

https://dspace.org/

Influence of satellite DNA molecules on severity of cassava begomoviruses and the breakdown of resistance to cassava mosaic disease in Tanzania

Ndomba, Osmund Aureus

2013-02-14
http://hdl.handle.net/10539/12427

Downloaded from a DSpace repository. https://dspace.org/

INFLUENCE OF SATELLITE DNA MOLECULES ON SEVERITY OF CASSAVA BEGOMOVIRUSES AND THE BREAKDOWN OF RESISTANCE TO CASSAVA MOSAIC DISEASE IN TANZANIA

Osmund Aureus Ndomba

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of

Doctor of Philosophy

Johannesburg

2012

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. The thesis has not been submitted before for a degree or other award to this or another University.



Osmund Aureus Ndomba

This the 19TH October 2012

Abstract

Cassava *Manihot esculenta* Crantz (*Euphorbiaceae*), is a source of food for more than 700 million people in developing countries and is cultivated in estimated global area of 18.6 million hectares with total annual production of 238 million tonnes. Diseases however, take a substantial toll of yield, with CMD being the most important disease and major constraint for cassava production in Tanzania and Africa. The disease causes an estimated loss of over US$ 14 million per annum.

A study was undertaken in 2006/2007 to investigate the influence of satellite DNA molecules on severity of cassava begomoviruses and the breakdown of resistance to cassava mosaic disease (CMD) in Tanzania. The goal was to appraise the nature of resistance to CMD in indigenous and improved cassava cultivars in the presence of resistance-breaking satellites. Three specific aims were earmarked: to identify and characterize cassava mosaic virus isolates and satellite DNA molecules in major cassava growing areas of Tanzania; to screen cassava cultivars for resistance to begomoviruses in presence and absence of the satellite DNA molecules; and to determine the nature of interaction between begomovirus DNAs and Satellite DNA molecules in *Nicotiana benthamiana*.

To achieve these aims, a survey was done in the major cassava growing areas of Tanzania to investigate occurrence of cassava mosaic begomoviruses and associated satellites namely, SatDNA-II and SatDNA-III. Stems from plants showing CMD symptoms were collected from field. The stems were re-planted in screenhouse to study more about the symptoms. Symptomatic leaves from sprouting cuttings were collected for DNA extraction to be used in two downstream assays - amplification of EACMV, ACMV, SatDNA-II and SatDNA-III by polymerase chain reaction (PCR) and sequencing.

In another experiment to evaluate cassava cultivars for resistance to CMD in presence of satellites, stem cuttings of the classical CMD-resistant cultivars were planted in greenhouse. Infectious clones of EACMV-TZ and EACMV-UG2 comprising both

DNA-A and DNA-B components of bipartite begomoviruses (EACMV-TZ and EACMV-UG2) as well as infectious clones of SatDNA-II and SatDNA-III were bombarded onto the greenhouse cassava plants using a gene gun. Emerging disease symptoms on inoculated plants were scored using standard procedure. Total nucleic acid extraction from the inoculated plants was done and PCR was performed to amplify *AC1* and *βC1* genes as well as full length SatDNA-II and SatDNA-III. Southern blot analysis was performed to determine the presence of *AC1*, *βC1*, SatDNA-II and SatDNA-III on the DNA.

In order to study interaction between cassava mosaic begomovirus (EACMV-TZ) and satellites, infectious clones of EACMV-TZ (DNA-A and DNA-B) and that of SatDNA-II and SatDNA-III were used. The clones (DNAs) were used to infect *Nicotiana benthamiana* by abrasion. Inoculated plants were covered with a plastic dome and placed in insect-free growth chamber for symptom development, which were scored on a standard scale of 1 to 5. Total DNA was extracted from the *N. benthamiana* leaves and used for Southern blot analysis.

Results from the field survey showed that disease incidences varied from 60 to 90% in the Lake Victoria Zone and from 10 to 90% in the Eastern Zone. Cultivar *Lyongo* had the highest disease symptom severity in the Lake Victoria Zone while in the Eastern zone plants with high severity levels were from cvs *Maiza* and *Tabora*. In the screenhouse, some sprouted cuttings remained healthy up to 16 days after planting (DAP) and others recovered from the disease. Reversion was also observed in some cultivars. Using PCR, *East African cassava mosaic Tanzania virus* (EACMV-TZ) was amplified from 72.8% of tested samples while *African cassava mosaic virus* (ACMV) was amplified from 4.3%. Five percent of plants had dual infection of the two viruses. While ACMV was detected in samples collected from Lake Victoria, EACMV-TZ was mostly found on samples from the Eastern zone. Sequencing showed the presence of two new virus isolates: EACMV-TZ [TZ113] and EACMV-TZ [TZ108]. Seventy five percent of plants, which showed reversion of symptoms, contained SatDNA-II. It was found that full length SatDNA-II occurred in both zones, while SatDNA-III was

exclusive to the Lake Zone. Multiple DNA bands were noted in PCR agarose gels, more so in SatDNA-II than SatDNA-III. For SatDNA-II, the multiple bands were more evident for samples collected from Eastern zone than for those from the Lake Zone. Using primers based on expressed sequence tags (EST-primers) for SatDNA-II (895 bp) and SatDNA-III (306 bp), genome integrated forms of the satellites were amplified from 68% and 71.17% of samples, respectively. Thirty percent of the samples showed co-infection of the satellites. While EST-primers for detection of the integrated forms of SatDNA-III produced single bands on gels, those of SatDNA-II still produced additional bands, most noteworthy being the closely spaced 'double bands'. Upon sequencing, the satellite DNA isolates showed similarity with sequences deposited in the genebank and bearing accession numbers AY836366 and AY836367 for SatDNA-II and SatDNA-III isolates, respectively. Alignment reports (Clustal W) revealed presence of GC-rich regions, TATA protein binding motifs (TATAAAT) and CAAT boxes as well as poly (A) signal. GC-rich regions in SatDNA-II were mostly trinucleotides (CGC) and hexanucleotides (CCGCCG) while in SatDNA-III the regions were trinucleotides (CGC) and pentanucleotides (CCGCC).

Following biolistic inoculation, five-week scoring for the symptoms showed that plants from cvs AR37-92, CR27-24 and AR16-3 remained symptomless while plants from cv T200 were symptomatic. PCR amplification of *βC1* gene five weeks post inoculation (*wpi*) gave PCR products in 19.6% of the samples while *AC1* was amplified from only two plants. Full-length SatDNA-II was amplified from 70% of DNA samples, mostly from plants in which a begomovirus was co-inoculated with SatDNA-II. Amplification of full-length SatDNA-III from bombarded plants was unsuccessful. Amplification of integrated fragments of SatDNA-II from bombarded plants using EST-primers gave a PCR product in 93.7% of the samples. PCR amplification of the fragments from DNAs extracted from plants of cvs AR17-5 and CR27-24 previously inoculated with EACMV-TZ + SatDNA-II and EACMV-UG2 + SatDNA-III, respectively, gave closely spaced bands on 13% of the DNA samples. Amplification of integrated forms of SatDNA-III gave bands in 52.4% of samples. Probing for full-length SatDNA-II, SatDNA-III and *AC1* from DNAs extracted from plants pre-inoculated with these DNAs using DIG-

labeled probes gave hybridization signals in 60%, 83% and 68% of the samples, respectively. Further analysis of the signals in the context of screening suggested that cvs AR37-92 and AR37-96 were highly resistant to CMD while cv AR40-10 was susceptible.

In the interaction experiment, *Nicotiana benthamiana* plants inoculated with an infectious clone of EACMV-TZ developed moderate CMD symptoms 7 days post inoculation (*dpi*) with symptoms consisting of leaf distortion and moderate stunting of plants. There were also plants which recovered from the symptoms by 35 *dpi*. Plants inoculated with EACMV-TZ + SatDNA-II produced similar symptoms with *N. benthamiana* plants developing symptoms 7 *dpi* that became severe by 14 *dpi* and without recovery even after 35 *dpi*. Very severe symptoms were also observed when *N. benthamiana* plants were inoculated with EACMV-TZ + SatDNA-II + SatDNA-III. Plants inoculated with SatDNA-II or SatDNA-III alone remained asymptomatic even after 35 dpi. Southern blot analysis showed clear increase in DNA accumulation when EACMV-TZ was inoculated together with both SatDNA-II and SatDNA-III as compared to when EACMV-TZ was inoculated alone or with SatDNA-II only and probed with EACMV-TZ.

In conclusion, symptom recovery and reversion of symptoms in screenhouse plants is associated with virus resistance. There is a wide occurrence of satellites (SatDNA-II and SatDNA-III) across the sampled regions consistent with distribution of their helper cassava begomoviruses. The satellites are of a wider occurrence and diversity in Eastern zone than elsewhere in the country. The occurrence of SatDNA-III was not confined to the Lake zone as previously thought. There is evidence for satellite sequence integration into host plant genome, a further indication that the satellites are wider spread in cassava germplasm than earlier conceptualized. In few instances, both SatDNA-II and SatDNA-III isolates co-existed in the same plant though its effect on symptom enhancement could not be immediately established. The observed recovery in screening studies is thought to result from resistance introduced in the plant materials involved. Since labeled probes for satellites that were used in hybridization had been prepared from satellite sequences

considered to be integrated, the hybridization signals did not depend on whether the leaf samples were picked from symptomatic or asymptomatic plants. From the study, three observations clearly suggest that SatDNA-II and SatDNA-III are biologically functional and that their effects on host plants are distinctly different. The study has demonstrated enhanced cassava begomovirus symptoms in *N. benthamiana* in the presence of satellite DNA molecules. This is the first detailed study undertaken to highlight the occurrence and role played by satellite DNA molecules in breaking the resistance to CMD of cassava cultivars grown in Tanzania.

Keywords: Cassava mosaic disease, Cassava mosaic begomoviruses, Satellite DNA molecules, Tanzania.

Dedication

The thesis is affectionately dedicated
to my daughter
Veronica O. Ndomba

Acknowledgements

Various institutions, bodies and individuals have helped in completion of this work. I am grateful to the Kirkhouse Trust (UK), for sponsoring the programme. I wish to accord a sincere gratitude to Tropical Pesticides Research Institute (TPRI) of Arusha in Tanzania for granting a study leave and for financial support. I thank the Faculty of Science, University of the Witwatersrand in Johannesburg for the overall coordination of this study program. I cave in to appreciate the International Center for Tropical Agriculture (CIAT) in Colombia for financial administration and for hosting me as a Visiting Researcher within their Marker Assisted Selection (MAS) Project based at Mikocheni Agricultural Research Institute (MARI) in Dar es Salaam. I am deeply indebted to MARI for the provision of 'bench' research facilities as well as some laboratory consumables. The staffs at both TPRI and MARI are also generously thanked for their encouragement!

Several people helped by commenting, supplying information or inspiring through their own work. Special thanks to my supervisors Prof. Chrissie Rey, Dr. Martin Fregene and Dr. Joseph Ndunguru who rendered their professional support whenever it was required. Prof. Rey in addition to technical support helped solicit additional guidance as the situation would require. Dr. Joseph Ndunguru, with very high level of humility, assisted me in various ways in coping with the hard-to-tell enduring research environment with respect to the availability of reagents and research equipment while in Tanzania. Dr. Ndunguru also sponsored some training sessions in mapping disease occurrence, sequence analysis and studying vector populations in the field. Dr. Martin Fregene is acknowledged for rolling down a series of important events prelude to the successful matriculation and enrolment into the University. He further facilitated the availability of some laboratory reagents and for financial cushioning between Kirkhouse and me, the student. Overall, I was particularly moved by the way these supervisors collaborated among themselves while giving technical advice to me. They suggested two courses at the University of the Witwatersrand in which I learnt so many molecular biology techniques and principles - I thank them!

It is difficult to have a comprehensive list of all the people who supported me in various ways and state how they did individually. However, I wish to mention some researchers and fellow students who, in their research institutions particularly made my supervisors' work an easy one. They include (though not limited to) – Dr. Fred Tairo (MARI), Mr. Rudolf Shirima (IITA Tanzania), Mr. Rajabu Kitona (MARI), Mrs. Imanu Mwaba (University of the Witwatersrand), Mrs. Tarryn Seagram (University of the Witwatersrand), Ms. Claudia Rossin (University of the Witwatersrand), Ms. Erica Pierce (University of the Witwatersrand) and Ms. Maabo Moralo (University of the Witwatersrand).

Research Outputs

No conference or research outputs to date.

Table of Contents

Page

Title Page ---------- i
Declaration ----------ii
Abstract ----------iii
Dedication ---------- viii
Acknowledgements ----------ix
Research Outputs ----------x
Table of Contents ----------xi
List of Figures ----------xviii
List of Tables ---------- xx
Acronyms and Abbreviations ----------xxii
List of References ----------xxv
List of Appendices ----------xxvi

CHAPTER 1: GENERAL INTRODUCTION AND OBJECTIVES ----------1

1.1 OVERVIEW OF PLANT VIRUS DISEASES IN AFRICA ----------1
1.2 CASSAVA MOSAIC DISEASE AS A CONSTRAINT ----------2
1.3 WHY CMD REMAINS A PROBLEM ----------3
1.4 ROLE OF PLANT RESISTANCE IN CURBING CMD ----------4
1.5 SATELLITES AND THEIR ROLE IN BEGOMOVIRUS EPIDEMIOLOGY ---5
1.6 SATELLITE DNA MOLECULES ----------7
1.7 PROBLEM JUSTIFICATION ----------9
1.8 PROBLEM HYPOTHESIS ----------9
1.9 OBJECTIVES OF THE STUDY ---------- 13
1.10 THESIS ORGANIZATION, SOURCES OF DATA AND METHODOLOGY - 13
1.11 REFERENCES----------15

CHAPTER 2: LITERATURE REVIEW ---------- 19

2.1 CASSAVA AND LIVELIHOOD ---------- 19

2.1.1 General Description and Uses of Cassava ---------- 19
2.1.2 Production of Cassava ---------- 20
2.1.3 Cassava Production and Use in Tanzania ---------- 23
2.1.4 Constraints to Cassava Production ---------- 25
2.1.5 Cassava Mosaic Disease as Constraint to Cassava Production ---------- 26
2.2 CMD – IMPORTANCE SPREAD AND PREVALENCE ---------- 29
2.2.1 Early Reports and Control Strategies ---------- 29
2.2.2 CMD Symptoms ---------- 30
2.2.3 Crop Losses Due to CMD ---------- 33
2.2.4 CMD Pandemic in Eastern and Central Africa ---------- 34
2.3 ETIOLOGY OF CMD ---------- 34
2.3.1 Early Consideration ---------- 34
2.3.2 Structure and Organization of Geminiviruses ---------- 35
2.3.3 Replication of Geminivirus Genomes ---------- 38
2.3.4 Features of Genus *Begomovirus* ---------- 39
2.3.5 Arrangement and Functions of *Begomovirus* Genes ---------- 41
2.3.6 Classification and Diversity of Cassava Geminiviruses ---------- 43
2.3.7 Diversity of Cassava Mosaic Begomoviruses ---------- 46
2.3.8 Begomoviruses Occurring in Tanzania ---------- 47
2.4 INVOLVEMENT OF SATELLITES IN BEGOMOVIRUS PATHOGENICITY ---------- 47
2.4.1 Molecular Epidemiology of Begomoviruses ---------- 47
2.4.2 Factors Influencing *Begomovirus* Diseases ---------- 49
2.4.3 Suppression of Gene Silencing by Begomoviruses ---------- 50
2.4.4 Capacity for Genetic Variation ---------- 51
2.4.5 Recombination Hotspots ---------- 52
2.4.6 Association of Begomoviruses with Satellites ---------- 53
2.4.7 What are Satellites ---------- 54
2.4.8 Main Features of Satellites ---------- 54
2.4.9 Classification of Satellites ---------- 55

2.4.10 TLCV Sat-DNA, Betasatellites and Alphasatellites ---------- 55
2.4.11 Recombination and Reassortment Between Satellites and Begomoviruses ---------- 59
2.4.12 Effects of Satellites on Host and Helper Virus ---------- 60
2.4.13 Infectivity of Pseudorecombinants ---------- 62
2.5 INTEGRATION OF VIRAL AND SATELLITE SEQUENCES INTO HOST GENOME ---------- 63
2.5.1 What is Integration ---------- 63
2.5.2 Integration of Viral Sequences ---------- 64
2.5.2.1 Noninfectious endogenous BSV ---------- 65
2.5.2.2 Infectious endogenous BSV ---------- 65
2.5.3 Integration of Subgenomic DNAs and Satellite Sequences ---------- 65
2.5.4 Impact of Sequence Integration in Cassava ---------- 66
2.6 ASSOCIATION OF SATELLITES WITH OLD WORLD BIPARTITE BEGOMOVIRUSES ---------- 67
2.6.1 Early CMD Situation in Tanzania ---------- 67
2.6.2 Occurrence of Satellites in Tanzania ---------- 68
2.6.3 Features of SatDNA-II and SatDNA-III ---------- 69
2.7 CMD MANAGEMENT ---------- 70
2.7.1 Overview of CMD Management ---------- 70
2.7.2 Prophylactic Control Measures ---------- 71
2.7.3 Conventional Breeding for Resistant Varieties ---------- 73
2.7.4 Genetic Engineering for Protection ---------- 74
2.8 RESISTANCE AS AN INTERVENTION TO CMD ---------- 76
2.8.1 Viral Systemic Infection in Plants ---------- 76
2.8.2 Components of Resistance to CMD ---------- 76
2.8.3 Molecular Implications of Satellite Involvement in CMD ---------- 77
2.8.4 Resistance Studies in Tanzania ---------- 79

2.9 REFERENCES----------81

CHAPTER 3: MOLECULAR IDENTIFICATION AND DISTRIBUTION OF CASSAVA GEMINIVIRUSES AND SATELLITE DNA MOLECULES IN MAJOR CASSAVA - PRODUCING AREAS OF TANZANIA ---------- 91

ABSTRACT----------91
3.1 INTRODUCTION ---------- 93
3.2 MATERIALS AND METHODS ---------- 94
3.2.1 Field Survey and Sample Collection ---------- 94
3.2.2 Sprouting Cassava Cuttings in Screenhouse ---------- 96
3.2.3 Total Nuclear DNA Extraction ---------- 97
3.2.4 PCR Amplification of CMGs ---------- 98
3.2.5 PCR Amplification of EACMV ---------- 98
3.2.6 PCR Amplification of ACMV ---------- 99
3.2.7 PCR Amplification of Full-length SatDNA-II ---------- 99
3.2.8 PCR Amplification of Full-length SatDNA-III ----------100
3.2.9 Cloning and Sequence Verification ----------100
3.2.10 Sequence Analysis and Comparison ----------102
3.2.11 Detection of SatDNA and Survey for Satellites ----------102
3.2.12 Sequencing Satellite Fragments ----------103
3.3 RESULTS ----------104
3.3.1 Field Survey for CMD-infected Plants ----------104
3.3.2 Screenhouse Observations ----------109
3.3.3 PCR Amplifications ----------112
3.3.4 Results of Sequence Analysis ----------116
3.3.5 Survey of Satellite-like Molecules in Tanzania ----------120
3.3.6 Sequence Analysis of Satellite Fragments ----------
3.3.6.1 SatDNA-II ----------121
3.3.6.2 SatDNA-III ----------123
3.4 DISCUSSION ----------125
3.4.1 Incidence and Severity of Infection in the Field ----------125
3.4.2 Variability of Cassava Cultivar Symptoms Phenotypes ----------126

3.4.2.1 Severe symptoms in the field but healthy in screenhouse -----126
3.4.2.2 Mild symptoms in the field but severe in screenhouse ----------128
3.4.2.3 Differences in symptom expression within same plant --------129
3.4.3 Characterization of EACMVs ----------131
3.4.4 Survey for Satellites in Tanzania ----------131
3.4.5 Sequence Relatedness Between Satellite Fragments ----------131
3.5 CONCLUSION ----------132
3.6 REFERENCES----------133

CHAPTER 4: SCREENING OF CASSAVA CULTIVARS FOR DISEASE RESPONSE TO BEGOMOVIRUSES AND SATELLITE DNA MOLECULES -136

ABSTRACT----------136
4.1 INTRODUCTION ----------138
4.2 MATERIALS AND METHODS ----------139
4.2.1 Survey of Inoculated Cultivars for Resistance/Susceptibility ----------139
4.2.2 Reagents and Infectious Clones ----------141
4.2.3 Preparation of Plasmids ----------141
4.2.4 Preparation of Gold Particles and Biolistic Inoculation ----------142
4.2.5 Disease Severity Scores ----------144
4.2.6 Nucleic Acid Extraction ----------145
4.2.7 Southern Blot Analyses ----------146
4.3 RESULTS ----------146
4.3.1 Growing of Cassava Cuttings ----------146
4.3.2 Particle Bombardment ----------147
4.3.3 Symptoms of Biolistically-infected Plants ----------149
4.3.4 PCR Amplifications ----------156
4.3.4.1 *βC1* gene fragment amplification ----------156
4.3.4.2 *AC1* gene ----------157
4.3.4.3 Amplification of full length SatDNA-II ----------157
4.3.4.4 Amplification of full length SatDNA-III ----------159

4.3.4.5 Amplification of satellite fragments ----------159
4.3.5 Southern Blot Analyses ----------162
4.4 DISCUSSION ----------165
4.4.1 Bombardment and Symptom Score ----------165
4.4.2 PCR for *ßCI*, *AC1*, SatDNA-II and SatDNA-III ----------167
4.4.3 Amplification of Satellite Fragments ----------169
4.4.4 Southern Blot Hybridization ----------169
4.5 CONCLUSION ----------171
4.6 REFERENCES----------173

CHAPTER 5: NATURE OF INTERACTION BETWEEN BEGOMOVIRUS DNAs AND SATELLITE DNA MOLECULES IN *NICOTIANA BENTHAMIANA* ------175

ABSTRACT----------175
5.1 INTRODUCTION ----------176
5.2 MATERIALS AND METHODS ----------177
5.2.1 *Begomovirus* and Satellite Infectious Clones ----------177
5.2.2 Treatments ----------177
5.2.3 Geminivirus and Satellite Infections ----------178
5.2.4 Scoring for Disease Symptoms ----------179
5.2.5 Total DNA Purification and DNA Accumulation Analysis ----------179
5.3 RESULTS ----------180
5.3.1 Symptom Monitoring ----------180
5.3.2 Southern Blot Analysis ----------181
5.4 DISCUSSION ----------182
5.5 CONCLUSION ----------183
5.6 REFERENCES----------183

CHAPTER 6: **CONCLUSIONS AND RECOMMENDATIONS** ----------185
6.1 Conclusions and significance of findings ----------185
6.2 Recommendations ----------189
6.2.1 Strengthening Plant Quarantine ----------190
6.2.2 Use of Genetic Engineering for Protection----------191

6.2.3 Marker-Assisted Detection of Resistance-Associated Genes in Cassava Germplasm----------191
6.3 References----------192

List of Figures

Page

Figure 2.1 Edible tubers produced by cassava plant, cv. Kiroba ---------- 19
Figure 2.2 Some of the commonest food uses of cassava ---------- 20
Figure 2.3 Cassava field with CMD ---------- 27
Figure 2.4 The sweetpotato whitefly *Bemisia tabaci* Genn. ---------- 27
Figure 2.5 Transfer and use of cuttings for planting new crop as one cause of CMD spread ---------- 28
Figure 2.6 Typical CMD symptoms ---------- 31
Figure 2.7 ACMV geminate particles ---------- 36
Figure 2.8 Genome structures of viruses in the Family *Geminiviridae* ---------- 37
Figure 2.9 Simplified version of geminivirus DNA replication cycle ---------- 39
Figure 2.10 Genome structure of ACMV ---------- 40
Figure 2.11 Genome organization of a DNA-ß satellite ---------- 56
Figure 2.12 Representative betasatellites showing conserved structures ---------- 57
Figure 2.13 Genome organization of SatDNA-II and SatDNA-III ---------- 69
Figure 3.1 Stem cuttings planted in screen house ---------- 96
Figure 3.2 Map of Lake Victoria cassava-growing zone to show survey sites -----105
Figure 3.3 Map of Eastern cassava-growing zone to show survey sites ----------107
Figure 3.4 Appearance of plant coded MU06S1 in the field and screen house -----109
Figure 3.5 Differential symptom expression in plant coded KN14S1 ----------110
Figure 3.6a PCR amplification of EACMV-TZ core coat protein ----------114
Figure 3.6b PCR amplification of ACMV ----------115
Figure 3.7 Amplification of full length SatDNA-II from Lake and Eastern zones -115
Figure 3.8 Amplification of full length SatDNA-III ----------116
Figure 3.9 Phylogenetic tree of EACMV-TZ [TZ113] and EACMV-TZ [TZ108] -116

Figure 3.10 Survey sites for integrated forms of satellites ----------117
Figure 3.11 Amplification of SatDNA-II and SatDNA-III fragments ----------118
Figure 3.12 PCR results showing double bands amplified DNA bands ----------118
Figure 3.13 Integrated SatDNA-II sequence amplification ----------119
Figure 3.14 Expected bands for integrated SatDNA-II ----------120
Figure 3.15 Additional bands on agarose gel in SatDNA-III ----------120
Figure 3.16 Phylogenetic tree to show sequence relationship among Tanzanian SatDNA-II isolates ----------121
Figure 3.17 Phylogenetic tree to show sequence relationship among Tanzanian SatDNA-III isolates ----------123
Figure 4.1 Loading cartridges into cartridge holder ----------143
Figure 4.2 Bombardment aimed to hit leaf blades and internodes ----------144
Figure 4.3 Mechanical damage on growing tips in cassava line AR40-10 ----------148
Figure 4.4 Leaf curling and distortion in cassava plant line AR37-92 ----------149
Figure 4.5 Naturally-occurring EACMV-TZ + SatDNA-infected plant collected from a field at Yombo Vituka, Dar es Salaam ----------149
Figure 4.6 Cassava lines showing CMD symptoms 5 days post-inoculation --------150
Figure 4.7 Mosaic symptoms on cassava plant of line AR21-2 one week after bombardment with EACMV-TZ + SatDNA-III ----------150
Figure 4.8 Cassava plant of line AR21-2 recovering from infection due to bombardment with EACMV-TZ + SatDNA-III ----------151
Figure 4.9 PCR amplification of *βC1* gene from experimentally biolistically-inoculated plants ----------156
Figure 4.10 Symptomatic leaf and asymptomatic leaf occurring in the same plant --157
Figure 4.11 Amplification of full length SatDNA-II from bombarded plants -------158
Figure 4.12 Re-amplification of SatDNA-II from total DNA ----------158
Figure 4.13 Amplification of SatDNA-II from total DNA 42 dpi ----------159
Figure 4.14 Amplification of SatDNA-II fragments using EST-based primers ------159
Figure 4.15 Amplification of SatDNA-III fragments ----------160
Figure 4.16 Results of Southern blot hybridization ----------163

Figure 4.17 Southern hybridization for SatDNA-III ----------163
Figure 4.18 Southern hybridization for EACMV-UG2 and EACMV-TZ ----------164
Figure 5.1 Insect-free chambers used to grow test plants ----------179
Figure 5.2 *Nicotiana benthamiana* plants inoculated with different virus/satellite combination 14 days after inoculation ----------180
Figure 5.3 Southern hybridization blots showing levels of DNA accumulation in *N. benthamiana* inoculated with EII III 24 *hpi* ----------181

List of Tables

		Page
Table 2.1	Top 10 cassava producing countries 2003 – 2008	21
Table 2.2	Twelve most earning staples in 7 countries of Sub-Saharan Africa	22
Table 2.3	Production of staple tuber crops 2003 – 2008	22
Table 2.4	Cassava production and use in 1993 and projection for 2020	23
Table 2.5	Proportion of households cultivating and planted area of food crops	24
Table 2.6	Cassava production trend in Tanzania	24
Table 2.7	Summary of CMD research milestones	30
Table 2.8	Cassava mosaic disease symptoms scale 1 – 5	32
Table 3.1	Cassava mosaic disease symptoms scale 1 – 5	95
Table 3.2	Districts and number of fields covered during survey	104
Table 3.3	CMD symptoms in Lake Victoria zone	106
Table 3.4	Average CMD incidence and symptom severity scores in Eastern zone	108
Table 3.5	Symptom development of cuttings in the screenhouse	111
Table 3.6	PCR results for EACMV-TZ, ACMV, SatDNA-II and SatDNA-III	112
Table 3.7	Distance table for SatDNA-II isolates	122
Table 3.8	Distance table for SatDNA-II isolates	124
Table 4.1	Cassava lines used in CMD resistance evaluation	140
Table 4.2	List of primers used to detect viral and satellite sequences	145
Table 4.3	Concentration and purity of plasmids before transformation	147
Table 4.4	Concentration and purity of plasmids after multiplication	148
Table 4.5	Symptom observations following biolistic inoculations	151
Table 4.6	Plasmid mixes means for three stages of observation	154
Table 4.7	ANOVA table for the observed variables ($P \leq 0.05$)	154
Table 4.8	Mean separation for disease development stages and virus/satellite mixes	155
Table 4.9	Summary of PCR amplification results	161

Acronyms and Abbreviations

asl	_	above sea level	DIG	_	Digoxigenin
AYVB	_	*Ageratum yellow vein betasatellite*	*dpi*	_	days post inoculation
AYVD	–	*Ageratum* yellow vein disease	DLR	–	DNA loading ratio
AYVV	–	*Ageratum yellow vein virus*	DSO	–	Double strand origin
ca.	-	*circa* =about	EA	–	EACMV-TZ DNA-A Component
CLCuD	–	Cotton leaf curl disease	EB	–	EACMV-TZ DNA-B component
CLCuMuB	–	*Cotton leaf curl Multan betasatellite*	EII	–	EACMV-TZ + SatDNA-II
CLCuMuV	–	*Cotton leaf curl Multan virus*	EIII	–	EACMV-TZ + SatDNA-III
CMG	–	Cassava mosaic geminivirus	EII III	_	EACMV-TZ + SatDNA-II+ SatDNA-III
CR	–	Common region	E0	–	EACMV-TZ alone
Cv	–	Cultivar	II0	_	SatDNA-II alone
DAP	–	Days after planting	III0	_	SatDNA-III alone
			EST	–	Expressed sequence tag

EpYVB	_	*Eupatorium yellow vein betasatellite*	*rpm*	–	revolutions per minute
EpYVD	–	*Eupatorium* yellow vein disease	ROS	-	Reactive oxygen species
EpYVV	–	*Eupatorium yellow vein virus*	SAR	_	Systemic acquired resistance
hpi	_	hours post inoculation	SatDNA-II	–	Satellite DNA-II
Hsp	_	Heat shock proteins	SatDNA-III	–	Satellite DNA-III
ibid.	_	*ibidem* = same place	SCR	–	satellite conserved region
kPa	–	kilopascal	sdH_2O	_	Sterile distilled water
LB	_	Lysogeny Broth	SOC	–	super optimal broth with catabolite repression
MLQ	–	Microcarrier loading quantity	*Syn.*	_	synonymous
Min	-	Minute	TNA	_	Total nucleic acid
nts	_	nucleotides	ToLCB	–	*Tomato leaf curl betasatellite*
NW	_	New World	TYLCCNB	–	*Tomato yellow leaf curl China betasatellite*
OW	_	Old World			
psi	–	pound per square inch			
RCR	-	rolling circle replication			

TYLCCNV		–*Tomato yellow leaf curl China virus*
UVA	_	Ultraviolet A
UII	–	EACMV-UG2 + SatDNA-II
UIII	–	EACMV-UG2 + SatDNA- III
U0	–	EACMV-UG2 alone
wpi	_	weeks post inoculation

List of Appendices

		Page
Appendix 3.1	The 2817 bp - pCR8/GW/TOPO Sequence	194
Appendix 3.2	Sequences of isolates obtained from alignment report	195
Appendix 3.3	The 27 homologous sequences for begomoviruses	195
Appendix 3.4	Occurrence of integrated forms of SatDNA-II and SatDNA-II in Tanzania	196
Appendix 3.5	Protocol for the cloning of PCR products	201
Appendix 3.6	Portion of SatDNA-II isolates alignment report to show GC-rich regions	204
Appendix 3.7	Portion of SatDNA-II isolates alignment report to show poly (A) signals	205
Appendix 3.8	Portion of SatDNA-II isolates alignment report to show TATA box	206
Appendix 3.9	Portion of SatDNA-III isolates alignment report to show GC-rich Regions	207
Appendix 4.1	Conditions which influenced growth of stakes in the green house	208
Appendix 4.2	Growth trend of plants toward stage of bombardment	212
Appendix 4.3	Cell transformation protocol	215
Appendix 4.4	Plasmid extraction protocol	216
Appendix 4.5	Preparation of plasmid mixes for bombardment	217
Appendix 4.6	Protocol for particle bombardment	218
Appendix 4.7	CTAB total nucleic acid extraction Protocol	220
Appendix 4.8	Protocol for Southern blot analysis	222
Appendix 6.1	Members of Genus: *Begomovirus-associated Satellite DNA*	227

CHAPTER 1

GENERAL INTRODUCTION AND OBJECTIVES

1.1 OVERVIEW OF PLANT VIRUS DISEASES IN AFRICA

Plant viruses present a threat to crop production and compromise livelihood strategies far more than other plant pathogens combined. Four emerging plant viruses - *Rice yellow mottle virus*, RYMV (Genus *Sobemovirus*), *Banana streak virus*, BSV (Genus *Badnavirus*) (Lockhart, 1995; Harper *et al*., 2004), *Maize streak virus*, MSV (Genus *Mastrevirus*) (Shepherd *et al*., 2010) and cassava mosaic begomoviruses (CMBs) (Genus *Begomovirus*) have particularly affected subsistence agriculture and rural livelihoods in tropical areas (Kouassi *et al*., 2005; Fargette *et al*., 2006; Gayral *et al*., 2008; Patil and Fauquet, 2009).

Begomoviruses have long been a problem in plant disease epidemics in many horticultural and field crops such as bean, cassava, cotton, tomato and cucurbits (Morales and Anderson, 2001; Jones, 2003). FAO statistics (FAO, 2008) show that crops affected by these viruses (rice, banana and plantain, maize and cassava) rank high as food staples and income providers in seven countries of Sub-Saharan Africa – Tanzania, Malawi, Kenya, Mozambique, Rwanda, Zambia and Uganda. Several mechanisms constitute the basis of epidemiology in most of these viruses: recombination between viruses themselves and with satellites, synergism between virus species, genome integration of the virus (Gayral *et al*., 2008), new vector biotypes, host adaptation and long-distance dispersal of the whitefly vectors. As a group, cassava mosaic begomoviruses have continued to cause major negative impact on the production of cassava.

This negative impact is manifested in three distinct areas:

(i) Yield reduction as a direct consequence of virus infection,
(ii) Limitation on the multiplication and distribution of vegetative planting material and,
(iii) Impact on conventional breeding programs.

Among the starchy staples indicated above, cassava is perhaps the most important (Tonukari, 2004). This is because cassava gives a carbohydrate production, which is about 40% higher than rice and 25% more than maize. In the five well-known tuber crops (potato, cassava, sweet potato, yams and taro), which together account for 90% of the World's total staple tuber production, cassava ranks second after potato and accounts for 26% (Shewry, 2003). Cassava demand is estimated to grow at 2.0% annually for food and 1.6% per year for feed in developing countries.

In Africa, the largest producers are Nigeria (36.6% of total African cassava production), DRC (19%), Ghana (8%), Mozambique (7%) and Tanzania (6%) (Hillocks, 2002; FAO, 2009). However, it is in the DRC and Tanzania (Lazaro *et al*., 2007) where cassava is the most important crop to the largest proportion of farming households. Considering the importance of cassava, some of the mechanisms through which viruses cause havoc to production are explored in a case study of cassava mosaic begomoviruses in Tanzania.

1.2 CASSAVA MOSAIC DISEASE AS A CONSTRAINT

Cassava mosaic disease (CMD) has consistently remained a single most important limiting factor to cassava production worldwide (Hillocks, 2002; Legg and Fauquet, 2004). The disease is caused by viruses of genus *Begomovirus* in the family *Geminiviridae* (Fauquet and Stanley, 2003). The disease is spread by whiteflies, *Bemisia tabaci* Genn. (*Homoptera*: *Aleyroidae*) and infected stem cuttings. CMD occurs all across Sub-Saharan Africa where cassava is grown. Total losses due to the disease are estimated at 15 – 24% (Hahn *et al*., 1980) and can reach 90% in some cultivars. In Kenya for example, yield loss incurred in 1998 due to CMD was estimated at 15,000 tons valued at US$ 10 million (Obiero *et al*., 2007). This

constitutes a substantial gap in food production, considering the ever increasing population growth.

As need for higher cassava production is on the increase, cassava is rapidly being transformed from a traditional staple into a market-oriented commodity to meet non-food industrial needs especially production of starch (Tonukari, 2004). Alliance for a Green Revolution in Africa (AGRA) for example, has launched a program to help 30,000 smallholder farmers in Kenya, Uganda and Tanzania increase cassava for food and industrial use by processing it at village level (The Guardian, 15 June 2010).

1.3 WHY CMD REMAINS A PROBLEM

Since CMD was first reported (Warburg, 1894), immense effort has been taken to control its effects. Such efforts have included phytosanitation and resistance breeding. The first source of resistance to CMD was derived from the Ceara rubber tree, *Manihot glaziovii* Müll. Arg. (Jennings, 1976). It was the most widely deployed source of resistance and was represented in farmers' fields by clones of the Tropical *Manihot* Species (TMS) series.

Despite the efforts, CMD has remained consistently a problem because of:

(i) The occurrence of mixed infections of viruses, which lead to synergistic infections;
(ii) The molecular characteristics of begomoviruses, especially the ability to form more virulent strains;
(iii) The propensity for the begomoviruses to associate with satellites through recombination and reassortment;
(iv) The possibility for the begomoviruses and/or associated satellites to integrate into host genome;
(v) The widespread, high fecundity of the whitefly vector and its ability to form biotypes;
(vi) The varied tropical climate which favours evolution of the whiteflies as well as the viruses; and

(vii) The widespread habit of farmers to use unimproved and often infected planting materials.

Nevertheless, plant resistance still remains a first-hand tool in the on-going efforts to control CMD and renew farmers' hopes.

1.4 ROLE OF PLANT RESISTANCE IN CURBING CMD

For many years, the approach to the control of plant virus diseases and CMD in particular, has been the deployment of resistant varieties (Thresh *et al*., 1997). These varieties have been developed by either classical breeding or by genetic engineering. However, the resistance-breaking begomoviruses have adversely complicated these approaches. Epidemiological models indicate that host resistance is always associated with selection of viruses with higher virus titres. Resistance is also associated with viruses for which virulent strains exist (*ibid.*). Higher virus population levels in hosts, coupled with mixed infections and the involvement of satellites increase the rate of evolution of begomoviruses. This process also provides a potential reservoir from which viruses can spread to neighboring crops and weeds.

Many instances exist in which resistance as part of integrated pest management (IPM) strategy has brought epidemics under control. This includes cassava mosaic epidemics in Madagascar in the 1930s (Legg *et al*., 2006), in Uganda in the 1990s and the recent use of tomato lines resistant to *Tomato leaf curl virus*, TLCV (Syn. *Indian tomato leaf curl virus*, ITmLCV) in India. Breeding for begomovirus resistance in cassava requires screening under conditions of high disease pressure. The screening must include standardization of inoculation protocols that address the phenomenon of reversion. The screening must also be done in suitable location to avoid the occurrence of symptomless plants from infected cuttings. The development of molecular markers linked to resistance genes has been valuable as it allows screening of symptomless propagation material.

1.5 SATELLITES AND THEIR ROLE IN BEGOMOVIRUS EPIDEMIOLOGY

As stated above, recombination events take place between virus genome components as well as between begomovirus DNA-A and satellite DNA circles. Recently, numerous begomoviruses in the Old World have been identified that require satellites to induce authentic symptoms in the hosts from which they were isolated (Briddon and Stanley, 2006). A satellite is a sub-viral agent composed of nucleic acid that depends on the co-infection of a host cell with a helper or master virus for their multiplication.

The classical definition of a satellite in plant virology involves three main features (Roossinck *et al*., 1992):

(i) They are not capable of replicating (within the host cell) in absence of the helper virus,
(ii) They are not required for the life cycle of the helper virus, and
(iii) They share little or no sequence similarity with the helper virus genome.

By virtue of their constitution and activities in host cells, satellites often modulate the replication of their helper viruses. The symptoms caused by the helper virus/satellite co-infection usually differ dramatically from those of the helper virus alone. In order for the begomoviruses to effectively interact with satellites and influence epidemiology, there must be some structural similarities among them.

Nawaz-ul-Rehman and Fauquet (2009) suggest four common features between begomoviruses and satellites:

(i) Rolling circle replication (RCR) mechanism (true for betasatellites),
(ii) Presence of a stem loop containing origin of replication and other sequence elements required for DNA replication,
(iii) Similar genetic architecture, localization, and length of individual genes at specific locus,
(iv) Replication in nuclei of differentiated phloem-associated cells.

Rolling circle replication is induced by an initiator protein encoded by the geminivirus DNA. It nicks one strand of the dsDNA molecule at a site called double strand origin (DSO). This is perhaps one of the reasons why satellites require helper viruses for their productivity. Thus, to successfully invade new hosts, break host resistance and move virus particles within and between plants, begomoviruses and their satellites have evolved a coordinated network of protein interactions, showing a possible evolutionary path (Cuong *et al*., 2008).

The first report for a viral satellite to be associated with a DNA virus was a 682-nt circular DNA satellite associated with *Tomato leaf curl virus* (TLCV) infection in Northern Australia (Dry, *et al*., 1997). TLCV sat-DNA is strictly dependent for replication on the helper virus *Rep;* it is encapsidated by TLCV coat protein; has no significant open reading frames (ORFs); and shows no sequence similarity to the 2766-nt helper-virus genome. TLCV sat-DNA is also trans-replicated by other taxonomically distinct geminiviruses, including *Tomato yellow leaf curl virus* (TYLCV), *African cassava mosaic virus* (ACMV), and *Beet curly top virus* (BCTV). TLCV sat-DNA has no apparent effect on symptom expression.

Begomoviruses fall in two groups based on geographical origin – Old World (Europe, Africa, Asia and Australia) and New World (North, Central and South America) begomoviruses. Most Old World monopartite begomoviruses are associated with one or more betasatellite DNAs (DNA-ß), which are required for the induction of typical disease symptoms.

DNA-β molecules are widespread in the Old World and all contain three conserved features (Briddon *et al*., 2003):

(i) They all encode a single gene *βC1*,
(ii) They have satellite conserved region (SCR), and
(iii) They have an adenine-rich region approximately 370 to 420-nt upstream of the SCR.

Betasatellite-requiring begomoviruses are now known to outnumber the bipartites and truly monopartites in the Old World. Since they were first identified in 2000, over 260 full length sequences have been deposited in databases, and the number is said to increase. Despite their recent discovery, DNA-ß satellites have existed for many centuries. The DNA-ß satellite associated with *Eupatorium yellow vein virus* (EpYVV) has been linked to disease symptoms described some 1250 years ago.

The majority of begomoviruses, which associate with DNA-ß have two things in common:

(i) They are poorly infectious to the hosts from which they were isolated, and
(ii) They induce atypical symptoms.

In addition to the begomoviruses and betasatellites, Old World begomovirus disease complexes are also associated with a third single stranded DNA component termed alphasatellites (*syn*. DNA-1). These are satellite-like, single stranded DNA molecules of approximately 1370-nt associated with begomoviruses that require the satellite molecule DNA-β to induce authentic disease symptoms in some hosts. Unlike betasatellites, the alphasatellites encode a replication-associated protein (Rep), which is required to initiate rolling circle replication (Shahid *et al*., 2009). In common with betasatellites, the alphasatellites have the adenine-rich region (*ibid*.). Disease complexes involving alphasatellites include Okra leaf curl disease (OLCD) observed in Mali and Pakistan (Kon *et al*., 2009) and Cotton leaf curl disease in Pakistan (Shahid *et al*., 2009).

1.6 SATELLITE DNA MOLECULES

Along the evolutionary line of satellites, there has been a recent discovery of satellites in Tanzania. They are satellite DNA molecules that have been shown to break down resistance to CMD (Ndunguru *et al*., 2008), even in cultivars which were thought to have modest levels of resistance such as TMS 60444. They are typical satellites - they cannot replicate on their own and they do not share any sequence similarity with the helper cassava begomoviruses with which they are associated. The satellites have been designated Satellite DNA II (SatDNA-II) and Satellite DNA III (SatDNA-III), and have sizes of 1032 and 1209 bp, respectively.

SatDNA-II was responsible for the very severe CMD symptoms in cassava (line TME 3) characterized by leaf distortion, yellowing and mosaic while SatDNA-III was responsible for the observed unique symptoms in cassava characterized by severe leaf narrowing and prominent leaf yellowing. SatDNA-II breaks resistance to CMD in cassava line TME 3 and other cultivars. SatDNA-III breaks resistance to EACMV-UG2. The satellites were mechanically transmitted to *Nicotiana benthamiana* but not to cassava. However, biolistic inoculation of SatDNA-II and SatDNA-III infectious clones produced symptom modulation of their associated helper begomoviruses in both cassava and *N. benthamiana*.

In terms of structure, the satellite molecules do not have the SCR containing nanonucleotide sequences TAATATTAC as in betasatellites or TAGTATTAC, which occurs in alphasatellites. Further, they do not have the adenine-rich region; instead, they have a GC-rich region. In another development (unpublished) SatDNA-II and SatDNA-III are thought to have parts of their sequences integrated into cassava genome. This follows the findings that the satellites sequences were able to find some unigenes resembling the satellites when blasted in cassava expressed sequence tags (EST). The length of the alignments varied from 112 to 346 nucleotides for SatDNA-II and 65 to 226 for SatDNA-III. This suggested the possibility of integration of the satellite sequences in host plant genome.
The process of integration has considerable implications for plant genome evolution:

(i) The integrated viral DNA could act as an insertional mutagen. The foreign DNA sequence disrupts the expression of the gene into which it is inserted but also acts as a marker for subsequent identification of the mutation.
(ii) The sequence could contribute strong constitutive promoters to neigbouring plant genes,
(iii) The DNA could accumulate to generate new repetitive sequence family.

Thus, the incorporation of viral sequences into host genome can have functional implications for the virus-host interactions. It can render the host susceptible or resistant to diseases (Bertsch *et al*., 2009).

There are two things, which are new about the recently discovered satellites:

(i) The association between satellite DNAs and bipartite begomoviruses originating from the Old World;
(ii) The occurrence of begomovirus/satellite association in cassava.

1.7 PROBLEM JUSTIFICATION

Since the discovery of CMD in 1894 (Warburg, 1894) to the year 2004, three main factors were known to aggravate CMD situation in Tanzania:

(i) Mixed infection of ACMV and EACMV, leading to synergistic severe CMD,
(ii) The discovery of *East African cassava mosaic Zanzibar virus* (EACMZV), and
(iii) The emergence of ACMV-EACMV recombinant designated EACMV-UG2.

The occurrence of SatDNA-II and SatDNA-III constitutes a fourth factor in the CMD pandemic. The discovery and characteristic features of these satellites clearly elicit a number of questions to plant virologists and breeders alike, regarding their role in begomovirus epidemiology. It has been shown that the satellites cannot be mechanically transmitted to cassava. It is not known whether the satellites undergo some changes in the whitefly vector that influence their replication in cassava. The satellites have neither origin of replication nor SCR containing the nicking site as in the known satellites like betasatellites, posing a question as to how they are replicated. Furthermore, the satellites do not contain the Rep ORF (*AC1*) associated with replication machinery of viruses, yet they aggravate CMD severity, similar to ACMV-EACMV mixed infections. This clearly shows that there is an interaction between the begomoviruses and the satellites and warrants a need for a thorough study.

1.8 PROBLEM HYPOTHESES

In composing pertinent hypotheses about possible ways in which the novel satellites could be involved in molecular epidemiology of cassava mosaic begomoviruses, two important aspects have been considered:

(i) SatDNA-II and SatDNA-III have been confirmed to be associated with resistance-breaking following Koch's postulates as amended for plant viral pathogens;

(ii) SatDNA-II and SatDNA-III have been shown to be satellites consistent with the classical definition in plant virology.

In the study on involvement of betasatellites in a begomovirus disease in West Africa (Chen *et al*., 2009), it was suggested that the findings could indicate the role of satellites in other begomovirus diseases. Thus, although the satellites (SatDNA-II and SatDNA-III) do not necessarily show the same properties as the previously described satellites, there could be some discernible similarities with the previously described satellites.

These similarities may include the following, from which present study hypotheses are deduced:

(a) In the context of the association with SatDNA-II and SatDNA-III, **the helper begomoviruses are likely to be occurring in mixed infections**. This is because most of the previously reported satellites' helper begomoviruses occurred in mixed infections with others. A recent and similar epidemic of a resistance-breaking strain of Okra leaf curl disease (OLCD) observed in Mali and Pakistan (Kon *et al*., 2009) was due to multiple virus infections. In the present study, mixed infections between ACMV and EACMV have been observed under field conditions as previously reported (Sserubombwe *et al*., 2008; Fondong *et al*., 2000);

(b) The predicted multiple infections provide pre-conditions for recombination, allowing for the appearance of more severe begomovirus strains or new species. **The satellites are then thought to enhance the viability of the recombinant(s), thereby exacerbating disease**. This situation has since been reported in the previously described disease situation in OLCD in Mali (Kon *et al*., 2009). In CMD in Tanzania, one of the helper begomovirus is indeed a recombinant, EACMV-UG2;

(c) One mechanism by which betasatellites enhance disease situation is by stimulating the division of terminally differentiated phloem-associated cells,

thereby generating more cells for viral infection (Chen *et al*., 2009). Considering the fact that replication of betasatellites is dependent on the begomovirus-encoded initiator protein (Nawaz-ul-Rehman and Fauquet, 2009), **the SatDNAs (SatDNA-II and SatDNA-III) could be replicated in the same manner, strongly supporting the thinking that they could stimulate cell division**;

(d) The previously reported satellites together with their respective helper begomoviruses replicated through RCR using the begomovirus-encoded Rep. In the present situation, **it is likely that one of the motifs donated by helper begomovirus to the satellite (SatDNA-II and SatDNA-III) is *AC1*/Rep, which in turn supports satellites replication**;

(e) In some previously reported satellite/begomovirus disease complexes, the helper virus was often found to donate its CR to the satellite in exchange with their SCR (Navas-Castillo *et al*., 2000; Chen *et al*., 2009). CR recombinant betasatellites include *Ageratum yellow vein betasatellite* (AYVB), *Tomato yellow leaf curl China betasatellite* (TYLCCNB) and *Cotton leaf curl Multan betasatellite* (CLCuMuB). The SCR contains a motif GGTDKN (D = A, G, T; K= G, T; N = A, C, G, T), resembling the conserved begomovirus iteron sequence. **These sequences possibly support the formation of viable recombinants**. Though the satellites in the present study do not have SCR *per se*, the same mechanism could apply;

(f) Most of the betasatellites involved in resistance-breaking epidemics owe their success to the ability of being trans-replicated by different viruses (Nawaz-ul-Rehman and Fauquet, 2009). Examples include *Tomato leaf curl betasatellite* (ToLCB) trans-replicated by ACMV and BCTV and CLCuMuB by AYVV (although AYVB cannot be trans-replicated by CLCuMuV). **The propensity of the present satellites to break resistance of cassava to CMD could be attributed to trans-replication by other cassava-infecting begomovirus strains yet to be described**;

(g) Some well-known begomovirus epidemics described are characterized by co-infection of viruses and satellites. For example, OLCD in Mali and Pakistan (Kon *et al*., 2009) involves okra virus-1, okra-virus-2, CLCuGB and an alphasatellite. A similar situation is reported for CLCuD in Pakistan and AYVD in Singapore. In the present study, **co-infection of EACMV and SatDNA-II**

and SatDNA-III has been demonstrated in Tanzania and could be the cause of severe CMD; and

(h) There could also be a possibility that the satellites are integrated in the cassava genome. The sequence integration process can render the host susceptible or resistant to diseases (Bertsch *et al*., 2009). If evidence for integration is sought and confirmed in the present study, **the integrated satellite sequence may influence the interaction between the helper cassava begomovirus (EACMV) and the host (cassava)**.

In the present satellite-involving CMD, the most critical issue is the tendency of the satellites to breakdown resistance to the disease despite breeders' efforts to improve resistance. The previously described satellites are diverse; beliefs that even the satellites in present study could be diverse. There is thus a need to find out the diversity of these satellite DNA molecules and their spread at least within Tanzania. There is also a need to find out whether there is any interaction between the satellites and the begomovirus and if it occurs at gene- or DNA- or protein-level. If the interaction is at gene-level, it is important to find out which genes are involved. Where the interaction is at DNA-level then on the part of the virus - is it through DNA-A or DNA-B or both? It may also be important to know if SatDNA-II and SatDNA-III encode for any viable protein. Finally, it may be imperative to investigate if the satellites are integrated into cassava and the nature of that integration.

Furthermore, does the interaction enhance or ameliorate CMD symptom severity? It is known for example, that ACMV enhances symptom severity in plants inoculated with EACMV through the transcriptional activator protein (*AC2*) in which the later suppresses posttranscriptional gene silencing (PTGS). There is a need to find out whether SatDNA-II or SatDNA-III enhances symptoms induced by the helper virus by activating the promoter of its *AC2* to result into its increased replication. Is there molecular interaction between begomovirus *AC2* and the satellites DNA-DNA, DNA-protein or protein-protein? Finally, but indeed important, is the need to find out if trans-replication of SatDNA-II and SatDNA-III occurs in begomoviruses other than cassava mosaic begomoviruses such as TYLCV, CLCV, TGMV, TGYMV, which encode *AC2* as well.

1.9 OBJECTIVES OF THE STUDY

The goal of this study is to appraise the nature of resistance in the indigenous cassava cultivars in the face of the resistance-breaking satellites. The study title of the thesis is presented as

'Influence of satellite DNA molecules on severity of cassava begomoviruses and the break down of resistance to cassava mosaic disease in Tanzania'.

Within the overall objective, there are three aims:

1. To identify and map the distribution of cassava geminiviruses and satellite DNA molecules in major cassava-producing areas of Tanzania;
2. To screen cassava cultivars for disease response to begomoviruses and satellite DNA molecules; and
3. To determine the nature of interaction between begomovirus DNAs and Satellite DNA molecules in *Nicotiana benthamiana*.

1.10 THESIS ORGANIZATION, SOURCES OF DATA AND METHODOLOGY

This thesis is organized in chapters, which each chapter covers an objective. The first chapter (**Chapter 1**) covers a general introduction and adequate acknowledgement of secondary data in a way of critical review of previous investigations and presents the objectives of the study. Topics related to the present study are presented as Literature Review (**Chapter 2**). The review has covered Cassava and Livelihood; CMD – Importance, Spread and Prevalence; Etiology of CMD; Involvement of Satellites in Begomovirus Pathogenicity; Integration of Viral and Satellite Sequences into Host Genome; Association of Satellites with Old World Bipartite Begomoviruses; CMD Management; and Resistance as an Intervention to CMD. The chapter also covers a comprehensive study of the influence of satellite DNA molecules on the severity of cassava mosaic viruses in Tanzania.

Chapter 3 is on the survey and etiology of cassava mosaic virus isolates and satellite DNA molecules in major cassava growing areas of Tanzania. The study under this chapter comprised primary field, screen house and laboratory data. **Chapter 4** involved the screening of cassava cultivars for resistance or susceptibility to EACMV-TZ or EACMV-UG2. The screening also involved SatDNA-II and SatDNA-III. To achieve the screening, some cassava lines were planted in a screenhouse and particle bombardment with infectious clones of satellites (SatDNA-II and SatDNA-III) and helper viruses (EACMV-TZ and EACMV-UG2). This was followed by visual observation of their response in the green house by scoring as well as laboratory investigations by PCR and Southern blot analyses.

Chapter 5 is on the survey at the probable nature of the interaction between begomovirus DNAs and Satellite DNA molecules in susceptible cassava cultivars. It involved the planting of test plants (*Nicotiana benthamiana*) in screen house and inoculating them with different mixtures of satellite infectious clones as well as infectious clones of the helper viruses. The thesis also comprises a conclusive chapter (**Chapter 6**) which gives a summary of the findings from the research chapters; the implications of the findings in the awe of managing cassava mosaic epidemic in Tanzania are indicated; and finally, the chapter proposes future work on the involvement of satellites on cassava begomovirus epidemiology. At the end of the thesis are references, list of figures and appendices.

1.11 REFERENCES

1. Bertsch C., Beuve M., Dolja V.V., Wirth M., Pelsy F., Herrbach E. and Lemaire O. (2009) Retention of the virus-derived sequences in the nuclear genome of grapevine as a potential pathway to virus resistance. *Biology Direct* 4: 21-24.
2. Briddon R.W., Bull S.E., Amin I., Idris A.M., Mansoor S., Bedford I.D., Dhawan P., Rishi N., Siwatch S.S., Abdel-Salam A.M., Brown J.K., Zafar Y., and Markham P.G. (2003) Diversity of DNA beta, a satellite molecule associated with some monopartite begomoviruses. *Virology* 312(1): 106–121.
3. Briddon R.W. and Stanley J. (2006) Subviral agents associated with plant single stranded DNA viruses. *J. Virol.* 344(1): 198-210.
4. Carter S. and Jones P. (1993) A model of the distribution of cassava in Africa. *Appl.Geogr*. 13: 353-371.
5. Chen L., Rojas M., Kon T., Gamby K., Xoconostle-Cazares and Gilbertson R.L. (2009) A severe symptom phenotype in tomato in Mali is associated by a reassortment between a novel recombinant begomovirus (TYLCMLV) and a betasatellite. *Molecular Plant Pathology* 10(3): 415–430.
6. Cuong H.A, Coombs S., Revill P., Harding R,. Man V. and Dale J. (2008) Molecular characterization of begomoviruses and DNA satellites from Vietnam: additional evidence that the New World geminiviruses were present in the Old World prior to continental separation. *Journal of General Virology* 89(1): 312-326.
7. Dry I.B., Krake L.R., Rigden J.E. and Rezaian M.A. (1997) A novel subviral agent associated with a geminivirus: The first report of a DNA satellite. Division of Horticulture, Commonwealth Scientific and Industrial Research Organization, Hartley Grove, Urrbrae, South Australia 5064, Australia.
8. FAO (2008) FAO statistics (FAOSTAT).
9. FAO (2009) FAO statistics (FAOSTAT).
10. Fauquet C.M. and Stanley J. (2003) Geminivirus classification and nomenclature: Progress and problems. *Annals of Applied Biology 142:165-189*.

