for 35-40 minutes (specific according to each of the primers PCR product size). The power supply was turned off after the completion of the run and the upper buffer chamber was partially emptied by attaching the connector to the drain port on the gel unit. The unit was disassembled and the larger plate was removed and small plate containing the gel was used for silver staining.

3.8.5 Silver staining

The glass plate containing gel was transferred into a large tray containing the fixer (Appendix IX) and placed on a shaker (RiVOTEK) for 20 min, ensuring that the gel surface faced upwards. Similarly, staining was performed using silver stain (Appendix X) after the gel was washed in another trough containing deionized water for 5 min. Subsequent to a further wash with deionized water for few seconds, the stained gel was developed by transferring into a trough containing the developer (Appendix XI) and gently rocking the trough in a to and fro motion. A white surface was placed under the gel to enable visualization during development. After the bands had visibly developed, the plate was immediately transferred into the fixer for a few min for proper fixing of the bands. Following a final wash step, ensure that the wavy nature on the glass plate due to the fixer residues should wash out. Gel was allowed to dry on open air for an overnight. Clear and reproducible bands were only selected for scoring.

3.9 Data analysis

3.9.1 Data scoring for PAGE

The SSR allele containing PAGE gels with consistent banding pattern were used for scoring. The alleles were scored individually based on comparison with the molecular ladder. The size of the amplicons was compared using a 100 bp ladder (M/s Bangalore Genei, India). Each band generated by SSR primers was considered as an independent locus. Bands were scored visually for their presence (1) or absence (0) in a binary data form.

3.9.2 Data scoring for capillary electrophoresis

The amplified fragments were viewed, sized and binned using the Gene mapper v.3.7 software (Applied Biosystems). This software performs allele calls

which include peak detection and fragment size matching. The data from the size standard (LIZ) is used to determine a standard curve plotting mobility of the fragments against the known size. Fragments arising from the PCR products were compared with the standard curve and their sizes determined. Based on the presence (1) or absence (0) of peak on each samples at proper base pair position the binary data were generated.

3.9.3 Cluster analysis

The scoring data in the form of binary values was used for the construction of dendrogram. The genetic associations between accessions were evaluated by calculating the Jaccard's similarity coefficient for pair wise comparisons based on the proportions of shared bands produced by the primers. Similarity matrix was generated using the SIMQUAL programme of NTSYS-pc software, version 2.02 (Rohlf, 1998). The similarity coefficients were used for cluster analysis and dendrogram was constructed by the Unweighted Pair-Group method (UPGMA) (Sneath and Sokal, 1973).

3.9.4 Principal component analysis (PCA)

PCA was done to obtain both 2-D and 3-D images, in order to visualize the difference between the individuals. They were generated using the NTSYS-PC software 2.0 (Rohlf, 1997).

3.9.5 Polymorphism information content (PIC)

PIC was calculated for the marker, in order to characterize the capacity of each primer to detect polymorphic loci among the accession. It was calculated based on the formula $PIC=1\Sigma pi2$ where pi is frequency for the ith allele (Nei, 1973).

Results

4. RESULTS

The present study on "molecular characterization of CMD resistant, susceptible and wild relatives of cassava using SSR and SNP markers" was taken up to quantify the extent of genetic diversity associated with CMD resistance in cassava genotypes. The results of the experiments conducted are described as below.

4.1 Selection of plant material

Based on the field trials conducted at CTCRI for the evaluation of CMD resistance, among 60 cultivars planted, 17 highly resistant plants with no symptoms, 8 resistant with mild symptoms and 16 highly susceptible with severe symptoms were selected (Plate 1, 2 and Table 4). Along with that 7 wild relatives were also selected from the wild collections of cassava in CTCRI for marker analysis.

Plate 1. CMD Resistant plot

CMD Resistant with no symptom

Plate2. CMD suceptible plot

CMD symptoms on leaf

Table 4. Cassava accessions selected for the study and their CMD reaction

Cultivars	Code	CMD reaction	Cultivars	Code	CMD reaction
TME3	R1	R, NS	CMR60	S8	R, MS
TME4	R2	R, NS	H165	S9	R, MS
Sree Padmanabha (MNga-1)	R3	R, NS	H266	S10	R, MS
30572	R4	R, NS	Sree Vishakam	S11	R, MS
96/1089A	R5	R, NS	Sree Sahya	S12	R, MS
0160	R6	R, NS	Sree Prakash	S13	R, MS
0304	R7	R, NS	Sree Harsha	S14	R, MS
Albert	R8	R, NS	Sree Jaya	S15	R, MS
CMR100	R9	R, NS	Sree Vijaya	S16	R, MS
CMR401	R10	R, NS	Sree Rekha	S17	R, MS
CMR402	R11	R, NS	Sree Prabha	S18	R, MS
CMR403	R12	R, NS	Sree Athulya	S19	R, MS
CMR404	R13	R, NS	Sree Apoorva	S20	R, MS
CMR405	R14	R, NS	CO2	S21	R, MS
CMR406	R15	R, NS	CO3	S22	R, MS
CMR407	R16	R, NS	Ambakadan	S23	R, MS
CMR408	R17	R, NS	BR 105	S24	R, MS
CMR409	S1	R, MS	*M. glaziovii*	*W1*	-
CMR410	S2	R, MS	*M. caerulescence*	W2	-
CMR411	S3	R, MS	*M. pseudo glaziovii*	W3	-
CMR412	S4	R, MS	*M. caerulescence* (BC3)	W4	-
CMR413	S5	R, MS	*M. peruviana*	W5	-
CMR414	S6	R, MS	*M. flabellifolia*	W6	-
CMR415	S7	R, MS	*M. tristis*	W7	-

R – Resistant, NS - No symptom and MS – Mild symptom

4.2 Genetic diversity analysis at molecular level

In the present study, 48 cassava accessions were studied for their genetic diversity at molecular level using 14 SSR markers and 2SNP markers. Based on experiment field trials conducted in CTCRI, 25 CMD resistant, 16 susceptible, and 7 wild relatives of cassava were selected for the isolation of genomic DNA for the SSR and SNP analysis.

4.2.1 Isolation of genomic DNA

Genomic DNA was extracted from fresh young leaf of samples using Dellaporta method. The quantity and quality of DNA was tested by spectrophotometer and agarose gel electrophoresis respectively to ensure the use of good quality DNA.

4.2.2 Quantification of DNA

The concentration of DNA present in each sample was determined by reading the absorbance at 260nm and 280nm in spectrophotometer. The quantity of DNA in different samples varied from 1500-2500ng μL^{-1}. After quantification, all the samples were diluted to a concentration of 50ng $\mu L^{-1}$and the uniformity in concentration of the genotype was checked in agarose gel electrophoresis.

4.2.3 Testing the quality of DNA

Molecular marker profile analysis requires intact, unsheared DNA sample of sufficient quantity. Quality of genomic DNA samples is checked on 0.8% agarose gel with ethidium bromide dye along with 100bp reference marker. The gels were documented using a gel documentation system. The results, showed genomic DNA of all the studied accessions were intact with no protein and RNA contamination with an expected size range 100 to 300 bp (Plate3).

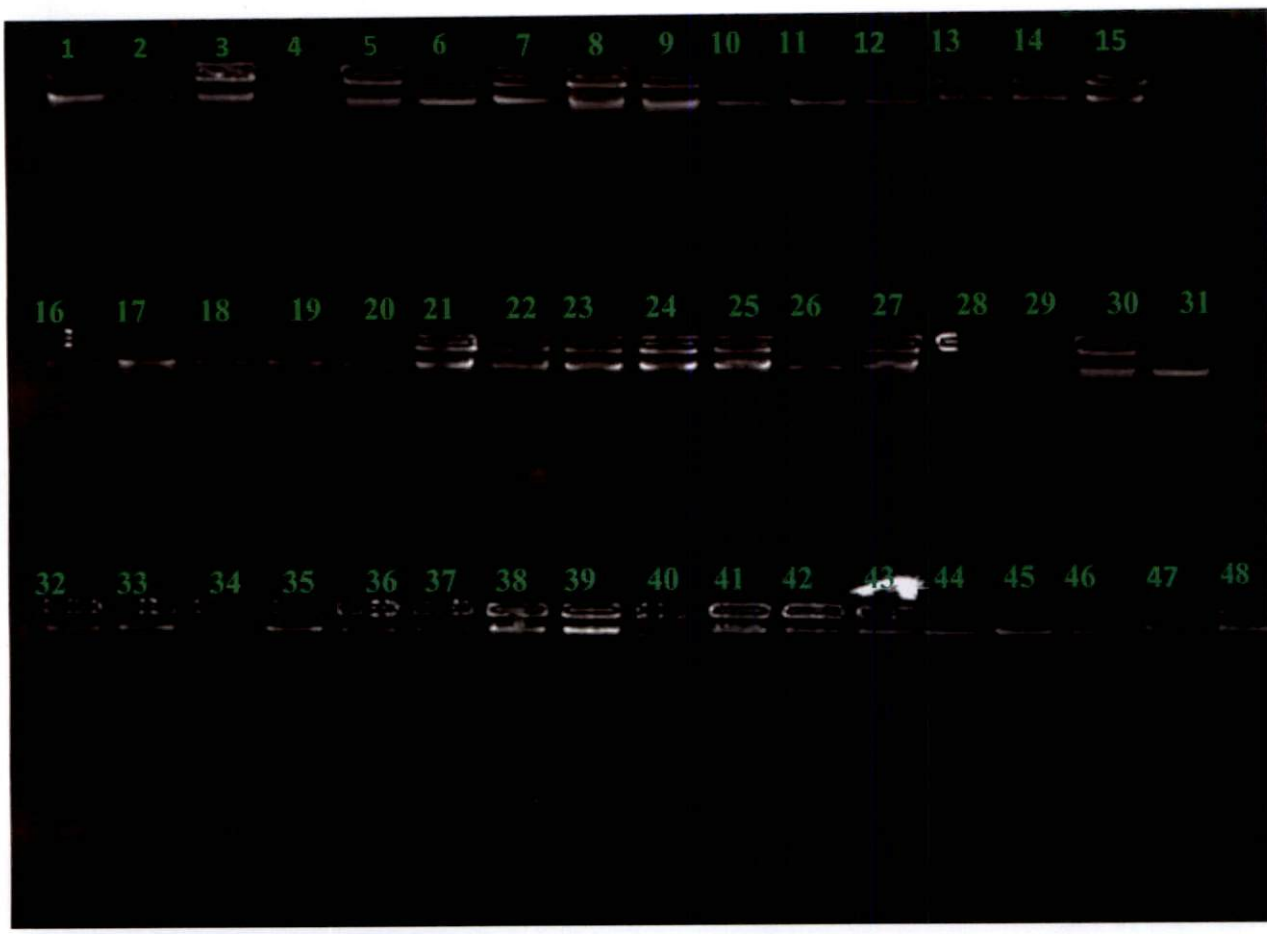


Plate 3. Checking the quality of isolated DNA by Agarose gel electrophoresis

4.2.4 Detection of PCR amplification on gel

PCR amplification of selected primers (most of them are associated with CMD in cassava) were performed under optimum conditions and the PCR products were run on 2% agarose gel along with the 100bp ladder. Among these, 14 SSR and 2SNP primers were produced good amplification of PCR product on gel (Plate 4 and Plate 5).

Plate 4. PCR product amplification of SSR primers

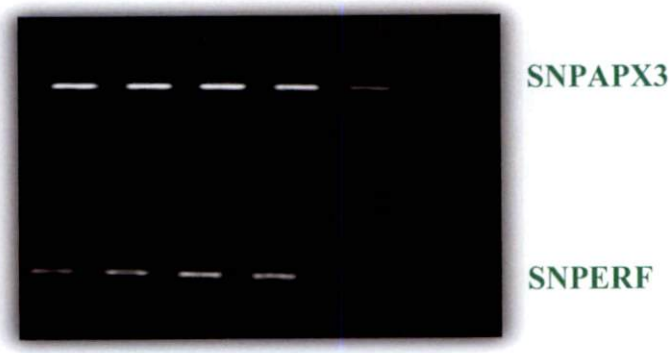


Plate 5. PCR amplification of SNP primers

4.2.5 Variability and polymorphism analysis

4.2.5.1 SSR and SNP analysis

All the SSR markers used were polymorphic and produced scorable, unambiguous results. The primers produced a total of 53 alleles across the 48 cassava accessions, the two SNP marker analysis on genetic analyzer namely SNPAPX3 and SNPERF for capillary electrophoresis revealed that, out of the two peaks generated one of the peak at a range of 650-690 and 500-530bp respectively was common to all the 48 accessions and the other peak was variable between accessions (Plate 6 and 7). Five SSR markers on genetic analyzer also produced high discriminative peaks on appropriate base pair locations (Plate 8, 9 and 10). The remaining nine SSR primer analysis on PAGE also generated clear and scorable bands (Plate 11,12,13 and 14). Among the primers used, NS198 was highly polymorphic in nature (6 allelles) followed by NS169, SSR36 and SSR39 (5 allelle). The details of banding patterns produced in the accessions are given in Table 5.The primer wise analysis of SSR and SNP markers has been described below

SSRY21: This primer pairs resulted in three amplified SSR loci, all the alleles were polymorphic and some unique bands were observed between the wild genotypes only. The SSR amplicons observed for this primer was at 180-192bp.

SSRY28: Produced three SSR loci at 180bp and all the loci were polymorphic with no unique bands in all the genotypes.

SSR32: Less polymorphic with only two SSR loci at 292bp.

SSR36: Highly polymorphic with five SSR loci at 120-134bp, CMD associated peaks were generated between most of the susceptible and resistant genotypes of cassava.

SSRY39: Highly polymorphic with five SSR loci at 290-310bp, some unique bands observed between *M. cearulescens* and its third backcross hybrid.

SSRY40: Three polymorphic SSR loci at 231bp, distinguished most of the highly resistant and susceptible genotypes.

SSR44: Three polymorphic SSR loci at 194bp, some unique bands observed between *M. pseudoglaziovii* and in some susceptible varieties.

