ASSOCIATION MAPPING FOR CASSAVA MOSAIC DISEASE (CMD) RESISTANCE IN CASSAVA USING SSR MARKER

by

APARNA T. K.

(2012-09-117)

THESIS

Submitted in partial fulfilment of the

Requirement for the degree of

B. Sc. - M. Sc. (INTEGRATED) BIOTECHNOLOGY

Faculty of Agriculture
Kerala Agricultural University, Thrissur

B. Sc. - M. Sc. (INTEGRATED) BIOTECHNOLOGY
DEPARTMENT OF PLANT BIOTECHNOLOGY
COLLEGE OF AGRICULTURE
VELLAYANI, THIRUVANANTHAPURAM - 695 522
KERALA, INDIA

2017

DECLARATION

I hereby declare that the thesis entitled **"Association mapping for cassava mosaic disease (CMD) resistance in cassava using SSR marker"** 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, associateship, fellowship or other similar title, of any other University or Society.

Place: Vellayani
Date: 03/11/2017

APARNA T. K.
(2012-09-117)

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

3

Dr. C. Mohan
Principal Scientist

CERTIFICATE

Certified that this thesis entitled **"Association mapping for cassava mosaic disease (CMD) resistance in cassava using SSR marker"** is a record of research work done independently by Ms. Aparna, T. K. (2012-09-117) 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 : ICAR-CTCRI, Trivandrum
Date : 03.11.2017

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
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. APARNA T. K., (2012-09-117) a candidate for the degree of B.Sc. - M.Sc. (Integrated) Biotechnology, agree that the thesis entitled **"Association mapping for Cassava Mosaic Disease (CMD) resistance in cassava using SSR marker"** may be submitted by Ms. APARNA T. K. in partial fulfilment of the requirement for the degree.

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

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

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

Dr. Jayaleksmy V. G.
(Member, Advisory Committee)
Professor
Dept. of Plant Breeding and Genetics
College of Agriculture, Vellayani,
Thiruvananthapuram-695 522

External Examiner

Dr. M. K. Rajesh
Principal Scientist
Division of Crop Improvement
ICAR-Central Plantation Crops Research Institute
Kasaragod-671 124

ACKNOWLEDGEMENT

With Boundless love and appreciation, I would like to extend my heartfelt gratitude and appreciation to the people who helped me in bringing the study into reality.

It is with my heartfelt feelings I would like to express my sincere gratitude to my beloved advisor Dr. C. Mohan, Principal Scientist, ICAR-CTCRI, Thiruvananthapuram for his continuous support, patience and motivation. In addition to his support for completing the project, his advice was valuable for me in every aspect.

My special thanks to the former Director of ICAR-CTCRI Dr. James George and the present Director Dr. Archana Mukherjee for allowing me to do my project work and for their support.

My sincere thanks to my advisory committee members: Dr. Swapna Alex, Dr. Jayaleksmy V.G., Dr. J. Sreekumar, Principal Scientists, ICAR-CTCRI not only for their insightful comments and encouragement, but also for their valuable counseling and constructive

My whole hearted thanks to Dr. K. B. Soni, Course Director for their valuable help, suggestions and advice during my research work. I would like to put on record my sincere thanks and gratitude to Dr. M. N. Sheela, Head, Division of Crop Improvement, ICAR-CTCRI for permitting me and extending all the facilities to complete my work.

My very sincere gratitude towards Mr. B. S. Prakash Krishnan, Technical Assistant, ICAR-CTCRI who helped me with the technical aspects of my research work. I wish to express my deep gratitude to all the scientists and staff members of ICAR-CTCRI, teachers in college, my seniors and juniors for their timely support.

my gratitude to my beloved friends, Revathi, Anjali, Babitha, Sreevidya, Aarathy, Ashna, Sadiq, Pooja Akshara, Abhilash, Shersha and all other classmates for their invaluable care, constant support, motivation and selfless help.

I owe this achievement to my parents, who always stood along my side and I will never forget the timely help, mental support, kindness and affection extended by other family members without them this work would have never seen light. Finally, I humbly thank the Almighty for showering His blessings and bestowing the wisdom, and physical ability to accomplish this work.

APARNA T. K

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

TABLE OF PLATES

TABLE OF APPENDICES

ABBREVIATIONS

°C	-	Degree Celsius
%	-	Percentage
µl	-	Micro litre
µM	-	Micro Molar
AFLP	-	Amplified Fragment Length Polymorphism
APS	-	Ammonium Per Sulphate
Bp	-	Base pair
BSA	-	Bulk Segregant Analysis
CMD	-	Cassava Mosaic Disease
CMV	-	Cassava Mosaic Virus
CTCRI	-	Central Tuber Crops Research Institute
DNA	-	Deoxy-ribonucleic acid
EDTA	-	Ethylene Di-amine Tetra Acetate
EST	-	Expressed Sequence Tag
et al.	-	And Other Co-Workers
Fig.	-	Figure
FOA	-	Food and Agricultural Organization
hr. (s)	-	Hour(s)
MAS	-	Marker Assisted Selection

ABBREVIATIONS

Min	-	Minutes
mg	-	Milli gram
mL	-	Milli liter
ng	-	Nano gram
OD	-	Optical Density
PAGE	-	Poly Acrylamide Gel Electrophoresis
PCA	-	Principal Component Analysis
PCR	-	Polymerase Chain Reaction
PIC	-	Polymorphism Information Content
PVP	-	Poly Vinyl Pyrrolidone
QTL	-	Quantitative Trait Loci
RAPD	-	Random Amplified Polymorphic DNA
RFLP	-	Restriction Fragment Length Polymorphism
SDS	-	Sodium Dodecyl Sulphate
Sec	-	Seconds
SNP	-	Single Nucleotide Polymorphism
SSR	-	Simple Sequence Repeats
TBE	-	Tris Borate EDTA
TEMED	-	Tetra Ethyl Methyl Ethylene Diamine

INTRODUCTION

14

Chapter 1

INTRODUCTION

Cassava (*Manihot esulenta* Cratz.) is one among the most important tropical tuber crops with high starchy roots and is using as a food material by millions of people in many countries. It was brought to India by Portuguese during seventeenth century. As per the survey result of specialized agency of the United Nations; Food and Agriculture Organization (FAO), cassava comes in fourth position after rice, wheat and maize as a food crop in developing countries.

Worldwide cassava production rate is 202.65 million tonnes, which is cultivated an area of 18.51 million ha with 10.95 t/ha. productivity. In India, cassava is mainly cultivated in an area of 216.66 thousand hectares. Kerala, Andhra Pradesh, Tamil Nadu, Meghalaya, Assam and Nagaland are the major cassava producing countries in India. According to the global cassava cultivation rate, India ranks first in the productivity (28 t/ha) and seventh in production (7 million tonnes).

Now a day's farmers are facing a lot of problems during the cultivation processes. One among the most serious problems is the invasion of Cassava mosaic disease (CMD) caused by gemini virus of the genus Begomovirus (Geminiviridae family) simply called cassava mosaic virus. (CMV). Whitefly (Bemisia tabaci) is the viral vector responsible for the disease. CMD was first reported in 1894 in Tanzania, it is now known to happen in various cassava cultivating countries of Africa, India, Sri Lanka and Indonesia causing 20 to 90 per cent of yield loss.

The best method to control the disease is the exploitation of host plant resistance. Polygenic resistance from Manihot glaziovii, a wild species of cassava introgressed into cultivating cassava species helps to protect them from CMD attack. Three major origin of host plant resistance in cassava are widely used by cassava breeders to resist CMD are *CMD1*, *CMD2* and *CMD3*.

15

Association mapping is a molecular method used to identify the disease-resistant loci and another quantitative trait loci within the chromosomal regions. So, association mapping also known as quantitative trait loci mapping. Molecular characterization of these loci helps to provide the detailed information related to the gene which is responsible for the disease resistance. Association mapping studies are based on the relationship of marker locus with trait locus. Therefore, some marker alleles travelling to the next generations along with the trait alleles mapping helps to identify single nucleotide polymorphism. single nucleotide polymorphism leads to the phenotypical vitiations are characterized by using different molecular markers. Molecular markers are DNA fragments which are present inside the genome.

In most association studies researchers must contend with the confusing effects of both family structure and population. TASSEL is a bioinformatical software that implements mixed linear model and general linear model approaches for controlling family structure and population. Tassel helps to calculate linkage disequilibrium statistics and the result visualized graphically. Browsing of databases and importation are assisted by integrated middleware. Diversity statistics calculation, insertion or deletion analysis, integration of genotypic and phenotypic data, imputing missing data, and principal components calculation are the other features of the software. Linkage mapping and association mapping are two important mapping processes in molecular biology. Linkage maps are widely using to identify genes at specific chromosomal regions. Association mapping is also known as QTL mapping'. Association mapping is more powerful with high resolution compared to the linkage mapping process.

Linkage mapping and Association mapping are two important mapping processes in molecular biology field. linkage maps are widely using to identify genes at specific chromosomal regions. Association mapping is also known as QTL mapping'. Association mapping is more powerful with high resolution compared to the linkage mapping process.

16

Wolfe *et al.*, (2016) conducted an association mapping study in cassava to identify the underlying genetic nature of CMD resistant cassava cultivars with molecular markers. They noticed a particular portion of chromosome 8, which confer about 66% of the disease resistance in African cassava varieties. multiple resistance allele was also identified within the chromosome number 8, by further dissection of quantitative trait loci (QTL).

Association mapping is also useful for studying trait of interest in various crops like potato, cotton, maize, etc. Association mapping studies helps to the development of stress tolerant rice varieties, wilt resistant cotton varieties etc. SSR markers are widely using for the association mapping studies. They are microsatellites, contain repeating sequence of 2-5 DNA bases.

The present study is an association mapping analysis to identify the markers which are closely associated with CMD resistance.

REVIEW OF LITERATURE

18

Chapter 2

REVIEW OF LITERATURE

2.1 Cassava

Cassava *(Manihot esculenta* Crantz*)* is the most important perennial tropical crop mainly cultivated in the regions where annual availability of rainfall is greater than 500 mm and the most suitable temperature for better plant growth is greater than 20°C. However, some varieties of cassava can also be grown in the altitude range between 1500-2000mm where the mean annual temperature is below 16°C. Cassava plants contain two major parts, woody shrub and leaves. Cassava is a woody shrub mainly propagating from stem cuttings. Base of the stem cuttings give rise to the adventitious roots. Some of the adventitious roots will produce from the nodal roots or axillary buds of the cuttings. Plant breeders uses cassava seeds as a propagating material. The main problem related to the production of cassava is the quick deterioration of starchy roots after harvesting process. This limitation prevents the storage of roots more than few days in fresh condition (Okezie and Kosikowski, 1982).

Among the genus *Manihot*, cassava is the only plant species which is used as a staple food. Cassava can produce large rate of soluble carbohydrate per unit area by converting solar energy Based up on the cyanogenic glycoside content and morphological traits cassava can be classified into various groups. However, the trait expression is mainly controlled by various environmental conditions. Plant height, stem colour, petiole colour, leaf shape, flower and inflorescence colour, were the major morphological traits used for the classification process. The plant contains both the male flowers and female flowers. Female flowers are seen at the base and male flowers are seen on tip of the single branched panicle. Duration between the planting and flowering is mainly depending on the environmental factors and the genotype of the plant selected. Usually the crop growth duration will be 8 -24 for months (Byrne, 1984).

19

2.1.1 Cassava mosaic disease (CMD)

Cassava mosaic disease (CMD) is one of the most dangerous cassava disease caused by cassava mosaic gemini virus coming under the family geminiviridae. These viruses belong to the genus begomoviral. (Fauquet and stately, 2003). CMD leads to the reduction of crop yield of over 1.5 billion US dollar a year (Thresh *et al.,* 1994). In Ghana CMD is a major constrain that adversely affect the cassava production (Lamptey *et al.,* 1998). In Ghana, 80% yield loss occurred in susceptible cultivars due to the attack of CMV (Moses *et al.,* 2007) Plant stunting and distortion are the major result of CMV attack (Lamptey *et al.,* 2000). Severe CMD disease leads to the formation of candlestick like sign on the CMD affected plants of cassava. Environmental factors are also having some effects on the expression of disease symptoms. Usually Symptom expression is more in the cool weather condition compare to the hotter weather condition (Gibson, 1994)

Development of plant varieties which are resistant to the CMD is quite important to improve crop yield in Ghana (Thresh *et al.*, 1997). Resistant varieties are developed by the process of interspecific hybridization. CMD resistant trait from *Manihot glaziovii* was integrated into the local varieties to confer disease resistance (Nicholas, 1947). Six various elements of CMD resistance was recognized by Fauquet and Fargette in 1990.These elements include field resistance (it is the total proportion of CMD attacked cassava plants in the field), virus resistance (it is the number of viral vector (whitefly) infected to each of the plants), virus and inoculation resistance, range of CMD sign, distribution resistance of virus (Thresh *et al.*, 1997).

2.1.2 Cassava mosaic virus (CMV)

Whitefly is the major vector accompanied with the virus spread. The virus can also be transferred from infected plants to normal plants. The plants which are produced from severely affected plant cutting was seriously affected with the CMD disease compare to the disease caused by the viral vectors. Viral infection in the late

stage of plant growth can be negligible compare to the productive phase (Thresh *et al.*, 1994).

There are mainly nine different viruses have been studied from the cassava plants which are attacked with CMD worldwide. Among the nine viruses seven viruses are characterized from sub Saharan-Africa. They are African cassava mosaic virus (ACMV), Indian cassava mosaic virus (ICMV), East African cassava mosaic virus (EACMV), Madagascar cassava mosaic virus (MCMV), Ugandan Variant of Eacmv (ugA), South African cassava mosaic virus (SACMV, and SriLankan cassava mosaic virus (SLCMV) (Fauquet and Stanley, 2003).

Most of the Gemini virus contains bipartite genome, which is single stranded, circular genome covered with a protein coat (30 KDa). Bipartite means the virus possess two types of DNA: DNAA and DNAB. The two elements share 200 base pair sequence commonly with length about 2700-2800bp.The DNAA codes for the encapsulation and viral replication related functions and which possess two open reading frames AV1 and AV2.AC1, AC2, AC3and AC4 are other complementary sense open reading frames along with the major open reading frames. DNAB is mainly for the movement related functions (Harrison and Robinson, 1999) and DNAB also two open reading frames. They are BV1 and BV2 contains EACMV-UG is a virus that linked with the serious epidemic in Uganda. EACMV-UG is the abbreviation of *East African cassava mosaic virus* Uganda2. Occurrence of a section in the coat protein AV1 gene of *ca* differentiate the genome of EACMV-UG from that of other viral strains. This section has high sequence identity with 500bp size (Deng *et al.*, 1997, Zhou *et al.*, 1997).

2.1.3 CMD dominant genes

Three major origin of host plant resistance are recognizing and widely used by cassava breeders to resist CMD.The first source of CMD was introgressed in Africa from *Manihot glaziovii* during initial breeding programmes (Nichols, 1947). Further

this methodology was found to be recessive along with polygenic in nature (Hahn *et al.,* 1980). After gene mapping process, this type of polygenic recessive resistance named as *CMD1* Cassava plants with this type of CMD resistance have been widely used in many of the African countries as tropical manihot species (TMS) series predominantly and they were produced by International Institute for Tropical Agriculture (IITA) Nigeria, through various breeding programmes. (Okogbenin *et al.*, 2013).

The second CMD host plant resistance source was identified from Nigeria and other West African countries during the period of 1980s and 1990s respectively. This type of resistance was discovered within the local cassava cultivars obtained from farmer's fields (Okogbenin *et al.*, 2013) and were prefixed with TME (Tropical *Manihot esculenta*). Which are originates from a dominant monogenic locus, named it as CMD*2* (Akano *et al.*, 2002). Cassava plants with *CMD2* host plant resistance were introduced into East Africa and commonly used by farmers in Uganda (Legg *et al.*, 2006). Development process of improved CMD2-type germplasm was supported by molecular markers (Akano *et al.*, 2002).CMD2 type germplasm widely favoured for African plant breeders due to the ease of single dominant locus introgression (Rabbi *et al.*, 2014)

The third CMD resistance mechanism named *CMD3* exist as quantitative trait locus (QTL). *CMD3* genes also provide high resistance capacity against CMD with little or without any sign of disease on plants (Okogbenin *et al.*, 2012).

Identification of molecular markers associating with these two CMD resistance genes help for the selection of germplasm with CMD resistance. (Bi *et al.*, 2010). The *CMD2* gene identification helps for the analysis of CMD resistance in genetic basis. Development of molecular markers for *CMD2* were facilitated to identify CMD resistant genotypes in Latin American cassava varieties and is validated for 4 years in the field and ensured the result. 64 Latin American genotypes were

identified as disease resistant after the field screening process. TMS 97/2205 and TMS 98/0505 are the two CMD resistant cassava varieties in Nigeria. These two varieties were analysed using specific primers followed by the field experiments.