11. Fargette D., Konate G., Fauquet C., Muller E., Peterschmitt M. and Thresh J.M. (2006) Molecular ecology and emergence of tropical plant viruses. *Ann.Rev. Phytopath*. 44: 235-260.
12. Fondong V.N., Pita J.S., Rey M.E.C., DeKochko, Beachy R.N. and Fauquet C.M. (2000) Evidence of synergism between *African cassava mosaic virus* and a new double recombinant geminivirus infecting cassava in Cameroon. *J. Gen. Virol*., 81: 287-297.
13. Gayral P., Noa-Carrazana J., Lescot M., Lheureux F., Lockhart B.E.L., Matsumoto T., Piffanelli P. and Iskra-Caruana (2008) A single *Banana streak virus* integration event in the banana genome as the origin of infectious endogenous pararetrovirus (EPRV). *J. Virol.* Doi: 10.1128/JVI.00212-08.
14. Hahn S.K., Terry E.R. and Leuschner K. (1980) Breeding cassava for resistance to cassava mosaic disease. *Euphytica* 29 (3).
15. Harper G., Hull R., Lockhart B. and Olszewski (2002) Viral sequences integrated into plant genomes. *Ann. Rev. Phytopath*. 40: 119-136.
16. Harper G., Hart D., Moult S., and Hull R. (2004) *Banana streak virus* is very diverse in Uganda. *In*: Plant Virus Epidemiology: First steps into the new millennium. *Virus Research* 100 (1): 51–56.
17. Hillocks R.J. (2002) Cassava in Africa. *In* Hillocks R.J., Thresh J.M. and Bellotti A.C. (eds). Cassava: Biology, Production and Utilization. CABI Publishing Oxon UK.
18. Jakowitsch J., Mette M.F., van der Winden J., Matzke M.A. and Matzke A.J.M. (1999) Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. *PNAS* 96 (23): 13241 – 13246.
19. Jennings D.L. (1976) Breeding for resistance to African cassava mosaic pp 39-44. *In* African cassava mosaic: Report of an interdisciplinary workshop held at Muguga, Kenya.19. Jones D.R. (2003) Plant viruses transmitted by whiteflies. *European Journal of Plant Pathology* 109: 195 – 219.
20. Kon T., Rojas M.R., Abdourhamane I.K., and Gilbertson R.L. (2009) Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *J. Gen. Virol.* 90: 1001–1013.

21. Kouassi N.K., N'Guessan P., Albar L., Fauquet C.M., and Brugidou C. (2005) Distribution and characterization of *Rice yellow mottle virus*: A threat to African Farmers. *Plant Disease* 89 (2): 124–133.
22. Legg J.P. and Fauquet C.M. (2004) Cassava mosaic geminiviruses in Africa. *Plant Molecular Biology* 56: 585–599.
23. Legg J.P., Owor B., Sseruwagi P., and Ndunguru J. (2006) Cassava mosaic virus disease in East and Central Africa: Epidemiology and Management of a Regional pandemic. *Advances in Virus Research* 67: 355–418.
24. Lockhart B.E. (1995) *Banana streak badnavirus* infection in *Musa*: Epidemiology, Diagnosis and Control. University of Minnesota.
25. Morales F.J. and Anderson P.K. (2001) The emergence and dissemination of whitefly-transmitted geminiviruses in America Latina. *Archives of Virology* 146: 415–441.
26. Moriones E. and Navas-Castillo J. (2008) Review. Rapid evolution of the population of begomoviruses associated with the tomato yellow leaf curl disease after invasion of a new ecological niche. *Spanish Journal of Agricultural Research* 6: 147-159.
27. Navas-Castillo J., Sanchez-Campos S., Noris E., Louro D., Accotto G.P. and Moriones E. (2000) Natural recombination between *Tomato yellow leaf curl-Is virus* and Tomato *leaf curl virus*. *J. Gen. Virol.* 81: 2797-2801.
28. Nawaz-ul-Rehman M.S. and Fauquet C.M. (2009) Evolution of geminiviruses and satellites. *FEBS Letters* 583: 1825–1832.
29. Ndunguru J. and Tairo F. (2009) Proceedings of the regional cassava virus disease diagnostic project first annual meeting Kigali, Rwanda 19th to 22nd October 2009.
30. Ndunguru J., Fofana B., Legg J.P., Chellappan P., Taylor N., Aveling T., Thompson G. and Fauquet C. (2008) Two novel satellite DNA molecules associated with bipartite cassava mosaic begomoviruses enhancing symptoms and capable of breaking high virus resistance in a cassava landrace. *In* Cassava: Meeting the Challenges of the New Millenium. First Scientific Meeting of the Global Cassava Partnership GCP-I, Ghent Belgium 21-25 July 2008. Pg 141.

31. Obiero H.M., Whyte J.A.B., Legg J.P., Akhwale M.S., Malinga J. and Magut T. (2007) Successful restoration of ccassava production in Western Kenya. Proceedings of the 13th ISTRC Symposium Arusha Tanzania pp. 682–685.
32. Ogbe F.O., Thottappilly G., Dixon A.G.O., Atiri G.I. and Mignouna H.D. (2003) Variants of *East African cassava mosaic virus* and its distribution in double infections with *African cassava mosaic virus* in Nigeria. *Plant Disease* 87 (3): 229–232.
33. Patil B.L. and Fauquet C.M. (2009) Cassava mosaic geminiviruses: actual knowledge and perspectives. *Molecular Plant Pathology* 10(5): 685–701.
34. Roossinck, M.J., Sleat D., and Paluakaitis P. (1992) Satellite RNAs of Plant Viruses: Structures and biological effects. *Microbiological reviews* 56(2): 265–279.
35. Shahid M.S., Ali L. and Wajid S. (2009) The function of the a-rich region of the alphasatelite associated with the cotton leaf curl disease in Pakistan. *EurAsia J. BioSci.* 3: 152–156.
36. Shepherd D.N., Martin D.P., van Der Walt E., Dent K., Varsan A. and Rybicki E.P. (2010) *Maize streak virus*: an old and complex 'emerging' pathogen. *Mol. Plant Pathol*. 11 (1): 1-12.
37. Shewry P.R. (2003) Tuber Storage Proteins. *Annals of Botany* 91: 755–769.
38. Sserubombwe W.S., Briddon R.W., Baguma Y.K., Ssemakula G.N., Bull S.E., Bua A., Alicai T., Omongo C., Otim-Nape G.W. and Stanley J. (2008) Diversity of begomoviruses associated with mosaic disease of cultivated cassava (*Manihot esculenta* Crantz) and its wild relative (*Manihot glaziovii* Mull. Arg.) in Uganda. *J. Gen. Virol.* 89: 1759–1769.
39. Tonukari N.J. (2004) Cassava and future of starch. *Electron. J. Biotechnol*. Vol. 7 No. 1 2004.
40. Thresh J.M., Otim-Nape G.W., Legg J.P. and Fargette D. (1997) African cassava mosaic virus disease: Magnitude of the problem. *African Journal of Root and Tuber Crops* 2: 13-19.
41. Warburg O. (1894) Die kulturpflanzen Usambaras. Mitteilungen aus den Deutschen Schutzgebieten. *Annals of Applied Biology* 25: 780–786.

CHAPTER 2

LITERATURE REVIEW

2.1 CASSAVA AND LIVELIHOOD

2.1.1 General Description and Uses of Cassava

Cassava, *Manihot esculenta* Crantz belongs to the spurge family, *Euphorbiaceae*. Plants of the *Euphorbiaceae* are characterized by lactiferous vessels composed of secretory cells. Most of the family members are herbs but some are shrubs or trees. Leaves are simple but where compound, they are always palmate and never pinnate (Gupta, 2006). Other commercially important plants in the family include the rubber tree *Hevea brasiliensis* Muell. Arg., medicinal plants like bellyache bush *Jatropha gossypiifolia* L., the castor oil plant *Ricinus communis* L., ornamentals such as christmas star *Poinsentia pulcherrima* Willd. and important weeds such as leafy spurge *Euphorbia esula* L. and copperleaf *Acalypha ostryifolia* Riddell.

Cassava is a shrubby, tropical, perennial plant that grows to 1 – 3m tall (Tonukari, 2004) and produces edible root tubers (Figure 2.1). It has a diploid number of 2n = 36 and genome size of 7.5 x 10^8 base pairs (bp) (Hughes, 1996).



Figure 2.1: Edible tubers produced by cassava plant, cv. Kiroba. (**Source:** Mikocheni Agricultural Research Station, Dar es Salaam Tanzania).

Cassava leaves are compound with palmate lamina; they have a petiole and phyllotaxis of the plant is a 2/5 spiral (Alves, 2002). The plant is monoecious with protogynous entomophilous flowers. The mature fruit is a capsule (*ibid*.). Centers of diversity of cassava are Brazil (major) and Central America (minor) (Okogbenin *et al*., 2007). Cassava is known by various other names – manioca, mandioca, manioc, yucca, tapioca, kaspe, and mihogo. The crop has a variety of both food- and non-food uses (Figure 2.2).



Figure 2.2: Some of the commonest food uses of cassava. The three packs on the left is cassava mealy meal used to prepare cassava meal. (**Source:** Mikocheni Agricultural Research Institute, Dar es Salaam Tanzania).

The non-food industrial uses of cassava include production of starch which is used in a wide range of products: paper, textiles and pharmaceuticals (Tonukari, 2004).

2.1.2 Production of Cassava

Grown in over 90 countries (http://www.lswn.it/en/nutrition/news), cassava is planted on about 16 million hectares, with about 50% in Africa, 30% in Asia, and 20% in Latin America. Total World root production is around 152 million tons. The crop provides about 500 million people with food and is therefore important to the livelihoods of poor people in the World (Westby *et al*., 2004). The World's top 10 cassava producing countries are shown in Table 2.1.

Table 2.1: Top 10 World cassava-producing countries 2003 – 2008 (million tons)

Country	2003	2004	2005	2006	2007	2008	Average
Nigeria	36.3	38.84	41.56	45.72	43.41	44.58	41.73
Brazil	21.96	23.92	25.87	26.63	26.54	25.87	25.13
Thailand	19.71	21.44	16.93	22.58	26.91	27.56	22.52
Indonesia	18.52	19.42	19.32	19.98	19.98	21.59	19.80
Congo, D.R.	14.94	14.95	14.97	14.98	15	15.01	14.97
Ghana	10.23	9.738	9.56	9.63	9.65	9.65	9.74
Angola	6.89	8.58	8.6	8.81	8.84	8.84	8.42
India	5.42	5.94	7.46	7.85	8.23	9.05	7.32
Mozambique	6.14	6.41	6.5	6.76	5.03	5.03	5.97
Tanzania, U.R.	3.96	4.44	5.53	6.15	6.6	6.6	5.54

Source: FAO Statistics Division, 2010. D.R = Democratic Republic, U.R. = United Republic.

These FAO statistics show that five countries: Nigeria, Brazil, Thailand, Indonesia and Congo D.R. account for 77% of the World's cassava production. The statistics also show that cassava production has been increasing in most countries in the last five years.

In the Sub-Saharan Africa (SSA), cassava still plays an important role as a food staple as well as income-generating crop. Statistics (FAO, 2008) from seven SSA countries show that cassava ranks high among other livelihood alternatives in the provision of income (Table 2.2).

Table 2.2: Twelve most earning staples in 7 countries of Sub-Saharan Africa (SSA)

Rank	Tanzania	Malawi	Kenya	Mozambique	Rwanda	Zambia	Uganda
1	Cassava	Cassava	Sugarcane	Cassava	Plantains	Sugarcane	Plantains
2	Maize	Maize	Cow milk	Sugarcane	Potato	Maize	Cassava
3	Banana	Potato	Maize	Maize	S/potato	Cassava	S/potato
4	Sugarcane	Sugarcane	Potato	Coconuts	Cassava	Vegetables	Sugarcane
5	Rice	Banana	Sweetpotato	Pulses	Beans	Wheat	Maize
6	Sweet potato	Plantains	Cabbage	Sorghum	Pumpkins	Cottonseed	Cow milk
7	Vegetables	Groundnuts	Vegetable	Fruits	Sorghum	Cow milk	Millet
8	Sorghum	Fruits	Banana	Cottonseed	Cocoyam	Fruits	Potato
9	Cow milk	Vegetables	Plantains	Vegetables	Sugarcane	Sweetpotato	Bananas
10	Potato	Pigeonpeas	Tomato	Rice	Cow milk	Groundnuts	Sorghum
11	Plantains	Beans	Beans	Groundnuts	Maize	Tobacco	Beans
12	Beans	Tobacco	Pineapples	Cotton lint	Fruits	Hen eggs	Vegetable

Source: FAO Statistics Division, 2008.

Table 2.2 shows that cassava (underlined), maize, banana and rice are the most income-earning staples in the seven representative SSA countries. Except for Kenya, cassava ranks very high as the most earning staple in the SSA countries above, a clear indication of the importance of cassava to the livelihood of many families in SSA.

Five well known tuber crops (potato, cassava, sweet potato, yams and taro) together account for 90% of World's total staple tuber production (Shewry, 2003). Among these tuber crops, FAO (2010) statistics show that cassava ranks second after potato, accounting for 30% of total tuber production (Table 2.3).

Table 2.3: Production of staple tuber crops 2003 to 2008 (million tons)

Tuber Crop	2003	2004	2005	2006	2007	2008	Average
Potatoes	314.2	336.3	325.1	305.6	323.5	314.1	319.8
Cassava	191.3	203.1	207.1	222.3	224.1	232.9	213.5
S/potato	129.8	129.6	126.7	105.7	99.7	110.1	116.9
Yams	44.2	46.9	49.0	52.2	46.6	51.7	48.5
Taro	10.7	10.9	11.2	11.7	11.3	11.8	11.3

Source: FAO Statistics Division, 2010. S/potato = Sweet potato.

The SSA is also expected to experience the most rapid growth in food demand for root and tubers averaging 2.6% per year through 2020 (Scott *et al*., 2000). Most of this growth in demand, which will account for nearly 122 million tons (*ibid*.), is expected to come from cassava (by about 80 million tons). This demand in cassava is equivalent to 66% of the total demand. Cassava demand is estimated to grow at 2.0%

annually for food and 1.6% per year for feed in developing countries, while total cassava production is projected to reach 168 million tons by 2020 based on the current production rate (Table 2.4).

Table 2.4: Cassava production and use in 1993, and projected to 2020

	Area (million ha)		**Yield (tons/ ha)**		**Production (million tons)**		**Total Use (million tons)**	
Area/Region	**1993**	**2020**	**1993**	**2020**	**1993**	**2020**	**1993**	**2020**
SSA	11.9	15.9	7.4	10.6	87.8	168.6	87.7	168.1
Latin America	2.7	2.7	11.3	15.6	30.3	41.7	30.3	42.9
South East Asia	3.5	3.5	12.1	13.7	42.0	48.2	18.9	24.4
India	0.2	0.2	23.6	28.4	5.8	7.0	5.7	7.3
Other South Asia	0.1	0.1	9.4	13.5	0.8	1.3	0.9	1.4
China	0.3	0.3	15.1	20.2	4.8	6.5	5.1	6.4
Other East Asia	Na	Na	na	na	Na	na	1.8	1.9
Developing	18.8	22.9	9.2	12.0	172.4	274.7	152.0	254.6
Developed	...	...	12.1	14.7	0.4	0.4	20.7	20.5
World	18.8	22.9	9.2	12.0	172.7	275.1	172.7	275.1

Legend: na = no recorded production; ha = hectares. **Source:** Scott *et al*., 2000.

Table 2.4 shows that area under cassava cultivation is expected to expand only in Sub-Saharan Africa and other developing countries. Production and use is expected to double in SSA by the year 2020. Cassava alone accounts for over 97 percent of the total in developing countries.

2.1.3 Cassava Production and Use in Tanzania

The Food Crops sub-sector of the Tanzania's Ministry of Agriculture Food and Cooperatives groups cassava under drought resistant crops. Other crops in the group are millet and sorghum. According to the Tanzania's 1998/1999 Agricultural Survey, cassava was planted to 73% of the total area under root and tuber crops as shown in Table 2.5.

Table 2.5: Proportion of households cultivating and planted area of food crops

Crop	Households (x 1000)	%	Planted area (1000' Ha)	%
Maize	3,233.6	40	3,010.6	55
Paddy	685.3	8	503.5	9
Sorghum	867.0	11	551.7	10
Bulrush millet	207.7	3	165.9	3
Finger millet	192.3	2	90.9	1.6
Wheat	60.9	0.8	35.8	0.6
Cassava	1,697.3	21	848.1	15
Sweet potato	1,017.0	12.6	266.9	5
Irish potato	76.4	0.9	28.4	0.5
Yam & cocoyam	64.1	0.8	11.4	0.2
Total	**8,101.6**	**100**	**5,513.4**	**100**

Source: National Bureau of Statistics, 2001.

Cassava farms have in general remained small, with majority of the households cultivating between 2 and 5 hectares. Most of the farmers use traditional production methods, with hand hoes being the main tools for land preparation, planting, weeding and harvesting (Lazaro *et al*., 2007). Cassava production trends in Tanzania are shown in Table 2.6.

Table 2.6: Cassava production trend in Tanzania

Year	Area harvested (1000 ha)	Yield (kg/ha)	Production (1000 tons)
1999	655.7	8,213.3	5,386.1
2000	809.7	6,597.6	5,342.1
2001	660.9	6,591.3	4,336.4
2002	660.2	7,839.5	5,176.1
2003	660.0	6,003.1	3,962.1
2004	660.0	6,728.1	4,440.6
2005	670.0	8,267.4	5,539.2
2006	670.0	9,191.4	6,158.3
2007	675.0	9,777.7	6,600.0
2008	675.0	9,777.7	6,600.0

Source: FAO Statistics Division, 2010.

Table 2.6 shows that area under cassava cultivation has remained the same from 2001 to 2008. The variation in yield that is observed between 2001 and 2008 may be attributed to climatic conditions, technology uptake by farmers especially the use of

improved varieties. According to these data, the average cassava production in the country is 7.9 tons/ha.

The main cassava producing regions of Tanzania include Mtwara, Mara and Ruvuma, which cultivated about 12.2%, 11.7% and 11.3% of total cassava-planted area during 1998/99 season, respectively (Lazaro *et al*., 2007). Large cassava-planted areas are also found in the following regions with percentage of the area in brackets: Mwanza (8.3%), Kigoma (6.5%), Lindi (6.1%), Coast (6.4%), Shinyanga (5.9%) and Kagera (5.9%). In terms of amount of cassava produced, three major growing zones in Tanzania are recognized: the Coastal strip, the Lake Victoria zone, and the Lake Nyasa zone constituting 48.8%, 23.7%, and 13.7%, of total cassava produced in the country, respectively (Mtunda *et al*., 2002; Massumba *et al*., 2005).

Uses of cassava tubers, which comprise about 80% starch dry weight, are variable. The tubers can be directly cooked or processed into flour. The processing, which involves cooking, grinding, drying, or fermenting, according to local custom is also necessary to neutralize the varying amounts of cyanide produced by the plant (Lazaro *et al*., 2007). Cynogenic glucosides are secondary metabolites, which are potentially toxic and their quantity depends on cassava cultivars and growing conditions. The processed roots are then eaten in various forms: e.g., boiled, baked, fried, and as meal or flour. The young tender leaves contain high levels of protein (7% fresh weight) and can be prepared in a similar manner as spinach. However, cassava is cultivated throughout the tropics mainly for the starchy tuberous roots. The crop offers the cheapest source of calories and highest yield per unit area making it a target for intervention.

2.1.4 Constraints to Cassava Production

Cassava production in Africa makes up to 54% of the World total, with the other two major producing continents being South America and Asia (FAO, 2003). This 54% is produced on only about 65% of the total area under cassava. Thus, there is an obvious mismatch between area under cultivation and yield in Africa (averaging 8.9 tons/ha),

being only 70% of those in South America and 61% of those in Asia. Under favourable experimental conditions, cassava can yield as much as 80 - 90 tons of fresh roots per hectare.

However, under field conditions characterized by poor soils, harsh climates and association with other crops, average yields in tons of fresh roots per hectare are much lower (FAO, 2001). Poor yields in SSA, can be attributed to a range of factors, but one of the most important is loss due to pests and diseases. Although there is much less diversity among the pests and diseases of cassava compared with those of Latin America, those that are present are particularly damaging.

Evaluation of the biotic constraints to cassava production in developing countries (Beatriz *et al*., 2011) has revealed that four constraints are outstanding: whiteflies (*Homoptera: Aleyroidae*), cassava green mites *Mononychellus tanajoa* Bondar (*Tetranychidae: Acarina*), cassava mosaic disease (CMD) and cassava brown streak disease (CBSD). Hotspots for potential cassava pest and disease outbreaks include the Mato Grosso in Brazil, northern South America, the African rift valley, the southern tip of India and much of Southeast Asia, where all four biotic constraints show high potential suitability.

2.1.5 Cassava Mosaic Disease as Constraint to Cassava Production

As far as diseases are concerned, the main attention in Africa has been on cassava mosaic disease (CMD) (Thresh *et al*., 1998; Hillocks, 2002). CMD (Figure 2.3) is considered to be the main biotic constraint to cassava production and the most important threat to food security in Sub-Saharan Africa.



Figure 2.3: Cassava field with CMD. (**Source**: Own field survey at Mwangika in Kwimba District, Tanzania).

CMD was first shown to be transmitted by the Sweetpotato whitefly, *Bemisia tabaci* Genn. (*Homoptera*: *Aleyrodidae*) (Fondong *et al*., 1998) in Africa in a persistent, circulative manner (*ibid*.). The whitefly (Figure 2.4) is also a vector in families *Potyviridae* and *Comoviridae* as well as genera *Carlavirus* and *Closterovirus*.



Figure 2.4: The Sweetpotato whitefly *Bemisia tabaci.*

However, it is now considered that CMD is transmitted by a whitefly complex, with different populations occurring in different regions (Brown *et al*., 1995). For example,

two other species of whitefly (*Bemisia afer* Priesner & Hosny and *Aleurodicus disperses* Russell) also infest cassava in Africa and India but have not been adequately tested as possible vectors. In India also, two distinct biotypes of whitefly, cassava biotype and the sweet potato biotype have been reported through sequencing of the mitochondrial cytochrome oxidase I gene (*mtCOI*) (Lisha *et al*., 2003).

Transmission by whiteflies is responsible for the spread of CMD between plants and can be rapid in some areas. CMD is also disseminated through infected stem cuttings (Figure 2.5) used for vegetative propagation, which is by far, the most popular method of propagation.



Figure 2.5: Transfer and use of cuttings for planting new crop is one course of CMD spread (**Source:** Dr J. Ndunguru).

The use of infected stem cuttings can lead to the introduction of CMD to new areas and accounts for the occurrence of the disease in areas where there is little or no spread by the whitefly vector. The estimation of whitefly numbers is thus an important component of CMD studies. The most common method used to assess *B. tabaci* populations on individual cassava plants involves direct counts of adults or sessile nymphs on representative shoots.

To this end, it is possible to distinguish between cutting-derived (C) and whitefly-derived (W) infections, provided that cassava fields are 3 – 6 months after planting (MAP) (Sseruwagi *et al*., 2004). The W infections cause disease symptoms on only the upper most leaves, whereas C infections cause symptoms on the lowest first-formed leaves.

2.2 CMD – IMPORTANCE, SPREAD AND PREVALENCE

2.2.1 Early Reports and Control Strategies

Symptoms of CMD were first reported in 1894 under the name 'Krauselkrankheit' (literally meaning 'curling') in what is now Tanzania (Warburg, 1894) and then in Uganda in 1928 (Martin, 1928). The disease was later reported in many other countries of the East, West and Central Africa. It is now known to occur in all the cassava-growing countries of Africa and the adjacent islands and also, in India and Sri Lanka. This spread has led to studies on means of spread and control.

Early CMD epidemics were controlled using sanitation and resistant varieties. CMD-resistant varieties were found to express less severe symptoms than susceptible ones. Resistance breeding began in the late 1940s in Tanzania, Madagascar and elsewhere.

A brief history of CMD research milestones is shown in Table 2.7.

Table 2.7: Summary of CMD research milestones

S/N	Year	Event
1	1894	CMD first reported in Tanzania
2	1926	Disease was first recorded in West Africa
3	1932	First whitefly transmission in Congo
4	1930s – 1960	Studies and resistance developed from inter-specific crosses
5	1971 – Present	Breeding programs
6	1983	Etiology determined
7	1970s/1980s	Epidemiology characterized (in Kenya and Cote d'voire)
8	1994	First distribution map in Africa
9	1990s	Pandemic of severe CMD recorded in Uganda
10	1997	Novel recombinant cassava mosaic geminivirus (CMG) described
11	1997	Mixed infections first described
12	1996 – 2000	Increased diversity observed for Africa
13	1995 - 2003	Spread of CMD pandemic in East and Central Africa recorded
14	1990s – 2003	Major deployment of host plant resistance
15	2001	New single dominant CMD resistance gene identified from local landraces
16	2003	Geminiviruses reclassified: Six CMG species recorded for Africa.
17	2004 – 2005	Diversity of CMGs and discovery of Satellites II & III (Tanzania)
18	2008	Geminiviruses classification updated to seven species in Africa: EACMV, EACMMV, EACMV-TZ, EACMKV, EACMZV, ACMV, and SACMV.

Source: Legg and Thresh (2000); Ndunguru *et al*. (2005); Fauquet *et al*., (2008).

2.2.2 CMD Symptoms

CMD symptoms are variable, but generally distinguished as leaf mosaic patterns (chlorosis) that affect discrete areas. The symptoms vary from leaf to leaf, shoot to shoot and plant to plant, even for the same variety and virus strain in the same locality. The chlorotic areas fail to expand fully so that the unequal expansion of leaf lamina causes malformation and distortion (Figure 2.6).



Figure 2.6: Typical CMD symptom. It is characterized by chlorotic areas and leaf distortion. (**Source**: Own field survey).

Severely affected leaves are reduced in size, misshapen and twisted, with yellow areas separated by areas of normal green colour. The plants are stunted and the young leaves absciss. The leaf chlorosis may be pale yellow or nearly white, or just discernibly paler than normal (Fondong *et al*., 1998). Distortion, reduction in leaflet size and general growth retardation, appear to be secondary effects associated with symptom severity.

Variation in symptoms may be due to differences in virus strain or variant, cassava cultivar (or variety), sensitivity of the host, plant age and environmental factors such as soil fertility, soil moisture, radiation and temperature. Some leaves situated between affected ones may seem normal and give the appearance of recovery. This behaviour depends on the ambient temperature and host plant resistance. Symptoms tend to reappear on axillary growth after the shoot tips are removed and this procedure is sometimes adopted to enhance symptom expression in screening for resistance.

The main difficulties that arise in recording CMD symptoms occur when pests or nutritional deficiency has affected the plants. Cassava green mites and zinc deficiency cause particular problems in disease assessment (Sseruwagi *et al*., 2004). A key distinction however, is that both mineral deficiency symptoms and mite damage are usually similar on each side of the midrib, whereas the symptoms of CMD are usually asymmetrical and the two sides differ.

In recording experiments and in screening for resistance to CMD, two parameters have been explored - incidence and severity. Disease incidence refers to the number of visibly diseased plants, usually in relation to the total number assessed and so is expressed as the percentage of plants in stand with symptoms. CMD surveys conducted in 18 countries of Africa have shown that the overall incidence exceeds 50% in 11 of the 18 countries and was greatest in Congo (79%), Ghana (72%), Western Kenya (up to 84%), Nigeria (up to 82%) and Senegal (83%) (Sseruwagi *et al*., 2004). In the surveys, the lowest incidences were in Chad, Malawi, Madagascar, Rwanda and South Africa.

Disease severity usually refers to the degree of symptom expression as assessed visually using an arbitrary scale. Scales of 0 – 5 (Fauquet and Fargette, 1990) and 1 – 5 (Hahn *et al*., 1980) have been used for CMD, where 0 or 1 represents no disease symptoms and 4 or 5 the most severe symptoms, including leaf distortion and stunting of plants, respectively. The scale of 1 – 5 (Table 2.8) has been used most commonly for individual plants.

Table 2.8: Cassava mosaic disease symptom scale of 1 – 5

Scale	Symptom Label	Symptom Description
1	Healthy	Unaffected shoots, no symptoms
2	Mild	Mild chlorosis, mild distortions at bases of most leaves, while the remaining parts of the leaves and leaflets appear green and healthy
3	Moderate	Pronounced mosaic pattern on most leaves, narrowing and distortion of the lower one-third of the leaflets
4	Severe	Severe mosaic, distortion of two thirds of most leaves and general reduction of leaf size and stunting of shoots
5	Very severe	Very severe mosaic symptoms on all leaves, distortion, twisting, mishappen and severe leaf reductions of most leaves accompanied by severe stunting of plants.

Source: Hahn *et al*., 1980.

Two contrasting methods have been used to calculate mean severity. One approach is to consider records for all the plants in the stand being assessed, including those free of disease, whereas the other only considers data for diseased plants. For instance, four plants having CMD severity scores of 3, 4, 4 and 1 on the 1 – 5 scale have a mean severity of 3.0 using the first method and 3.7 using the second. The second

method has been used in recent surveys and field experiments in Uganda and elsewhere (Sseruwagi *et al*., 2004), and is recommended for use more widely as it provides a true evaluation of disease severity in the stand assessed.

2.2.3 Crop Losses Due to CMD

Recent reports suggest that CMD has affected an area of 2.6 million square kilometers and causing an estimated annual economic loss of US$1.9 – 2.7 billion and it has been termed the most damaging plant virus disease in the World causing famine and deaths to thousands of people (Legg *et al*., 2006). Great differences exist between countries in the overall prevalence of CMD and in the severity of the losses caused as determined in field experiments with diverse landraces and improved cultivars.

The losses range from negligible to almost total, depending on factors such as:

(i) Virulence of the virus species present, with virulence defined as the degree of damage caused to a host by parasite infection (Sacristan and Garcia-Arenal, 2008);
(ii) The susceptibility of the host;
(iii) The stage of growth when infection occurs;
(iv) Soil fertility and other growth conditions.

Available information from surveys and yield loss assessments has been summarized by Thresh *et al*. (1997), in which it is estimated that total losses in Africa are 15 – 24%. This is equivalent to 15 - 28 million tons based on assumption of an overall CMD incidence of 50 – 60% and a loss of 30 – 40% in the yield of diseased plants. An important limitation of the information available is that with few exceptions they have been obtained in trials at experimental stations with cassava grown at regular spacing and without intercrops. In the trials, the differences in growth and yield have been determined from representative 5m x 5m quadrats of the stands assessed. Notwithstanding, there is a general agreement in many African countries that CMD is the most important disease of cassava.

2.2.4 CMD Pandemic in Eastern and Central Africa

Recent assessment of pests and disease problems have distinguished three contrasting CMD situations referred to as epidemic, endemic and benign (Thresh *et al*., 1997). In epidemic situations, the whitefly vector is spreading CMD rapidly and the symptoms are prevalent and severe. The situation encountered in much of Uganda in 1990s and which then spread to adjacent areas of Western Kenya, Tanzania and Rwanda is rated as epidemic. Endemic areas have a high incidence of CMD but the symptoms are not usually very severe. The benign situation is characterized by the areas experiencing low CMD incidence and seldom exceeds 20%. This was formerly (before 1990s) a situation in much of Uganda and Western Kenya.

Recent surveys have indicated that over 80% of cassava plants in fields show severe CMD symptoms with cassava in the Lake Victoria basin expressing the most severe symptoms followed by Southern regions of Tanzania (Ndunguru *et al.*, 2005).

2.3 ETIOLOGY OF CMD

2.3.1 Early Consideration

CMD was first proposed to be a viral disease by Zimmermann (1906). The virus was shown to consist of geminate particles and the genetic material was found to comprise of two components of circular ssDNA. The virus was first isolated in East Africa and was named *Cassava latent virus,* CLV (Bock *et al*., 1981, Bock *et al*., 1978) because in the early experiments it was not possible to infect cassava and produce similar symptoms. The first sequence of the DNA-A was published in 1983 (Bock and Woods, 1983) and shortly after that, successful infection back to cassava was achieved from *Nicotiana benthamiana*. This fulfilled Koch's postulates and led to the naming of the causal virus as *African cassava mosaic virus* (ACMV).

It is noteworthy however, that Koch's postulates in its original text have been ignored in plant virology. This is because viruses cannot be grown in pure culture, on cell-free artificial media (Bos, 1999).

In order to apply Koch's postulates to establish a virus is a cause of a particular disease the postulates require that the virus must be:

(i) Shown to be concomitant (or associated) with the disease,
(ii) Isolated from the diseased plant and separated from contaminants by selective propagation in a differential host,
(iii) Purified and identified for its intrinsic properties,
(iv) Demonstrated to reproduce the disease when inoculated into a healthy specimen of the host species, and
(v) Confirmed to occur in, and be re-isolated from, the experimental host.

2.3.2 Structure and Organization of Geminiviruses

Although CMD was first shown to be caused by ACMV, it is now known to be caused by members of genus *Begomovirus* belonging to the family *Geminiviridae* (Fauquet, 2008). It is a large family of plant DNA viruses posing severe threats to economically important crops worldwide. Infection by geminiviruses can produce leaf mottling that interferes with photosynthesis, decreasing yield of starchy foods such as cassava and they disrupt flower and fruit formation in crops such as tomato, pepper and cotton (Moffat, 1999). In fact, cassava mosaic geminiviruses are one of the economically most important tropical plant viruses (Fargette *et al*., 2006).

Geminiviruses have four distinctive features:

(i) The virion particle has a twin icosahedral (geminate) capsid structure, ~ 18-30 nm in size (Figure 2.7),
(ii) Their genetic material consist of one or two covalently closed, circular, single stranded DNA (ssDNA), varying in size from 2.5 kb to 3.0 kb in length, depending on the virus,
(iii) The genomic ssDNA is replicated in the nucleus of the host cell by a rolling-circle mechanism using double stranded DNA (dsDNA) intermediates,
(iv) All geminivirus genome components have non-coding intergenic region (IR) of

about 200 nt from which viral genes diverge in both the viral and complementary sense. The IR contains an invariant nanonucleotide sequence TAATATTAC.

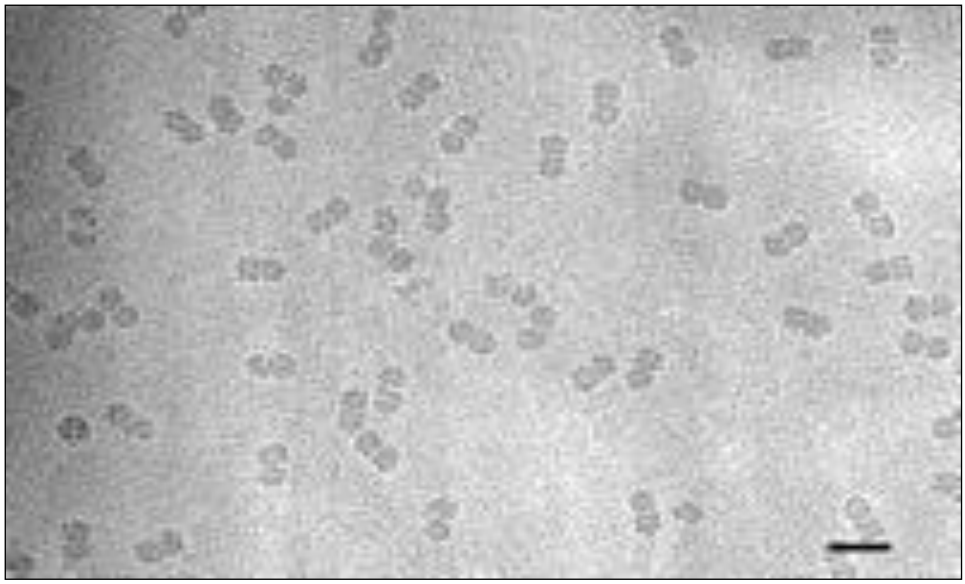

Figure 2.7: ACMV geminate particles. Bar = 50 nm. (**Source:** Bottcher *et al*., 2004).

The IR contains divergent RNA polymerase II-type promoters responsible for the expression of viral genes. It also contains sequence elements, including two TATA motifs and multiple copies of *cis*-elements known as iterons, which are binding sites for *Rep* (Patil and Fauquet, 2009).

Different types of iterons have been identified in CMGs and they can be classified into three groups:

(i) The ACMV type with isolates of ACMV, *East African cassava mosaic Zanzibar virus* (EACMZV), and *Sri-Lankan cassava mosaic virus* (SLCMV);

(ii) The *East African cassava mosaic virus* (EACMV) type encompassing all the other EACMV-like viruses and *South African cassava mosaic virus* (SACMV); and

(iii) The *Indian cassava mosaic virus* (ICMV) type with ICMV isolates alone.

A striking feature within the IR is a conserved inverted repeat that is capable of forming a hairpin structure.

There are four genera within the *Geminiviridae* distinguished on the basis of genome organization, insect vector and host range: *Mastrevirus* (Type species *Maize streak virus*, MSV), *Begomovirus* (Type species *Bean golden mosaic virus*, BGMV), *Topocuvirus* (Type species *Tomato pseudo-curly top virus*, TPCTV) and *Curtovirus* (Type species *Beet curly top virus*, BCTV) (Fauquet and Stanley, 2003; Fauquet *et al*., 2008). The genome structures of these geminivirus species are shown in Figure 2.8.

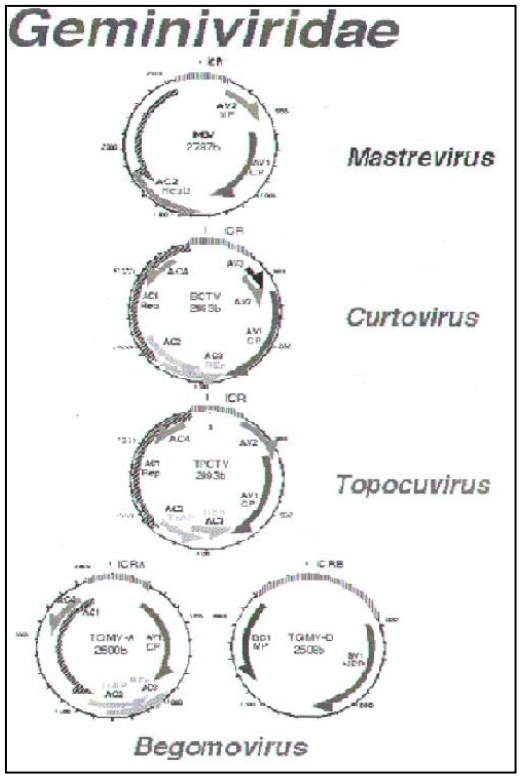

Figure 2.8: Genome structures of viruses in the Family *Geminivirdae*.

2.3.3 Replication of Geminivirus Genomes

Geminiviruses replicate their circular single-stranded DNA (ssDNA) genome via double-stranded DNA (dsDNA) intermediates either by rolling-circle mechanism or a recombination-dependent mechanism in the nuclei of actively dividing infected cells (Jeske *et al*., 2001). Rolling-circle replication (RCR) (Jeske *et al*., 2001) describes a process of nucleic acid replication that can rapidly synthesize multiple copies of circular molecules of DNA or RNA.

Virus particles are introduced into the cell during whitefly feeding and are uncoated. Then the RCR occurs in three stages. In the first stage (Figure 2.9), the viral ssDNA (SS) enters the nucleus and is converted into a covalently closed dsDNA replicative form (RF) in a process involving host-directed, RNA-primed synthesis of a complementary strand. This is thus an SS → RF synthesis stage. The second stage of the RCR is RF → RF synthesis. The stage generates more RF, which serves as template for viral transcription and for further replication.

The RF → RF synthesis stage (the rolling-circle phase), requires the action of Rep protein. Rep nicks the plus strand of the dsDNA at a specific site called double strand origin (DSO), usually between positions 7 and 8 of the nanomeric motif (*ibid*.; Khan and Dijkstra, 2001). The third stage of RCR is RF → SS synthesis. It involves the conversion of dsDNA replicative form into ssDNA, which is then assembled into virus particles that accumulate in the nucleus. The stage occurs late in the replication cycle and is responsible for the accumulation of viral genomes for encapsidation. This stage is similar to the RF - RF synthesis, except that priming is prevented and ssDNA is the predominant product.

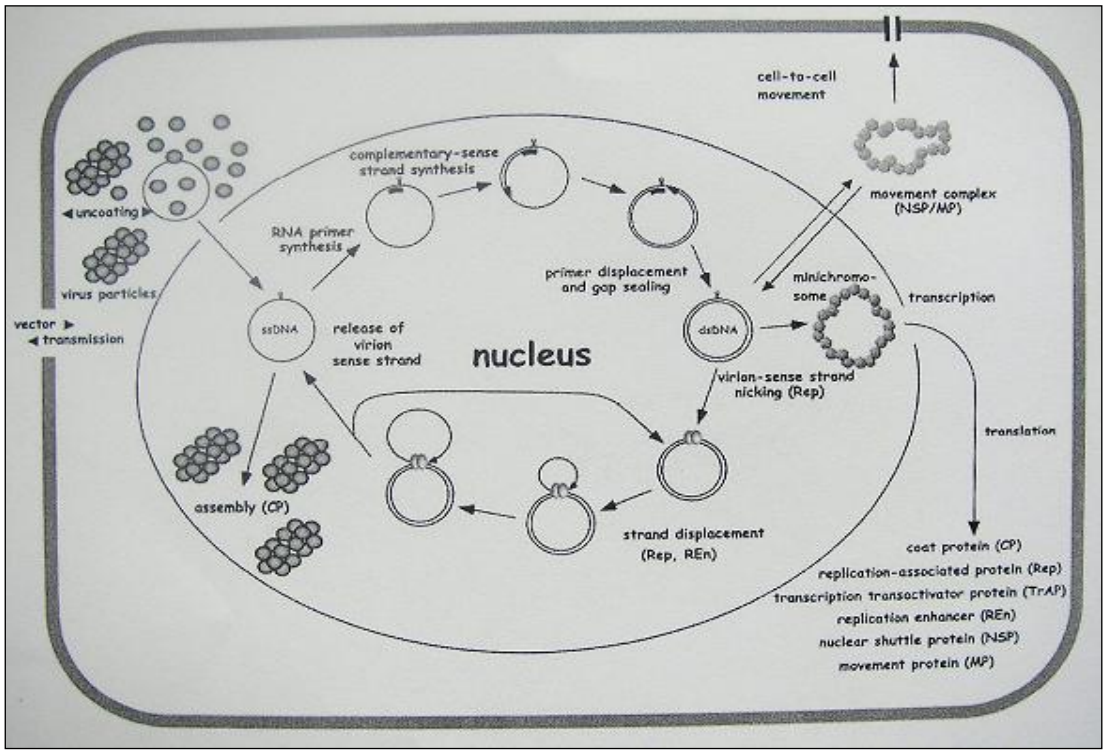

Figure 2.9: A simplified version of the geminivirus DNA replication cycle and intercellular movement of viral DNA. CP = coat protein, MP = movement protein. (**Source:** Guiterrez, 2000).

The fate of the ssDNA products formed in the RCR process is threefold:

(i) They can re-enter the DNA replication pool,
(ii) They can associate with the CP, or
(iii) They can be transported outside the nucleus and to the neighbouring cell, through the plasmodesmata with the help of viral MPs.

The characteristic feature of rolling-circle replication is the involvement of a replication initiator protein (Rep) with a nicking-closing activity similar to that found in topoisomerases. Geminiviruses can also replicate by a recombination strategy (Jeske *et al*., 2001).

2.3.4 Features of Genus *Begomovirus*

The overwhelming majority of geminiviruses belong to the genus *Begomovirus*. Fauquet *et al*., 2008 suggest that virus species in the genus are characterized by:

(a) The presence or absence of a DNA-B component,

(b) The presence or absence of ORF AV2,

(c) The hairpin structure of the IR contains a 31bp-consensus sequence GGCCAT/ACCGNT/AA/TTAATATTACCGGA/TTGGCC (Lazarowitz, 1992)

(d) Transmission by whiteflies; and

(e) The viruses infect dicotyledonous hosts.

Thus, begomovirus genomes may be monopartite or bipartite. Most of the described begomoviruses are bipartite containing both DNA-A and DNA-B molecules each about 2.6 - 2.8 kb (Fauquet and Stanley, 2003). Genome structure of a typical begomovirus is shown in Figure 2.10.

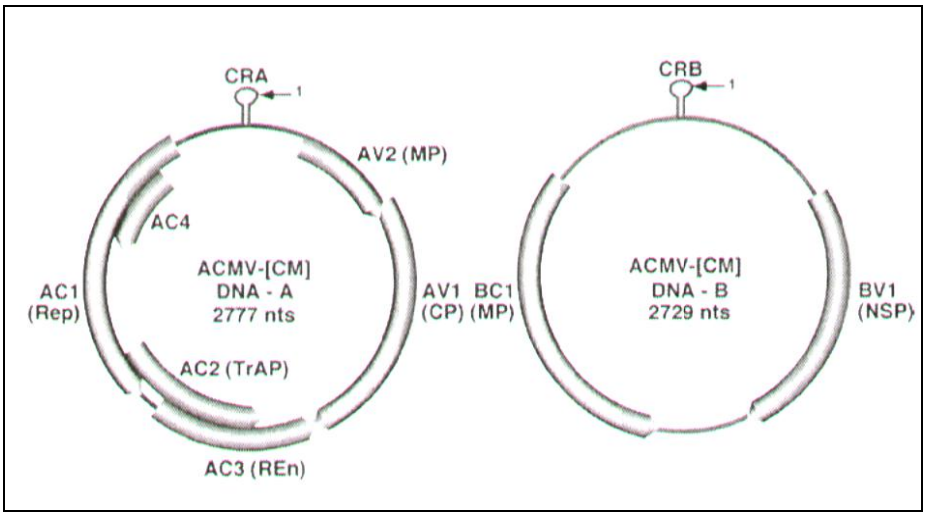

Figure 2.10: Genome structure of ACMV-[CM: 98], a typical geminivirus in genus *Begomovirus* (**Source:** Chellappan *et al*., 2005).

Serological tests have shown that all begomoviruses are related. However, there is a group of epitopes unique to the begomoviruses, which infect crops in the Old World (Europe, Africa, Asia and Australia) and a distinct set of epitopes is shared by begomoviruses that infect crops in the New World (North, Central and South America) (Padidam *et al*., 1995). While Old World (OW) begomoviruses are represented by ACMV, the New World (NW) begomoviruses are represented by BGMV. The NW members have bipartite genomes whereas most of the OW members

have monopartite genomes that has homology with the DNA-A of the bipartite members. The replication of begomoviruses follows the rolling-circle mechanism (RCR) in common with other geminiviruses.

2.3.5 Arrangement and Functions of Begomovirus Genes

In common with other geminiviruses, begomoviruses contain open reading frames (ORFs) that are bidirectionally transcribed. An ORF is a nucleotide sequence from a start codon to stop codon, without an intron in between and capable of expressing a protein of 10kDa or more (Khan and Dijkstra, 2001). The begomovirus DNA-A component carries on its virion-sense strand, which is the encapsidated strand, ORF *AV1* encoding the viral coat protein, CP (about 27 – 30kDa), which is also the most highly conserved begomovirus gene. The CP of begomoviruses is specifically adapted for transmission by the local whitefly population, which explains the antigenic similarity of the CPs of begomoviruses from the same area (Harrison and Robinson, 1999) and co-adaptation between begomoviruses and their local whitefly populations. DNA-A of the OW begomoviruses also encode an additional ORF *AV2*/*V2* that have been implicated in virus accumulation and movement (Padadim *et al*., 1995; Rigden *et al*., 1993). It possibly encodes the smallest protein of approximately 13kDa.

There are four ORFs *AC1* to *AC4*, on the complimentary sense strand. *AC1* ORF encodes Rep, a replication initiation protein of about 40kDa, which is also required for cell cycle regulation. *AC2* ORF encodes a transcriptional activator protein, TrAP (15 – 20kDa) required for the rightward gene expression, particularly the activation of *AV1* (CP) and *BV1* (NSP) promoters. TrAP functions in a virus-nonspecific way, suggesting that DNA elements present in the CP and *BV1* promoters and involved in the transactivation process could be conserved. It was later confirmed that the conserved DNA motif, named conserved late element, CLE was GTGGTCCC.

AC3 ORF encodes the replication enhancer protein, REn, (14 – 16kDa) that regulates the virus replication rate, possibly via the activation of an early gene (*AV1*) required for DNA synthesis. Thus, REn increases replication efficiency of the virus. Some DNA-A of bipartite viruses and of all monopartite viruses encode ORF *AC4* that participates in cell cycle control (Briddon and Stanley, 2006). The ORF determines

symptom expression in the monopartite begomoviruses through RNA silencing. These virus genes are sufficient for virus replication and formation of virion particles.

Bipartite begomoviruses usually require *BV1* (nuclear shuttle protein, NSP) and *BC1* (cell-to-cell movement protein, MP) genes encoded on DNA-B component for inter- and intra-cellular movement, respectively (Hull, 2002). The product of *BV1* is located in the nucleus and binds ssDNA, allowing the newly formed virus genome to be transported to the cytoplasm. The *BC1* product has been extracted from cell wall and cellular membrane fractions. Its function is to increase the exclusion limit of plasmodesmata (pd) to facilitate cell-to-cell movement. However, some bipartite begomoviruses can infect plant hosts systemically with only DNA-A component. The nucleotide sequences of DNA-A and DNA-B are different except for the CR that shares more than 90% nucleotide sequence identity. The CR carries regulatory sequences essential for viral DNA replication.

Geminivirus DNA for example, is synthesized in the nucleus, requiring that genomes destined for intercellular transport must first traverse the nuclear-cytoplasmic boundary (Carrington *et al*, 1996). Plant viruses mainly use two principal strategies for cell-to-cell movement. One involves binding of MP or MP complexes with the viral genome, which are either DNA or RNA, thereby increasing the size exclusion limit of the Pd. The other one is dependent on tubule formation, the example of which is found in cowpea (Lent *et al*., 1991). In cowpea plant cells infected with cowpea mosaic virus, tubular structures containing virus particles are formed in the plasmodesmata between adjacent cells. The structures are involved in cell to cell spread of the virus.

Depending on the particular begomovirus, five different proteins are involved in mediating virus transport in host plants: the CP (Khan and Dijkstra, 2001), the nuclear shuttle protein, the movement protein, the C4 protein, and the V1 protein. Unlike monopartite begomoviruses, the CP of bipartite begomoviruses is not required for cell-to-cell or long-distance movement. DNA-A of Old World begomoviruses also encode an additional ORF *AV2*/*V2* that has been implicated in virus accumulation and symptom development (Padadim *et al*., 1995).

2.3.6 Classification and Diversity of Cassava Geminiviruses

One important feature of CMD pandemic in Africa is the diversity of geminiviruses involved, comprising species, strains and variants. In earlier studies, virus species causing CMD fell into three categories on the basis of their reactions with monoclonal antibodies (MAbs) raised against purified virus particles (Thomas *et al*, 1986). The three virus species, which all belong to the genus *Begomovirus*, were named ACMV, *East African cassava mosaic virus* (EACMV) and *Indian cassava mosaic virus* (ICMV) (Berrie *et al*., 1998; Hong *et al*., 1993).

These viruses were reported to have different, scarcely overlapping distributions:

- ACMV in Africa West of the Great Rift Valley and South Africa,
- EACMV in Africa East of the Great Rift Valley and Madagascar, and
- ICMV in India and Sri-Lanka.

Recently, taxonomic guidelines have been developed to provide a framework on which to base the definitions of species and strains. Based on this new approach, in which the sequence identity demarcation between members of different species has been set at 89% for the DNA-A component of begomoviruses, seven African and two Indian CMG species have been recognized (Fauquet *et al*., 2003; Fauquet *et al.*, 2008): ACMV, EACMV, East Africa cassava mosaic Tanzania virus (EACMV-TZ), East African cassava mosaic Kenya virus (EACMKV), East African cassava mosaic Malawi virus (EACMMV), East African cassava mosaic Zanzibar virus (EACMZV) and South African cassava mosaic virus (SACMV), reported from Africa; and ICMV and Sri Lankan cassava mosaic virus (SLCMV), identified from the Indian subcontinent.

Classification taxa below the species level have been restricted to strains and variants. Strains within a species level have been associated with a specific phenotype, host range or geographical distribution. Several strains have been recognized: *East African cassava mosaic virus-Uganda* (EACMV-UG), *-Kenya* (EACMV-KE), *-Tanzania* (EACMV-TZ); *South African cassava mosaic virus-South Africa* (SACMV-ZA), *-Madagascar* (SACMV-MG); East *African cassava mosaic Cameroon virus-Cameroon* (EACMCV-CM), *-Tanzania* (EACMCV-TZ); *Indian cassava mosaic*

virus-India (ICMV-IN), -*Kerala* (ICMV-Ker); *Sri Lankan cassava mosaic virus-India* (SLCMV-IN and -*Sri Lanka* (SLCMV-LK) (Ogbe *et al*., 2003).

Virus variants are viruses with slightly different genome, symptom or mode of transmission. The percentage nucleotide identity demarcations are 85 – 93% and 94 – 100% for strains and variants, respectively (*ibid*.).

In the early 1990s, a new epidemic of severe CMD occurred in Uganda and progressed steadily from the northeast towards the south-west (Legg and Thresh, 2000). The subsequent discovery of the association of a novel recombinant virus, the EACMV Uganda strain (EACMV-UG), in severely affected plants from the epidemic area boosted research interest in CMD (Zhou *et al.*, 1997). This was followed by an extensive series of surveys in different parts of the African continent during which the co-occurrence of ACMV and EACMV was discovered in western Kenya and north-western Tanzania.

From these surveys, many virus isolates were sequenced and new species related to EACMV (EACMV-like viruses) were identified from East and West African countries. The initial surveys conducted during the early period of the epidemic in central-southern Uganda showed that ACMV occurred more frequently than EACMV-UG, although there were some incidences of mixed infections (Harrison *et al.*, 1997). However, subsequent surveys revealed more frequent occurrences of EACMV-UG than ACMV and a significantly decreased proportion of mixed infections (Legg *et al.*, 2006).

Such a drastic change has been attributed to two factors:

(i) The selection of planting material by the farmers, in which they avoided severely diseased plants that were usually dually infected;
(ii) The hypothesis (though not proven), that EACMV-UG DNA-A, which contains a fragment of the ACMV coat protein (CP), might be more efficiently transmitted by whitefly vectors, thus leading to its widespread distribution in Africa.

In addition to the transmission of CMD by whiteflies, human intervention can be another important cause for rapid geographical movement and exchange of geminiviruses. Recombination between members of different virus species has been a driving force for biodiversity, as exemplified by the EACMV/SACMV gradient between East Africa and South Africa. Recently, Bull *et al.* (2006) conducted an exhaustive survey of the genetic diversity of CMGs across the major cassava-growing areas of Kenya, mainly representing EACMV and EACMZV, as well as a novel begomovirus named *East African cassava mosaic Kenya virus* (EACMKV). The DNA-B components of these EACMV-like viruses were much less diverse than their corresponding DNA-A components, but clustered as Western and Eastern (coastal) groups.

Comparison of the CMG sequences from Africa indicates that all isolates of ACMV, irrespective of their geographical origin, are clustered together with little variation in their genomic sequence. In contrast, the genomes of EACMV-like viruses are more genetically diverse as a result of the frequent recombination between members of different species. In addition, there is a large range of phenotypic symptom variation for each of these virus isolates, irrespective of their location and their belonging to a particular species (Bull *et al.*, 2007).

Thus, the explosion of the cassava mosaic pandemic in Africa can currently explained by a conjunction of several mechanisms:

(i) The key synergistic interaction between two viruses, leading to a huge increase in viral titre in the top of the plants, thereby enhancing whitefly transmission capacity;

(ii) The CP recombinant fragment that may provide an advantage to EACMV-UG, as it has been shown that the whitefly transmission epitopes are coded by this fragment and that there has been co-adaptation, as far as the rate of transmission is concerned, between the local whiteflies and viruses (Maruthi *et al.*, 2002); and

(iii) The tremendous increase in whitefly populations adapted to cassava in Uganda, where the epidemic started.

The second major distinguishing biological feature of the CMD pandemic in Africa, particularly at the epidemic 'front' between severely affected and relatively unaffected areas is the abundance of whiteflies and association with the viruses (Legg and Fauquet, 2004). CMD was first shown to be transmitted by *Bemisia* sp. in Africa and it is now considered to be a whitefly complex, with different populations occurring in different regions. The CP of a geminivirus is specifically adapted for transmission by the local whitefly population. This explains the antigenic similarity of the CPs of begomoviruses from the same area (Harrison and Robinson, 1999) and co-adaptation between CMGs and their local whitefly populations.

2.3.7 Diversity of Cassava Mosaic Begomoviruses

The host range of viruses is also considered to be an important criterion to differentiate between geminiviruses and to better understand their epidemiology. Each CMG has different preferences for different hosts in addition to their natural host, cassava.

Classical studies have shown that the experimental host range of ACMV is largely restricted to the Solanaceae family, within which it is more readily transmitted to members of the genera *Nicotiana* and *Datura* (Bock and Woods, 1983). Both ICMV and SLCMV also infect other species of *Nicotiana* (for example, *N. tabacum* and *N. glutinosa*), and SLCMV, in particular, is highly virulent with a very broad host range extending to *Arabidopsis*, as reported recently (Mittal *et al.*, 2008), and even to *Ageratum conyzoides* in association with *Ageratum yellow vein betasatellite* (Saunders *et al.*, 2004). The host range of EACMV has not been studied extensively, but it is not known to infect *Nicotiana* species, except for the commonly used experimental host *N. benthamiana*, and this is also the case with SACMV (Berrie *et al.*, 2001).

Thus, CMGs can be classified into two major groups based on their host range:

(a) ACMV, ICMV and SLCMV forming one class infecting all species of *Nicotiana* in addition to cassava, although with different levels of virulence; and
(b) EACMV-like viruses and SACMV in another class, which infects only *N. benthamiana* and cassava.

2.3.8 Begomoviruses Occurring in Tanzania

A study conducted to characterize cassava mosaic begomoviruses in Tanzania (Ndunguru *et al*, 2005) showed that there is higher molecular diversity among EACMVs than ACMVs. In the study, any cassava mosaic begomovirus (CMB) that gave Restricted Fragment Length Polymorphism (RFLP) patterns that was neither ACMV- nor EACMV-like following digestion with both *Eco*RV and *Mlu*I were provisionally designated as EACMV-[TZ] type, with the addition of name of the location from which it was sampled.

In addition to the EACMV-[TZ] types, 11 EACMV-like viruses were identified and were designated as EACMV-[TZ1] to [TZ11]. ACMV-[KE] and ACMV-[NG] were the main ACMV isolates detected in Tanzania. Another isolate of ACMV was detected that was different from these two isolates and was designated ACMV Tanzania (ACMV-[TZ]. EACMV-[TZ] types occurred infrequently in the Lake zone but were widely distributed in the Coastal zone. ACMV-[KE] and ACMV-[NG] were only found in the Northwestern part of the country. Most of the EACMV-like viruses had localized distribution. EACMV-[TZ1] for example, was more common in the coastal regions but particularly widespread in the Ruvuma region of southern Tanzania.

2.4 INVOLVEMENT OF SATELLITES IN BEGOMOVIRUS PATHOGENICITY

2.4.1 Molecular Epidemiology of Begomoviruses

Molecular epidemiology explains those processes that lead to a successful state of disease in a host plant. It primarily enables identification of a weak link in the biotic

and molecular interactions or processes that may be capitalized upon to achieve disease control (Khan and Dijkstra, 2001). Two terminologies have traditionally been used in plant pathology to describe the epidemiological processes: pathogenicity and virulence.