SSRY106: Four polymorphic SSR loci at 272bp,

SSRY235: Three polymorphic SSR loci at 216bp, distinguished resistant and susceptible genotypes

NS158: Four polymorphic SSR loci at 166bp, unique bands observed in wild collections

NS169: Highly polymorphic with five SSR loci at 319bp, CMD associated peaks were generated between most of the susceptible and resistant genotypes of cassava, unique peaks observed for some wild collections.

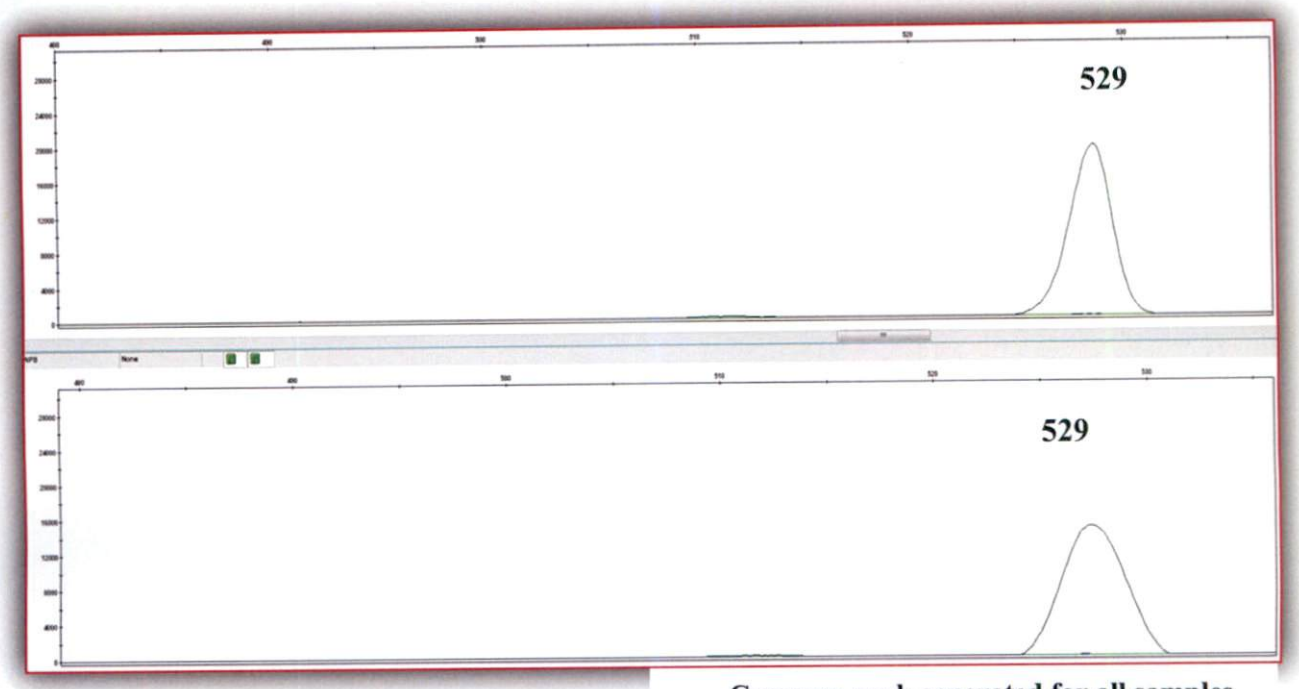
NS198: Most efficient, highly polymorphic marker with six SSR loci at 170-210bp in this study. Distinguished the resistant and susceptible cultivars by producing distinct peaks on capillary electrophoresis.

SSR324: Three polymorphic loci at 206bp, unique bands generated between wild and some susceptible samples.

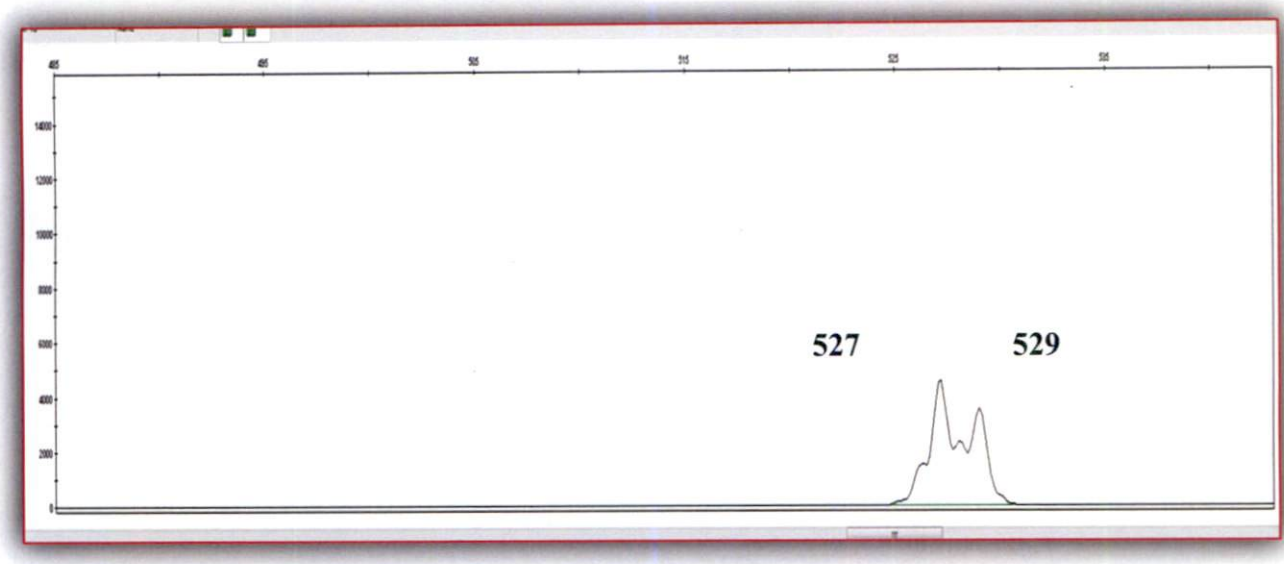
RME-1: Four polymorphic alleles were generated at 650bp, CMD associated peaks were generated between the susceptible and resistant cultivars of cassava

SNPERF: SNP marker generated one monomorphic peak in all 48 samples at 529bp and one highly polymorphic peak at 527bp only for few genotypes

SNPAPX3: SNP marker generated one monomorphic peak in all 48 samples at 688bp and one highly polymorphic peak at 692bp only for few genotypes.