Fregene and Puonti-Kaerlas identified the first CMD association markers in 2002. Subsequently, other markers associating with *CMD2* genes are also identified by Scientists (Carmo *et al.*, 2015). DNA sample isolated from various cassava varieties do not show occurrence of any CMD virus strain. PCR amplification with specific SSR primers (SSR 28, NS158, RME-1) after the amplification process with CMD virus stain specific primers followed by the Electrophoresis helps to study the presence of *CMD2* resistant gene. 38 genotypes of cassava were analysed to identify CMD associated resistant genes and to find cassava mosaic gemini virus strain that infecting Ghana cassava varieties. Screening of these 38 genotypes were done by several selected molecular and morphological markers. *CMD2* gene was identified in a Nigerian cassava variety TME3. Simple sequence repeat (SSR) and sequence characterized amplified region (SCAR) markers were used to construct the genetic map of this farmer preferred landrace. Molecular markers assist the finding and analysis of the disease resistant genes and other trait of interest. (Asare *et al.*, 2014).

2.2 Molecular markers

Different molecular techniques are available to analysis genetic diversity in biological system including different plants. These molecular technologies are not only applicable to study the genetic diversity, but also behavior and structure of genome of the species. Markers are used for the trait identification and there by characterize germplasm. On DNA based methods were used in the early techniques. Later, these techniques were replaced by DNA based molecular methods. Development of PCR based methods displaced early hybridization based methods. These developments were played a major role the crop improvement program. (Henry, 2008).

23

Naturally occurring DNA polymorphism is the real basis for the use of molecular markers. A polymorphic marker will exist in different forms, so that it helps to distinguish the mutant gene from the normal gene within the chromosome. DNA markers have several advantages over protein markers. DNA markers are do not affected by the environment and their segregation is in the single gene format. Genomic DNA isolation from the plant tissue is also an easy process, labor and cost effective. A best molecular marker has several ideal characteristics. These properties include codominant inheritance, high reproducibility, highly polymorphic nature, frequent occurrence in genome, easy access and easy exchange of information between laboratories. It is tough to find a molecular marker with all these ideal criteria. At least few of them are required according to the type of molecular study (Weising *et al.*, 1995).

Molecular markers are of two types. Polymerase chain reaction (PCR)-based markers and hybridization-based markers. In PCR based method in vitro amplification of template DNA occurs with the help of thermostable DNA polymerase and specific oligonucleotide primer. The newly synthesized amplicon was separated by electrophoresis method. In hybridization based marker method visualization of DNA profiles occur by hybridizing the restriction digested DNA to a labelled probe (Saiki *et al.*, 1985).

There are mainly 11 molecular markers. They are simple sequence repeats (SSRs) (Akkaya *et al.*, 1992), restriction fragment length polymorphism (RFLP) (Grodzicker *et al.*, 1975), single nucleotide polymorphisms (SNPs) (Jordan and Humphries, 1994), random amplified polymorphic DNA (RAPD) (Williams *et al.*, 1990) cleaved amplified polymorphic sequences (CAPS) (Akopyanz *et al.*, 1992) amplified fragment length polymorphism(AFLP) (Vos *et al.*, 1995), inter-simple sequence repeats (ISSRs) (Zietkiewicz *et al.*, 1994), expressed sequence tags (ESTs) (Adams *et al.*, (1991), diversity arrays technology (DArT) (Akbari *et al.*, 2006),

24

sequence characterized amplified regions (SCAR) (Paran and Michelmore, 1993), sequence tag sites (STSs) (Olsen *et al.*, 1989).

2.2.1 SSR marker

SSR markers are widely using for plant genetics studies over the past 20 years. SSR markers are also known as short tandem repeats, microsatellites and sequence tagged microsatellite sites. These markers have been extensively used for plant genotyping because of their improved characters. They are codominant markers and so the heterozygous can be differentiated. They are also highly informative and multiallelic genetic markers. The major distinctive characteristics of SSR markers are their hyper variability and their high degree of allelic variation. Among closely related species these markers are easily transferable and reproducible. (Mason, 2015). SSR markers are PCR based and easy to identify through PAGE or AGE.

SSR markers have been specifically useful for the preparation of maps for different plant species. Sub-families of these plant species are used for the linkage map construction process (Garcia *et al.*, 2006; Souza *et al.*, 2013; Pereira *et al.*, 2013). These markers are also using in the field of population structure analysis, genetic map construction process and evolutionary studies. Parentage analysis and forensic studies are also done by using SSR markers with the length of basic repeats 3 to 5 nucleotides. (Benson, 1999).

2.2.2 SSR marker study on CMD

SSR marker study was performed in cassava accessions to identify the CMD resistant plant varieties. The accessions selected for the analysis included a breeding stock with 58308 clones, 62 CMD resistant accessions,10 CMD susceptible accessions and five improved lines. Five cluster groups were generated from UPGMA evaluation on data obtained from 18 SSR primer pairs. Genetic variation calculated by Fst and Gst were 12.1% and 9.6% respectively, representing a feeble

25

genetic structure. Average gene diversity value was higher in all the five clusters obtained with an average heterozygosity value ranges 0.591 ± 0. 061.. 51.4% diversity on genes were observed within all the selected cassava genotypes and 46.6% of genetic variation was noticed among the cluster groups (Lokko *et al.*, 2006)

Multiple marker analysis was conducted on collected cassava germplasm to identify the presence of *CMD2* gene in the progenies and progenitors Occurrence of CMD resistant gene on parental cultivars need alleles which are closely linked with *CMD2* and segregating into next generation with CMD resistant character. Scientists were used four markers for the study. Markers included three SSR markers include SSR169, SSRY28, and SSR158 one SCAR marker (RME1). Three SSR markers of the four markers shown alleles connected with the *CMD2* gene for TMEII cassava germplasm used. And only 1 SSR marker associated *CMD2* gene in Dabodabo. All four marker alleles linked with *CMD2* for two genotypes (AR14-10 and CR52A-31) collected from CIAT.The result revealed that that it was the feasible basis of CMD resistance in these cultivars (Bi *et al.*, 2010)

A study was done to analyze the genetic diversity based on SSR markers, and morphological traits to characterize the SSR markers closely associated with cassava mosaic disease-resistant QTL region. Three SSR markers and 28 morphological traits were used for the genetic diversity study among the chosen cassava varieties. PIC analysis was done and the result provides great information for all the selected SSR markers. 13 and 10 clusters were formed for SSR markers and morphological traits with a similarity coefficient 0.75, similarity coefficient value of 0.75 and 2.514 respectively. Total SSR2, SSR28 and SSR324 were molecular markers used for the study. Polymorphic allele generated by the primers SSR324 and SSR2 ranging from 189bp and 138bp respectively.PIC obtained for SSR 28 marker was higher than other SSR primers. Low PIC was obtained for SSR2. Scientists were calculated Jaccard's similarity coefficient for alleles of SSRR primer and they created phylogenetic tree using UPMA, NTSYS-Clusters (II and XII) comprised of cassava accessions with

yield contributing traits and accessions with CMD susceptibility and low yield were assembled together into cluster IX. The morphological traits diversity that Genotypes ME460 and ME209 were two selected accessions contain high starch content and higher yield with CMD resistance. There varieties were coming under cluster VI (Duraisamy *et al.*, 2011)

SSR markers were used to characterize potato cultivars in molecular level.38 cassava accessions were used for the molecular characterization. 46 alleles generated from 10 alleles were used for the study as binary data. Scientists were used UPGMA method and Jaccard's similarity coefficient for cluster analysis with the help of NTSYSpc software. Only five alleles of 46 were occurred in all the analyzed cultivars, Rest of the 41 alleles were polymorphic for all the cultivars selected for the study, therefore 89.1% polymorphism was present (Favoretto *et al.*, 2011).

A mapping population was produce by the cross between local CMD resistant variety TME3 and the susceptible hybrid variety TMS3055.The first codominant inheritance marker SSR-28 was used to construct the genetic map of newly synthesized mapping population (Akano *et al.*, 2002). TME3 was the actual source for the resistance to CMD disease. But scientists were used TME14 variety for the study, because TME-14 is the duplicate of TME-3 variety. Genotyping of these variety revealed that the *CMD2* gene was related to the allele size for all the markers. (Rabbi *et al.*, 2014)

A study was conducted by Federal University of Bahia Reconcavo to find the cassava varieties with CMD resistance. They identified the resistant accessions by using MAS process. MAS is a method to selected a trait of interests with the help of suitable molecular markers. Scientists were genotyped thousand two hundred twenty-four cassava accession. Genotyping was done by five selected CMD associated SSR markers. They were RME1(Fregene *et al.*, 2006), NS169 (Okogbenin *et al.*, 2012), SSRY40 (Mba *et al.*, 2001), SSRY28 (Akano *et al.*, 2002) and NS158 (Fregene *et al.*,

27

2006). From the selected cassava accessions scientists were identified that only 5, 4 and 5 percentage accessions possess *CMD2* gene and NS169+RME1, NS158+RME1 and SSRY28+RME1 were the markers flanking respectively. From the study noticed that only seven cassava accessions were presented all the selected markers associated with CMD. A specific study was conducted to identify the existence of alleles that are closely linked with CMD2 for the markers NS158, RME1, SSRY28 and NS169. These markers contribute16,27,23 and 19 percentage of cassava individual had CMD resistance associated alleles respectively (Carmo *et al.*, 2015).

A collaborative work was done by scientists at International Institute of Tropical Agriculture (IITA) Nigeria, Georgetown University Medical Centre, Washington, DC USA. Ibadan, Nigeria and Department of Crop Science, University of Ghana, Accra, Ghana to analyze the molecular markers related to the CMD resistance. The study was done with the help of F1 progenies originated from the cross between CMD resistant and CMD susceptible variety TME7 and TMS30555 respectively. MAB was the first step they done for the identification of appropriate markers. Bulk sergeant analysis (BSA) with the help of selected SSR and RFLP marker revealed that, SSRY28-180 contributed by the CMD resistant parent was associated with CMD resistance. Association between trait and marker was identified by the regression analysis. The results revealed that the selected molecular marker accounts about 57.42 percentage of the total variation in phenotypes for resistance (Lokko *et al.*, 2005).

28

Table 1. SSR markers used for cassava diversity studies.

Accessions used	References
69 accessions from Kenyan germplasm,7SSR primers	Ndungu *et al.*, 2014
43 accessions from Ghana ,20 SSR markers	Asare *et al.*, 2011
36 Cassava genotypes,16 SSR markers	Moyib *et al.*, 2007
138 Nigerian resistant accessions	Lokko *et al.*, 2005
60 accessions from Brazilian geographical areas,11 SSR markers	Siquera, *et al.*, 2009
38 cassava accessions,10 SSR primers	Elias *et al.*, 2004
521 cassava accessions	Chavarriaga-Aguirre *et al.*, 1998
60 genotypes 14 polymorphic microsatellite markers	Mezette *et al.*, 2013
24 central Kerala collections and 12 varieties released,36 SSR markers	Lekha *et al.*, 2011
18 cassava accessions,3 SSR markers	Bi *et al.*, 2010

2.3 Association mapping and linkage disequilibrium

The major strategy behind the entire plant breeding progress is the study of complex quantitative traits with economic importance. Linkage mapping has been used traditionally for reading phenotypic variations in plants in the form of sequence variations in DNA and their position within the genome (Myles *et al.*, 2009) Linkage mapping is mainly based on the development of a family with known parentage by the crossing process and the next step is the identification of co-segregating genetic markers with target trait within the family (Al-Maskari *et al.*, 2012).

Initially association mapping was widely applied for the study of human genetics. Now a day it shows great opportunity for plant genetic studies. However, both the two approaches are mainly based on the association of genetic markers with a phenotypical trait of interest like crop yield, stress tolerance etc. Association mapping using natural population to identify association between marker and trait using linkage disequilibrium (LD) (Garcia *et al.*, 2003). LD is the reduced level of crossing over or minimum level of recombination. It also refers to the non-random organization of alleles at different loci. Various statistical methods are used to measure the LD level. The range has been used to construct genetic map and lastly produce gene copies controlling complex traits. (Risch and Merikangas, 1996, Weiss and Clark, 2002, Chapman *et al.*, 2003, Taniguchi *et al.*, 2006). In 2001 this method was extended to plants and the resolution of map was reported as high compared to the F1 derived mapping population. Association mapping provides numerous benefits over conventional family based linkage mapping. A large genetic variability occurs in nature as germplasm. These germplasm offers large coverage of alleles and helps to avoid cost and save time required for the expensive biparental mapping populations. (Thornsberry *et al.*, 2001).

There are mainly four types of association mapping population. They are germplasm bank Natural and synthetic population and elite breeding material. This

classification is mainly based up on the source of material. These different populations differ from each other with respect to the structure of population, phenotypic diversity, genotypic diversity and extent of linkage disequilibrium and the relationship of families (Breseghello and Sorrells, 2006).

Linkage disequilibrium and linkage equilibrium are two major concepts in population genetics to describe the allelic linkage in various loci. Linkage disequilibrium is the non-random pattern of alleles on different loci and is using to define non-equal haplotype frequencies. PAB $\neq$ PA x PB is the equation using to define linkage disequilibrium. A and B are two alleles at two different loci. PAB is the frequency of haplotype contains both the two alleles. PA denote the frequency of haplotype having A allele and PB is he frequency of haplotype contains B allele only (Gupta *et al.,* 2005). Linkage disequilibrium was first described in 1917 by the scientist Jennings and quantified in 1964 by Lewtonin (Abdurakhmonov and Abdukarimov, 2008).

Genome wide association study was conducted by Wolfe *et al.,* (2016) to examine the genetic nature of CMD resistance in cassava using 6125 cassava genotypes and 42,113 single-nucleotide polymorphism (SNP) markers. They identified a specific region on chromosome 8 that provides 30 to 66% of genetic resistance against CMD in cassava cultivars of Africa. Thirteen regions with minute effects were also noticed. Multiple resistance alleles or two epistatic loci were identified by further examination of quantitative trait loci(QTL) in chromosome 8.

CMD resistant cassava line TME3 and the CMD susceptible cassava line TMS30555 along with 18 selected cassava accessions were analysed by *CMD2* gene associated markers comprised of simple sequence repeats (SSR) NS158 SSRY28 and SCAR marker. RME1.Only 3 varieties named 11Q, T7 and N13 were displayed the expected banding pattern same as TME3 (Bi *et al*., 2010).

31

Genome wide association mapping study was conducted by Esuma *et al.*, (2016) to analyze the genomic region which involves in the control of carotenoid production in cassava. 591 genotypes were used for the study. A total of 179,310 SNPs was identified and they were located across the genome of cassava. A genomic region within a stretch of 1.37 Mb situated on chromosome 1 was noticed with four SNPs, which are associated with carotenoid production in cassava genotypes. Purposeful annotation of this specific gene connects its biological action to the *protein synthase* enzyme, this enzyme is related to the biosynthesis of carotenoid in plants (Goodstein *et al.*, 2012)

Association mapping study was performed in hundred and fifty-eight elite cotton (*Gossypium hirsutum* L.) varieties using 212 whole genome-wide markers and phenotypic data collected by greenhouse screening and nursery screening method. They were analyzed for verticillium wilt resistance by the evaluation of association mapping, linkage disequilibrium and population structure. 480 alleles which are ranging from two to four loci were recognized from all the collected varieties. In the association group significant LD (P,0.001) was shown by only 8.2% of linked marker pair (Zhao *et al.*, 2014)

Association study was conducted by Swamy *et al.*, (2017) for the development of stress tolerant rice varieties. Rice germplasm collected from Malaysia. The study was done for yield and yield-related traits and noticed significant trait marker associations through the structured association mapping program. Three subgroups with admixtures were identified in Malaysian rice germplasm used for the study. LD analysis helps to confirm the marker trait association. LD analysis showed that the LD value increases with decrease in distance between marker pairs and the LD decay value varied from 5 to 20 cM. A total 80 marker trait association were identified by mixed linear model based structured association process for three quantitative characters. They were plant height (PH), days to flowering (DTF) and grain yield (GY). Seven marker trait associations were noticed for the yield of grains under drought condition.

32

Most of them are located on the chromosome number 2, 5, 10, 11 and 12 phenotypic variance (PV) value varied among the population 5% to 19%. Several drought-responsive genes are also identified by in-silico examination of distinct QTL region related to drought tolerance. The study results revealed that the structure based association mapping is a best method to identify major influence of QTLs for drought tolerance-related traits in rice.

Certain association mapping studies were performed in potato varieties at Max Planck Plant Breeding Research Institute. The experiments were based on the candidate gene genotyping method (Gebhardt *et al.* 2004, Li et al. 2005, Malosetti *et al.* 2007). Candidate gene approach was an appropriate method to identify association between marker and trait. It is a biased method. Here, candidate genes are chosen according to the details taken from physiological, biochemical and genetic studies in both model and non-model species of plants. The candidate gene method is only forthright when limited to traits, of which the molecular and biochemical details are understood well (Hall *et al.*, 2010). High density genotyping is required for genome wide association (GWAS) studies. SolCAP SNP array was a method used for the genotyping process of eight diploid and thirty-six tetraploid breeding using 8303 clones. The major objective of the study was the identification of population structure, genetic diversity, degree of polymorphism in European germplasm for SolCAP SNPs and genome-wide coverage of linkage disequilibrium. A clear separation of diploid from tetraploid genotypes were observed by principal coordinate analysis. Separate sub categories or subgroups were not noticed in tetraploid varieties (Stich *et al.*, 2013).