Pathogenicity is the ability of a pathogen to cause disease on a particular host (i.e. a qualitative property) (Sacristan and Garcia-Arenal, 2008) and virulence is the degree of damage caused to the host (i.e. a quantitative property) assumed to be negatively correlated with host fitness (Fraser, 2006). Virulence describes the ability of a virus strain to overcome a specific host resistance gene.

Stated briefly, conditions necessary for wide-scale development of a damaging disease are:

(i) The virulent race of the pathogen must be present in low frequency in the host;
(ii) The host that is susceptible to this race must be widely distributed in the region;
(iii) Environmental conditions must be favourable for development of the pathogen (Altieri, 1995).

It is widely accepted that virulent parasites are able to infect and damage a host. Therefore, virulence is the key property of pathogens (Harrison, 2002), and understanding the evolution of virulence has been a major goal in plant pathology, for a long time. The evolution of virulence may determine important phenomena, which compromise the success of control strategies for infectious diseases in plants.

These phenomena include:

- The emergence and re-emergence of pathogens,
- Host switch and host range expansion, and
- Overcoming host resistance.

Notwithstanding, plant pathologists, particularly those involved with diseases of crops, have devoted much effort to understanding the evolution of pathogenicity,

largely because it determines the success and durability of resistant cultivars for the control of plant diseases.

Despite the efforts made in the last few decades to improve resistance of plants to diseases, begomoviruses have become an important group of plant viruses causing devastating losses on many crops. In West Africa and the Nile Basin, diseases caused by begomoviruses have emerged in crops such as okra, pepper, bean, cassava, cotton, and tomato (Fauquet *et al*., 2005).

Evidence for the involvement of begomoviruses in these diseases includes one or more of the following:

(i) Leaf curling, narrowing and crumpling symptoms;
(ii) The absence of normal green colour, often replaced by yellowing;
(iii) The presence of its vector, the *Bemisia* whitefly; and
(iv) Molecular detection of begomovirus DNA in the tissues.

Begomoviruses are probably the most ancient and timely plant viral pathogens. This is because of their wide distribution, wide host range, unique molecular features, virus-vector interactions and the co-evolution with their dicotyledonous hosts (Nawaz-ul-Rehman and Fauquet, 2009). The beauty (yellowing) of eupatorium plants described in a short poem by a Japanese Empress Koken way back in 752 AD is considered to be the earliest known record of a plant virus disease. Some 1,257 years later, scientists confirmed that the virus was a begomovirus, *Eupatorium yellow vein virus* (EpYVV). It is ssDNA geminivirus vectored by whiteflies.

2.4.2 Factors Influencing Begomovirus Diseases

Several factors, most of them unsuspected until recently, have influenced the emergence of begomovirus epidemics, the main ones (Fargette *et al*., 2006; Urbino *et al*., 2008) being:

(i) Agricultural intensification and extensification, that has led to the increased populations and diversity of the highly fecund and polyphagus whitefly vector,
(ii) Increased and uncontrolled movement of infected planting material,

(iii) The introduction of the highly susceptible plant varieties,

(iv) Ability of the viruses to suppress host gene silencing machinery,

(v) Unusually high capacity of the begomoviruses for genetic variation, probably resulting from mobilization of begomoviruses from other crops and weeds, and

(vi) Enhanced capacity of the begomoviruses to associate with single stranded replicons such as satellite DNAs (Patil and Fauquet, 2009).

The response of the host to virus infection is characterized by an initial onset of systemic symptoms from which the plant may or may not recover. Based on this, have distinguished two distinct infection types in cassava: a rapid onset from which plants recover (ACMV and SLCMV), and a slow onset from which plants do not recover (EACMV and EACMV-TZ).

2.4.3 Suppression of Gene Silencing by Begomoviruses

The symptom recovery phenotype exhibited by ACMV and SLCMV in cassava has been correlated with the production of virus-derived small-interfering RNAs (SiRNAs) through posttranscriptional gene silencing (PTGS). Begomoviruses have evolved suppressor proteins that counteract this host defense response. *AC2* of ACMV was first shown to be the suppressor of PTGS (Voinnet *et al*., 1999). The *AC2* protein was thought to suppress silencing by two mechanisms: one which targets cytoplasmic RNA silencing and the other siRNA-directed methylation. *AC4* has also been found to be capable of suppressing gene silencing for some CMGs, enhancing disease symptoms.

Mixed infections of ACMV and EACMV have been shown to be an important feature of the severe CMD first reported from Uganda (Harrison *et al*., 1997). A synergistic effect in a mixed infection occurs when there is a significant difference in *AC2* and *AC4* functions of these individual viruses. Therefore, having more than one type of PTGS suppressor provides an advantage to viruses synergistically interacting in mixed infections (Fondong *et al*.) and thus helps in the establishment of a severe/successful disease.

2.4.4 Capacity for Genetic Variation

Molecular diversity among begomoviruses results from recombination, pseudo-recombination and transreplication (Patil and Fauquet, 2009). Recombination is the process by which discrete segments from one nucleotide strand become incorporated into that of a different individual strand during replication (Seal *et al*, 2006). Recombination in begomoviruses represents a normal rather than exceptional evolutionary mechanism. Recombination between begomoviruses was first revealed with the molecular analysis of EACMV-UG, which is a recombinant between ACMV and EACMV and the virus responsible for the pandemic on cassava in Africa (Zhou *et al*., 1997). Although ACMV donated 550 nts to an EACMV isolate to create EACMV-UG, EACMV-like viruses seem to be highly prone to recombination than ACMV. Recombination has later been described in many other diseases such as Tomato yellow leaf curl disease (TYLCD), which devastated tomato crops in Spain in the early 1990s (Moriones and Navas-Castillo, 2008).

Pseudorecombination (or regulon grafting) is a term used to describe the situation in which the DNA-A and DNA-B of the two components of a geminivirus originate from two different geminiviruses (Patil and Fauquet, 2009). In begomoviruses, the process occurs when there is re-assortment of genomic DNA-A and DNA-B of isolates of the same begomovirus species, but not those of distinct begomovirus species. Reassortment, in which the DNA-A of one virus coexists with and transreplicates the DNA-B of another virus has been reported for begomoviruses from both the Old and New Worlds. It has been described between EACMV-UG2 DNA-A (which contains a recombinant fragment between ACMV and EACMV-UG1) and EACMV-UG3 DNA-B in Uganda (Pita *et al*., 2001). EACMV and SACMV recombined to form members of a new species, EACMKV (Bull *et al*., 2006).

The viability of recombinant genomes could be influenced by the degree to which their co-evolved protein-protein and protein-nucleotide and nucleotide-nucleotide interactions are disreputable by recombination (Martin *et al*., 2011). Possession of identical iteron sequences seems to be one of the pre-conditions for the formation of viable recombinants. Both EACMV and EACMKV have the same iteron sequences

(GGGGG) and form viable recombinants (Bull *et al*., 2007). Surprisingly, ACMV and EACMZV did not form viable pseudorecombinant despite having identical iteron sequences (GGAGA), indicating that there could be additional factors required for transreplication.

Notwithstanding, recombination among different DNA-A components remains the main source of molecular variation among geminiviruses and often results in gain of virulence for the helpers and acquisition of new satellite molecules. The frequency of recombination in these ssDNA viruses can partially be explained by two facts: one, a recombination-dependent replication mechanism that occurs in several begomoviruses and the satellite molecules they support; two, mixed virus infections (Pita *et al*., 2001) which provide preconditions for recombination.

2.4.5 Recombination Hotspots

Factors such as local degrees of sequence similarity strongly influence where recombination breakpoints occur (Martin *et al*., 2011). The major recombination hotspot in begomoviruses appears to be around IR containing the *ori* (Navas-Castillo *et al.*, 2000; Martin *et al*., 2011). This hot spot may be because DNA-A donates frequently its *ori* to other DNAs. In *Tomato yellow leaf curl Mali virus* (TYLCMLV), a recombination event occurred between *Tomato yellow leaf curl mild* strain virus (TYLC-Mld) and *Hollyhock leaf crumple virus* (HoLCrV) and involved the exchange of LIR (Chen *et al*., 2009). Recombination within the CP (*AV1*) genes as shown by ACMV and EACMV to produce EACMV-UG2 is the most common in begomoviruses. The complementary-sense genes show relatively higher frequency of recombination than the virion-sense genes.

Selection of EACMV recombinants with different CP genes appears to be favoured by prevailing local vector populations. In the absence of vector selection, the most powerful selection pressure by different host species will be on genes *AC1* to *AC4* that interact predominantly with host factors. Recent reports show that *AC1* (Rep) at position around 2000nt is particularly vulnerable to recombination. Recombination

hot spots other than the IR, *AV1* and *AC1* gene, appears to be centered on overlapping gene regions such as the *AV2*/*AV*1, *AV1*/*AC3*, and *AC3*/*AC2* junctions.

According to findings by Zhou *et al* (1997), the IR, the *AV1* (CP) and the *AC1* regions are parts of the DNA-A which differs significantly among whitefly transmitted geminiviruses - evidence that they were acquired by a recombination with other viruses. It follows that great similarity between IRs of two geminivirus isolates constitutes strong evidence that they are closely related. Also, the CP of whitefly transmitted geminiviruses is conserved to a greater extent than any other gene product.

The CPs of EACMV and ACMV have an identical 8-nt sequence. EACMV-UG CP and ACMV CP are identical but differ from EACMV CP. Thus, although only 16% of their DNA-A is ACMV-like and 84% EACMV-like in the recombinant EACMV-UG2, the UG isolates would be grouped with ACMV in routine antigenic analyses. Hillocks (2002) has indicated that the hybrid recombinant, EACMV-UG2 has been reported in Uganda, Kenya, Tanzania, Sudan and the DRC.

2.4.6 Association of Begomoviruses with Satellites

In the context of epidemiology, the organization of begomovirus genomes has conventionally been grouped into three categories (Briddon *et al*., 2008):

(i) Monopartite begomoviruses in which the DNA-A component alone has been shown to cause wild-type disease symptoms such as *Tomato yellow leaf curl virus* (TYLCV) and *Tomato leaf curl virus* (ToLCV);
(ii) Monopartite begomoviruses, which are infectious to their respective hosts but are unable to induce typical disease symptoms. These include *Ageratum yellow vein virus* (AYVV), *Cotton leaf curl Multan virus* (CLCuMuV) and *Tomato yellow leaf curl China virus* (TYLCCNV); and
(iii) Bipartite begomoviruses, which require both DNA-A and DNA-B components for the induction of typical disease symptoms and which represent an overwhelming majority of the begomoviruses.

Majority of the Old World begomoviruses have monopartite genomes and are usually associated with satellites while the New World begomoviruses are exclusively bipartite and are not associated with satellites. Thus, recombination events are not restricted to virus genome components but takes place between begomovirus DNA-A and satellite DNA circles (Seal *et al*., 2006). In fact, the evolution of begomoviruses in the Old World is intimately linked and dependent upon the existence and evolution of their satellites.

2.4.7 What are Satellites

A satellite is a sub-viral agent composed of nucleic acid that depends on the co-infection of a host cell with a helper or master virus for their multiplication (Xie *et al*., 2010). Nucleotide sequences in the nucleic acids usually have lower density from the bulk nucleic acid such that they form a second 'satellite' band when genomic nucleic acid is separated along a density gradient on agarose gels. When a satellite encodes a coat protein in which its nucleic acid is encapsidated it is referred to as satellite virus.

2.4.8 Main Features of Satellites

The classical definition of a satellite in plant virology involves three main features:

(i) They are not capable of replicating (within the host cell) in absence of the helper virus,
(ii) They are not required for the replication cycle of the helper virus, and
(iii) They share little or no sequence similarity with the helper virus genome.

Satellites depend on the helper virus for their replication within a host cell. The basis of this dependence is not completely known but is thought that satellite genomes are replicated by RNA-dependent RNA polymerase (RdRP) encoded by the helper virus. Satellites differ from other sub-viral entities such as viroids, which are capable of independent replication in a host cell and from defective interfering (DI) RNAs, which are comprised of helper virus genetic sequences. Satellites are commonly associated with plant viruses (Briddon *et al*., 2003; Briddon and Stanley, 2006).

2.4.9 Classification of Satellites

In the last decade, fascinating biological and molecular characteristics of many satellite systems have been elucidated (Briddon *et al*., 2008). These have shown that satellites can occur in complex mixtures and can include intermediate sub-viral forms, which do not fit perfectly into the classical definition described above. Satellites have been implicated as regulators of gene expression or as part of *bona fide* ORFs.

The early described biologically potent satellites had ssRNA genomes ranging from 194 – 1376-nt long. They were well known in infectious RNA systems and have been found to be associated with about 28 different plant viruses. Such satellites include *Strawberry latent ringspot virus satellite* RNA and *Tomato black ring virus satellite* RNA. These satellites showed little or no nucleotide sequence similarity to the viral or host genomes, but were completely dependent on the helper virus for replication. Indeed, the satellites were found to exacerbate viral symptoms or induce symptoms quite distinct from those induced by the helper virus alone.

2.4.10 TLCV Sat-DNA Betasatellite and Alphasatellite

TLCV Sat-DNA: A small, circular DNA satellite found to be associated with a geminivirus in 1997. The 682-nt circular DNA satellite was associated with TLCV infection in Northern Australia (Dry *et al*., 1997). The DNA satellite (TLCV sat-DNA) was strictly dependent on the helper virus *Rep* for replication and was encapsidated by TLCV coat protein. TLCV sat-DNA does not encode any protein, and shows no significant sequence similarity to the 2766-nt helper-virus genome except for TAATATTAC and a sequence motif GGTGTCT, essential for binding the viral Rep *in vitro*. TLCV sat-DNA has no apparent effect on symptom expression.

Betasatellites: Most Old World monopartite begomoviruses are associated with one or more betasatellite DNA(s), which are required for the induction of typical disease symptoms (Saunders *et al*., 2004; Cui *et al*., 2004). They are small circular ssDNAs about 1350-nt in length. They are pathogenicity determinant molecules required by their helper begomoviruses in some host plants. First identified in 2000, betasatellites

are completely dependent on the helper component (DNA-A) for their replication, encapsidation and transmission by whitefly vectors.

DNA-β molecules (Figure 2.11) are widespread in the Old World and all contain three conserved features:

i) They all encode a single gene βC1
ii) They have a satellite conserved region (SCR),
iii) They have adenine-rich region approximately 370 to 420-nt upstream of SCR.

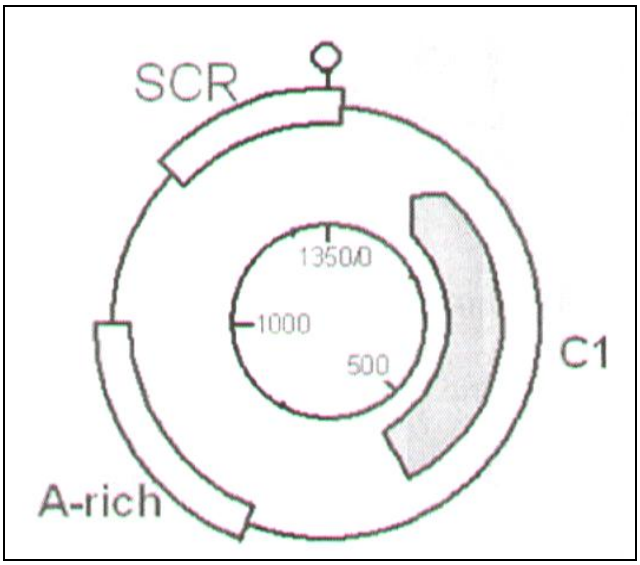

Figure 2.11: Genome organization of DNA-ß satellites. The position and orientation of the conserved C1 gene is shown as an arrow. The relative position of the satellite conserved region (SCR) and A-rich region are shown as shaded slices.

Due to these characteristic features, the majority of begomoviruses, which associate with betasatellites have one thing in common – they are poorly infectious in the host from which they were isolated and they induce atypical symptoms. Betasatellites-requiring begomoviruses are now known to outnumber the bipartites and truly monopartites in the Old World. Since they were first identified in 2000, over 260 full length sequences (Briddon *et al*., 2003) have been deposited with databases, and the number is said to increase on daily basis. All the satellites described so far have a similar genetic structure (Figure 2.12).

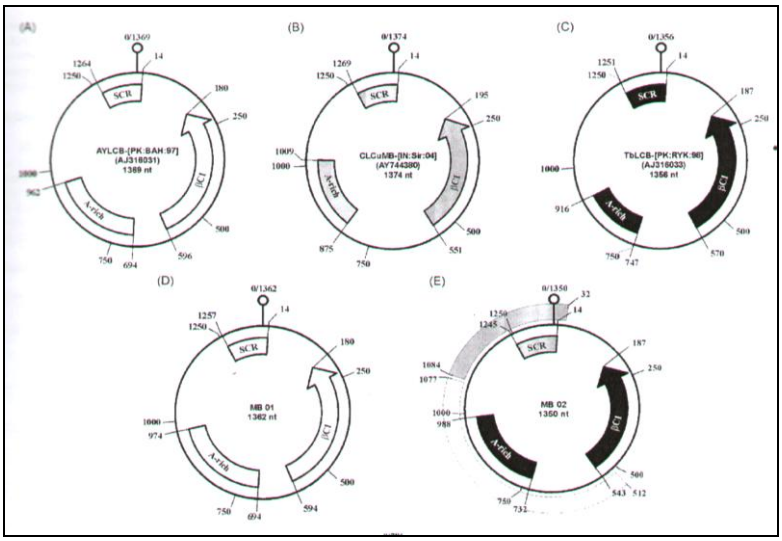

Figure 2.12: Representative betasatellites showing conserved structures. Mb 01 is an isolate of AYLCB closely related to AYLCV. MB 02 is a recombinant TbLCB derived from CLCuMB.

In Okra leaf curl disease (OLCD) for example, the associated betasatellite is 1346-nt and has the following characteristic features (Seal *et al*., 2006; Kon *et al*., 2009):

- Has a single complimentary-sense ORF (βC1),
- Contains an adenine (A)-rich region (nt 720 – 877 with 60% A residues), and
- Has a satellite conserved region (SCR) with predicted stem loop structure containing the geminivirus nanonucleotide sequence TAATATTAC.

The βC1 protein is multifunctional in the sense that it is a pathogenicity determinant, it is a suppressor of PTGS, it raises virus DNA levels in plants and is involved in virus movement in plants.

Within the SCR are also sequence motifs GGTDKN (D = A, G, T; K= G, T; N = A, C, G, T), which are involved in Rep binding. The propensity of betasatellites to be trans-replicated by more than one begomoviruses is possibly a contributing factor to their success. This is because betasatellites are clearly associated with their specific helper component irrespective of hosts and geographical distribution (Briddon *et al*., 2003; Jose and Usha, 2003).

As a result, there exists a loose specificity of trans-replication of betasatellite by a different helper component. For example, CLCuMuV cannot trans-replicate *Ageratum yellow vein betasatellite* (AYVB), whereas AYVV can trans-replicate *Cotton leaf curl Multan betasatellite* (CLCuMuB). On the other hand, *Tomato leaf curl betasatellite* (ToLCB) has been reported to be trans-replicated by the ACMV and BCTV, which are highly diverse in terms of their geographical distribution and molecular diversity. Despite their recent discovery, betasatellites have existed for many centuries. The betasatellite associated with *Eupatorium yellow vein virus* (EpYVV) has been linked to disease symptoms described some 1,250 years ago.

Alphasatellites: Alphasatellites (previously called DNA-1) (Xie *et al*., 2010) are satellite-like, single stranded DNA molecules of approximately 1370-nt associated with begomoviruses that require the satellite molecule DNA-β to induce authentic disease symptoms in some hosts. A typical example of a disease complex involving alphasatellite is Okra leaf curl disease (OLCD) observed in Mali and Pakistan (Kon *et al*., 2009). In the Mali OLCD case, two monopartite begomoviruses (okra virus-1 and okra virus-2), a betasatellite and an alphasatellite were involved. Other disease complexes are Cotton leaf curl disease (CLCuD) in Pakistan and *Ageratum* yellow vein disease (AYVD) in Singapore.

Alphasatellites have a highly conserved genome organization encompassing a conserved major common region (CR-M), a potential TATA box 3'of the stem loop, have a single product in the virion-sense with similarity to the replication associated protein (Rep, a rolling circle replication initiator protein) of nanoviruses, a region of ≈200nts sequence rich in adenine, a predicted hairpin structure containing an origin of replication (Ori) (including a conserved nanonucleotide TAGTATTAC), with similarity to the origin of virion-sense DNA replication of nanoviruses (typical example of which is *Banana Bunchy Top virus*, BBTV. Consequently, alphasatellites are capable of autonomous replication but are dependent on the helper virus for movement, encapsidation and vector transmission. Typical examples of alphasatellites include *Potato leaf curl alphasatellite* (PotLCA), *Tomato yellow leaf curl Yunnan alphasatellite* and *Tomato yellow leaf curl China alphasatellite* (Cui *et al*., 2004).

2.4.11 Recombination and Reassortment between Satellites and Begomoviruses

Geminiviruses and satellites can undergo recombination or reassortment. For the case of recombination, the begomovirus CR can be exchanged with the SCR, though not all betasatellites show this structural modification. CR recombinant betasatellite have been characterized for AYVV and TYLCCNV. The recombinant betasatellites lack the normal SCR but encode a functional *BetaC1* gene responsible for pathogenicity. ToLCV-associated satellite from Australia, for example, lacks BetaC1 gene but comprise an A-rich and SCR. Systemic infection in this case does not require the BetaC1 gene. The recombination in which BetaC1 is deleted or in which the betasatellite SCR is replaced by a geminivirus CR usually result in the most virulent betasatellite comprising an SCR, an A-rich region and a BetaC1 gene.

In some Old World begomoviruses, the DNA-B component is not essential and is exchangeable among geminiviruses and with betasatellites. This is called pseudorecombination (or reassortment) and results in a severe infection phenotype. In principle, pseudorecombination involves one viral protein (Rep) and sequences in the CR, especially the iterons. Iterons are virus-specific binding sites for cognate Rep protein to initiate virus replication. All begomovirus iterons are composed of an invariant GG sequence followed by three nts named N1, N2 and N3 that vary among species (Andrade *et al*., 2006). Examples of iteron sequences for some viruses are as indicated in brackets: *Tomato yellow spot virus*, ToYSV (GGTGA), *Tomato crinkle leaf yellows virus*, TCrLYV (GGTGA), *Sida mottle virus*, SiMoV (GGAGT) and EACMZV (GGAGA). In theory, iteron divergence between viruses does not allow pseudorecombination while conservation between the iterons allows the formation of pseudorecombinants.

Cases in which satellites form reassortants with begomoviruses also exist. The interaction between TYLCMLV and *Cotton leaf curl Gezira betasatellite* (CLCuGB) (Chen *et al*., 2009), produce a reassortant that is introduced into tomato by the polyphagus whitefly vector. In a bipartite begomovirus *Tomato leaf curl New Delhi virus* (ToLCNDV), both DNA A and DNA B are required for systemic infection. The inoculation of tomato plants with DNA A alone induced local but not systemic

infection while co-inoculation with DNA A and DNA-ß associated with cotton leaf curl disease resulted in systemic infection (Saeed *et al*., 2007). This clearly indicates that the DNA-ß substitutes for DNA B (reassortment) in the bipartite begomovirus ToLCNDV to permit systemic infection.

Sri Lankan cassava mosaic virus (SLCMV) may also have been a satellite-requiring begomovirus. The SLCMV has exchanged the satellite for a DNA B component, most likely captured from ICMV, but nevertheless is able to productively interact with betasatellite. The frequency with which reassortants occur in the field depends on whitefly population densities and feeding habits. Further, it depends on the proximity of tomato fields and malvaceous crops and weeds.

2.4.12 Effect of Satellites on Host and Helper Virus

By virtue of their constitution and activities in host cells, satellites often modulate the replication of their helper viruses. The symptoms caused by the helper virus/satellite co-infection can differ dramatically from those of the helper virus alone. One of the possible mechanism by which satellites control cellular processes (DNA replication, recombination and repair, transcription and viral assembly) is through interaction of proteins with helper virus DNA. Begomoviruses and satellites show some similar features (as indicated below), no wonder they form effective infectious components.

Nawaz-ul-Rehman and Fauquet (2009) point out four common features between begomoviruses and satellites:

(v) Recombination-dependent replication mechanism – the rolling circle replication (RCR) mechanism,

(vi) Presence of a stem loop containing a nanonucleotide origin of replication TAATATTAC or TAGTATTAC, and

(vii) The genetic architecture, localization, length of individual genes at specific locus, for example both have single stranded circular genomes. Also, function of individual proteins is highly conserved in both begomoviruses and satellites.

(viii) Begomoviruses and their associated DNA satellites replicate in differentiated cells that are in G phase and have shut down most of their DNA replication activities.

Thus, to successfully invade new hosts, break host resistance, move virus particles within and between plants, geminiviruses and their satellites have evolved a coordinated network of protein interactions, showing a possible evolutionary path.

Betasatellites depend on begomovirus for their replication, movement in plants, and insect transmission between plants. In turn, the helper begomovirus depends on the betasatellite for efficient infection of hosts, possibly by the *BC1*-mediated PTGS suppression of a host defense response. In course of their interaction, disease symptoms are enhanced. In order for a viable pseudorecombinant to be formed, the factors responsible for replication and movement must be interchangeable between the non-cognate DNA components. One of the sequence elements determining this interchange are the iterons located within the CR.

2.4.13 Infectivity of Pseudorecombinants

Andrade *et al* (2006) suggests the infectivity (or viability) of pseudorecombinants to be of four main categories as follows:

(a) Non-viable pseudorecombinants, which do not induce symptoms and viral DNA cannot be detected in newly emerged leaves. Examples are pseudorecombinants between ToYSV DNA-A/ToRMV DNA-B and ToYSV DNA-A/*Tomato golden mosaic virus* (TGMV) DNA-B. In these reassortants there is no conservation among the iterons. Also, the iterated recognition domain (IRD) sequences show poor conservation, with the specific amino acids potentially responsible for iteron recognition not fully conserved;

(b) Pseudorecombinants, which do not induce symptoms but viral DNA (DNA-A) can be detected in newly emerged leaves. Examples of such pseudorecombinants include ToRMV DNA-A/ToYSV DNA-B and ToYSV DNA-A/*Tomato chlorotic mottle virus* (ToCMoV) DNA-B. The capacity of ToRMV DNA-A to infect *N. benthamiana* in absence of cognate DNA-B has

been reported. However, ToYSV DNA-A alone did not infect *N. benthamiana* indicating that a very low amount of ToCMoV DNA-B is sufficient to allow the DNA-B encoded proteins to move ToYSV DNA-A from cell to cell. Thus, symptom determinants in begomovirus infections are encoded by DNA-B.

(c) Pseudorecombinants, which induce severe mosaic, leaf curling and epinasty in *N. benthamiana*. Examples include ToYSV DNA-A/TCrLYV DNA-B and TGMV DNA-A/ToYSV DNA-B. In this case, CR alignment revealed that ToYSV and TCrLYV have identical iterons (GGTGA), which contributed to the viability of this pseudorecombinant. However, ToYSV and TGMV iterons are not identical (GGTGA and GGTAG, respectively). This suggests that TGMV iterons might be more versatile and recognize a heterologous cognate DNA component with different iterons. Analysis of the amino acid sequence of Rep proteins indicated that the IRDs of TGMV and ToYSV are the most similar among all analysed begomovirus. The sequence conservation allowed iteron recognition by the heterologous Rep protein and that the gene products encoded by the two heterologous components were capable of interacting efficiently, allowing the development of a systemic infection with severe symptoms.

(d) Pseudorecombinants, which enhance symptoms in *N. benthamiana* in the presence of betasatellite. These include CbLCuV in the presence of CLCuMuB and AYVV.

There are four possibilities about the mechanisms through which satellites enhance infectivity and systemicity of begomoviruses:

(i) Different begomovirus DNA-A sequences possess different iteron sequences in their IR region, which facilitates sequence-specific Rep binding to initiate RCR (Bull *et al*., 2007). All betasatellites have SCR which contains motifs GGTDKN [D = A, G, T; K = G, T; N = A, C, G, T]. The GG motif relates to the conserved begomovirus iteron sequences. Thus, it is possible that the satellites contain iteron sequences of their helper begomoviruses, thereby enhancing the viability of their recombinants;

(ii) The reassortant is made highly viable by the promiscuous nature of betasatellites and by the recombinant region derived from the malvaceous begomovirus (Chen *et al*., 2009); and

(iii) The betasatellites are thought to stimulate the division of terminally differentiated phloem-associated cells thereby generating more cells for viral infection.

(iv) For bipartite begomoviruses as in ToLCNDV and SLCMV, the virus undergoes recombination in which DNA-B is replaced by a betasatellite (Saeed *et al*., 2007), enhancing its movement.

2.5 INTEGRATION OF VIRAL AND SATELLITES SEQUENCES INTO HOST GENOME

2.5.1 What is Integration

Integration is a process whereby foreign sequences are expressed in plants (Porta and Lomonossoff, 2002). This expression can occur either through stable genetic transformation (transgenic approach) or through plant-virus based vectors. In stable genetic transformation heterologous genes are integrated into chromosomes of a host plant in which expression of the integrated gene is under the control of a promoter immediately upstream of the coding region. The integrated sequence is heritable.

2.5.2 Integration of Viral Sequences

This is situation whereby foreign genes of viral origin or genes incorporated into viral genome are concomitantly amplified in plants when viruses multiply within infected cells (Porta and Lomonossoff, 2002). Many plant viruses tend to have their sequences integrated into host genomes. Some sequences identical to an RNA dependent RNA polymerase (RdRP) gene of *Cardamom mosaic virus* (CdMV) for example, have been found in cardamom plant genome (Jebsingh *et al*., 2011). Of the three families of plant viruses, which have DNA genomes: *Geminiviridae*, *Circoviridae* and *Caulimoviridae* (Harper *et al*., 2002), members of two families (*Caulimoviridae* and *Geminiviridae*) are involved in integration.

Integrated *Caulimoviridae* viruses have included *Banana streak virus* (BSV) in *Musa* spp (Gayral *et al*., 2008), *Tobacco vein-clearing virus* (TVCV) in tobacco, *Petunia vein-clearing virus* (PVCV) in petunia (Harper *et al*., 2002; Harper *et al*., 2003; Hohn *et al*., 2008) and *Cauliflower mosaic virus* (CaMV) in cauliflower (Porta and Lomonossoff, 2002). In the *Geminiviridae*, begomoviruses such as ACMV (*ibid*.) have been able to replicate when expressed from integrated sequences. This happens for example, when ACMV's CP gene is replaced by bacterial chloromphenicol acetyltransferase (CAT).

Integrants can be of two forms - those that can form episomal viral infections and those that cannot (Ndowora *et al.* 1999; Harper and Hull 1999; Harper *et al*., 2002). In the *Caulimoviridae*, most of the reported plant pararetroviruses are defective and do not trigger virus infection. A few examples however, exist that have now been determined that the genome is integrated in the genome of the host plant in such a way that after activation (stress) the integrated genome fragments can recombine to produce an episomal virus infections in certain plant hosts. This includes PVCV, TVCV and such distinct BSV species as *Banana streak OL virus* (BSOLV), *Banana streak Imove virus* (BSImV), and *Banana streak GF virus* (BSGFV) (Harper and Hull, 1998). Different abiotic stresses have been identified that trigger the production of viruses from infection from integrated viruses: temperature differences, water stress, *in vitro* culture and interspecific crosses (Cote *et al*., 2010), which leads to the production of hybrids.

2.5.2.1 Noninfectious endogenous BSV

The majority of cultivated banana lines originated mainly from inter- and intraspecific crosses between two wild species *Musa acuminata* (AA genome) and *M. balbisiana* (BB genome) (Simmonds and Shepherd, 1955). A series of diploid, triploid and tetraploid hybrid genomes (AA, AB, AAA, AAB, ABB, AABB, AAAB and ABBB) resulting from such crosses is used in three types of food:

(a) As dessert banana fruit,
(b) In banana for cooking (including plantains); and
(c) In banana for beverages (beers).

Banana streak disease is caused by a complex of BSV, all transmitted in a semipersistent manner by three species of mealybug, of which *Planoccocus citri* is the most prevalent (Meyer *et al*., 2008). BSV infections cause characteristic discontinuous chlorotic areas turning into necrotic streaks on leaves, and a split pseudostem. A severe form develops heart rot of the pseudostem which, eventually leads to death of the plants.

A range of variation in severity is reported depending on several factors such as virus species, host genotype, level of management, and environmental conditions such as temperature and water stress. Field observations suggest that natural dissemination of BSV by mealybugs is limited and does not play a major role in epidemiology. The most significant transmission is likely to occur through vegetative propagation by exchanges of suckers and plantlets (Lockhart and Jones, 2000). This is a classic triangular pathogenic interaction.

2.5.2.2 Infectious endogenous BSV (eBSV)

Recently, there have been outbreaks of the disease in banana among micropropagated interspecific *Musa* hybrids in areas with no external virus source as the origin. The use of BSV detection tools such as PCR indicated the widespread presence of viral sequences in *Musa* plants. The origin of these field outbreaks was correlated with the presence of infectious endogenous BSV (eBSV) sequences within *M. balbisiana* genome (Cote *et al*., 2010).

2.5.3 Integration of Subgenomic DNAs and Satellite Sequences

Until recently, the integration of subgenomic and satellite sequences in host genomes have been *naïve* and have often been adapted in molecular strategies to control viruses (Fauquet and Beachy, 1992). Subgenomic DNAs and satellite RNA have been used as sense approaches in replication strategy for virus control. Subgenomic DNAs are produced when full-fledged viruses such as the ACMV infect host plants. Satellite (SAT) RNAs are associated with several viruses in which they are dependent upon a helper virus for their replication and spread in infected plants. The presence of SAT

RNAs in *Cucumber mosaic virus* (CMV)-infected tobacco reduced disease symptoms. The same has been observed for SAT *Tobacco ringspot virus* (SAT TobRSV).

It is noteworthy however, that not all satellite sequences provide symptom attenuation and sometimes they can cause necrosis, leading to dramatic symptoms when naturally infected by the helper virus. The case of unusual severity of cassava mosaic disease can be linked to the occurrence of infectious endogenous virus sequences. Much as it is known that integration can involve both single-stranded as well as double-stranded DNAs as reported in the BSV case (Ashby *et al*, 1997), there could be a possibility of satellite DNA molecules being integrated into the cassava genome.

The discovery of SatDNA-II and SatDNA-III came up before cassava genome was sequenced (Dr Joseph Ndunguru *pers. Comm.*). When sequences of begomovirus-associated SatDNA-II and SatDNA-III were blasted onto the expressed sequence tags (ESTs) for cassava, sequences with homology to the satellites were found to be integrated into the cassava genome (Ndunguru *et al*., 2005). An assembly of 7 cassava unigenes partially resembling SatDNA-II and 11 others sharing sequence similarity with SatDNA-III were found in the ESTs. The fact that the severe forms of cassava mosaic disease were found in the Lake Victoria zone could suggest the prevalence of certain environmental conditions, which constituted a sufficient stress to illicit the symptom-modulating satellites.

2.5.4 Impact of Sequence Integration in Cassava

Plant cell growth and development depend on continuous cell proliferation, which is restricted to meristems. Infecting geminiviruses and/or the satellites they support depend on replicative factors supplied by the host cell to complete their replicative cycle. Since most of the replicative factors are present in the proliferating cells, geminivirus infection is believed to induce a cellular state permissive for viral DNA replication, such as the S-phase or at least some specific S-phase function (Gutierrez, 2000).

Where some virus sequences or those of their satellites become integrated into the genome, the said cellular state remains a characteristic of the plant, i.e. the chromosomes carrying the integrant replicate with it. Just as deliberate integration of

virus sequences is expected to make the plant resistant to viruses, spontaneous integration of infectious episomes can spontaneously leave the host plant more susceptible to infection.

Thus, the process of integration has considerable implications for plant genome evolution:

(i) The integrated viral DNA could act as an insertional mutagen. The foreign DNA sequence disrupts the expression of the gene into which it is inserted but also acts as a marker for subsequent identification of the mutation.
(ii) The sequence could contribute strong constitutive promoters to neighboring plant genes,
(iii) The DNA could accumulate to generate new repetitive sequence family.

The incorporation of viral sequences into host genome can have functional implications for the virus-host interactions. It can render the host susceptible or resistant to diseases (Bertsch *et al*., 2009). Sometimes, the integration process may be latent. Virus latency (or viral latency) is the ability of a pathogenic virus to lie dormant within a cell.

2.6 ASSOCIATION OF SATELLITES WITH OLD WORLD BIPARTITE BEGOMOVIRUSES

2.6.1 Early CMD situation in Tanzania

Until 1997, only ACMV and EACMV were known to occur in Tanzania, with the former occurring in the Western part of the country (Ogbe *et al*., 1997). Disease symptoms due to EACMV were always mild except when the plant was concurrently infected with ACMV. The synergistic severe mosaic disease situation was observed in cassava and tobacco and was characterized by the increase in viral-DNA accumulation of both viruses compared to single infection.

It was later confirmed that ACMV TrAP encoded by gene *AC2* played the role of circumventing the host defense mechanism characterized primarily by an increase in viral infectivity (Hao *et al*., 2003). *AC2* as well as *AC4* proteins of cassava

geminiviruses play a role in synergism and the regulation of anti-PTGS activity. It has been demonstrated that ACMV-[CM] *AC4* and EACMV-TZ *AC2* in *trans* are able to suppress PTGS induced by green fluorescent protein (GFP) and eliminate siRNAs associated with PTGS with a concomitant increase in mRNA accumulation (Vanitharani *et al*., 2005).

2.6.2 Occurrence of Satellites in Tanzania

By the year 2004, two main factors were known to aggravate CMD situation in Tanzania – the discovery of EACMZV and the emergence of ACMV-EACMV recombinant designated EACMV-UG, which was spreading from Uganda into Tanzania. Recent findings have indicated that EACMV alone has been able to cause very severe CMD symptoms (Ndunguru *et al.*, 2005) contrary to what was known before. Further investigations revealed the presence of two subgenomic constructs with characteristics typical of satellites. Their sizes were ranging between 1.0 and 1.5kb; their replication depended on a helper virus; they had no sequence similarity to the helper virus.

The satellites were named satellite DNA II (SatDNA-II) and III (SatDNA-III) and both were associated with bipartite begomoviruses (Ndunguru *et al*., 2008). In the Genebank, the satellites have been referred to as African cassava mosaic virus-associated DNA-II and DNA-III, both belonging to the Genus *Begomovirus-associated Satellite DNA*. The satellites have been shown to exacerbate CMD severity caused by cassava mosaic begomoviruses. SatDNA-II induced very severe symptoms characterized by leaf distortion, yellowing and mosaic. SatDNA-III caused unique symptoms in cassava characterized by severe leaf narrowing and prominent leaf yellowing.

In the epidemiology of CMD, SatDNA-II was found to be associated with EACMV-UG while SatDNA-III was associated with EACMV-TZ. SatDNA-II was found along the Eastern coast and the Lake Victoria zone while SatDNA-III was found in the Lake

Victoria basin alone. Thus, the discovery of these satellite DNA molecules in Tanzania is considered as a third factor in the intricacy to manage CMD.

2.6.3 Features of SatDNA-II and SatDNA-III

The satellites (Figure 2.13) are circular ssDNAs, distinct from each other (23% nts identity) and possess putative open reading frames suggesting their possible ability to code for functional proteins.

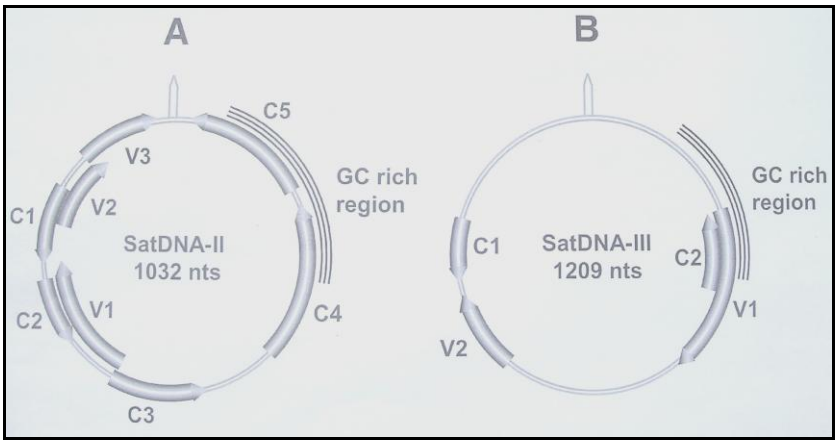

Figure 2.13: Genome organization of SatDNA-II and SatDNA-III (**Source:** Ndunguru, 2005).

SatDNA-II has eight putative ORFs, three on the virion sense and five on the complementary sense. The longest ORF is 159nt with potential coding capacity of 6 kDa. SatDNA-III has 4 ORFs two on either sense. The most important characteristic feature of the SatDNAs however, is the breaking down of resistance even in the long-known resistant cultivars such as TME3 (Ndunguru *et al*, 2008). The satellites, which were only associated with bipartite begomoviruses are dependent on the begomoviruses for replication and movement within the plant. They are mechanically transmissible into *Nicotiana benthamiana* but not in cassava. Transmission into cassava was only successful through particle bombardment.

The genomes of satellite DNA molecules associated with cassava begomoviruses are small, 1032 bp for SatDNA-II and 1209 bp for SatDNA-III (NCBI Accession Numbers AY836366 and AY836367.1, respectively). Genomic structures of the satellites do not possess the SCR and hence TAATATTAC, which is found in all the begomoviruses and SatDNA-β or TAGTATTAC found in SatDNA-1 and nanoviruses.

The subviral agents do not encode coat protein or any other ORF, indicating that they are not typical satellite viruses, but simply satellite nucleic acid molecules. There is no origin of replication that can be distinguished from the satellites. The insidious molecules have a characteristic GC-rich region characterized by the presence of direct repeats of short hexanucleotides CCGCCG, pentanucleotides CCGCC and trinucleotides CGC. Ndunguru *et al* (2005) found that both SatDNA-II and SatDNA-III have putative TATA Binding Protein (TBP) motifs, one in SatDNA-II (TATAAAT) and three in SatDNA-III (GATATAAATA, TACATATATAT and TCTGTATATA).

The genomes of both SatDNAs were found to have putative consensus transcription poly (A) signal AATAAA, with a motif TTGTA positioned upstream. This shows that the poly (A) and hence the SatDNAs are biologically functional. Sequence search for recognizable *Rep* binding element similar to the one found in EACMV isolates (GGTGGAATGGGG) and ACMV isolate (GGAGACA) was unsuccessful in both satellites. Furthermore, the nucleotide and amino acid sequences of these molecules did not match any known geminivirus cassettes deposited in gene banks.

2.7 CMD MANAGEMENT

2.7.1 Overview of CMD Management

Disease management refer to deliberate interventions on crop plants aimed at reducing devastating effects by diseases to levels which do not cause economic yield losses. For the case of CMD, three stages of management are involved:

(i) Monitoring the emergence, spread and distribution of cassava mosaic begomovirus (CMB) strains and associated satellites, through regular diagnostic surveys and molecular characterization;

(ii) The forecasting of future spread of CMD and satellites in cassava growing areas as well as other areas, achieved by the diagnostic surveys to establish spread of the disease in spatial and temporal dimensions; studying the population and distribution of the whitefly vector; and by studying the CMD epidemic characteristics (i.e. the amount and relative proportion of whitefly-based and cuttings-based infections);

(iii) The designing of appropriate control measures

It is considered (Thresh *et al*., 1998; Fraser, 2006; Mallowa *et al*., 2006) that control measures of plant virus diseases fall into three broad categories:

(i) Prophylactic measures, which includes a broad span of approaches to avoid or exclude pathogen, including phytosanitation and use of virus-free seeds and vegetative propagules. The measure also involves preventing and reducing virus dissemination.

(ii) Breeding for resistant varieties, which include search for and characterization of virus resistances levels and durability.

(iii) Genetic engineering for plant protection, which involves various strategies of gene transfer at different levels.

Although these methods are available for use either singly or in combination, host plant resistance is the most widely used approach in both national and regional CMD management programmes. However, the other methods are described with particular focus to CMD management.

2.7.2 Prophylactic Control Measures

Prophylactic (cultural) disease control measures cover an array of measures aimed at reducing virus dissemination or preventing infection altogether. Prophylactic measures also involve the use of virus-free seeds and vegetative propagules. Unlike cultural approaches to the control of diseases in other crops, little attention has been

given to the possible control of CMD through managing populations of its whitefly vector, due to one or more of the following reasons:

(i) The whiteflies transmit viruses in non persistent manner and because of their high mobility, the principle of the economic injury level (EIL) does not work;
(ii) The difficulty associated with the control of whitefly populations in cassava throughout its growth period;
(iii) The absence of clear correlation between populations of whiteflies supported by a given variety and the spread of CMD.

Another whitefly-related concern in the management of CMD is that some of the newly-released CMD-resistant varieties being used to tackle the pandemic appear to be particularly favourable hosts of *B. tabaci* (Thresh *et al*., 1998). The varieties tend to support populations sufficiently large to cause physical damage.

This raises two important implications:

- Direct damage due to vector feeding may have an economically important effect on productivity of these new varieties than the effect of damage by disease (CMD).
- The increased number of vector-virus-host contact that occurs may increase the probability of resistance break down in those cultivars.

In a broad sense, use of virus-free planting material refers to phytosanitation, which comprises all those techniques that aim to keep plants of the crop or variety being grown in a virus-free condition.

Thresh *et al*. (1998) and Mallowa *et al*. (2011) describe three main features of phytosanitation for the control of CMD:

(i) Crop hygiene, involving the removal of all diseased cassava plants from within and immediately around areas to be used for new plantings;
(ii) The use of CMD-free stem cuttings as planting material; and
(iii) The removal (roguing) of diseased plants from within crop stands.

The most commonly advocated phytosanitation approaches are the removal of CMD-diseased plants from within a crop stand and the selection of symptom-free cassava stems for planting subsequent crop.

Phytosanitation also involves more sophisticated approaches to the provision of virus-free germplasm involving the cleanup of tissue culture material through meristem tip culture and thermotherapy. It further involves the screening of imported germplasm in post entry quarantine support facilities, consistent with requirements of the International Plant Protection Convention (IPPC) (Meester *et al*., 1999; Ndomba, 2005).

2.7.3 Conventional Breeding for Resistant Varieties

Although CMD was reported in as early as 1894, it did not become a serious problem until the late 1920s and early 1930s when agriculturists in East, Central and West Africa and also in Madagascar became aware of the need for CMD resistant varieties (Thresh *et al*., 1998). From the early years of research into CMD, it was apparent to researchers that cultivars varied in their response to the disease (Okogbenin *et al*., 2007).

CMD-resistant varieties selected in Tanzania, Madagascar, IITA and elsewhere in Africa had several important characteristics, which seem to show the same basic virus resistance mechanism:

(i) They developed CMD symptoms that were less conspicuous than those of sensitive (intolerant) varieties;
(ii) The symptoms were often restricted to certain shoots or branches and became inconspicuous or disappear as the plants aged;
(iii) CMD viruses such as ACMV were less systemic in resistant plants than in susceptible ones. As a result, a large proportion of the cuttings collected from infected stands were free from ACMV even if taken from plants that were infected. This is called reversion.

(iv) Resistant varieties contained lower concentration of ACMV than susceptible ones.

Host plant resistance is a major component in the control of virus diseases in crop plants (Fraser, 2006) and also of integrated pest management (IPM) strategies for the control of virus epidemics. However, the sustainability of host plant resistance is often limited by the emergence of resistance-breaking virus strains (Urbino *et al*., 2008). For cassava begomoviruses, the resistance-breaking phenomenon is made worse by the association of viruses with satellites.

Overall though, resistance breeding is ought to include:

(a) Germplasm introduction of CMD-resistant material using open quarantine procedure in recently affected areas;
(b) The development of a 'fast-track' approach to evaluate new CMD-resistant germplasm through on-station and on-farm trials; and
(c) Participatory evaluation of new germplasm to technology transfer centers and within farmer field schools.

2.7.4 Genetic Engineering for Protection

Since 1986, the ability to confer resistance against an otherwise devastating virus by introducing a single pathogen-derived or virus targeted sequence into the DNA of a potential host plant has had a marked influence on much of the research effort, focus and short term objectives of plant virologists. Much of the genetic engineering approaches have focused on pathogen-derived resistance (PDR) (Sansford and Johnston, 1985). The PDR proposes that the expression of certain genes of a pathogen, in this case a virus, in a host would disrupt the normal balance of viral components and thereby interfere with the virus life cycle. Three molecular strategies - ribozyme strategy (Huttner *et al*., 2001; Prins *et al*., 2008), translation strategy and replication strategy (Dale and Harding, 2003) are known to be potential in the management of plant viruses (Fauquet and Beachy, 1992). Several groups have

attempted this approach and have led to the production of transgenic plants that are resistant to several CMGs.

A novel approach for CMD resistance has been to genetically transform cassava with genes conferring resistance to the begomoviruses. In a recent twist, reports *The EastAfrican* weekly (September 11 – 17, 2006), a leading American research facility, the Donald Danforth Plant Science Center, has admitted that varieties of genetically-modified cassava that it had considered to be disease-resistant are declared to be vulnerable to the devastating CMD. The statement further indicated that though resistance to CMD had been established through genetic engineering seven years before, the resistance was subsequently lost due to the virulence of CMD viruses.

Development of transgenes in cassava and many other crops has never been without difficulties. Several potential impacts of virus resistance transgenes on the environment have been highlighted since the early 1990s. Thanks to the improved scientific understanding of the underlying phenomena, some of the risks thought to be serious are now considered to be relatively unimportant. Two public concerns however, deserve a mention.

The first serious potential risk was considered to result from recombination between viral transgene mRNA and the genomic RNA of a non target virus (Prins *et al*., 2008). However, it was recently shown that in cucumoviruses that similar populations of recombinant viruses appear in transgenic plants expressing a CMV CP proteins infected by another cucumovirus and equivalent non transgenic ones infected simultaneously with the two cucumoviruses. This indicated that transgenic plants do not contribute to the generation of recombinant viruses that would not have been generated in natural double infections.

A second of the original concerns (*ibid*.) is whether sexual out-crossing between the transgenic virus-resistant plant and a wild relative could have significant effect in the wild plant's fitness and invasiveness/weediness. This concern has been subdued by the fact that there quite a few examples in which virus infection has a strong impact on plant's fitness. Also, if indeed change in fitness ensued, the increased fitness does not necessarily lead to increased population size (implying invasiveness).

2.8 RESISTANCE AS AN INTERVENTION TO CMD

2.8.1 Viral Systemic Infection in Plants

A successful viral pathogen must necessarily multiply, a stage which requires 'a replication-susceptible' host. This is followed by a stage which requires a host permissive to invasion of cells adjacent to the primary infection court, and ultimately systemic spread (Khan and Dijkstra, 2001). Systemic plant infection is a critical step in viral pathogenicity, during which the virus spreads from the initial site to distal parts of the plant through the vascular system. The CP is the viral protein to which a function in vascular movement has been most frequently ascribed.

The dispensability of the CP for begomovirus systemic spread is dependent on both viral-specific and host-specific functions. It is important to note that CP is not required for systemic infection by bipartite begomoviruses indicating that virus DNA may be the principal form of transport in host plants. In bipartite begomoviruses NSP substitutes for the function of CP. Because of these events, it appears that the most effective means of controlling CMD has been through the deployment of resistant varieties (Thresh *et al*., 1997).

2.8.2 Components of Resistance to CMD

In the realm of plant infection by viruses, plants can be susceptible, resistant or immune. A susceptible plant allows the replication of virus particles as well as the movement of the virus through the entire plant. In resistant plants, virus activity is restricted to inoculated cells.

Fauquet and Fargette (1990) have identified six different components of resistance to CMD:

(d) Field resistance (= percentage of infected plants in the field);
(e) Vector resistance (= average number of adult whiteflies per plant);
(f) Inoculation resistance (= minimum concentration of inoculum required to elicit infection);
(g) Virus resistance (= virus content in host tissues as estimated by ELISA);

(h) Symptom severity (= number of symptomatic leaves and leaf areas affected); and

(i) Virus movement resistance (= duration of symptom development on uninoculated leaves).

There are at least four features as observed in the field, which are indications of CMD resistance (Sseruwag *et al.*, 2004; Fraser, 2006) in cassava plants:

(a) Recovery of plants from infection,
(b) Reversion of disease symptoms,
(c) Occurrence of mild symptoms on few leaves, and
(d) Localization of symptoms on young shoots.

To avoid inconsistencies in the terminology relating to host resistance, a host plant is resistant if it can suppress the multiplication of a virus, and consequently suppress the development of disease symptoms. Cassava mosaic begomoviruses are not always fully systemic. Uninfected cuttings can be obtained from some branches of infected plants, especially those of resistant varieties not expressing symptoms.

2.8.3 Molecular Implications of Satellite Involvement in CMD

Because of the need for parasites to infect hosts for their survival and as parasite infection limits host fitness, pathogenicity in parasites and resistance in hosts have been targets of selection. Plants resist disease through a variety of pre-performed and induced barriers to infection and pathogens use virulence factors to overcome plant defenses and make infection possible.

Selection on both the resistance proteins that recognize the pathogenicity effectors and the pathogenicity factors usually leads to an antagonistic host-pathogen co-evolution. The most relevant model of host-parasite interaction determining the success of infection in plants is the gene-for-gene (GFG) model (Flor, 1971; Keen, 1990). The GFG hypothesis explains host-pathogen recognition events on the assumption that single genes at particular loci in the plant and single genes at particular loci in the pathogen together determine whether the plant-pathogen interaction will be compatible

(resulting in infection and disease development) or incompatible (resulting in resistance to the disease).

In the model, plant resistance proteins (*R* proteins) in the host recognize corresponding avirulence (*Avr*) factors in the pathogen. The recognition triggers defense responses leading to the limitation of spread of the pathogen from the infection site, often associated with hypersensitive response (HR). In absence of the *Avr* allele in the parasite or if the host does not have the *R* allele, the parasite is not recognized by the host, resistance is not triggered and the host is infected.

More than 30 genes have been characterized from different plant species that provide resistance to a variety of different pathogens (Lopez *et al*., 2009). At least 14 *R* gene-like sequences have been identified in cassava. Most of the described classes of plant *R* proteins contain nucleotide binding site (NBS) and leucine-rich repeat (LRR) domains and the later is known to be involved in pathogen recognition.

Ever since capsid protein (CP) of *Tobacco mosaic virus* (TMV) was identified as the first *Avr* factor shown to elicit HR response triggered by *'N'* resistance gene in *Nicotiana spp*, virus-encoded proteins each with role in life cycle of the virus have been shown to act as *Avr* factors. Several (perhaps potentially all) viral genes can encode an avirulence factor that elicits resistance controlled by a cognate dominant host gene: such factors include viral coat protein, RNA polymerase, movement protein and proteinase (Harrison, 2002). In a bipartite *Tomato leaf curl New Delhi virus* (ToLCNDV), nuclear shuttle protein (NSP) gene has been described as *Avr* factor (Malik *et al*., 2011).

Some resistance-breaking virus variants have merely a single nonsynonymous nucleotide replacement in their avirulence gene but, with more durable resistances, virulence necessitates two or even multiple replacements. The probability of a resistance breaking variant appearing and spreading also depends on its biological fitness in the absence of the host resistance gene, and on the type of resistance and number of resistance genes to be overcome. Where satellites are involved in the epidemic, they stand a big chance modifying functional structure of the avirulence genes.

In molecular terms, and as experienced in the case of cassava, the development of resistance against viruses need to take into considerations several features, which raise the need to broaden options for plant virus resistance:

(a) The number of viruses causing problems in plants is large. In cassava for example, seven virus species are known to cause CMD;
(b) Many viruses are capable of infecting a multitude of host plants;
(c) Strains of the viruses vary in their virulence on different crops and even within varieties of the same crop (Scholthof *et al*., 1993);
(d) Classical genetic sources of resistance to viruses are scarce;
(e) Because of the high plasticity of viral genomes, caused by involvement of satellites and climate change leading to mutations and recombinations, the resistances are often not durable in the field (Gruissem, 2010). This necessitates periodic field evaluation.

2.8.4 Resistance Studies in Tanzania

The gene pools for breeding cassava in Africa currently contain only a fraction of the existing genetic variation found in Latin America where the crop originates. Research in Africa has aimed at broadening the genetic base by introducing Latin American germplasm.

According to Okogbenin *et al*. (2007), there have been two sets of introduction:

(i) The introduction of sexual seeds for evaluation in which several seedlings were evaluated in Nigeria between 1990 and 1994,
(ii) The introduction of *in vitro* cultures, where a dominant *CMD2* gene for CMD resistance was introgressed into Latin America germplasm through marker assisted selection (MAS). The MAS genotypes were then pre-selected for the gene and evaluated in Nigeria between 2004 and 2006.

Results from the first set of introduction indicated that the Latin American germplasm was highly susceptible to CMD, minimizing their usefulness in African cassava-breeding programs. Introgression of *CMD2* gene in the second set of introductions

resulted in high CMD resistance under African field conditions. African breeding programs are now concerned with combining this CMD resistance with high yield and in some programs, with other local preferred traits.

A substantial headway has been made in Tanzania in regard to breeding for resistance in cassava. CIAT and IITA have been collaborating with NARS researchers in Tanzania in search for CMD-resistant plants. In the studies, cassava landraces have been used at different levels. Studies by Massumba *et al*. (2005) involving farmer-prefered genotypes collected from the Eastern, Southern and Lake Victoria regions of the country have recommended some varieties in CMD-affected areas: Kiroba, Kigoma-Mafia, Nachinyanya, Kalulu, Kitumbua, Namikonga, Naliendele, TMS30001, I96/1089A and AR40-6. To achieve this, leaves were collected from healthy plants for MAS using available markers for *CMD2* gene.

There has also been progress in the positional cloning of the *CMD2* gene which confers high level of resistance to CMD (Moreno *et al*., 2005). In another development, improved varieties with good resistance to CMD and cassava green mite (CGM) were introduced from CIAT and evaluated for resistance to pests and diseases in Tanzania (Kullaya *et al*., 2005).

Recently, the Cassava Transformation Project based at Mikocheni Agricultural Research Institute in Dar es Salaam is bringing some hope in the fight against the devastating cassava mosaic disease. Some constructs have been designed to be introgressed into preferred Tanzanian cassava landraces to confer them with resistance against CMD and cassava brown streak disease (CBSD). The critical problem though (Ndunguru, pers. Communication) has been the difficulties associated with transforming the landraces.