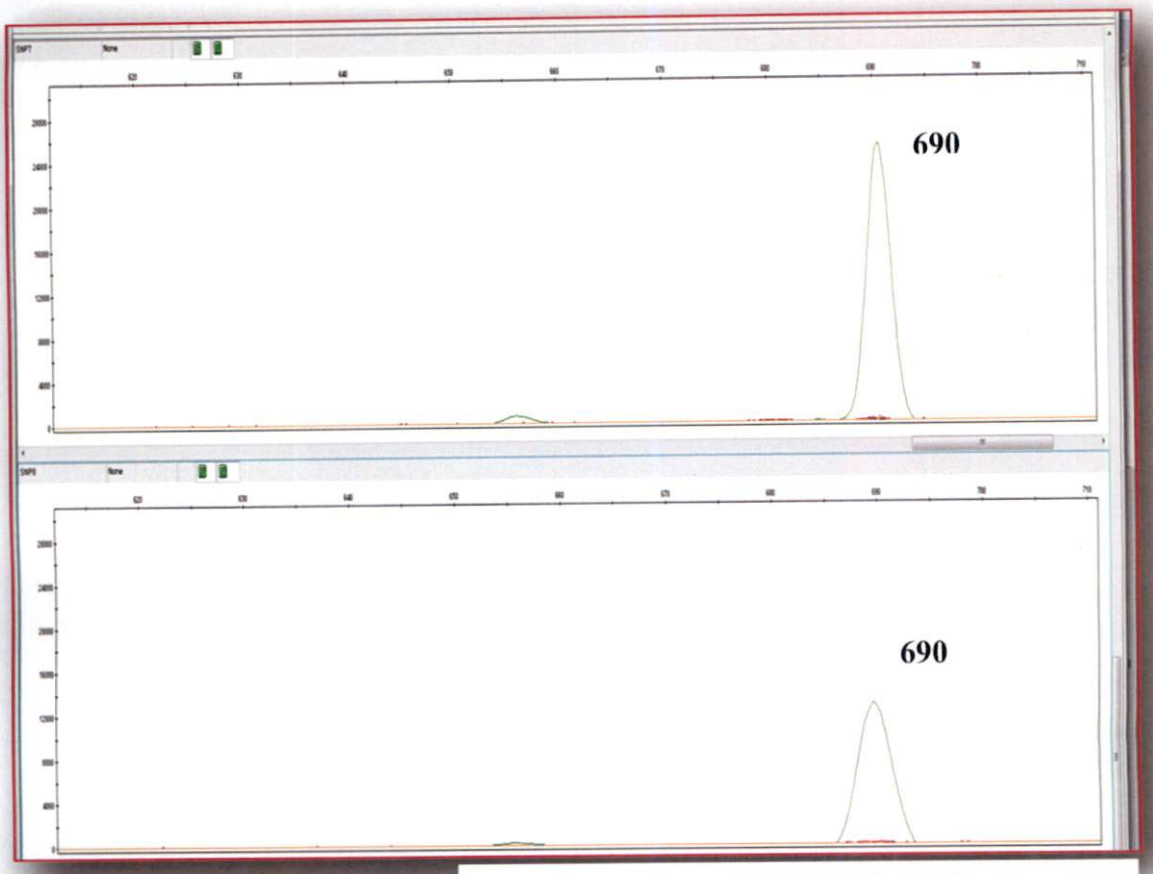


Common peak generated for all samples



Unique peak obtained for CMR 406

Plate 6. SNP ERF Analysis on genetic analyzer



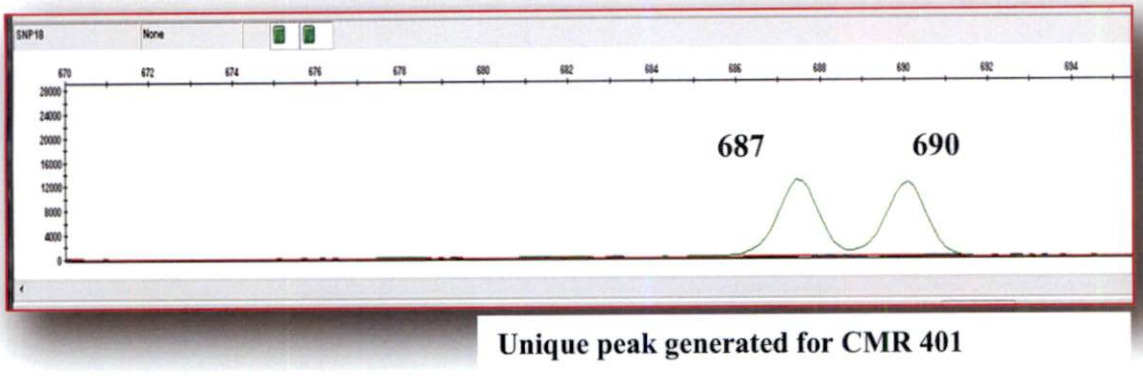


Plate7. SNPAPX3 Analysis on genetic analyzer

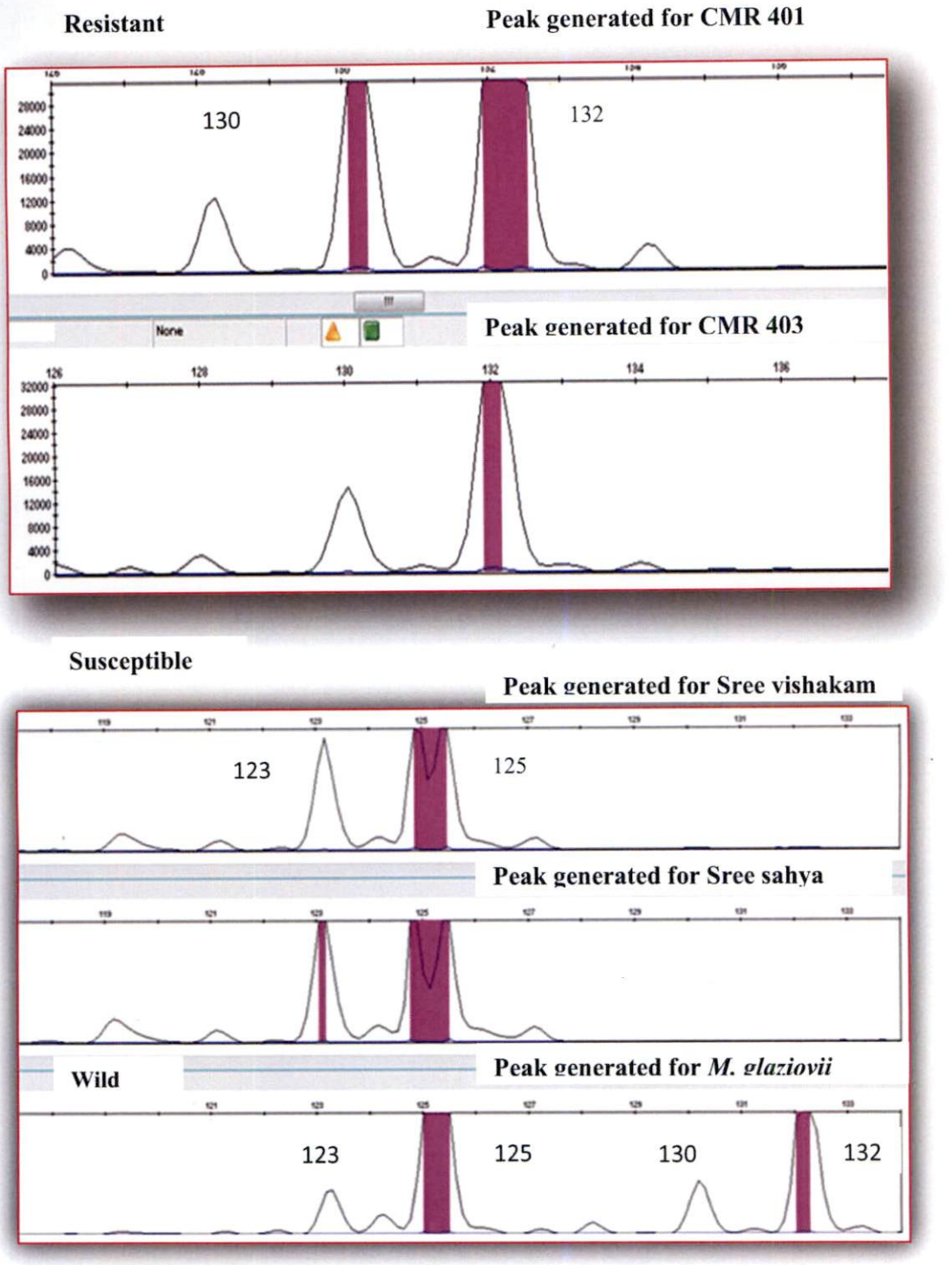


Plate 8. SSR36 Analysis on Genetic analyzer

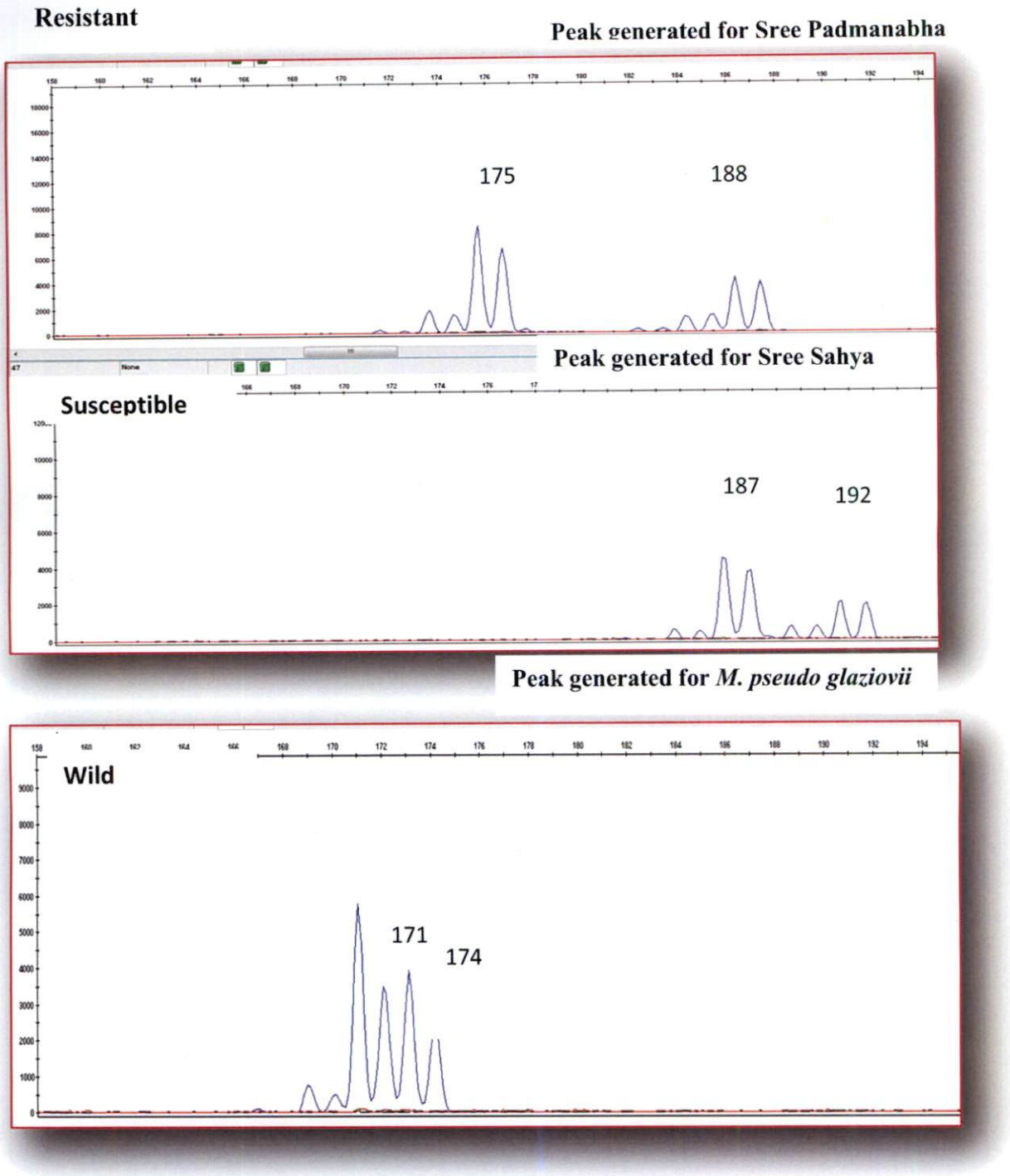
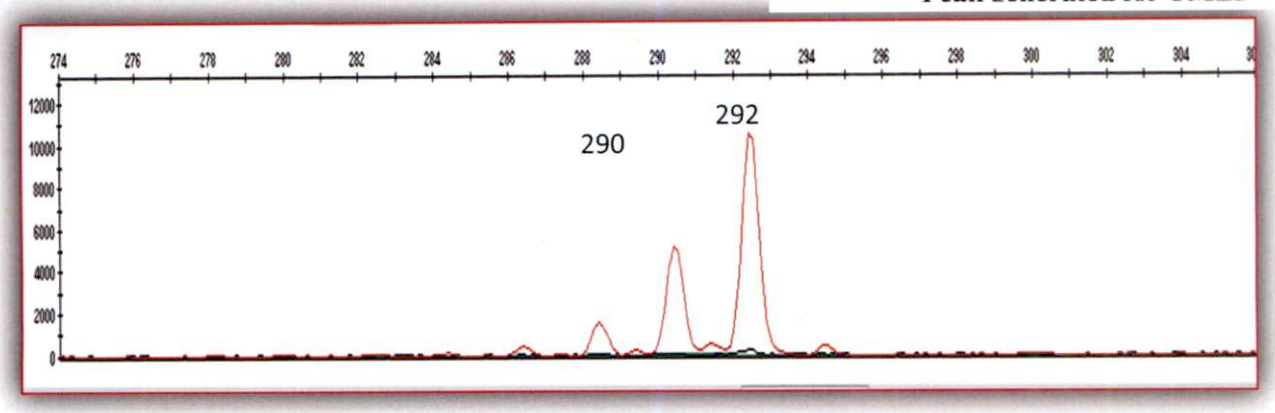