So, by the association mapping studies we can identify the markers which are closely associated with CMD resistance.

MATERIALS AND METHODS

Chapter 3

34

MATERIALS AND METHODS

This study was conducted at ICAR-Central Tuber Crop Research Institute Sreekariyam, Thiruvananthapuram during 2016-2017. The laboratory facilities were utilized from Division of Crop Improvement of this Institute.

I. Materials

3.1 Planting.

From the cassava breeding project, 30 CMD resistant and 25 CMD susceptible cassava lines were selected and all are planted in the field of ICAR-CTCRI farm, for the phenotypic and genotypic analysis with the objective of identifying CMD resistance associated marker.

3.2 Sample collection

Fresh young leaves were collected from different varieties of cassava plants growing in the field. Collected leaf samples consist of 30 CMD resistant (1-30), 25 CMD susceptible (31-55) accessions (Table 2). The CMD susceptible plants were identified by the presence of mosaic symptom on the leaves.

Table 2. Cassava varieties collected for the experiment

Sl. No	Cultivars	Code	Sl. No	Cultivars	Code
1	CMR19	R1	29	CMR197	R29
2	CMR55	R2	30	CMR213	R30
3	CMR73	R3	31	CI732	S1
4	CMR196	R4	32	CO2	S2
5	CMR218	R5	33	H165	S3
6	CMR21	R6	34	H226	S4
7	96/1089A	R7	35	11-S-33	S5
8	CR 2042	R8	36	Sree Rekha	S6
9	CR 24	R9	37	Sree Vijaya	S7
10	TME3	R10	38	M4	S8
11	8W5	R11	39	BR 2	S9
12	R15	R12	40	BR 5	S10
13	CR 43/11	R13	41	BR 6	S11
14	CMR 195	R14	42	BR 10	S12
15	CMR13	R15	43	GO31	S13
16	CMR9	R16	44	Kollam local	S14
17	CMR65	R17	45	849	S15
18	CMR138	R18	46	C 32	S16
19	CMR205	R19	47	C 50	S17
20	CMR14	R20	48	C 77	S18
21	CMR18	R21	49	Farm Local	S19
22	CMR22	R22	50	Sree Athulya	S20
23	CMR 26	R23	51	Sree Pavithra	S21
24	CMR29	R24	52	Sree Swarna	S22
25	CMR30	R25	53	Sree Jaya	S23
26	CMR51	R26	54	CI 848	S24
27	CMR55	R27	55	Sree PrabhaTCH2	S25
28	CMR109	R28			

II. Methods

36

3.3 Genomic DNA isolation

1. Genomic DNA was isolated from collected young fresh leaves of cassava using Dellaporta *et al.,* (1983) DNA isolation method.
2. 3-4-gram cassava leaves were weighed and ground into fine powder using sterilized mortar and pestle with the help of liquid nitrogen.
3. Powder transferred into sterilized centrifuge tubes using sterilized spatula.
4. Added 15ml extraction buffer.
5. 100-200 mg PVP and 20µl β-mercaptoethanol were added to the mixture. Mixed well and Kept at 4°C on ice/fridge.
6. Added 1ml of 20% SDS into the centrifuge tubes and mixed well. Kept it in water bath for 30minutes to 1 hours. Mixed well in between 15 minutes once.
7. Added 5ml of 5M Potassium acetate to the tubes and mixed well
8. Incubated on ice at 4°C for 30 minutes
9. Centrifuged tubes at 10,000 rpm for 30 minutes and transferred the upper layer into new sterilized centrifuge tubes. Discarded the residues.
10. Added 10 ml of isopropanol into the supernatant. The tube was inverted gently to precipitate the DNA and incubated in fridge at 4°C of overnight.
11. The precipitate DNA pooled out from the centrifuge tubes and transferred into sterilized eppendorf tubes.
12. Centrifuged the tubes at10,000 rpm for few minutes to remove the solution.
13. Discarded the supernatant. Resuspended the DNA in 500µl TE buffer /sterilized water. Incubated in water bath at 65°C for 10 minutes.
14. 3µl RNase was added to the tubes and incubated at 37°C for 1 hour to degrade the RNA
15. Added equal volume of 24:1 chloroform: isoamyl alcohol. Mixed by gentle inversion to remove proteins.

37

16. Centrifuged at 10,000 rpm for 20 minutes. Transferred the supernatant into new sterilized eppendorf tubes and added ice cold ethanol to precipitate the DNA.
17. Incubated at -20°C for 2 hours/ 4°C overnight
18. Centrifuged at10,000rpm for 5 minutes. DNA pellet was washed with 500µl 70% alcohol. Washed twice and centrifuged for few Seconds.
19. Resuspended the DNA in 500µl TE buffer.

3.4 DNA quality checking

Agarose gel electrophoresis

Quality of DNA was checked in 0.8% agarose gel electrophoresis

1. 0.8g agar powder weighed and dissolved in 100ml 1X TBE buffer. Agarose was dissolved by keeping it in microwave oven.
2. After cooling to 50-60°C 2µl ethidium bromide (10 mg/ml) was added to and mixed well.
3. Poured the agarose solution into the gel casting tray where the comb was placed. After the gel solidification, the gel tray was transferred into the electrophoresis unit filled with 1x TBE buffer.
4. Removed the comb and 2µl DNA along with 3µl loading dye was loaded into the wells. The DNA was run at 80-100 volts for 15 min.
5. Gel was documented using gel documentation system (Syngene) and bands were visualized.

3.5 DNA quantification

38

Nano-spectrophotometry

DNA concentration and purity were checked using Nano-spectrophotometer by taking the absorbance at 260/280nm.

1. 1 µl distilled water placed on to the pedestal of Nano-spectrophotometer system using pipet to remove the dust particle. Wiped pedestal with Kim wipes.
2. Nano Drop system was calibrated using 1µl TE buffer. Again, wiped with kimwipes thoroughly.
3. 1µl sample was placed directly onto the pedestal using pipet.
4. Clicked the 'Measure' button on the screen of Nano-spectrophotometer. Noted the values displayed on the screen like concentration of DNA in ng/µl, A260, A280 A260/A230 and A260/280.

3.6 PCR Amplification for SSR primers

PCR amplifications was done in controlled condition (Table 3) for 20 SSR primers (Table 4) were performed in the DNA isolated from cassava verities.

1. DNA was diluted to 10ng/µg according to the concentration value obtained from the Nano Drop system
2. 20 SSR primers were used for the molecular study. Primers were diluted in nuclease free water.
3. 2µM primer stock was prepared by dissolving 4µl of each forward and reverse primer in 192µl of sterilized water.
4. 20 µl reaction volume was prepared contains DNA, forward and reverse primer, 1X buffer, DNA taq polymerase dNTPs and nuclease free pure water.
5. Cocktail was prepared according to the procedure
6. Amplification process was done by using Bio Rad C1000^{TM} thermal cycler.

39

Table 3. PCR Program

Sl. No	Reaction	Temperature	Duration (min.)
1.	Initial denaturation	94°C	5
2.	Denaturation	94°C	1
3.	Annealing	58°C	2
4.	Extension	72°C	2
5.	Final extension	72°C	5
6.	Holding	4°C	∞ times

Table 4. SSR primers used for molecular characterisation of cassava

Locus	Forward primer (5'-3') and Reverse primer	Product size (bp)	Source
SSRY40	GCATCATGGTCCACTCACT CATTCTTTTTCGGCATTCCAT	293	Mba, *et al.*, (2001)
NS158	GTGCGAAATGGAAATCAATG TGAAATAGTGATACATGCAAAAGGA	166	Akano et *al.*, (2002)
NS169	GTGCGAAATGGAAATCAATG GCCTTCTCAGCATATGGAGC	319	Rabbi *et al.*, (2014)
SSRY28	TTGACATGAGTGATATTTTCTTGAG GCTGCGTGCAAAACTAAAAT	180	Mba *et al.*, (2001)
RME1	ATGTTAATGTAATGAAAGAGC AGAAGAGGGTAGGAGTTATGT	700	Rabbi *et al.*, (2014)
SSRY21	CCTGCCACAATATTGAAATGG CAACAATTGGACTAAGCAGCA	192	Mba *et al.*, (2001)
SSRY6	TTTGTTGCGTTTAGAAAGGTGA AACAAATCATTACGATCCATTTGA	298	Mba *et al.*, (2001)
SSRY7	TGCCTAAGGAAAATTCATTCAT TGCTAAGCTGGTCATGCACT	250	Mba *et al.*, (2001)
SSRY42	TTCCTCCAAAGTTATCTAGAACCA CAATCCTTGTAGTAGCCAGTCTCA	221	Mba *et al.*, (2001)
SSRY77	CAGGAGGTGGCAGATTTTGT GCATGTTCCACCTGCATAAG	275	Mba *et al.*, (2001)
SSRY106	GGAAACTGCTTGCACAAAGA CAGCAAGACCATCACCAGTTT	270	Mba *et al.*, (2001)
SSRY136	CGACTGCATCAGAACAATGC AGCATGTCATTGCACCAAAC	296	Mba *et al.*, (2001)
SSRY44	GGTTCAAGCATTCACCTTGC GACTATTTGTGATGAAGGCTTGC	194	Mba *et al.*, (2001)
SSRY32	CAAATTTGCAACAATAGAGAACA TCCACAAAGTCGTCCATTACA	298	Mba *et al.*, (2001)
SSR36	CAACTGTTTCAACCAACAGACA ATTCTCGTGAACTGCTTGGC	134	Mba *et al.*, (2001)
SSRY39	TCAATGCATAGGATTTTGAAAGTA AATGAAATGTCAGCTCATGCT	293	Mba *et al.*, (2001)
NS198	TGCAGCATATCAGGCATTTC TGGAAGCATGCATCAAATGT	170-120	Asare *et al.*, (2004)
SSRY235	CAGCTTTGCCATCCAATTTT CAGCAAAATGACATGAGTGTATCTC	166	Fregene *et al.*, (1997)
SSRY324	CGCTTACAAACCACCTTCA GCTTGATCTCAGCCATGTCA	206	Lokko *et al.*, (2005)
SSRY9	ACAATTCATCATGAGTCATCAACT CCGTTATTGTTCCTGGTCCT	278	Mba *et al.*, (2001)

3.7 Polyacrylamide gel electrophoresis (PAGE)

41

3.7.1 Cleaning and assembling glass plates

1. Small glass plates were dipped in 2% sodium hydroxide overnight and the plates were cleaned thoroughly and washed in tap water.
2. Cleaned both large plate and small plate with laboline to remove dust particles and oil contaminant. Then washed with water and removed by wiping with tissue paper.
3. Cleaned both the plates using absolute alcohol with kimwipe.
4. Wiped small plate with 2ml of bind silane and spread uniformly with kimwipes.
5. 2ml repellent (Laboline) was applied on the large plate and spread uniformly with kimwipes.
6. Wiped spacers, side clamps, bottom assembly, caster gasket, comb with alcohol.
7. Spacers were kept on the both side of large and placed small plate above the large plate carefully. Coated side was towards the spacer.
8. The plates together were lifted vertically on the table and clamped on both sides using side lever lamb set and locked them by pressing the levers close to the IPC panel.
9. Assembled unit was then placed into a bottom precision caster with a bottom edge resting against the gray gasket inside. The cam pegs were pushed inside and tightened the unit. And locked the handles by turning in perpendicular direction
10. Then the unit was laid down on a flat slab.

3.7.2 Gel preparation and casting

1. 15ml of 40% acrylamide was filtered in a beaker
2. Weighed 42gm urea and added 15 ml 1 X TBE buffer and little water in it and dissolved by keeping it in microwave oven.
3. Filtered the solution into the same beaker containing acrylamide solution with same filter paper and made up the volume into 100 ml.

4. Added 600µl (10%) APS and 60µl TEMED just before casting the gel.
5. A syringe was used to suck the solution and injected solution into unit through the notched region on the caster base where both the glass plates are aligned together. Placed comb above the glass plate after reaching the gel on the top position.
6. Allowed it to solidify for 30 min in an appropriate position without any disturbance

3.7.3 Gel running

1. Precision caster base was dislodged and the assembled apparatus with casted gel was then fitted vertically into the universal base using stabilizer bar.
2. The upper and lower buffer chambers filled with 1 X TBE buffer. Removed the comb and allowed it to pre-run for 20 min at 100 W.
3. After pre-run, the power supply was stopped and the wells were cleaned to remove urea using 100µl pipette.
4. PCR DNA samples were then denatured along with the PAGE dye at 95°C for 5min in thermal cycler. After denaturation, the PCR tubes were placed in a chiller and 3-4µl samples were loaded in each well along with 100bp ladder.
5. Samples were electrophorized at 100W power supply for 30-40 min.
6. After completion of gel run power supply was turned off and collected TBE buffer partially from upper buffer chamber by connecting a connector to the drain port on the gel unit.
7. The unit was disassembled and separated large and small plates.

3.7.4 Silver staining

After the removal of small plate from the assembly, the small plate was carried into large tray containing 2 L fixer solution and placed the tray in a shaker for 20 min. The gel should be faced upward. After the gel fixing process, the glass plate was washed in water for 5min in another tray. After washing process, the gel plate was dipped in a tray containing silver nitrate ($AgNO_3$) solution and placed the tray in a shaker for 20 min. Plate was washed in sterile water for few seconds and removed. The gel plate was transferred into developer solution. A white surface was placed under the plate to visualize the bands. After few seconds bands were appeared and the plate was transferred into fixer solution for 5 min for the proper fixating of DNA. Fixer residue was washed out in sterile water. Gel plate was allowed to dry.

3.8 Data analysis

3.8.1 PAGE data scoring.

Banding pattern obtained from the PAGE result was scored based upon the presence (1) or absence (0) of individual bands. Each of the alleles scored separately by comparison with 100 bp molecular ladder. Each band pattern generated by the SSR primers are considered as single locus.

3.8.2 Data analysis using TASSEL software

Tassel software is used to identify markers which are closely associated with CMD resistance and relationship between genotype and phenotype. The scored data converted into SNP and was used for the analysis. Dissimilarity matrix and cladogram were constructed by the software. Dissimilarity matrix used to identify the genetic dissimilarity between selected cassava accessions. Cladogram was used to identify inter relationship between accessions.

3.8.3 Data analysis using STRUCTURE software

44

STRUCTURE is a population analysis tool developed by Pritchard *et al.*, (2000).The input file for STRUCTURE analysis was obtained from the scored data of SSR marker experiment result. STRUCTURE software, was used to analysis distinct genetic population in the experiment samples and identified whether there is any admixed population or migrant. The details were obtained from a graph between 'K' value and "Mean of est. Ln prob of data" obtained from the structure harvester program and the alpha value obtained from the software.

RESULTS

46

Chapter 4

RESULTS

The present study on "Association mapping for cassava mosaic disease (CMD) resistance in cassava using SSR markers" was carried out to identify the markers which are associated with CMD resistance in cassava varieties. The results of the experiments conducted are described below.

4.1 Plant material selection

Cassava young leaves were collected from different varieties of cassava plants growing in the field of ICAR-Central Tuber Crop Research Institute (CTCRI), Sreekariyam. The collected plant materials include leaves of 30 resistant varieties and 25 susceptible varieties. The susceptible plant varieties are selected based upon the expressed disease symptoms, and resistant varieties are selected based on the healthy appearance and based on the previous CMD resistant data (Plate 1, 2).

4.2 Genomic DNA isolation

Genomic DNA was isolated from fresh young leaves of cassava using Delleporta DNA isolation method. Quality of isolated DNA was checked in 0.8% Agarose gel and bands were visualized using gel documentation system and the quantity was determined in nano spectrophotometer.

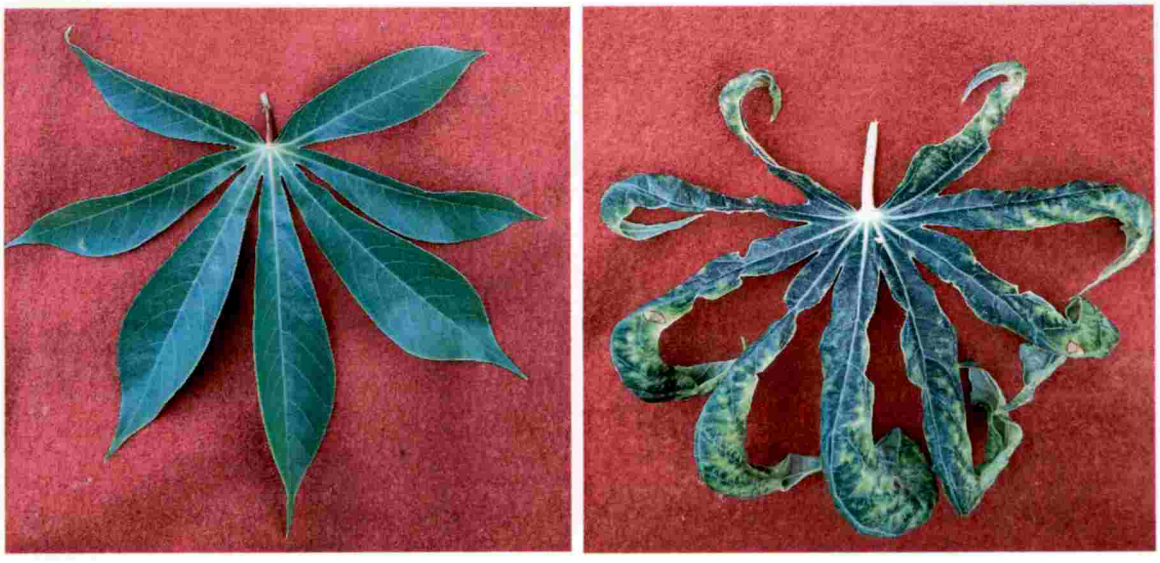
4.3 Checking the quality of DNA

Quality of isolated DNA was checked in 0.8% agarose gel. The dye used for the quality checking was ethidium bromide. The gel was documented in gel documentation system (Syngene). The results showed that the isolate genomic DNA from selected cassava varieties was intact without protein or RNA contamination (Plate 3).

a)

b)

Plate 1. a) CMD resistant cassava plant b) CMD susceptible cassava plant



a) b)

Plate 2. a) Cassava leaf without CMD symptoms b) Leaf with CMD symptoms

4.4 Quantification of DNA

The concentration of the DNA present in each of the isolated sample was determined through the analysis of the absorbance reading at 260nm and 280nm.The quality of DNA in different sample was varied from 500-4000ng/µl. After the quantification process, each sample was diluted into 50ng/µl according to the concentration value obtained from the nano spectrophotometer system and the diluted DNA wad loaded in 0.8% agarose to check the uniformity in concentration.