2. 9 REFERENCES

1. Altieri M.A. (1995) Agroecology: The science of sustainable agriculture. Second Edition. Westview Press Inc., Colorado. Pp. 307-308.
2. Andrade E.C., Manhani G.G., Alfenas P.F., Calegario R.F., Fontes E.P.B. and Zerbini F.M. (2006) *Tomato yellow spot virus*, a tomato-infecting begomovirus from Brazil with a closer relationship to viruses of *Sida* spp. forms pseudorecombinants with begomoviruses from tomato but not from *Sida*. *J. Gen. Virol*. 87: 3687–3696.
3. Berrie L.C., Rybick E.P. and Rey M.E.C. (2001) Complete nucleotide sequence and host range of SACMV: Further evidence for recombination amongst begomoviruses. *J. Gen. Virol.* 82: 53-58.
4. Berrie L.C., Palmer K.E., Rybick E.P., and Rey M.E.C. (1998) Molecular characterization of a distinct South African cassava infecting geminivirus. *Arch.Virol.* 143: 2253 –2260.
5. Bertsch C., Beuve M., Dolja V.V., Wirth M., Pelsy F., Herrbach E. and Lemaire O. (2009) Retention of the virus-derived sequences in the nuclear genome of grapevine as a potential pathway to virus resistance. *Biology Direct* 4: 21.
6. Bock K.R., Guthrie E.J. and Meredith G. (1978) Distribution, host range, properties and purification of *Cassava latent virus*, a geminivirus. *Ann. of Appl. Biol*. 90: 361–367.
7. Bock K.R., Guthrie E.J. and Figueiredo G. (1981) A strain of cassava latent virus occurring in coastal districts of Kenya. *Ann. Appl. Biol.* 99: 151–159.
8. Bock K.R. and Woods R.D. (1983) The etiology of African cassava mosaic disease. *Plant Disease* 67: 994–995.
9. Bos L. (1999) Plant Viruses: unique and intriguing pathogens. A textbook of Plant Virology. Backhuys Publishers Leiden. Pp. 13, 17–19.
10. Briddon R.W., Bull S.E., Amin I., Idris A.M., Mansoor S., Bedford I.D., Dhawan P., Rishi N., Siwatch S.S., Abdel-Salam A.M., Brown J.K., Zafar Y. and Markham P.G. (2003) Diversity of DNA-β, a satellite molecule associated with some monopartite begomoviruses. *Virology* 312: 106-121.
11. Briddon R.W. and Stanley J. (2006) Subviral agents associated with plant single stranded DNA viruses. *J. Virol.* 344 (1): 198-210.

12. Briddon R.W., Brown J.K. Moriones E., Stanley J., Zerbini M., Zhou X. and Fauquet C.M. (2008) Recommendations for the classification and nomenclature of the DNA-β satellites of begomoviruses. *Arch. Virology* 153: 763-781.
13. Bull S.E., Briddon R.W., Sserubombwe W.S., Ngugi K., Markham P.G., Stanley J. (2007) Infectivity, pseudorecombination and mutagenesis of Kenyan cassava mosaic begomoviruses. *J. Gen. Virol*. 88 (5): 1624-33.
14. Carrington J.C., Kasschau K.D., Mahajan S.K. and Schaad M.C. (1996) Cell-to-cell and long distance transport of viruses in plants. *The Plant Cell* 8: 1669–1681.
15. Chellappan P., Vanitharani R., Ogbe F. and Fauquet C.M. (2005) Effect of temperature on geminivirus-induced RNA silencing in plants. *Plant Physiol*. 138: 1828–1841.
16. Chen L., Rojas M., Kon T., Gamby K., Xoconostle-Cazares and Gilbertson R.L. (2009) A severe symptom phenotype in tomato in Mali is associated by a reassortment between a novel recombinant begomovirus (TYLCMLV) and a betasatellite. *Mol. Plant Pathol*. 10 (3): 415–430.
17. Cote F.X., Galzi S., Folliot M., Lamagnere Y., Teycheney P.Y. and Iskra-Carauna M.L. (2010) Micropropagation by tissue culture triggers differential expression of endogenous *Banana streak virus* (eBSV) in the B genome of natural and synthetic interspecific banana plantain. *Mol. Plant Pathol*. 11: 137–144.
18. Cui X., Tao X., Xic Y., Faquet C.M. and Zhou X. (2004) A DNA-β associated with *Tomato Yellow Leaf Curl China Virus* Is required for symptom induction. *Journal of Virology* 78 (24): 13966-13974.
19. Dale J. and Harding R. (2003) Strategies for the generation of virus resistant bananas. *In* Atkinson H., Dale J., Harding R., Kiggundu A., Kunert K., Muchwezi J.M., Sagi L. and Viljoen A. (eds). Genetic transformation strategies to address the major constraints to banana aand plantains production in Africa. INBAP Montpellier (France) p 130
20. Dry I.B., Krake L.R., Rigden J.E. and Rezaian M.A. (1997) A novel subviral agent associated with a geminivirus: The first report of a DNA satellite. Division of Horticulture, Commonwealth Scientific and Industrial Research Organization, Hartley Grove, Urrbrae, South Australia 5064, Australia.

21. FAO (2001) Cassava production statistics. http://www.fao.org.
22. FAO (2003) Cassava production statistics.
23. FAO (2008) FAO Statistics Division. http:/www.fao.org.
24. FAO (2010) FAO Statistics Division.
25. Fargette D., Konate G., Fauquet C., Muller E., Peterschmitt M. and Thresh J.M. (2006) Molecular ecology and emergence of tropical plant viruses. *Ann. Rev. Phytopath*. 44: 235–260.
26. Fauquet C.M. and Fargette D. (1990) *African cassava mosaic virus*: Etiology, epidemiology and control. *Plant Disease* 76: 404-411.
27. Fauquet C.M. and Beachy R.N. (1992) Virus engineered resistance: concepts, efficacy and stability. *In* G. Tottapilly, L.M. Monti, D.R. Mohan Raj and A.W. Moore (Eds). Biotechnology: enhancing production on tropical crops in Africa. Pp. 273–286. Ibadan, Nigeria CTA/IITA.
28. Fauquet C.M. and Stanley J. (2003) Geminivirus classification and nomenclature: progress and problems. *Ann. Appl. Biol.* 142: 165-189.
29. Fauquet C.M., Sawyer S., Idris A.M. and Brown J.K. (2005) Sequence analysis and classification of apparent recombinant begomoviruses infecting tomato in the Nile and Mediterranean basins. *Phytopathology* 95: 549-555.
30. Fauquet C.M., Briddon R., Brown J., Moriones E., Stanley J., Zerbini M. and Zhou X. (2008) Geminivirus strain dermacation and nomenclature. *Arch. Virol.* 153: 783–821.
31. Flor H.H. (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology* 9: 275-296.
32. Fondong V.N., Pita J.S., Rey C., Beachy R.N. and Fauquet C.M. (1998) Disease Notes: First report of the presence of *East African cassava mosaic virus* in Cameroon. *Plant Dis.* 82 (10): 1172.
33. Fondong V.N., Pita J.S., Rey M.E.C., de Kochko A., Beachy R.N. and Fauquet C.M. (2000) Evidence of synergism between *African cassava mosaic virus* and a new double-recombinant geminivirus infecting cassava in Cameroon. *J. Gen. Virol.* 81: 287-297.
34. Fraser R.S.S. (2006) Virus recognition and pathogenicity: Implications for resistance mechanisms and breeding. *Pest Management Science* 22 (3): 267-275.

35. Gayral P., Noa-Carrazana J., Lescot M., Lheureux F., Lockhart B.E.L., Matsumoto T., Piffanelli P. and Iskra-Caruana (2008) A single *Banana streak virus* integration event in the banana genome as the origin of infectious endogenous pararetrovirus (EPRV). *J. Virol.* Doi: 10.1128/JVI.00212-08.
36. Gruissem W. (2010) The Battle Against Plant Viruses. Sixth World Conference on the Future of Science "Virus: the invisible enemy" – Venice, September 19-21, 2010.
37. Gupta R. (2006) A textbook of systematic botany. Cbs publishers and Distributors.
38. Guiterrez, C. (2000) DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J.* 19 (5): 792–799.
39. Hahn S.K., Terry E.R. and Leuschner K. (1980) Breeding cassava for resistance to cassava mosaic disease. *Euphytica* 29 (3).
40. Hao L., Wang H., Sunter G. and Bisaro M. (2003) Geminivirus AL2 and L2 proteins interact with and inactivate SNF1 kinase. *The Plant Cell* 15: 1034-1048.
41. Harrison B.D. (2002) Virus variation in relation to resistance-breaking in plants. *Euphytica* 124: 181–192.
42. Hillocks R.J. (2002) Cassava in Africa. *In* Hillocks R.J., Thresh J.M. and Bellotti A.C. (eds). Cassava: Biology, Production and Utilization. CABI Publishing Oxon UK.
43. Hohn T., Richert-Pogeler K.R., Harper G., Schawarzacher T, Teo C., Teycheney P.Y. Iskra-Caruana M.L. and Hull R. (2008) Evolution of integrated plant viruses. *In* Plant virus evolution. M. Roosinck (ed.), Academic Springers, Heidelberg, Germany. Pp 54-76.
44. Hong Y.G., Robinson D.J. and Harrison B.D. (1993) Nucleotide sequence evidence for the occurrence of three distinct whitefly transmitted geminiviruses in cassava. *J. Gen. Virol.* 74: 2437-2443.
45. Hull R. (2002) Matthew's Plant Virology Fourth Edition. Academic Press London. Pp. 171, 180.
46. Huttner E., Tucker W., Vermeulen A., Ignart F., Sawyer B. and Birch R. (2001) Ribozyme genes protecting transgenic melon plants against potyviruses. *Curr. Issues Mol. Biol.* 3 (2): 27-34.

47. Iskra-Caruana M.L., Baruens F.C., Gayral P. and Chabannes M. (2010) A four-partner plant-virus interaction: Enemies can also come from within. *Molecular Plant-microbe Interactions* 23 (11): 1394–1402.
48. Jakowitsch J., Mette M.F., van der Winden J., Matzke M.A. and Matzke A.J.M. (1999) Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. *PNAS* 96 (23): 13241–13246.
49. Jebasingh T., Backiyarani S., Manohaari C. and Usha R. (2011) Detection of Cardamom mosaic-related sequences in plant geneomes. *Indian Journal of Biotechnology* 10: 369–371.
50. Jeske H. (2009) Geminiviruses. *Curr. Top. Microbiol. Immunol.* 331: 185-226.
51. Jeske H., Lutgemeier M. and Preiss W. (2001) DNA forms indicate rolling circle and recombination-dependent replication of *Abutilon mosaic virus*. *EMBO J.* 20: 6158–6167.
52. Jose J. and Usha R. (2003) Bhendi Yellow Vein Mosaic Disease in India is caused by association of a DNA-β satellite with a begomovirus. *Virology* 305: 310-317.
53. Keen N.T. (1990) Gene-for-gene complementarity in plant-pathogen interactions. *Annual Review of Genetics* 24: 447-463.
54. Khan J.A. and Dijkstra J. (2001) Plant viruses as molecular pathogens. Haworth Press Inc. Pp. 267–280.
55. Kon T., Rojas M.R., Abdourhamane I.K. and Gilbertson R.L. (2009) Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *J. Gen. Virol.* 90: 1001–1013.
56. Kullaya A., Mtunda K., Massumba E., Kulembeka H., Ferguson M., Ospina C., Hurtado P., Guterrez J., Barrera E., Marin J., Castro A., Morrilo A., Alzate A., Morante N., Ceballos H., Tohme J. and Fregene M. (2005) Molecular Marker-Assisted selection (MAS) for improvement of local cassava germplasm in Tanzania for pest and disease resistance.
57. Lazaro E.A., Lisasi H. and Assenga A. (2007) Cassava in Tanzania: A famine or commercial reserve? *In* Kapinga R. Kingamkono R., Msabaha M., Ndunguru J., Lemaga B. and Tusiime G. (eds). Proceedings of the 13th ISTRC Symposium 10–14 November 2003 Arusha Tanzania Pp. 614–621.

58. Legg J.P. and Thresh J.M. (2000) Cassava mosaic disease in East Africa: a dynamic disease in a changing environment. *Virus Res.* 71: 135-149.

59. Legg J.P. and Fauquet C.M. (2004) Cassava mosaic geminiviruses in Africa. *Plant Mol. Biol*. 56: 585–599.

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

61. Lent J., Storms M., Meer F., Wellink J. and Goldbach R. (1991) Tubular structures involved in movement of *Cowpea mosaic virus* are also formed in infected cowpea protoplasts. *J. Gen. Virol.* 72: 2615–2623.

62. Lockhart B.E.L. and Jones D. (2000) Banana streak. Pp. 263–274. *In* Diseases of Banana, Abaca and Enset. D. Jones, ed. CAB International, Wallingford, UK.

63. Malik A.H., Briddon R.W. and Mansoor S. (2011) Infectious clones of *Tomato leaf curl Palampur virus* with a defective DNA B and their pseudorecombination with *Tomato leaf curl New Delhi virus*. *Virology Journal* 8:173 – 180.

64. Mallowa S.O., Isutsa D.K., Kamau A.W., Obonyo R. and Legg J.P. (2006) Current characteristics of cassava mosaic disease in postepidemic areas increase the range of possible management options. *Ann. Appl. Biol*. 149 (2): 137-144.

65. Martin E.F. (1928) Report of the mycologist. Annual Report of the Department of Agriculture, Uganda, p31. Entebbe, Ugandan Government.

66. Martin D.P., Lefeuvre P., Varsani A., Hoareau M., Semegni J.Y. (2011) Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. *PLoS Pathog* 7(9): e1002203. doi:10.1371.

67. Massumba E., Kulembeka H., Mtunda K., Kullaya A., Fregene M., Kanju E. and Ferguson M. (2005) Asessment of disease severity of local cassava collections in Tanzania. *In* van Houten H. (ed). Biotechnology, Breeding and seed systems for African crops: Research and development that reaches farmers. Proceedings of the 2nd General Meeting of the Rockfeller Foundation-supported program. Nairobi Kenya 24-27 January 2005. Pg. 118.

68. Meester G., Woittiez R.D. and de Zeeuw A. (1999) Plants and Politics – On the occasion of the Centenary of the Netherlands Plant Protection service. Wageningen Pers., Wageningen. Pp. 149-158, 180–181.

69. Mittal D., Borah B.K. and Dasgupta I. (2008) Agroinfection of cloned *Sri Lankan cassava mosaic virus* DNA to *Arabidopsis thaliana*, *Nicotiana tabacum* and cassava. *Arch. Virol*. 153: 2149–2155.

70. Moffat A.S. (1999) Geminiviruses emerge as serious crop threat. *Science* 286: 1835.

71. Moreno H., Tomkins J., Ariyo O.A., Okogbenin O. and Fregene M. (2005) Progress in positional cloning of *CMD2*: the gene that conferes a high level of resistance to cassava mosaic disease. *In* van Houten H. (ed.). Biotechnology, Breeding and seed systems for African crops: Research and development that reaches farmers. Proceedings of the 2nd General Meeting of the Rockfeller Foundation-supported program. Nairobi Kenya 24-27 January 2005. Pg. 126.

72. Moriones E and Navas-Castillo J (2008) Review. Rapid evolution of the population of begomoviruses associated with the tomato yellow leaf curl disease after invasion of a new ecological niche. Spanish Journal of Agricultural Research 6 (Special issue): 147-159.

73. Mtunda K.J., Muhanna M., Raya M.D., Kanju E.E. (2002) Current status of cassava brown streak virus disease in Tanzania. *In*: Legg J.P. and Hillock R.J. eds). Cassava brown streak virus disease – Past Present and Future. Proceedings of an international Workshop, Mombasa Kenya 27-30 October 2002.

74. Navas-Castillo J., Sanchez-Campos S., Noris E., Louro D., Accotto G.P. and Moriones E. (2000) Natural recombination between *Tomato yellow leaf curl-Is* and *Tomato leaf curl virus*. *J. Gen. Virol.* 81: 2797-2801.

75. Nawaz-ul Rehman M.S. and Fauquet C.M. (2009) Minireview: evolution of geminiviruses and their satellites. *FEBS Letters* 583: 1825–1832.

76. Ndomba, O.A. (2005) Towards developing a systems approach for pest risk management in Tanzania: Challenges and Opportunities. *Trop. Pest. Mgt. Bull.* 3(6): 92–96.

77. Ndowora T., Dahal G., LaFleur G., Hull R., Olezewski N. and Lochhart B. (1999) Evidence that badnavirus infection in Musa can originate from integrated pararetroviral sequences. Virology 225: 214-220.

78. Ndunguru, J., Legg, J.P. Aveling, T.A.S., Thompson, G. Fauquet, C.M. (2005) Molecular biodiversity of cassava begomoviruses in Tanzania: evolution of cassava geminiviruses in Africa and evidence for East Africa being a center of diversity of cassava geminiviruses. *J. Virol.* 2: 21.

79. Ndunguru J., Legg J.P., Ismael B.F., Aveling T.A.S., Thompson G. and Fauquet C.M. (2005). Identification of a defective molecule derived from DNA-A of the bipartite begomovirus of *East African cassava mosaic virus*. *Plant Pathol*. 55: 2-10.

80. Ndunguru J., Fofana B., Legg J.P., Chellappan P., Taylor N., Aveling T., Thompson G., Fauquet C. (2008) Two novel satellite DNA molecules associated with bipartite cassava mosaic begomoviruses enhancing symptoms and capable of breaking high virus resistance in a cassava landrace. *In* Cassava: Meeting the Challenges of the New Millenium. First Scientific Meeting of the Global Cassava Partnership GCP-I, Ghent Belgium 21-25 July 2008. Pg. 141.

81. Ogbe F.O., Legg J., Raya M.D., Muimba-Kankalongo A., Theu M.P., Kaitisha G., Phiri N.A. and Chalwe A. (1997) Diagnostic survey of cassava mosaic viruses in Tanzania, Malawi and Zambia. *Roots 4: 12*-15.

82. Okogbenin E., Porto M.C.M., Egesi C., Mba C., Espinosa E., Santos L.G., Ospina C., Marin J., Barrera E., Gutierrez J., Ekanayake I., Iglesias C. and Fregene M.A. (2007) Marker assisted introgression of resistance to cassava mosaic disease into Latin American germplasm for the genetic improvement of cassava in Africa. *Crop Sci.* 47: 1895–1904.

83. Padadim M., Beachy R.N. and Fauquet C.M. (1995) Classification and identification of geminiviruses using sequence comparisons. *J. Gen. Virol*. 76: 249–263.

84. Patil B.L. and Fauquet C.M. (2010) Differential interaction between cassava mosaic geminiviruses and geminivirus satellites. *J. Gen. Virol.* 91: 1871-1882.

85. Patil B.L. and Fauquet C.M. (2009) Cassava mosaic geminiviruses: Actual knowledge and perspectives. *Molecular Plant Pathology* 10 (5): 685-701.

86. Pita J.S., Fondong V.N., Sangaré A., Otim-Nape G.W., Ogwal S., Fauquet C.M. (2001) Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J. Gen Virol.* 82 (3): 655-665.

87. Porta C. and Lomonossoff G.P. (2002) Viruses as vectors for the expression of foreign sequences in plants. Biotechnology and Genetic Engineering Reviews 19: 245–291.

88. Prins M., Laimer M., Noris E., Schubert J., Wessenegger M. and Tepfer M. (2008) Strategies for antiviral resistance in transgenic plants. A review. *Mol. Plant Pathol.* 9 (1): 73–83.

89. Sacristan S. and Garcia-Arenal F. (2008) Challenges for molecular plant pathology over the next ten years: the evolution of virulence and pathogenicity in plant pathogen populations. *Mol. Plant Pathol.* 9 (3): 369–384.

90. Saeed M., Zafar Y., Randles J.W. and Rezaian M.A. (2007) A monopartite begomovirus-associated DNA-ß satellite substitutes for the DNA-B of a bipartite begomovirus to permit systemic infection. *J. Gen. Virol* 88: 2881–2889.

91. Sansford J.C. and Johnston S.A. (1985) The parasite-derived resistance–deriving resistance genes from the parasite's own genome. *J. Theor. Biol.* 113: 395–405.

92. Saunders K., Norman A., Gucciardo S. and Stanley J. (2004) The DNA-β satellite component associated with Ageratum yellow vein disease encodes an essential pathogenicity protein (βC1). *Virology* 324: 37-47.

93. Scholthof K.C., Scholthof H.B. and Jackson A.O. (1993) Control of plant virus disease by pathogen-derived resistance in transgenic plants. *Plant Physiol*. 102: 7-12.

94. Scott G.J. Rosegrant M.W. and Ringler C. (2000) Roots and tubers for the 21st Century: Trends, projections and policy options Pp. 1-71. http://www.ifpri.org/2020/dp/2020dp31.

95. Seal S.E., vandenBosch F. and Jeger M.J. (2006) Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. *Crit. Rev. Plant Sci*. 25 (1): 23–46.

96. Shewry P.R. (2003) Tuber storage proteins. *Ann. Botany* 91: 755–769.

97. Sserubombwe W.S., Briddon R.W., Baguma Y.K., Ssemakula G.N., Bull S.E., Bua A., Alicai T., Omongo C., Otim-Nape G.W. and Stanley J. (2008) Diversity of begomoviruses associated with mosaic disease of cultivated cassava

(*Manihot esculenta* Krantz) and its wild relative (*Manihot glaziovii* Mull. Arg.) in Uganda. *J. Gen. Virol.* 89: 1759–1769.

98. Sseruwagi P., Sserubombwe W.S., Legg, J.P., Ndunguru J. and Thresh J.M. (2004) Methods of surveying the incidences and severity of cassava mosaic disease in Africa: A Review: *Virus Res.* 100: 129–142.
99. Thresh J.M., Otim-Nape G.W., Legg J.P. and Fargette D. (1997) African cassava mosaic virus disease: Magnitude of the problem. *Afr. J. Root and Tuber Crops.* 2: 13-19.
100. Thresh J.M., Otim-Nape G.W. and Fargette D. (1998) The control of African cassava mosaic disease: Phytosanitation and/or Resistance?. *In* Hadid A., Khetarpal R.K., Koganezawa H. (eds). Plant Virus Disease Control. American Phytopathological Society Press Pp. 670–677.
101. Tonukari N.J. (2004) Cassava and future of starch. *Electron. J. Biotechnol*. Vol. 7 No 1 2004.
102. Urbino C., Thebaud G., Granier M., Blanch S. and Peterschmitt M. (2008) A novel cloning strategy for isolating genotyping and phenotypig genetic variants of geminiviruses. *J. Virol*. 5: 135.
103. Vanitharani R., Chellappan P. and Fauquet C.M. (2005) Geminiviruses and RN silencing. *Trends in Plant Science* 10 (3): 144-151.
104. Warburg O. (1894) Die kulturpflanzen Usambaras. Mitteilungen aus den Deutschen Schutzgebieten. *Ann. Appl. Biol*. 25: 780–786.
105. Westby A., Lankford B.A., Coulter J.F., Ochard J.E. and Morton J.F. (2004) Rural infrastructure to contribute to African development: The cases of irrigation and post harvest. Natural resource Institute and Overseas Development Group.UK.
106. Xie Y., Wu P., Liu P., Gong H., Zhou X. (2010). Characterization of alphasatellites associated with monopartite begomovirus/betasatellite complexes in Yunnan, China. *Virology Journal* 7: 178.
107. Zhou X., Liu Y., Calvert L., Munoz D., Otim-Nape G.W., Robinson D.J. and Harrison B.D. (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *J. Gen. Virol.* 78: 2101-2111.
108. Zimmermann A. (1906) Die Krauselkrankheit des Maniok. *Pflanzer*, 2, 145.

CHAPTER 3

MOLECULAR IDENTIFICATION AND DISTRIBUTION OF CASSAVA GEMINIVIRUSES AND SATELLITE DNA MOLECULES IN MAJOR CASSAVA–PRODUCING AREAS OF TANZANIA

ABSTRACT

A field survey was conducted in three cassava-growing areas in Tanzania in the 2006/2007 growing season to investigate occurrence of cassava mosaic begomoviruses and SatDNA-II and SatDNA-III. Incidence and severity scores were recorded for all samples in all geographical regions. Cassava stems were collected from diseased field plants, and were re-planted in screenhouse to reproduce and verify disease symptoms, to study temporal variation in symptom development, and to obtain fresh leaf samples for laboratory analyses. Total nucleic acids were extracted from the leaves and subjected to PCR amplification of near full-length DNA A, in order to confirm the presence of begomoviruses. Sequencing of PCR products was subsequently undertaken. Full-length SatDNA-II and SatDNA-III sequences were PCR-amplified and the PCR products sequenced. Integrated satellite-like sequences were amplified using primers designed by German Plata (CIAT, Colombia) to amplify a 895 bp and 306 bp fragment from SatDNA II and SatDNA III, respectively.

Results revealed that CMD incidences were higher in Lake Victoria zone (with fields having 70 - 90% incidence) compared with the Eastern zone, where some fields had as low as 10% incidence. Severity scores of the field samples were categorised as healthy, mild or severe and samples of the different categories exhibited different symptom trends when re-planted in screen house compared with the original field samples. The CMD incidence patterns and those of the satellites were not correlated with temperature or altitude of the locations.

Screen house observations revealed changes in symptom trends from original symptoms observed in the field, including reversion and recovery from symptoms. PCR amplifications showed presence of EACMV (100%), ACMV (5%), SatDNA-II (31.6%) and SatDNA-III (33.0%). No co-infection of the satellites was recorded in individual leaf samples. Full-length SatDNA-II was amplified primarily from samples collected from Eastern zone, while SatDNA-III was prevalent in samples collected from the Lake zone. Nucelotide sequencing results of the EACMV PCR products indicated distinct isolates, EACMV-TZ [TZ113] and EACMV-TZ [TZ108], which, in the phylogenetic tree, were shown to be closely related to EACMZV-[K18].

The survey and amplification of integrated sequences revealed a high occurrence of the sequences (68% for SatDNA-II and 71.1% for SatDNA-III) in plant leaf samples tested. The integrated sequences were variable in size for SatDNA-II, but uniform for SatDNA-III. The occurrence of other amplified DNA bands, notably the closely spaced 'double bands' in SatDNA-II and SatDNA-III, suggest multiple locational insertions of different sizes of these sequences in cassava germplasm. Sequences of the integrated SatDNA-II isolates resembled *Begomovirus-associated DNA-II satellite* (AY836366) by 96 – 98% nucleotide identity and *Mentha leaf deformity-associated satellite DNA-II* (EU862815) by 85 – 92%. Sequences of integrated SatDNA-III isolates resembled *Begomovirus-associated DNA-III satellite* (AY836367) by 76 – 94% nucleotide identity. Full-length SatDNAs were found to occur more widely than reported earlier.

The present study still seems exploratory and it is evident that more research is required to have a comprehensive picture of the occurrence and influence of SatDNA-II and SatDNA-III on severity of CMD.

3.1 INTRODUCTION

Cassava is important to the livelihoods of millions of poor people in the World (Westby *et al*., 2004). It plays a food security role in war prone areas. In Tanzania, 51.6% of farmers in cassava growing areas ranked cassava as the most important crop followed by maize (Hillocks *et al*., 2002). The major growing areas in the country include the coastal strip, the Lake Victoria, and the Lake Nyasa zones constituting 48.8%, 23.7% and 13.7% of total cassava produced in the country, respectively (Mtunda *et al*., 2002).

Cassava mosaic disease (CMD) is considered to be the main biotic constraint to cassava production and the most important threat to food security in Tanzania. The disease is spread by whiteflies, *Bemisia tabaci* Genn. (*Homoptera*: *Aleyrodidae*), in a non-persistent manner. CMD symptoms include green and yellow mosaic, distortion of leaves and plant stunting (Fondong *et al*., 1998). Severe chlorosis is also associated with premature leaf abscission and decrease in growth and yield. CMD epidemics have been controlled by using resistant cassava varieties and sanitation.

Seven distinct geminivirus species, with variants, belonging to the genus *Begomovirus*, family *Geminiviridae*, are responsible for CMD in Africa: *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *East African cassava mosaic Cameroon virus* (EACMCV), *East African cassava mosaic Malawi virus* (EACMMV), *East African cassava mosaic Zanzibar virus* (EACMZV), *East African cassava mosaic Kenya virus* (EACMKV) and *South African cassava mosaic virus* SACMV (Fauquet *et al*., 2003; Fauquet *et al*., 2008).

Earlier reports of severe CMD epidemics in Tanzania were associated with either mixed ACMV/EACMV infections or the occurrence of EACMV Uganda variant (EACMV-UG2) (Patil and Fauquet, 2009). Present reports of CMD severity in the Lake Victoria zone and along the Eastern coast of Tanzania showed concurrent occurrence of EACMV, ACMV and satellite-like DNA molecules (Ndunguru *et al*., 2008). This was the first report of the association of satellites with cassava mosaic geminiviruses. Two satellite-like DNAs, satellite DNA-II (SatDNA-II) and satellite DNA-III (SatDNA-III) were isolated from a few diseased plants collected from the

Eastern zone and Lake Victoria zones, respectively. The size of SatDNA-II and SatDNA-III were 1032 bp and 1209 bp, respectively. The satellites appear to increase symptoms in resistant cassava cultivars to CMD (Ndunguru, 2005).

The extent to which these satellites have spread in the cassava growing zones is not known. Whether these are the only satellites associated with the cassava mosaic begomoviruses also remains obscure. Furthermore, the pattern of disease development in different temperature regimes is also not known with certainty. This study reports a survey carried out in 2006/2007 around the Lake Victoria and Eastern cassava growing zones (Tanga, Dar es Salaam, Coast, Lindi, Mtwara and other southern regions) of Tanzania to explore the epidemiology of CMD in the presence of satellite DNA molecules.

3.2 MATERIALS AND METHODS

3.2.1 Field Survey and Sample Collection

A survey was conducted in three major cassava growing areas of Tanzania – the Lake Victoria zone (1°-3°S 31°-34°E), the Eastern Coast zone ($4^{o}30'$-$11^{o}30'$S $37^{o}30'$-$40^{o}30'$E) and the Southern zone (9°–11°S 34°–37°E) in 2006/2007 growing season to investigate the occurrence of cassava mosaic geminiviruses (CMGs) and associated satellite DNA molecules SatDNA-II and SatDNA-III. Cassava plants showing typical symptoms of CMD, and plants showing unique symptoms of extreme leaf narrowing (filiform) associated with the presence of SatDNAs were selected. Stems were collected from the symptomatic plants by a zig-zag transect in the field, and observing and recording disease severity of every ninth plant on a scoring scale 1 – 5 (Hahn *et al*., 1980) described in Sseruwagi *et al*., 2003, as depicted in Table 3.1.

Table 3.1: Cassava mosaic disease symptom scale of 1 – 5

Scale	Symptom Label	Symptom Description
1	Healthy	Unaffected shoots, no symptoms
2	Mild	Mild chlorosis, mild distortions at bases of most leaves, while the remaining parts of the leaves and leaflets appear green and healthy
3	Moderate	Pronounced mosaic pattern on most leaves, narrowing and distortion of the lower one-third of the leaflets
4	Severe	Severe mosaic, distortion of two thirds of most leaves and general reduction of leaf size and stunting of shoots
5	Very severe	Very severe mosaic symptoms on all leaves, distortion, twisting, misshapened, and severe leaf reductions of most leaves accompanied by severe stunting of plants.

Source: Hahn *et al*., 1980.

The number of stems collected from each field was depended upon CMD incidence and severity varied from 2 to 4 per field and 2 to 3 fields per district, depending on level of severity of the disease.

The pattern of diseased plants in the fields, i.e. disease incidence as defined by Delp *et al*. (1986) were also characterized by considering the proportion of plants infected by disease relative to the wholesome plants in the field (Madden and Hughes, 1995). Disease incidence was assessed as the proportion or percentage of plants with CMD symptoms on a scale of 0-1 (P) or 0-100 (%) (Fargette, 1985; Sseruwagi *et al*., 2003; Adriko *et al*., 2011).

In order to determine if altitude and latitude had an effect on disease incidence, the survey was conducted in several districts across several altitudes and latitudes. In the Lake Victoria zone, the altitude varied from 1660m above sea level (*asl*) at Nyakashushu (1°36.948'S 31°08.243'E) in Karagwe through 1499m *asl* at Butiama (1°45.385'S 33°58.169'E) in Musoma to 1139m *asl* at Buhima (2°36.163'S 32°53.003'E) in Mwanza. In the Eastern zone, there were 35 sites varying in altitude

from 1221m *asl* at Kongei in Lushoto through 74m *asl* at Chai Bora in Dar es Salaam to 0m *asl* at Hamjambo in Kilwa Masoko. For the Eastern zone, the area covered was from Kwasunga (38°20.144'E) in Bagamoyo to Mnazi Bay (40°28.500'E) in Mtwara. In the field, the severity of CMD (on the scale of 1 – 5) was taken into account by scoring plant samples showing different symptom phenotype levels. A disease incidence percentage in each field where a sample was taken was shown.

3.2.2 Sprouting Cassava Cuttings in Screen House

A total of 152 cassava stem cuttings were re-planted in a screen house at Mikocheni Agricultural Research Institute (06°45.408'S 39°14.043'E, 14m) in Dar es Salaam to reproduce and verify field disease symptoms. They were also re-grown for the purpose of studying variations in symptom development over time and for obtaining young leaf samples for laboratory analyses. Each cutting was cut into two, ≈20-cm length pieces and both pieces were planted in an improvised FG Pail 4-Liter pot (Simba Plastics Company Ltd., Dar es Salaam) containing thermo-sterilized soil (Figure 3.1). The screen house was occasionally sprayed with Decis 2.5% EC (Deltamethrin 25g/l; Bayer East Africa Ltd., Nairobi) to control whiteflies.



Figure 3.1: Stem cuttings planted in screen house.

As seen in Figure 3.1, each cutting was cut into two pieces and both were planted in the same pot. The aim was to see if there were any differences in disease symptom phenotype within the same plant. The planted cuttings were monitored for symptom development for up to 35 days after planting (DAP). At times, the sprouted cuttings

were cut back to allow re-shooting in order to verify certain temporal symptom development trends. The screen house symptoms for each sprouted cutting were then compared with symptoms observed in the field for the same plant.

3.2.3 Total Nucleic DNA Extraction

The top second fully expanded symptomatic leaves from the cuttings with clear and characteristic symptoms were collected for DNA extraction. Crude total nucleic DNA (TNA) was extracted from the leaves by the method described by Dellaporta *et al*., (1983). Each leaf sample (0.1 - 1g) was ground by mortar and pestle in 700 µL of Dellaporta grinding buffer. The buffer contained 100 mM Trisma base (Tris hydroxymethyl methylamine, $C_4H_{11}NO_3$) in HCl pH 8.0, 500 mM NaCl and 50 mM EDTA (tetra-acetate disodium salt dihydrate, $C_{10}H_{14}N_2O_8Na_2.2H_2O$), in distilled and autoclaved water. On the day of use, 98% β-mercaptoethanol (C_2H_6OS) was added to the buffer at rate of 2.3 µL/mL.

The ground tissue was transferred into a 1.5 mL microfuge tube (Eppendorf AG, Hamburg, Germany). Forty two microlitres (42 µL) of cold 20% Sodium Dodecyl Sulphate (SDS, $C_{12}H_{25}NaO_4S$) was added, the solution was mixed by inverting the tubes and incubated in a water bath at 65^oC for 10 min to lyse and disperse tissues. The actual isolation was accomplished by the addition of 160 µL of cold 0.5M Potassium acetate (CH_3COOK) to each tube, and the solution mixed by inverting and placed on ice for 10min to precipitate cellular proteins and polysaccharides. The tubes were then spun in a Biofuge 13 microcentrifuge (Heraeus, Germany) at 13000 *rpm* for 10min to separate the aqueous phase (containing nucleic acids) from the organic phase. Six hundred and fifty microlitres of the supernatant in each tube was transferred to a new tube.

An equal volume (650 µL) of ice-cold propan-2-ol (isopropanol) was added to the microtube to precipitate the DNA. The tubes were spun at 13000 *rpm* for 10 min to pellet the DNA and the supernatant was decanted. The DNA pellets were purified by first washing with 500µL of cold 70% ethanol by inverting the microtubes followed by spinning at 13000 *rpm* for 5 min. The ethanol supernatant was removed and the DNA pellets air dried by inverting the microtubes over blotting paper. The DNA

pellets were re-suspended in 250 µL of sterile distilled water and then 200 µL of the resulting DNA supernatant transferred to new tubes and stored at –20°C.

3.2.4 PCR Amplification of CMGs

The DNA-A components of CMGs (*ca*. 2760 - 2780 bp) from the sample DNAs were amplified by PCR using primer pair UniF (5'-KSG GGT CGA CGT CAT CAA TGA CGT TRT AC-3') and UniR (5'-AAR GAA TTC ATK GGG GCC CAR ARR GAC TGG C-3') (Integrated DNA Technologies Inc., USA) (Briddon and Markham, 1994; Pita *et al*., 2001) [K = G, T; R = A, G; S = C, G]. A 20 µL PCR reaction mix was made (8.16 µL sterile water, 2.0 µL 10x PCR buffer, 2.0 µL 5% Tween 20, 1.6 µL 2.5 mM dNTPs,1.2 µL 25 mM $MgCl_2$, 10 mM Primers and 0.04 µL 5U/µL *Taq*). Sixteen microlitres (16 µL) of the mixture was added to labelled PCR tubes (Abgene House, Epsom UK) and 4.0 µL of respective DNA samples were added to the tubes. The tubes were loaded into GeneAmp PCR System 9700 thermocycler (Applied Biosystem Inc., Singapore). The PCR program was: Initial denaturation 94°C 2 min, denaturation 94°C 1 min, primer annealing 59°C 1 min, extension 72°C 2min, 30 cycles).

3.2.5 PCR Amplification of EACMV-TZ

Detection of viral DNAs was done by PCR using primer pair EAB555F (5'-TAC ATC GGC CTT TGA GTC GCA TGG-3') and (5'-CTT ATT AAC GCC TAT ATA AAC ACC-3') (Integrated DNA Technologies Inc., USA) (Pita *et al*., 2001) designed to amplify the *BC1*/CR (550 – 560 bp) region on the DNA-B component. The lyophilized primers were diluted to 100 mM to make stock solution according to manufacturer's instructions. A 20 µL-PCR total reaction volume was made as described above, replacing the primer pair as appropriate. Sixteen microlitres (16.0 µL) of the mixture was added into labeled 0.2 µL Thermo-Tubes (Abgene House, Epsom UK) and 2.0 µL of respective DNA samples were added to the tubes as described above. The tubes were loaded into a thermocycler as stated in section 3.2.4 (above) and the contents amplified according to the amplification profile: Initial

denaturation 94°C 2 min, denaturation 94°C for 1 min, primer annealing 55°C 1 min, extension 72°C 1 min and final extension 72°C 10 min., 35 cycles.

3.2.6 PCR Amplification of ACMV

Near full-length ACMV from the leaf sample TNAs were amplified by PCR using primers JSP001 (5'-ATG TCG AAG CGA CCA GGA GAT-3') and JSP002 (5'-TGT TTA TTA ATT GC CAA TAC T-3') (Integrated DNA Technologies Inc., USA). A 20 µL PCR total reaction mixture (2.0 µL of 10x PCR buffer, 2.0 µL 5% Tween 20, 1.6 µL of 2.5 mM dNTPs, 1.2 µL 25 Mm $MgCl_2$, 0.5 µL JSP001, 0.5 µL JSP002, 0.04 µL *Taq* DNA polymerase and 8.16 µL cold sterile distilled water) was made and mixed well. The reaction mixture was then calculated to accommodate the number of samples at hand. Sixteen microlitres (16.0 µL) of the mixture was added into labeled PCR tubes and 4.0 µL of respective DNA samples were added to the tubes. The amplification profile was: Initial denaturation 94°C 2 min, denaturation 94°C for 1 min, primer annealing 46°C 1 min, extension 72°C 1 min and final extension 72°C 10 min, 30 cycles. An 1% agarose gel was prepared by boiling 1g agarose in 100 mL 1x TAE buffer in Sanyo Contempo Cook Microwave (Sanyo Microwave Co. Ltd, Uk) for 3 min. 1.3 µL of ethidium bromide (EtBr) was added to chelate the gel. The PCR products were separated on 1.0% agarose gel in 1x TAE against 1 kb ladder as described above and bands visualized using imaging software.

3.2.7 PCR Amplification of Full-length SatDNA-II

Full-length SatDNAs were amplified using primer pair Beta01 (5'-GGT ACC ACT ACG CAG CAG CC-3') and Beta02 (5'-GGT ACC TAC CCT CCC AGG GGT ACA C-3') (Integrated DNA Technologies Inc., USA) (Ndunguru, 2004). A 20 µL PCR reaction mixture was prepared per sample as described above. Sixteen microlitres (16.0 µL) of the mixture was added into labeled PCR tubes and 4.0 µL of respective DNA samples were added to the tubes. The contents were thermocycled according to the amplification profile: Initial denaturation 94°C 1min, denaturation 94°C for 1 min, primer annealing 49°C 1 min, extension 72°C 2 min and final extension 72°C 10 min, 35 cycles.

3.2.8 PCR Amplification of Full-length SatDNA-III

Full-length SatDNA-III from the sample DNAs were amplified by PCR using primers DNA-1F (5'-TGG GGA TCC TAG GAT ATA AAT AAC ACG TC-3') and DNA-1R (5'-CTA GGA TCC GGA CAA ATT ACA AGC GTA-3') (Integrated DNA Technologies Inc., USA) (Ndunguru, 2004). A 20 µL PCR total reaction mixture (2.0 µL of 10x PCR buffer, 2.0 µL of 5% Tween 20, 1.6 µL of 2.5 mM dNTPs, 1.2 µL of 25 mM $MgCl_2$, 0.5 µL DNA-1F, 0.5 µL DNA-1R, 0.04 µL *Taq* DNA polymerase and 8.16 µL cold sterile distilled water) was made and mixed well. The reaction mixture was then prepared to accommodate the number of samples available. Sixteen microliter (16.0 µL) of this mixture was added to labeled PCR tubes and 4.0 µL of respective DNA samples were added to the tubes. The contents amplified according to the amplification profile: Initial denaturation 94°C 1 min, denaturation 94°C for 1 min, primer annealing 49°C 1 min, extension 72°C 2 min and final extension 72°C 10 min, 35 cycles.

All PCR products were run on a 1% agarose gel. A 1% agarose gel was prepared by boiling 1g agarose in 100 mL 1x TAE buffer in a microwave (Sanyo Microwave Co. Ltd, Uk) for 3 min, and adding 1.3 µL of ethidium bromide (EtBr) to the gel mixture. The gel was cooled, and amplified DNA electrophoresed in an A6 Electrophoresis tube (Owl Separation Systems Inc., USA) at 120V, using 0.6 µL of 1 kb DNA ladder. Bands on the gel were visualized using GeneSnap Image acquisition Software (InGenius Bioimaging Technologies Inc., USA) against 1.0 kb DNA marker (Sigma-Aldrich, UK).

3.2.9 Cloning and Sequence Verification

Two EACMV DNA B fragments (550-560bp) from Muheza and Handeni districts were amplified by PCR using primers: EAB555F (5'-TAC ATC GGC CTT TGA GTC GCA TGG-3') and EAB555R (5'-CTT ATT AAC GCC TAT ATA AAC ACC-3'). Conditions of the amplification were pre-denaturation 94°C 2min; denaturation 94°C 2 min; primer annealing 55°C 1 min; extension 72°C 1 min; final extension 72°C

10 min. 30 cycles. The PCR product was purified and cloned into a 2817 bp-plasmid cloning vector (Invitrogen Life Technologies Ltd., Carlsbad) (Appendix 3.1). The plasmid has an antibiotic resistance sequence for selection with spectinomycin and attachment sites *att*L1 and *att*L2 for the integration of gene cassettes.

The amplicon (above) was ligated into the vector using a ligation mix prepared according to manufacturer's instruction (4 μL amplicon, 1 μL cloning vector and 1 μL salt solution) and incubated at room temperature. Five microlitres of the ligation mix was added to 50 μL of DHα chemically competent *Escherichia coli* cells and the mixture incubated on ice for 30 min for transformation. The contents were heat-shocked in water bath at 42°C for 20 seconds and tubes incubated on ice for 2 min. To the mixture was then added 250 μL S.O.C media (Appendix 3.5) and incubated at 37°C with shaking at 225 rpm for 1 hour. The tubes were centrifuged at 10000 rpm for 5 minutes, and half of the S.O.C medium discarded and vortexed.

The bacterial cells were plated in 100 μL of LB agar (containing NaCl, Tryptone, Yeast extract and agar) in petri dishes and incubated at 37°C overnight without shaking. Bacterial colonies surviving the antibiotic (cream white) selection were cultured in LB broth media (containing NaCl, Yeast extract, and Tryptone) and the tubes incubated at 37°C with shaking overnight. To extract the plasmids, 1.5 mL of the bacterial cultures were spun at 10000 rpm for 5 min and the media discarded. Re-suspension buffer (250 μL) was added to the tubes and the tubes vortexed to re-suspend the pellets. The cells were prepared for selection by adding 250 μL of lysis buffer and the tubes inverted four times to mix.

Precipitation buffer (350 μL) was added to the tubes and centrifuged at 10000 rpm for 10 minutes. The precipitates were discarded and supernatant transferred into spin columns. The columns were centrifuged for 1 minute and the flow-through discarded. Wash buffer (750 μL) was added to the tubes, centrifuged for 1 minute and the flow-through discarded. The blank columns were centrifuged for 1 minute and transferred into new tubes. To the tubes were added 50 μL of elution buffer and spun for 1 minute. The columns were discarded and the collected plasmid stored at -20°C. The plasmids containing the inserts were verified by digesting the transformants with *Eco*RI to release the inserts. In doing the selection, the reaction contents were 14.0μL

distilled water, 3 μL plasmid DNA, 2 μL buffer and 1 μL *Eco*RI. The reaction mixture was incubated for 1.30 hours at 37°C without shaking. To select positive transformants, the mixture was electrophoresed in 1% agarose gel and the bands visualized. The released DNA inserts were then sequenced in an automatic sequencer using primers GW1 (5'-GTT GCA ACA AAT TGA TGA GCA ATG C-3') and GW2 (5'- GTT GCA ACA AAT TGA TGA GCA ATT A-3').

3.2.10 Sequence Analysis and Comparison

Sequences of the two virus isolates (Appendix 3.2) were BLASTED against published sequences deposited in NCBI Server for comparison with published DNA sequences. About 27 orthologous DNA sequences (Appendix 3.3) were compared using Cluster option of the Multiple Sequence Alignment ClustalW (slow/Accurate, IUB) program based on bit scores from 904 (closely related) to 343 (least related) and E-values from 0.0 (closely related) to 1e-96 (least related). This is further elaborated on the results section below.

3.2.11 Detection of Integrated SatDNA Sequences and Survey for Satellites in Cassava Germplasm

In the light of new information that satellites might be integrated into the cassava genome (Rey and Ndunguru, pers. comm.) in discrete short sequences, primers were designed from cassava Expressed Sequence Tags (ESTs) (by German Plata, CIAT, Columbia), which amplify the random individual integrated sequences. For SatDNA-II, the primers SatIIF (5'-GCC GCA CGA CTG GAT CTC–3') and SatIIR (5'-CAG CAG CCA GTC AGG AAG TT-3') were designed to amplify a 895 bp fragment. For SatDNA-III fragments, the primers were designed to amplify a 306 bp DNA fragment: SatIIIF (5'-AGG CCT CGT TAC TAA AAG TGC-3') and SatIIIR (5'-ACC TGA CGG CAG AAG GAA T-3'). The amplification profile for both pairs of primers were 94°C 1min, 94°C 1min, 55°C 1min, 72°C 2min, 72°C 10min in 30 cycles. Using these sets of primers, an attempt was made to amplify satellites from 50 DNA samples from different infected cassava germplasm in each zone. A few non-symptomatic leaf samples were also included as controls. PCR products were purified

by precipitation with isopropanol and sent out to Biotechnology East and Central Africa (BeCA) facility in Nairobi for sequencing.

In order to have a better understanding of the occurrence of the satellite fragments, a second survey was performed, not only in the two cassava growing zones surveyed previously but country wide. Many cassava fields were visited in the major cassava growing zones and other areas alike, collecting leaf samples from diseased plants. In doing this, emphasis was placed on plants showing distinct CMD symptoms such as leaf distortion, leaf curling and leaf mosaic. The integrated forms of both SatDNA-II and SatDNA-III, were subjected to PCR amplification with a view of further characterizing them and mapping their spatial distribution in the country.

3.2.12 Sequencing Satellites Fragments

PCR amplification of SatDNA-II and SatDNA-III were carried out using primer sets; SatIIF (5'-GCC GCA CGA CTG GAT CTC–3') and SatIIR (5'-CAG CAG CCA GTC AGG AAG TT-3'); and SatIIIF (5'-AGG CCT CGT TAC TAA AAG TGC-3') and SatIIIR (5'-ACC TGA CGG CAG AAG GAA T-3'), respectively designed for the amplification of cassava genome integrated satellitte forms using amplification profile: 94°C 1 min, 94°C 1 min, 55°C 1 min, 72°C 2 min, 72°C 10 min, in 30 cycles. PCR products were analysed by electrophoresis on 1% agarose gel. PCR products were precipitated using isopopanol, and then re-suspended in 15 μL of sdH_20. The PCR products were cloned into pGEM T Easy vector (Promega, USA) following the protocol indicated in Appendix 3.5. The PCR products were ligated into the vector and multiplied in JM109 High Efficiency chemically competent cells (Stratagene, USA). Plasmid isolation was done using Qiagene Miniprep kit (Qiagene, USA). The transformed plasmids were sent out to Biosciences eastern and central Africa (BecA), in Nairobi for sequencing. The sequences were compared with published sequences using BLAST (Basic Local Alignment Search Tool) on the NCBI website.

3.3 RESULTS

3.3.1 Field Survey for CMD-infected Plants

A total of 147 cassava cuttings were collected from 60 sites in 26 randomly selected districts in the zones as shown in Table 3.2.

Table 3.2: Districts and number of fields covered during survey

Zone	Districts	Number of Fields (Name of sites in bracket)
Lake Victoria	Karagwe	2 (Kihanga and Nyakashushu)
	Bukoba Urban	2 (Ibura and Bukoba town)
	Bukoba Rural	2 (Rwizanduru and Buturage)
	Muleba	2 (Kiziramuyaga and Chamyorwa)
	Biharamulo	2 (Mihongora and Kikomakoma)
	Ilemela	2 (Buswelu and Isenga)
	Nyamagana	2 (Lwanhima and Buhima)
	Sengerema	2 (Nyamasale and Kizugwangoma)
	Ukerewe	2 (Msozi and Buramba)
	Kwimba	2 (Mwangika and Manda)
	Magu	2 (Kinango and Mkula)
	Musoma	3 (Bukanga, Kisamwene and Butiama)
Eastern Zone	Lushoto	1 (Kongei)
	Korogwe	3 (Kwasunga, Mgobe and Hale)
	Handeni	2 (Komsala and Msolwa)
	Tanga	1 (Mgwisha)
	Muheza	2 (Kwabastola and Mkinga)
	Pangani	1 (Masaika)
	Bagamoyo	3 (Kiwangwa, Kidomole, Kimange and Chambezi)
	Kibaha	3 (Visiga and Tumbi)
	Kinondoni	2 (Chaibora and Kimara-Suca)
	Temeke	5 (Mikwambe)
	Kilwa	3 (Kinyonga, Hamjambo and Nangurukuru)
	Lindi	2 (Mingoyo and Mahumbika)

	Masasi	2 (Mwena)
	Mtwara	5 (Mifugo, Chuno, Chikongola, Likonde, Moma)
Total	**26 Districts**	**60 Sites**

Source: Own Survey Data.

For Lake Victoria zone, the survey sites are indicated (with 'x') in Figure 3.2.

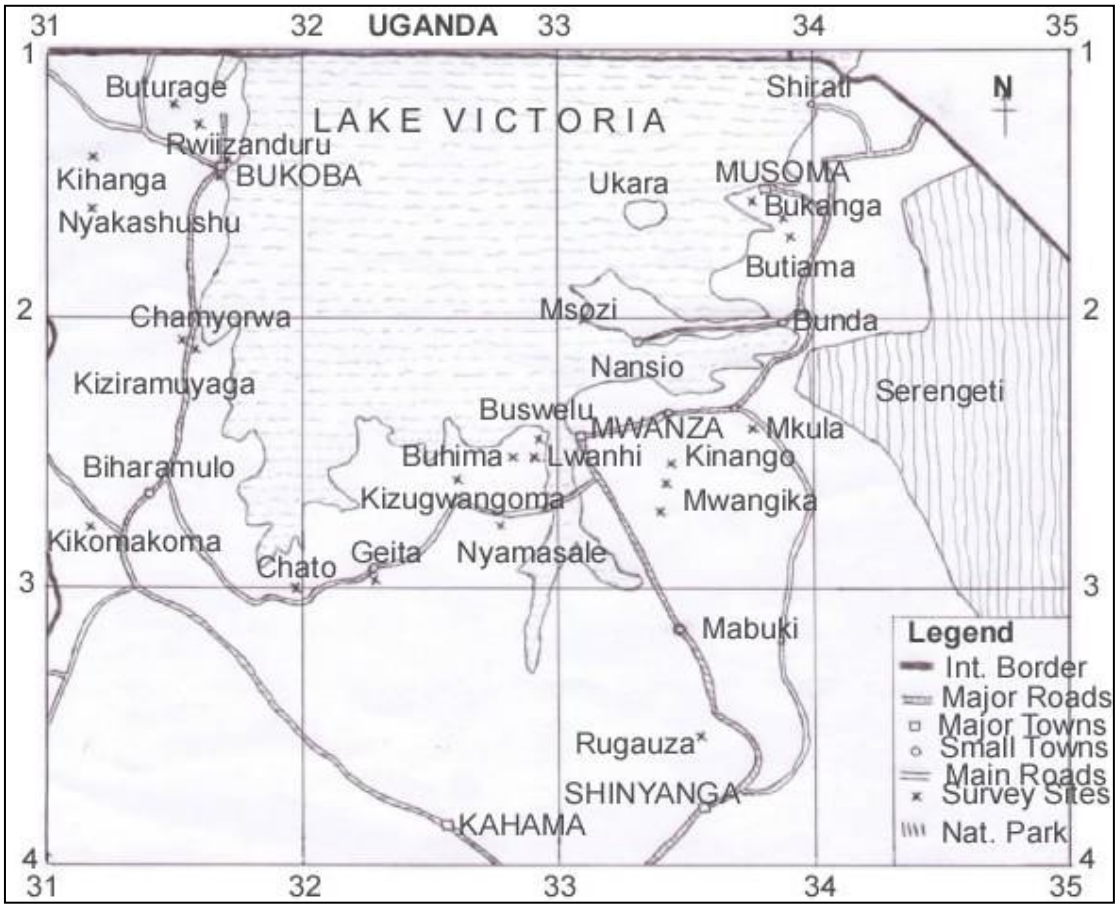

Figure 3.2: Map of the Lake Victoria cassava growing zone to showing survey sites (marked 'x').

Average CMD incidence results from the survey carried out in Lake Victoria Cassava growing zones are presented in Table 3.3.

Table 3.3: Average CMD incidence scores in Lake Victoria zone

No.	*Site*	*Variety*	**CMD Incidence (%)**	**Number of Stems Collected for replanting**
1	Kinango	Lyonjo	80	2
2	Mkula	Rangimbili	60	3
3	Msozi	Kitangaza	70	3
4	Buramba	Rwabukindo	60	2
5	Bukanga	Wamasuka	70	3
6	Kisamwene	Lumara	70	3
7	Butiama	Rwanguhira	60	2
8	Mwangika	Lyonjo	100	2
9	Manda	Lyonjo	60	2
10	Nyamasale	Rufaili	80	2
11	Kizugwangoma	Ngarabuto	80	2
12	Kiziramuyaga	Rusura	70	2
13	Chamyorwa	Lumara	60	2
14	Rwizanduru	Mporogoma	70	2
15	Buturage	Bukarasa	80	2
16	Ibura	Unnamed	70	2
17	Bukoba Town	Unnamed	80	3
18	Kihanga	Unnamed	80	2
19	Nyakashushu	Kaitampunu	70	2
20	Mihongora	Rubona	70	2
21	Kikomakoma	Unnamed	80	2
22	Lwanhima	Lyonjo	80	2
23	Buhima	Unnamed	70	2
24	Buswelu	Lyonjo	90	3
25	Isenga	Mapananzala	70	2

Table 3.3 shows that all the screened varieties grown in the different areas are affected by CMD. However, *cv* Lyonjo was far more susceptible to CMD than other

cultivars grown in those areas. Also most of the fields in the Lake Victoria zone had high disease incidence levels (60-90%).

Areas surveyed in the Eastern cassava growing zones are indicated in Figure 3.3

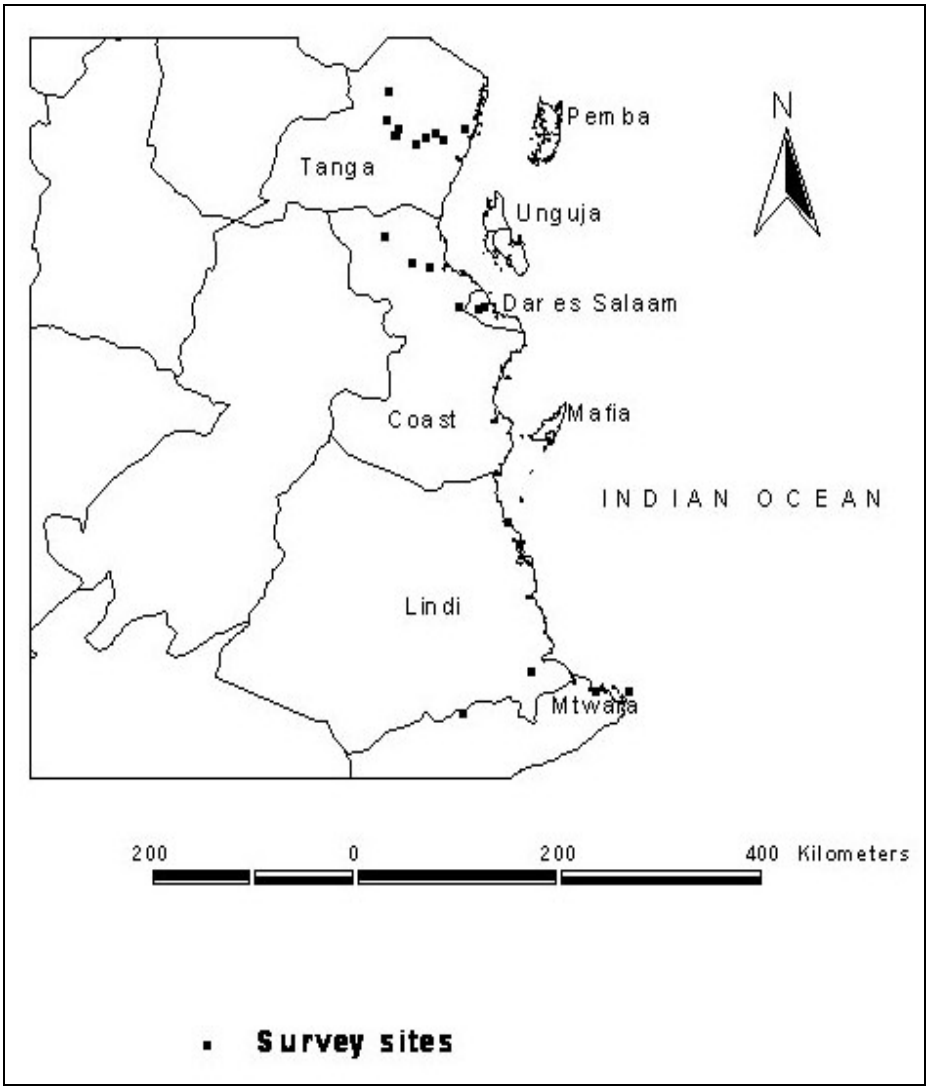

Figure 3.3: Map of Eastern Cassava-growing Zone to showing survey sites.