Plate 9. NS198 Analysis on Genetic analyzer

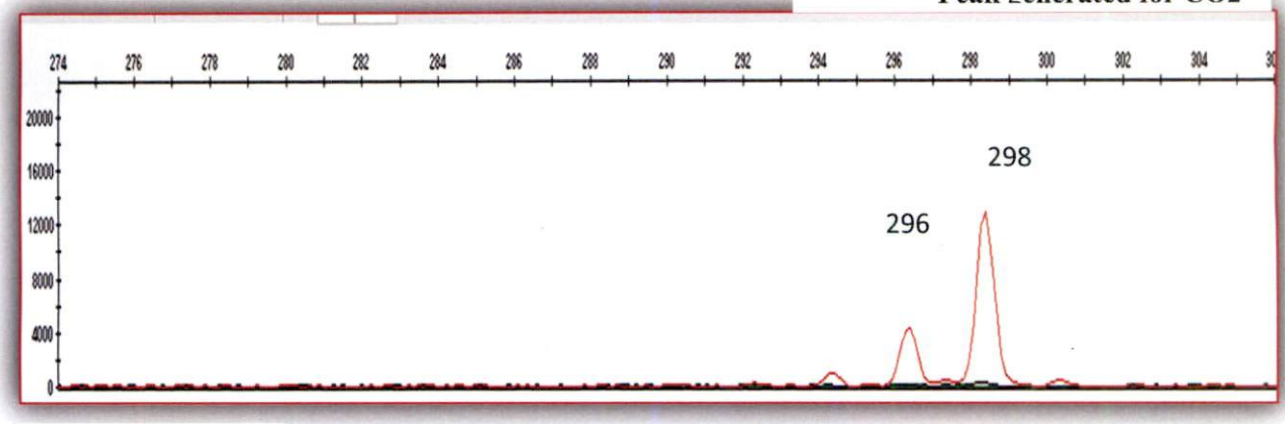
Resistant

Peak generated for TME3



Susceptible

Peak generated for CO2



Wild

Peak generated for *M. caerulescence*

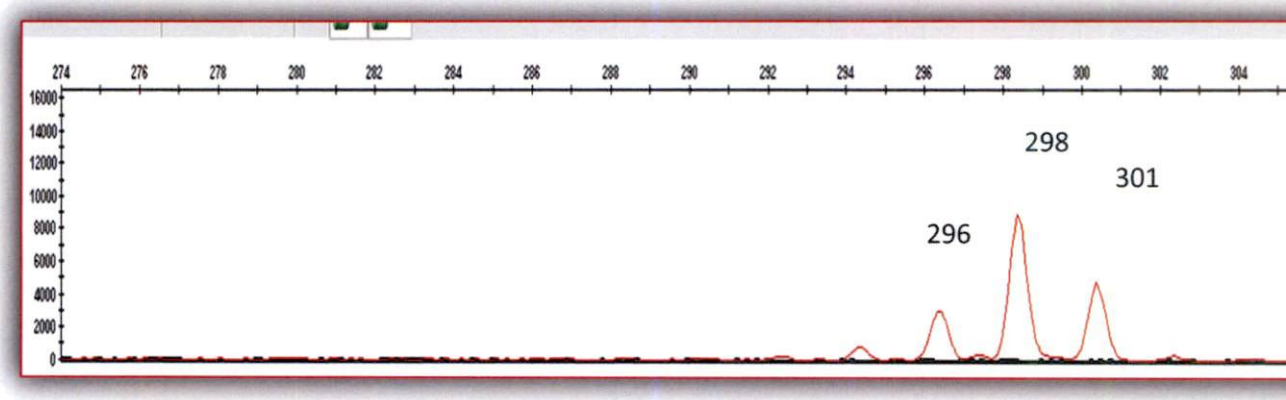


Plate 10. SSR39 analysis on genetic analyzer

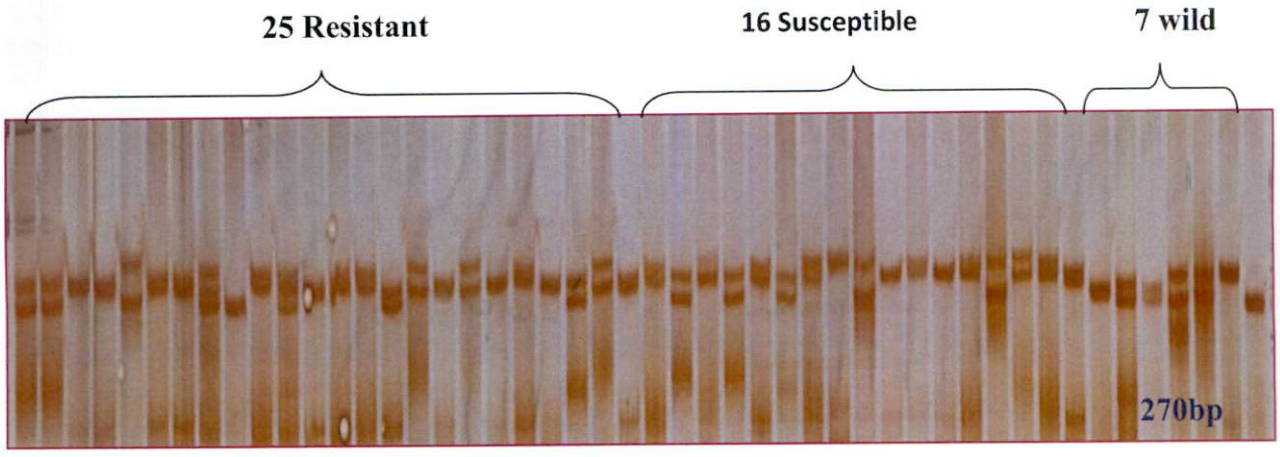


Plate 11. Segregation pattern of SSR106 on PAGE

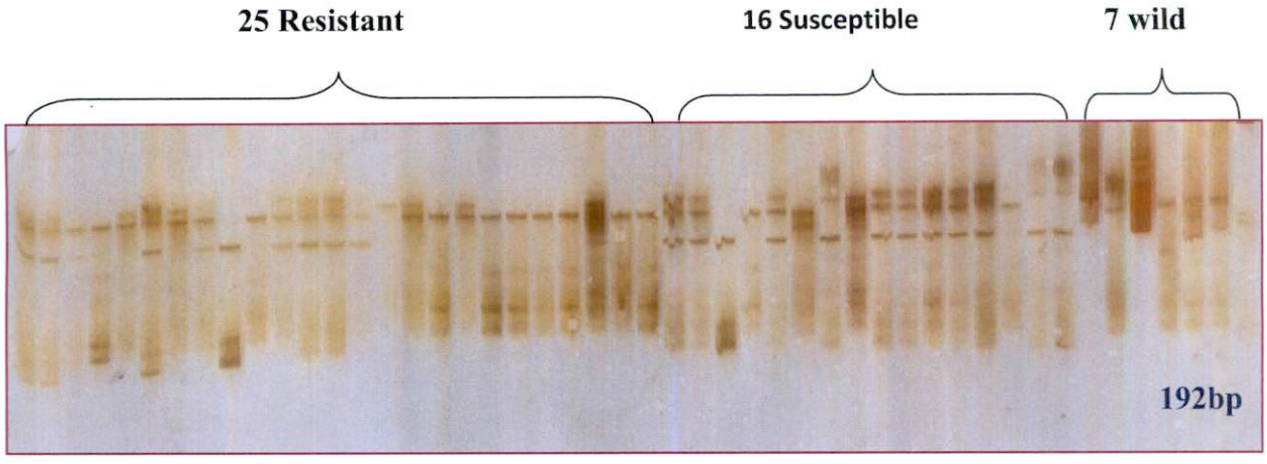


Plate 12. Segregation pattern of SSR21 on PAGE

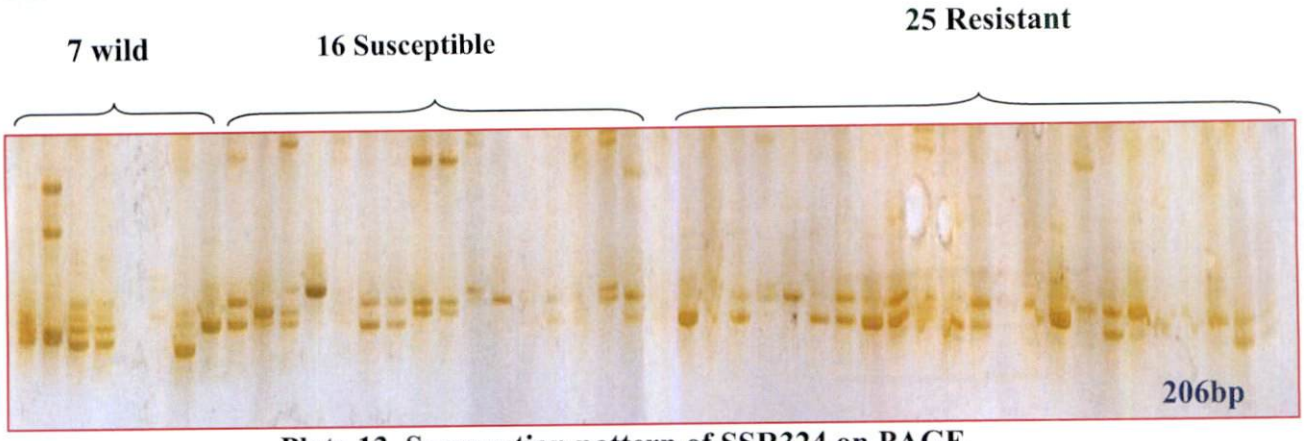


Plate 13. Segregation pattern of SSR324 on PAGE

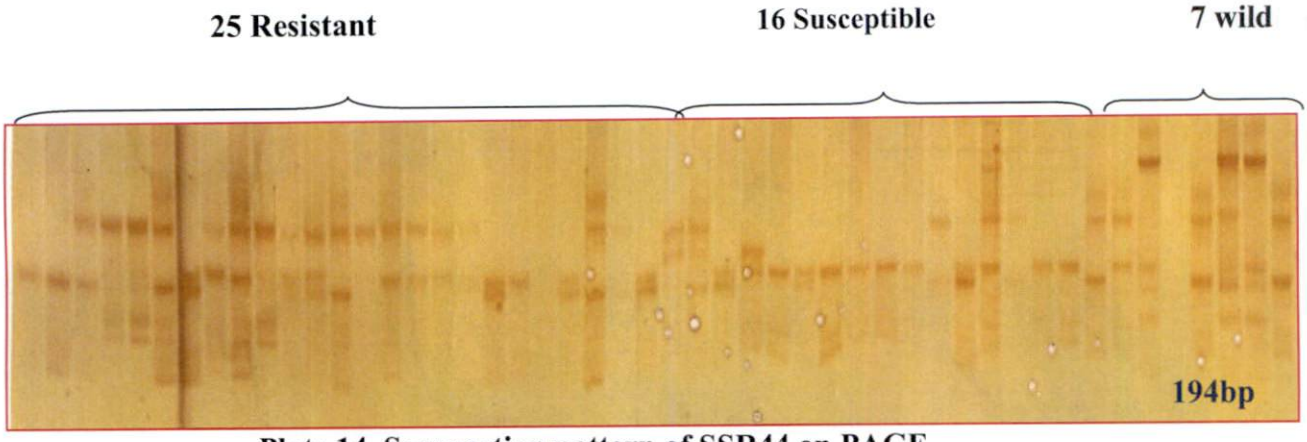


Plate 14. Segregation pattern of SSR44 on PAGE

Table 5. Percentage of polymorphism generated by the primers

Primer used	Number of alleles	Number of polymorphic bands	Percentage of polymorphism
SSRY21	3	3	100%
SSRY28	3	3	100%
SSRY32	2	2	100%
SSRY36	5	5	100%
SSRY39	5	5	100%
SSRY40	3	3	100%
SSRY44	3	3	100%
SSRY106	4	4	100%
SSRY235	3	3	100%
NS158	4	4	100%
NS169	5	5	100%
NS198	6	6	100%
SSRY324	3	3	100%
RME-1	4	4	100%

4.2.5.2 Cluster analysis

4.2.5.2 .1 Similarity index

The similarity index values obtained for each pair wise comparison among the 48 samples of cassava based on the combined SSR and SNP marker data is given in the Table 6. The similarity coefficient based on these markers ranged from 0.37-0.98. Among the 48 genotypes of cassava the lowest similarity index (0.37) was observed between wild genotype *M. cearulescens* and resistant cultivar 0304 and highest similarity index (0.98) was between two resistant cultivars, TME3 and TME4. Based on similarity index data, using Jaccard similarity coefficients a frequency graph was plotted in Fig.1. From the graph high similarity coefficients were between 0.65 and 0.75.

Fig.1 Frequency bar diagram for Jaccard similarity coefficient

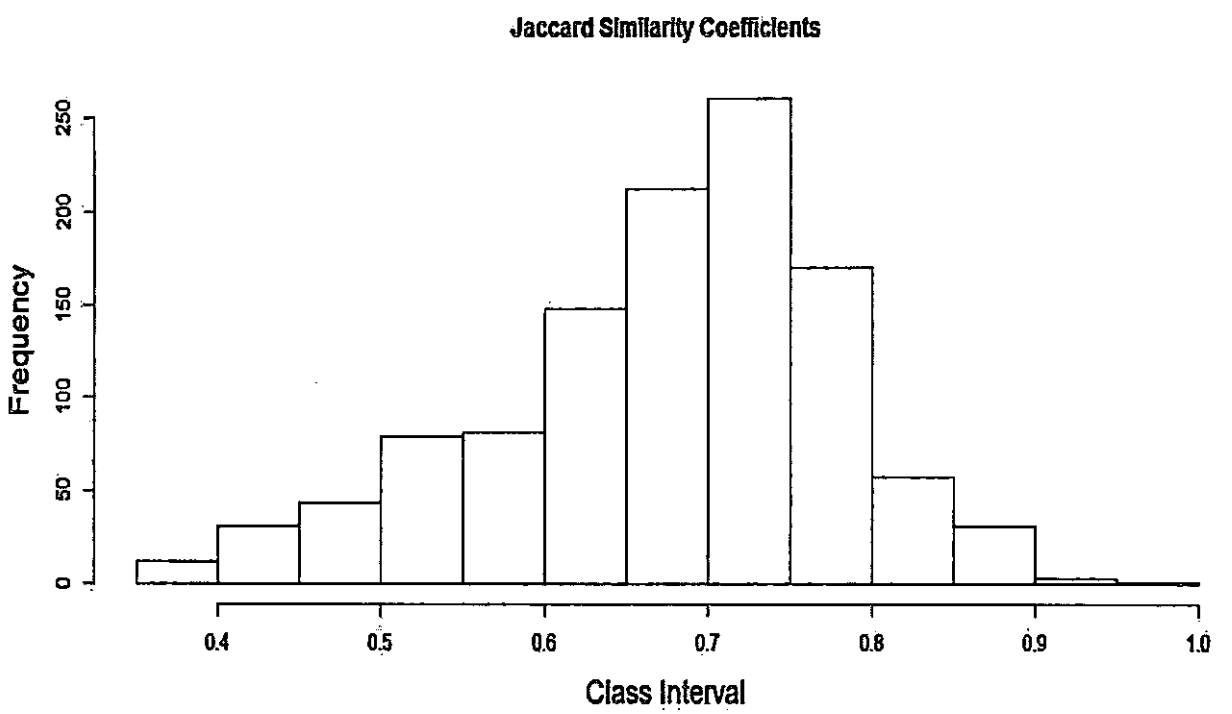
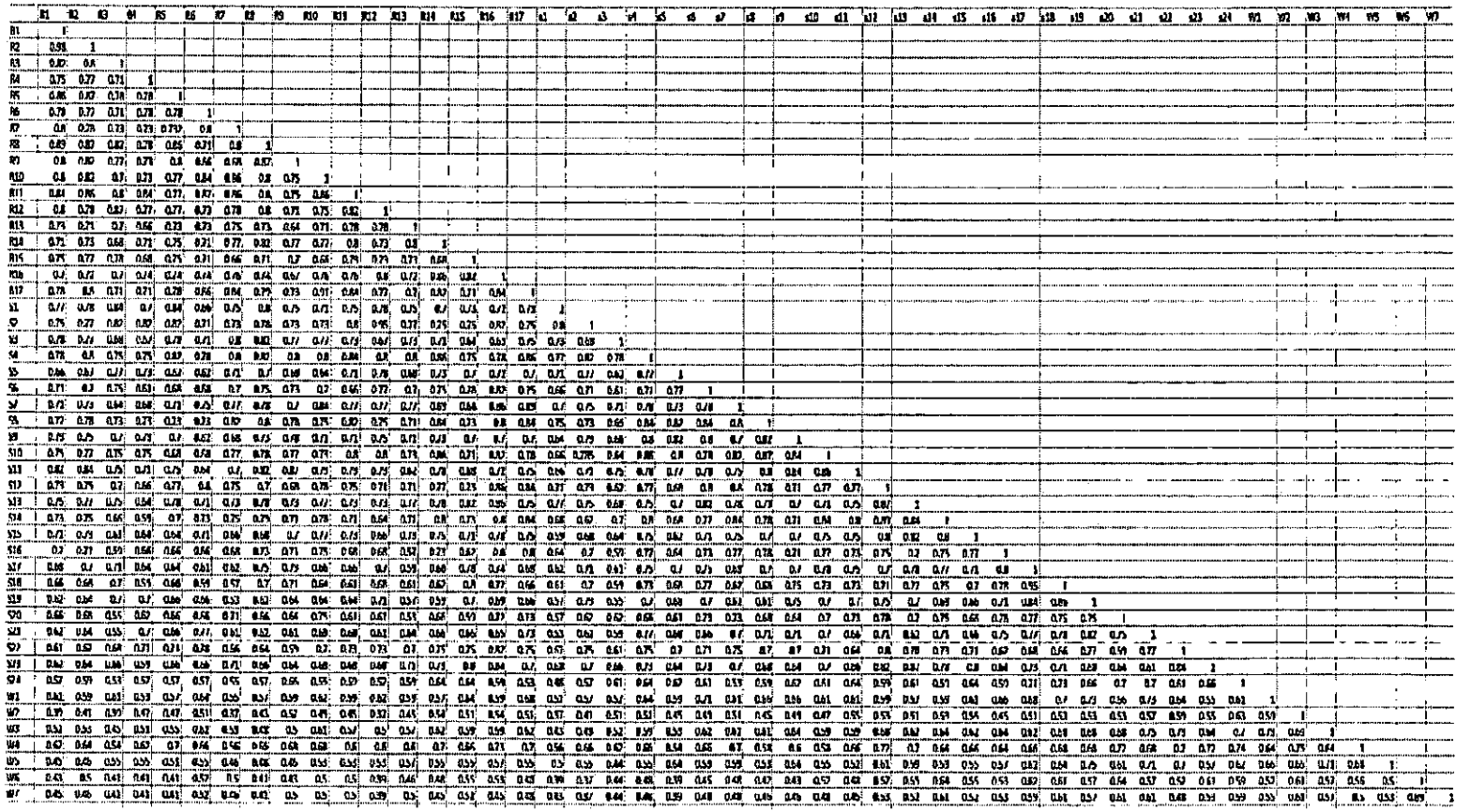


Table 6. Similarity index values for each pair wise comparison among 48 cassava accessions