4.5 PCR amplification using SSR primers.

PCR amplification of 55 DNA samples were done using 20 CMD associated SSR primers under optimum conditions. The PCR products were loaded in 2% Agarose gel with 100 bp ladder. The bands were visualized using gel documentation system (Plate 4).

4.5 Analysis of PCR products using PAGE

PCR product separated by PAGE and documented using computer scanner. All 20 selected SSR primers showed the presence of bands in each sample. Polymorphic bands are scored for the analysis.

4.6 SSR marker analysis

The 20 SSR marker gels were used for data scoring. Monomorphic bands were not selected for the data analysis. Polymorphic bands were used to assign loci for each primer and sored as presence (1) or absence (0) of bands. RME1 primer showed double banding pattern in 700bp size for all the resistant accessions. In the case of susceptible accessions, only single bands were observed (Plate 5). The polymorphic segregating bands were noticed for all the 20 primers used in the study (Plate 6).

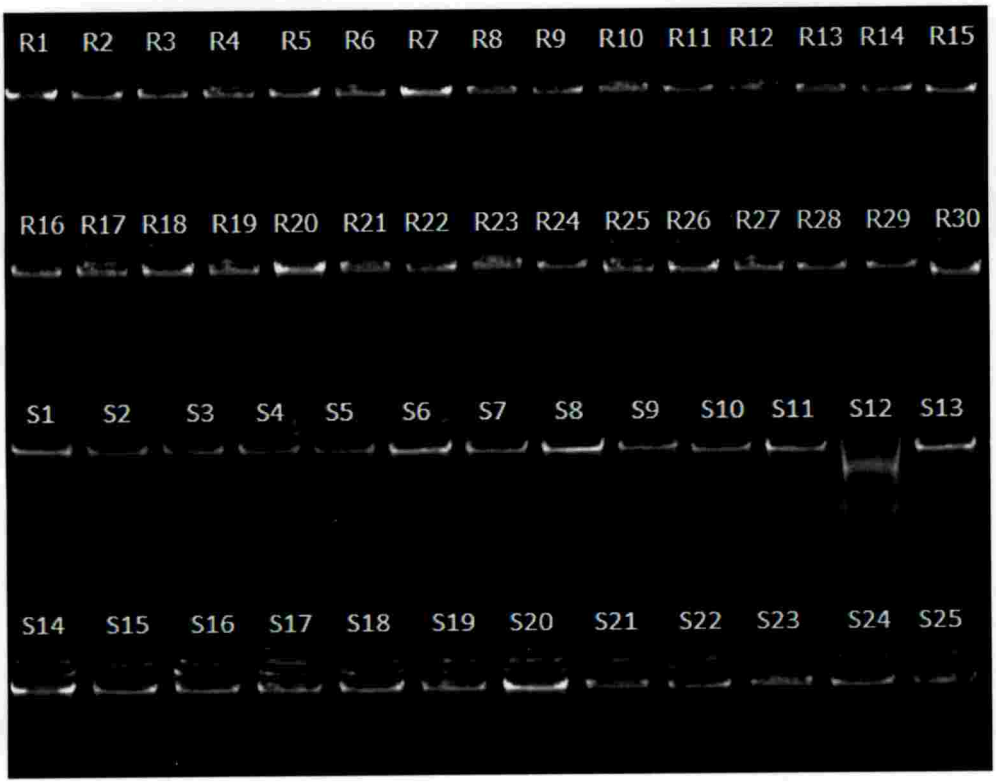


Plate 3. Checking the quality of isolated DNA by agarose gel electrophoresis

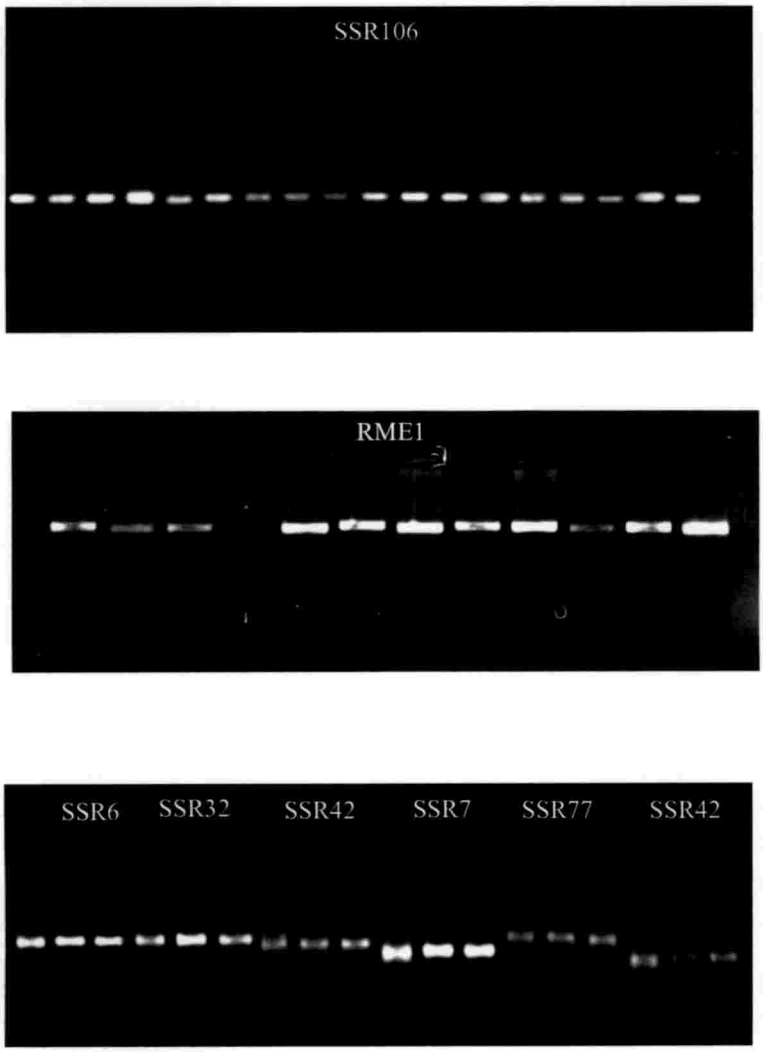


Plate 4. PCR amplification product of SSR primers

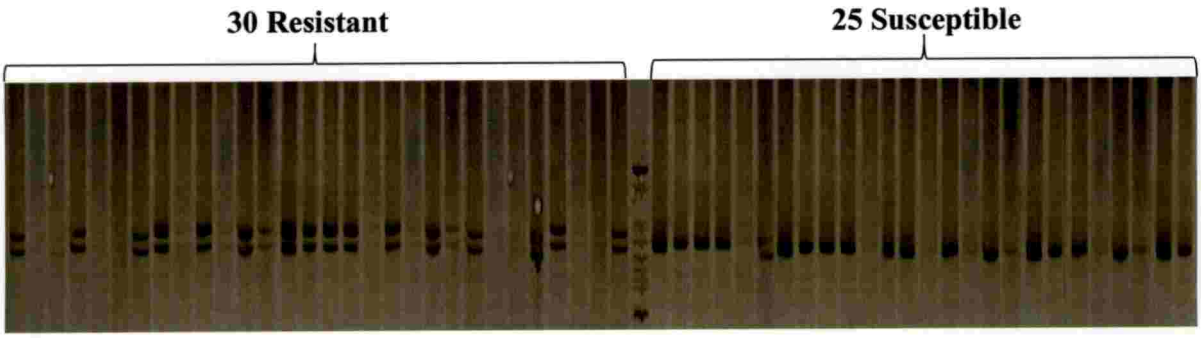


Plate 5. Segregation pattern of RME1 on PAGE

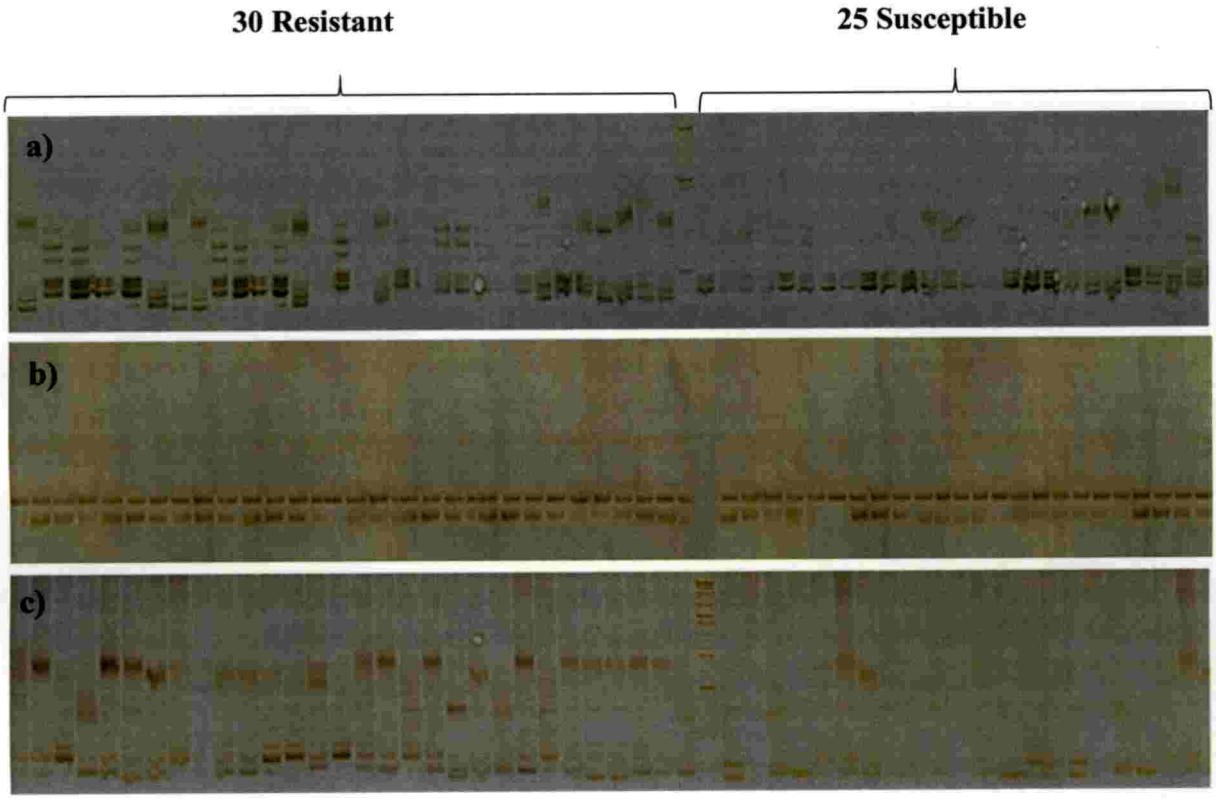


Plate 6. a) Segregation pattern of SSR28 on PAGE b) segregation pattern of SSR21 on PAGE c) segregation pattern of SSR32 on PAGE

53

4.6.1 Genotypic scoring

The 20 SSR marker gels were used for data scoring. Monomorphic bands were not selected for the data analysis. Polymorphic bands were used to assign loci for each primer and scored as presence (1) or absence (0) of individual bands. Based on the banding pattern of the RME1 primer clearly differentiated CMD resistant and susceptible accessions.

4.6.2 Phenotypic scoring

The varieties were scored 1-5 based on expressed phenotypical characters. Highly resistant varieties were coded with 1 and highly susceptible varieties were coded with 5 and rest of them are coded in-between 1-5 (Table 5). Graph was plotted based on the scored data (Fig. 1). Scored phenotypical data revealed that out of 55 cassava varieties 22 varieties were highly resistant and 12 varieties were highly susceptible to CMD.

54

Table 5. Phenotypic data collection

Sl. No	Variety	Score	Sl. No	Variety	Score
1	CMR19	1	29	CMR197	1
2	CMR55	2	30	CMR213	1
3	CMR73	1	31	CI732	4
4	CMR196	1	32	CO2	5
5	CMR218	1	33	H165	4
6	CMR21	2	34	H226	5
7	96/1089A	1	35	11-S-33	5
8	CR 2042	1	36	Sree Rekha	4
9	CR 24	1	37	Sree Vijaya	5
10	TME3	1	38	M4	5
11	8W5	1	39	BR 2	4
12	R15	1	40	BR 5	4
13	CR 43/11	1	41	BR 6	4
14	CMR 195	2	42	BR 10	5
15	CMR13	2	43	GO31	4
16	CMR9	1	44	Kollam local	4
17	CMR65	1	45	849	5
18	CMR138	2	46	C 32	4
19	CMR205	1	47	C 50	4
20	CMR14	2	48	C 77	4
21	CMR18	1	49	Farm Local	5
22	CMR22	1	50	Sree Athulya	4
23	CMR 26	1	51	Sree Pavithra	5
24	CMR29	1	52	Sree Swarna	5
25	CMR30	1	53	Sree Jaya	5
26	CMR51	2	54	CI 848	5
27	CMR55	2	55	TCH2	4
28	CMR109	1			

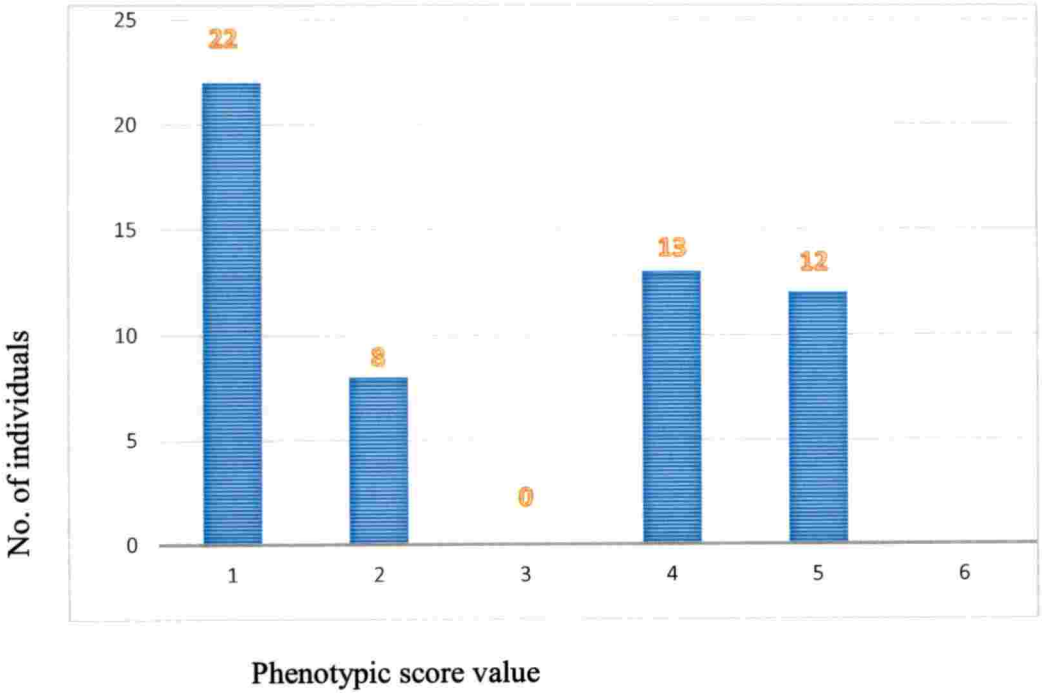


Fig. 1 Graph plotted using phenotypic score data

Highly resistant	-22 accessions
Moderately resistant	-8 accessions
Highly susceptible	-12 accessions
Moderately susceptible	-13 accessions

56

4.7 Molecular data analysis

4.7.1 Analysis of genetic diversity using TASSEL software

Genetic diversity of the 55 cassava accessions was analysed based on the microsatellite allele distribution among them. The genetic closeness of the clones was estimated as identical by dissent (IBS) similarity between every pair of them. The dissimilarity of each genotype is set to zero IBS. IBS is the probability in which any random allele chosen at any random locus is similar for any genotype pair. Therefore, the genotypes that are genetically closer will have a IBS value closer to zero, while for dissimilar genotypes the IBS value will tend towards one.