Just like for the Lake Victoria Cassava-growing zone, the symptoms observed differed over space and cultivars as shown in Table 3.4.

Table 3.4: Average CMD incidence scores in Eastern Zone

No.	*Site*	*Variety*	**CMD incidence (%)**	**Number of Leaf Samples Collected**
1	Kimange	Tabora	10	1
2	Kwasunga	Tabora	70	4
3	Kongei	Unnamed	40	5
4	Mgobe	Tingisha	60	6
5	Komsala	Kaniki, Tabora	90	3
6	Msolwa	Unnamed	90	4
7	Hale	Unnamed	80	4
8	Kwabastola	Unnamed	40	2
9	Mkinga	Unnamed	50	2
10	Masaika	Tingisha, Maiza	50	2
11	Mgwisha	Maiza, Tingisha, Kibandameno	60	3
12	Kiwangwa	Mfaransa	30	3
13	Kidomole	Mwanatabu, Kalolo	30	2
14	Chaibora	Maiza	80	3
15	Kimara-Suca	Unnamed	65	4
16	Tumbi	Unnamed	35	4
17	Chikongola	Sheria	40	2
18	Mifugo	Sheria	40	3
19	Chuno	Badi, Muhogo-Mpira	30	3
20	Likonde	Sheria	30	4
21	Mnazi Bay	Unnamed	20	2
22	Mahumbika	Sheria	10	2
23	Moma	Namikonga, Badi, Kigoma	40	4
24	Moma	Kigoma-Mafia	10	3
25	Moma	Nachinyanya, Kigoma-Zamani	10	2
26	Mwena	Unnamed	20	1
27	Hamjambo	Unnamed	25	2
28	Nangurukuru	Unnamed	25	2
30	Kinyonga	Unnamed	20	2
31	Mingoyo,	Unnamed	20	2
32	Visiga	Kigoma-Mtoto, Kiguru-cha-ninge	15	3
33	Visiga	Kiguru-cha-ninge	20	3

Source: Own Survey Data.

Table 3.4 shows that disease incidences in Eastern zone are generally lower than those for the Lake Victoria zone (Table 3.3)**.** In terms of the varieties grown, all were affected by CMD except *cv* Nachinyanya. This cultivar, which is widely grown in Mtwara, was not found to have disease symptoms in any of the sites surveyed.

3.3.2 Screen House Observations

The sprouted cuttings revealed a variety of symptoms on the newly emerging leaves (Table 3.5). Some of the cuttings whose mother plants were severely infected in the field were found to be healthy in the screen house even at 16 days after planting (DAP) (Figures 3.4A and 3.4B).

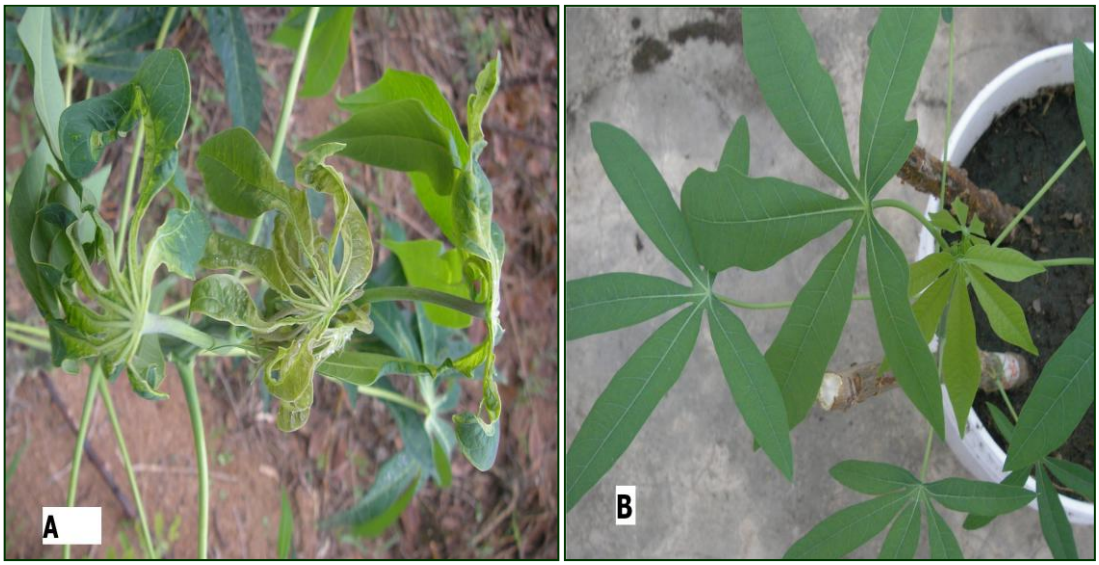

Figure 3.4: **A**. Appearance of plant labeled MU06S1 in the field. CMD symptoms are obvious in the field. **B**. Appearance of the same plant when sprouted in screen house at 16 DAP. The plant in the screen house is healthy (no CMD symptoms). [MU06S1 = sample code: MU = Musoma District; 06 = Field number 6; S1= sample 1].

On approaching 28 DAP and beyond, some of plants, similar to MU06S1 remained healthy as others became symptomatic. Other cassava plants, which exhibited mild symptoms in the field showed very severe leaf distortions in the screen house upon

sprouting. Some of the plants from different districts and sites exhibiting different symptoms are indicated in the Table 3.5.

Another observation was that two sections of the same cutting, planted in the same pot, each showed different symptoms: one had a severity scale of 3 in the screen house while the other section remained healthy up to 16 DAP (Figures 3.5A and 3.5B).

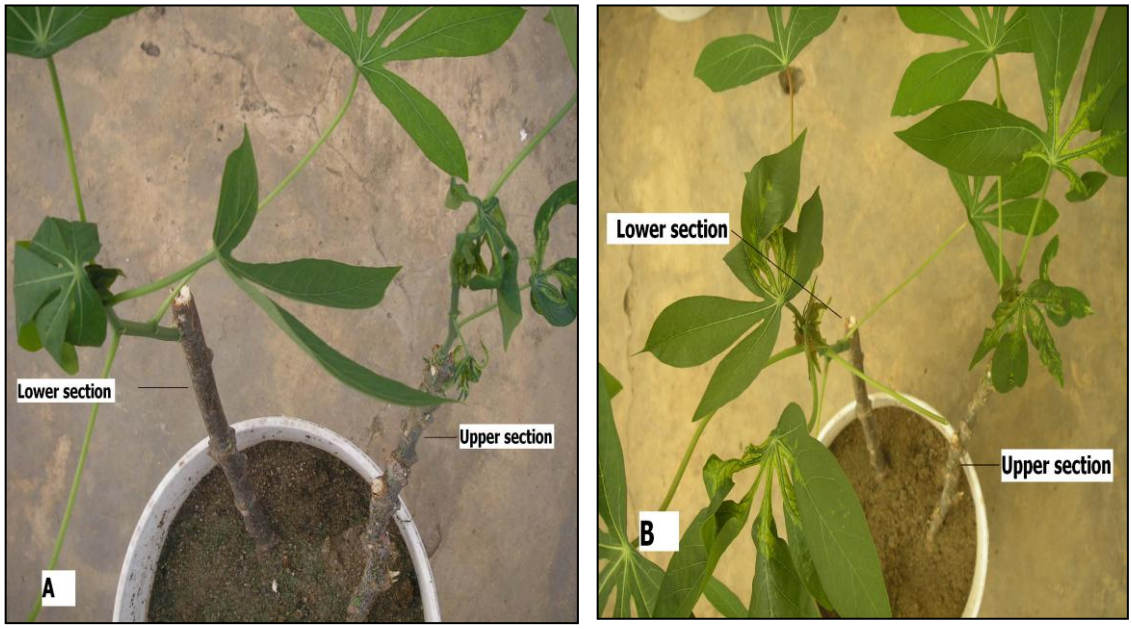

Figure 3.5: **A**. Leaf sample from plant labeled KN14S1 at 16 DAP (the lower section is healthy). **B**. The same sections at 29 DAP (both sections are symptomatic).

At 16 DAP (Figure 3.5A), only leaves growing from the upper section showed disease symptoms while at 29 DAP both leaves growing from both sections showed disease symptoms.

Such variability in symptoms derived from a single cutting was recorded in five cassava cuttings from the Eastern zone (Table 3.5). Except for a cutting coded KR07S3 collected from Korogwe District, in all the four remaining samples the younger (upper) section of the cutting became symptomatic earlier than the lower (older) part of the cutting. On keeping the cuttings longer (beyond 2 weeks) in the green house both parts of the cutting finally became infected, some showing mild and

others severe symptoms. When the plants were cut back and allowed to re-sprout, the same trend was observed 14 DAP.

Table 3.5: Symptom development of cuttings in screen house

Codes of samples which remained healthy to 16 DAP	Codes of samples which only became symptomatic at 28 DAP	Codes of plants which showed variation within same cutting	Codes of plants which recovered from symptoms
MG02S2, BG12S2, KN14S2	SE11M0, BR17M0	TG11S3, KR07S3	TG11S1
MG02S2, MT02S3, LU03S3	BU16M0, KG18M0	HD06S1, KN14S1	MH09S2
MG02S2, MT09S1, KR04S5	KG18M0, BH21M0	KB02S2	PG10S2
MG02S2, KL13S2, MG02S2	IL24M0		

Legend: DAP = days after planting; In the label, MG02S2: MG = District code (Magu); 02 = Field code (Field code for Mkula site in Magu District); S2 = Symptom severity code (severe, stem sample number 2 from that field); M0 = Symptom severity code (mild, the only stem in that field); BG = Bagamoyo, KN = Kinondoni, MT = Mtwara, LU = Lushoto, KR = Korogwe, KL = Kilwa, SE = Sengerema, BR = Bukoba Rural, BU = Bukoba Urban, KG = Karagwe, BH = Biharamulo, IL = Ilemela, TG = Tanga, HD = Handeni, KB = Kibaha, MH = Muheza, PG = Pangani.

In Table 3.5, plants in the first column remained healthy while plants in the second column started showing symptoms at 28 DAP. Some plants (Table 3.5) exhibited the phenomenon of recovery from disease symptoms. Newly formed leaves showed symptoms for a while but subsequently recover from the symptoms and become normal up to 18 DAP.

3.3.3 PCR Amplifications

The results of PCR amplifications for EACMV, ACMV, episomal SatDNA-II, and SatDNA-III are shown in Table 3.6.

Table 3.6: PCR results for EACMV, ACMV, SatDNA-II and SatDNA-III

No.	Leaf Sample	EACMV	ACMV	SatDNA-II	SatDNA-III
1	**ML13M0**	+	-	-	-
2	**UK03S2**	+	-	-	-
3	**BH20M0**	+	-	-	-
4	**NY22M0**	+	-	-	+
5	**MU06S1**	+	-	-	+
6	**BU16S1**	+	-	-	-
7	**BU15S0**	+	-	-	-
8	**UK03S1**	+	-	-	-
9	**SE11M0**	+	-	-	-
10	**ML13S0**	+	-	-	-
11	**BR14S0**	+	-	-	+
12	**BR14M0**	+	+	-	-
13	**KG18M0**	+	-	-	+
14	**BU16S2**	+	-	-	+
15	**BU16M0**	+	-	-	+
16	**BU15M0**	+	-	-	+
17	**KG19S0**	+	-	-	+
18	**BH21M0**	+	-	-	+
19	**BR17M0**	+	-	-	+
20	**SE10M0**	+	+	-	+
21	**BH21S0**	+	-	-	+
22	**MG01S0**	+	-	-	+
23	**MG01M0**	+	-	-	+
24	**MG02S1**	+	-	-	+
25	**MU05S1**	+	-	-	+
26	**MU06S2**	+	-	-	+
27	**MU06M0**	+	-	-	+
28	**ML12S0**	+	-	-	+
29	**ML13S0**	+	-	-	-
30	**KW08S0**	+	-	-	+

31	**BR17S0**	+	+	-	-
32	**NY23M0**	+	-	-	-
33	TG11S1	+	-	-	-
34	TG11S3	+	-	+	-
35	KR07S2	+	-	+	-
36	KR07S3	+	-	+	-
37	PG10S2	+	-	-	-
38	KR04S3	+	-	+	-
39	KR04S4	+	-	-	-
40	KN15S2	+	-	-	-
41	KN14S1	+	-	+	-
42	KR02S2	+	-	+	-
43	MH09S1	+	-	-	-
44	MH09S2	+	-	-	-
45	BG01S0	+	-	+	-
46	MT09S1	+	-	+	-
47	MT01S1	+	-	+	-
48	MT05M1	+	-	+	-
49	MT07S2	+	-	-	-
50	KL12S2	+	-	-	-
51	MT07S1	+	-	+	-
52	LD11M1	+	-	-	-
53	MT08S0	+	-	+	-
54	LU03S4	+	-	+	-
55	PG10S1	+	-	+	-
56	HD06S3	+	-	+	-
57	HD05S3	+	-	+	-
58	KN15S3	+	-	+	-
59	BG13S2	+	-	+	-
60	MS10S2	+	-	+	-
Total		**60**	**3 (= 5%)**	**19 (= 31.6%)**	**20 (= 33.3%)**

Source: Own Survey Data. **Legend:** District codes: ML = Muleba, UK = Ukerewe, NY = Nyamagana, MU = Musoma, LD = Lindi, MS = Masasi, KW = Kwimba, MT = Mtwara-Mikindani. Codes from the Lake zone are in bold.

EACMV was detected in all symptomatic leaf samples that were submitted to the laboratory for PCR analysis. As EACMV has since been reported to be highly diverse in Tanzania (Ndunguru, 2005), some of its strains could be supporting the replication

of SatDNA-II and others SatDNA-III. Table 3.6 shows that SatDNA-III is associated with leaf samples originating from the Lake Victoria zone, so the helper virus is likely to be EACMV-UG because it prevalent in the area. SatDNA-II, which was detected from leaf samples from the Eastern zone, could be replicated by EACMV. Only three cassava leaf samples out of 60 (5%) were found to be infected with ACMV. These leaf samples were collected at Rwizanduru (Bukoba Rural District), Nyamasale (Sengerema) and Buturage (Bukoba Rural) in the Lake Zone. Furthermore, these three samples infected with ACMV were also co-infected with EACMV. From Table 3.6, it seems unlikely for either SatDNA-II or SatDNA-III to be replicated by ACMV since very few leaf samples were positive for ACMV.

The laboratory investigations in general indicate that EACMV was encountered more frequently (89.5% of all tested leaf samples) than ACMV (Table 3.6), which was detected in only 5% of the leaf samples. Figure 3.6 demonstrates PCR results for some of the samples.



Figure 3.6: PCR amplification of EACMV core coat protein (540 – 560 bp) fragments. C = Positive control (Positive leaf); Last lane right: Molecular weight marker.

Of the total 60 leaf samples from both zones that were tested for SatDNA-II and SatDNA-III, 19 samples (31.6%) tested positive for SatDNA-II while 20 samples (33.3%) tested positive for SatDNA-III (Table 3.6). Some of the PCR results from some of the leaf samples are shown in Figures 3.7 and 3.8



Figure 3.7: Amplification of full-length SatDNA-II from the Lake & Eastern zones. 1 – Sengerema, 2 – Magu, 3 - Biharamulo, 4 – Ukerewe, 5 – Muleba, 6 – Bukoba Rural, 7 – Bukoba Urban, 8 – Bukoba Urban, 9 – Karagwe, 10 – Musoma, 11 – Kibaha, 12 – Korogwe, 13 – Mtwara, 14 – Lushoto, 15 – Handeni, 16 – Kilwa, 17 – Kinondoni, 18 – Tanga, 19 – Lindi, 20 – Bagamoyo, 21 – Mtwara. Lane L – DNA marker.

In Figure 3.7 leaf samples 1 to 10 were collected from the Lake Victoria zone while samples 11 to 21 were collected from the Eastern zone. On the agarose gel, samples from the Eastern zone have many other bands, in addition to the expected band (of about 1 kb).

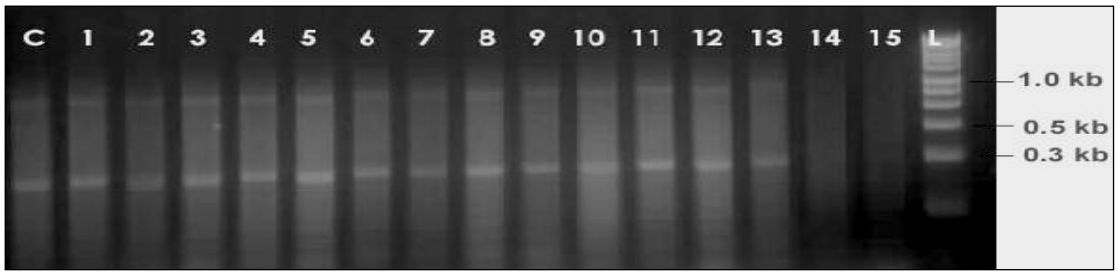

Figure 3.8: PCR amplification of full length SatDNA-III from the Lake zone.

Legend: C – Control, positive leaf sample; 1 – 15, leaf samples collected from different sites in the Lake Victoria zone, with samples 14 and 15 being negative. L – DNA marker

3.3.4 Results of Sequence Analysis

The consensus sequences of full-length DNAs of TZ 113 and TZ 108 (Appendix 3.2) were aligned with other begomovirus sequences (Appendix 3.3) using Clustal W and a phylogenetic tree constructed (Figure 3.9).

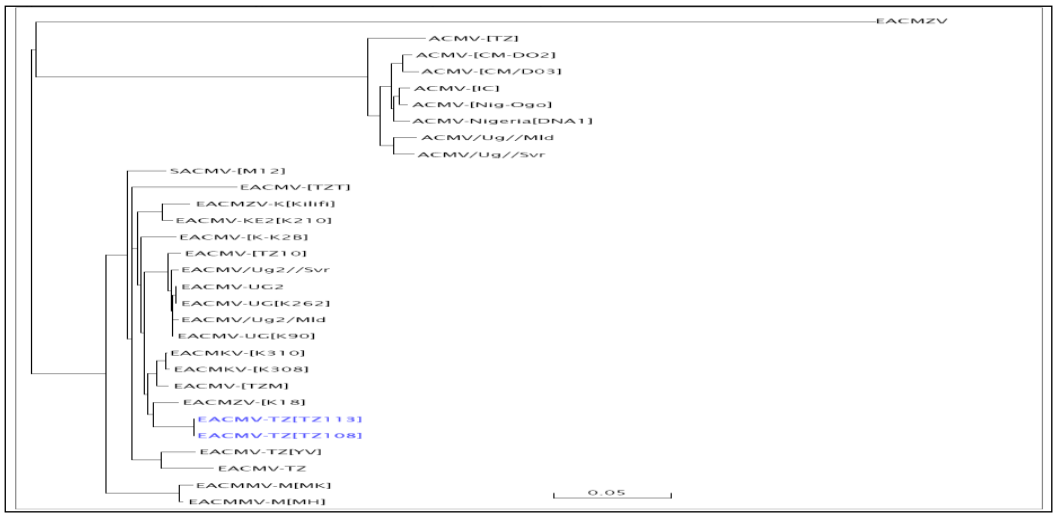

Figure 3.9: Phylogenetic tree of EACMV-TZ[TZ113] and EACMV-TZ[TZ108] showing the relationship with other cassava begomoviruses. In the tree, sequences for TZ113 and TZ108 group together, being more closely related to EACMZV-[K18].

Due to the number of genetic recombinations now known to occur among CMGs, individual ORF sequences were not considered as criteria for the establishment of CMG species. Instead, sequence fragment of DNA-A component was used. The CR sequences were considered for this analysis due to their importance for geminivirus replication.

3.3.5 Survey of Satellite-like molecules in Tanzania

Results from the second survey show that the occurrence of the integrated form of satellites over the country seems to be a widespread occurrence (Figure 3.10).

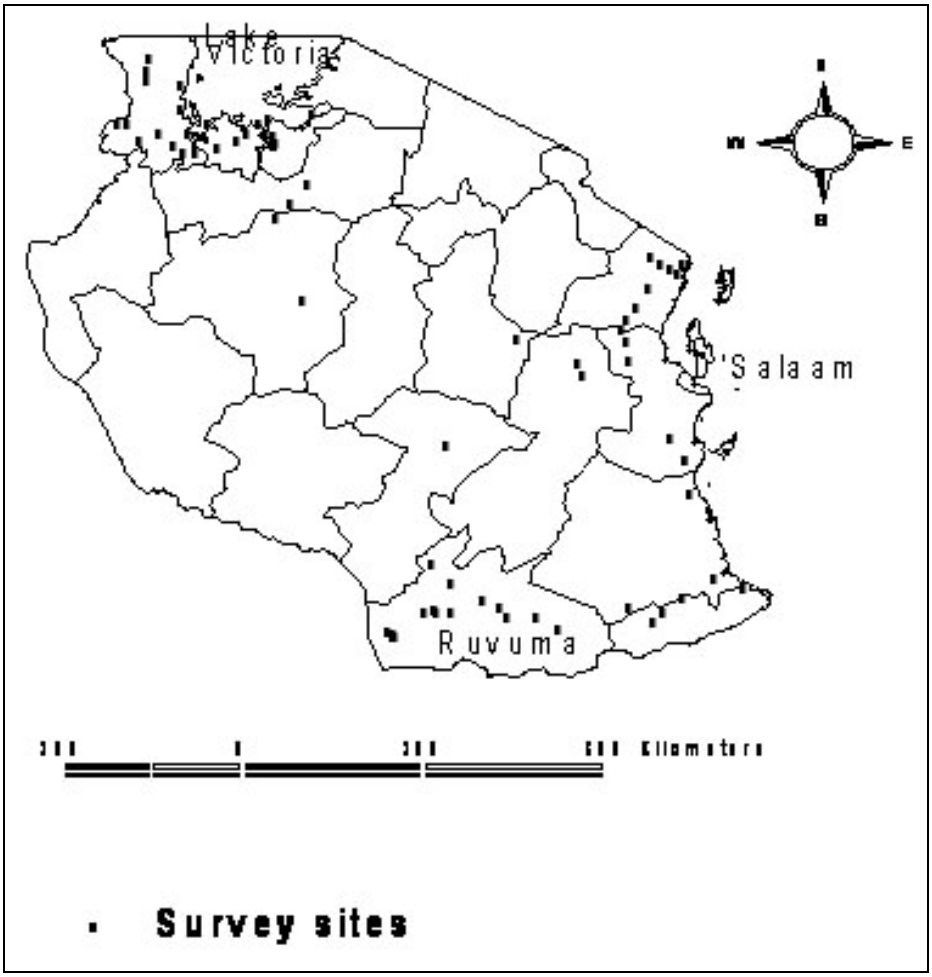

Figure 3.10: Survey sites where integrated forms of satellites were documented.

The results of the country survey for the occurrence of integrated satellite sequences in cassava genome are presented in Appendix 3.4. The area covered in this survey was not limited to major cassava growing zones but included other areas to explore the occurrence of these satellite-like molecules as wide as possible. Out of 40 samples collected from Mbinga District (Ruvuma Region) in 2007 for example, 32 tested positive for SatDNA-II, equivalent to 80%, while 18 tested positive for SatDNA-III equivalent to 45%.

From this survey, both SatDNA-II and SatDNA-III integrated sequences occurred widely in the survey area, as seen in Appendix 3.4. Out of 222 leaf samples collected in the survey area and submitted to the laboratory for PCR amplification, 151 tested positive for SatDNA-II fragments (equivalent to 68.0%) while 158 tested positive for SatDNA-III fragments (equivalent to 71.1%). PCR amplification of the integrated

satellite sequences also show that the sequences are present in almost every sampled cassava plant, whether diseased or symptomless. Some of the PCR results are shown in Figure 3.11.

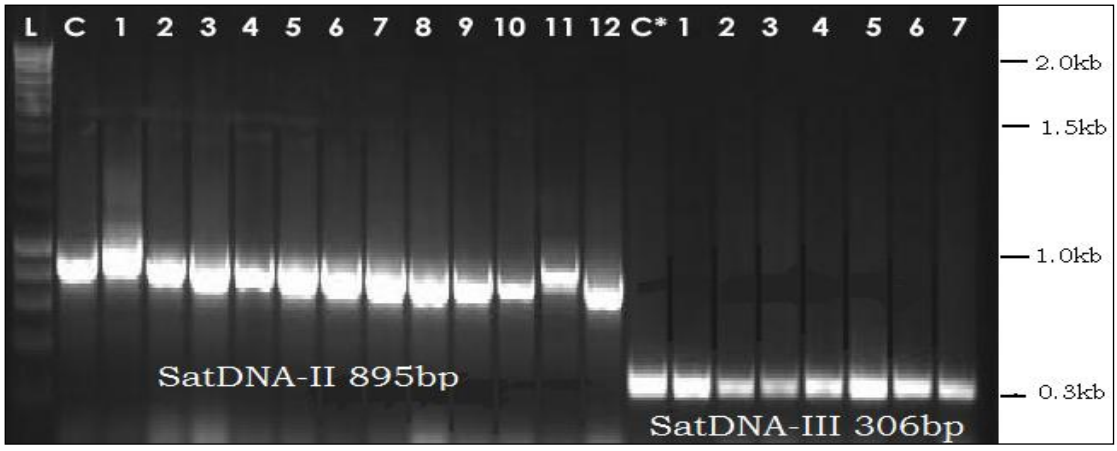

Figure 3.11: The amplification of SatDNA-II and SatDNA-III fragments from diseased and symptomless samples raised in the screen house. C = positive control for SatDNA-III; C* = Positive control for SatDNA-II. Lane on the extreme right = DNA marker.

SatDNA-II bands on the left (lanes 1 – 12) are variable in size while SatDNA-III bands on the right (lanes 1 – 7) are rather uniform in size along the gel. The variation of the gel band sizes for amplified SatDNA-II also was sometimes accompanied by the occurrence of a second band as shown in Figure 3.12.

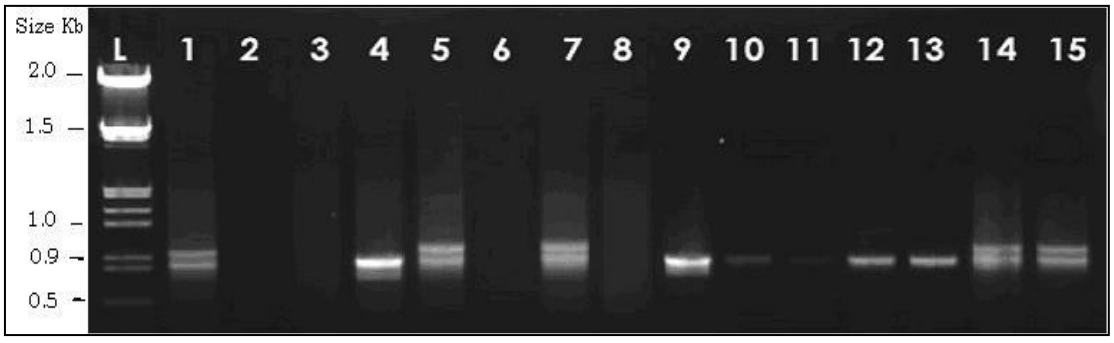

Figure 3.12: PCR results showing double amplified DNA bands when SatDNA-II is separated on 1% agarose gel. Sample 1: *cv* Mwaya from Milonde in Tunduru District; Sample 5: *cv* Nakalai from Mapera in Mbinga District; Sample 7: *cv* Mkongo-Mwaya

collected from Peramiho in Songea District. Samples 14 and 15: in *cvs* Local and Shija, collected from Madaba (Songea) and Nzega (Tabora), respectively.

The occurrence of a second band other than the expected band for the integrated satellite sequences was also seen in some leaf samples collected from the Lake Victoria Zone (Figure 3.13).

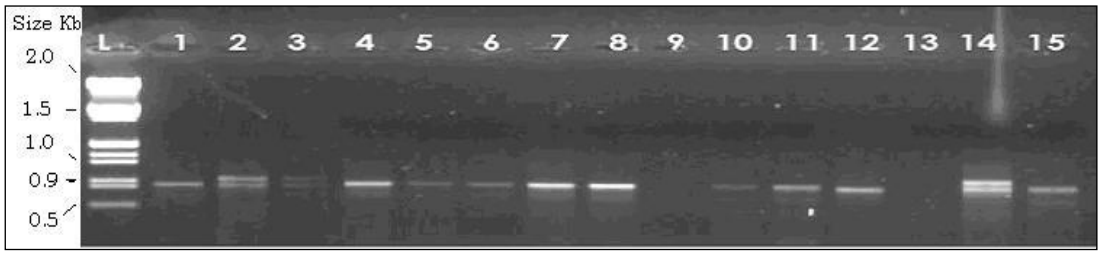

Figure 3.13: Integrated SatDNA-II sequence amplification showing double bands. Lane 2: cv Lyonjo from Misungwi District in Mwanza; Lane 3: *cv* Lyonjo from Geita District in Mwanza; and Lane 14: *cv* Kajaga from Kimwani, Muleba District in Kagera. L = DNA 2 kb Ladder

Appearance of additional bands on agarose gel was not the only variation observed during PCR amplification of integrated SatDNA-II sequences. Figure 3.14 shows the occurrence of distantly spaced bands lower than 500bp, which is not the expected ~900bp fragment seen in lanes 2 and 6. The stems from which the leveas were picked were all collected from Mbinga District in the Southern zone.



Figure 3.14: Lanes 2 and 6 were the expected bands for integrated SatDNA-II. The other lanes are distantly spaced (size below 0.5 kb). Sample in lane 1 was collected from cv Tupuka at Kingerikiti; Lane 3: *cv* Nakalai (at KwaTai); Lane 5: *cv* Nakalai (at Makuli-Mbalama); Lane 7: *cv* Tupuka (at Kinzege); Lane 9: *cv* Nakalai (at Mpapa); Lane 10: *cv* Tupuka (at Ilela); Lane 13: cv Kausa (at Kingerikiti-Pisi); Lane 14: *cv* Nakalai (at Luhalala), Lane 15: *cv* Nakalai (at Iringa-Litindo); Lane 17: *cv* Nakalai (at Mikalanga-Baraza); Lane 18: *cv* Kausa (at Matarawe-Lumeme); Lane 19: *cv* Nakalai (at Mapera-Bula); Lane 20: *cv* Nakalai (at Mbalama); and Lane 21: *cv* Tupuka (at Tugutu Kilimani). L = DNA 2kb Ladder.

About 87.5% of the bands are out of the expected size position. Samples 2 and 6 are the expected SatDNA-II bands (895bp) but the other lanes show smaller bands. These samples were mostly from Eastern part of Mbinga District where surveys had never been done before. Upon cloning (Protocol indicated as Appendix 3.5) of the PCR products and sequencing, the patterns for the sequences obtained are described under section 3.3.6.

3.3.6 Sequence Analysis of Satellites Fragments

Upon sequence comparison (BLAST) on NCBI, the sequences of integrated SatDNA-II isolate fragments closely resembled *Begomovirus-associated DNA-II satellite* (Accession Number AY836366) (96-98%) available in the GenBank. Some of the isolates also shared an 85-92% nucleotide identity with *Mentha leaf deformity-associated satellite DNA-II* (Accession Number EU862815). SatDNA-III isolate sequences resembled *Begomovirus-associated DNA-III satellite* (Accession Number AY836367), with a 76-94% nucleotide sequence identity.

3.3.6.1 SatDNA-II

Sequence editing was done using DNASTAR software, followed by sequence alignment using CLC Sequence Viewer 6.4, yielding a phylogenetic tree as shown in Figure 3.16.

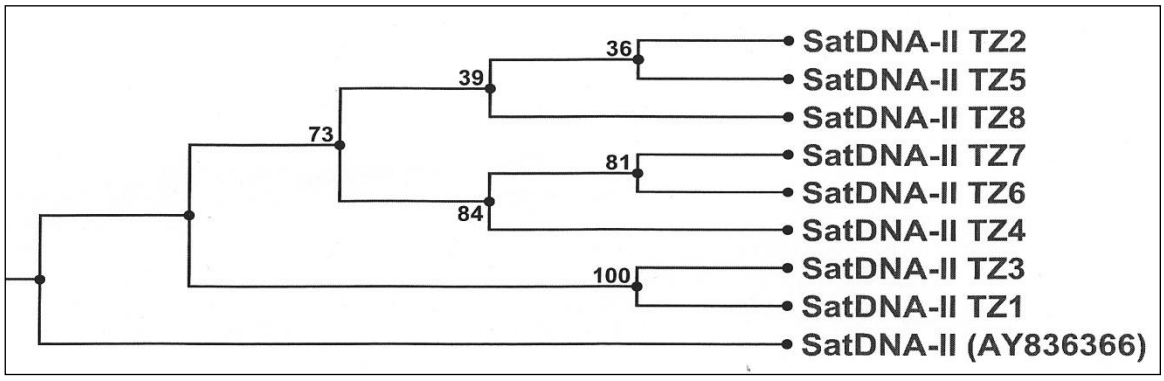

Figure 3.16: Phylogenetic tree based on a neighbour-joining analysis showing nucleotide sequence relationships among eight Tanzanian SatDNA-II isolates. The tree was rooted with the original sequenced SatDNA-II isolate (AY836366).

Locations from which the isolates were collected were: SatDNA-II TZ1 from Tanga, SatDNA-II TZ2 (Tanga), SatDNA-II TZ3 (Mwanza), SatDNA-II TZ4 (Mwanza), SatDNA-II TZ5 (Muleba), SatDNA-II TZ6 (Morogoro), SatDNA-II TZ7 (Tabora) and SatDNA-II TZ8 (Shinyanga). Sequence length of the isolates was variable: SatDNA-II TZ1 was 526 nts, (SatDA-II TZ2 (834 nts), SatDNA-II TZ3 (524 nts), SatDNA-II TZ4 (876 nts), SatDNA-II TZ5 (861), SatDNA-II TZ6 (866), SatDNA-II TZ7 (861 nts) and SatDNA-II TZ (8). Sequence alignment for SatDNA-II isolates using DNASTAR software (Clustal W, Slow/Accurate, IUB) gave sequence pair distances as shown in Table 3.7

Table 3.7: Distance Table for SatDNA-II isolates showing percentage nucleotide sequence identities (upper part) and divergences (lower part of the table).

	SatDNA-II (AY836366)	**SatDNA-II TZ1**	**SatDNA-II TZ2**	**SatDNA-II TZ3**	**SatDNA-II TZ4**	**SatDNA-II TZ5**	**SatDNA-II TZ6**	**SatDNA-II TZ7**	**SatDNA-II TZ8**
SatDNA-II (AY836366)	-	71.8	**98.4**	**72.6**	**95.4**	**93.0**	**97.5**	**96.0**	**96.0**
SatDNA-II TZ1	35.5	-	70.0	**97.0**	69.2	67.6	69.8	68.0	67.9
SatDNA-II TZ2	1.6	38.0	-	70.6	**95.1**	**92.6**	**97.0**	**95.5**	**92.9**
SatDNA-II TZ3	34.9	3.2	37.4	-	69.1	68.2	70.8	68.6	68.1
SatDNA-II TZ4	4.5	40.6	5.4	40.8	-	**90.4**	**93.7**	**93.7**	**92.7**
SatDNA-II TZ5	6.5	42.3	7.3	41.6	10.2	-	**90.7**	**90.5**	**88.2**
SatDNA-II TZ6	2.4	39.1	3.7	38.5	6.1	9.1	-	**95.8**	**97.7**
SatDNA-II TZ7	4.5	41.5	5.3	41.2	7.1	10.4	4.3	-	**92.6**
SatDNA-II TZ8	3.8	41.7	5.3	41.0	6.9	9.8	5.3	6.7	-

Table 3.7 shows that the isolate sequence that showed maximum nucleotide sequence similarity with the reference sequence [SatDNA-II (AY836366)] was SatDNA-II TZ2 (98.4%) while isolate SatDNA-II TZ1 showed the least nucleotide sequence identity (71.8%). Among the isolates themselves, the maximum similarity was observed among isolate pairs SatDNA-TZ1 and SatDNA-II TZ3 and SatDNA-II TZ2 and SatDNA-II TZ6, both with nucleotide sequence identity of 97%. These isolates also grouped closely in the phylogenetic tree (Figure 3.16). The least nucleotide identity (67.9%) with the reference SatDNA-II were SatDNA-II TZ1 and SatDNA-II TZ8 isolates.

Some of the nucleotide sequences examined showed several features including the GC-rich regions comprising of trinucleotides (CGC) at nts 2-3, 59-61, 87-89, 118-120, 122-124 and 159-161 and hexanucleotides (CCGCCG) at nts 99-104 and 110-

115. These are shown in Appendix 3.6. There is also a Poly (A) signal (AATAAA) at nts 530 – 535 (Appendix 3.7) and a TATA box Protein binding (TBP) motif TATAAAT at nts 817 – 825 (Appendix 3.8). The TATA box is preceded by nucleotides TG and upstream of the TATA box also appears the CAAT box at nts 808 to 811.

3.3.6.2 SatDNA-III

Editing for SatDNA-III isolate sequences was done using DNASTAR. A neighour-joining phylogenetic tree was generated using CLC Sequence Viewer 6.4, and a phylogenetic tree (Figure 3.17) was generated.

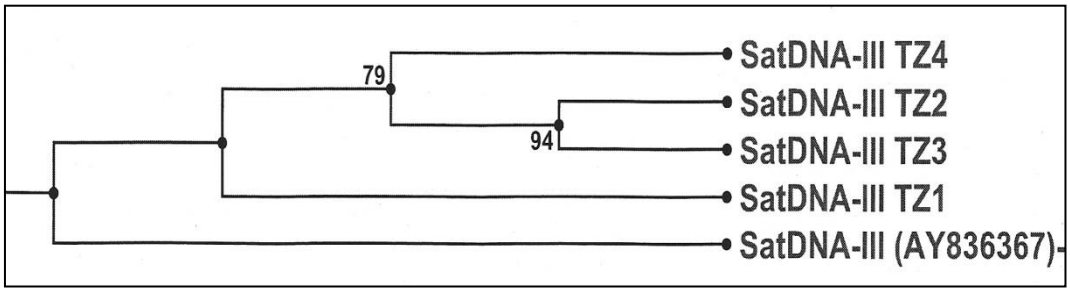

Figure 3.17: Phylogenetic tree showing nucleotide sequence relationships among four Tanzanian SatDNA-III isolates. The tree was rooted using the original sequenced SatDNA-III (AY836367).

Multiple sequence alignment for SatDNA-III isolates was performed using DNASTAR software (Clustal W, IUB) gave nucleotide sequence pair distances as shown in Table 3.8 below.

Table 3.8: Distance Table for SatDNA-III isolates showing percentage nucleotide sequence identity (upper part) and divergence (lower part of the table)

	SatDNA-III (AY836367)	**SatDNA-III TZ1**	**SatDNA-III TZ2**	**SatDNA-III TZ3**	**SatDNA-III TZ4**
SatDNA-II (AY836366)	-	**78.1**	**79.3**	**80.2**	**75.3**
SatDNA-III TZ1	26.0	-	**70.9**	**78.3**	**83.3**
SatDNA-III TZ2	24.2	43.8	-	71.4	68.5
SatDNA-III TZ3	22.2	29.5	34.0	-	72.1
SatDNA-III TZ4	33.8	17.9	51.7	35.5	-

Note: High percentages are in bold.

Table 3.8 shows that an isolate SatDNA-III TZ3 was closely related to the reference sequence [SatDNA-III (AY836367)] with a 80.2% nucleotide sequence identity. SatDNA-III TZ4 was 75.3% related to the SatDNA-III (AY836367)]. Among the isolates themselves, the maximum similarity observed was 83.3% for isolates SatDNA-III TZ1 and SatDNA-III TZ4. The least nucleotide identity was shown by the isolate pair SatDNA-III TZ1 and SatDNA-III TZ2 (70.9%). Locations from which the isolates were collected are: SatDNA-III TZ1 from Mtwara, SatDNA-III TZ2 (Mtwara), SatDNA-III TZ3 (Songea) and SatDNA-III TZ4 (Geita). Sizes of these isolate sequences are SatDNA-III TZ1 (563 nts), SatDNA-III TZ2 (558 nts), SatDNA-III (530 nts), and SatDNA-III (513 nts).

The GC-rich regions for SatDNA-III comprised pentanucleotides (CCGCC) and trinucleotides (CGC) located at nts 146 – 150 and 188 – 190, respectively (Appendix 3.9). Poly (A) signal, TATA box Protein binding (TBP) motif, and CAAT boxes were not found in SatDNA-III.

3.4 DISCUSSION

3.4.1 Severity of Infection and CMD Incidence in the Field

There is one important observation for each of the estimates of CMD in this study, i.e. severity and incidence. For all the fields irrespective of cassava growing zone, plants in those fields along roadsides and within urban areas had high severity scores than those plants in fields that were far away from roadsides and towns. In the Lake Victoria zone for example, a field at Mwangika in Kwimba District was along the road and had severity score of 4 and 5. The appearance of plants in this field is shown as Figure 2.3 in Chapter 2. This was also observed in plants grown in backyard gardens in urban areas as was noted in Bukoba town (Lake Zone) and at Chikongola in Mtwara (Eastern Zone).

For the case of incidence, cassava in the fields in the Lake zone showed higher CMD incidences than those of the Eastern Zone. Cassava in the Bukoba and Sengerema districts in the Lake Victoria Zone exhibited high disease incidences. Also, fields along the roadsides and closer to towns had higher incidences than those situated in remote from town and roadsides. The observation suggests that the spread of CMD was associated with the movement of infected stakes, implicating human influence in the spread of the disease. The incidence of CMD symptoms was variable, depending on the cassava variety. In the Lake zone, cv. *Lyonjo* was more severely infected (incidence 80-100%) compared to other varieties. In one field in Kwimba District, the whole field planted to *Lyonjo* was affected. In Mtwara, cvs *Kigoma-Mafia* and *Nanchinyanya* both encountered at Moma location had low level of disease incidence (less than 15%). Within individual plants, some branches showed symptoms while other branches remained symptomless.

These disease severity and incidences presented in this study suggest that researchers need to be more careful when collecting diseased leaf samples from cassava fields for incidence and severity measurements. There has been a tendency to pick samples from the roadsides and in towns because of accessibility, and the results obtained are not accurate in describing the disease situation. Conclusions could therefore be made

about disease based on human movement patterns as opposed to natural (whitefly) infection in the fields, thereby inflating disease incidence figures. Deliberate efforts should be made to sample more remote areas in order to obtain more accurate disease incidence estimates. Because of these trends in the occurrence of CMD symptoms it is also difficult to pre-determine survey sites because some targeted areas may have low levels of disease incidence or may even be free from disease.

3.4.2 Variability of Cassava Cultivar Symptom Phenotypes

3.4.2.1 Severe Symptoms in Field but Healthy in Screen House

It has been shown in some studies that replication and movement of viruses can be influenced by external environmental conditions, such as temperature, thereby influencing symptom expression and severity (Chellappan *et al*., 2005). In virus–infected plants, high temperature is frequently associated with attenuated symptoms (heat masking) and with low virus titer. Recovery of some field cassava cuttings in the screen house with regards to symptoms was shown in the present study, and has also been demonstrated in other studies involving geminiviruses (Hagen *et al*., 2008).

The expression of virus symptoms in cassava plants is a gradual process characterized by discrete stages such as infection, multiplication of the pathogen and symptom expression. The stakes were re-grown in the screen house with maximum temperature ranging 30 to 40°C. It is therefore postulated that some of the virus infection stages may be temperature-dependent, and are down-regulated by higher temperature in some varieties. In so doing, the host plant is prevented from reaching higher levels of infection.

Recovery can also be due to the fact that some branches exclude spread of virus which may be related to expression of certain host mechanisms later on leading to virus-free branches. Viruses may not be evenly distributed between branches from a cutting and some branches may have lower virus titer due to patterns in virus movement, i.e. the plants are ‘chimeric’ or ‘mosaic’. So, if some stakes are cut that do not have leaves you will not be sure if they have symptoms so you may choose a virus-free or low-virus branch although not if later than 28 DAP. Some branches may

be free from the virus through reversion. Such branches are likely to grow healthy in screen house because they are virus-free or have low levels of virus titer. The hypothesis of restricted movement is also supported by the existence of a variety of biochemical responses by plants to curtail replication and systemic movement of viruses (Patil and Fauquet, 2009). For example, cassava plants encode proteins such as the highly conserved ubiquitin in response to begomovirus infection, which cause selective degradation of the viral proteins (Kemp *et al*., 2005). In the process, viral gene products selected for degradation are marked by linking them to ubiquitin. Plants, including cassava, target selected protein to the proteasome, which recognizes ubiquitiated proteins induced by pathogen and are then degraded. So, it is possible that the proteasome is up-regulated at higher temperatures in particular varieties.

Cassava is also known to produce salicyclic acid and reactive oxygen species (ROS) in response to wounding or infection (Reilly *et al*., 2004). ROS and salicyclic acid can also be induced by abiotic stress (heat and drought). ROS is induced when plants are exposed to ultraviolet A (UVA) (320 – 380nm wavelength) leading to the splitting of water. Although the presence of ROS has been demonstrated in tubers, it could as well occur in stems and leaves from cuttings since they also suffer from wounds during the cutting process. So, it could be that the cultivars exhibited this phenomenon under screen house conditions.

Plants can also be healthy in the screen house due to cross-protection as described by Owor *et al* (2004) in Uganda. Viruses in mixed infection lead to subdued symptom expression in some cultivars. There is thus a possibility that the cutting had been infected by two viruses with a strong tendency of cross-protection at temperatures higher than 30°C. This may be possible since a stake coded BR17S0 for example, taken from the same field at Buturage in Bukoba Rural District as one of the screen house-healthy samples, was confirmed in the laboratory to be co-infected by EACMV and ACMV. However synergistic infections with more than one begomoviruses in cassava may not always lead to reduced infections (Patil and Fauquet, 2009). In fact, there have been many reports of enhanced symptoms when more that one virus is present (references), and there are no reports on mild cassava begomoviruses causing attenuated disease symptoms.

While symptom recovery may not be related to several interacting factors such as temperatures, abiotic stress-induced systemic acquired resistance (SAR), or cultivar dependence, the most convincing approach to this reversion of symptom expression however, is the concept of RNA interference (Gibson *et al*., 1997). Plants tend to curtail the effect of infecting viruses by the production of small interfering RNAs (siRNAs) by the enzyme dicer, which splice viral mRNAs. The production of siRNAs in plants challenged or containing various geminiviruses is affected by temperature (Nelson and Citovsky, 2005). RNA silencing was reported to increase as the temperature was raised from 25°C to 30°C. Cassava cultivars exhibiting the recovery phenotype may express RNA silencing in later stages of infection.

Viruses may not be evenly distributed between branches and some branches may have lower virus titer due to patterns in virus movement, i.e. the plants are 'chimeric' or 'mosaic'. Hence, it is not possible to predict the disease status of cassava stakes without leaves, as one stake from a single branch may have a higher virus titer compared to another branch. Some branches may even be virus-free even though the plant is infected, leading to the planting of virus-free or low-virus stakes in the screen house. Some branches may be free from the virus through reversion, and are likely to grow healthy in screen house because they have no viruses.

3.4.2.2 Mild Symptoms in field but severe in screen house

The progression of disease development tends to go from mild to severe in most cases. In the present case, it may be that the plants had just been infected by the virus and that they were in the process of developing symptoms (from mild to severe). If being mild was the characteristic of that cultivar, then the influence of host heat shock proteins (Hsp70) may have played a role in the screen house. Hsp70 are highly conserved ≈70-kD monomeric proteins in eukaryotes whose rate of synthesis greatly increases at elevated temperatures (Voet and Voet, 2004). Nelson and Citovsky (2005) have reported that in soils with abundant cadmium, host Hsp70 and virus encoded Hsp-related proteins are likely to help viral and/or host macromolecular transport through plasmodesmata. The higher temperature (up to 41°C) in the screen

house could be a sufficient stress to trigger the production of Hsp70, thereby causing the otherwise mild symptoms to become severe.

3.4.2.3 Differences in symptom expression within same plant

Viruses are transported through the phloem alongside photoassimilates. In the three cassava plants showing differences in symptom expression (i.e. plants labeled TG11S1, PG10S2 and MH09S2), cultivar-specific-characteristics may have allowed a more efficient virus movement through sieve elements. Older cuttings from the lower sections of TG11S1, PG10S2 and MH09S2 may have influenced sieve element translocation efficiency and therefore virus symptom development. It is suggested that it is not only the cultivar age, but vascular tissue morphological differences that might have contributed to this observation. For another plant (KR07S3) collected from Hale in Korogwe District, there may be two possibilities for the observed difference in symptom expression. One, the plant was infected by a different virus strain compared to the three plants (TG11S1, PG10S2 and MH09S2) collected from different and distant districts (Tanga, Handeni and Kinondoni) as revealed by the GPS coordinates.

The phenomenon of recovery (the plant producing fewer and fewer symptoms over time following infection) may be associated with gene silencing in some cultivars. Nelson and Citovsky (2005) have indicated that increase in siRNA was most striking by 3-to 6-fold for non-recovery-type geminiviruses (EACMV and ICMV) compared to those that were associated with recovery. It is further elaborated that the dramatic increase in siRNAs was correlated with the presence of one of two viral suppressors in these geminiviruses.

In another study in cassava, Chellappan *et al* (2005) supports the idea of recovery through PTGS. Biolistically inoculated cassava and *N. benthamiana* developed disease symptoms and reached maximum severity in 2 to 3 weeks post inoculation. Later, the newly emerged leaves of infected plants showed recovery from symptoms and the recovery was correlated with the abundance of virus-derived siRNA in both plants. It was further observed that the extent of recovery was more dramatic at high temperature (30°C) than plants grown at lower temperature (25°C). The fact that not all plants recovered from disease symptoms could mean that the phenomenon is cultivar related.

PCR results show that most of the CMD cases were attributed to EACMCV. As indicated in literature, ACMV occurs on the Eastern part of the great East African Rift Valley (Berrie *et al*., 2001). All the three plants in which ACMV was detected were collected from the Lake zone, which lies on the East of the Rift Valley. SatDNA-II was encountered in samples collected from the Eastern zone and SatDNA-III on samples collected from the Lake zone.

Relating the studies above to the screen house observations it shows that out of the four plants exhibiting variation in symptoms in the same cutting (reversion), three (equivalent to 75%) were infected with SatDNA-II. It is therefore possible that the reversion phenomenon may be associated with the presence of SatDNA-II. The satellite (SatDNA-II) was detected in leaves collected from each of the three plants. Characteristic symptoms included papery leaf lamina; a sickle-shaped leaf when one side of the midrib is infected and s-shaped leaves where both sides of the midrib is infected by the satellite; narrowing (tapering) of leaflet bases; and down cupping. Symptoms on samples in which SatDNA-III was amplified had filiform appearance. It is possible that the occurrence of the satellites is associated with the occurrence of these symptoms.

In summary, variation in symptom phenotype and severity are usually dependent on a number of factors, and it would be difficult, without further studies, to conclude in this work what are the contributing factors to these symptom variations. It would seem from other studies (Fraser, 2006; Patil and Fauquet, 2009) that symptom recovery from viral infection or reversion of severity is due to RNA silencing, which is similar to PTGS, restriction of virus movement, or due to the presence of other synergistic viruses (mixed infection) or satellites/defective molecules. Evidence that symptom phenotype changes in this study may be related to the presence of these partially characterized satellite-like molecules, SatDNAII and SatDNAIII, is the observation to support the fact that there had been a change in symptoms in TME 3 (Ndunguru, unpublished data).

3.4.3 Characterization of the EACMVs

The phylogenetic tree reveals two main groups, with the isolates EACMV-TZ[Tz108] and EACMV-Tz[Tz113] grouping closely with majority EACMV group. In the phylogenetic tree, the isolates are distinct from the sequences deposited in the genebank. The isolates are closely related to EACMZV-[K18].

3.4.4 Survey for Satellites in Tanzania

Results from the country survey for satellites showed various trends which deserve emphasis. Many of the full length satellites are found associated with cassava and their helper EACMV in the major cassava growing regions, especially the Lake Zone and the Eastern zone. This observation may reflect the occurrence of mixed virus infections, which are reported to favor the occurrence of satellites. The unequal distribution of SatDNA-II and SatDNA-III over the two major growing zones may suggest that the two satellites are not easily trans-replicated by different helper viruses, suggesting also that the two zones could have different strains of the helper viruses. The multiple bands associated with the satellites, especially SatDNA-II may imply that more isolates exist, especially for SatDNA-II. The occurrence of integrated satellites sequences in cassava posses a potential huge epidemic threat since this may be a contributing factor to the breakdown of cassava cultivar resistance exacerbation. In line with this therefore, tight controls need to be put in place in regards the importation of plants and plant products.

3.4.5 Sequence Relatedness Between Satellite Fragments

Based on the constructed phylogenetic tree and distance table, the nucleotide sequences for SatDNA-II isolates are closely related despite being collected from distant geographical regions of Tanzania. Isolates SatDNA-II TZ1 and SatDNA-II TZ2 did not group together although they were collected from the same district. This may imply that the satellite isolates are widely distributed and not geographically isolated. For example, satellite isolates SatDNA-II TZ1 and SatDNA-II TZ8 were closely related despite their occurrence in distant locations and presence in different cassava cultivars. For SatDNA-III, phylogenetic tree shows that isolates which occur

in relatively adjacent regions group together. SatDNA-III TZ2 (from Mtwara) and SatDNA-III TZ3 (from Songea) grouped together.

The satellites nucleotide sequences (satDNA-II and SatDNA-III) showed the presence of features such as CAAT boxes and GC regions, indicating that they are probably similar to those described by Ndunguru *et al.* (2005).

3.5 CONCLUSION

During re-sprouting of cassava plants in a screen house, symptoms may take a long time (14 to 18 days) to develop and this may vary from one cultivar to another. In addition, symptom recovery (total or partial) in the screen house may occur frequently and should be recorded as this could prove significant with regard to experiments involving screening for virus resistance or tolerance. The recovery phenomenon may also be influenced by temperature. The distribution of satellites (SatDNA-II and SatDNA-III) appeared to be widespread across the surveyed regions consistent with the distribution of their helper cassava begomoviruses. The observation was also that the satellites had a wider occurrence and diversity in the Eastern zone compared to elsewhere in the country. The occurrence of SatDNA-III was not confined to the Lake zone as previously reported. In some instances, both SatDNA-II and SatDNA-III isolates co-existed in the same cassava plant though their effect on symptom enhancement could not be immediately established. The differential symptom phenotypes due to the presence of these satellites, may be due to interactions between host cultivar, virus, satellite and the environment, which may determine the outcome of symptoms, but further work needs to be done on this.

3.6 REFERENCES

1. Adriko J., Sserubombwe W.S., Adipala E., Bua A., Thresh J.M. and Edema R. (2011) Response of improved cassava varieties in Uganda to cassava mosaic disease (CMD) and their inherent resistance mechanisms. *Afr. J. Agric. Res.* 6 (3): 521-531.
2. Berrie L.C., Palmer K.E., Rybick E.P., and Rey M.E.C. (1998) Molecular characterization of a distinct South African cassava infecting geminivirus. *Archives of Virology* 143: 2253 –2260.
3. Briddon R.W. and Markham P.G. (1994) Universal primers for the PCR amplification of dicot-infecting geminiviruses. *Molecular biotechnology* 2: 202–205.
4. Chellappan P., Vanitharani R., Ogbe F. and Fauquet C.M. (2005) Effect of temperature on geminivirus-induced RNA silencing in plants. *Plant physiology* 138: 1828–1841.
5. Dale J. and Harding R. (2003) Strategies for the generation of virus resistant bananas. *In* H. Atkinson J., Dale R., Harding A., Kiggundu K., Kunert J.M. Muchwezi L., Sagi A. and Viljoen (eds). Genetic transformation to address the major constraints to banana and plantain production in Africa. INBAP Montpellier. Pg. 130.
6. Delp B.R., Stowell L.J. and Marois J.J. (1986) Evaluation of field sampling techniques for estimation of disease incidence. *J. Phytopathol*. 76 (12): 1299-1305.
7. Fargette, D. 1985. *Epidéimiologie de la Mosaique Africaine du Manioc en Côte d'Ivoire*. University Thesis, USTL de Montpellier. Pp. 203.
8. Fondong V.N., Pita J.S., Rey C., Beachy R.N. and Fauquet C.M. (1998) Disease Notes: First report of the presence of *East African cassava mosaic virus* in Cameroon. *Plant Disease* 82 (10): 1172.
9. Gibson R.W. and Otim-Nape G.W. (1997) Factors determining recovery and reversion in mosaic affected African cassava mosaic resistant cassava. *Ann. Appl. Biol.* 131 (2): 259-271.

10. Gruissem W (2010) The Battle Against Plant Viruses. In *Sixth World Conference on the Future of Science "Virus: the invisible enemy"* – Venice, September 19-21, 2010.

11. Hagen C, Rojas M. R., Kon T., and Gilbertson R. L. (2008) Recovery from *Cucurbit leaf crumple virus* (Family *Geminiviridae*, Genus *Begomovirus*) infection is an adaptive antiviral response associated with changes in viral small RNAs. *J. Virol*. 98 (9) : 1029 - 1037
12. Hahn S.K., Terry E.R., Leuschner K. (1980) Breeding cassava for resistance to cassava mosaic disease. *Euphytica* 29: 673–683.
13. Hillocks R.J. (2002) Cassava in Africa. *In* Hillocks R.J., Thresh J.M. and Bellotti A.C. (eds). Cassava: Biology, Production and Utilization. CABI Publishing Oxon UK.
14. Kemp B.P., Beeching J.R. and Cooper R.M. (2005) cDNA-AFLP reveals genes differentially expressed during the hypersensitive response of cassava. *Molecular Plant Pathology* 6 (2): 113–123.
15. Madden L.V. and Hughes G. (1995) Plant disease incidence: Distribution, heterogeneity and temporal analysis. *Annu. Rev. Phytopathol*. 33: 529-564.
16. Mtunda K.J., Muhanna M., Raya M.D., Kanju E.E. (2002) Current status of cassava brown streak virus disease in Tanzania. *In*: Legg J.P. and Hillock R.J. (eds). Cassava brown streak virus disease – Past Present and Future. Proceedings of an international Workshop, Mombasa Kenya 27-30 October 2002.
17. Nelson R.S. and Citovsky V. (2005) Plant Viruses. Invaders of cells and pirates of cellular pathways. *Plant Physiol*. 138: 1809–1814.
18. Ndunguru J. (2004) Characterization of cassava mosaic geminiviruses from Tanzania. PhD Thesis University of Pretoria, South Africa.
19. Ndunguru J., Legg J.P., Ismael B.F., Aveling T.A.S., Thompson G., and Fauquet C. (2008) Two novel satellite DNA molecules associated with bipartite cassava mosaic begomoviruses enhancing symptoms enhancing symptoms and capable of breaking high virus resistance in a cassava land race. *In* Cassava: Meeting the challenges of the new millennium. First scientific meeting of the global cassava partnership GCP-I, Ghent Belgium pg 141.