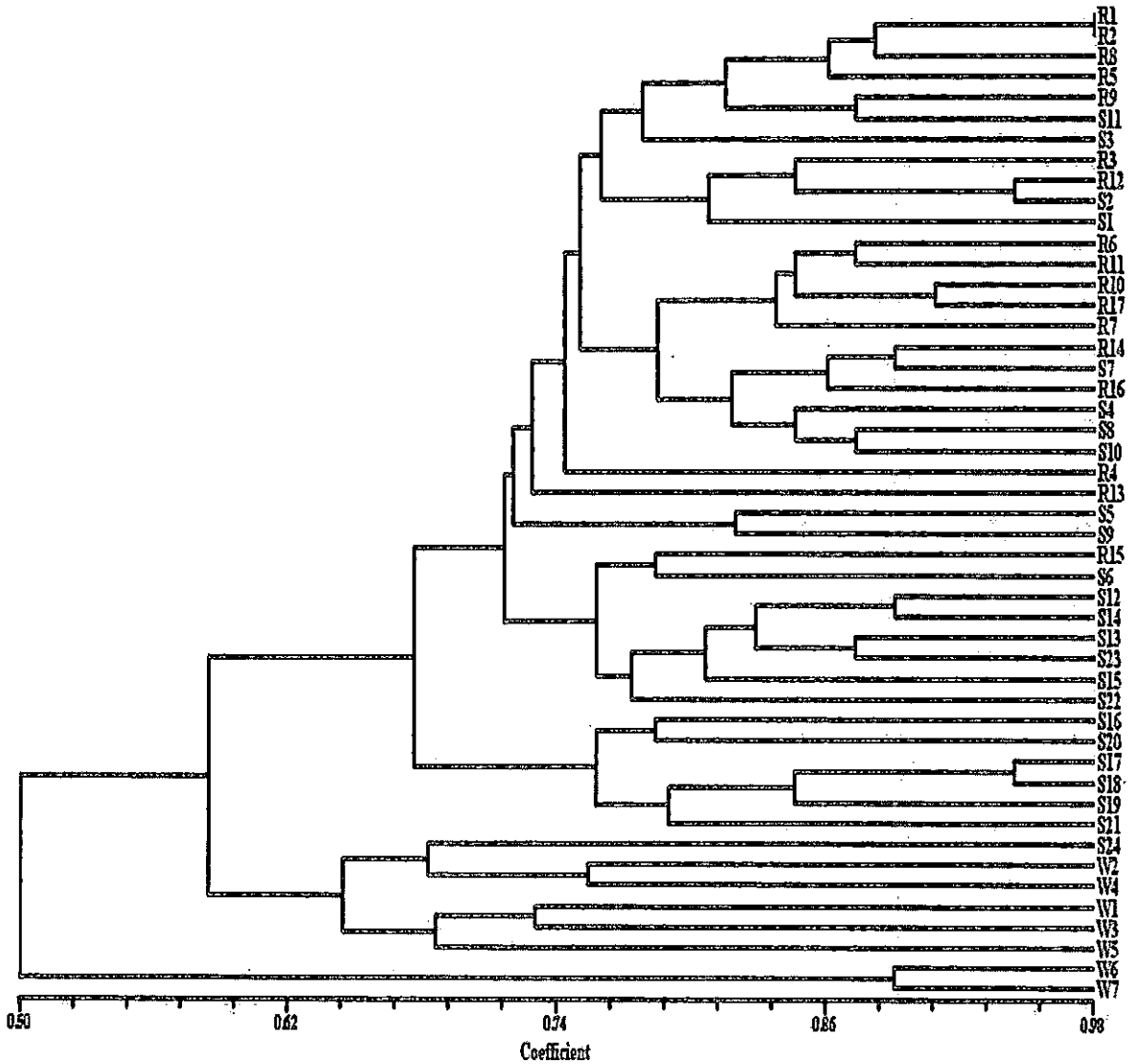
4.2.5.2.2 Cluster based dendrogram

UPGMA dendrogram was constructed using Jaccard's similarity coefficient based on SSR and SNP marker data. Dendrogram of combined SSR and SNP genotypic data (Fig. 2) resulted in four major clusters (Cluster 1, Cluster II, Cluster III and Cluster IV) at a similarity coefficient value of 0.668. All the seven wild collection of cassava except one susceptible variety BR105 was found to be under the first three major clusters. Cluster I included two wild species *viz.*, *M. flabellifolia* and *M. tristis*. Whereas cluster II and cluster III grouped remaining five wild accessions including *M. glaziovii* and *M. psuedoglaziovii* in one sub cluster. The remaining cluster IV included all the cassava cultivars taken for the study. Here the cluster IV is again divided in to two major sub clusters at a similarity coefficient of 0.677 viz., cluster IV A with most of the susceptible genotypes and cluster IV B with two sub cluster which arrange some of the susceptible genotypes and almost all of the resistant genotypes separately. Cluster IV A is further divided into sub cluster IV A1 and cluster IV A2 sub clusters at a similarity coefficient of 0.758, which grouped six susceptible varieties in to one group (Sree Vijaya, Sree Apoorva, Sree Rekha, Sree Prabha, Sree Athulya and CO2). Sub- sub clusters under cluster IVA1 and clusterA2 arrange these susceptible varieties in the order of the genotypes Sree Rekha and Sree Prabha closely together when compare to Sree Athulya and CO2, which were divided to another sub clusters and the genotypes Sree Vijaya and Sree Apoorva under the same sub cluster in the clusterA2.

The major cluster IVB with 34 cassava cultivars having two main sub clusters at a similarity coefficient of 0.719 viz., cluster IVB1 and cluster IVB2, which grouped almost all of the resistant variety except few and most of the remaining susceptible varieties from other clusters respectively in to distinct groups. The sub-sub clusters within these cluster IVB1 (seven sub cluster) and cluster IVB2 (two sub cluster) again grouped resistant and susceptible genotypes respectively

according to their similarity between the genotypes. One of the close observations on the sub clusters within the resistant cluster IVB1 revealed possible duplication between TME3 and TME4.

Fig. 2 Dendrogram based cluster analysis



4.2.5.2.3 Principle Coordinate Analysis (PCA) Analysis

2-D and 3-D principal co-ordinate analysis was performed (Fig.3 and Fig.4) for 48 genotypes of cassava using NTSYS-PC software 2.0 (Rohlf, 1997). Accordingly the resistant, susceptible and wild genotypes of cassava were categorized in to distinct groups in both 2-D and 3D-dimensions.

Fig.3 2-D Principal coordinate analysis plot for cassava resistant, susceptible and wild genotypes of cassava

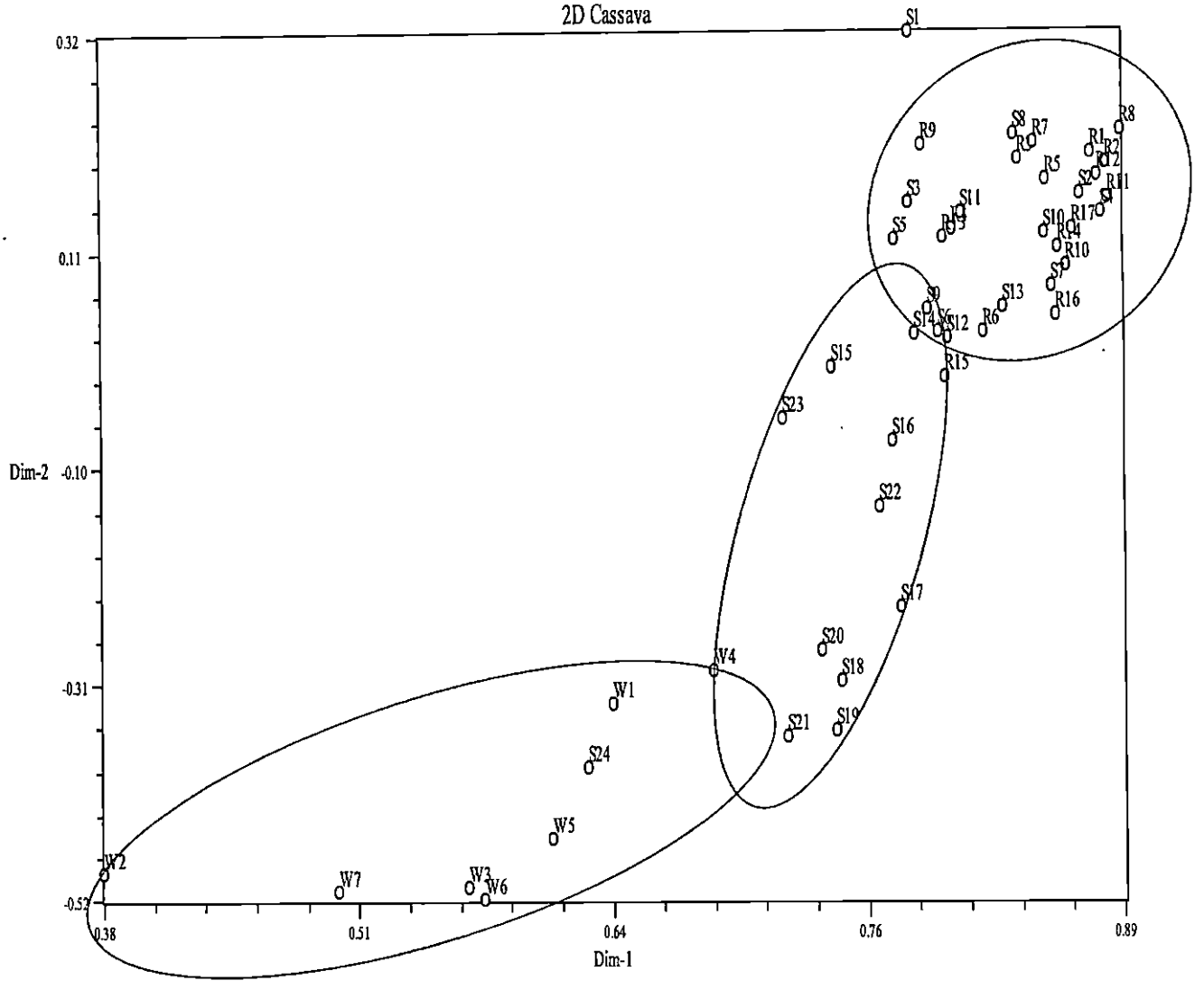
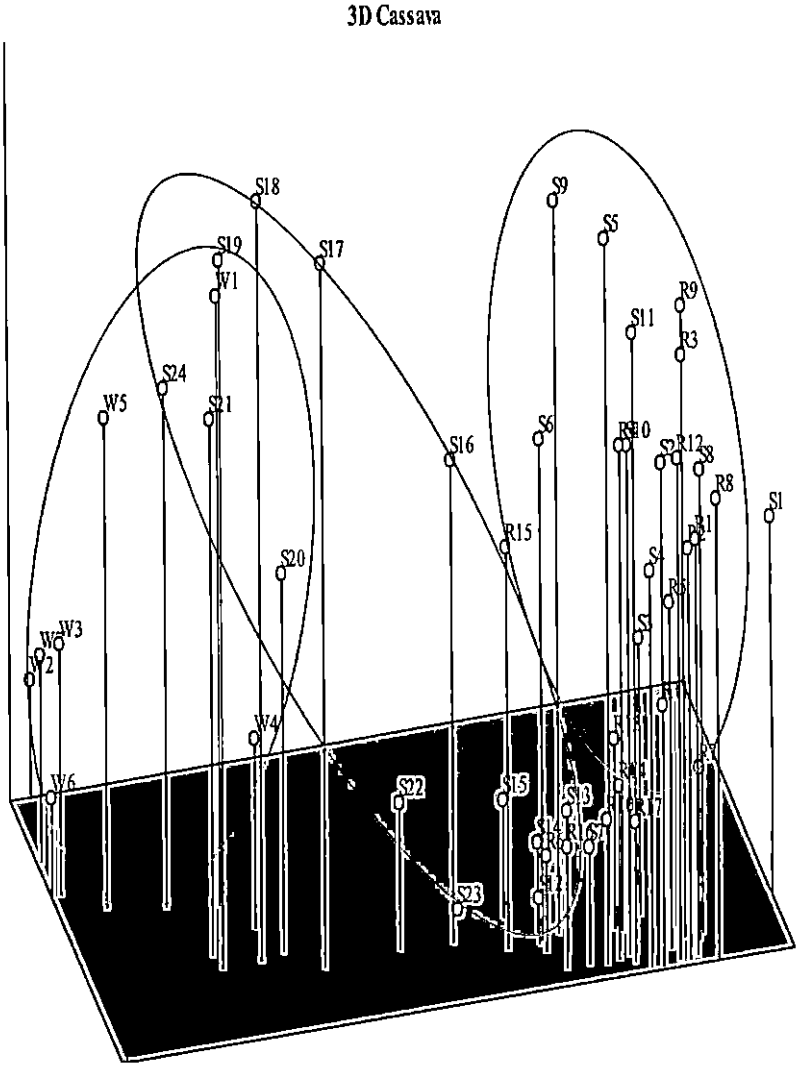


Fig.4. 3-D Principal coordinate analysis plot for cassava resistant, susceptible and wild genotypes of cassava



4.2.5.2.4 Polymorphic Information Content (PIC)

PIC values that revealed the ability of each primer to distinguish the 48 cassava genotypes were calculated for each of the 16 SSR primers assessed. Comparison of PIC value for all the 16 primers including both SSR and SNP primers revealed that the PIC value for individual SSR markers is higher than the individual SNP PIC value at a range of 0.19 - 0.24 (Table 7). SSR PIC values ranged from 0.347 in SSR32 with observed heterozygosity (He) 0.447 to 0.72 in NS198 with He 0.76. Based on the PIC ranges NS198, NS169, SSRY39, RME-1, SSR106 and NS158 were selected as highly polymorphic markers. This high heterozygosity observed values were significant, which substantiate the heterozygote nature of most of the genotypes and the fact that cassava is largely cross-pollinated.

Table 7. Polymorphic information content and heterozygosity of primers

Primer used	He	PIC
SSRY21	0.4644	0.3833
SSRY28	0.6608	0.5947
SSRY32	0.4478	0.3475
SSRY36	0.6068	0.551
SSRY39	0.7052	0.658
SSRY40	0.5662	0.50
SSRY44	0.5494	0.4472
SSRY106	0.6408	0.57710
SSRY235	0.517	0.4211
NS158	0.6149	0.5374
NS198	0.7636	0.7249
NS169	0.7413	0.699
SSRY324	0.5425	0.4373
RME-1	0.6443	0.5804
SNPAPX3	0.221	0.1975
SNP ERF	0.2854	0.2447

Discussion

5. DISCUSSION

Extensive use of closely related cultivars could result in vulnerability to pests and diseases. The use of diverse parental combinations in breeding may provide a greater supply of allelic variation that can be used to create new favorable gene combinations and increase in the levels of genetic variation, heterosis with marked increase in transgressive segregants in a population (Cowen and Frey 1987; Van Esbroeck and Bowman, 1998). Durable resistance to CMD that would be difficult for the pathogens to circumvent in the long term is an important breeding objective to ensure the sustainability of cassava production in Africa. Magoon and Krishnan (1977) showed that for maximum efficiency and reliability of breeding programmes in cassava, parents should be genetically diverse.

To ensure that durable resistance is maintained within the cassava germplasm there is a need to increase the levels of resistance within the gene pool using additional sources of resistance with a wider genetic base. Thus, this study was conducted to determine the extent of genetic diversity among CMD resistant cassava accessions and their relationship with susceptible accessions, including wild collections of cassava using SSRs and SNPs, for further extend the information to predict sources of resistance to CMD, for breeding of resistance to and management of CMD.

In this study the 48 genotypes, 25 resistant, 16 susceptible and 7 wild collections of cassava showed 53 alleles across the 48 cassava accessions with an average of 3.78 alleles per primer (ranging between 2 and 6 alleles) when analyzed with 14 SSR markers and 2 SNP marker. Out of the 16 primers, analyzed NS198 with 6 polymorphic allele was considered as one of the best SSR marker with high polymorphism, followed by NS169, SSR36 and SSR39 with 5 alleles. When analyzing 283 accessions of cassava from various countries with 67 SSR loci Fregene *et al.* (2003) found an average of 5.02 alleles per locus, ranging between 3 and 17 alleles per locus. Ragu *et al.*, (2007) assessed the molecular diversity by using

15 SSR primers, number of allele produced by different primers ranged from 2 to 6 with an average of 4 alleles per primer and level of polymorphism was found to be 50 to 100%. Among the primers used SSRY28, SSRY235 and SSRY324 produced a maximum of 6 alleles.