4. 7.1.1 Dissimilarity matrix

Pair wise distances between 55 genotypes are calculated (Table 6). The dissimilarity coefficients ranged between 0.08 to 0.56. The highest genetic similarity as shown by the lowest dissimilarity index (0.07) was observed between CMD resistant variety CMR14 and CMR21. The genetic dissimilarity of 0.55 was the highest among the genotypes TCH2 and CMR55. Among the 1485 pair-wise distances among 55 genotypes, only 8 pairs (0.5%) showed high genetic similarity having distance less than 0.1, followed by 247 pairs (16.6%) having genetic dissimilarity between 0.1 and 0.2. Almost 80.2% of the genotype pairs had dissimilarity values >0.2 and <0.5. The most diverse genotypes formed 39 pair wise distances accounting for 2.6% of the total pair-wise distances (Figure 2).

Taxa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	0.00																										
2	0.15	0.00																									
3	0.15	0.27	0.00																								
4	0.13	0.21	0.25	0.00																							
5	0.15	0.23	0.27	0.25	0.00																						
6	0.15	0.19	0.19	0.21	0.23	0.00																					
7	0.27	0.35	0.27	0.33	0.27	0.31	0.00																				
8	0.13	0.25	0.17	0.23	0.29	0.17	0.29	0.00																			
9	0.21	0.25	0.25	0.19	0.29	0.29	0.37	0.19	0.00																		
10	0.23	0.19	0.27	0.29	0.31	0.27	0.38	0.25	0.29	0.00																	
11	0.17	0.17	0.25	0.19	0.21	0.21	0.33	0.27	0.31	0.17	0.00																
12	0.17	0.17	0.17	0.23	0.29	0.29	0.33	0.27	0.31	0.13	0.15	0.00															
13	0.17	0.17	0.33	0.19	0.21	0.29	0.37	0.23	0.23	0.25	0.19	0.27	0.00														
14	0.15	0.27	0.15	0.21	0.19	0.19	0.31	0.17	0.21	0.23	0.25	0.25	0.25	0.00													
15	0.12	0.19	0.23	0.17	0.27	0.15	0.35	0.17	0.29	0.27	0.17	0.25	0.25	0.19	0.00												
16	0.12	0.15	0.19	0.21	0.19	0.19	0.31	0.21	0.25	0.23	0.25	0.21	0.21	0.19	0.15	0.00											
17	0.15	0.23	0.19	0.13	0.23	0.19	0.27	0.21	0.29	0.27	0.21	0.21	0.25	0.15	0.15	0.19	0.00										
18	0.27	0.31	0.23	0.29	0.27	0.35	0.35	0.33	0.33	0.27	0.29	0.29	0.33	0.27	0.27	0.19	0.23	0.00									
19	0.12	0.23	0.19	0.17	0.23	0.23	0.31	0.17	0.25	0.23	0.17	0.21	0.21	0.19	0.15	0.19	0.19	0.23	0.00								
20	0.19	0.23	0.19	0.25	0.27	0.08	0.35	0.21	0.33	0.23	0.17	0.29	0.33	0.19	0.15	0.23	0.19	0.27	0.19	0.00							
21	0.13	0.21	0.21	0.23	0.25	0.17	0.33	0.19	0.27	0.33	0.19	0.23	0.27	0.25	0.13	0.17	0.21	0.29	0.13	0.17	0.00						
22	0.10	0.17	0.17	0.15	0.17	0.21	0.25	0.19	0.27	0.21	0.15	0.15	0.23	0.17	0.17	0.17	0.10	0.17	0.10	0.17	0.15	0.00					
23	0.19	0.27	0.19	0.13	0.31	0.19	0.35	0.21	0.25	0.23	0.21	0.25	0.29	0.19	0.19	0.27	0.15	0.23	0.19	0.12	0.25	0.17	0.00				
24	0.19	0.23	0.23	0.21	0.31	0.31	0.35	0.21	0.29	0.23	0.21	0.17	0.33	0.27	0.19	0.27	0.23	0.23	0.15	0.23	0.17	0.13	0.15	0.00			
25	0.23	0.23	0.23	0.25	0.35	0.23	0.35	0.25	0.29	0.31	0.21	0.29	0.29	0.27	0.19	0.23	0.23	0.27	0.15	0.19	0.13	0.21	0.23	0.19	0.00		
26	0.17	0.21	0.21	0.19	0.25	0.25	0.37	0.27	0.23	0.21	0.19	0.19	0.31	0.13	0.17	0.25	0.17	0.21	0.17	0.17	0.23	0.12	0.13	0.13	0.21	0.00	
27	0.19	0.27	0.23	0.17	0.35	0.23	0.35	0.21	0.29	0.31	0.25	0.25	0.33	0.23	0.12	0.19	0.19	0.27	0.23	0.23	0.25	0.25	0.19	0.23	0.23	0.21	0.00

Taxa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
28	0.13	0.17	0.21	0.15	0.29	0.17	0.37	0.19	0.23	0.17	0.19	0.19	0.27	0.21	0.10	0.17	0.13	0.21	0.17	0.17	0.19	0.15	0.13	0.17	0.21	0.12	0.13	0.00
29	0.15	0.23	0.19	0.17	0.27	0.19	0.35	0.17	0.25	0.23	0.21	0.21	0.25	0.19	0.15	0.19	0.15	0.23	0.19	0.15	0.17	0.13	0.12	0.19	0.23	0.13	0.19	0.10
30	0.15	0.15	0.23	0.21	0.31	0.19	0.35	0.17	0.25	0.27	0.21	0.21	0.25	0.27	0.15	0.15	0.27	0.31	0.19	0.19	0.13	0.17	0.23	0.15	0.15	0.21	0.19	0.21
31	0.31	0.35	0.35	0.37	0.46	0.42	0.42	0.40	0.40	0.38	0.40	0.29	0.44	0.42	0.35	0.31	0.35	0.38	0.35	0.42	0.33	0.29	0.42	0.38	0.42	0.37	0.35	0.33
32	0.33	0.40	0.29	0.38	0.44	0.40	0.40	0.31	0.38	0.37	0.38	0.31	0.42	0.40	0.37	0.37	0.40	0.44	0.40	0.40	0.42	0.35	0.40	0.40	0.52	0.42	0.37	0.38
33	0.27	0.31	0.27	0.37	0.38	0.38	0.31	0.33	0.40	0.31	0.33	0.21	0.40	0.35	0.31	0.27	0.27	0.35	0.35	0.38	0.29	0.25	0.38	0.31	0.38	0.33	0.35	0.29
34	0.29	0.40	0.29	0.35	0.40	0.40	0.40	0.38	0.38	0.40	0.38	0.31	0.46	0.37	0.33	0.29	0.33	0.37	0.33	0.37	0.31	0.27	0.37	0.37	0.40	0.35	0.33	0.35
35	0.35	0.46	0.23	0.37	0.42	0.42	0.38	0.40	0.37	0.42	0.40	0.33	0.48	0.38	0.42	0.38	0.38	0.35	0.31	0.38	0.37	0.29	0.38	0.38	0.38	0.37	0.42	0.40
36	0.37	0.40	0.33	0.42	0.48	0.40	0.44	0.31	0.38	0.40	0.42	0.38	0.46	0.40	0.33	0.37	0.37	0.37	0.37	0.37	0.31	0.31	0.40	0.33	0.40	0.38	0.37	0.35
37	0.33	0.44	0.29	0.35	0.44	0.40	0.37	0.38	0.38	0.44	0.42	0.35	0.46	0.44	0.37	0.33	0.37	0.37	0.29	0.40	0.31	0.31	0.40	0.40	0.37	0.42	0.37	0.35
38	0.40	0.44	0.37	0.42	0.52	0.48	0.48	0.42	0.38	0.40	0.50	0.35	0.50	0.44	0.44	0.37	0.37	0.48	0.40	0.48	0.38	0.38	0.48	0.40	0.44	0.46	0.44	0.42
39	0.37	0.48	0.37	0.42	0.48	0.48	0.44	0.46	0.50	0.48	0.50	0.35	0.54	0.44	0.37	0.40	0.37	0.40	0.40	0.44	0.38	0.35	0.44	0.37	0.48	0.38	0.33	0.38
40	0.31	0.42	0.27	0.33	0.38	0.42	0.31	0.37	0.37	0.42	0.37	0.29	0.44	0.38	0.38	0.35	0.35	0.38	0.27	0.38	0.29	0.25	0.38	0.31	0.35	0.37	0.38	0.40
41	0.40	0.48	0.33	0.42	0.52	0.44	0.40	0.42	0.38	0.48	0.42	0.42	0.50	0.48	0.37	0.40	0.44	0.37	0.37	0.44	0.35	0.42	0.48	0.44	0.33	0.46	0.37	0.38
42	0.38	0.50	0.35	0.37	0.46	0.46	0.42	0.44	0.40	0.42	0.40	0.40	0.48	0.42	0.38	0.38	0.35	0.31	0.38	0.42	0.44	0.37	0.38	0.46	0.42	0.40	0.35	0.37
43	0.33	0.44	0.29	0.31	0.40	0.37	0.40	0.38	0.38	0.37	0.35	0.35	0.46	0.37	0.33	0.37	0.33	0.37	0.33	0.33	0.42	0.31	0.33	0.40	0.44	0.35	0.33	0.31
44	0.37	0.52	0.37	0.42	0.44	0.52	0.40	0.50	0.50	0.48	0.42	0.38	0.50	0.48	0.44	0.44	0.44	0.40	0.33	0.44	0.42	0.35	0.44	0.40	0.44	0.42	0.40	0.46
45	0.40	0.44	0.44	0.38	0.52	0.44	0.44	0.42	0.50	0.44	0.38	0.38	0.50	0.48	0.29	0.40	0.37	0.40	0.37	0.40	0.35	0.35	0.44	0.37	0.37	0.38	0.33	0.35
46	0.40	0.44	0.33	0.42	0.48	0.37	0.37	0.42	0.46	0.44	0.42	0.38	0.50	0.44	0.40	0.40	0.40	0.40	0.37	0.33	0.35	0.35	0.40	0.40	0.37	0.42	0.44	0.42
47	0.33	0.44	0.25	0.35	0.40	0.40	0.40	0.31	0.35	0.33	0.35	0.31	0.42	0.37	0.33	0.37	0.37	0.37	0.25	0.37	0.38	0.31	0.37	0.33	0.40	0.38	0.33	0.35
48	0.40	0.52	0.33	0.42	0.44	0.40	0.37	0.42	0.46	0.40	0.38	0.35	0.54	0.44	0.40	0.44	0.40	0.40	0.33	0.37	0.38	0.35	0.37	0.37	0.40	0.38	0.44	0.38
49	0.37	0.52	0.33	0.38	0.44	0.44	0.44	0.46	0.46	0.48	0.46	0.38	0.50	0.44	0.44	0.44	0.40	0.48	0.40	0.44	0.46	0.38	0.44	0.48	0.52	0.46	0.44	0.42
50	0.38	0.46	0.31	0.44	0.50	0.42	0.46	0.37	0.40	0.42	0.44	0.40	0.52	0.46	0.38	0.38	0.46	0.38	0.31	0.38	0.37	0.37	0.42	0.35	0.31	0.40	0.42	0.40
51	0.42	0.50	0.46	0.33	0.46	0.46	0.42	0.44	0.44	0.46	0.44	0.40	0.48	0.46	0.42	0.46	0.46	0.54	0.42	0.54	0.44	0.44	0.46	0.38	0.46	0.48	0.42	0.44
52	0.48	0.52	0.40	0.42	0.52	0.52	0.44	0.50	0.42	0.40	0.46	0.46	0.54	0.48	0.52	0.44	0.52	0.33	0.44	0.44	0.50	0.42	0.40	0.40	0.40	0.42	0.44	0.46
53	0.42	0.46	0.42	0.40	0.54	0.46	0.54	0.44	0.44	0.42	0.44	0.44	0.52	0.54	0.42	0.42	0.50	0.46	0.38	0.42	0.48	0.40	0.42	0.42	0.42	0.44	0.38	0.40
54	0.44	0.52	0.44	0.38	0.52	0.44	0.37	0.46	0.42	0.48	0.46	0.46	0.54	0.52	0.44	0.52	0.48	0.52	0.40	0.48	0.46	0.46	0.44	0.44	0.44	0.46	0.40	0.38
55	0.48	0.56	0.44	0.46	0.56	0.52	0.48	0.50	0.50	0.52	0.50	0.50	0.54	0.56	0.48	0.48	0.56	0.40	0.40	0.48	0.42	0.46	0.48	0.44	0.44	0.54	0.48	0.50

Taxa	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	34	55
29	0.00																										
30	0.19	0.00																									
31	0.35	0.35	0.00																								
32	0.33	0.37	0.17	0.00																							
33	0.31	0.31	0.12	0.17	0.00																						
34	0.29	0.33	0.10	0.15	0.17	0.00																					
35	0.35	0.38	0.15	0.17	0.23	0.13	0.00																				
36	0.33	0.29	0.17	0.23	0.17	0.19	0.21	0.00																			
37	0.33	0.37	0.17	0.27	0.29	0.15	0.13	0.27	0.00																		
38	0.44	0.40	0.21	0.27	0.21	0.15	0.25	0.23	0.27	0.00																	
39	0.37	0.40	0.13	0.23	0.21	0.15	0.17	0.19	0.23	0.23	0.00																
40	0.35	0.31	0.15	0.17	0.15	0.13	0.08	0.21	0.17	0.21	0.17	0.00															
41	0.44	0.40	0.25	0.27	0.29	0.23	0.17	0.27	0.19	0.27	0.27	0.21	0.00														
42	0.38	0.46	0.27	0.29	0.27	0.25	0.19	0.29	0.29	0.29	0.29	0.27	0.21	0.00													
43	0.33	0.40	0.25	0.15	0.25	0.15	0.17	0.31	0.23	0.27	0.27	0.17	0.23	0.25	0.00												
44	0.44	0.44	0.25	0.27	0.33	0.19	0.17	0.35	0.23	0.31	0.19	0.17	0.19	0.29	0.19	0.00											
45	0.40	0.33	0.21	0.31	0.21	0.27	0.29	0.19	0.31	0.35	0.23	0.21	0.27	0.33	0.23	0.31	0.00										
46	0.37	0.33	0.29	0.23	0.25	0.23	0.17	0.27	0.27	0.27	0.23	0.13	0.19	0.33	0.19	0.23	0.27	0.00									
47	0.37	0.37	0.25	0.15	0.25	0.19	0.13	0.23	0.19	0.23	0.23	0.13	0.23	0.25	0.12	0.23	0.27	0.23	0.00								
48	0.37	0.40	0.33	0.27	0.29	0.27	0.21	0.35	0.27	0.35	0.31	0.17	0.23	0.37	0.19	0.23	0.27	0.15	0.23	0.00							
49	0.40	0.48	0.25	0.23	0.29	0.23	0.13	0.31	0.23	0.27	0.19	0.17	0.27	0.25	0.15	0.19	0.31	0.23	0.19	0.23	0.00						
50	0.42	0.31	0.31	0.29	0.31	0.29	0.19	0.33	0.25	0.29	0.33	0.19	0.25	0.31	0.25	0.29	0.29	0.21	0.17	0.25	0.29	0.00					
51	0.46	0.42	0.35	0.33	0.38	0.33	0.31	0.40	0.37	0.37	0.33	0.27	0.33	0.42	0.33	0.37	0.33	0.33	0.33	0.29	0.33	0.38	0.00				
52	0.44	0.40	0.33	0.27	0.33	0.27	0.21	0.35	0.35	0.35	0.35	0.21	0.23	0.25	0.23	0.27	0.35	0.19	0.27	0.27	0.35	0.25	0.29	0.00			
53	0.42	0.35	0.27	0.25	0.38	0.21	0.23	0.33	0.25	0.29	0.33	0.27	0.29	0.27	0.21	0.25	0.33	0.29	0.21	0.33	0.29	0.15	0.31	0.21	0.00		
54	0.44	0.48	0.33	0.35	0.37	0.35	0.25	0.38	0.27	0.35	0.31	0.25	0.23	0.29	0.27	0.31	0.35	0.27	0.27	0.23	0.19	0.29	0.25	0.31	0.29	0.00	
55	0.52	0.44	0.29	0.31	0.37	0.27	0.21	0.31	0.27	0.31	0.31	0.21	0.15	0.29	0.27	0.19	0.31	0.19	0.27	0.27	0.27	0.21	0.29	0.19	0.21	0.23	0.00

60

4.7.1.2 Clustering of genotypes

Based on the genetic distances calculated as Euclidean distances, a cladogram was constructed using neighbour joining method to graphically visualize the relationship between selected cassava accessions. Empirical grouping of genotypes from the centre point, separated 55 accessions into at least four clusters (cluster I and cluster IV). Closely related individuals were located closer together than those who were distantly related. It was observed that the grouping separated CMD resistant and susceptible varieties into separate groups. Cluster I and Cluster II comprised of 5 accessions each and included most closely related genotypes. Cluster III comprised of 9 accessions. The remaining 36 genotypes fell into cluster IV. Among the clusters, cluster IV had mostly intermediate (admixture) genotypes (Fig 3).