20. Owor B., Legg J.P., Okao-Okuja G., Obonyo R., Kyamanywa S. and Ogenga-Latigo M.W. (2004) Field studies of cross protection with cassava mosaic geminiviruses in Uganda. *J. Phytopathology* 152: 243-249.
21. Patil B.L. and Fauquet C.M. (2009) Cassava Mosaic Geminiviruses: Actual Knowledge and Perspectives. *Molecular Plant Pathology* 10 (5): 685-701.
22. Reilly K., Gomez-Vasquez R., Buschmann H., Tohme J. and Beeching J.R. (2004) Oxidative stress responses during cassava post harvest deterioration. *Plant Molecular Biology* 56: 625–641.
23. Pita J.S., Fondong V.N., Sangare A., Otim-Nape G.W., Ogwal S. and Fauquet C.M. (2001) Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidermic of cassava mosaic disease in Uganda. *Journal of General Virology* 82: 655–665.
24. Prins M., Laimer M., Noris E., Schubert J., Wassenegger M. and Tepfer M. (2008) Strategies for antiviral resistance in transgenic plants: A review. *Molecular Plant Pathology* 9 (1): 73–83.
25. Sseruwagi P., Sserubombwe W.S., Legg J.P., Ndunguru J., and Thresh J.M. (2003) Methods of surveying the incidence and severity of cassava mosaic disease and whitefly vector populations on cassava in Africa: a review. *Virus Research* 100 (1): 129–142.
26. Slater A., Scott N. and Fowler M. (2005) Plant Biotechnology – The genetic manipulation of plants. Oxford University Press, Oxford Pp. 6, 15.
27. Sseruwagi P., Otim-Nape G.W., Osiru D.S.O. and Thresh J.M. (2003). Influence of NPK fertiliser on populations of the whitefly vector and incidence of cassava mosaic virus disease. *African Crop Science Journal* 11 (3): 171-179.
28. Voet D. and Voet J.G. (2004) Biochemistry. 3rd Edition. John Wiley & Sons, Inc., USA. Pp. 291, 1353–1355.
29. Westby A., Lankford B.A, Coulter J.F., Ochard J.E., and Morton J.F. (2004) Rural infrastructure to contribute to African development: The cases of irrigation and post harvest. Natural resource Institute and Overseas development Group.UK.

CHAPTER 4

SCREENING OF CASSAVA CULTIVARS FOR DISEASE RESPONSE TO BEGOMOVIRUS AND SATELLITE DNA MOLECULES

ABSTRACT

Seventeen cassava lines bred for resistance to CMD were evaluated for resistance to satellite SatDNA-II and SatDNA-III, with cultivar T200 used as susceptible control. Cassava stem cuttings were planted in the greenhouse and maintained at temperature range 15°C - 35°C. Head-to-tail dimers of DNA-A and monomers of DNA-B of *East African cassava mosaic virus Tanzania* (EACMV-TZ) and *East African cassava mosaic virus-Uganda* variant (EACMV-UG) and infectious clones of SatDNA-II and SatDNA-III were used to bombard greenhouse plants. The infectious clones were diluted to 1μg/μL and used to prepare six plasmid mixes [EACMV-TZ + SatDNA-II, EACMV-TZ + SatDNA-III, EACMV-TZ (control), EACMV-UG + SatDNA-II, EACMV-UG + SatDNA-III and EACMV-UG (control)], denoted respectively as EII, EIII, E0, UII, UIII and U0. The plasmids were coated onto 1.0 μm gold particles at a ratio of 25 mg gold to 50 μg DNA to make microcarriers adequate for 50 shots, giving microcarrier loading quantity (MLQ) of 0.5 mg/shot and DNA loading ratio (DLR) of 2μg/mg gold. The microcarriers were used to coat an inner side of Gold-Coat^{TM} tubing and the tube cut into 0.5in-length microprojectiles. The microprojectiles were fitted into Helios Gene Gun and bombarded into apices of the greenhouse plants applying 3 shots per plant at Helium pressure of 240 *psi* and at a distance 4.5 cm. The biolistically-inoculated plants were kept in the greenhouse within a temperature range of 15°C (night) – 35°C (day) and disease symptoms scored for 35 days. Total nucleic acid extractions and PCR were performed to amplify *AC1*, *BC1*, SatDNA-II and SatDNA-III. Integrated forms of SatDNA-II and SatDNA-III were also amplified using appropriate primers. Southern blot analysis was performed using DIG-labeled probes to determine the presence of *AC1*, *BC1*, SatDNA-II and SatDNA-III in the biolistically inoculated plants. The inoculated plants suffered mechanical damage due

pressure of the gun. CMD symptoms were observed as early as 5 days post inoculation (dpi) mostly on plants inoculated with either SatDNA-II or SatDNA-III. ANOVA ($P \leq 0.05$) for mean scores for temporal disease development did show significant differences among inoculated cultivars. While the scores showed significant differences at 0.05 probability level among stages of temporal observation, the plasmids containing different viruses and satellites were highly significantly different ($P = 0.000$). *BC1* gene was amplified from 12 out of 61 symptomatic and non symptomatic leaf samples, giving PCR product of size range 500 bp - 700 bp. *AC1* gene was amplified in 2 out of 54 DNA samples tested. PCR amplification of SatDNA-II revealed the presence of multiple bands, including closed spaced bands. The amplification of full length SatDNA-III was unsuccessful, even at 42 dpi. Integrated forms of the satellites were amplified from DNA samples irrespective of whether SatDNA-II or SatDNA-III was inoculated. Southern blots for SatDNA-II and helper EACMV-TZ and EACMV-UG gave positive hybridization signals in 7 out of 15 total DNA samples tested, with strongest signals recorded on plants bombarded with EII. Blots for SatDNA-III and helper EACMV-TZ and EACMV-UG gave positive hybridization signals in 10 out of 12 total DNA samples tested. Strong signals for *AC1* gene was detected in plants inoculated with plasmids containing satellites, i.e. either EII or EIII. Based on symptom score, PCR amplification and Southern blots, and consistent with the purpose of evaluation for resistance, highly resistant cassava lines were AR37-92, CR27-24, and AR16-3 while moderately resistant lines were Bwana Mrefu AR37-96, AR40-10, AR37-6 and AR17-5. The susceptible cultivar was T200.

4.1 INTRODUCTION

Cassava (*Manihot esculenta* Crantz) plays an important role as a food staple as well as income-generating crop in Sub-Saharan Africa (SSA) countries. Statistics from seven SSA countries – Tanzania, Malawi, Kenya, Mozambique, Rwanda, Zambia and Uganda (FAO, 2008) show that cassava ranks high among other livelihood alternatives in the provision of income.

The most important disease affecting cassava production in Africa is Cassava mosaic disease (CMD), caused by several whitefly-transmitted begomoviruses that lead to yield reductions of up to 90% (Sseruwagi *et al*., 2003). Measures to control the disease are considered to fall into three broad categories (Thresh *et al*., 1998; Fraser, 2006; Mallowa *et al*., 2006): phytosanitation; use of virus-free seeds and vegetative propagules; or use of resistant varieties, which includes screening for and characterization of virus resistances levels and durability in cultivars. The most favourable way to reduce begomovirus damage is to breed cassava varieties resistant or tolerant to the devastating disease, either by classical breeding or by genetic engineering.

Unfortunately, the resistance developed in cassava cultivars so far has never been durable. The begomoviruses have been constantly evolving through recombination or in response to changing host environment, producing more virulent species/strains. Furthermore the present occurrence of satellite sequences has critically compromised efforts to breed for CMD-resistant cassava cultivars. Efforts to breed for resistance are further hampered by the reliance on CMD screening, under natural infection conditions and accurate virus identity at a given time and location. Furthermore, the wide diversity of begomoviruses and occurrence of mixed infections makes resistance screening difficult. Symptoms of CMD can be suppressed at temperatures above 35°C, and symptom severity can vary under other environmental conditions, such as altitude.

So, while infection of cassava by whiteflies or cuttings represents the natural situation in the field, biolistic inoculation is a useful method for cassava inoculation to evaluate resistance in establishing resistance or susceptibility under controlled conditions. In

this study, biolistic inoculation of seventeen farmer preferred cassava varieties and improved varieties was performed in order to evaluate resistance to CMD.

4.2 MATERIALS AND METHODS

4.2.1 Screening of Inoculated Cultivars for Resistance and Susceptibility

A total of 17 cassava lines bred at International Center for Tropical Agriculture (CIAT) in Cali, Colombia for resistance to CMD (11 AR series and five CR series lines) and 1 local variety (Table 4.1), were taken from a cassava experimental block in Bagamoyo Tanzania ($06^{o}33.306'S$ $38^{o}55.047'E$, 45m) for resistance evaluation. A South African landrace T200 susceptible to CMD (Source: Cassava Biotechnology Laboratory, University of the Witwatersrand, Johannesburg, South Africa) was included as control.

Table 4.1: Cassava lines used in CMD resistance evaluation

S/N	Name	Description	Origin
1	Bwana Mrefu	Landrace resistant to CMD	Tanzania
2	AR30-3	CMD and green mite resistant	CIAT
3	AR37-96	CMD and green mite resistant	CIAT
4	AR21-2	CMD and green mite resistant	CIAT
5	AR37-1	CMD and green mite resistant	CIAT
6	AR40-10	CMD and green mite resistant	CIAT
7	AR37-92	CMD and green mite resistant	CIAT
8	CR44-6	CMD resistant	CIAT
9	CR27-24	CMD resistant	CIAT
10	CR45-3	CMD resistant	CIAT
11	AR37-6	CMD and green mite resistant	CIAT
12	CR25-4	CMD resistant	CIAT
13	AR16-3	CMD and green mite resistant	CIAT
14	AR17-5	CMD and green mite resistant	CIAT
15	AR37-73	CMD and green mite resistant	CIAT
16	AR14-2	CMD and green mite resistant	CIAT
17	CR45-9	CMD resistant	CIAT
18	T200	Landrace susceptible to CMD	South Africa

CR series = cassava lines obtained after crossing TME 3 (containing *CMD2* resistance gene) to elite Nigerian cassava parents. AR series = cassava lines combining CMD and green mite resistance (Okogbenin *et al*., 2007).

The cuttings were sprayed with Karate (Dimethoate 40 EC), at a rate of 30mL/20L to kill incidental insects prior to transfer to South Africa (SA) (consistent with recommendations by the Directorate of Plant Health in SA). Four- to five-node stem cuttings of the cassava lines were planted in plastic pots filled with potting soil (Cultera Ltd., Muldersdrift) in a quarantine greenhouse at the University of the Witwatersrand ($26^{o}10.866$'S $28^{o}02.199$'E, 27m) in Johannesburg in April 2008. The greenhouse was fitted with heaters and wet walls for temperature control. The plants were maintained in the green house at a temperature range 15 - 35°C (Appendix 4.1). Improvised yellow sticky cards (Ndomba, 2005) were also put in the greenhouse to monitor the occurrence of whiteflies and aphids. While growing, the plants were

monitored for CMD symptoms and suitable growth stage for particle bombardment as indicated in Appendix 4.2.

4.2.2 Reagents and Infectious Clones

Infectious clones, used in the resistance study were kindly provided by Dr. Joseph Ndunguru of Mikocheni Agricultural Research Institute, Dar es Salaam. The clones were head-to-tail dimers of DNA-A and monomers DNA-B of *East African cassava mosaic virus Tanzania* (EACMV-TZ) and *East African cassava mosaic virus-Uganda* variant (EACMV-UG). Two infectious clones (dimmers) of SatDNA-II and SatDNA-III were cloned into *E. coli*. Concentration of plasmid DNA in the infectious clones were determined by a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Worcester USA). The plasmids were transformed into competent *Escherichia coli* 5Hα strain (kindly provided by Imanu Kahozi of the University of the Witwatersrand). Fifty nanograms of plasmid DNA was used (volume varying from 1μL to 5μL depending on original concentration of the plasmids) for transformation (Protocol presented as Appendix 4.3). The transformed bacterial suspensions were grown on LB Agar medium supplemented with ampicillin at 1μl/mL at 37°C for 16 hours. The cells were cultured on LB broth with ampicillin 1μL/mL at 37°C with shaking at 170 rpm for 16 hours. Plasmid extraction was done by Alkaline Lysis Miniprep (Sambrook *et al*., 1989) (Appendix 4.4). The DNA concentration was determined using a NanoDrop spectrophotometer. Ten microlitre with a concentration of 100ng/μL was prepared from each plasmid DNA sample and used in gel electrophoresis (1% agarose in 1x TAE buffer) to verify integrity of the plasmids.

4.2.3 Preparation of Plasmids

All the plasmid samples were diluted to 1μg/μL, consistent with recommendations by manufacturers of the biolistic device (Helios Gene Gun, Biorad). There were six plasmid mixes as follows: EACMV-TZ + SatDNA-II, EACMV-TZ + SatDNA-III, EACMV-TZ (control), EACMV-UG + SatDNA-II, EACMV-UG + SatDNA-III and EACMV-UG (control). These mixes were denoted for simplicity respectively as EII, EIII, E0, UII, UIII, and U0. In forming the plasmid mixes, DNA components A and B

were prepared in a ratio of 1:1 (w/w) for each begomovirus (Ariyo *et al.*, 2006) and kept in separate microfuge tubes. The ratio between begomoviruses and satellites in the mixes were also 1:1 (w/w). Details on these formulations are shown in Appendix 4.5.

4.2.4 Preparation of Gold Particles and Biolistic Inoculation

A stock solution of 20mg/mL polyvinylpyrrolidone (PVP) in ethanol was prepared and diluted in ethanol to a working concentration of 0.05mg/mL. Twenty five milligrams of 1.0 μm gold particles were weighed into 1.5 mL microfuge tube and mixed with 100 μL of 0.05M spermidine according to the procedure indicated below. The gold was deemed to be adequate for 50 shots giving a microcarrier loading quantity (MLQ) of 0.5mg/shot. Fifty micrograms of the plasmid DNAs (infectious viral clones) was added to the gold and spermidine mixture, according to requirement of the plasmid mix as explained above, to give a DNA loading ratio (DLR) of 2μg/mg gold. The microcarrier coating process was completed by the addition of 100 mL of 1M $CaCl_2$. Full particle bombardment protocol is shown in Appendix 4.6.

Seventy five centimeters of Gold-Coat^{TM} tubing (Bio-Rad, Hercules) was cut and washed with absolute ethanol. The tubing was then inserted into Tubing Prep Station (Bio-Rad, Hercules) and dried with N_2 at 50kPa for 15 minutes. The microcarrier/DNA suspension was vortexed and drawn into the tubing using syringe, avoiding bubbles. The tube was immediately brought to horizontal position, and with syringe attached and fitted to the support cylinder. The microcarrier suspension was allowed to settle for 5 minutes and then ethanol removed using the syringe. The syringe was detached and the tube immediately rotated through 180° while in the groove and the gold was allowed to coat the inside surface of the tubing for 3-4 seconds. The coated tube was cut into 0.5 inch-length using tubing cutter (Ariyo *et al*., 2006) and according to instructions from manufacturers of the biolistic device.

The tubing was washed with fresh absolute ethanol using a syringe to suck in the ethanol from one of its ends. This was then inserted into the Tubing Prep Station and dried for 15 minutes with N_2 connected to the Tubing Prep Station. The DNA microcarrier solution was vortexed and immediately drawn into the Gold-Coat Tubing

with aid of the syringe. The tube was then immediately rotated in the Tubing Prep Station to evenly coat the DNA/microcarrier solution onto the inner wall of the tubing. The remaining particles in solution were drawn out slowly with the syringe followed by immediate rotation of the Gold-Coat tubing for a few seconds. The tubing was dried with Nitrogen by opening the gas tap on one end of the rotor machine containing the tubing. The tubing was then cut into 0.5 in-length (Ariyo *et al*., 2006) cartridges using the tubing cutter (Bio-Rad, Hercules). Inserted into the cartridge holder (Figure 4.1) of the Helios Gene Gun (Bio-Rad, Hercules NC USA), these cartridges were the projectiles carrying the DNA into target cells by helium discharge.



Figure 4.1: Loading cartridges into cartridge holder. Prepared cartridges were stored at 10°C until time of biolistic inoculation.

Biolistic inoculation was carried out on plants which survived the green house conditions (Appendix 4.2) using the Helios Gene Gun; applying 3 shots per plant onto the sprouted stem cuttings at three- to five-leaf stage (20 cm tall). Helium pressure was set at 240*psi* (1654.7kPa) (Ariyo *et al*., 2003). The plant pots were previously covered (Figure 4.2A) to prevent soil splash in the laboratory on application of the high pressure. The plants were also prepared by removing aged leaves.

The plants were bombarded at the apices (Figure 4.2B), including immature leaves and apical young stem and meristem (Zhang *et al.*, 2005), by applying 3 shots per plant at a distance of about 4.5 cm (Schopke *et al*., 1993; Ariyo *et al*., 2003) when the shoot was 5 – 10 cm tall and with about five leaves.



Figure 4.2: A = plant pot covered to prevent soil splash in the laboratory. B = Bombarding plants from apices. The bombardment aimed to hit the leaf blades and internodes.

As virus-only controls, some sets of plants were bombarded with EACMV-TZ and EACMV-UG without SatDNA-II and SatDNA-III, and a known highly CMD susceptible variety, South African landrace T200, was included as a positive susceptible control. After particle bombardment, the biolistically-inoculated plants parts were marked with a blue tag, so as to identify leaves from which to score for disease symptoms. The plants were returned to the green house and kept within temperature range of 15 – 35°C until symptoms developed. The symptom development process was recorded for 35 days.

4.2.5 Disease Severity Scores

The biolistically inoculated plants were examined for their reaction to CMD by assigning Mean Disease Severity Scores (MDSS) on a standard 1 - 5 point scoring scale for CMD (Hahn *et al.* 1980; Raji *et al*., 2007) on a weekly basis for 35 days. On the scale, score 1 was used for unaffected shoots with no obvious CMD symptoms; score 2 for mild chlorosis, mild distortions at bases of most leaves, while the remaining parts of the leaves and leaflets appear green and healthy; score 3 for pronounced mosaic pattern on most leaves, narrowing and distortion of the lower one third of the leaflets; score 4 for severe mosaic, distortion of two thirds of most leaves

and general reduction of leaf size and stunting of shoots; score 5 for very severe mosaic symptoms on all leaves, distortion, twisting, misshapen and severe leaf reduction on most leaves accompanied by severe stunting of plants. Since the symptoms could not be uniform on the leaves, symptom score was based on the whole plant as done by Ogbe *et al* (2003). The observed MDSS were analyzed using SPSS 16.0 statistical program (IBM SPSS Statistics, Chicago) to produce and compare means for the tested variables.

4.2.6 Nucleic Acid Extraction

Total nucleic acid (TNA) extraction was done by the CTAB method (Appendix 4.7) 35 days post inoculation (dpi). PCR was performed to amplify helper begomovirus full-length *AC1* (of EACMV-TZ), *BC1* using primer pairs EAC1-F/R, EAB555-F/R, and SatDNA-II and SatDNA-III using primer pairs Beta01/Beta02 and DNA-1/DNA-2, respectively (Table 4.2). During the course of this study, integrated satellite sequences were discovered in cassava genome of expressed sequence tags (ESTs) (Ndunguru *et al*., unpublished). So, additional primers designed from cassava ESTs were also used to amplify integrated sequences of the satellites.

Table 4.2: List of primers used to detect viral and satellite sequences

Name	**Sequence (5' → 3')**	**Target Sequence**	**Band Size (*bp*)**
EAC1F	GGA ATT CCA TAT GCC GAG AGC CGG TCGT TTT CAA ATA	*AC1*/Rep	1000 - 1077
EAC1R	CCG CCT CGA GGC TTG CCT GTG CCT GGC T		
EAB555F	TAC ATC GGC CTT TGA GTC GCA TGG	*BC1*/CR	550 - 560
EAB555R	CTT ATT AAC GCC TAT ATA AAC ACC	*BC1*/CR	
Beta01	GGT ACC ACT ACG CAG CAG CC	SatDNA-II	1009
Beta02	GGT ACC TAC CCT CCC AGG GGT ACA C		
DNA-1	TGG GGA TCC TAG GAT ATA AAT AAC ACG TC	SatDNA-III	1032
DNA-2	CTA GGA TCC GGA CAA ATT ACA AGC GTA		
SatII F	GCC GCA CCA CTG GAT CTC	Integrated SatII	895
SatII R	CAG CAG CCA GTC AGG AAG TT		
SatIII F	AGG CCT CGT TAC TAA AAG TGC	Integrated SatIII	306
SatIII R	ACCTGACGGCAGAAGGAAT		

4.2.7 Southern Blot Analyses

The concentration of total DNA extracted from leaf tissue of biolistic-inoculated cassava plants were first measured using a Nanodrop Spectrophotometer and then diluted to 500ng/μL. The DNAs were then fractionated by electrophoresis in 1% (w/v) agarose gels in TAE buffer and transferred to nylon membrane (Roche Diagnostics, Mannheim, Germany) by capillarity following alkali denaturation and neutralization as described in Appendix 4.8. Cross-linking was done in a UV Cross-linker at 120 mJ for 5 minutes. DNA fragments representing SatDNA-II, SatDNA-III, *AC1* and *BC1* - produced by PCR were labeled by a steroid hapten DIG, a chromogenic dye containing random primed DNA digoxigenin-dUTP (according to manufacturer's instructions) and used in hybridization separately as described in Appendix 4.8. After prehybridization and hybridization, the blots were washed with 2x Standard Saline Citrate (SSC) and 0.5x SSC [1x SSC is 0.15M NaCl plus 0.015 M sodium citrate]. Hybridization signals were detected colorimetrically by the addition of 200 μL of NBT/BCIP stock solution to freshly prepared detection buffer.

4.3 RESULTS

4.3.1 Growing of Cassava Cuttings

The growth of the plants, grown from field cuttings, was generally slow due to low night winter temperatures of 15°C in the green house (Appendix 4.9), despite use of heaters to supplement heat, while some plants did not grow at all in the green house. Plants showing typical CMD symptoms as indicated in Appendix 4.2, were excluded from the study, as symptom-free plants were selected for biolistic inoculation. All stakes from a cultivar AR37-73 did not sprout in the green house. In cultivar CR45-9, a few cuttings sprouted but later died. Some plants from variety CR44-6 had typical CMD symptoms and were also excluded from the study. No whiteflies were found on the yellow traps. However, fungus gnats, *Exechia nugatori*a (*Diptera: Sciaridea*) were found in the green house. They were sprayed with VectoBac WG (*Bacillus*

thuringiensis subsp. *israelensis*, 3000 IUT/mg, Valent Biosciences USA) at a rate of 3g/L.

4.3.2 Particle Bombardment

In order to standardize the inocula, the concentration and purity of the plasmids were determined. Table 4.3 shows the concentration of various inocula used in the experiment.

Table 4.3: Concentration and purity of plasmids before transformation into *E. coli*

Sample	**ng/μL**	**A_{260}**	**A_{230}**	**A_{260}/A_{280}**	**A_{260}/A_{230}**
SatDNA-II	48.54	0.971	0.516	1.88	2.43
SatDNA-III	51.49	1.03	0.898	1.15	1.53
EACMV-UG DNA-A	19.38	0.388	0.21	1.8	2.33
EACMV-UG DNA-B	12.9	0.258	0.143	1.8	2.19
EACMV-TZ DNA-A	15.97	0.319	0.171	1.86	2.23
EACMV-TZ DNA-B	10.4	0.208	0.106	1.96	2.67

Since the concentration of the plasmids was far too low to allow biolistic inoculation, they were therefore transformed into *E.coli* DHα for amplification. After transformation, the concentrations of the extracted plasmid were determined again (Table 4.4).

Table 4.4: Concentration and purity of plasmids after multiplication

Sample	ng/μL	A_{260}	A_{280}	A_{260}/A_{280}	A_{260}/A_{230}
SatDNA-II	5248.8	104.9	96.2	1.09	1.22
SatDNA-III	5203.1	104.6	91.1	1.14	1.29
EACMV-UG DNA-A	5206.4	104.1	89.1	1.17	1.31
EACMV-UG DNA-B	5196.8	103.9	91.8	1.13	1.26
EACMV-TZ DNA-A	5196.0	103.9	80.7	1.29	1.46
EACMV-TZ DNA-B	5152.7	103.0	91.0	1.13	1.23

The concentrations of plasmids DNAs were then suitable for bombardment and the purity, based on the A_{260}/A_{280} values were within acceptable values. The values (in Table 4.4) were the concentrations of plasmids which were used to prepare plasmid mixes for inoculation. Nine varieties were bombarded with the six mixes of infectious clones – E0, U0, EII, EIII, UII, and UIII.

Following the bombardment process, some plants had their tips mechanically damaged as a result of intense pressure from the gene gun as shown in Figure 4.3



Figure 4.3: Mechanical damage on growing tips in cassava line AR40-10 after bombardment with EACMV-UG + SatDNA-II.

The mechanical damage also sometimes resulted in the shattering of leaves. The most affected were those leaves from cv. T200, which had been propagated through tissue culture. Therefore, the pressure of biolistic inoculation was adjusted to a lower value of 240 *psi*.

4.3.3 Symptoms of Biolistically Infected Plants

CMD symptoms observed on the bombarded plants with satellites included leaf distortion and curling (Figure 4.4). The symptom compares well with typical CMD symptoms involving satellite infection (Figure 4.5) as observed in the field in Dar es Salaam and proved by PCR to contain SatDNA-II. This observation was not as severe though, because it did not persist for an extended period of time.



Figure 4.4: Leaf curling and distortion from cassava plant line AR37-92, 7 dpi with EACMV-UG + SatDNA-II. The symptom is typical of CMD symptoms involving satellite infection.



Figure 4.5: Naturally-occurring EACMV-TZ + SatDNA-II-infected plant collected from a field at Yombo Vituka in Dar es Salaam Tanzania.

Four plants exhibited these CMD symptoms as early as 5 days after biolistic inoculation (Figure 4.6 A-D), most of them inoculated with plasmid mixes involving either SatDNA-II or SatDNA-III. The landrace T200 also showed CMD symptoms.

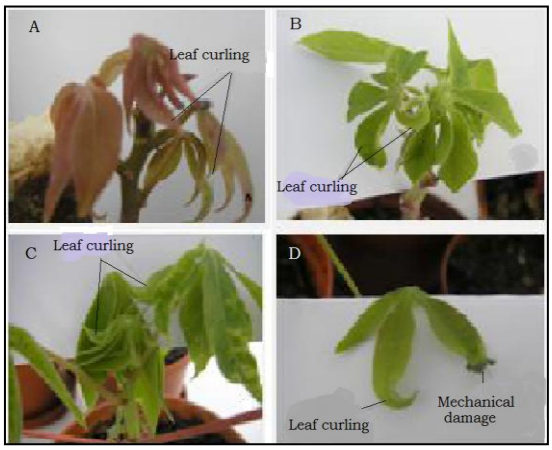

Figure 4.6 A - D: A = CMD symptoms on cassava line Bwana Mrefu, 5dpi with EACMV-TZ + SatDNA-II; B = CMD symptoms on cassava line AR37-92, 5dpi with EACMV-TZ + SatDNA-III; C = CMD symptoms on cassava line AR37-92, 5dpi with EACMV-UG + SatDNA-II; D = CMD symptoms on susceptible cassava cv. T200, 5dpi with EACMV-TZ alone. In all these cases, leaf curling is evident.

Other typical CMD symptoms observed on the bombarded plants were green mosaic (Figure 4.7), but only a few plants displayed this symptom.



Figure 4.7: Mosaic symptoms observed on cassava plant line AR21-2, one week after bombardment with EACMV-TZ + SatDNA-II. Only a few plants displayed this symptom.

At 10 days' post infection (10 dpi), most plants recovered from the infection as in Figure 4.8 below.



Figure 4.8: Cassava plant from line AR21-2 recovering from infection due to bombardment with EACMV-TZ + SatDNA-III.

A summary of the symptoms observed over 5 weeks, following virus or virus and satellite biolistic inoculations are shown in Table 4.5.

Table 4.5: Symptom observations following biolistic inoculations

Virus/satellite Mix	Symptom description			Mean Disease Severity Score		
	Early Symptoms (5 - 7dpi)	Mid Symptoms (8 - 21dpi)	Late Symptoms (22 - 35dpi)	5 - 7dpi	8 - 21dpi	22 - 35dpi
EACMV-TZ (E0)						
Bwana Mrefu	Mild leaf distortion	Leaf distortion	Recovery	2	2.5	2
AR37-96	Healthy	Mild leaf distortion	Mild leaf distortion	1	2	2
AR21-2	Dead	Dead	Dead	0	0	0
AR40-10	Healthy	Tip death	Tip regrow	1	2	2
AR37-92	Dead	Dead	Dead	0	0	0
CR27-24	Healthy	Tip death	Tip sprouting	1	1	1
AR37-6	Healthy	Mild leaf distortion	Mild leaf distortion	1	2	2
AR16-3	Healthy	Tip death	Tip sprouting	1	2	2
AR17-5	Healthy	Dead	Dead	1	0	0
T200	Tip damage	Tip regrow	Tip regrow and stunting	1	2	2

EACMV-TZ+ SatII (EII)

Bwana Mrefu	Mild leaf distortion	Leaf distortion	Recovery	2	3	2
AR37-96	Plant wilting	Plant regrowth	Recovery	2	2	1
AR21-2	Mild leaf distortion	Leaf distortion	Recovery	2	2.5	2
AR40-10	Mild leaf distortion	Tip death	No regrowth	2	2	1
AR37-92	Healthy	Tip death	Tip sprouting	1	1	1
CR27-24	Healthy	Tip death	No regrowth	1	1	1
AR37-6	Mild leaf distortion	Leaf distortion	Recovery	2	3	2
AR16-3	Healthy	Tip death	Recovery	1	2	1
AR17-5	Healthy	Leaf distortion	Recovery	1	2.5	2

EACMV-TZ+ SatIII (EIII)

Bwana Mrefu	Healthy	Mild leaf distortion	Healthy	1	1.5	1
AR37-96	Healthy	Tip death	Plant dead	1	1	0
AR21-2	Mild leaf distortion	Mild leaf distortion	Recovery	2	2	2
AR40-10	Mild leaf distortion	Severe leaf distortion	Recovery	2	3.5	3.5
AR37-92	Mild leaf distortion	Mild leaf distortion	Recovery	2	2	2
CR27-24	Healthy	Tip death	Tip sprouting	1	2	1
AR37-6	Tip damage	Tip regrowth	Recovery	2	2	1
AR16-3	Tip damage	Mild leaf distortion	Recovery	2	2	1
AR17-5	Plant wilting	Dead	Dead	2	0	0

EACMV-UG (U0)

Bwana Mrefu	Mild leaf distortion	Leaf distortion	Recovery	2	3	2.5
AR37-96	Plant dead	Plant dead	Plant dead	0	0	0
AR21-2	Healthy	Plant dead	Plant dead	1	0	0
AR40-10	Plant dead	Plant dead	Plant dead	0	0	0
AR37-92	Plant dead	Plant dead	Plant dead	0	0	0
CR27-24	Healthy	Tip death	Tip death	1	2	1
AR37-6	Healthy	Leaf distortion	Mild leaf distortion and mosaic	1	3	2
AR16-3	Plant dead	Plant dead	Plant dead	0	0	0
AR17-5	Mild leaf distortion	Leaf distortion	Recovery	2	2.5	1

18U0	Healthy	Mild leaf distortion	Mild leaf distortion	1	2	1
EACMV-UG+SatII (UII)						
Bwana Mrefu	Mild leaf distortion	Mild leaf distortion	Recovery	2	2	1
AR37-96	Mild leaf distortion	Leaf distortion	Recovery	2	3	2
AR21-2	Tip damage	Healthy	Healthy	1	1	1
AR40-10	Mild leaf distortion	Tip death	Tip death	2	1	1
AR37-92	Leaf distortion	Leaf distortion	Recovery	3	3	2
CR27-24	Healthy	Tip weak	Tip death	1	1	1
AR37-6	Mild leaf distortion	Plant weak	Plant dead	2	2	0
AR16-3	Tip damage	Tip regrow	Tip regrow	2	1	1
AR17-5	Healthy	Mild leaf distortion	Recovery	1	2	2
EACMV-UG+SatIII (UIII)						
Bwana Mrefu	Mild leaf distortion	Leaf distortion	Recovery	2	3	2
AR37-96	Healthy	Mild leaf distortion	Recovery	1	2	2
AR21-2	Mild leaf distortion	Mild leaf distortion	Recovery	2	2	1
AR40-10	Mild leaf distortion	Mild leaf distortion	Tip death	2	2	1
AR37-92	Mild leaf distortion	Mild leaf distortion	Leaf distortion and mosaic	2	2	3
CR27-24	Tip damage	Tip death	Dieback	1	2	2
AR37-6	Plant wilting	Plant dead	Plant dead	2	1	1
AR16-3	Mild leaf distortion	Mild leaf distortion	Recovery	2	2	1
AR17-5	Leaf distortion	Leaf distortion	Recovery	2	2.5	2.5

Many of the plants suffered some mechanical damage following bombardment but they subsequently showed a marked tip re-growth. A single plant in cultivar AR40-10, bombarded with EIII, was the only plant that reached a score 4 of CMD severity (severe CMD: leaf curling and chlorosis in 1 – 3 leaves of the plant).

Mean severity scores for disease, in the tested cultivars at 3 dpi stages of observation, and for helper virus-satellite combinations, are indicated in Table 4.6.

Table 4.6: Mean severity scores for 3 stages of observation and 6 helper virus-satellite combinations

	Stages of Observation			**Virus, Virus/satellite**					
Label	1	2	3	1	2	3	4	5	6
Description	5-7dpi	8-21dpi	22-35dpi	E0	EII	EIII	U0	UII	UIII
Score	1.41	1.69	1.27	1.13	1.70	1.57	0.89	1.59	1.85

E0, EII, EIII, U0, UII and UIII refer to the plasmid mixes as described in Table 4.5.

ANOVA at $P \leq .05$ level showed that the tested cultivars were not significantly different (Table 4.7).

Table 4.7: ANOVA table for the observed variables ($P \leq .05$)

Variable	**Levels**	**df**	**TrSS**	**EMS**	**F-Value**	**Std Error**	**P-Level**
Cultivars	9	8	10.670	1.334	1.822	0.28523	0.077ns
Plasmids	6	5	18.327	3.665	5.479	0.22262	0.000**
Stages	3	2	5.096	2.548	3.445	0.16551	0.034*

df = degree of freedom, TrSS = Treatment sum of squares, EMS = Error mean square, Std Error = standard error, ns = non significant, * = significant, ** = highly significant.

While in terms of CMD severity, the three stages of temporal observation of the development of disease were significantly different at 0.05 probability level, the plasmids containing different viruses and satellites were highly significantly different ($P = 0.000$). Since variable levels stages of temporal observation and plasmid mixes were not many (i.e. less than 10), mean separation was performed using least significant differences (LSD). The output for mean separation is presented in Table 4.8.

Table 4.8: Mean separation for disease development stages and virus, virus/satellite mixes (Post Hoc = LSD)

I	J	Mean Difference (I – J)	Significance
Stage	**Stage**		
5 – 7dpi	8 – 21dpi	-0.28704	0.085ns
	22 – 35dpi	0.13889	0.403ns
8 – 21dpi	22 - 35dpi	0.42593	0.011**
Plasmid Mix	**Plasmid Mix**		
E0	EII	-0.57407	0.011**
	EIII	-0.44444	0.048*
	U0	0.24074	0.281ns
	UII	0.46296	0.039*
	UIII	-0.22222	0.001**
EII	EIII	0.12963	0.561ns
	U0	0.81481	0.000**
	UII	0.11111	0.618ns
	UIII	-0.14815	0.507ns
EIII	U0	0.68519	0.002**
	UII	-0.01852	0.934ns
	UIII	-0.27778	0.214ns
U0	UII	-0.70370	0.002**
	UIII	-0.96296	0.000**
UII	UIII	-0.25926	0.246ns

ns = non significantly different, * = significantly different at $P \leq .05$, ** = significantly different at $P \leq .01$

The separation of the means showed that disease development differed remarkably between 8 – 21 dpi and 22 – 35 dpi stages. Mean separation for the plasmid mixes showed significant differences between plasmid mix 1 (E0) and plasmid mix 2 (EII), between E0 and EIII, E0 and UII, E0 and UIII, EII and U0, EIII and U0, UII and U0, and UIII and U0. The analysis presented shows that the two satellites were not significantly different between themselves.

4.3.4 PCR Amplification

The DNA concentration was determined, standardized by diluting them to 500ng/µL and used for PCR.

4.3.4.1 *BC1* gene fragment amplification of EACMV-UG

From the DNAs extracted 35 dpi, *BC1* gene was amplified from 12 out of a mixture of 61 symptomatic and non symptomatic leaf samples submitted in the laboratory, giving a PCR product of between 500 bp and 700 bp. PCR results from selected DNA samples are presented in Figure 4.9.

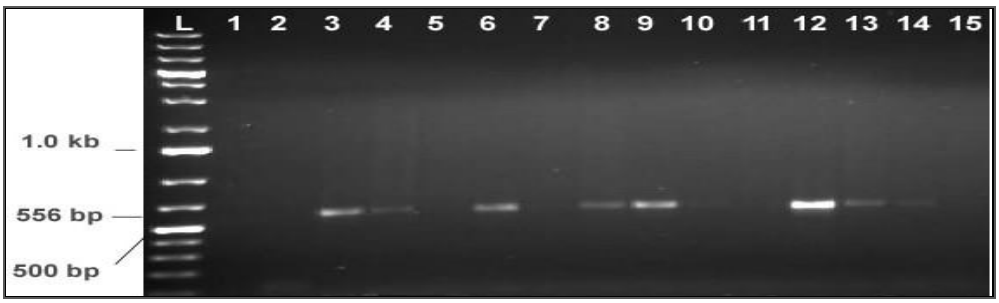

Figure 4.9: PCR amplification of *BC1* gene of EACMV-UG from experimentally biolistically-inoculated plants (expected product size is 540 – 560 bp). Lane 1 = negative control (line C98-7), Lane 2 = Bwana Mrefu inoculated with E0, 3 = AR21-2 (EII), 4 = AR17-5 (EII), 5 = AR37-92 (UII) (leaf sample picked from asymptomatic leaf), 6 = AR37-96 (UII), 8 = CR27-24 (UII), 9 = T200 (U0), 12 = AR17-5 (UII), 13 = AR40-10 (UIII), 14 = AR37-6 (EII) (leaf sample picked from symptomatic leaf), 15 = AR37-6 (EII) (leaf sample picked from asymptomatic leaf).

In Figure 4.9, with the exception of T200 inoculated with U0, all the other positive results were obtained from plants bombarded with EII, UII, and UIII. The EII-bombarded plants were of lines AR21-2 and AR17-5 while the UII-bombarded plants were AR37-96 and again AR17-5. The UIII-bombarded plant was AR40-10. As can be seen in this case, the positive amplification was obtained from plants in which either SatDNA-II or SatDNA-III was included in the inoculations.

For plants in which leaf samples were collected from both symptomatic and asymptomatic leaves (Figure 4.10), the *BC1* gene could only be amplified from symptomatic leaves in most plants.



Figure 4.10: Symptomatic leaf (left) and asymptomatic leaf (right) occurring on the same plant

4.3.4.2 *AC1* gene of EACMV-TZ

An *AC1* gene fragment was amplified to screen for the presence of EACMV-TZ. Positive amplification was only obtained in two samples out of 54 DNA samples tested. These were cv. T200 bombarded with EACMV-TZ alone and line AR40-10 bombarded with EACMV-TZ + SatDNA-III.

4.3.4.3 Amplification of full-length SatDNA-II

The amplification of full-length SatDNA-II yielded three amplified bands on the agarose gel: the expected band (1.0kb), another one at 700bp and a third between 400bp and 500bp, with the 700bp-band being particularly well-defined (Figure 4.11).

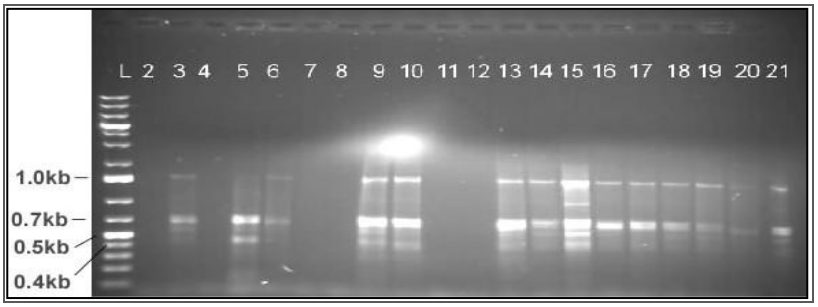

Figure 4.11: Amplification of full-length SatDNA-II from bombarded plants. Lane 2 = Bwana Mrefu inoculated with EII, 3 = AR16-3 inoculated with E0, 5 = CR27-24 (E0), 6 = AR37-6 (E0), 9 & 10 = AR17-5 (U0), 13 = AR21-2 (EII, symptomatic leaves), 14 = AR21-2 (EII, asymptomatic leaves), 15 = AR40-10 (EII), 16 = AR37-92 (EII), 17 = CR27-24 (EII), 18 = AR37-6 (EII, symptomatic leaf), 19 = AR37-6 (EII, asymptomatic leaf), 20 = AR16-3 (EII, symptomatic leaf) and 21 = AR16-3 (U0).

In Figure 4.11, multiple bands appear on the gel. Apart from the expected size of 1kb, there were other bands with size 500 bp, 700 bp and 800 bp. Most of the positive samples were from plants bombarded with EACMV-TZ + SatDNA-II. Because of appearance of multiple bands, attempt to re-amplify SatDNA-II was done. Annealing temperature was raised from 49°C to 51°C and annealing time was increased from 1minute to 1minute 30 seconds. This gave two bands, 1.0kb (expected size) and other band at 700 bp (Figure 4.12). Of the two bands, the 700 bp band was more prominent than the 1.0kb band. Fifteen out of same 19 samples had positive amplification.

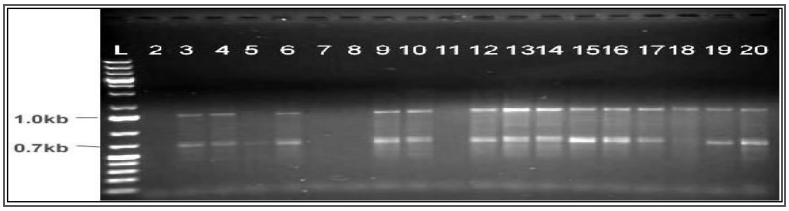

Figure 4.12: Re-amplification of SatDNA-II. There were two prominent bands, one around 700 bp and the other at 1.0kb.

PCR amplification of SatDNA-II from the DNAs extracted from the same plants after allowing them to grow for 7 more days after the 35 days still revealed the presence of multiple bands (Figure 4.13).



Figure 4.13: Amplification of SatDNA-II from total DNA 42dpi. Multiple bands were still observed on the gel. The expected band is 1.0kb. Other bands, including a pair of closely spaced (double) bands, occurred at around 700bp.

4.3.4.4 Amplification of full-length SatDNA-III

The amplification of full length SatDNA-III was unsuccessful from any of the inoculated plants, even at 42 dpi.

4.3.4.5 Amplification of integrated satellite fragments

PCR amplification was carried out on the same DNA samples, as above, using primers designed from cassava Expressed Sequence Tags (ESTs) (CIAT, Colombia) to amplify integrated satellite sequences. Positive results (895 bp) were obtained in many samples as shown in Figure 4.14. This suggested the integration of SatDNA-II in the cassava genome, even in CMD-resistant cassava cultivars which were used in this study.



Figure 4.14: Amplification of SatDNA-II fragments using EST-based primers. PCR amplified bands appear at 895 bp. Double bands occur in lanes 1 and 8. Lane L – Molecular weight marker; C: – Line AR37-6 bombarded with EACMV-TZ alone (Control);

Lane 1: AR17-5 bombarded with EACMV-TZ + SatDNA-II (i.e. plasmid mix EII); Lane 2: AR37-96 bombarded with EII; Lane 3: AR21-2 (EII); Lane 4: AR40-10 (EII); Lane 5: AR37-92 (EII); Lane 6: CR27-24 (UII); Lane 7: AR37-6 (UII); Lane 8: AR37-6 (UII) (asymptomatic leaves); Lane 9: AR16-3 (EII); Lane 10: AR37-6 (EII); Lane 11: AR17-5 (EII); Lane 12: AR37-96 (EII); Lane 13: AR21-2 (EII) (symptomatic leaves); Lane 14: AR37-92 (UIII); and Lane 16: AR37-6 (EII) (symptomatic leaves).

The integrated form was amplified from the DNA samples irrespective of whether SatDNA-II or SatDNA-III was inoculated.

The amplification of integrated SatDNA-III fragments (size 306 bp) from DNA extracted from bombarded plants gave positive results in all samples, whether SatDNA-III was included in the biolistic inoculation or not (Figure 4.15).

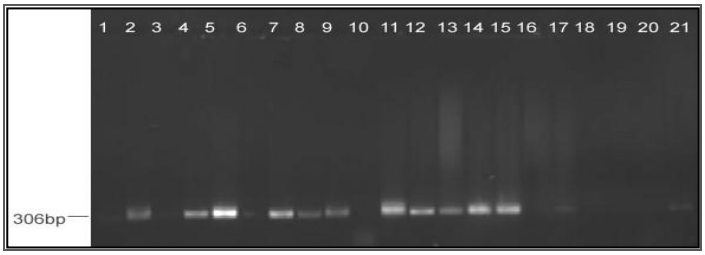

Figure 4.15: Amplification of SatDNA-III fragments. Lane 1: AR16-3 bombarded with E0; Lane 2: AR21-2 (EII); Lane 3: AR21-2 (UIII); Lane 4: AR40-10 (UIII); Lane 5: AR37-92 (UIII); Lane 6: CR27-24 (UIII); Lane 7: AR16-3 (UIII); Lane 8: AR17-5 (UIII); Lane 9: AR21-2 (EIII); Lane 10: AR40-10 (EIII), symptomatic leaves; Lane 11: AR40-10 (EIII), asymptomatic leaves; Lane 12: AR37-92 (EIII); Lane 13: Bwana Mrefu (UIII); Lane 14: AR37-96 (UIII); Lane 15AR21-2 (UIII); Lane 16: AR40-10 (UIII); Lane 17: AR37-92 (UIII); Lane 18: AR37-92 (EIII); Lane 19: AR37-6 (EIII); and Lane 20: AR16-3 (EIII).

These results suggest the integration of SatDNA-II and SatDNA-III in the cassava genome, even in CMD-resistant cassava cultivars that were used in this study. The overall summary of PCR amplification results is indicated in Table 4.9.

Table 4.9: Summary of PCR amplification results

Cultivar and virus/satellite mix in brackets	**PCR Amplification**				**EST-Primers**	
	SatDNA-II	**SatDNA-III**	***AC1***	***BC1***	**SatII**	**SatIII**
Bwana Mrefu (E0)	-	-	-	+	-	-
AR37-96 (E0)	-	+	-	-	-	-
AR40-10 (E0)	+	+	-	+	-	-
AR37-92 (E0)	-	-	+	-	-	-
CR27-24 (E0)	+	-	-	-	-	-
AR37-6 (E0)	+	+	-	+	+	-
AR16-3 (E0)	+	+	-	-	-	-
T200 (E0)	-	-	+	+	-	-
Bwana Mrefu (U0)	-	+	-	-	-	-
T200 (U0)	-	+	-	-	-	-
AR37-6 (U0)	+	+	-	-	-	-
AR17-5 (U0)	+	+	-	+	++	-
Bwana Mrefu (EIISy)	-	+	-	+	+	-
Bwana Mrefu (EIIAsy)	-	-	-	-	-	-
AR37-96 (EII)	+	+	-	+	-	-
AR21-2 (EIISy)	+	+	-	+	-	-
AR21-2 (EIIAsy)	+	-	-	-	+	-
AR40-10 (EII)	+	+	-	-	++	-
AR37-92 (EII)	+	-	-	-	-	-
CR27-24 (EII)	+	-	-	-	-	-
AR37-6 (EII)	+	+	-	-	-	-
AR16-3 (EIISy)	+	+	-	-	+	-
AR16-3 (EIIAsy)	+	-	-	-	-	-
AR17-5 (EII)	+	-	-	+	++	-
Bwana Mrefu (IISy)	+	-	-	-	+	-
Bwana Mrefu (UIIAsy)	+	-	-	-	-	-
AR37-96 (UIISy)	+	+	-	+	-	-
AR37-96 (UIIAsy)	+	-	-	-	++	-
AR21-2 (UII)	+	-	-	-	+	-
AR40-10 (UII)	+	+	-	+	-	-
AR37-92 (UIISy)	+	-	-	-	-	-

AR37-92 (UIIAsy)	+	-	-	-	+	-
AR17-5 (UII)	+	-	-	+	-	-
Bwana Mrefu (UIII)	-	+	-	+	-	+
AR37-96 (UIII)	+	-	-	-	-	-
AR21-2 (UIII)	+	-	-	-	-	+
AR40-10 (UIII)	+	-	-	+	-	-
AR37-92 (UIII)	+	+	-	+	-	+
CR27-24 (UIII)	+	+	-	+	-	+
AR16-3 (UIII)	+	-	-	-	-	+
AR17-5 (UIII)	+	-	-	-	-	+
Bwana Mrefu (EIII)	-	-	-	-	-	+
AR21-2 (EIII)	+	-	-	-	-	-
AR40-10 (EIIISy)	+	+	+	-	-	++
AR40-10 (EIIIAsy)	-	+	-	-	-	-
AR37-92 (EIII)	-	+	-	-	-	-
CR27-24 (EIII)	-	+	-	-	-	-
AR37-6 (EIII)	-	+	-	-	-	++
AR16-3 (EIII)	-	+	-	-	-	+

+ positive amplification, - Negative amplification, ++ multiple bands on the lane, Sy = sample taken from symptomatic leaf, Asy = Sample taken from asymptomatic leaf.

4.3.5 Southern Blot Analyses

Southern blots to detect the presence of SatDNA-II sequences from plants bombarded with SatDNA-II and helper EACMV-TZ and EACMV-UG gave positive hybridization signals (Figure 4.16) in seven out of 15 total DNA samples tested.

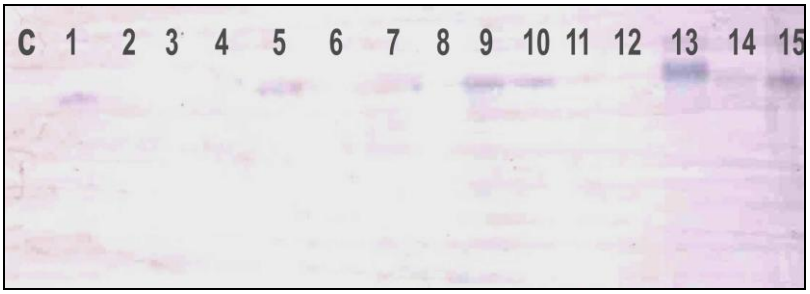

Figure 4.16: Results of Southern hybridization. Lane C: non inoculated control line C98-7; Lane 1: Bwana Mrefu (EII), sample taken from symptomatic leaves; Lane 2: Bwana Mrefu (EII), sample taken from symptomatic leaves; Lane 3: Bwana Mrefu (UII), sample taken from asymptomatic leaves AR37-96; Lane 4: AR37-96 (EII); Lane 5: AR37-96 (UII), sample taken from symptomatic leaves; Lane 6: AR37-96 (UII), sample taken from asymptomatic leaves; Lane 7: AR21-2 (EII); Lane 8: AR21-2 (UII); Lane 9: AR37-92 (EII); Lane 10: AR37-92 (UII), sample taken from symptomatic leaves; Lane 11: AR37-92 (UII), sample taken from asymptomatic leaves; Lane 12: AR37-6 (EII); Lane 13: AR16-3 (EII); Lane 14: AR17-5 (EII); and Lane 15: AR17-5 (UII).

The highest hybridization signals were recorded in cassava lines AR37-96 (lane 5) bombarded with EACMV-UG + SatDNA-II, AR37-92 (lane 9) bombarded with EII and line AR16-3 (lane 13) bombarded with EACMV-TZ + SatDNA-II.

Probing for SatDNA-III in plants bombarded with SatDNA-III and helper begomoviruses EACMV-TZ and EACMV-UG using SatDNA-III as a labeled probe, gave positive hybridization signals as presented in Figure 4.17.

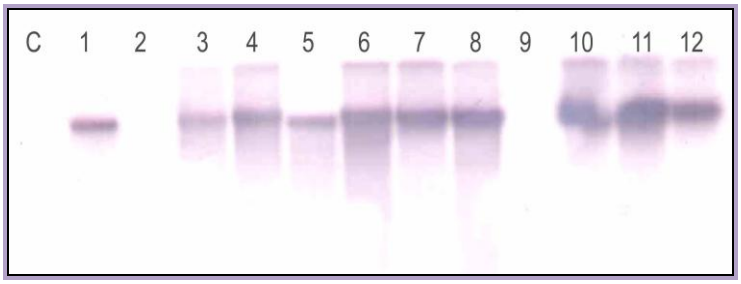

Figure 4.17: Southern hybridization for SatDNA-III. Lane C: Negative control C98-7; Lane 1: Bwana Mrefu bombarded with UIII; Lane 2: AR37-96 (UIII); Lane 3: AR21-2 (UIII); Lane 4: AR37-92 (UIII); Lane 5: CR27-24 (UIII); Lane 6: AR16-3 (UIII); Lane 7: AR17-5 (UIII); Lane 8: Bwana Mrefu (EIII); Lane 9: AR21-2 (EIII); Lane 10: AR40-10 (EIII), showed double band during amplification; Lane 11: AR37-6 (EIII), showed double band during amplification; and Lane 12: AR16-3 (EIII).

SatDNA-III was not detected in lanes 2 (cassava line AR37-96) and 9 (AR21-2) in which EACMV-UG and EACMV-TZ were respectively co-inoculated with SatDNA-III. In line AR21-2 (lanes 3 and 9), EACMV-UG was the only helper virus that could replicate SatDNA-III. The two helper viruses could equally replicate SatDNA-III in cassava line Bwana Mrefu (lanes 1 and 8).

In studying the influence of satellites on replication of helper viruses EACMV-TZ and EACMV-UG in different cassava lines as probed by a labeled *AC1* gene, the results are shown in Figure 4.18

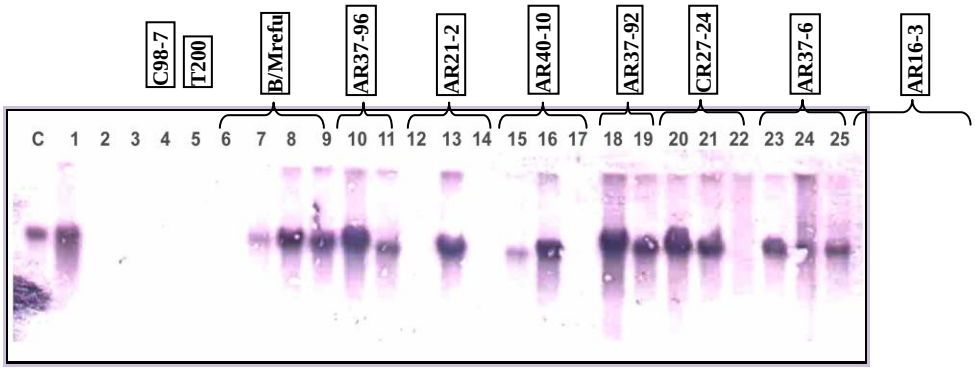

Figure 4.18: Southern hybridization for EACMV-UG and EACMV-TZ. Lane C: Uninoculated control, C98-7; Lane 1: Susceptible control, cv. T200 bombarded with E0; Lanes 2 to 5: Bwana Mrefu bombarded with E0, EII (from symptomatic leaves, EIISy), EII (from asymptomatic leaves, EIIAsy) and EIII, respectively; Lanes 6 and 7: AR37-96 (E0 and EII, respectively); Lanes 8 to 10: AR21-2 (EIISy, EIIAsy, and EIII, respectively); Lanes 11 to 13: AR40-10 (E0, EII, and EIII, respectively); Lanes 14 and 15: AR37-92 (EII and EIII, respectively); Lanes 16 to 18: CR27-24 (E0, EII, and EIII, respectively); Lanes 19 to 21: AR37-6 (E0, EII, and EIII, respectively); and lanes 22 to 25: AR16-3 (E0, EIISy, EIIAsy, and EIII, respectively).

In Figure 4.18 above, the negative or faint blots such as cv Bwana Mrefu (lane 2 – 5) and line AR37 – 92 (lanes 14 & 15) show that the lines are resistant to CMD. The causative virus (EACMV did not replicate). Positive blots such as the T200 (lane 1), line AR21-2 (lanes 8 – 10), line AR37 – 6 (lanes 19 – 21) are susceptible to CMD. The causative virus replicated enormously, leading to the accumulation of AC1. Ambiguous blots, in which blots from the same cultivar are such that some are positive and others negative, such as line AR40-10 (lanes 11 – 13) and line CR27 – 24 (lanes 16 – 18) could probably imply problems in workmanship. The control line C98-7 (lane C) was supposed to be negative. The positive signal in it means though the plant was asymptomatic, it still had some EACMV.

4.4 DISCUSSION

4.4.1 Bombardment and Symptoms Score

The death of tips and whole plants in some cases in the greenhouse was experienced, and may have resulted from pressure impact following bombardment, as the microprojectiles were accelerated from the gun at 240 *psi*. The wilting of some of the plants that was observed followed with tip death in either the second or third week, is thought to be caused by the stress of inoculation. In cassava plants inoculated with cassava mosaic begomoviruses, it is possible to observe symptoms as early as five days after inoculation as supported by findings by Chellappan *et al* (2004), although symptoms usually take 7-14 days, compared to helper virus only, in which symptoms appeared from 14 days after inoculation. In such situations, maximum severity can possibly be reached in 10 days post inoculation. In the present study, the appearance of symptoms five days post inoculation was recorded only in plants in which SatDNA-II or SatDNA-III were co-inoculated with helper begomoviruses. This is perhaps the impact of the satellites' influence on the epidemiology of cassava mosaic begomoviruses.

Recovery from infection also depends on the type of helper viruses involved. Viruses such as EACMV-TZ and EACMV-UG do not usually induce recovery (Chellappan *et al.,* 2004). So, the observed recovery in inoculated plants may have resulted due to resistance to CMD that had been introduced in the plant germplasm through breeding. In resistant plants, recovery from infection may be a gradual process, with newly emerging leaves displaying very mild symptoms. In the present study, a plasmid mix of EACMV-TZ + SatDNA-III had maximum infection in cassava line AR40-10. This line, although it also finally recovered from the infection after two weeks, its infection gradually increased over time from score 2 to 4. It could be thought that SatDNA-III enhanced the replication of EACMV-TZ in this line. The susceptible cultivar T200 also had marked stunting in week 5.