The internal branches of the dendrogram generated in this study were short while the external branches were long,indicating that within group variability was higher than that between groups. The resulted dendrogram clearly distinguished the resistant, susceptible and wild relatives of cassava into distinct clusters, revealed that most of the primers used in this study was highly associated with CMD resistance. The resistant cultivars with mild CMD symptoms used for the study also grouped under the resistant cluster. Epidemiological studies and field evaluation of the resistant genotypes revealed that they are not readily infected (Hahn *et al.,* 1980), and if infected, show mild symptoms, which may be restricted to some shoots (Jennings, 1960; Fargette *et al.,* 1996). Some of the resistant genotypes are characterized by transient or mild symptoms when infected by CMBs (Jennigs, 1994; Thresh *et al.*, 1994), while others develop conspicuous symptoms that are restricted to a few leaves or shoots (Thresh *et al.*, 1998). Virus concentration in resistant genotypes was reported to be low and a significant correlation was shown between symptom severity and CMD titer among resistant genotypes (Fargette *et al.*, 1996). However, the severity of symptoms expressed was not necessarily a reflection of virus concentration in some of the genotypes (Obge *et al.,* 2003).

The resistant/ tolerant genotypes often have two categories of plants: symptomless and symptomatic plants (Jennings, 1994; Fargette *et al,*. 1996). Among the latter, at least some branches of infected plants of resistant varieties are free from virus through 'reversion' under natural conditions (Storey and Nichols, 1938; Njock *et al.,* 1996). The extent of reversion depends on the genotype, and is regarded as a component of the resistance of cassava to CMBs (Fargette *et al.,* 1996; Thresh *et al.*,1998; Fondong *et al.,* 2000). Reversion has been exploited to select and produce

healthy cuttings for CMD epidemiological studies in Côte d'Ivoire (Fargette *et al.*, 1985, 1988). Recent studies showed that restriction of virus movement into axillary buds is an important aspect of resistance in CMD (Ogbe *et al.*, 2002). This probably explains reversion phenomenon in which infected stem of resistant genotypes could sprout into healthy plants in subsequent generation.

The similarity index values obtained for each pair wise comparison among the 48 genotypes of cassava based on the combined SSR and SNP marker showed highest similarity index (0.98) between two resistant cultivars, TME3 and TME4. This reflects their similarity in genetic base. Fregene *et al.* (2000) identified another duplication of genotype *viz.*, TME4 and TME6. Detection of duplicates in a collection is critical for effective management of germplasm.

Even though all the wild genotypes taken for this study were grouped in to one cluster, the sub cluster within that revealed their extent of diversity. Olsen *et al.*, 2001 also found out moderate-to-high level of population differentiation among the wild populations, through the analysis of genetic variation at five microsatellite loci to investigate the evolutionary and geographical origins of cassava (*M. esculenta* subsp. *esculenta*) and the population structure of cassava's wild relatives. Their estimates of variations among populations are greater for *M. esculenta* subsp. *flabellifolia* than for *M. pruinosa.*

The PCA scatter plot obtained here with a high discrimination between the selected resistant, susceptible and wild genotypes with spatial representations of genetic distances among them, revealed three major cluster groups along the second and third principal component axis, here the first group on one extreme was made up most of the resistant cultivars, the second group comprised the susceptible and third had all the wild collections near to the axis.

High heterozygosity values associated with the primers *viz.*, 0.7636, 0.7413, and 0.7052 for NS198, NS169 and SSRY39 respectively in this study revealed, that they are well suitable to estimate the extent of diversity between the genotypes.

Lokko *et al.*, (2006) also observed high average heterozygosity of 0.591 ± 0.061, when assessing the genetic diversity among 62 CMD resistant, 10 susceptible, clone 58308 and five improved lines in five cluster groups derived from UPGMA analysis on data from 18 SSR primer pairs.

The higher PIC values of SSRs, (0.7249) derived here was due to their multi-allelism, rapid mutation rate and low probability of being affected by the narrowing influence of selection, appeared to provide greater discrimination among resistant and susceptible cultivars, where as the individual SNPs had lower PIC values than SSRs resulting from their bi-allelism (0.1975- 0.2447). Kawuki *et al.*, 2009 also identified the same results of lower PIC value for individual SNP's than SSR (PIC = 0.228).

This may have been compounded by the fact that SNPs were closely associated within genes, although evidence was found for recombination within gene fragments. Genic regions have a higher probability of being affected by directional selection than non-genic regions. In addition, only a small proportion of the genome was represented in this assessment. The above factors will tend to decrease the discriminatory power of SNPs. It is likely that as a result of their biallelic as opposed to multi-allelic nature, a larger number of SNPs than SSRs will be required for the same level of discrimination. If the objective is to study intragenic variation, however, then SNPs may be more appropriate than SSRs due to their far higher frequency but the individual SNPs had lower PIC values than SSRs. For this reason a larger number of SNPs will be necessary to achieve the same level of discrimination provided by SSRs (Kawuki *et al.*, 2009).

Results of this study, thereby established a collection of these 5 highly polymorphic SSR primers, NS198, NS169, SSR36, SSR39 and RME-1 that could readily be used molecular characterization of CMD resistant, susceptible, and wild genotypes of cassava and the efficacy of SSR in genetic diversity analysis. Therefore

application of few number of highly polymorphic CMD associated SSR markers is possible for genetic variation studies in cassava to introduce additional sources of resistance associated with CMD resistance. This reduces the stress of applying many SSR primers for the identification associated with CMD resistance in cassava cultivars, hence saves time, ambiguous data, and also cuts the cost of research studies for genotype identification and genetic diversity studies.

Summary

6. SUMMARY

Cassava mosaic disease is one of the most globally damaging plant virus disease caused great yield loss in cassava. Host plant resistance is the most effective control strategy to tackle the disease. Due to the dynamic evolution of the virus, with the emergence of new recombinants, there is a need to identify new sources of resistance. High CMD resistance in cassava genotypes is often indistinguishable phenotypically for the different resistance genes making it unavoidably important to employ molecular tools in breeding for the disease. Recent studies on cassava revealed a better performance of SSR and SNP marker when compare to other molecular markers on genetic diversity studies due to its specific characters such as highly abundant, reproducible, co-dominant, and widely distribution throughout the genome of SSR and significant variation made to discriminate between closely related individual by SNP. With this view the present study was made for, molecular characterization of CMD resistant, susceptible and wild relatives of cassava using SSR and SNP markers.

Following are the summary of research work done in this study:

- To estimate the extent of genetic diversity between CMD resistant, susceptible and wild relatives of cassava a total of 48 cassava samples, 25 resistant(highly resistant, resistant with mild symptom of CMD),16 susceptible and 7 wild relatives of cassava were selected according to the observations made on the basis of field evaluation for CMD symptoms.
- The genomic DNA of selected samples was isolated using Dellaporta extraction method. Quantity and quality of the samples analysed using spectrophotometer and 0.8% agarose gel electrophoresis. The intact DNA with no protein and RNA contamination used for the PCR amplification of 14 SSR and 2 SNP primers. Out of the 16 primers 9 of them are analysed on PAGE and 7 primers on genetic analyzer for capillary electrophoresis.

- Based on the polymorphic bands/peaks obtained, the data were scored in the binary form. The polymorphism analysis of SSR and SNP generated a total of 53 alleles across the 48 cassava accessions, the NS198 is found to be highly polymorphic with 6 alleles followed by NS169, SSR36 and SSR39.
- The UPGMA derived dendrogram using Jaccard's similarity coefficient had four major clusters which clearly distinguished the resistant, susceptible and wild relatives of cassava in to distinct groups.
- The 2D and 3D principal coordinate analysis on the resistant, susceptible and wild relatives of cassava also made same level of discrimination of samples on the dendrogram.
- The similarity index values obtained for each pair wise comparison among the 48 samples of cassava based on the combined SSR and SNP marker showed, two resistant varieties TME3 and TME4 were closely related with a similarity index of 0.98.
- The PIC value calculated for each primer to check their ability to distinguish the samples revealed NS198, NS169, SSRY39, RME-1, SSR106 and NS158 are the most efficient polymorphic markers which clearly distinguished the resistant, susceptible and wild collections of cassava taken for the study.
- But the PIC value calculated for the SNP primers was very low which couldn't make any significant discrimination between genotypes. So larger number of SNPs will be necessary to achieve the same level of discrimination provided by SSRs as well as to make a significant discrimination between closely related genotypes.

References

7. REFERENCE

Akano, A., Dixon, A., Mba, C., Barrera, E. and Fregene, M. 2002. Genetic mapping of a dominant gene conferring resistance to *cassava mosaic disease. Theor. Appl. Genet.* pp. 105: 521–525.

Allen, A. C. 2002. The origin and taxonomy of cassava. In Hillocks, R. J., Thresh, J. M., Bellotti, A. C. (eds.) Cassava, biology, production and utilization. CABI, New York, pp. 1-15.

Appleby, N., Edwards, D. and Batley, J. 2009. New technologies for ultra-high throughput genotyping in plants. *Mol. Biol.* 513: 19-39.

Asare, P. A., Galyuon, I. K. A., Asare-Bediako, E., Sarfo, J. K. and Tetteh, J. P. 2014. Phenotypic and molecular screening of cassava (*Manihot esculentum Crantz*) genotypes for resistance to *cassava mosaic disease. J. Gen. Mol. Biol.* pp. 6 (2): 6-18.

Atwell, S., Huang, Y. S., Vilhjalmsson, B. J., Willems, G. and Horton, M. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465: 627-631.

Best, R. and Henry, G. 1992. Cassava: Towards the year 2000. In Report of the first meeting of the international network for cassava genetic resources. Roca, W. M. and Thro, A. M. (eds.), Cali, Colombia: *Centro Int. de Agric. Trop.* pp. 3-11.

Bonierbale, M. W., Maya, M. M., Claros, J. L. and Iglesias, C. 1995. Application of molecular markers to describing the genetic structure of cassava gene pools. In: *The Cassava Biotechnology Network: Proceedings of the Second International Scientific Meeting. Bogor, Indonesia, 22-26 August 1994. Cali, Colombia. Centro Int. de Agric. Trop.* 2v. (Working document No. 50).

Carvalho, L. J. C. B. and Schaal, B. A. 2001. Assessing genetic diversity in the cassava (*Manihot esculenta Crantz*) germplasm collection in Brazil using PCR-based markers. *Euphytica* 120: 133-142.

Ceballos, C. A., Iglesias, J. C. and Pérez, A. G. 2004. Dixon Cassava breeding: opportunities and challenges. *Plant Mol. Biol.* pp. 56: 503-516.

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. *Theo. Appl. Gen.* 97: 493-501.

Cock, J. H. 1985. Cassava: New potential for a neglected crop. London: Westfield.

Colombo, C., Second, G. and Charrier. 2000. Diversity within American cassava germplasm based on RAPD markers. *Genet. Mol. Biol.* 23 (1): 189-199.

Colombo, C., Second, G., Charrier, A. and Valle, T. L. 1998. Genetic diversity characterization of cassava cultivars (*Manihot esculenta Crantz*). RAPD markers. *Genet. Mol. Biol.* 21: 105-113.

Cowen, N. M. and Frey, K. J. 1987, Relationships between genealogical distance and breeding behavior in oat (*Avena sativa* L.). *Euphytica* 36: 413-424.

Dellaporta, S. L., Wood, J. and Hicks, J. R. 1983. A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* 1: 19-21.

Dixon, A. G. O., Whyte, J. B. A., Mahungu, N. M. and Ng, S. Y. C. 2001. Tackling the *cassava mosaic disease* (CMD) challenge in sub-saharan Africa: The role of host-plant resistance and germplasm deployment. In: Cassava, an ancient crop for modern times: Food, health, culture, N. J. Taylor, F. Ogbe and Fauquet, C. M. (eds.), Donald Danforth Plant sciences Center, St. Louis, USA. pp. S8-05.

Edmond, K. K., Boni, N., Guy, B. K., Modeste, K. K. and Koua, K. A. 2015. Enzymatic polymorphism of genetic diversity in cassava (*Manihot esculenta Crantz*) accessions in Côte d'Ivoire Greener. *J. Biochem. Biotech.* 2 (1): pp 009-017.

Edwards, D., Forster, J. W., Chagn, D. and Batley, J. 2007. What are SNPs? in *Association Mapping in Plants.* Oraguzie, N. C. , Rikkerink, E. H. A., Gardiner, S. E. and De Silva, H. N. (eds.), Springer , NY , 41-52 .

Elias, M., Panauda, O. and Robert, T. 2000. Assessment of genetic variability in a traditional cassava (*Manihot esculenta Crantz*) farming system, using AFLP markers *Heredity.* 85: 219-230.

Elias, M., Mühlen, G. S., McKey, D., Roa, C. A. and Tohme, J. 2004. Genetic diversity of traditional South American landraces of cassava (*Manihot esculenta Crantz*): An analysis using microsatellites. *Eco. Bot.* 58 (2): 242-56.

Elibariki, G., Njahira, M., Wanjala, B., Hosea, K. and Ndunguru, J. 2013. Genetic diversity and identification of duplicates in selected Tanzanian farmer preferred cassava landraces using simple sequence repeat (SSR) markers. *Int. J. Res. Plant Sci.* 3 (4): 81-87.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A. and Kawamoto, K. 2011. A Robust, Simple Genotypingby-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE 6: e19379.

FAO. 2000. www.fao.org.

FAO. 2013. http://faostat3.fao.org

FAO. 2014. www.fao.org.

Fargette, D., Fauquet, C. and Thouvenel, J. C. 1985. Field studies on the spread of African cassava mosaic. *Ann. Appl. Biol.* 106: 285-294.

Fargette, D., Colon, L. T., Bouveau, R. and Fauquet, C. 1996. Components of resistance of cassava to *African casava mosaic virus*. *Eur. J. Plant Pathol.* 102: 645-654.