4.7.2 Analysis of population structure

Population structure of the cassava accessions were analysed using a model based Bayesian approach in which the Hardy-Weinberg equilibrium is tested within the assumed sub-populations to identify optimum sub-populations as far as possible. The algorithm implemented in the software STRUCTURE, allows estimation of log probability of the data for each assumed sub-population number (K) which can be further used to identify the optimum sub-population structure. The analysis assumes a co-ancestry model for all the genotypes in the populations, with correlated allele frequencies to establish Hardy-Weinberg equilibrium. The simulated data for all the assumed K was then used to estimate an adhoc statistic ΔK (delta K), which can help in identifying the optimum K easily. The ΔK is the estimate of rate of change in log probability of the data between successive K values, such that the optimum K has the highest values of ΔK (Evanno *et al.*, 2005). The estimated ΔK values for the cassava samples, indicated the presence of at least eight sub-populations and admixtures in the tested population (Table 7). A graph was plotted between' Delta K' and Number

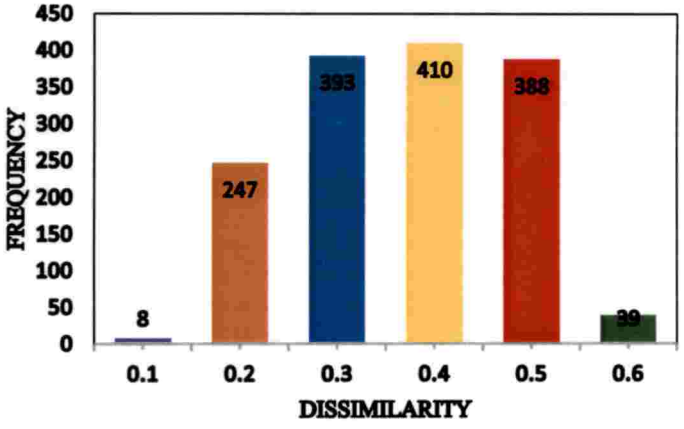


Fig. 2 A graph was plotted using dissimilarity coefficient value

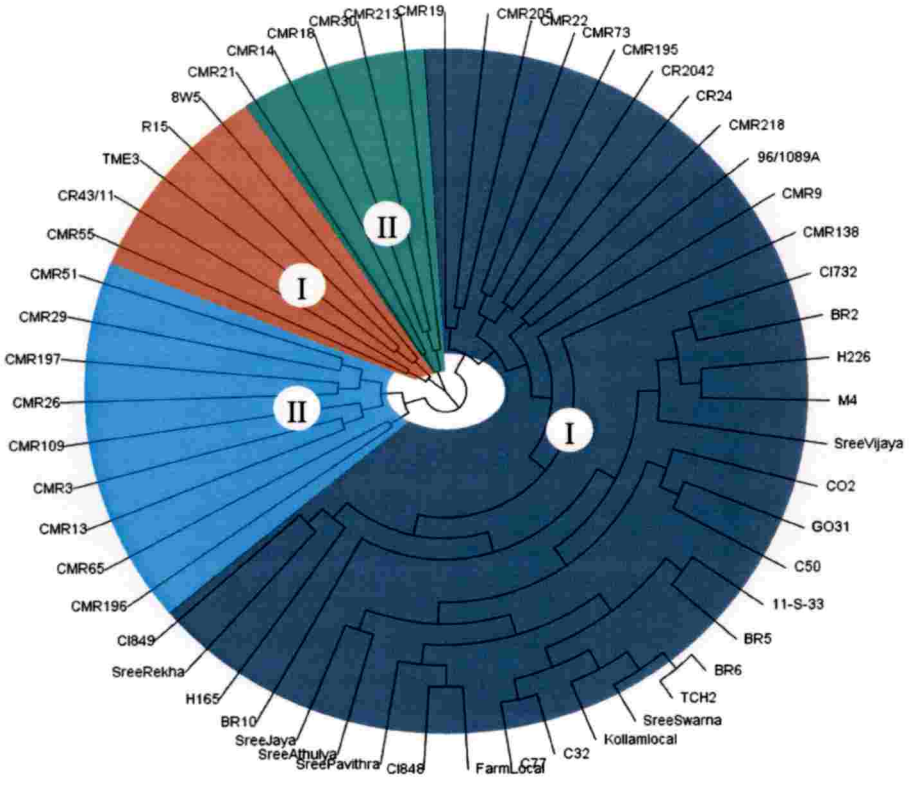


Fig 3. Cladogram of genotypes showing distinct grouping based on Euclidean distances

62

of sub populations (k) (Fig. 4) showed a peak value of 628.3 for K=8, indicating the total sub-populations in the sample.

Table 7. Evanno's table output

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Δ K
1	3	-1133.5	1.0	—	—	—
2	3	-1056.8	0.6	76.67	64.30	100.01
3	3	-1044.5	14.6	12.37	23.50	1.61
4	3	-1055.6	9.6	-11.13	3.47	0.36
5	3	-1063.3	9.0	-7.67	3.33	0.37
6	3	-1074.3	16.6	-11.00	537.00	32.45
7	3	-1622.3	903.5	-548.00	1101.17	1.22
8	**3**	**-1069.1**	**0.9**	**553.17**	**558.43**	**628.29**
9	3	-1074.4	12.2	-5.27	0.20	0.02
10	3	-1079.4	10.2	-5.07	—	—

LnP(K), log probability of K; Stdev LnP(K), standard deviation of LnP(K); Ln'(K), mean deviation of LnP(K) for adjacent K; |Ln" (K)|, standard deviation of the rate of change; ΔK, ratio between Ln'(K) and Stdev LnP(K).Mean simulation parameters: Estimated Ln Prob of data = -1068.4; Mean Ln likelihood = -1002.7; Variance of Ln likelihood = 131.4; Mean value of alpha = 0.0510

4.7.2.1 Sub population identification

The eight sub-populations identified among the 55 cassava genotypes, showed almost equal distribution of members among themselves. Population membership varied between 10.8% (POP3) and 20.6% in POP4 (Table 8). Expected heterozygosity, also known as gene diversity, almost showed similar pattern across the sub-populations, and ranged narrowly between 0.27 (POP4) to 0.34 (POP5). Gene diversity indicated the probability of getting a heterozygote at any given locus in the population, therefore, the studied population of 55 cassava genotypes has an average heterozygosity of 0.33, indicating that there was 33% probability for identifying any locus in the population to be heterozygous. The population

differentiation estimated using Fst, indicated that POP2 and POP4 as more diverse accounting for 33.8% and 41.3% of the total variation in the population respectively. These were followed by POP8 which accounted 16.6% variation. The least diverse sub-population was POP5 with an Fst value of 0.006.

Table 8. Overall proportion of membership

Parameters	Inferred clusters							
	POP1	POP2	POP3	POP4	POP5	POP6	POP7	POP8
Membership (%)	11.2	12.8	10.8	20.6	11.0	10.9	11.0	11.7
Expected heterozygosity	0.342	0.296	0.343	0.268	0.344	0.342	0.342	0.341
Mean Fst	0.020	0.338	0.014	0.413	0.006	0.024	0.019	0.166

4.7.2.2 Allele frequency divergence

Allele frequency divergence (Table 9) or the inter population genetic diversity indicated that POP2 and POP4 distinctly varied from rest of the population with POP4 showing higher distance from the rest followed by POP2. The average distance between POP2 and POP4 was the maximum among the sub-populations (0.098).

Table 9. Allele frequency divergence among the 8 clusters

K	Inferred sub-populations						
	POP1	POP2	POP3	POP4	POP5	POP6	POP7
POP2	0.041	-					
POP3	0.000	0.040	-				
POP4	0.092	0.098	0.091	-			
POP5	0.000	0.040	0.000	0.090	-		
POP6	0.000	0.040	0.000	0.092	0.000	-	
POP7	0.000	0.041	0.000	0.091	0.000	0.000	-
POP8	0.008	0.012	0.008	0.085	0.008	0.008	0.008

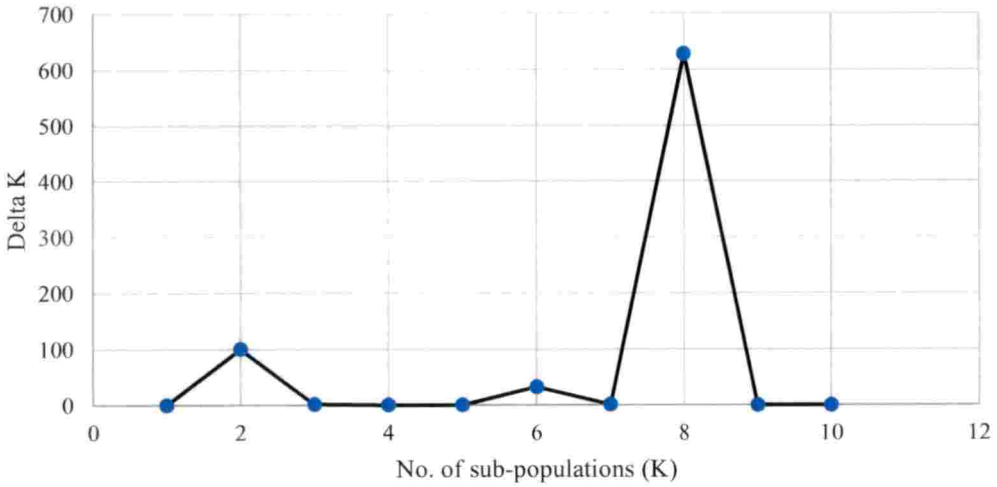


Fig 4. The graphical plot of ΔK values showing the peak at K=8

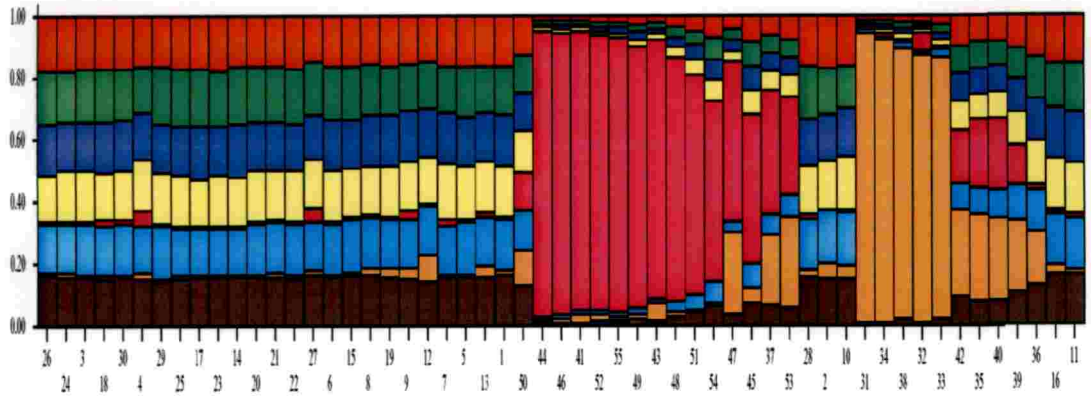


Fig 5. Bar plot showing the distribution of genotypes in different sub-populations

65

4.7.2.3 The inferred ancestry coefficients

The inferred sub-populations showed highly admixed pattern of allele distribution except two sub-populations (Fig 5). The indistinct pattern of allele conservation among the six sub-populations which varied only is some specific loci indicated that they may be breeding lines that were closely related but distinct for certain allelic combinations. Nevertheless, two larger sub-populations as indicated earlier as POP2 and POP4 were largely distinctly different from the rest of the genotypes in the population. Q values of the genotypes indicated that most diverse genotypes included Kollam Local, BR6, C32, Farm Local, Sree Swarna, TCH2, GO31, C77, Sree Pavithra and CI848 in one cluster and CI732, CO2, H165, H226 and M4 in the other. These genotypes were unique in their allelic pattern for most of the loci tested in the germplasm. The inferred ancestry values (Q values) of each of the genotypes are given (Table 10).

Table 10. The inferred ancestry coefficients of the 55-cassava accession (Q matrix) for eight sub-populations

No.	Genotype	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
1	CMR19	0.162	0.156	0.163	0.165	0.011	0.162	0.016	0.165
2	CMR55	0.167	0.149	0.149	0.157	0.006	0.168	0.047	0.156
3	CMR73	0.171	0.167	0.161	0.159	0.011	0.161	0.012	0.157
4	CMR196	0.164	0.145	0.155	0.16	0.049	0.151	0.019	0.158
5	CMR218	0.164	0.162	0.158	0.169	0.01	0.168	0.006	0.164
6	CMR21	0.162	0.17	0.161	0.165	0.012	0.163	0.008	0.16
7	96/1089A	0.163	0.149	0.163	0.176	0.023	0.156	0.011	0.159
8	CR_2042	0.157	0.16	0.168	0.152	0.009	0.162	0.026	0.166
9	CR_24	0.152	0.155	0.162	0.156	0.029	0.154	0.041	0.151
10	TME3	0.16	0.136	0.159	0.167	0.011	0.168	0.046	0.152
11	8W5	0.151	0.161	0.163	0.166	0.009	0.165	0.016	0.168
12	R15	0.15	0.145	0.157	0.152	0.009	0.152	0.086	0.149
13	CR_43/11	0.161	0.149	0.158	0.166	0.013	0.159	0.032	0.163
14	CMR_195	0.164	0.179	0.177	0.156	0.008	0.149	0.008	0.158
15	CMR13	0.161	0.168	0.163	0.154	0.01	0.167	0.011	0.165
16	CMR9	0.154	0.146	0.16	0.165	0.016	0.164	0.025	0.17
17	CMR65	0.169	0.183	0.175	0.15	0.006	0.146	0.013	0.157

66

No.	Genotype	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
18	CMR138	0.17	0.168	0.162	0.155	0.022	0.156	0.015	0.151
19	CMR205	0.16	0.159	0.164	0.16	0.011	0.157	0.029	0.159
20	CMR14	0.16	0.177	0.16	0.164	0.007	0.167	0.006	0.159
21	CMR18	0.163	0.174	0.162	0.155	0.01	0.16	0.015	0.161
22	CMR22	0.171	0.172	0.157	0.16	0.007	0.164	0.017	0.153
23	CMR26	0.172	0.18	0.159	0.161	0.009	0.152	0.006	0.162
24	CMR29	0.172	0.17	0.157	0.161	0.007	0.16	0.012	0.16
25	CMR30	0.17	0.187	0.155	0.16	0.008	0.155	0.008	0.157
26	CMR51	0.175	0.174	0.159	0.155	0.006	0.157	0.008	0.165
27	CMR3	0.147	0.171	0.145	0.158	0.043	0.152	0.019	0.165
28	CMR109	0.162	0.168	0.153	0.156	0.012	0.169	0.019	0.162
29	CMR197	0.159	0.189	0.155	0.162	0.01	0.17	0.008	0.148
30	CMR213	0.168	0.166	0.16	0.163	0.011	0.162	0.011	0.158
31	CI732	0.009	0.009	0.008	0.009	0.007	0.009	0.941	0.009
32	CO2	0.015	0.013	0.014	0.013	0.056	0.014	0.864	0.013
33	H165	0.026	0.025	0.024	0.024	0.007	0.024	0.846	0.024
34	H226	0.013	0.014	0.012	0.014	0.013	0.011	0.912	0.012
35	11-S-33	0.084	0.086	0.082	0.082	0.22	0.084	0.281	0.08
36	Sree_Rekha	0.135	0.134	0.135	0.142	0.015	0.136	0.169	0.134
37	Sree_Vijaya	0.06	0.058	0.059	0.062	0.397	0.064	0.234	0.065
38	M4	0.019	0.019	0.018	0.018	0.01	0.019	0.875	0.021
39	BR_2	0.101	0.103	0.105	0.109	0.127	0.118	0.224	0.113
40	BR_5	0.083	0.076	0.088	0.083	0.234	0.09	0.263	0.083
41	BR_6	0.011	0.013	0.011	0.011	0.903	0.011	0.027	0.011
42	BR_10	0.093	0.091	0.093	0.088	0.171	0.087	0.284	0.092
43	GO31	0.018	0.019	0.018	0.017	0.837	0.017	0.056	0.018
44	Kollam local	0.011	0.011	0.011	0.011	0.925	0.011	0.009	0.011
45	CI_849	0.08	0.08	0.078	0.082	0.479	0.077	0.046	0.079
46	C_32	0.014	0.013	0.014	0.014	0.909	0.013	0.01	0.013
47	C_50	0.037	0.035	0.039	0.035	0.514	0.034	0.266	0.04
48	C_77	0.034	0.032	0.032	0.034	0.789	0.031	0.015	0.033
49	Farm Local	0.023	0.025	0.025	0.023	0.84	0.026	0.015	0.024
50	Sree_Athulya	0.125	0.12	0.124	0.133	0.121	0.13	0.118	0.13
51	Sree_Pavithra	0.047	0.044	0.045	0.05	0.713	0.044	0.013	0.045
52	Sree_Swarna	0.017	0.014	0.016	0.015	0.883	0.015	0.024	0.016
53	Sree_Jaya	0.072	0.058	0.062	0.067	0.319	0.068	0.291	0.062
54	CI_848	0.068	0.069	0.065	0.075	0.576	0.069	0.013	0.064
55	TCH2	0.016	0.018	0.017	0.018	0.882	0.019	0.012	0.019

67

4.7.2.4 Identification of markers associated with resistance to CMD

To identify if any relation exists between the markers used to survey the germplasm set, with CMD resistance, a mixed linear model association was carried out. The genotype data was filtered out of rare and infrequent alleles to make the data fit Hardy-Weinberg equilibrium. The Q matrix was used as the covariate in the model to account for the population structure in the association analysis. A kinship matrix derived from the IBS similarity accounted by the marker alleles across the members of the population, was also used as an additional co-variate in the model. The mixed linear model (MLM) was run with a moderate cut-off threshold of 0.05 for claiming the valid marker-trait associations. The MLM is based on strong genetic principles and hence can predict reliable associations. The use of population structure (Q) and the kinship (K) increases the statistical power of truly associated markers. The model contains genotypic data and the Q coefficients as fixed effects, and genetic data and the QTL effects are used as random in the model. The model used restricted maximum likelihood (REML) and efficient mixed-model association (EMMA) to improve statistical power of detection of associations.