Recovery from virus infection as observed in this study is common with CMD-infected plants, depending on ambient temperature and host plant resistance (Lapidot and Friedmann, 2002). In the present study, ambient temperature is unlikely to affect

the screening process. This is because the study was carried out in subtropical, cool conditions where temperature rarely went above 35°C (Appendix 4.1). In resistant plants, high mean recovery percentages ranging from 49 to 89% can be recorded.
Consistent with the goal of screening cassava cultivars for resistance based on the cumulative symptom scores for 35 days (Table 4.5), the cultivars' levels of resistance were compared:

(i) those plants, which did not show disease symptoms seven *dpi* with the helper viruses alone or in combination with satellites were described as highly resistant;
(ii) Moderately resistant lines were those showing severity score 3 in the third week;
(iii) In the susceptible category were plants, which showed disease symptoms after inoculation with either EACMV-TZ or EACMV-UG (i.e. without satellites) and which became stunted 35 *dpi* instead of recovering from the infection.

Using the categories above, consensus categories for each cultivar in terms of resistance were evaluated as described below:

- Highly resistant lines – AR37-92, CR27-24, and AR16-3;
- Moderately resistant lines – Bwana Mrefu, AR37-96, AR40-10, AR37-6 and AR17-5;
- Susceptible line – T200.

ANOVA Table for the observed variables (Table 4.7) showed that two variables, plasmid mixes (E0, EII, EIII, U0, UII and UIII) and growth stages after bombardment (5-7dpi, 8-21dpi and 22-35dpi) were significantly different (P-Levels 0.000** and 0.034*, respectively). This means that severity of disease in the plants was changing depending on which plasmid mix was applied and also depending on which stage of growth after inoculation when observations (records) were taken. In order to know which particular plasmid mixes among the ones indicated above caused the differences or which growth stages caused the difference in disease development, mean separation had to be performed. In doing this, it was found that the growth stages that differed were between 8-21dpi and 22-35dpi, with P-values of 0.011**, at

least based on least significant differences LSD. The critical (optimum) point of disease development perhaps lies between the second and third growth stages, after which the amount of disease significantly dropped constituting recovery.

Mean separation for the plasmid mixes implied that the satellites enhanced the amount of disease (CMD) and led to breakdown of resistance (Table 4.8). Both the control plasmid mixes (E0 and U0) were significantly different from those mixes in which satellites were included. Except for the differences between E0 and EIII and that between E0 and UII, which showed significant differences at $P \leq .05$, the rest of the differences were highly significantly different at $P \leq .01$ level. The implication of this finding is that among the plasmid mixes, which involved satellites, EIII and UII caused the least level of infection. The two control plasmid mixes EACMV-TZ (E0) and EACMV-UG (U0) were not significantly different between themselves ($P \leq .281$) as revealed by this mean separation process. This finding suggests that any difference observed between plants infected by EACMV-TZ and EACMV-UG was due to the involvement of satellites, not the helper begomoviruses. This study further shows that the two satellites (SatDNA-II and SatDNA-III) involved were not significantly different between themselves.

4.4.2 PCR of *BC1, AC1,* SatDNA-II and SatDNA-III

The relatively few PCR products obtained may be due to slow growth of the plants and high resistance levels, resulting in a low rate of virus replication. The failure to amplify full length SatDNA-III may also be due to the same reason of compromised helper virus replication, culminating into failure of the later to facilitate replication of the satellite molecules.

BC1/CR is a 540 – 560 bp sequence stretch on begomovirus DNA-B covering part of the coding sequence for the movement protein and part of the common region. The gene is therefore responsible for cell to cell movement of the virus, and indicates that the virus is replicating and moving efficiently. In the present study only a few samples had positive amplification for this gene, meaning that the PCR possibly did not work.

AC1, with a putative sequence with size of about 1077 bp, is the only gene that is indispensable for viral replication and codes for the replication initiator protein (rep). In most of the samples tested in this study, there was no replication of *AC1*. A similar situation as with BC1. This could be due to PCR failure. The few cultivars from which *AC1* could be amplified suggest detectable levels of virus replication i.e. susceptibility despite the fact that these cultivars were bred for resistance. T200 is a susceptible cultivar and cassava line AR40-10 (inoculated with EACMV-TZ + SatDNA-III) was the least resistant lines.

The amplification of full-length satellites in this study has two main controversial results. One, as shown in Table 4.9, the amplification of SatDNA-II and SatDNA-III from total plant DNA occurred even when the satellites were not inoculated with the helper viruses, and also in uninoculated healthy cassava germplasm. PCR amplification of SatDNA-III was weak, giving faint bands on the agarose gels. The reason for this was not known at the time, but is now known that integrated SatDNA-II and SatDNA-III sequences occur and this may explain this amplification. The occurrence of bands other than the expected size, even after adjusting amplification profile was another unexpected result. Again, the integrated SatDNA sequences may yield different PCR-sized fragments as they are integrated more than once, and also different sizes of satellite sequences have been demonstrated (unpublished results).

The cassava cultivars used for inoculation experiments in this study were grown in areas with high prevalence of satellites. So, based on the possibility of integration, it is possible that at the time of biolistic inoculation, the stem cuttings already had satellites in them. The occurrence of multiple PCR bands in the present study is thought to be due to multiple integrations of different sized satellites in the cassava germplasm. The difficulty encountered in the amplification of full-length SatDNA-III could mean that the bombarded SatDNA-III was not replicated by the helper virus. The greater severity could thus be attributed to SatDNA-II

And this also confirms that SatDNA-III is distinctly different from SatDNA-III in terms of structure and therefore biochemical activity in begomovirus epidemiology. The putative ORFs in SatDNA-II could mean that they are coding for proteins, for

example in SatDNA-II, ORFs V1, C4 and C5 are long enough to code for functional proteins.

4.4.3 Amplification of Satellite Fragments

As mentioned earlier, the anomalies associated with amplification of full-length satellites using episomal primers suggested integration of satellite fragments. It has been indicated on Table 4.9, that during the amplification of the integrated sequences using EST-based primers, PCR products were obtained on the agarose gels even if the satellite sequences was not inoculated. This suggested integration of SatDNA-II and SatDNA-III sequences into the genome of cassava. The successful amplification of SatDNA-II and SatDNA-III sequences from cassava total DNA using EST-based primers was an indication that the satellites were present as integrated form even in CMD-resistant cultivars. This is why in the previous finding in this work, SatDNA-II could be PCR amplified even from plants where it had not been inoculated. It would appear that most of the SatDNA-III, which did not amplify with the full length satellites primers, was present in an integrated form that is why full-length SatDNA-III could not be amplified by PCR. In SatDNA-II, some of the circular forms are partly episomal and replicated by the helper virus and hence detected by PCR. For SatDNA-III, all of the sequences may be present as integrated form.

4.4.4 Southern Blot Hybridization

Negative hybridization during Southern blotting suggests poor or no replication of the helper virus sequence that is being probed. In other words, the inoculated plant did not allow replication of the introduced sequence, implying resistance. On the other hand, positive hybridization signals indicate susceptibility of the inoculated plant to the introduced sequence.

Based on this assumption, the bombarded cassava lines can be arbitrarily grouped into three categories based on Southern blots for SatDNA-II fragments (Figure 4.16). Highly resistant cassava lines, with inoculated plasmid mixes indicated in brackets include: Bwana Mrefu (EIIAsy, lane 2), Bwana Mrefu (UIIAsy, lane 3), AR37-96 (EII, lane 4), AR37-96 (UIIAsy, Lane 6), AR21-2 (UII, lane 8), AR37-92 (UIIAsy,

lane 11), and AR37-6 (EII, lane 12). As most of these hybridization signals were observed on DNAs extracted from asymptomatic leaf samples, it could mean that the signals were due to the presence of integrated forms of the satellites. This assumption is partly in agreement with conclusive remarks drawn for symptom scores in which cassava line AR37-92 was grouped as highly resistant line. As observed on data analysis for mean disease severity scores (MDSS), the use of EACMV-UG or EACMV-TZ as helper virus seems to cause no difference in hybridization signals. The landrace cultivar Bwana Mrefu for example, inoculated with plasmid mix EII and that inoculated with UII had similar hybridization signals.

Moderately resistant cultivars are those which gave faint hybridization signals and include Bwana Mrefu (EIISy, lane 1), AR37-96 (UIISy lane 5), AR21-2 (EII, lane 7) and AR37-92 (UIISy, lane 10). By far, this moderately resistant cultivars category agrees with the MDSS categories in which cassava line AR37-92 was also rated as moderately resistant line. This observation was mostly recorded on DNA samples extracted from symptomatic leaves. Where the hybridization signals were strong, as observed for cassava lines AR37-96 (UIISy, lane 5), AR37-92 (EII, lane 9), AR16-3(EII, lane 13) and AR17-5 (lane 15), this would imply susceptibility. There is one point to note, though. Cassava line Bwana Mrefu (lanes 1 and 3) and line AR37-92 (lanes 9 and 11) were only susceptible to SatDNA-II replication when EACMV-TZ was the helper virus involved rather than EACMV-UG.

Southern blotting (SB) of SatDNA-III and using the same definitions as for SatDNA-II above, the grouping is as follows: Highly resistant cassava lines – AR37-96 (UIII, Lane 2), and AR21-2 (EIII, lane 9); moderately resistant lines – AR21-2 (UIII, Lane 3), and CR27-24 (UII, lane 5). Susceptible lines – AR37-92 (UIII, lane 4), AR16-3 (UIII, lane 6), AR17-5 (UIII, lane 7), Bwana Mrefu (EIII, lane 8), AR40-10 (EIII, lane 10), AR37-6 (EIII, 11), and AR16-3 (EIII, lane 12). The use of EACMV-TZ or EACMV-UG as helper viruses does not make any difference. For example, AR16-3 in lanes 6 and 12 inoculated with UIII and EIII respectively had similar signal strength.

Several observations on the influence of satellites on replication of helper viruses EACMV-TZ and EACMV-TZ, in different cassava cultivars using *AC1* gene as the Southern blotting probe (Figure 4.18), can be made. The susceptible control agreed

with the idea that the local cultivar T200 (lane 1) is highly susceptible to CMD, giving a strong signal. The negative (resistant control) was positive for the hybridization signal, but this is not surprising as we expect resistant germplasm to be infected but to recover at a later stage. Highly resistant cassava lines were regarded as such if there were negative hybridization signals. This includes Bwana Mrefu (lanes 2 – 5) inoculated respectively with E0, EIISy, EIIAsy and EIII; AR37-96 (lane 6) inoculated with E0; AR40-10 (lane 12) inoculated with EII, AR37-92 (lane 14) inoculated with EII; and CR27-24 (lane 17) inoculated with EII.

Moderately resistant lines are AR37-96 (lane 7) inoculated with EII, AR40-10 (lane 11) inoculated with E0, AR37-92 (lane 15) inoculated with EIII and AR16-3 (lane 22) inoculated with E0. In susceptible cassava lines, the blotting gave strong hybridization signals. In Figure 4.18, this includes the local cultivar T200 (lane 1) inoculated with E0, AR21-2 (lanes 8 to 10), AR40-10 (lane 13) inoculated with EIII, CR27-24 (lane 18) inoculated with EIII, and AR37-6 (lanes 19 to 21) inoculated with E0, EII, and EIII, respectively.

The helper virus *AC1* gene could not be detected in plants where the virus was co-inoculated with SatDNA-II and SatDNA-III. In cassava line AR37-96 (lanes 6 and 7), *AC1* could only give signals in a plant where EACMV-TZ was co-inoculated with SatDNA-III. The overall resistance ranking based on blots could be slotted as follows: Highly resistant lines – AR37-96, AR37-92 and AR21-2. Moderately resistant lines – Bwana Mrefu and CR27-24. Susceptible lines are T200, AR40-10, AR16-3 and AR17-5.

4.5 Conclusion

One of the effects of particle bombardment on cassava raised from stakes was death of growing tips. Disease symptoms were observed as early as five days after inoculation with helper virus and satellites. The observed recovery is thought to result from resistance introduced in the plant materials involved through breeding. The highest disease score was recorded in cassava line AR40-10, inoculated with EIII. Based on symptom score, and consistent with evaluation for resistance based on symptoms and positive/negative Southern blot hybridizations, highly resistant cassava

lines were identified as AR37-92, CR27-24, and AR16-3 while moderately resistant lines were Bwana Mrefu AR37-96, AR40-10, AR37-6 and AR17-5. The susceptible cultivar was T200.

Statistical analysis however showed that the nine tested cassava lines were equally resistant. A much longer period than two weeks is suggested for optimum infection when working with CMD-resistant plant materials. The analysis showed that both SatDNA-II and SatDNA-III enhanced disease severity, but the two satellites were not significantly different between themselves. The two control plasmid mixes (E0 and U0) did not contain SatDNA-II or SatDNA-III. Plasmid mixes EIII (EACMV-TZ + SatDNA-III) and UII (EACMV-UG + SatDNA-II) caused the least level of infection (compared to EII and UIII).

The amplification of full-length episomal SatDNA-III was not successful from plant total DNA. The amplification of full-length SatDNA-II gave multiple bands with episomal primers. Amplification of satellite fragments using EST-primers showed multiple bands for both satellites irrespective of whether the corresponding satellite was included in the inoculation plasmid mix. However, for the case of SatDNA-II, the multiple bands included two closely spaced bands. DNA fragments representing SatDNA-II, SatDNA-III and *AC1* and *BC1*, produced by PCR were labeled by a steroid hapten DIG, a chromogenic dye containing random primed DNA digoxigenin-dUTP (according to manufacturer's instructions) and used in hybridization separately as described in Appendix 4.8. In this resistance study, the highly CMD-resistant cassava lines are AR37-92 and AR37-96. Moderately resistant is the local cultivar Bwana Mrefu while T200, AR16-3 and AR40-10 are susceptible lines. Since the labeled probes for satellites were derived from the integrated sequences, the hybridization signals did not depend on whether leaf samples were picked from symptomatic or asymptomatic leaves. In this resistance study, the highly CMD-resistant cassava lines are AR37-92 and AR37-96. Moderately resistant is the local cultivar Bwana Mrefu while T200, AR16-3 and AR40-10 are susceptible lines.

The amplification of episomal SatDNA-II and SatDNA-III sequences proved problematic, due to the discovery after the experiments in this chapter that integrated forms of these two satellites occur widely in both infected and healthy cassava germplasm. Furthermore,

these satellites can integrate as full-length sequences or as smaller fragments, thereby potentially leading to the amplification of multiple bands with EST-derived primers. Therefore no unequivocal positive conclusions can be made until the effects of the integrated versus the episomal forms of these satellites are evaluated in integrated satellite-free germplasm.

4.6 References

1. Ariyo O.A., Koerbler M., Dixon A.G.O., Atiri G.I. and Winter S. (2003) Development of an efficient transmission technique to screen cassava genotypes for resistance to cassava mosaic disease. Proceedings of a Conference on International Agricultural Research for Development held at Gottingen, Germany October 8–10, 2003.
2. Ariyo O.A., Atiri G.I., Dixon A.G.O. and Winter S. (2006) The use of biolistic inoculation of cassava mosaic begomoviruses in screening cassava for resistance to cassava mosaic disease. *J. Virol. Methods* 137: 43–50.
3. Chellappan P., Vanitharani R., Ogbe F. and Fauquet C.M. (2004) Effect of temperature on geminivirus-induced RNA silencing in plants. *Plant Physiol*. 138: 1828–1841.
4. FAO (2008) FAO Statistics Division. http:/www.fao.org.
5. Fraser R.S.S. (2006) Virus recognition and pathogenicity: Implications for resistance mechanisms and breeding. *Pest Management Science* 22 (3): 267- 275.
6. Hahn S. K., E. R. Terry and K. Leuschner (1980) Breeding cassava for resistance to cassava mosaic disease. *Euphytica* 29 (3): 673-683.
7. Mallowa S.O., Isutsa D.K., Kamau A.W., Obonyo R. and Legg J.P. (2006) Current characteristics of cassava mosaic disease in post-epidemic areas increase the range of possible management options. *Ann. Appl. Biol*. 149 (2): 137-144.
8. Ndomba, O.A. (2005) Towards developing a systems approach for pest risk management in Tanzania: Challenges and Opportunities. *Trop. Pest. Mgt. Bull.* 3(6): 92–96.
9. Ogbe F.O., Atiri G.I., Dixon A.G.O. and Thottappilly G. (2003) Symptom severity of cassava mosaic disease in relation to concentration of *African cassava mosaic virus* in different cassava genotypes. *Plant Pathology* 52: 84–91.

10. Okogbenin E., Porto M.C.M., Egesi C., Mba C., Espinosa E., Santos L.G., Ospina C., Marin J., Barrera E., Gutierrez J., Ekanayake I., Iglesias C. and Fregene M.A. (2007) Marker assisted introgression of resistance to cassava mosaic disease into Latin American germplasm for the genetic improvement of cassava in Africa. *Crop Sci.* 47: 1895–1904.
11. Raji A., Ladeinde O. and Dixon A. (2007) Screening landraces for additional sources of resistance to cassava mosaic disease and green mite for integration into the cassava improvement program. *Journal of Integrated Plant Biology* 50 (3): 311-318.
12. Schopke C., Franche C., Bogusz D., Chavarriaga P., Fauquet C. and Beachy R.N. (1993) Transformation in cassava (*Manihot esculenta* Crantz). Biotechnology in Agriculture and Forestry Vol 23.
13. Sseruwagi P.P., Legg J.P., Rey C.M.E., Colvin J. and Brown J.K. (2003) Molecular variability of cassava *Bemisia tabaci* genotypes in East and Central Africa. *In* 3rd International *Bemisia* workshop Barcelona 17 – 20 March, 2003.
14. Thresh J.M., Otim-Nape G.W. and Fargette D. (1998) The control of African cassava mosaic disease: Phytosanitation and/or Resistance?. *In* Hadid A, Khetarpal R.K., Koganezawa H. (eds). Plant Virus Disease Control. American Phytopathological Society Press Pp. 670–677.
15. Zhang P., Vanderschuren H., Futterer J. and Gruissem W. (2005) Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. *J. Plant Biotech.* 3: 385–397.

CHAPTER 5

INTERACTION BETWEEN BEGOMOVIRUS DNAs AND SATELLITE DNA MOLECULES IN NICOTIANA BENTHAMIANA

ABSTRACT

Cassava mosaic begomovirus, *East African Cassava mosaic virus Tanzania* (EACMV-TZ) DNA-A and DNA-B, SatDNA-II and SatDNA-III clones previously proven to be infectious through biolistic inoculation, were prepared for interaction studies. Different combinations of the satellite DNAs and helper virus EACMV-TZ (treatments) were inoculated (concentration of 500 ng/μL) by the abrasion method onto *Nicotiana benthamiana* Domin. (*Solanaceae*). The six treatments were as follows: EACMCV alone (E0), EACMCV + SatDNA-III (EIII), EACMCV + SatDNAII (EII); EACMCV + SatDNA-II + SatDNA-III (EII III), SatDNA-II alone (II), SatDNA-III alone (III), and an uninoculated plant (negative control). To inoculate the DNA mixes into *Nicotiana benthamiana*, carborandum was evenly dispersed onto target leaves, followed by the inoculation of 1–5 g of plasmid DNA in 20 μL sdH_2O. Five plants were inoculated for each treatment, with two to three young leaves larger than 1.5 cm in length used for abrasion. The leaves were then washed with a distilled water to remove excess carborundum. The inoculated plants were covered with a plastic dome and placed in an insect-free growth chamber for symptom development. The inoculated plants were observed for symptom development. The symptoms were scored on a severity scale of 1 – 5. Total DNA was extracted from *N. benthamiana* leaves and used for analysis of DNA accumulation by Southern blot analysis. In the blotting process, total DNA (4 μg) was electrophoresed on a 1% agarose gel without ethidium bromide and transferred to nylon membranes. Viral DNA was detected by using specific non-radioactively DIG-labelled probes of EACMCV DNA-A, as well as SatDNA-II and SatDNA-III according to manufacturer's instructions.

Nicotiana benthamiana plants inoculated with helper virus only (E0) developed moderate symptoms 7 days after inoculation (dpi), which comprised of leaf distortion and moderate plant stunting and in which plants recovered 35 dpi. Plants inoculated with EII produced symptoms resembling those on plants inoculated with E0. *N. benthamiana* plants inoculated with EIII, developed symptoms 7 dpi, that became severe by 14 dpi and there was no recovery even after 35 dpi. Severe symptoms (score>3) were also observed when *N. benthamiana* plants were inoculated with EACMCV- and both satellites (EII III). All plants inoculated with SatDNA-II or SatDNA-III alone remained asymptomatic even after 35 dpi. Southern blot analysis using total DNA extracted from *N. benthamiana* 14 dpi inoculated with EIII alone, but not E11, revealed a clear increase in EACMCV DNA accumulation as compared plants inoculated with EACMCV only or EACMCV with SatDNA-II only (EII) and probed with EACMCV. This study showed enhanced virus symptoms on *N. benthamiana* when EACMCV-was co-inoculated with both SatDNA-II and SatDNA-III. This suggests that the satellite molecules play a role in symptom modification, and that EACMCV is involved in replicating the satellites albeit using unknown mechanisms.

5.1 INTRODUCTION

Begomoviruses belong to family *Geminiviridae* (Fauquet and Stanley, 2003) and are known to cause serious epidemics in many crop plants and ornamentals such as tomato, cassava, cotton, and tobacco. Transmitted by whitefly vector *Bemisia tabaci*, begomoviruses have either bipartite or monopartite genomes and infect dicotyledonous hosts. The successful state of begomovirus epidemic in plants owes to successful establishment of its whitefly vector, the presence of susceptible plant varieties, propensity of the begomoviruses for genetic variation and their ability to associate with single stranded replicons, especially the satellite DNAs. The association of begomovirus pathogens with betasatellites is well-known, where the role of ß-satellites in movement facilitation of cognate begomovirus helper, and the dependence of the satellites on the helper begomovirus for replication, is established (Briddon and Stanley, 2006).

In cassava, begomoviruses cause cassava mosaic disease (CMD), which is the most important biotic factor affecting cassava production. Resistance bred into the local cassava cultivars, by conventional breeding, appears in some cases to be broken down in the field. In Tanzania, the breakdown of resistance has been attributed to several factors, such as the synergistic association between ACMV and EACMV (Sserubombwe *et al*., 2008; Fondong *et al*., 2000) and the occurrence of a recombinant EACMV-U2. Recently, the discovery of satellite-like DNA molecules which occur with their cognate helper begomovirus, has been reported (unpublished). They breakdown of resistance is accompanied by exacerbated severe disease symptoms in cassava. The satellites have been designated SatDNA-II and SatDNA-III.

Knowledge of these helper begomoviruses and associated satellites infecting cassava are crucial step if management of CMD by developing resistant cultivars is to be achieved, and this study aimed to evaluate the presence of SatDNAII and III on levels of helper virus replication and disease severity.

5.2 MATERIALS AND METHODS

5.2.1 Begomovirus and Satellite Infectious Clones

Cassava mosaic begomoviruses, *East African Cassava mosaic virus Cameroon* (EACMCV-) DNA-A and DNA-B, SatDNA-II and SatDNA-III infectious clones were kindly provided by Dr. Joseph Ndunguru of Mikocheni Agricultural Research Institute (MARI), Dar es Salaam and were previously proven to be infectious by biolistic inoculation (Ndunguru, 2005; Ndunguru *et al*., 2006).

5.2.2 Treatments

Different combinations of the satellite DNAs and helper virus EACMCV (treatments) were prepared from DNAs (concentration of 500 ng/μL) and abraded onto *Nicotiana benthamiana* Domin. (*Solanaceae*). There were 6 treatments as follows: EACMCV-alone (E0), EACMCV + SatDNA-III (EIII), EACMCV + SatDNA-II (EII); EACMCV + SatDNAII + SatDNAIII (EII + III), SatDNA-II alone (II), SatDNA-III alone (III) and uninoculated (Negative control). E0 was formulated by adding 10 μL

of each of EACMCV DNA A and DNA B to 10 μL of sterile distilled water (sdH_2O) in a microfuge tube. EII or EIII was made by using 6 μL of each of EACMCV DNA-A and DNA-B components, 5 μL of either Sat II or SatDNA-III and 10 μL of sdH_2O. EII+ III was made by using 6 μL of each of EACMCV DNA-A and DNA-B, 5 μL of SatDNA-II, 5 μL of SatDNA-III and 10 μL of sdH_2O. Treatment II was made by adding 6 μL of SatDNA-II to 5 μL of sdH_20. Treatment III was made by adding 6 μL of SatDNA-III, respectively, to 5 μL of sdH_20. The negative control was not inoculated.

5.2.3 Geminivirus and Satellite infection

To infect *Nicotiana benthamiana*, DNA abrasion method previously described by Trinidad *et al*. (2007) was used. Carborundum (glass powder has also been used successfully) was evenly dispersed onto target leaves from a 15 mL conical tube through 12 layers of cheesecloth, taking care that the abrasive should be almost invisible on the leaf. This was followed by the inoculation of 1–5 μg of plasmid DNA in 20 μL sdH_2Or or up to 20 μg of total DNA from infected plants. The 20 μL DNA solution was placed at the leaf base and spread by rubbing with 10 soft but firm strokes from the base to tip with a gloved finger. Each plant was inoculated on two to three young leaves larger than 1.5 cm in length. At the end of the manipulations, the leaves were gently washed with a stream of distilled water to remove excess carborundum. Plants were covered with a plastic dome for 24 h and placed in an insect-free growth chambers (Figure 5.1). A total of 5 plants were inoculated for each treatment.



Figure 5.1: Insect-free chambers used to grow test plants in green house.

5.2.4 Scoring for Disease Symptoms

The inoculated plants were observed for symptom development and scored on scale 1 – 5 (Sseruwagi *et al*., 2003) as follows: Score 1 = The plant is healthy; Score 2 = Slight infection (CMD-like infection), at least one leaflet is malformed; Score 3 = Mild infection, mild curl in entire leaf; 4 = Severe infection and leaf curl in 1 – 3 leaves; Score 5 = very severe infection and leaf curl/chlorosis in more than 3 leaves, coupled with severe stunting. The plants were scored for virus symptoms over a period of 35 days.

5.2.5 Total DNA Purification and DNA Accumulation Analysis

Total DNA was extracted from *N*. *benthamiana* leaves as described by Dellaporta *et al*. (1983) and used for analysis of DNA accumulation in Southern blot analysis. For Southern blot analysis, total DNA (4 µg) was electrophoresed on a 1% agarose gel without ethidium bromide and transferred to nylon membranes. Viral DNA was detected by using specific non-radioactively DIG-labelled probes of EACMCV DNA-A, as well as SatDNA-II and SatDNA-III according to manufacturer's instructions.

5.3 RESULTS

5.3.1 Symptom Monitoring

Nicotiana benthamiana plants inoculated with EACMCV infectious clone (E0) developed moderate symptoms 7 days after inoculation (dpi). Symptoms consisted of leaf distortion and moderate plant stunting which by 35 dpi plants recovered (Figure 5.2). Plants inoculated with EACMCV together with SatDNAII (EII) produced symptoms that resembled those displayed when inoculated by E0 (figure 5.2). However, *N. benthamiana* plants that were inoculated with EACMCV together with SatDNA III (EIII), developed symptoms 7dpi that became severe by 14 dpi and there was no recovery even after 35 dpi. Very severe symptoms were also observed when *N. benthamiana* plants were inoculated with EACMCV and SatDNAII and SatDNAIII (EII+III). However, all plants inoculated with SatDNA II or III alone remained asymptomatic even after 35 dpi (Figure 5.1).

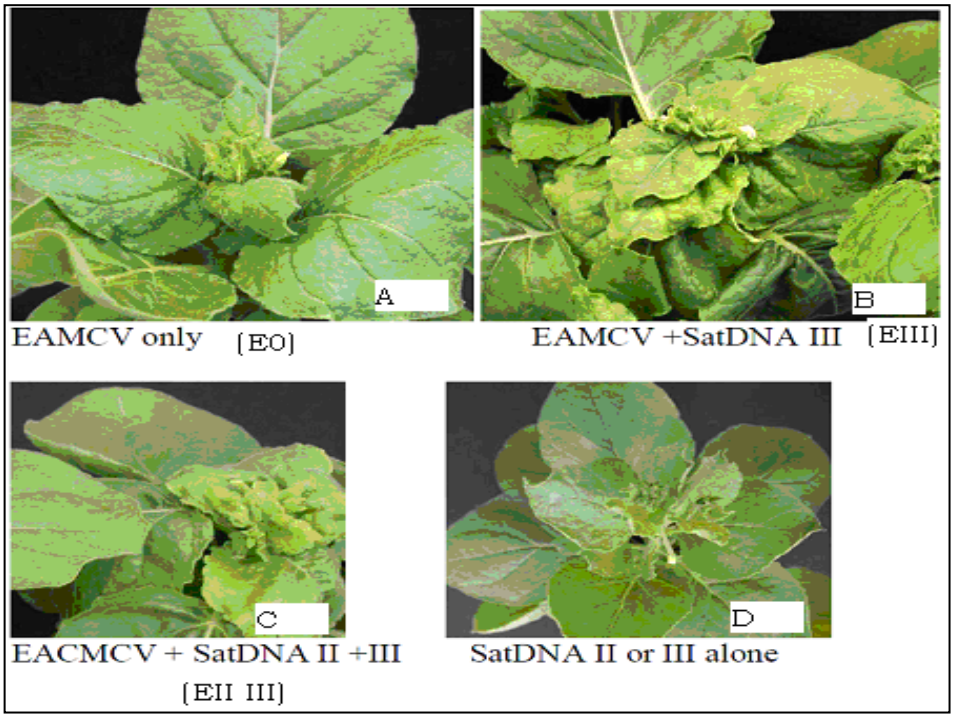

Figure 5.2: *N. benthamiana* plants inoculated with different virus/satellite combinations 14 days after inoculation.

5.3.2 Southern Blot Analysis

To examine the effect of SatDNA-II and SatDNA-III on cassava begomovirus DNA accumulation as well as the impact of cassava begomovirus on satellite replication, total DNA was extracted from *N. benthamiana* 14 dpi and used for Southern blot analysis in a non-radioactively DIG-labeled system using probes specific for EACMCV DNA-A and SatDNA-II and SatDNA-III. Clear increase of DNA accumulation was noted when EACMCV was inoculated with SatDNAIII or together with both SatDNA-II and SatDNA-III as compared to when EACMV-TZ was inoculated alone or with SatDNA-II onlyand probed with EACMCV (Figure 5.3). DNA accumulation analysis also showed signals that were observed only when satellite II or III were inoculated with EACMCV but not when they were inoculated alone (Figure 5.3).

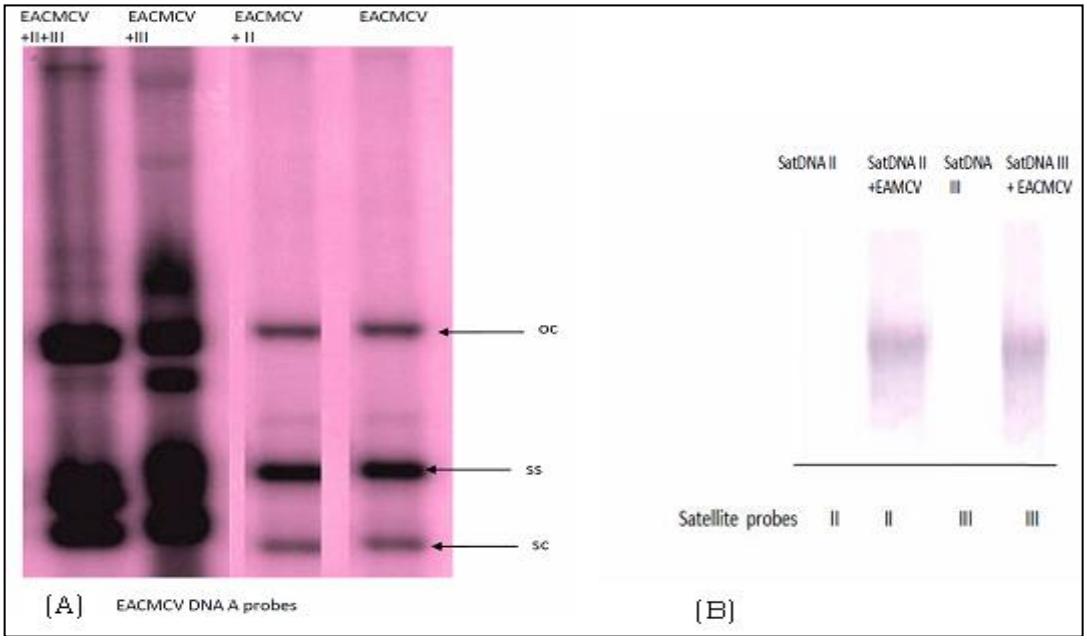

Figure 5.3: Southern Blot hybridization blots showing levels of DNA accumulation in *N. benthamaina*. (A) Levels of EACMCV DNA-A accumulation showing the three replicative forms (oc = open circular, ss = single-stranded DNA and sc = super coiled) in the presence of the two satellite combinations and probed with EACMCV DNA-A; (B) Levels of satellites DNA accumulation in the presence and absence of EACMCV and probed with SatDNA-II or SatDNA-III.

5.4 DISCUSSION

This study showed enhanced virus symptoms on *N. benthamiana* when EACMCV was co-inoculated with SatDNA-III or both SatDNA-II and SatDNA-III. However, no discernable increase in symptom severity was observed when EACMCV was co-inoculated with SatDNA-II alone, confirming an earlier report (Ndunguru, 2005). Symptom enhancement by co-infection of the two satellites used here has also been demonstrated in cassava previously (Ndunguru, 2005). The phenotypes of symptoms observed in the present study correlated those observed in the earlier study (Ndunguru, 2005), which comprised severe leaf deformation. The mechanism involved in the symptom enhancement is not clear but it is possible that the satellites encode putative proteins that overcome host defense to allow the begomoviruses to establish infection more readily than in plants singly infected by the helper begomovirus. It is also possible, though not established, that the satellites encode a gene silencing suppressing factors therefore allow viruses to establish infections more easily.

Virus and satellite DNA accumulation analysis showed increased viral DNA levels when plants were co-inoculated with the two satellites as compared with when EACMCV was inoculated alone, or when EACMCV was incoculated with satDANAII. Two possibilities can be used to explain this: Firstly, it is possible that the increased virus accumulation in the presence of satellites is resulting from more plant cells getting infected in the presence of SatDNAIII, which could be enhancing movement, as shown with the betasatellites (Briddon and Stanley, 2006). Secondly, SatDNAIII may be enhancing begomovirus replication. The two possibilities are subject to further investigation to generate a knowledge base. Recent knowledge that the satellite DNA fragments and full-length satDNAII and III are able to integrate into the cassava genome (unpublished) may shed some light on these novel molecules and their potential role in pathogenicity. However it has not been demonstrated that these satellites occur in the *N. benthamiana* genome, suggesting that the episomal forms play a role in symptom enhancement. Both satellites II and III DNA accumulated only when they were co-inoculated with EACMV-TZ but not when they were

inoculated alone. This suggests that EACMV-TZ could play a role in replicating the satellites using unknown mechanisms.

5.5 CONCLUSION

This study has observed enhanced cassava begomovirus symptoms in *N. benthamiana* in the presence of SatDNA-III and when the virus was inoculated with a combination of SatDNA-II and SatDNA-III, but not with satDNAII alone. This suggests that satDNAII may not play a role in symptom phenotype alteration. These satellites were only replicated in the presence of EACMCV suggesting some interaction between the helper virus and its associated cognate satellite molecules, in an unknown mechanism worthy investigating in future.

5.6 REFERENCES

1. Ascencio-Ibanez J.T. and Settlage S.B. (2007) DNA abrasion onto plants is an effective method for geminivirus infection and virus-induced gene silencing. *J. Virol. Meth.* 142: 198–203.
2. Briddon R.W. and Stanley J. (2006) Subviral agents associated with plant single stranded DNA viruses. *J. Virol.* 344 (1): 198-210.
3. Fauquet C.M. and Stanley J. (2003) Geminivirus classification and nomenclature: Progress and problems. *Annals of Applied Biology 142:165-189.*
4. Fondong V.N., Pita J.S., Rey M.E.C., DeKochko, Beachy R.N., Fauquet C.M. (2000) Evidence of synergism between African cassava mosaic virus and a new double recombinant Geminivirus infecting cassava in Cameroon. *J.Gen. Virol*., 81: 287-297.
5. Ndunguru J., Legg J.P., Fofana I.B.F., Aveling T., Thompson G. and Fauquet C. (2006) Identification of a defective molecule derived from DNA-A of the bipartite begomovirus of East African cassava mosaic virus. *Plant Path.* 55: 2-10.
6. Ndunguru J. (2005) Molecular characterization of cassava mosaic geminiviruses in Tanzania. PhD thesis. University of Pretoria, South Africa.

7. Samuel G (2008) Some experiments on inoculating methods with plant viruses and on local lesion. *Annals of Applied Biology* 18 (4): 494–507.
8. Sserubombwe W.S., Briddon R.W., Baguma Y.K., Ssemakula G.N., Bull S.E., Bua A., Alicai T., Omongo C., Otim-Nape G.W. and Stanley J. (2008) Diversity of begomoviruses associated with mosaic disease of cultivated cassava (*Manihot esculenta* Crantz) and its wild relative (*Manihot glaziovii* Mull. Arg.) in Uganda. *J. Gen. Virol* 89: 1759–1769.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions and significance of findings

Satellites are composed of a nucleic acid and depend on co-infection of host cell with a helper virus for their multiplication. They are usually associated with plant viruses and many of them modulate disease symptoms. The earliest satellites were associated with pathosystems that involved RNA viruses, but recently satellites associated with DNA viruses, such as geminivirsues, have been discovered. Four categories of such satellites have been described – TLCV Sat-DNA, betasatellites, alphasatellites, and Satellite DNAs. Recent, intensive research on satellites has ushered in fresh knowledge about their association with viruses. For example, positive interaction has been demonstrated between betasatellites and the NW bipartite begomovirus *Cabbage leaf curl virus* (CbLCuV) (Nawaz-ul-Rehman *et al*., 2009) contrary to the popular belief and knowledge. Also, the recent discovery of satellite DNAs proves that the Old World (OW) bipartite begomoviruses can indeed be associated with satellites.

The association between satellites and viruses has caused a huge economic loss on crop yields through exacerbated disease symptoms. Most of the serious disease epidemics in plants – OLCD, CLCD, and now the severe CMD, are a result of the association and interaction of satellites with begomoviruses in afflicted plants. The association has been traced and studied over temporal and spatial components across different crop plants. A cross section of members of Genus *Begomovirus-associated satellite DNA* comprising 65 members (Appendix 6.1) shows that 62.5% are found in tomato, 20.8% in tobacco and 16.6% in each of the crops cotton and okra. Only two satellites, *African cassava mosaic virus-associated DNA-II (Sat DNA-II)* and *African cassava mosaic virus-associated DNA-III* (comprising 8% of the cross section) are associated with cassava. Of the 65 members, betasatellites comprise 66%, alphasatellites 26% and satellite DNAs 3%. By and large, the DNA satellites are unique. Unlike the TLCV Sat-DNA, satellite DNAs interact with the helper viruses and modulate disease symptoms; the satellites do not completely resemble their helper virus sequences; they do not autonomously replicate because they do not encode *Rep*; and they are not known to be trans-replicated by other viruses. Further, the SatDNAs

do not have the GG-rich repeats (TGGTGGT) instead they have GC-rich repeats (CCGCCG, CCGCC and CGC). Satellite DNAs are similar to TLCV Sat-DNA by not being transcriptionally active due to small and insignificant ORFs, thus relying on geminiviral or host factors for replication and systemic spread. SatDNAs also have evidence of integration into host genome while the previously reported satellites were not known to be integrated into genome.

From the list of satellites (Appendix 6.1), it is evident that comparatively little is known about the satellite DNAs, and therefore this group deserves a thorough and continued investigation in an effort to understand the role of satellites in diseases, and in order to develop strategic interventions to reduce losses due to begomovirus diseases. The first record of the association of satellite-like DNAs with OW bipartite begomoviruses was with cassava in the major cassava growing areas of Tanzania. A satellite-like DNA was reported in the coastal regions of Tanzania, associated with enhanced disease symptoms (data not published). With specific reference to this study, the goal of this study has been to survey wider areas in Tanzania for satellites, and to evaluate the validity and reliability of the various levels of resistance to CMD that have been developed in local cassava cultivars when satellites are present. Furthermore the study aimed to further elucidate the nature of the interaction between helper begomovirus and satellite DNA in cassava cultivars in relation to disease development and severity.

One of the most important findings of this study was that the satellites are of a wide occurrence within the country, including areas where their presence was not previously reported. Concurrent occurrence of the full length versions of both satellites (II and III) within a host was observed in many plant samples across the country. PCR products of full length satellites produced multiple bands on stained agarose gels, particularly SatDNA-II. This could imply that the satellites exist in different replicative forms.

In terms of disease resistance and susceptibility in cassava germplasm, screen house and laboratory studies showed that the satellites (especially SatDNA-II) possibly cause the observed reversion of resistance and increase in disease symptoms in the some cassava cultivars. This is a significant finding, as this has enormous

implications for genetic modification of cassava for begomovirus resistance, and secondly for cassava breeding. Two other important implications emerge from this observation in this study. Firstly, the satellites will likely be a focal point for future resistance studies in cassava, and secondly it highlights the importance of searching for DNA markers to these widespread DNA molecules to monitor their movements;

The occurrence of reversion could have three important implications in the implementation of Plant Protection and Quarantine by the National Plant Protection Organizations (NPPOs). First, cassava should not be moved unreasonably across borders. Second, the satellite DNAs should be included in the list of regulated non-quarantine pests for Tanzania. Areas of occurrence should be mapped and a plan for official control be implemented. Third, cassava should be included in the list of high risk plant material imported into Tanzania.

Another hugely significant and exciting finding was that integrated forms of the satellites were observed across many cassava germplasms widely spread across Tanzania. Genome integrated forms were amplified from 68% and 71.17% of samples for SatDNA-II and SatDNA-III, respectively. Thirty percent of the samples showed co-infection of the satellites. Upon sequencing, fragments (306 and 875bp) of the integrated satellite DNA-II and SatDNA-III isolates, respectively, showed high nucleotide sequence similarity with cognate regions of the full-length satellite sequences deposited in the Genebank (accession numbers AY836366 and AY836367 for SatDNA-II and SatDNA-III, respectively). Sequence alignments revealed presence of GC-rich regions, TATA protein binding motifs (TATAAAT) and CAAT boxes as well as poly (A) signal. GC-rich regions in SatDNA-II were mostly trinucleotides (CGC) and hexanucleotides (CCGCCG) while in SatDNA-III the regions were trinucleotides (CGC) and pentanucleotides (CCGCC). The role of these integrated forms is not known, but is currently the subject of further investigation. The potential interaction between the integrated and episomal satellites DNAs with their cognate geminivirus could prove not only to be fascinating in terms of molecular interactions and functions, but also will impact on developing virus resistance or tolerance in cassava.

Bombardment of the satellites (with helper geminivirus) into cassava cultivars bred for resistance to CMD (recovery from severe to mild symptoms indicate that the bred-cultivars are resistant to CMD) showed that levels of resistance differ among cultivars when satellites are included. However, most cultivars showed a breakdown of resistance/tolerance, leading to enhanced symptoms. Further, it was found that many but not all resistant cassava varieties, tested in this study, have satellites integrated into their genomes. This suggests that somehow the presence of the satellite is causing a reversal of innate immunity (RNA silencing), which has been suggested as a possible mechanism of resistance. It is thought that the integrated satellite sequences in resistant cultivars target incoming homologous satellite sequences, thereby exacerbating disease symptoms. In the absence of the satellite the resistant cultivars display partial or full RNA silencing of the incoming viral DNAs. This has grave implications for future breeding plans for resistance to CMD, and an undertanding of these episomal and intergarted satellite forms are urgently needed.

The amplification of full length satellites from the bombarded plants show that SatDNA-II is more easily replicated than SatDNA-III. This is an indication that the two satellites are distinctly different as described in earlier studies. In studying the nature of the interaction, it was found that the satellites alone, or in combination with their cognate helper did not enhance disease symptoms indicating that interaction between the two satellites themselves may be inexistent. Co-inoculation of satellites and begomoviruses led to greater severity, in comparison to the helper begomoviruses alone, and this supports a role for these intriguing molecules in the cassava mosaic disease complex.

The question remains as to what is the role of these unusual satellite DNAs, and what mechanisms lead to changes in disease phenotype and severity? The role of the long-known betasatellites in begomovirus epidemiology has been twofold, firstly, the betasatellite replace the movement function of DNA-B component of bipartite begomoviruses (Saeed *et al*., 2007); secondly, the betasatellite affects the replication of its helper virus by suppressing host gene silencing. In this study, the satellites also enhance symptoms, and in resistant cultivars, resistance is broken down. The

interaction of geminivirus, satellite DNAs (episomal and integrated) and host cultivar, and the proposed role of RNA silencing need further investigation.

Mixed virus infection in hosts is thought to influence resistance-breaking. Mixed infections make it possible for recombination to take place. Alongside the presence of satellites, recombination facilitates the appearance of the more severe begomovirus strains. The satellites are then thought to enhance the viability of the recombinant, thereby exacerbating disease. Such a situation has since been reported in the previously described disease situations in OLCD in Mali (Seal *et al*., 2006; Kon *et al*., 2009) and in CLCuD in Pakistan (Mubin *et al*., 2009). In the Tanzania's CMD situation, one of the helper begomovirus is indeed a recombinant, EACMV-UG. This indicates that perhaps the satellite DNAs enhance disease symptoms in a similar manner. However, disease severity of both the different helper viruses (EACMV-TZ and EACMV-UG), led to symptom enhancement. Cassava viruses are all chimeric fragments due to recombination, so this cannot be the only possible mechanism.

Another possible thinking is that the direct repeat of sequence elements CCGCCG, CCGCC and CGC in the GC-rich region of the cassava-associated satellite DNAs may interact with the replication associated protein *Rep* of the cognate helper begomovirus. This has previously been observed with TLCV Sat-DNA (Dry *et al.*, 1997), in which direct repeats of the sequence element TGGTGGT was thought to have a role in the capacity of the satellite to interact with the heterologous replication-associated protein. A similar mechanism can be responsible in the tendency of satellite DNA molecules to enhance disease symptoms.

6.2 Recommendations

The battle against the influence of satellites to exacerbate CMD symptoms cannot be won due to the limitations in the knowledge of these satellite DNAs. However some measures can be made in an attempt to reduce the effect and spread in the present pandemic, and prevent chances for occurrence of new epidemics similar to CMD in other crops. Strengthening of quarantine, the adoption of molecular markers to identify endogenous resistance genes in cassava germplasm, and the use of genetic

engineering to engineer resistance to the satellite and helper begomovirusvirus can be useful tools to control and monitor the pandemic.

6.2.1 Strengthening Plant Quarantine

Many of the earlier described satellites were associated with monopartite begomoviruses, which were widely spread in the Old World. In the New World only bipartite begomoviruses exist. The satellites may be present in the New World but they lack the helper viruses to replicate them. Also, long before the discovery of the satellite DNA molecules, the severe CMD pandemic in Tanzania was known to be restricted to the North-Western part of Tanzania (i.e. around Lake Victoria region) (Legg and Thresh, 2000). With the discovery of the satellites and the way the integrated form is widely spread, this could suggest that the responsible strain of the helper virus was inexistent in areas where CMD cases were mild. Much as it is known that infections of cassava plants by cassava mosaic begomoviruses may be cutting-derived or whitefly-derived (Sseruwagi, *et al*., 2004); and much as it is long known that the occurrence of CMD in areas where it was not known to occur was caused by use of infected planting material, a legislative strategy to disease management becomes indispensable. Domestic plant quarantine has served to a great extend in limiting the spread of the virulent EACMV UG from spreading to other areas of Tanzania. The detection of viruses and their associated satellites in imported or exported plant consignments including cassava, can potentially contribute to effective disease control.

Plant protection and quarantine (PPQ) need to strive to adapt to movement of plant material, not only with goal of avoiding disguised restriction to international trade (Meester *et al*., 1999), but also that which takes place locally. It is the critical role of National Plant Protection Organizations (NPPOs) in each country, including Tanzania, to re-align their policies with other countries to control the spread of diseases. The level of risk of spread of satellite DNAs and begomoviruses need to be managed by the country at a national level, to reduce the risk of spread between uninfected areas and provinces.

For a country as vast as Tanzania, the proposed model (Ndomba, 2005) comprises three key components:

- ⊕ Adherence to pre-export measures, with emphasis on pre-clearance inspection focusing on the capacity of the exporting country's NPPO to screen for plant viruses;
- ⊕ A thorough consignment inspection, comprising physical inspection for the presence of live whiteflies and presence of virus disease symptoms in the material;
- ⊕ Emphasis in post entry quarantine procedures, especially the capacity of the quarantine station to screen for viruses and associated satellites.

6.2.2 Use of Genetic Engineering for Protection

One of the popular strategies for virus resistance is the antisense or invert-repeat (RNA hairpins) technology, which targets incoming homologous viral sequences and silences them via TGS or PTGS (Prins *et al*., 2008). In the case of co-infection with begomovirus and satellites, the resistance must be durable and reliable. Antisense or RNA interference- mediated resistance can be an answer to the resistance-breaking satellites in cassava. This will involve careful designing of genetic constructs that will express negative sense RNA molecules or hairpin RNA to target the undesirable pathogenic molecules, helper virus and satellite DNA. The antisense molecule is expected to hybridize with the infecting virus sequence thereby interfering with replication process, while the hairpin RNA is cleaved to generate siRNAs (silencing RNAs). Once replication of the master virus is impeded, the replication of the associated satellites is also curtailed. However, since the role of the integrated satellite DNAs are not known, the GM strategy needs to be approached with caution.

6.2.3 Marker-Assisted Detection of Resistance-Associated Genes in Cassava Germplasm

The disease or resistance status of a crop, such as cassava, is controlled by the genetic interaction between the virus, satellite DNA and host. In gene-for-gene hypothesis, plants contain a single dominant resistance (R) gene, or a few (oligogernic resistance),

which specifically recognize viral pathogens that contain avirulence (*avr*) genes (McDowell and Woffenden, 2003; Collinge *et al*., 2008). Breakdown of dominant resistance usually requires a genetic change in either the host or the virus. The nature of resistnace in cassava has been speculated to be either recovery (innate immunity or RNA silencing), or specific dominant genes CMD2 or CMD 3 (Lokko *et al*., 2006; Bi *et al*., 2010). Molecular markers could be developed to locate desirable resistance genetic traits or indicate specific genetic differences between susceptible and resistance cassava cultivars. In order to use a specific marker to follow a specific trait, the marker must be found close enough to the gene of interest so that variation of both the marker and the gene can be inherited together during breeding. It is speculated that the integrated satellite DNAs may play a direct role in pathogenicity, or indirectly through effects on the genes with which they are associated.

Since these novel SatDNA-II and SatDNA-III molecules have only recently been discovered, and even more recently been found to be integrated in a large number of cassava germplasm, the most important recommendation is that fundamental research urgently be undertaken to understand these unique molecules and their role in pathogenicy of CMD of cassava.

6.3 References

1. Bi H., Aileni M. and Zhang P. (2010) Evaluation of cassava varieties for cassava mosaic disease resistance jointly by agro-inoculation screening and molecular markers. *African Journal of Plant Science* 4(9): 330-338.
2. Collinge B.C., Lund O.S., Thordal-Christensen H. (2008) What are the prospects for genetically engineered, disease resistant plants? *Eur. J. Plant Pathol*. 121:217–231.
3. Dry I.B., Krake L.R., Rigden J.E. and Rezaian M.A. (1997) A novel subviral agent associated with a geminivirus: The first report of a DNAsatellite. Division of Horticulture, Commonwealth Scientific and Industrial Research Organization, Hartley Grove, Urrbrae, South Australia 5064, Australia.
4. Kon T., Rojas M.R., Abdourhamane I.K. and Gilbertson R.L. (2009) Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *J. Gen virol.* 90: 1001–1013.

5. Legg, J.P. and Thresh, J.M. (2000) Cassava mosaic disease in East Africa: a dynamic disease in a changing environment. *Virus Res.* 71: 135-149.
6. Lokko Y., Dixon A.G.O., Offei S.K. and Danquah E.Y. (2006) Gene complimentarity of resistance to the cassava mosaic disease among African cassava accessions. *African Crop Science Journal* 14(3): 207 – 220.
7. McDowell J.M. and Woffenden B.J. (2003) Plant disease resistance genes: recent insights and potential applications. *Trends in Biotechnology* 21(4): 178 – 183.
8. Meester G., Woittiez R.D. and de Zeeuw A. (1999) Plants and Politics – On the occasion of the Centenary of the Netherlands Plant Protection Service. Wageningen Pers., Wageningen.
9. Mubin M., Briddon R.W. and Mansoor S. (2009) Diverse and recombinant DNA betasatellites are associated with begomovirus disease complex of *Digera arvensis*, a weed host. *Virus Research* 142: 208–212.
10. Nawaz-ul-Rehman M.S., Mansoor S., Briddon R.W. and Fauquet C.M. (2009) Maintenance of an Old World betasatellite by a New World helper begomovirus and possible rapid adaptation of the betasatellite. *J. Virol*. 1128: 795–809.
11. Ndomba O.A. (2005) Towards developing a systems approach for pest risk management in Tanzania: Challenges and Opportunities. *Trop. Pest. Mgt. Bull.* 3(6): 92–96.
12. Prins M., Laimer M., Noris E., Schubert J., Wassenegger M. and Tepfer M. (2008). Strategies for antiviral resistance in transgenic plants: A review. *Molecular Plant Pathology* 9 (1): 73–83.
13. Saeed M., Zafar Y., Randles J.W. and Rezaian M.A. (2007) A monopartite begomovirus-associated DNA-ß satellite substitutes for the DNA B of a bipartite begomovirus to permit systemic infection. *J. Gen Virol* 88: 2881 - 2889.
14. Seal S.E., vandenBosch F. and Jeger M.J. (2006) Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control. *Crit. Rev. Plant Sci*. 25 (1): 23–46.
15. Sseruwagi P., Sserubombwe W.S., Legg, J.P., Ndunguru, J. and Thresh, J.M. (2004) Methods of surveying the incidences and severity of cassava mosaic disease in Africa: A Review: *Virus Res.* 100: 129 – 142.