Fargette, D., Fauquet, C. and Thouvenel, J.C. 1988. Yield losses induced by *African cassava mosaic virus* in relation to the mode and the date of infection. *Trop. Pest Manag.* 34: 89-91.

Ferguson, M., Hearne, S., Close, T., Wanamaker, S., Moskal, W., Town, C., de Young, J., Marri, P., Rabbi, I. and de Villiers, E. 2012. Identification, validation and high-throughput genotyping of transcribed gene SNPs in cassava. *Theor. Appl. Genet.* Mar 124 (4): 685-695.

Ferreira, C. F., Alves, E., Pestana, K. N., Junghans, D. T., Kobayashi, A. K., Santos, V., de, J., Silva, R. P., Silva, P. H., Soares, E. and Fukuda, W. 2007. Molecular characterization of Cassava with yellow-orange roots for beta-carotene improvement. *J. Crop Breed. Appl. Biotechnol.* pp. 8 (1): 23-29.

Fondong, V. N., Thresh, J. M. and Fauquet, C. 2000. Field experiments in Cameroon on *cassava mosaic virus* disease and the reversion phenomenon in susceptible and resistant cassava cultivars. *Int. J. Pest Manag.* 4: 211-217.

Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavarriaga, P., Roca, W. M., Tohme, J. and Bonierbale, M. W. 1997. A molecular genetic map of cassava (*Manihot esculenta Crantz*). *Theor. Appl. Genet.* 95: 431-441.

Fregene, M., Bernal, A., Duque, M., Dixon, A. G. O. And Tohme, J. 2000. AFLP analysis of African cassava (*Manihot esculenta* Crantz) germplasm resistant to the *cassava mosaic disease* (CMD). *Theor. Appl. Genet.* 100: 678-685.

Fregene, M., Okogbenin, E., Mba, C., Angel, F., Suarez, M. C., Janneth, G., Chavarriaga, P., Roca, W. M., Bonierbale, M. W. and Tohme, J. 2001. Genome mapping in cassava improvement: challenges, achievements and opportunities. *Euphytica.* 120: 159-165.

Fregene, M. A., Suarez, M., Mkumbira, J., Kulembeka, H., Ndedya, E., Kulaya, A., Mitchel, S., Gullberg, U., Rosling, H., Dixon, A. G., Dean, R. and Kresovich, S. 2003. Simple sequence repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. *Theor. Appl. Genet.* 107: 1083-1093.

Gedil, M., Kumar, M. and Igwe, D. 2012. Isolation and characterization of resistant gene analogs in cassava, wild *Manihot* species, and castor bean (*Ricinus communis*). *Afr. J. Biotech.* 11 (85): 15111-15123.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P. and Le Proust, E. M. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* 27: 182-189.

Gomez, R., Angel, F., Bonierbale, M., Rodriguez, F., Tohme, J. and Roca, W. 1994. Selection of heterozygous parents and single dose markers for genetic mapping in cassava. In: Roca, W. and Thro, A. M. (eds.) Proc 2nd CBN meeting. CIAT, Cali, Colombia. Pp. 113-124.

Gore, M. A., Wright, M. H., Ersoz, E. S., Bouffard, P. and Szekeres, E. S. 2009. Large-Scale Discovery of Gene-Enriched SNPs. *Plant Gen.* 2: 121-133.

Grattapaglia, D. E., Nassar, N. M. A. and Dianese, J. 1986. Biossistemática de espécies brasileiras do Gênero *Manihot* baseada em padrões de proteína da semente. Cienc. Cult. 39: 294-300.

Gupta, M., Chyi, Y. S., Romero-Severson, J. and Owen, J. L. 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theor. Appl. Genet.* 89: 998-1006.

Hahn, S. K., Terry, E. R. and Leuschner, K. 1980. Breeding cassava for resistance to *cassava mosaic disease. Euphytica.* 29: 673-683.

Hamblin, M. T. and Rabbi, I. Y. 2014. The effects of restriction-enzyme choice on properties of genotyping-by-sequencing libraries: A study in cassava (*Manihot esculenta*). *Crop Sci.* 54: 1-6.

Henry, G. and Hershey, C. 2002. Cassava in South America and Caribbean. In Hillocks, R. J., Thresh, J. M. and Bellotti, A. C. (eds.), Cassava biology, production and utilization. CABI, UK, pp. 17-40.

Hahn, S. K. and Keyser, J. 1985. Cassava: a basic food of Africa. Outlook *Agric.* 4: 95-100.

Jacobson, A. and Hedrén, M. 2007. Phylogenetic relationships in Alisma (*Alismataceae)* based on RAPDs, and sequence data from ITS and trnL. *Plant. Syst. Evol.* 265: 27-44.

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

Jansson, A., Westerbergh, J., Zhang, X. and Hu, C. 2009. Sun Cassava, a potential biofuel crop in (the) People's Republic of China. *Appl. Energy.* pp. 86: S95–S99.

Jennings, D. L. 1976. Breeding for resistance to *African cassava mosaic* disease: progress and prospects. In: Nestel, B. L. (eds.) *African cassava mosaic*: report of an interdisciplinary workshop. Muguga, Kenya. Report no. IDRC 0171e. IDRC (International Development Research Centre), Ottawa, Canada. Jones, W.O. 1959. Manioc in Africa. California. Stanford University Press, USA, p. 39-44.

Jennings, D. L. 1960. Observations on virus diseases of cassava in resistant and susceptible varieties. *Emp. J. Exp. Agric.* 28: 23-34.

Jennings, D. L. 1994. Breeding for resistance to *African cassava mosaic geminivirus* in East Africa. *Trop. Sci.* 34: 110-122.

Jones, W. O. 1959. Manioc in Africa. Stanford University Press, Starford, pp. 315.

Kabeya, M. J., Kabeya, U. C., Bekele, B. D. and Ingelbrecht, I. L. 2012. Genetic Analysis of Selected Cassava (*Manihot esculenta*) Genetic Pool in Africa Assessed with Simple Sequence Repeats. *World J. Agric. Sci.* 8 (6): 637-641.

Kashi, Y., King, D. and Soller, M. 1997. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.* 13: 74-78.

Kawuki, R., Ferguson, M., Labuschagne, M., Herselman, L. and Kim, D. J. 2009. Identification, characterisation and application of single nucleotide polymorphisms for diversity assessment in cassava (*Manihot esculenta Crantz*). *Mol. Breed.* 23: 669-684.

Kizito, E. B., Ronnberg-Wastljung, A. C., Egwang, T., Gullberg, U., Fregene, M., Westerbergh, A. 2007. Quantitative trait loci controlling cyanogenic glucoside and dry matter content in cassava (*Manihot esculenta Crantz*) roots. *Hereditas*. 144: 129-136.

Lancaster, P. A. and Brooks, J. E. 1983. Cassava leaves as human food. *Econ. Bot.* 37 (3): 331 -348.

Legg, J. P. and Fauquet, C. M. 2004. *Cassava mosaic gemini viruses* in *Afr. Plant Mol. Biol. 56: 585-599.*

Lefèvre, F. and Charrier, A. 1993. Isozyme diversity within African Manihot germplasm. *Eupliytica.* 66: 73-80.

Lekha, S., Jaime, A., Teixeira, da Silva. and Santha, V. P. 2011. Genetic variability studies between released varieties of cassava and central Kerala cassava collections using SSR markers. *J. Stored Products and Postharvest Res.* 2 (4): pp. 79-92.

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

Lokko, Y., Gedil, M. and Dixon, A. G. O. 2004. QTLs associated with resistance to the *cassava mosaic disease*: new directions for a diverse plant. In *Proceedings of the 4th International Crop Science Congress*. 26th September-1st October, 2004, Brisbane, Australia.Retrievedfromhttp://www.cropscience.org.au/icsc2004/poster/3/7/4/1154_1okko.htm.

Lokko, Y., Danquash, E. Y., Offei, S. K., Dixon, A. G. O. and Gedil, M. A. 2005. Molecular markers associated with a new source of resistance to the *cassava mosaic disease*. *Afr. J. Biotechnol.* 4 (9): 873-881.

Lokko, Y., Dixon, A. G. O., Offei, S. K., Danquah, E. and Fregene, M. 2006. Assessment of genetic diversity among African cassava *Manihot esculenta Crantz* accessions resistant to the *cassava mosaic virus* disease using SSR markers. *Gen. Resour. Crop Evol.* 53: 1441-1453.

Lopez, C., Pie'gu, B., Cooke, R., Delseny, M., Tohme, J. V. 2005. Using cDNA and genomic sequences as tools to develop SNP strategies in cassava (*Manihot esculenta Crantz*). *Theor. Appl. Genet.* 110: 425–431.

Magoon, M. L. and Krishnan, R. K. 1977. Extending the frontiers of genetic improvement in cassava. In Leaky, L. A. (eds.). *Proceedings, Third Symposium of the International Society for Tropical Root Crops*, Ibadan, Nigeria, 2-9 Dec. 1973, IITA, Ibadan, Nigeria, pp. 14-18.

Marmey, P., Beeching, J. R., Hamon, S. and Charrier, A. 1994. Evaluation of cassava (*Manihot esculenta Crantz*) germplasm collections using RAPD markers. *Euphytica.* 74 : 203-209.

Mba, R. E. C., Stephenson, P., Edwards, K., Melzer, S., Nkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, T. and Fregene, M. 2001. Simple sequence repeat (SSR) markers survey of the cassava genome: towards a SSR based molecular genetic map of Cassava. *Theor. Appl. Genet.*, 102:21-31.

Mignouna, H. D. and Dixon, A. G. O. 1997. Genetic relationships among cassava clones with varying levels of resistance to the *African mosaic disease* using RAPD markers. *Afr. J. Root and Tuber Crops.* 2 (1and2): 28-32.

Mohan, C., Shanmugasundaram, P., Maheswaran, M., Senthil, N., Raghu, D., and Unnikrishnan, M. 2013. Mapping New Genetic Markers Associated with CMD Resistance in Cassava (*Manihot esculenta Crantz*) using Simple Sequence Repeat Markers. *J. Agric. Sci.* 5

Mortimer, J., Batley, J., Love, C., Logan, E. and Edwards, D. 2005. Simple sequence repeat (SSR) and GC distribution in the *Arabidopsis thaliana* genome. *J. Plant Biotechnol.* 7: 17-25.

Montero-Rojas, M., Correa, A. M. and Siritunga, D. 2011. Molecular differentiation and diversity of cassava (*Manihot esculenta*) taken from 162 locations across Puerto Rico and assessed with microsatellite markers. AoB Plants, v. plr010, 2011. DOI: 10.1093/aobpla/plr010.

Moyib, O. K., Odunola, O. A. and Dixon, A. G. O. 2007. SSR markers reveal genetic variation between improved cassava cultivars and landraces within a collection of Nigerian cassava germplasm. 6 (23): 2666-2674.

Nassar, N. M. A. and Dorea, G. 1982. Protein content in some cassava cultivars and its hybrid with wild *Manihot* species. *Turrialba.* 32: 429-432.

Nassar, N. M. A. and Grattapaglia, D. 1986. Variabilidade de clones de mandioca em relação a fertilidade e aspectos morfológicos. *Turrialba* 36: 555-559.

Nassar, N. M. A. and O'Hair, S. 1985. Variation among cassava clones in relation to seed germination. *Indian J. Genet. Plant Breed.* 45: 394-398.

Nassar, N. M. A. 1986. Genetic variation of wild *Manihot* species native to Brazil and its potential for cassava improvement. *Field Crops Res.* 13: 177-184.

Nassar, N. M. A. 1985. *Manihot neusana* Nassar: A new species native to Paraná, Brazil. Can. *J. Plant Sci.* 65: 1097-1100.

Nassar, M. A., Nassar, H. N. M., Carvalho, G. G. and Vieira, C. 1996. Induction of a productive aneuploid in cassava, *Manihot esculenta Crantz. Braz. J. Genet.* 19:123-125.

Ndungu, J. N., Wachira, F. N., Okwaro, H., Obiero, H., Lelgut, D. K. and Kinyua, M. G. 2014. Genetic diversity study of Kenyan cassava germplasm using simple sequence repeats. *Afr. J. Biotechnol.* 13 (8): pp. 926-935.

Njock, T. E., Atiri, G. I., Thottappilly, G. and Thresh, J. M. 1996. Disease development and recovery in resistant and susceptible cassava genotypes infected by African cassava mosaic geminivirus. *Afr. Plant Prot.* 2: 97-102.

Ocampo, C., Hershey, C., Iglesias, C. and Iwanaga, M. 1993. Esterase isozyme fingerprinting of the cassava germplasm collection held at CIAT. In: Roca, W. and Thro, A. M. (eds.) *Proceedings of 1st International Science Meeting on Cassava Biotechnological work.* CIAT, Cali, Colombia. pp. 81-89.

Ogbe F.O., Dixon, A. G. O., Atiri, G. I. and Thottappilly, G. 2002. Restriction of virus movement into axillary buds in an important aspect of resistance in cassava to *African cassava mosaic virus. J. Phytopath.* 150: 546-552.

Ogbe, F. O., Atiri, G. I., Dixon, A. G. O. and Thottappilly, G. 2003. Symptom severity of cassava mosaic disease in relation to concentration on *African cassava mosaic virus* in different cassava genotypes. *Plant Pathol.* 52: 84-91.

Okogbenin, E., Egesi, C. N., Olasanmi, B., Ogundapo, O., Kahya, S., Hurtado, P., Marin, J., Akinbo, O., Mba, C., Gomez, H., de Vicente, C., Baiyeri, S., Uguru, M., Ewa, F. and Fregene, M. 2012. Molecular Marker Analysis and Validation of Resistance to *Cassava Mosaic Disease* in Elite Cassava Genotypes in Nigeria. *Crop Sci.* 52 (6): 2576-2586.

Okogbenin, E., Egesi, C., Mba, C., Espinosa, E., Santos, L. G., Ospina, C. and Porto, M. C. M. 2008. Marker-Assisted Introgression of resistance to *cassava mosaic disease* into Latin American germplasm for the genetic improvement of cassava in Africa. *Crop Sci.* 47 (5): 1895-1904.