The association analysis resulted in identification of two markers, having associated with CMD resistance in the cassava population studied (Table 11). The markers were SSR106.4 and SSR36, which indicated one of the alleles of these markers showed association with CMD resistance pattern among the genotypes.

Table 11. Significant marker trait association detected by mixed linear model

Trait	Marker	Chromosome	Frequency	Probability	Marker R^2
CMD	SSR106.4	1	6.60	0.01	13.19
CMD	SSR36.1	1	4.29	0.04	8.57

The SSR106 showed association with an R^2 value of 13.19%, with its allele explaining 13.19% of the CMD resistance, followed by SSR36 allele with R^2 of 8.57%.

DISCUSSION

69

Chapter 5

DISCUSSION

Cassava is the most important perennial tropical woody shrub mainly propagating from stem cuttings. Among the genus *Manihot*, cassava is the only plant species which is used as a staple food (Nwokoro *et al.*, 2002). Cassava plants produce large rate of soluble carbohydrate per unit area by converting solar energy. Cassava mosaic disease is one among the most dangerous cassava disease causing by cassava mosaic gemini virus coming under the family geminiviridae. It belongs to the *Begomovirus* genus (Fauquet and Stately, 2003). The major idea of resistance is to development CMD resistant varieties. CMD resistant trait from *Manihot glaziovii* was integrated into the local varieties to confer disease resistance. (Nicholas, 1947).

Association mapping study was conducted to identify the molecular markers which are closely linked with cassava mosaic disease resistance using SSR markers. 35 CMD resistant and 25 CMD susceptible cassava varieties were collected from the field were analyzed with 20 selected SSR markers. Genomic DNA was isolated from the selected cassava plants and PCR amplification was done using these SSR markers. The PCR amplification product was separated by polyacrylamide gel electrophoresis (PAGE). These markers include SSR40, SSR28, NS158, RME1, SSR6, SSR7, SSR21, SSR42, SSR77, SSR324, SSR28, SSR106, SSR235, SSR44, NS136, SSR9, SSR235, SSR198, SSR39 and SSR36. Out of twenty SSR markers, most of the markers were CMD associated markers based on earlier CMD mapping studies.

Based on the banding pattern of the RME1 primer clearly differentiated CMD resistant and susceptible accessions. RME1 primer showed double banding pattern in 700bp size for all the resistant accessions. In the case of susceptible accessions, only single bands were observed. RME1 and NS158 were two SSR markers which are tightly associated with CMD was identified by Tomkins and Fregene in 2004. A

collaborative work was conducted by CIAT and IITA in Ibadan, Nigeria to develop several markers for the source of CMD resistant, revealed to be controlled by single dominant gene *CMD2* (Akano et al., 2002). Five CMD association markers were developed. The closest being NS158 and RME1.They were at the distance of seven and four cM respectively.

According to the statistical data SSR106 and SSR36 were highly associated with CMD resistance. These two markers showed higher association frequency. SSR106 and SSR36 markers were already reported CMD association markers by different studies. The SSR106 showed association with an R^2 value of 13.19%, with its allele explaining 13.19% of the CMD resistance, followed by SSR36 allele with R^2 of 8.57%.

Mohan *et al.*, (2013) identified four SSR markers (SSRY28, SSRY235, SSRY44 and NS136.) associated with CMD resistance using CO2 x MNga-1 population by single marker analysis (SMA).

Genome wide association study was conducted by Wolfe *et al.* (2016) to study the genetic nature of CMD resistance in cassava. In CIAT, Scientists used 6125 cassava genotypes and 42,113 single nucleotide polymorphism (SNP) markers. They identified a specific sequence on chromosome 8 that provides 30 to 66 per cent of genetic resistance against CMD in cassava cultivars of Africa.

TASSEL and STRUCTURE were the two-softwares used for the association analysis and population structure prediction. TASSEL software was used to identify the markers which are closely associated with CMD resistance based on LD value. Dissimilarity matrix was generated based on the genotypic values among the selected cassava lines and multivariate evolutionary relationship among the collected cassava accessions were obtained from the cladogram. The linkage disequilibrium value was obtained using software. The LD value helps to identify the degree of association between marker and trait. Genome-wide LD patterns can help to identify the history

of changes in populations (Slatkin, 2008). However, the power of AM can be strongly decreased because of population structure (Balding, 2006).

Dissimilarity matrix value was obtained by the pairwise comparison of each cassava species. The similarity co-efficient values were ranges from 0.08 to 0.56. Lowest dissimilarity value was obtained by the pair wise comparison of two CMD resistant cassava varieties CMR21 and CMR 14 and the highest value was obtained by the pair wise comparison of CMD susceptible variety TCH2 and resistant variety CMR55. Shahriar *et al.*, (2014) conducted a work to analyses the genetic relationship among the 35 rice varieties. The dissimilarity matrix constructed was used to identify the level of dissimilarities among the studied genotypes. Pair-wise estimates of dissimilarity ranged from 0.00 to 1.00 and the average dissimilarity among all 34 cultivars was 0.36.

The resulted cladogram clearly distinguished the CMD resistant and susceptible varieties into two distinct clusters. The resistant varieties are then divided into three sub clusters. The external branches of cladogram were generated in this analysis were longer than the internal branches, indicating that within group variability was higher than that between groups. Favoretto *et al.*, (2011) done a cluster analysis to analyze the genetic diversity among the collected cassava accessions. 13 and 10 clusters were formed for SSR markers and morphological traits respectively. Clusters (II and XII) comprised of cassava accessions with yield contributing traits and accessions with CMD susceptibility and low yield cassava varieties were assembled together into cluster IX.

STRUCTURE software was used to identify the sub-populations and admixtures in the collected cassava accessions. Input file for structure analysis was taken from the scored genotypic data obtained from 55 cassava accessions. Analysis was carried out in ten replications with 100,000 burns. The total number of sub-populations (K) was calculated according to the posterior probability values (LnP(D)

and ΔK). The values of LnP(D) and ΔK increased continuously in all ten replicates and the highest ΔK of 628.29 was reached when K was 8 with mean LnP(K) value of -1069.1. Thus, there were eight sub-populations in the 55 accessions used for the study. The similar Association Mapping wok was conducted by Swamy *et al.*, (2017) in rice for yield and yield related traits. They were done a structural analysis of 75 rice genotypes using 119 markers and used 10 replicates. But the length of burin period used was 500,000 burns. They observed only 3 sub- populations in the selected 75 rice varieties.

Similar population structure analysis work was conducted in cassava using SSR molecular markers by Costa *et al.* (2013) the results obtained from the Structure were subsequently analyzed with the software Structure Harvester. Only two distinct groups of cassava varieties were observed. For this analysis, a burn-in of 100 000 Monte Carlo Markov Chain (MCMC) for each K, was applied. The K varies from 2 to15. A total of 14 clustering simulations were conducted for each K value, and analyzed in relation to probabilities P(K) of the individuals to belong to a K-th group admixture model (Pritchard *et al.*, 2000, Evanno et al., 2005 and Kwak and Gepts, 2009). The threshold probability for a certain individual to belong to a K-th group was determined after the analysis of ΔK parameters with the software Structure Harvester. Q matrix of eight subpopulation was constructed.

The population structure studies on 96 genotypes of soybean showed the occurrence of seven distinct subpopulations. Genome wide association study was done with 121 SSR markers. From the 121 SSR markers 66 Polymorphic SSR markers were selected for the population structure and kindship studies. The *r*2 value between marker pairs ranged from 0 to 0.35, while *D* values ranged from 0 to as high as 1.00 (Kumar et al., 2014). K value selected was 2 to 15 and, for each K package was run with 3 independent runs. Number of iterations and Length of burn-in period were set 1, 50,000. 7 sub populations were identified in 96 soybean genotypes.

Based on the structure and association mapping done using the population structure. The population structure was very interesting; it shows population has several closely related breeding lines and few diverse lines. Except few lines, almost all are admixtures, but interestingly they have distinct genetic structure. Based on the Association mapping two markers (SSR36 and SSR106) were associated with CMD trait.

SUMMARY

Chapter 6

75

SUMMARY

Cassava is a major tuber crop generally growing in tropical regions under various environmental conditions and using as a stable food crop. Cassava is a heterozygous plant with 36 chromosomes. High heterozygous nature is due to the out-crossing character and is a monoecious plant, female and male flowers are born in same plant.

There are several pests and diseases that affect cassava plants. Cassava mosaic disease is one among the most severe disease the affect the plant growth and productivity.CMD is a viral disease caused by Gemini virus and transmitted by white fly (*Bemisia tabaci.*) But the major transmission occurs when using disease infected stem cuttings as propagating material. There are several approaches to control CMD, such as user of virus free planting materials and development of CMD resistant varieties. Sometimes it is difficult to distinguish resistant varieties phenotypically. Now a days SSR markers widely using for genetic diversity studies and the identification of CMD resistant varieties.

The present study on "Association mapping for cassava mosaic disease (CMD) resistance in Cassava using SSR markers" was carried out to identify the markers which are associated with CMD resistance in cassava varieties. The present study includes 30 CMD resistant and 25 susceptible cassava accessions. The study was conducted at ICAR-Central tuber crop research institute (CTCRI), Sreekariyam. Following are the summary of research work done in the study.

Genomic DNA was isolated from leaves of 30 CMD resistant and 25 susceptible varieties using Dellaporta method. Isolated DNA was good with no protein and RNA contamination. Quality and quantity of isolated DNA was checked by nano-spectrophotometer and 0.8% agarose gel electrophoresis respectively. The DNA was diluted into 50ng/µl and used for the PCR amplification of 20 selected SSR primers. After the PCR amplification, they were analyzed on Polyacrylamide gel electrophoresis. Bands obtained from PAGE

were scored in the binary form. RME1 primer showed double banding pattern in 700bp size for all the resistant accessions .But in case of susceptible accessions only single bands were observed. The scored data were then used for the bioinformatics analysis using STRUCTURE and TASSEL software.

TASSEL software was used to identify relationship between genotype and phenotype. Dissimilarity matrix value was obtained by pairwise comparison of cassava species among 55 accessions and a graph was plotted using frequency of dissimilarity coefficient value.Cladogram was constructed using software. CMD resistant and CMD susceptible varieties were grouped separately. Marker trait association result revealed that two markers were associated with CMD.They were SSR36 and SSR106.

STRUCTURE software was used to identify the subpopulations and admixtures in the collected cassava accessions. A graph was plotted between 'Delta K' and number of sub populations (K) using structure harvester by Evanno's method for the identification of total subpopulation in the population used. Result showed that there are 8 sub population in the populations used. Bar diagram was constructed, which shows the distribution of genotypes in different sub- populations.

REFERENCES

7. REFERENCES

Abdurakhmonov, I. and Abdukarimov, A. 2008. Application of association mapping to understanding the genetic diversity of plant germplasm resources. *Int. J. Pl. Genomics*. 200: 1-18.

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubrick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu. A., Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie, W. R., and Venter, J. C. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Sci.* 252: 1651-1656.

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

Akbari, M., Wenzl, P., Caig, V., Carlig, J., Xia, L., Yang, S., Uszynski, G., Mohler, V., Ehmensiek, A., Howes, N., Sharp, P., Huttner, E., and Kilian, A. 2006. Diversity arrays technology (DArT) for high throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.*113: 1409-1420.

Akkaya, M. S., Bhagwat, A. A., and Cregan, P. B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genet.* 132: 1131-1139.

Akopyanz, N., Bukanov, N. O., Westblom, T. U., Berg, D. E. 1992. PCR-based RFLP analysis of DNA sequence diversity in the gastric pathogen Helicobacter pylori. *Nucleic Acid Res.* 20: 6221-6225.

Al-Maskri, A.Y., Sajjad, M. and Khan, S.H. 2012. Association mapping: a step forward to discovering new alleles for crop improvement. *Int. J. Agric. Biol.* 14: 153–160.

19

Asare, P. A., Labuschagne, M. T., Herselman L., and Mahungu, N. 2011. Morphological and molecular based diversity studies of some cassava (Manihot esculenta Crantz) germplasm in Ghana. *Afr. J. Biotechnol.* 10: 13900-13908.

Asare, P. A., Galyuon, K. A., Bediako, E. A., Safo, J. K., and Teeth, J. P. 2004. Ph enotypic and molecular screening of cassava (*Manihot esculenta* Crantz) genotypes for resistance to cassava mosaic disease. *J. Gen. Mol. Biol.* 6(2): 6-8.

Balding, D. 2006. A tutorial on statistical methods for population association studies. *Nat. Rev. Genetics.* 7: 781-791.

Benson, G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27: 573-580.

Bi, H., Aileni, M., Zhang, P. 2010. Evaluation of cassava varieties for cassava mosaic disease resistance jointly by agro-inoculation screening and molecular markers. *Afr. J. Pl. Sci.* 4(9): 330-338.

Breseghello, F. and Sorrells, M. 2006 Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci.* 46:1323-1330.

Byrne, D. 1984. Breeding Cassava. *Pl. Breed. Rev.* 2: 73-133.

Carmo, C. D., Silva, M. L., Oliveira, G. A. F., and Oliveira, E. J. 2015. Molecular-assisted selection for resistance to cassava mosaic disease in *Manihot escullenta Crantz. Sci. Agric.*72(6): 520-527.

Chapman, J.M., Cooper, J.D., Todd, J. A., and Clayton, D. G. 2003. Detecting disease associations due to linkage disequilibrium using haplotype ags: a class of tests and the determinants of statistical power. *Hum. Heredity.* 56: 18-31.

Chavarriaga-Aguirre, P. P., Maya, M. M., Bonierble, M. W., Kresovich, S., Fregene, M. A., Tohme, J., and Kochert, G. 1998. Microsatellites in cassava ((*Manihot esculenta* Crantz): discovery, inheritance and variability. *Theor. Appl. Genet.* 97(3): 493-501.

da Costa, T. R., Filho, P. S. V., Vidigal, M. C. G., Galván, M. Z., Lacanallo, G. F., da Silva, L. I., and Kvitschal, M. V. 2013. Genetic diversity and population structure of sweet cassava using simple sequence repeat (SSR) molecular markers. *Afr. J. Biotechnol.* 12(10): 1040-1048.

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

Deng, D., Otim-Nape, G. W., Sangare, A., Ogwal, S., Beachy, R.N., and Fauquet, C.M. 1997. Presence of a new virus closely associated with cassava mosaic outbreak in Uganda. *Afr.J. Root Tuber Crops.* 2: 23-28.

Duraisamy, R. Rathinasamy, A. S., Antisana, S., Muthurajan, R., Ramineni, J. J., Karuppusamy, N., Lakshmanan, P., Mohan , C., and Gandhi, K. 2011. Starch Content and Cassava Mosaic Disease Genetic Diversity with Relation to Yield in South Indian Cassava (Manihot esculenta Crantz) Germplasm. *J. Crop Sci. Biotech.* 14(3): 179-189.

Elias, M., Muhlen, G.S., Roa, A.C., and Thome, J. 2004. Genetic diversity of traditional south American landraces of cassva (*Manihot esculenta* Crantz): an analysis using microsatellites. *Econ. Bot.* 58(2): 242-256.

Esuma, W., Herselman, L., Labuschagne, M. T., Ramu, P., Baguma, F. L. Y., Buckler, E. S., and Kawuki, R. S. 2016. Genome-wide association mapping of provitamin A carotenoid content in cassava. *Euphytica*. 212:97–110.

Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14: 2611-2620.

Fauquet, C. and Stanley, J. 2003. Geminivirus classification and nomenclature progress and problems. *Ann. Appl. Biol.* 142: 165-189.

Favoretto, P., Easey, E. A., and Melo, P. C. T. 2011. Molecular characterization of potato cultivars using SSR markers. *Hortic. Brasileira*. 29: 542-547

Flint-Garcia, S., Thornsberry, J. and Bukler, E. 2003. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54: 357-374.

Fouquet, C. and Fargette, D. 1990. African cassava mosaic virus. Etiology, Epidemiology and Control. *Pl. Dis.* 74(6): 404-411.

Fregene, M. and Puonti-Kaerlas, J. 1997. Cassava biotechnology. In: Cassava: Biology, Production and Utilization (Hillocks, R.J. and Thresh, J.M., eds), pp. 179-208.