Oliveria, E.J., Ferreria, V.S., Santos and Oliveria, G. A. F. 2014. Development of a cassava core collection based on single nucleotide polymorphism markers. *Genet and Mol. Res.* 13(3): 6472-6485.

Olsen, K. M. and Schaal, B. A. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta. Proceed. Natl. Acad. Sci.* USA. 96: 5586-5591.

Olsen, K. and Schaal, B. 2001. Microsatellite variation in cassava *Manihot esculenta, Euphorbiaceae* and its wild relatives: evidence for a southern Amazonian origin of domestication. *Am. J. Bot.* 88: 131-142.

Olsen, K. M. 2004. SNPs, SSRs and inferences on cassava's origin. *Plant Mol. Biol.* 56: 517-526.

Oscar, J., Maria, E., Zaldivar., Castro, L., Castro, E. and Barrantes, R. 2008. Microsatellite variation of cassava (*Manihot esculenta Crantz*) in home gardens of Chibchan Amerindians from Costa Rica. *Conserv Genet.* 9: 107-118.

Pariyo, A., Tukamuhabwa, P., Baguma, Y., Kawuki, R. S., Alicai, T., Gibson, P., Kanju, E., Wanjala, B. W., Harvey, J., Nzuki, I., Rabbi, I. Y. and Ferguson, M. 2013. Simple sequence repeats (SSR) diversity of cassava in South, East and Central Africa in relation to resistance to cassava brown streak disease. *Afr. J. Biotechnol.* 12: 4453-4464.

Paterson, A. H. 1996. Making genetic maps, Genome Mapping in Plants R. G. Landes Company, San Diego, California; Academic Press, Austin, Texas. 23-39.

Peroni, N., Kageyama, P. Y. and Begossi, A. 2007. Molecular differentiation, diversity, and folk classification of sweet and bitter cassava (*Manihot esculenta*) in Caicara and Caboclo management systems (Brazil), *Genet. Resour Crop Evol.* 54: 1333-1349.

Pootakham, W., Shearman J. R., Ruangareerate, P., Sonthirod, C., and Sangsrakru, D. 2014. Large scale SNP discovery through RNA sequencing and SNP Genotyping by targeted enrichment sequencing in cassava (*Manihot esculenta*). PloS ONE 9(12): e116028.doi:10.1371/Journel.phone.0116028.

Powell, W., Machray, G. C. and Provan, J. 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* 1: 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*). *Theor. Appl. Genet.* 125 (2): 329-342.

Rabbi, I. Y., Hamblin M. T., Kumar P. L., Gedil, M. A., Ikpan, A. S. and Jannink, J. L. 2014. High-resolution mapping of resistance to *cassava mosaic gemini viruses* in cassava using genotyping-by-sequencing and its implications for breeding. *Virus Res.* 186: 87-96.

Rabbi I, Y., Peter, A., Kulakow, Joseph, A., Manu-Aduening, Ansong, A., Dankyi, James, Y. and Asibuo. 2015. Tracking crop varieties using genotyping by sequencing

markers: a case study using cassava (*Manihot esculenta Crantz*). *BMC enetics* 16:115 doi:10.1186/s12863-0150273-1.

Raghu, D., Senthil, N., Saraswathi, T., Raveendran, M. and Gnanam, R. 2007. Morphological and simple sequence repeats (SSR) based DNA fingerprinting of south Indian cassava germplasm. *Int. J. Integr. Biol.* 1: 141-148.

Raji, A. A., Anderson, J. V., Kolade, O. A., Ugwu, C. D. and Dixon, A. G. 2009. Ingelbrecht Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility BMC. *Plant Biol.* 9: 118.

Roa, A. C., Chavarriaga-Aguirre, P., Duque, M. C., Maya, M. M., Bonierbale, M.W., Iglesias, C., and Tohme, J. 2000. Crossspecies amplification of cassava (*Manihot esculenta*) (*Euphorbiaceae*) microsatellites: allelic polymorphism and degree of relationship. *Am. J. Bot.* 87: 1647-1655.

Raji, A. A. J., Fawole, I., Gedil, M. and Dixon, A. G. O. 2009. Genetic differentiation analysis of African cassava (*Manihot esculenta*) landraces and elite germplasm using amplified fragment length polymorphism and simple sequence repeat markers. *Ann. Appl. Biol.* 155: 187-199.

Rogers, D. and Appan, C. 1973. *Manihot, Manihotoides, Euphorbiaceae.* Flora *Neotropica* Hafner Press, NY, pp. 272

Rogers, D. J. 1963. Studies of *Manihot esculenta Crantz* and related species. Bulletin Torrey Botanical Club. 90: 43-54.

Rogers, D. and Appan, C. 1973. *Manihot, Manihotoides, Euphorbiaceae. Flora Neotropica.* Hafner Press, New York, NY.

Rohlf, F. J., 1998. NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 2.02. Exter Software, Setauket, NY.

Sauer, J. D. 1993. Historical Geography of Crop Plants: a select Roster: CRC press, Inc., Boca Raton, FL, USA, pp. 309.

Sneath, P. H. A. and Sokal, R. R. (1973). Numerical taxonomy. WH Freeman, San Francisco. *Theor. Appl. Genet.* 93: 613-617.

Siqueira, M. V., Pinheiro, B. M., Borges T. T., Valle, A., T. L. and Zatarim, M. 2010. Microsatellite Polymorphisms in Cassava Landraces from the Cerrado Biome, Mato Grosso do Sul, Brazil Biochemical. *Genetics* 48: 879-895.

Soto, J. C., Ortiz, J. F., Jiménez, L. P., Vásquez, A. X., Lavalle, L. A., Mathew, B., Léon, J., Bernal, A. J., Ballvora, A. and López, C. E. 2015. A genetic map of cassava (*Manihot esculenta Crantz*) with integrated physical mapping of immunity-related genes. BMC *Genomics.* (16:190 DOI 10.1186/s12864-015-1397-4)

Storey, H. H. and Nichols, R. F. W. 1938. Studies of the mosaic disease of cassava. *Ann. Appl. Biol.* 25: 790-806.

Sumarani, G. O., Pillai, S. V., Harisankar, P. and Sundaresan, S. 2004. Isozyme analysis of indigenous cassava germplasm for identification of duplicates. *Genet. Resour. Crop Evol.* 51: 205-209.

Syvanen, A. C. 2001. Genotyping single nucleotide polymorphisms. *Nat. Rev. Genet.* 2: 930-942.

Tautz, D. 1989. Hypervariability of simple sequences as a general source of polymorphic DNA markers. *Nucl. Acid Res.* 17: 6463-6471.

Thresh, J. M., Otim-Nape, G. W., Legg, J. P. and Fargette, D. 1997. *African cassava mosaic virus* disease: The magnitude of the problem. *Afr. J. Root Tuber Crops.* 2: 13-18.

Thresh, J. M., Otim-Nape, G. W. and Jennings, D. L. 1994. Exploiting resistance to *African cassava mosaic virus. Aspects Appl. Biol.* 39: 51-60.

Thresh, J. M., Otim-Nape, G. W. and Fargette, D. 1998. The component and deployment of resistance to *cassava mosaic virus* disease. Integrated Pest Manag. Rev. 3: 209-224.

Turner, E. H., Lee, C., Nickerson, D. A. and Shendure, J. 2009. Massively parallel exon capture and library-free resequencing across 16 genomes. *Nat. Methods* 6: 315-316.

Ugent, D., Pozorski, S. and Pozorski, T. 1986. Archaeological manioc (*Manihot*) from Coastal Peru. *Econ. Bot.* 40: 78-102.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting, *Nucl. Acid Res.* 23: 4407-4414.

Van Esbroeck, G. and Bowman, D. T. 1998. Cotton improvement, cotton germplasm diversity and its importance to cultivar development. *J. Cotton Sci.* 2:121-129.

Welsch, R., Arango, J., Bar, C., Salazar, B., Al-Babili, S., Beltran, J., Chavarriaga, P., Ceballos, H., Tohme, J. and Beyer, P. 2010. Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *Plant Cell* 22: 3348-3356.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucl. Acid Res.* 8: 7213-7218.

Williams, J. G. K., Kuberlik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acid Res.* 18: 6531-6535.

Winter, P. and Kahl, G. 1995. Molecular marker technologies for plant improvement. *World J. Microbiol. Biotechnol.* 11: 438-448.

Zayed, E. M. and Shams, A. S. 2012. Genetic diversity in introduced cassava using inter simple sequence repeat markers (ISSRs). *Egypt. J. Agric. Res.* 90 (4): 225-236.

Appendices

8. APPENDIX I

a. DNA extraction buffer (Dellaporta *et al.*, 1983)

Tris- HCl (pH 8.0)	100mM	
EDTA (pH 8.0)	50mM	
NaCl	500mM	
β-mercaptoethanol	0.2 % (v/v)	freshly added prior to DNA extraction
PVP	2% (w/v)	

b. SDS 20%

c. Potassium acetate 500mM

d. Ice-cold Isopropanol

e. Sodium acetate(pH 5.2) 300mM

f. RNase A

10 mg/ml (RNase A was dissolved in TE buffer and boiled for 15 minutes at 100^0 C to destroy DNase and stored at –20 o C).

g. Chloroform:Isoamyl alcohol (24:1)

h. Ethanol 70%

APPENDIX II

TE buffer (10X)

Tris- HCl (pH 8.0)	10 mM
EDTA	1 mM

Final volume made upto 100ml with distilled water.

APPENDIX III

a. TBE buffer (10X)

Tris base	107 g
Boric acid	55 g
0.5 M EDTA (pH 8.0)	40 ml

Final volume made up to 1000 ml with distilled water and autoclave before use.

APPENDIX IV

a. Gel loading dye

Formamide	50ml
Xylene cyanol	50mg
EDTA (0.5M)	1ml
Bromophenol blue	50mg

b. Empty well dye

Loading dye (Appendix IV)	50µl
Sterile distilled water	50µl

c. 100bp marker

100bp marker	5µl
Loading dye	40µl
Sterile distilled water	55µl

APPENDIX V

PCR Cocktail for SSR and SNP primers

	Stock Concentration	Volume taken	Final concentration
DNA	50ng	2.0µl	5.0 ng
Primer	2.0µM	2.0 µl	0.2 µM
dNTPs	40µM	0 .2 µl	0.4 µM
Taq bu er	10x	2.0 µl	1x
Taq DNA polymerase	3.0 unit	0.2 µl	0.03 Unit
Sterile water		13.6 µl	

20 µl

APPENDIX VI

Acrylamide

Acrylamide - 38g

Bis acrylamide - 2g

Made up the final volume up to 100 ml using distilled water.

APPENDIX VII

6% Polyacrylamide gel containing 7 M urea

42 g urea was dissolved in a beaker containing 10 ml TBE buffer (10X) and 15 ml distilled water by heating in a microwave oven for 30-40 s. 15 ml acrylamide solution() was filtered and added to a measuring cylinder followed by the urea solution. The final volume was made up to 100 ml using distilled water and stored in dark till use. 60 µl TEMED and 600 µl APS (100 mg/ml) was added and mixed just before casting the gel.

APPENDIX VIII

Bind silane

Absolute ethanol	99.5%
Acetic acid	.5%
Bind silane	1.0 µl

APPENDIX IX

Fixer

Acetic acid	200 ml
Distilled water	1800 ml

APPENDIX X

Silver stain

2 g silver nitrate dissolved in distilled water to a final volume of 2000 ml and 3 ml formaldehyde added.

APPEND XI

Developer

60 g sodium carbonate dissolved in distilled water to a final volume of 2000 ml and stored at -20°C. 3 ml formaldehyde and 4 ml sodium thiosulphate (10 mg/ml) was added and mixed thoroughly before use.

Abstract

MOLECULAR CHARACTERIZATION OF CASSAVA MOSAIC DISEASE (CMD) RESISTANT VARIETIES AND WILD RELATIVES OF CASSAVA (*Manihot esculenta* Crantz) USING SSR AND SNP MARKERS

DHANYA O.G.

(2010-09-114)

Abstract of the thesis

Submitted in partial fulfilment of the

requirement for the degree of

MASTER OF SCIENCE (INTEGRATED) IN BIOTECHNOLOGY

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Department of Plant Biotechnology

B. Sc. - M.Sc. (Integrated) Biotechnology

COLLEGE OF AGRICULTURE

VELLAYANI, THIRUVANANTHAPURAM - 695 522

KERALA, INDIA

2015

9. ABSTRACT

In the present study an attempt was made to estimate the extent of genetic diversity between 25 CMD resistant, 16 susceptible and seven wild relatives of cassava using 14 SSR and two SNP primers. Out of the 16 primers, nine primers was analysed using PAGE and remaining seven using capillary electrophoresis on genetic analyzer. The primers produced a total of 53 alleles across the 48 cassava accessions. NS198 was found to be highly polymorphic with 6 allelles followed by NS169, SSR36 and SSR39 (5allelles). The two SNP marker analysis on genetic analyzer namely SNPAPX3 and SNP ERF revealed that out of the two peaks generated, one of the peak at a range of 650-690 and 500-530bp respectively was common to all the 48 accessions and the other peak was variable between samples.

The dendrogram constructed with 16 primers using UPGMA had four major clusters which clearly distinguished the resistant, susceptible and wild collections of cassava. The close observation made on one of the sub sub cluster within major resistant cluster revealed the resistant cultivars TME3 and TME4 were closely related with a similarity coefficient of 0.98. Clustering analysis was well supported by PCA, made representation of distinct location of CMD resistant, susceptible and wild relatives of cassava on 2D and 3D dimensions.

Comparison of PIC value for all the 16 primers found out the PIC value for individual SSR markers is higher than the individual SNP PIC value at a range of 0.19 - 0.24. SSR PIC values ranged from 0.347 in SSR32 with observed heterozygosity (He) 0.447 to 0.72 in NS198 with He 0.76. Based on the PIC ranges NS198, NS169, SSRY39, RME-1, SSR106 and NS158 were selected as highly polymorphic markers.