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

Garcia, F. A. A., Kido, A. E., Meza, N. A., Souza, B. H. M., Pinto, R. L., Pastina, M. M., Leite, S. C., Silva, G. J. A., Ulian, C. E., and Figueira, A. 2003. Development of an integrated genetic map of a sugarcane (*Saccharum* spp.) commercial cross, based on a maximum likelihood approach for estimation of linkage and linkage phases. *Theor. Appl. Genet.* 112: 298-314.

Gebhardt, C. A., Ballvora, B., Walkemeier, P., Oberhagemann, J., and Schuler, K. 2004. Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Mol. Breed.* 13: 93-102.

Gibson, R.W. 1994. Long-term absence of symptoms in heat treated African cassava mosaic geminivirus-infected resistant cassava plants. *Trop. Sci.*34: 154-8.

Goodstein, D. M., Shu, S., Howson, R. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40: 1178-1186.

Grodzicker, T., Williams, J., Sharp, P., and Sambrook, J. 1975. Physical mapping of temperature sensitive *mutants of adenovirus. Cold Spring Harbor Symp. Quant. Biol.* 39: 439-446.

Gupta, P., Rustgi, S., and Kulwal, P. 2005. Linkage disequilibrium and association studies in higher plants: present status and prospects. *Pl. Mol. Biol.* 57: 461-485.

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

Hall, D., Tegstrcom, C., and Ingvarsson, P. K. 2010. Using association mapping to dissect the genetic basis of complex traits in plants. *Brief. Funct. Genomics.* 9: 157-165.

Harrison, B. D. and Robinson, D. J. 1999. Natural genomic and antigenic variation in whitefly transmitted geminiviruses (begomoviruses). *Ann. Rev. Phytopathol.* 37: 369-398.

Henry, R.J. 2008. *Plant Genotyping II: SNP Technology*, CABI Publishing, Wallingford, pp. 280-288.

Jordan, S. A. and Humphries, P. 1994. Single nucleotide polymorphism in exon 2 of the BCP gene on 7q31-q35. *Hum. Mol. Genet.* 3: 1915-1918.

Kumar, B., Talukdar A., Bala I., Verma K., Lal S. K., Sapra R. L., Namita B., Chander S. and Tiwari R. 2014 Population structure and association mapping studies for important agronomic traits in soybean. *J. Genet.* 93: 775-784

Kwak, M. and Gepts, P. 2009. Structure of genetic diversity in two major gene pools of common bean (Phaseolus vulgaris L., Fabaceae). *Theor. Appl. Genet.* 118: 979-992.

Lamptey, J. N. L., Okoll, O. O., 1998. Frimpong-Manso Incidence and severity of African Cassava mosaic disease (ACMD) and cassava bacterial blight (CBB) on some local and exotic cassava varieties in different ecological Zones of Ghana. *Ghana J. Agric. Sci.* 31(1): 35-443.

Lamptey, J. N. L., Okoll, O. O., Rossel, H. W., Frimpong-Manso, P. P. 2000. A method for determining tolerance of cassava genotype to African cassava mosaic disease in sacreenhouse. Ghana *J. Agric. Sci.* 33: 29-32.

Legg, J. P., Owor, B., Sseruwagi, P., Ndunguru, J., 2006. Cassava mosaic virus disease in East and Central Africa: epidemiology and management of a regional pandemic. *Adv. Virol. Res.* 67: 355-418.

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

Li, L., Strahwald, J., Hofferbert, H. R., Lubeck, J., Tacke, E., Junghans, H., Wunder, J., and Gebhardt, C. 2005. DNA variation at the invertase locus invGE/GF is associated with tuber quality traits in populations of potato breeding clones. *Genet.* 170: 813-821.

Lokko, Y., Danquah, E.Y., Offei, S. K., Dixon, A.G.O., and Gedil, M. A. 2004. Molecular markers associated with a new source of resistance to the cassava mosaic disease. *Afr. J. Biotechnol.* 4: 873-881.

Lokko, Y., Dixon, A., Offei, S., Danquah, E., and Fregene, M. A, 2006. Assessment ofgenetic diversity among African cassava accessions resistant to Cassava Mosaic Virus Disease using SSR marker. *Genet. Crop Resour. Evol.* 53: 1441-1453.

Malosetti, M., Linden, C. G., Vosman, B., and Eeuwijk, F. A., 2007. A mixed-model approach to association mapping using pedigree information with an illustration of resistance to Phytophthora infestans in potato. *Genet.* 175: 879-889.

Mason, A. S. 2015 SSR Genotyping. In: Batley J (ed) Plant Genotyping. Springer, New York, NY, pp. 77-89.

Mba, R. E.C., Stephenson, P., Edwards, K., Melzer, S., Nkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, J., and Fregene, M. 2001 Simple Sequence Repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theo.Appl. Genet.* 102: 21-31.

Mezzete, T. F., Blumer, C. P., and Veasey, E. A. 2013 Morphological and molecular diversity among cassava genotypes. *Pesq. Agropec Bras.* 48: 510-518.

Mohan, C., Shanmugasundharam, P., Kaneswaran, 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.

Moreno, M., Tomkins, J., and Fregene, M. 2004. Positional cloning of CMD2 the gene that confers high level of resistance to the cassava mosaic disease (CMD). Sixth International Scientific Meeting of the Cassava Biotechnology Network (CBN-VI) CIAT, Cali, Colombia (pp. 163).

Moses, E., Asafu-Agyei, J. N., Adubofour, K., and Adusei, A. 2007. Guide to identification and control of cassava diseases. Kumasi, Ghana. *CSIR Crop Res. Inst.* 33: 7-41.

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. *Afr. J. Biotechnol.* 6: 2666-2674.

Myles, S., Peiffer, J., Brown, P., Ersoz, E., Zhang, Z., Costich, D., and Buckler, E. 2009. Association mapping: critical considerations shift from genotyping to experimental design. *Pl. Cell.* 21: 2194-2202.

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

Nicholas, R. F. W. 1947. Breeding cassava for virus resistance. *East Afr. Agric. J.* 15: 154-160.

Nwokoro, S.O., Orheruata, A. M., and Ordiah, P. I. 2002. Replacement of maize with cassava sievates in cockerel starter diets: effect on performance and carcass characteristics. *Trop. Anim Health Prod.* 34: 163-167.

Okezie, B. O. and Kosikowski, F.V. 1982. Cassava as a food. *Crit. Rev. food Nutr.* 17: 259-275.

Okogbenin, E., Egesi, C. N., Olasanmi, B., Ogundapo, O., Kahya, S., Hurtado, P., Marin, J., Akinbo, O., Mba, C., Gomez H., 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: 2576-2586.

Okogbenin, E., Setter, T. L., Ferguson, M., Mutegi, R., Ceballos, H., O*lasanmi*, B., and Fregene, M. 2013 Phenotypic approaches to drought in cassava: review. *frontiers Physiology*. 4: 1-15.

Olsen, M., Hood, L., Cantor, C., and Botstein, D. 1989. A common language for physical mapping of the human genome. *Sci.* 245: 1434-1435.

Oscar, J., Maria, E., Castro, L., Castro, E., and Barrantes, R.2008. Microsatellite variation of cassava (Manihotesculenta Cratz) in home gardens of chibchan Amerindians from Costs Rica. *Conserv Gent.* 9: 107-118.

Paran, I. and Michelmore, R.W. 1993. Development of reliable PCR based markers linked to downy mildew resistance genes in lettuce. *Theor. Appl. Genet.* 85: 985-993.

Pereira, G. S., Nunes, E. S., Laperuta, L. D. C., Braga, M. F., Penha, H. A., Diniz, A. L., Munhoz, C. F., Gazaffi, R., Garcia, A. A. F., and Vieira, M. L. C., 2013. Molecular polymorphism and linkage analysis in sweet passion fruit, an outcrossing species: Molecular map in sweet passion fruit. *Ann. Appl. Biol.* 162:347–361.

Pritchard, J. K., Stephensand, M., and Donelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. 155: 945-959.

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

Risch, N. and Merikangas, K. 1996. The future of genetic studies of complex human diseases. *Sci.* 273: 1516-1517.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich., H. A., and Arnheim, N. 1985. *Sci.* 230: 1350-1354.

Shahriar, M. H., Robin, A. H. K., Begum S. N., and Hoque, A. 2014. Diversity analysis of some selected rice genotypes through SSR based molecular markers. *J. Bangladesh Agril. Univ.* 12(2): 307-311.

Siqueira, M.V., Pinheiro, B. M., Borges, T. T., Valle, A., Zatarim, T. L., and Veasey, E. A. 2010. Microsatellite polymorphisms in cassava landraces from the Cerrado biome, Mato Grosso do Sul, Brazil. *Biochemical Genet.* 48: 879-895.

Slatkin, M. 2008. Linkage disequilibrium: Understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genetics*. 9: 477-485.

Souza, L. M., Gazaffi, R., Mantello, C. C., Silva, C. C., Garcia, D., Le Guen, V., Cardoso, S. E. A., Garcia, A. A. F., and Souza, A. P. 2013. QTL Mapping of growth-related traits in a full-sib family of rubber tree (*Hevea brasiliensis*) evaluated in a sub-tropical climate. *PLoS One*. 8: 61238-61240.

Stich, B., Urbany, C., Hoffmann, P. 2013. Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the Sol CAP SNP array. *Pl. Breed.* 13: 718-724.

Swamy, B. P. M., Shamsudin, N. A. A., Noorzuraini, S., Rahman, A., Mauleon, R., Ratnam, W., Cruz, M. T. S., and Kumar, A. 2017. Association Mapping of Yield and Yield Related Traits Under Reproductive Stage Drought Stress in Rice (Oryza sativa L.). *Rice.* 10: 21-25.

Taniguchi, H., Lowe, C. E., Cooper, J. D., Smyth, D. J., Bailey, R., Nutland, S., Healy, B. C., Lam, A. C., Burren, O., Walke, N. M., Smink, L. J., Wicker, L. S., and Todd, J. A. 2006. Discovery, linkage disequilibrium and association analyses of polymorphisms of the immune complement inhibitor, decay accelerating factor gene (DAF/CD55) in type 1 diabetes. *BMC Genet.*7: 22-28.

Thornsberry, J. M., Goodman, M. M., Doebley, J., Kresovich, S., Nielsen D., and Buckler, E. S. 2001. *Dwarf* polymorphisms associate with variation in flowering time. *Nat. Genet.*, 28: 286-289.

Thresh, J. M., Otim-Nap, G. W., and Nicholas, R. F. W. 1994. Strategies for controlling African Cassava mosaic geminivirus. *Adv. Dis. Vector. Res.* 10: 215-236.

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

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De L, T., Hornes, M., Frijters, A, Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. 23: 4407- 4414.

Weising, K., Nybom, H., Wolff, K., Meyer, W. 1995 *DNA Fingerprinting in Plants and Fungi* (ed. Arbor, A.) CRC Press, Boca Raton, pp. 1-3.

Weiss, K. M. and Clark, A. G. 2002. Linkage disequilibrium and mapping of human traits. *Trends Genet*. 18: 19-24.

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

Wolfe, D. M., Rabbi, I. Y., Egesi, C., Hamblin, M., Kawuki, R., Kulakow, P., Lozano, R., Dunia Carpio, D. P. D., Ramu, P., and Jannink, J. L. 2016 Genome-Wide Association and Prediction Reveals Genetic Architecture of Cassava Mosaic Disease Resistance and Prospects for Rapid Genetic Improvement. *Pl. Genome*. 9(2): 1-13.

Zhao, Y., Wang, H., Chen, W., and Li, Y. 2014. Genetic Structure, Linkage Disequilibrium and Association Mapping of Verticillium Wilt Resistance in Elite Cotton (*Gossypium hirsutum L.)* Germplasm Population. *PLoS One*. 9(1): 86-308.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G. W., Robinson, D. J., and Harrison B. D., 1997. Evidence that DNAA of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J. Virol.* 78: 2101-2111.

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

APPENDICES

92

aAPPENDIX I

Chemical reagents used in genomic DNA extraction

a. **DNA extraction buffer (*Dellaporta et al.*, 1983)**
   - Tris -HCl (Ph 8.0) :100Mm
   - EDTA :5mM
   - NaCl :500mM
   - B-Mercaptoethanol :2%(v/v)-Freshly added
   - PVP :2%(W/V) Freshly added

b. SDS-20%

c. Potassium acetate-500Mm

d. Ice-cold ethanol

e. Sodium acetate (p H5.2)-300Mm

f. RNase A-10mg/ml

   RNase A was dissolved in TE buffer and boiled for 15min at 100 °C to destroy DNase and stored at-20 °C

g. Chloroform: Isoamyl alcohol - 24:1

h. Ethanol - 70%

i. **Preparation of 10X TE buffer**
   - Tris - HCl (Ph 8.0) :10mM
   - EDTA :1Mm

   Final volume was made up to 100ml with distilled water.

93

APPENDIX-II

Chemicals required in Agarose Gel Electrophoresis

a. Preparation of 10XTBE buffer

- Tris base :107g
- Boric acid : 55g
- 0.5M EDTA (Ph 8.0) :40 ml

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

b. Ethidium bromide

10mg of ethidium bromide powder (Sigma-Aldrich) and dissolve in 1 ml of distilled water

c. Preparation of gel loading dye

- Formamide :50ml
- Xyline cyanol :50mg
- Bromophenol blue :50mg
- EDTA :0.5M

d. Empty well dye

- Loading dye :50µl
- Sterile water :50µl

APPENDIX-III

PCR cocktail for SSR primers

	Stock concentration	volume taken concentration	final
DNA	20ng	2.0 µl	2.0ng
Primer	2.0 µM	2.0 µl	0.2µM
DNTPs	40 µM	0.2 µl	0.4 µM
Taq buffer	10 X	2.0 µl	1 X
Taq DNA polymerase	3.0 unit	0.2 µl	0.03 Unit
Sterile water		13.6 µl	

95

APPENDIX IV

Chemical reagents and solutions for DNA-PAGE

a. Bind silane

- Absolute ethanol :99.5%
- Acetic acid 0.5%
- Bind silane :1µl

b. 4% Acrylamide

- Acrylamide :38g
- Bis -Acrylamide :2g

Final volume made up to 100ml using distilled water.

c. 6%Denaturing polyacrylamide gel containing 7M Urea

42g Urea weighed and dissolved in a beaker containing 10ml TBE buffer(10X) and 15 ml distilled water by heating in a microwave oven for 30-40s.15ml acrylamide solution (19:1) was filtered and added to measuring cylinder followed by the melted urea solution. The final volume was made up to 10ml using distilled water and stored in dark till use.60µl TEMED (1:10) and 600µl APS(1000mg/ml) was added and mixed just before casting the gel

d. Fixer

- Acetic acid: 200ml
- Distilled water:1800ml

Final volume made up to 2000ml

e. Silver stain

- Silver nitrate:2g
- Formaldehyde:3ml

Final volume made up to 2000ml

f. Developer

- Sodium carbonate:60g

Final volume made-up to 2000ml and stored in -20 C until use.

- Formaldehyde:3ml
- Sodium thiosulphate(10mg/ml):400µl

The above reagents were freshly added and mixed thoroughly with the chilled sodium carbonate solution before use.

ABSTRACT

ASSOCIATION MAPPING FOR CASSAVA MOSAIC DISEASE (CMD) RESISTANCE IN CASSAVA USING SSR MARKER

by

APARNA T. K.

(2012-09-117)

Abstract of Thesis

Submitted in partial fulfilment of the

Requirement for the degree of

B. Sc. - M. Sc. (INTEGRATED) BIOTECHNOLOGY

Faculty of Agriculture
Kerala Agricultural University, Thrissur

B. Sc. - M. Sc. (INTEGRATED) BIOTECHNOLOGY
DEPARTMENT OF PLANT BIOTECHNOLOGY
COLLEGE OF AGRICULTURE
VELLAYANI, THIRUVANANTHAPURAM - 695 522
KERALA, INDIA

2017

99

ABSTRACT

The present study" Association mapping for Cassava Mosaic Disease (CMD) resistance in cassava using SSR marker" was conducted to identify the marker which are closely associated with CMD resistance in cassava. 30 CMD resistant and 25 CMD susceptible cassava varieties were analyzed with 20 selected SSR primers. Genomic DNA isolated from the leaves and PCR amplification was done with SSR primers. PCR products were separated by PAGE. Polymorphic bands were used to assign loci for each primer and score as presence (1) or absence (0) of bands. RME1 primer showed double banding pattern in 700bp size for all the resistant accessions .But in case of susceptible accessions only single bands were observed.

TASSEL and STRUCTURE software were used to analyses the data obtained from the PAGE. Cladogram was constructed using genotypic data. Different clusters were formed. Clusters clearly distinguished CMD resistant and CMD susceptible varieties of cassava.CMD resistant varieties sub divided into three sub clusters. Dissimilarity matrix constructed by the software used for the diversity study of collected cassava accessions. Two markers were associated with CMD resistance was identified by the software. SSR106 and SSR36 are the two markers associated with CMD resistance. Result showed that there are eight sub populations in the cassava population used. Based upon the output obtained from Evanno's method a bar diagram was constructed. Which shows the distribution of genotypes in different sub populations.

174185