L I S T O F A P P E N D I C E S

Appendix 3.1: The 2817 bp - pCR8/GW/TOPO Sequence

CTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTC
GCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGC
AAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA
AAGCGGGCAGTGAGCGCAACGCAATTAATACGCGTACCGCTAGCCAGGAAGAGTTTGTAGAAACGC
AAAAAGGCCATCCGTCAGGATGGCCTTCTGCTTAGTTTGATGCCTGGCAGTTTATGGCGGGCGTCCTG
CCCGCCACCCTCCGGGCCGTTGCTTCACAACGTTCAAATCCGCTCCCGGCGGATTTGTCCTACTCAGG
AGAGCGTTCACCGACAAACAACAGATAAAACGAAAGGCCCAGTCTTCCGACTGAGCCTTTCGTTTTA
TTTGATGCCTGGCAGTTCCCT M13 Forward Site
ACTCTCGCGTTAACGCTAGCATGGATGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTCTTAA
GC GW1 Priming Site
TCGGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCGTTGCAACAAATTGATGAGCAA
TG *Eco*RI Insert *Eco*RI
CTTTTTTTATAATGCCAACTTTGTACAAAAAAGCAGGCTCCG↓AATTCGCCCTT↓AAGGGCG↓AATTCG
ACCCA GW2 Priming Site
GCTTTCTTGTACAAAGTTGGCATTATAAAAAATAATTGCTCATCAATTTGTTGCAACGAACAGGTCAC
TA T7 Promoter Site M13
TCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCCTATAGTGAGTCGTATTACATGGTC
AT Reverse Site
AGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATA
TATCATCATGCCTCCTCTAGACCAGCCAGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCC
GGGTGACGCACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCCTGTTCGGTTCGT
AAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTG
GTAACGGCGCAGTGGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATGCCTCGG
GCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTA
CGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAAACATCATGAGGGAAGCGGTGATCGCCGAAGTAT
CGACTCAACTATCAGAGGTAGTTGGCGTCATCGAGCGCCATCTCGAACCGACGTTGCTGGCCGTACA
TTTGTACGGCTCCGCAGTGGATGGCGGCCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTG
ACCGTAAGGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAAACTTCGGCTTCCC
CTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGTCACCATTGTTGTGCACGACGACATCATTCCGTG
GCGTTATCCAGCTAAGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGCAGGTATC
TTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGCTGACAAAAGCAAGAGAACATAGCGTTG
CCTTGGTAGGTCCAGCGGCGGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGCT
AAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGCGATGAGCGAAATGTAGTGCTT
ACGTTGTCCCGCATTTGGTACAGCGCAGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGACT
GGGCAATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAGACAGGCTTATCTTGG
ACAAGAAGAAGATCGCTTGGCCTCGCGCGCAGATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGC
GAGATCACCAAGGTAGTCGGCAAATAACCCTCGAGCCACCCATGACCAAAATCCCTTAACGTGAGTT
ACGCGTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTT
CTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATC
AAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGC
TAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACG
ATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGA
GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGA
AGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGC
TTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGA
TTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGT
TCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTT

Appendix 3.2: Sequences of isolates obtained from alignment report

EACMV-TZ[Tz108]

ATGCCACCTTCATCACCCTCCACGAGCCATTGTTCTCTAGTGCCCATCAAAGTCCAACACC
GCACAGCGAAGACCAGGGCCGTCAGACGTAGGCGGGTAGACCTCGAATGCGGCTGCTCGT
TTTATCTCCATATCGACTGCATCAACCATGGATTCTCGCACAGGGGCACTCATCACTGCGC
CTCAAGCAACGAATGGCGTTTTTATCTGGGAAATAATAAATCCCCTATATTTCGAAATCAC
CAACCACGACAAGCGGCCAGGGAACATGAACCACGACATCATCATACTCCAGATACGGTT
CAACCACAACCTCCGGAAGGCATTGGGGATTCACACGTGTTTTCTCAACTTCAAGGTCTGG
ACGACCTTACGGCCTCAGACTGGTCGTTTCTTAAGAGTATTTAGATATCAAGTGCTCAAGT
ATTTAGATATGATAGGCGTTATTTCCATTAACACTGTACTTCACGCTGTTGATCATGTTCTG
TACAATGTATTACTAAACACACTCCAAGTTACGGAGCAACATGC

EACMV-TZ [Tz113]

ATGCCACCTTCATCACCCTCCACGAGCCATTGTTCTCTAGTGCCCATCAAAGTCCAACACC
GCACAGCGAAGACCAGGGCCGTCAGACGTAGGCGGGTAGACCTCGAATGCGGCTGCTCGT
TTTATCTCCATATCGACTGCATCAACCATGGATTCTCGCACAGGGGCACTCATCACTGCGC
CTCAAGCAACGAATGGCGTTTTTATCTGGGAAATAATAAATCCCCTATATTTCGAAATCAC
CAACCACGACAAGCGGCCAGGGAACATGAACCACGACATCATCATACTCCAGATACGGTT
CAACCACAACCTCCGGAAGGCATTGGGGATTCACACGTGTTTTNTCAACTTCAAGGTCTGG
ACGACCTTACGGCCTCAGACTGGTCGTTTCTTAAGAGTATTTAGATATCAAGTGCTCAAGT
ATTTAGATATGATAGGCGTTATTTCCATTAACACTGTACTTCACGCTGTTGATCATGTTCTG
TACAATGTATTACTAAACACACTCCCAGTTACGGAGCAACATGC

Appendix 3.3: The 27 homologous sequences for begomoviruses

Bit score	% Identity	Name of isolate (all from segment DNA-A)
904	97	*East African cassava mosaic virus* isolate EACMV-[TZM]
904	97	*East African cassava mosaic Kenya virus* isolate EACMKV-[K308]
898	97	*East African cassava mosaic Kenya virus* isolate EACMKV-[K310]
893	96	*East African cassava mosaic Zanzibar virus* isolate EACMZV-[K18]
893	96	*East African cassava mosaic virus-Uganda variant* isolate EACMV-UG[K90]
887	96	*East African cassava mosaic virus-Uganda2 Mild* strain EACMV/Ug2//Mld
887	96	*East African cassava mosaic virus-Uganda variant* isolate EACMV-UG[K262]
881	96	*East African cassava mosaic virus-Uganda2 Severe* strain EACMV/Ug2//Svr
876	96	*East African cassava mosaic virus* isolate TZ10
876	96	*East African cassava mosaic virus-Uganda2* variant EACMV-UG2
870	96	*East African cassava mosaic virus* - [Kenya – K2B]
870	96	*East African cassava mosaic virus-KE2* segment DNA A, complete sequence, isolate EACMV-KE2[K201]
865	96	*East African cassava mosaic Zanzibar virus*
848	95	*East African cassava mosaic Zanzibar virus* - Kenya [Kilifi]
846	95	*South African cassava mosaic virus* isolate SACMV-[M12]
826	94	*East African cassava mosaic virus* isolate EACMV-[TZT]
826	94	*East African cassava mosaic virus* -[Tanzania] EACMV-TZ

821	94	*East African cassava mosaic virus* isolate EACMV-TZ[YV]
769	92	*East African cassava mosaic Malawi virus*-Malawi[MH], EACMMV-M[MH]
758	92	*East African cassava mosaic Malawi virus* isolate MK, EACMMV-M[MK]
354	74	*African cassava mosaic virus-Uganda* Severe strain ACMV/Ug//Svr
352	74	*African cassava mosaic virus*, ACMV-Nigeria[DNA1]
351	74	*African cassava mosaic virus* isolate CM/DO3 ACMV-[CM/D03]
349	74	*African cassava mosaic virus*-[Nigeria-Ogo] ACMV-[Nig-Ogo]
349	74	*African cassava mosaic virus*-[Ivory Coast], ACMV-[IC]
349	74	*African cassava mosaic virus*-[Cameroon-DO2] strain ACMV-CM/D02
343	74	*African cassava mosaic virus* isolate TZ ACMV-[TZ]
343	74	*African cassava mosaic virus-Uganda Mild* DNA-A strain ACMV/Ug//Mld

Appendix 3.4: Occurrence of integrated forms of SatDNA-II and SatDNA-III in Tanzania

S/N	Field Sample Identifiers	Location	CV	COORDINATES		Altitude (m. asl)	Satellite Score	
				Latitude	Longitude		Sat DNA II	Sat DNA III
1	1	Kikundi, Morogoro	Local	06º43.075'S	37º38.948'E	536	+	+
2	2	Kikundi	Local	06º43.075'S	37º38.948'E	536	+	+
3	3	Kikundi	Kibandameno	06º43.075'S	37º38.948'E	536	+	+
4	4	Kikundi	Kibandameno	06º43.075'S	37º38.948'E	536	+	+
5	5	Sokoine	Mwarusha	06º31.378'S	37º33.931'E	386	+	+
6	6	Sokoine	Kibandameno	06º31.378'S	37º33.931'E	386	+	+
7	7	Sokoine	Kibandameno	06º31.378'S	37º33.931'E	386	+	+
8	8	Kibaigwa, Dodoma	Improved	06º04.917'S	36º40.110'E	1226	+	+
9	9	Kibaigwa	Improved	06º04.917'S	36º40.110'E	1226	+	+
10	10	Kibaigwa	Improved	06º04.917'S	36º40.110'E	1226	+	+
11	11	Kibaigwa	Improved	06º04.917'S	36º40.110'E	1226	+	+
12	12	Nzega Ndogo, Tabora	Kaumba	04º06.054'S	33º03.984'E	1161	+	+
13	13	Nzega Ndogo	Kaumba	04º06.054'S	33º03.984'E	1161	+	+
14	14	Nzega Ndogo	Shija	04º06.054'S	33º03.984'E	1161	+	+
15	15	Singita, Shinyanga	Local	03º50.196'S	33º16.538'E	1152	+	+
16	16	Singita	Local	03º50.196'S	33º16.538'E	1152	+	+
17	17	Singita	Jatropha	03º50.196'S	33º16.538'E	1152	+	+
18	18	Rugauza	Lyongo-Red	03º29.777'S	33o33.144'E	1161	_	-
19	19	Rugauza	Lyongo-White	03º29.777'S	33º33.144'E	1161	+	+
20	20	Rugauza	Marekani	03º29.777'S	33º33.144'E	1161	+	+
21	21	Misungwi, Mwanza	Lyongo	02º48.408'S	33º03.968'E	1191	+	+
22	22	Misungwi	Lyongo	02º48.408'S	33o03.968'E	1191	+	+
23	23	Misungwi	Kigoma	02º48.408'S	33º03.968'E	1191	+	+
24	24	Misungwi	Lyongo	02º48.408'S	33º03.968'E	1191	_	+
25	25	Kamanga	Lyongo- Red	02º31.781'S	32º49.344'E	1143	+	+
26	26	Kamanga	Lyongo- Red	02º31.781'S	32º49.344'E	1143	+	+
27	27	Kamanga	Lyongo-Red	02º31.781'S	32º49.344'E	1143	_	+
28	28	Kachuko	Pananzala	02º32.069'S	32º49.056'E	1144	Nd	nd

29	29	Singaholo	Pananzala	02°40.121’S	32°36.840’E	1228	Nd	nd
30	30	Singaholo	Pananzala	02°40.121’S	32°36.840’E	1228	+	+
31	31	Ibanda	Lyongo Nyeupe	02°47.641’S	32°28.678’E	1191	+	+
32	32	Ibanda	Lyongo Nyeupe	02°47.641’S	32°28.678’E	1191	_	+
33	33	Ibanda	Lyongo-Nyeupe	02°47.641’S	32°28.678’E	1191	+	+
34	34	Geita, Geita	Lyongo-White	02°54.914’S	32°11.034’E	1255	+	+
35	35	Geita		02°54.914’S	32°11.034’E	1255	+	+
36	36	Geita		02°54.914’S	32°11.034’E	1255	+	+
37	37	Buseresere, Chato	Local	03°01.356’S	31°52.924’E	1243	+	+
38	38	Buseresere	Local	03°01.356’S	31°52.924’E	1243	+	+
39	39	Bwera	Lyongo- white	02°59.357’S	31°40.580’E	1212	+	+
40	40	Bwera	Lyongo-White	02°59.357’S	31°40.580’E	1212	+	+
41	41	Bwera	Lyongo-White	02°59.357’S	31°40.580’E	1212	+	+
42	42	Isambala	Local	02°52.315’S	31°32.834’E	1263	+	+
43	43	Isambala	Local	02°52.315’S	31°32.834’E	1263	+	+
44	46	Biharamulo-Lupondo	Local	02°39.615’S	31°19.743’E	1437	+	+
45	47	Biharamulo-Lupondo	Rushura	02°39.615’S	31°19.743’E	1437	+	+
46	48	Biharamulo-Lupondo	Local	02°39.615’S	31°19.743’E	1437	+	+
47	49	Ngalakambi	TMS 0001	02°47.481’S	31°01.667’E	1300	_	+
48	50	Ngalakambi	Rubona-Rumala	02°47.481’S	31°01.667’E	1300	+	+
49	51	Ngara-Kumuyange	Local	02°30.758’S	30°41.341’E	1474	_	-
50	52	Ngara-Kumuyange	Local	02°30.758’S	30°41.341’E	1474	+	+
51	53	Ngara-Kumuyange	Local	02°30.758’S	30°41.341’E	1474	+	+
52	54	Nguvu kazi	Rushura	02°31.203’S	30°51.371’E	1522	+	+
53	55	Nguvu kazi	Rushura	02°31.203’S	30°51.371’E	1522	+	+
54	56	Nguvu kazi	Ngunda	02°31.203’S	30°51.371’E	1522	Nd	nd
55	56	Nyaishozi, Karagwe	Rushura	01°46.315’S	31°08.098’E	1657	+	+
56	57	Nyaishozi	TMS 3001	01°46.315’S	31°08.098’E	1657	_	+
57	58	Omulushaka	Rushura	01°37.183’S	31°08.930’E	1682	+	+
58	59	Omulushaka	Local	01°37.183’S	31°08.930’E	1682	+	+
59	60	Kishoja	Local	01°24.541’S	31°11.806’E	1188	+	+
60	61	Kishoja	Local	01°24.541’S	31°11.806’E	1188	+	+
61	62	Kishoja	Local	01°24.541’S	31°11.806’E	1188	Nd	nd
62	62	Muleba-Kishoju	Rushura	01°52.424’S	31°38.752’E	1213	_	0
63	63	Muleba-Kishoju	Rushura	01°52.424’S	31°38.752’E	1213	+	+
64	64	Muleba-Kishoju	Rushura	01°52.424’S	31°38.752’E	1213	+	+
65	65	Kimwani, Mwanza	Kajaga	02°15.680’S	31°39.027’E	1156	+	+
66	66	Kimwani	Kajaga	02°15.680’S	31°39.027’E	1156	+	+
67	67	Itale	Pananzala	02°40.782’S	31°44.798’E	1177	+	_
68	68	Itale	Pananzala	02°40.782’S	31°44.798’E	1177	+	+
69	69	Itale	Pananzala	02°40.782’S	31°44.798’E	1177	+	-
70	70	Itale	Sumaili	02°40.782’S	31°44.798’E	1177	_	+
71	71	Itale	Sumaili	02°40.782’S	31°44.798’E	1177	+	+
72	72	Usagara	Lyongo	02°42.303’S	33°00.390’E	1224	_	-
73	73	Usagara	Lyongo	02°42.303’S	33°00.390’E	1224	_	+
74	74	Kigongo ferry	Local	02°42.849’S	32°53.623’E	1157	_	+
75	75	Kigongo ferry	Local	02°42.849’S	32°53.623’E	1157	+	+
76	76	Igembe, Ilemela	Lyongo	02°24.687’S	32°56.672’E	1133	_	+
77	77	Igembe	TMS 30001	02°24.687’S	32°56.672’E	1133	_	+
78	78	Nyakaboja	Lyongo	02°20.824’S	33°37.034’E	1139	+	+
79	79	Nyakaboja	Kachaga	02°20.824’S	33°37.034’E	1139	+	-

80	80	Nyakaboja	Kachaga	02°20.824’S	33°37.034’E	1139	+	+
81	81	Sitakishari, Mkinga	Kibandameno	04°52.161’S	39°06.661’E	965	+	+
82	82	Sitakishari	Kibandameno	04°52.161’S	39°06.661’E	965	+	+
83	83	Pitanazako	Mwarabu	04°59.656’S	39°04.652’E	66	+	+
84	84	Gombe	Mwarabu	04°57.298’S	38°57.411’E	123	+	+
85	85	Gombe	Kibandameno	04°57.298’S	38°57.411’E	123	+	+
86	86	Gombe	Maiza	04°57.298’S	38°57.411’E	123	+	+
87	87	Maramba	Maiza	04°51.246’S	38°48.463’E	222	+	+
88	88	Maramba	Maiza	04°51.246’S	38°48.463’E	222	Nd	nd
89	89	Maramba	Maiza	04°51.246’S	38°48.463’E	222	+	+
90	90	Maramba	Maiza	04°51.246’S	38°48.463’E	222	+	+
91	91	Kijungu moto, Korogwe	Paja la mzee	04°45.458’S	38o39.226’E	123	+	+
92	92	Kijungu moto	Kibandameno	04°45.458’S	38°39.226’E	123	+	+
93	93	Maguzoni, Muheza	Kibandameno	05°15.537’S	38°38.077E	261	+	+
94	94	Maguzoni	Kibandameno	05°15.537’S	38°38.077’E	261	+	+
95	95	Changombe, Handeni	Maiza	05°28.350’S	33°29.563’E	372	+	+
96	96	Changombe	Maiza	05°28.350’S	33°29.563’E	372	+	+
97	97	Changombe	Maiza	05°28.350’S	33°29.563’E	372	+	+
98	98	Changombe	Maiza	05°28.350’S	33°29.563’E	372	+	+
99	99	Kisaza	Maiza	05°35.105’S	38°25.882’E	339	+	+
100	100	Kisaza	Maiza	05°35.105’S	38°25.882’E	339	+	+
101	101	Mkata	Kibandameno	05°47.680’S	38°17.465’E	421	_	+
102	102	Mkata	Namikonga	05°47.680’S	38°17.465’E	421	Nd	nd
103	103	Kwang'andu, Bagamoyo	Local	05°57.578’S	38°13.456’E	380	+	+
104	104	Kwang'andu	Local	05°57.578’S	38°13.456’E	380	_	-
105	105	Kwadisenyi	Mnazi	06°07.499’S	38°17.899’E	337	_	-
106	106	Kwadisenyi	Kibandameno	06°07.499’S	38°17.899’E	337	_	-
107	107	Lugoba	Local	06°27.892’S	38°19.632’E	291	_	+
108	108	Lugoba	Local	06°27.892’S	38°19.632’E	291	+	+
109	109	Lugoba	Local	06°27.892’S	38°19.632’E	291	+	+
110	110	Kibiti, Lindi	Kiroba	07°47.462’S	38°56.992’E	65	+	-
111	111	Kibiti	Local	07°47.462’S	38°56.992’E	65	+	-
112	112	Kibiti	Local	07°47.462’S	38°56.992’E	65	+	-
113	113	Muhoro	Kiroba	08°08.370’S	39°09.273’E	4	Nd	nd
114	114	Muhoro	Kiroba	08°08.370’S	39°09.273’E	4	_	-
115	115	Somanga, Kilwa	Kigoma	08°08.366’S	39°09.275’E	5	_	-
116	116	Somanga	Kiroba	08°08.366’S	39°09.275’E	5	+	+
117	117	Somanga	Mbuyu	08°08.366’S	39°09.275’E	5	_	-
118	118	Matandu	Kiroba	08°42.882’S	39°15.569’E	30	_	-
119	119	Matandu	Kiroba	08°42.882’S	39°15.569’E	30	+	-
120	120	Matandu	Kiroba	08°42.882’S	39°15.569’E	30	+	-
121	121	Mtwara town, Mtwara Mikindani	Local	10°15.996’S	40°03.370’E	124	+	-
122	122	Mtwara town	Local	10°15.996’S	40°03.370’E	124	_	+
123	123	Mtwara town	Local	10°15.996’S	40°03.370’E	124	+	-
124	124	Mtwara town	Local	10°15.996’S	40°03.370’E	124	+	-
125	125	Mtwara town	Mreteta	10°07.767’S	39°35.916’E	56	_	-
126	126	Mtwara town	Mreteta	10°07.767’S	39°35.916’E	56	+	-
127	127	Mtwara town	Mreteta	10°07.767’S	39o35.916’E	56	_	+
128	128	Mumbulu, Masasi	Kigoma-mafia	10°25.747’S	39°07.782’E	56	+	+
129	129	Mumbulu	Kigoma-mafia	10°25.747’S	39°07.782’E	56	+	+
130	130	Mumbulu	Kigoma-mafia	10°25.747’S	39°07.782’E	56	_	-
131	131	Malamba	Chimaje	10°41.718’S	38°50.644’E	395	_	+
132	132	Malamba	Liumbukua	10°41.718’S	38°50.644’E	395	_	-

133	133	Namatumbuso, Nanyumbu	Liumbukua	10°49.465’S	38°41.308’E	396	_	+
134	134	Namatumbuso	Liumbukua	10°49.465’S	38°41.308’E	396	_	-
135	135	Mangaka	Local	10°36.623’S	38°18.683’E	408	+	+
136	136	Mangaka	Local	10°36.623’S	38°18.683’E	408	+	+
137	137	Mangaka	Local	10°36.623’S	38°18.683’E	408	+	+
138	138	Nangomba, Tunduru	Dide	10°58.537’S	37°16.576’E	647	_	+
139	139	Nangomba	Dide	10°58.537’S	37°16.576’E	647	+	-
140	140	Milonde	Yuda	10°45.434’S	36°55.723’E	813	_	-
141	141	Milonde	Mwaya	10°45.434’S	36°55.723’E	813	+	+
142	142	Milonde	Malisaidi	10°45.434’S	36°55.723’E	813	+	+
143	143	Kilimasela, Namtumbo	Dide	10°45.334’S	36°30.704’E	972	_	-
144	144	Kilimasela	Mweda	10°45.334’S	36°30.704’E	972	+	+
145	145	Namtumbo	Dide	10°28.254’S	36°08.882’E	910	_	+
146	146	Namtumbo	Dide	10°28.254’S	36o08.882’E	910	+	+
147	147	Msumbiji, Songea	Mkuzu	10°40.709’S	35°28.472’E	1152	_	-
148	148	Msumbiji	Farajala	10°40.709’S	35°28.472’E	1152	_	+
149	149	Maposeni-peramiho	Mwaya	10°34.848’S	35°85.942’E	1016	+	+
150	150	Maposeni-peramiho	Mwaya	10°34.848’S	35°85.942’E	1016	+	+
151	149	Peramiho	Mkongo mwaya	10°37.411’S	35°27.190’E	1029	_	-
152	150	Peramiho	Mkongo mwaya	10°37.411’S	35°27.190’E	1029	+	+
153	151	Peramiho A	Mwaya	10°37.780’S	35°26.229’E	1000	_	-
154	152	Peramiho A	Mwaya	10°37.780’S	35°26.229’E	1000	+	+
155	153	Matogoro	Mwaya	10°41.304’S	35°41.097’E	1228	+	+
156	MB01M0	Kilimani, Mbinga	Tupuka	10°58.579’S	35°02.185’E	1264	Nd	nd
157	MB02S0	Kilimani	Tupuka	10°58.579’S	35°02.185’E	1264	+	+
158	MB03S0	Myangayanga	Tupuka	10°56.642’S	34°57.508’E	1347	+	-
159	MB04S0	Unyoni	Nakalai	11°04.038’S	34°52.694’E	1616	_	-
160	MB05S0	Mawono	Tupuka	11°04.060’S	34°51.186’E	1651	Nd	nd
161	MB06S0	Mawono	Tupuka	11°04.060’S	34°51.186’E	1651	Nd	nd
162	MB07S1	Bula (Mapera)	Nakalai	11°04.318’S	34°51.042’E	1571	Nd	nd
163	MB07S2	Bula (Mapera)	Nakalai	11°04.318’S	34°51.042’E	1571	+	+
164	MB07S3	Bula (Mapera)	Nakalai	11°04.318’S	34°51.042’E	1571	Nd	nd
165	MB08S0	Mpapa - Arusha	Tupuka	11°08.025’S	34°56.131’E	1407	Nd	nd
166	MB09S1	Gulio (Kingerikiti)	Tupuka	11°13.610’S	34°58.822’E	1401	Nd	nd
167	MB09S2	Gulio (Kingerikiti)	NA				+	-
168	MB10S1	Pisi (Kingerikiti)	NA	11°13.665’S	34°58.914’E	1299	+	+
169	MB10S2	Pisi (Kingerikiti)	NA				_	-
170	MB11S1	Mbanga (Lumeme)	Nakalai	11°15.109’S	35°00.223’E	1278	+	+
171	MB11S2	Mbanga (Lumeme)					+	-
172	MB12S1	Matarawe (Lumeme)	Kausa	11°16.714’S	35°00.778’E	1313	+	+
173	MB12S2	Matarawe (Lumeme)	Kausa	11°16.714’S	35°00.778’E	1313	+	-
174	MB13S1	Litindo	Nakalai	11°11.263’S	34°59.756’E	1385	+	+
175	MB13S2	Litindo	Tupuka	11°11.263’S	34°59.756’E	1385	+	+
176	MB13S3	Litindo	Tupuka	11°11.263’S	34°59.756’E	1385	+	-
177	MB13S4	Litindo	Tupuka	11°11.263’S	34°59.756’E	1385	+	+
178	MB14S1	Luhangarasi	Tupuka	11°08.137’S	34°58.102’E	1422	Nd	nd
179	MB14S2	Luhangarasi					+	-
180	MB14S3	Luhangarasi					-	-
181	MB15S0	Luhangarasi	Tupuka	11°08.137’S	34°58.102’E	1422	+	-
182	MB16S1	Mpapa	Tupuka	11°08.043’S	34°57.917’E	1369	-	+
183	MB16S2	Mpapa					+	+
184	MB16S3	Mpapa					-	+
184	MB16S4	Mpapa					+	+

186	MB17S0	Mikalanga	Nakalai	11°06.673'S	34°50.284'E	1517	+	+
187	MB18S0	Ilela	Tupuka	11°09.786'S	34°52.319'E	1595	+	-
188	MB19S0	Ilela					+	-
189	MB20S1	Mikalanga (Baraza)					+	+
200	MB20S4	Mikalanga (Baraza)					+	+
201	MB21S1	Mikalanga (Baraza)					-	-
202	MB21S2	Mikalanga (Baraza)					+	+
203	MB22S0	Makuli (Mapera)					+	-
204	MB24S1	Mbalama (Mapera)	Nakalai	11°04.722'S	34°50.629'E	1507	+	+
205	MB24S2	Mbalama (Mapera)					+	+
206	MB25M0	Kinzege (Matekela)					+	-
207	MB26M0	Kwa Kabanila					-	1
208	MB27S1	Kwa Tai	Nakalai	11°04.932'S	34°50.536'E	1515	-	-
209	MB27S2	Kwa Tai	Nakalai				1	-
210	MB29S1	Luhalala	Nakalai	11°05.334'S	34°49.827'E	1431	-	+
211	MB29S2	Luhalala	Nakalai	11°05.334'S	34°49.827'E	1431	+	+
212	MB30S0	Luhalala	Nakalai	11°05.334'S	34°49.827'E	1431	-	-
213	MB31S1	Luhalala	Nakalai	11°05.334'S	34°49.827'E	1431	-	+
214	MB31S2	Luhalala	Nakalai	11°05.334'S	34°49.827'E	1431	+	+
215	MB32S1	Malendi	Nakalai	11°05.568'S	34°49.291'E	1373	+	+
216	MB32S2	Malendi	Nakalai	11°05.568'S	34°49.291'E	1373	Nd	nd
217	MB32S3	Malendi	Nakalai	11°05.568'S	34°49.291'E	1373	-	-
218	154	Mapera	Nakalai	11°04.473'S	34°50.012'E	1472	+	+
219	155	Mapera	Nakalai	11°04.473'S	34°50.012'E	1472	+	+
220	156	Mapera	Nakalai	11°04.473'S	34°50.012'E	1472	+	+
221	157	Kibandai	Nakalai	11°01.738'S	34°47.712'E	1543	+	+
222	158	Kibandai	Nakalai	11°01.738'S	34°47.712'E	1543	+	+
223	159	Kibandai	Nakalai	11°01.738'S	34°47.712'E	1543	+	+
224	160	Kipapa	Mwaya	11°00.604'S	34°44.412'E	1518	+	+
225	161	Kipapa	Mwaya	11°00.604'S	34°44.412'E	1518	-	+
226	162	Kipapa	Mwaya	11°00.604'S	34°44.412'E	1518	+	+
227	163	Maguu	Nakalai	11°02.746'S	34°47.643'E	1499	-	+
228	164	Maguu	Nakalai	11°02.746'S	34°47.643'E	1499	-	+
229	165	Maguu	Nakalai	11°02.746'S	34°47.643'E	1499	+	+
230	167	Mbanda	Nakalai	11°04.397'S	34°47.906'E	1506	+	+
231	168	Lipokela, Songea	Gamani	10°42.364'S	35°16.802'E	1506	+	+
232	169	Lipokela	Gamani	10°42.364'S	35°16.802'E	1506	-	-
233	170	Hanga	Local	10°12.088'S	35°40.279'E	851	-	+
234	171	Hanga	Kagunila	10°12.088'S	35°40.279'E	851	-	-
235	172	Hanga	Kagunila	10°12.088'S	35°40.279'E	851	-	+
236	174	Madaba	Local	09°52.426'S	35°23.297'E	1172	+	+
237	175	Madaba	Local	09°52.426'S	35°23.297'E	1172	-	-
238	176	Tanangozi, Iringa	Asunami	07°53.456'S	35°35.947'E	1668	+	+
			TOTAL				**151/222**	**158/222**

nd = not determined

Appendix 3.5: Protocol for the cloning of PCR products

Requirements

- Microfuge tubes (1.5 μL),
- DNA template (PCR product),
- Plasmid vector (pGEM-T Easy) (Promega, USA),
- Ligation buffer,
- DNA ligase,
- Frozen competent cells,
- SOC medium,
- IPTG solution (Prepared as indicated below),
- X-Gal solution (Prepared as indicated below).

Preparation of IPTG (20%)

Dissolve 2 g of IPTG in 8 mL of sterile distilled water and adjust volume to 10 mL. Sterilize by passing it through 0.22 μn disposable filter. Divide into 1 mL aliquots and store at -20°C.

Preparation of X-Gal (2%)

Dissolve 0.1g X-gal into 5 mL N,N-dimethylformamide. Aliquot and store at -20°C.

Procedure

(i) Prepare LB agar containing ampicillin (100 μL/200 mL media), IPTG (100 μL/200mL media) and X-gal (400 μL/200mL media).

(ii) Label microfuge tubes up to the number of samples desired. Set ligation reaction by adding the following components individually (do not prepare cocktail mix):

- 2x Rapid Ligation buffer 7.5 μL
- pGM-T Easy vector 1.0 μL
- PCR product 4.0 μL
- T4 DNA ligase 1.0 μL
- Sterile distilled water 0.5 μL

Incubate reactions at 4°C overnight.

(iii) Shortly centrifuge tubes containing ligation reaction to collect contents at the bottom of the tube.

(iv) Add 3 μL of each ligation reaction to a sterile 1.5 μL microcentrifuge tube on ice. Set up an extra tube for the determination of transformation efficiency of the competent cells.

(v) Remove tubes of the frozen JM109 High Efficiency Competent cells (Stratagene, USA) from the -80°C storage and place them on water bath until just thawed. Mix cells by gently flicking the tube.

(vi) Carefully transfer 30 μL of JM109 competent cells into each reaction tube. (Serve some 100 μL of the mix for determination of transformation efficiency). Gently flick the tubes to mix and place them on ice for 20 minutes.

(vii) Heat shock the cells for 45 seconds in water bath at exactly 42°C (do not shake). Immediately return to ice and keep there for 2 minutes.

(viii) Add 950 μL of room temperature SOC medium to the tubes and 900 μL to the tube containing cells transformed with uncut plasmid.

(ix) Incubate tubes for 1.30 minutes at 37°C with shaking at 150 rpm to enhance aeration.

(x) Remove tubes from incubator and centrifuge them at maximum speed (10,000 rpm) for 10 minutes.

(xi) Discard some amount of supernatant to leave about 150 μL in the tube. Vortex briefly to re-suspend the pellet.

Culturing procedure and blue-white colour screening

1. Plate 100 μL of each transformation culture onto duplicate LB/ampicillin/X-Gal plates. The uncut control is diluted with SOC media in 1:10 ratio before plating. Prepare also a master plate.
2. Incubate plates at 37°C overnight without shaking. The master plate is stored at 4°C.
3. Pre heat the master plate (LB agar) at 37oC.
4. Add (2%) ampicillin to LB broth (400 μL per 200 mL broth). Aliquot 3 mL of LB broth into capped test tubes

5. Using sterile pipette tips transfer a single white colony from bacterial culture plate, strip onto the master plate and dip the tube into test tubes containing LB broth and ampicillin.
6. Incubate the LB broth tubes at 37°C with maximum shaking over night. Incubate the master plate at 37°C without shaking overnight.
7. The remaining bacteria on the culture plate are stored at 4oC for further use.

Plasmid DNA purification using Qiagen Spin miniprep kit

a. Pellet bacterial cells by centrifugation (13000 rpm for 10 minutes) and discard supernatant.
b. Resuspend pelleted bacterial cells in 250 μL of suspension buffer (P1) by votexing.
c. Add 250 μL of P2 buffer and mix thoroughly by inverting the tube 4 – 6 times until a clear solution appears.
d. Add 350 μL of N3 buffer and mix immediately and thoroughly by inverting the tube 4–6 times to obtain a cloudy solution.
e. Centrifuge tubes for 10 minutes at 13,000 rpm. Serve the supernatant to QIAprep spin column by decanting or pipetting.
f. Centrifuge for 1minute and discard flow-through.
g. Wash QIAprep spin column by adding 750 μL of PE buffer, centrifuge for 1 minute and discard flow through.
h. Centrifuge for additional 1 minute to remove residual wash buffer
i. Transfer QIAprep column into a clean 1.5 mL microcentrifuge tube. Elute DNA by adding 50 μL of elution buffer (EB) (10 mM Tris HCl, pH 8.5 or water) to the center of each QIAprep spin column. Allow to satand for 1 min and centrifuge for 1 minute.
j. Verify presence of insert in the plasmid by running the plasmid DNA onto agarose gel.

Restriction analysis

A single digest with recommended restriction enzyme (EcoRI) was used to release inserts cloned into the pGEM T Easy vector. Label clean eppendorf tubes according to the number of samples. Prepare cocktail mix as follows:

Component	Amount per reaction (µL)	x 25 reactions (µL)
sdH_2O	14.0	350
Buffer	2.0	50
EcoRI	1.0	25
DNA	3.0	-

Add 17 µL of the reaction mix to each tube and 3 µL of DNA. Incubate at 37°C for 1 hour. Run products on 1% agarose gel.

Appendix 3.6: Portion of SatDNA-II isolates alignment report to show GC-rich regions

```
    G T A T C A C G C G A T T G T G A G C C          Majority
    ----------------      +-----------------      +
                          90                    100
    ----------------      +-----------------      +
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II (AY836366)
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ1
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ2
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ3
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ4
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ5
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ6
 81 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ7
 80 G T A T C A C G C G A T T G T G A G C C          SatDNA-II TZ8

    G C C G A C T G C C C G C C G C A C G C          Majority
    ----------------      +-----------------      +
                          110                   120
    ----------------      +-----------------      +
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II (AY836366)
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ1
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ2
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ3
100 G C C G A C A G C C C G C C G C A C G C          SatDNA-II TZ4
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ5
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ6
101 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ7
100 G C C G A C T G C C C G C C G C A C G C          SatDNA-II TZ8

    C C G C A C C T C T G C A T C A A C T G          Majority
    ----------------      +-----------------      +
                          130                   140
```

```
     ----------------       +-----------------       +
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II (AY836366)
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ1
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ2
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ3
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ4
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ5
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ6
121  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ7
120  C C G C A C C T C T G C A T C A A C T G          SatDNA-II TZ8
```

Appendix 3.7: Portion of SatDNA-II isolates alignment report to show Poly A Signal

```
     C A T G A G T T T G A G T G C T T C C A          Majority
     ----------------       +-----------------       +
                          510                     520
     ----------------       +-----------------       +
497  C A T G A G T T T G A G T G C T T C C A          SatDNA-II (AY836366)
489  C A T C T G C T G G T C T G C T T C G A          SatDNA-II TZ1
494  C A T G A G T T T G A G T G C T T C C A          SatDNA-II TZ2
489  C A T C T G C T G G T C T G C T T C G A          SatDNA-II TZ3
494  C A T G A G T T T G A G T G C T T C C A          SatDNA-II TZ4
492  C A T G A G A T T G A G A G CG T C C A           SatDNA-II TZ5
494  C A T G A G T T T G A G T G C T T C C A          SatDNA-II TZ6
495  C A T G A G T T T G A G T G C T T C C A          SatDNA-II TZ7
494  C A T G A G T T T G A G T G C T T C C A          SatDNA-II TZ8

     A A A T T T T T T A A T A A A A G C T T          Majority
     ----------------       +-----------------       +
                          530                     540
     ----------------       +-----------------       +
517  A A A T T T T T T  A A T A A A A G C T T         SatDNA-II (AY836366)
509  C A A T A G T T G G A T A C G A G C G A          SatDNA-II TZ1
514  A A A T T T T T T  A A T A A A A G C T T         SatDNA-II TZ2
509  C A A T A G T T G G A T A C G A G C G A          SatDNA-II TZ3
514  A A A T T T T T T  A A T A A A A G C T T         SatDNA-II TZ4
512  A A A T T T T T T  A A T A T A A G C T T         SatDNA-II TZ5
514  A A A T T T T T T  A A T A A A A G C T T         SatDNA-II TZ6
515  A A A T T T T T T  A A T A A A A G C T T         SatDNA-II TZ7
514  A A A T T T T T T  A A T A A A A G C T T         SatDNA-II TZ8

     T A C A A A G C T C A G C T T G G A T C          Majority
     ----------------       +-----------------       +
                          550                     560
```

```
     ------------------      +-------------------      +
537  T A C A A A G C T C A G C T T G G A T C          SatDNA-II (AY836366)
529  T A C A A A A C T C C G A T G A G A T C          SatDNA-II TZ1
534  T A C A A A G C T C A G C T T G G A T C          SatDNA-II TZ2
529  T A C A A A G C T C C G A T G A G A T C          SatDNA-II TZ3
534  T A C A A A G C T C A G C T T G G A T C          SatDNA-II TZ4
532  T A C A C A G C T C A C C T T G T A T C          SatDNA-II TZ5
534  T A C A A A G C T C A G C T T G G A T C          SatDNA-II TZ6
535  T A C A A A G C T C A G C T T G G A T C          SatDNA-II TZ7
534  T A C A A A G C T C A G C T T G G A T C          SatDNA-II TZ8
```

Appendix 3.8: Portion of SatDNA-II isolates alignment report to show TATA box

```
     T T A T T T C T C T T G A A T T T - - G          Majority
     ------------------      +-------------------      +
                          790                     800
     ------------------      +-------------------      +
759  T T A T T T C T C T T G A A T T T - - G          SatDNA-II (AY836366)
755  T T A T T T C T C T T G A A T T T - - G          SatDNA-II TZ1
756  T T A T T T C T C T G G A A T T T - - G          SatDNA-II TZ2
753  T T A T T T C T C T T G A A T T T -  - G         SatDNA-II TZ3
771  T T A T T T C T C T T G G A A T T T G G          SatDNA-II TZ4
754  A T A T T T C T C T C G A G T T T - - G          SatDNA-II TZ5
765  T T A T T T C T C T T G A A T T T  - - G         SatDNA-II TZ6
765  T T A T T T C T C T T G A A T T T - -  G         SatDNA-II TZ7
756  T T A T T T C T C T T G A A T T T - -  G         SatDNA-II TZ8

     G T A T C T G C A A T T T A T G T A T A          Majority
     ------------------      +-------------------      +
                          810                     820
     ------------------      +-------------------      +
777  G T A T C T G C A A T T T A T G T A T A          SatDNA-II (AY836366)
773  G T A T C T G C A A T T T A T G T A T A          SatDNA-II TZ1
774  G T A T C T G C A A T T T A T G T A T A          SatDNA-II TZ2
771  G T A T C T G C A A T T T A T G T A T A          SatDNA-II TZ3
791  G T A T C T G C A A T T T A T G T A T A          SatDNA-II TZ4
772  G G A T C T G C G A T T T A T G T G T A          SatDNA-II TZ5
783  G T A T C T G C A A T T T A T G T A T A          SatDNA-II TZ6
783  G T A T C T G C A T T T A T  T G T A T A         SatDNA-II TZ7
774  G T A T C T G C A A T T T A T G T A T A          SatDNA-II TZ8

     A A T C C C - T A G C A G A A T A T T T          Majority
     ------------------      +-------------------      +
                          830                     840
     ------------------      +-------------------      +
797  A A T C C C - T A G C A G A A T A T T T          SatDNA-II (AY836366)
793  A A T C C C - T A G C A G A A T A T T T          SatDNA-II TZ1
```

```
794 A T T C C C - T A G C C G A A T A T T T      SatDNA-II TZ2
791 A A T C C C - T A G C A G A A T A T T T      SatDNA-II TZ3
811 A A T C C C - T A G C A G A A T A T T T      SatDNA-II TZ4
792 A A T C C C C T A G C A G A A T A T A T      SatDNA-II TZ5
803 A A T C C C - T A G C A G A A T A T T T      SatDNA-II TZ6
803 A A T C C C - T A G C A G A A T A T T T      SatDNA-II TZ7
794 A A T C C C - T A G C A G A A T A T T T      SatDNA-II TZ8
```

Appendix 3.9: Portion of SatDNA-III isolates alignment report to show GC-rich regions

```
    A A C T T C C G C C T C C C G T C C G A      Majority
    ----------      +----------      +
                   150              160
    ----------      +----------      +
138 G A C G T C C G C C T T C C G  T C C G A     SatDNA-III (AY836367)
 97 A A C T C C C  G C C C C C C G  T C C A A     SatDNA-III TZ1
 54 C A C C T C C  C C C T G G  C G T C A G A     SatDNA-III TZ2
 58 A A C T T C C  G C C T C C  C C  C C C G A     SatDNA-III TZ3
 55 A A C T T C C  C C C C C C  C C  T C C A A     SatDNA-III TZ4
    T T G T C T C  T C T G T T C T T C A T G      Majority
    ----------      +----------      +
                   170              180
    ----------      +----------      +
158 T T G T C C C T C T G C T C T T C A T  G      SatDNA-III (AY836367)
117 T T G T C T T T G T G T G C T T C T G  G      SatDNA-III TZ1
 74 T T G C C T C T C C G C T C T T A T T  G      SatDNA-III TZ2
 78 T T G G C T C T T T G G T G C T C A T G      SatDNA-III TZ3
 75 T G G T C C C T C G G T T C T T C A G G      SatDNA-III TZ4

    C T G C C A A C G C C A T T G C T G C A      Majority
    ----------      +----------      +
                   190              200
    ----------      +----------      +
178 C T G T C A A C G C C A T T G C T G C A      SatDNA-III (AY836367)
137 T T G C C A C  C G C C T T T G T G G C T      SatDNA-III TZ1
 94 C C C C C C A G C G C A T T G C G G C A      SatDNA-III TZ2
 98 C T G T C A A G C G C A T T G C T G C A      SatDNA-III TZ3
 95 T T G C C A A  C G C C T T G G T T G C T      SatDNA-III TZ4

    T C C G G G G C T C G T T G C T G C G T      Majority
    ----------      +----------      +
                   210              220
    ----------      +----------      +
198 T C C G G T G C T C G T T G C T G C G T      SatDNA-III (AY836367)
157 T C G G G G G C C C G T G G C T G C G T      SatDNA-III TZ1
114 T C A G G T G C T C G T T G C T G C C T      SatDNA-III TZ2
118 T T C G G G G G T C G T T G G T G C G T      SatDNA-III TZ3
```

115 T C C G G G G T C C G T G G C T G C G T SatDNA-III TZ4

Appendix 4.1: Conditions which influenced growth of stakes in the greenhouse

DATE	TIME (GMT+2)	TEMPERATURE (°C)			REMARKS			
		MIN	MAX	ACTUAL	WATERING	HEATER	FAN	WETWALL
2.04.08	09h00	11	45	29		-		
	15h17	25	46	45	No	No	No	No
3.04.08	07h48	21	46	18	Yes	No	No	No
	14h43	21	45	33	No	No	No	No
4.04.08	12h30	15	45	42	Yes	No	No	No
	15h00	34	43	37	No	No	No	No
5.04.08	11h00	16	43	38	Yes	No	No	No
7.04.08	08h00	14	47	25	No	No	No	No
7.04.08	13h55	25	44	40	No	No	No	No
8.04.08	10h15	12	44	40	Yes	No	Yes	No
	15h40	31	40	31	No	No	No	No
09.04.08	09h15	13	37	25	No	No	No	No
10.04.08	09h43	15	45	35	Yes	No	Yes	No
	14h22	13	35	24	No	No	No	No
	15h52	24	35	28	No	Yes	No	No
11.04.08	10h00	15	38	38	No	No	Yes	No
	16h52	19	38	19	No	Yes	No	No
12.04.08	09h15	15	30	30	Yes	No	Yes	No
	15h32	22	34	21	Yes	Yes	No	No
13.04.08	12h05	12	42	42	No	No	No	No
14.04.08	12h05	12	45	35	No	Yes	No	No
15.04.08	08h53	12	47	13	No	No	No	No
16.04.08	08h44	11	45	29	Yes	No	No	No
17.04.08	09h19	14	48	36	Yes	No	No	No
18.04.08	11h50	14	49	42	No	No	Yes	Yes
	17h24	18	35	18	No	Yes	No	No
19.04.08	10h37	15	39	39	Yes	No	Yes	No
	15h16	31	43	41	No	No	No	No
21.04.08	12h38	8	41	21	No	No	No	No
	16h46	9	41	24	Yes	Yes	No	No
22.04.08	08h40	8	24	23	No	No	No	No
	13h23	25	28	24	Yes	No	Yes	Yes
	16h25	16	25	16	Yes	No	No	No
23.04.08	09h53	7	37	37	No	No	Yes	Yes
	15h34	25	37	25	No	Yes	No	No
24.04.08	09h22	9	34	31	Yes	Yes	No	Yes
	16h30	24	36	29	Yes	No	No	No
25.04.08	09h46	9	35	35	No	No	Yes	Yes
	16h11	23	29	23	Yes	Yes	No	No
26.04.08	08h45	24	40	40	No	No	Yes	Yes
	17h07	19	40	19	No	Yes	No	No
27.04.08	08h50	19	45	45	No	No	Yes	Yes

	16h11	23	32	23	No	No	No	No
28.04.08	09h19	13	31	31	No	No	Yes	Yes
	15h40	24	33	26	No	No	No	No
29.04.08	08h22	25	35	31	No	No	Yes	Yes
	13h25	22	31	30	No	No	No	No
	16h30	20	32	20	No	Yes	No	No
30.04.08	09h57	20	35	35	No	No	Yes	Yes
	13h58	24	35	30	No	No	No	No
	15h39	30	36	34	No	Yes	No	No
1.05.08	09h45	28	40	40	No	No	No	No
	16h18	30	41	32	Yes	Yes	No	No
2.05.08	11h01	19	34	25	No	No	No	No
	16h00	27	32	22	No	No	No	No
3.05.08	12h55	18	28	22	Yes	No	No	No
4.05.08	14h11	17	22	21	No	No	No	No
5.05.08	09h00	17	34	34	No	No	No	No
	14h55	34	42	41	Yes	No	No	No
	15h41	38	41	38	No	Yes	No	No
6.05.08	09h14	23	42	39	No	No	No	No
	16h00	37	45	39	No	Yes	No	No
7.05.08	08h00	23	39	31	No	No	No	No
	10h18	31	39	39	No	No	Yes	No
	14h48	31	38	35	No	No	No	No
8.05.08	09h40	12	35	35	Yes	No	Yes	No
	13h15	30	38	30	No	No	No	No
	10h24	31	44	31	No	Yes	No	No
9.05.08	07h31	23	37	29	No	No	Yes	No
	14h41	22	31	30	No	No	No	No
	16h10	26	30	26	No	Yes	No	No
10.05.08	08h10	25	35	35	Yes	No	Yes	No
	14h12	23	39	25	No	Yes	No	No
11.05.08	09h22	24	38	38	No	No	Yes	No
	13h13	29	38	33	Yes	No	No	No
12.05.08	07h35	10	37	10	No	No	No	No
	11h37	11	40	40	No	No	Yes	No
	13h52	30	40	30	No	No	No	Yes
	17h22	15	30	15	No	Yes	No	No
13.05.08	08h57	17	39	39	Yes	No	Yes	No
	15h20	26	38	29	No	No	No	No
	17h20	26	34	26	No	Yes	No	No
14.05.08	13h20	26	50	50	Yes	No	Yes	Yes
	15h00	30	41	29	No	No	No	No
	16h06	26	29	27	No	No	No	No
15.05.08	09h42	13	34	34	Yes	No	Yes	No
	16h57	20	34	20	No	Yes	No	No
16.05.08	08h39	22	35	35	No	No	Yes	No
	14h03	28	30	30	Yes	No	No	No
	14h57	22	40	33	No	Yes	No	No
17.05.08	08h15	27	36	34	No	No	Yes	No
	14h02	25	34	29	No	No	No	No
18.05.08	10h02	12	30	20	No	No	No	No
	14h00	20	25	25	No	No	No	No
19.05.08	08h09	14	37	19	No	No	No	No

	12h00	22	36	36	No	No	Yes	No
	17h15	18	27	18	No	No	No	No
20.05.08	09h20	11	26	26	Yes	No	No	No
	14h47	26	38	31	No	No	No	No
	17h36	18	35	18	Yes	Yes	No	No
21.05.08	06h56	19	22	20	No	No	No	No
	12h38	18	43	36	No	No	No	No
	17h10	21	26	22	No	Yes	No	No
22.05.08	06h08	22	26	22	Yes	No	No	No
	09h15	15	32	32	No	No	Yes	No
	12h17	25	32	25	No	No	Yes	No
	14h45	23	30	27	No	No	No	No
	16h40	27	34	27	No	Yes	No	No
23.05.08	07h10	22	30	24	No	No	No	No
	11h50	20	42	42	No	No	Yes	Yes
	12h52	21	42	20	No	No	No	No
24.05.08	10h25	10	35	19	No	No	No	No
	12h20	20	34	30	No	No	No	No
25.05.08	10h03	10	38	13	No	Yes	No	No
	13h20	13	30	29	No	Yes	No	No
26.05.08	08h19	16	31	25	No	No	No	No
	12h15	20	31	30	No	No	No	No
	15h35	19	34	19	No	No	No	No
	16h26	16	19	16	No	Yes	No	No
27.05.08	08h31	16	26	26	No	No	No	No
	11h37	20	26	25	No	No	No	No
	14h30	21	41	40	No	No	No	No
	16h50	15	40	15	No	Yes	No	No
28.05.08	09h00	15	35	35	No	No	No	No
	12h24	32	37	37	No	No	Yes	No
	15h15	27	38	26	No	Yes	No	No
	17h00	24	30	28	Yes	Yes	No	No
	19h03	29	34	34	No	Yes	No	No
29.05.08	06h08	22	34	22	No	Yes	No	No
	06h52	22	30	30	No	No	No	No
	08h20	34	36	36	No	No	No	No
	10h44	24	36	36	No	No	Yes	No
30.05.08	07h54	24	30	30	No	No	No	No
	10h10	25	31	31	No	No	No	No
	11h55	26	30	30	No	Yes	No	No
	16h00	25	30	24	No	No	No	No
	18h38	15	26	15	No	Yes	No	No
31.05.08	08h10	16	30	30	No	No	No	No
	09h44	16	30	30	No	No	Yes	No
01.06.08	13h36	13	30	26	Yes	No	No	No
02.06.08	10h00	12	36	30	No	No	Yes	No
	15h43	28	29	26	No	Yes	No	No
03.06.08	08h28	18	27	26	No	No	No	No
	10h53	26	33	33	No	No	Yes	No
	14h50	21	34	22	No	Yes	No	No
04.06.08	08h35	18	29	20	No	Yes	No	No
	12h18	26	37	34	No	No	No	No
	15h00	23	37	37	No	Yes	No	No

	15h45	27	31	31	No	No	No	No
05.06.08	08h20	16	29	25	No	No	Yes	No
	11h45	18	26	26	No	No	No	No
	15h19	26	30	26	No	No	No	No
	16h54	26	31	26	No	Yes	No	No
06.06.08	08h10	19	27	27	No	No	No	No
	10h35	26	36	36	No	No	Yes	No
	15h42	26	30	26	No	No	No	No
07.06.08	09h53	8	31	31	No	No	No	No
	13h54	28	32	31	Yes	No	No	No
09.06.08	08h10	10	41	10	No	Yes	No	No
	13h09	12	37	37	No	No	Yes	Yes
	15h50	19	37	21	No	No	No	No
10.06.08	08h40	15	24	20	No	No	No	No
	11h24	20	30	27	No	No	No	No
	16h46	21	32	21	No	Yes	No	No
11.06.08	08h47	19	29	29	Yes	No	No	No
	12h21	34	33	33	No	No	Yes	No
	15h46	22	33	22	No	Yes	No	No
12.06.08	08h49	20	32	32	No	No	No	No
	12h05	31	39	39	No	No	Yes	No
	16h05	24	39	24	No	Yes	No	No
13.06.08	08h04	23	30	30	No	No	Yes	No
	15h43	20	26	30	No	No	No	No
14.06.08	08h41	18	35	35	No	No	Yes	No
	16h41	20	36	20	No	No	No	No
15.06.08	09h30	18	31	31	No	No	Yes	No
	15h53	25	31	27	No	No	No	No
16.06.08	10h35	10	33	33	No	No	Yes	No
	17h04	17	33	17	No	Yes	No	No
17.06.08	08h24	18	26	26	Yes	No	No	No
	14h39	22	28	22	No	Yes	No	No
18.06.08	09h06	20	26	22	No	No	No	No
	15h44	20	36	20	No	No	No	No
19.06.08	10h11	16	30	26	No	No	No	No
	14h18	26	35	35	Yes	No	Yes	No
	15h31	24	35	24	No	No	No	No
	17h02	21	26	21	No	Yes	No	No
20.06.08	08h50	17	27	27	No	No	No	No
	15h56	26	35	29	No	No	No	No
	18h47	15	29	15	No	Yes	No	No
21.06.08	08h08	16	22	22	No	No	No	No
	09h51	23	28	28	No	No	No	No
	17h14	20	36	20	No	Yes	No	No
22.06.08	09h00	20	36	36	Yes	No	Yes	No
	15h38	20	36	23	No	No	No	No
	18h40	14	26	14	No	Yes	No	No
23.06.08	09h19	15	35	35	No	No	No	No
	17h33	15	34	15	No	Yes	No	No
24.06.08	09h00	16	36	36	No	No	Yes	No
	13h50	20	41	41	No	No	Yes	No
	14h55	25	41	25	No	No	No	No
	17h22	22	36	22	No	Yes	No	No

25.06.08	08h34	20	31	31	No	No	No	No
	12h33	21	31	27	No	No	No	No
	15h51	25	28	25	No	No	No	No
	17h29	19	29	19	No	No	Yes	No
26.06.08	08h56	19	34	34	No	No	No	No
	17h00	23	42	26	No	No	No	No
27.06.08	08h01	21	27	26	No	Yes	No	No
	16h50	20	28	20	No	Yes	No	No
28.06.08	09h34	21	40	40	Yes	No	No	No
	15h37	25	40	25	No	No	No	No
	17h07	20	30	20	No	Yes	No	No
29.06.08	08h50	20	37	37	No	No	Yes	No
	16h45	19	37	19	No	Yes	No	No
30.06.08	08h51	20	30	30	No	No	Yes	No
	17h52	14	24	14	No	Yes	No	No
01.07.08	08h21	14	30	30	Yes	No	Yes	No
	13h10	20	30	27	No	No	No	No
	17h44	14	29	14	No	Yes	No	No
02.07.08	08h55	15	30	30	No	No	No	No
	16h33	21	30	21	No	Yes	No	No

Appendix 4.2: Growth trend of plants towards stage for bombardment

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX U0

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU				*	L	L	I	
2	AR30-3		*	*	L	I	I	I	ABANDONED
3	AR37-96			*	L	I	I	I	
4	AR21-2			*	L	L	L	I	
5	AR37-1					*	L	I	ABANDONED
6	AR40-10		*	L	L	L	L	I	
7	AR37-92				*	L	I	S	
8	CR44-6								ABANDONED
9	CR27-24		*	L	L	I	I	I	
10	CR45-3		*	*	Dead				ABANDONED
11	AR37-6					L	I	S	
12	CR25-4		*	L	L	I	I	S	ABANDONED
13	AR16-3		*	*	L	L	L	I	
14	AR17-5			*	L	I	I	CMD	
15	AR37-73								ABANDONED
16	AR14-2				*	I	I	I	ABANDONED
17	CR45-9			*	*	L	L		ABANDONED

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX E0

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU					*	L	L	
2	AR30-3			*	*	L	I	I	ABANDONED
3	AR37-96				*	L	L	?	
4	AR21-2				*	L	I	I	

5	AR37-1				*	L	L	I	ABANDONED
6	AR40-10			*	L	I	I	I	
7	AR37-92				L	I	I	I	
8	CR44-6								ABANDONED
9	CR27-24		*	*	L	I	I	I	
10	CR45-3		*	L	?	Dead		Dead	ABANDONED
11	AR37-6		*	*	*	L	L	S	
12	CR25-4			*	L	I	I	S	ABANDONED
13	AR16-3			*	L	L	I	I	
14	AR17-5				*	*	*	S	
15	AR37-73								ABANDONED
16	AR14-2			*	L	L	I	I	ABANDONED
17	CR45-9					?	?	Dead	ABANDONED

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX EII

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU			*	L	L	I	I	
2	AR30-3			?			?	?	ABANDONED
3	AR37-96				*	L	L	I	
4	AR21-2				*	*	I	I	
5	AR37-1					*	*	L	ABANDONED
6	AR40-10					L	I	I	
7	AR37-92		*	*	L	I	I	I	
8	CR44-6							?	ABANDONED
9	CR27-24		*	*	L	L	I	I	
10	CR45-3		L	L	L	L	I	I	ABANDONED
11	AR37-6				*	L	I	I	
12	CR25-4		*	*	L	L	?	?	ABANDONED
13	AR16-3		*	L	L	L	I	I	
14	AR17-5				*	L, R	I	I	
15	AR37-73							?	ABANDONED
16	AR14-2			*	*	L	I	I	ABANDONED
17	CR45-9				?	*	*	?	ABANDONED

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX EIII

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU					*	L	I	
2	AR30-3			*	*	L	I	I	ABANDONED
3	AR37-96			*	L	I	I	I	
4	AR21-2					L	L	I	
5	AR37-1								ABANDONED
6	AR40-10				*	L	I	CMD	
7	AR37-92			*	L	I	I	I	
8	CR44-6								ABANDONED
9	CR27-24		*	L	L	I	I	I	
10	CR45-3		*	L	L	?	?	Dead	ABANDONED
11	AR37-6				*	L	I	S	
12	CR25-4			*	L	I	I	S	ABANDONED
13	AR16-3				*	L	I	I	
14	AR17-5				*	L	I	I	
15	AR37-73								ABANDONED

16	AR14-2			*	L	I	I	S	ABANDONED
17	CR45-9								ABANDONED

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX U0

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU				*	L	L	I	
2	AR30-3		*	*	L	I	I	I	ABANDONED
3	AR37-96			*	L	I	I	I	
4	AR21-2			*	L	L	L	I	
5	AR37-1					*	L	I	ABANDONED
6	AR40-10		*	L	L	L	I	I	
7	AR37-92				*	L	I	S	
8	CR44-6								ABANDONED
9	CR27-24		*	L	L	I	I	I	
10	CR45-3		*	*	Dead				ABANDONED
11	AR37-6					L	I	S	
12	CR25-4		*	L	L	I	I	S	ABANDONED
13	AR16-3		*	*	L	L	L	I	
14	AR17-5			*	L	I	I	CMD	
15	AR37-73								ABANDONED
16	AR14-2				*	I	I	I	ABANDONED
17	CR45-9			*	*	L	L		ABANDONED

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX UII

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU					*	*	L	
2	AR30-3						*	?	ABANDONED
3	AR37-96		*	L	L	L	I	I	
4	AR21-2					L	I	I	
5	AR37-1					*	L	I	ABANDONED
6	AR40-10					L	I	I	
7	AR37-92		*	L	L	I	I	I	
8	CR44-6								ABANDONED
9	CR27-24			*	L	L	I	I	
10	CR45-3					Dead			ABANDONED
11	AR37-6			*	*	I	I	I	
12	CR25-4		*	*	L	I	I	I	ABANDONED
13	AR16-3			L	L	I	I, ?	Dead	
14	AR17-5					L	I	I	
15	AR37-73								ABANDONED
16	AR14-2			*	L	I	I	I	ABANDONED
17	CR45-9						?		ABANDONED

PLANTS PREPARED FOR BOMBARDMENT WITH PLASMID MIX UIII

S/N	LINE	02.04.08 (PLANTING)	7 DAP	14 DAP	21 DAP	28 DAP	35 DAP	42 DAP	REMARKS
1	B/MREFU					?	*	I	
2	AR30-3		*	*	L	I	I	I	ABANDONED
3	AR37-96			*	L	I	I	I	
4	AR21-2					L	I	I	
5	AR37-1					*	L	I	ABANDONED

6	AR40-10	L	L	L	I	I	I	
7	AR37-92	L	L	I	I	I	I	
8	CR44-6						?	ABANDONED
9	CR27-24	*	L	L	?	L	?	
10	CR45-3	*	L	L	I	I	I	ABANDONED
11	AR37-6	*	*	L	I	I	I	
12	CR25-4		?	Dead	Dead	Dead	Dead	ABANDONED
13	AR16-3	*	L	L	L	I	I	
14	AR17-5		?	*	L	I	I	
15	AR37-73					?	?	ABANDONED
16	AR14-2		*	L	L	I	I	ABANDONED
17	CR45-9					?	?	ABANDONED

LEGEND: * = Growing bud protruding;

L = Leaves emerging;

I = Growing shoot at size suitable for bombardment

R = Tendency of young leaves to roll

? = Status not exactly known, i.e. whether dead or alive.

S = Heat stress

Appendix 4.3: Cell transformation protocol

1. Add 3 μL of about 50 ng plasmid DNA into a tube of 5Hα competent cells. Mix by gently flicking with a finger and incubate on ice for 5 minutes.
2. Heat shock cells in a heating block at 42°C for 90 seconds, transfer back to ice and keep there for 5 minutes.
3. Pipette 50 μL of cell suspension and spread on LB agar (with ampicilin at 1 μL/mL) in duplicate selection plates. Leave plates to absorb cells for about 10 minutes.
4. Incubate plates at 37°C for 16 h without shaking.

Culturing Procedure

1. Put 50mL LB broth into Nunc tube and add 50μL ampicilin. Mix by inverting the tube.
2. Divide this volume into 6 Nunc tubes. Label them appropriately according to the samples to work with.

3. Transfer one colony of each sample into the tube, using pipette tip. Suck up and down in the broth.
4. Cover tubes with their caps and incubate at 37°C with shaking at 170rpm for 16hours.
5. Wrap selection plates with parafilm and store at 4°C.

Appendix 4.4: Plasmid extraction protocol

[Adopted from Sambrook J., Fritsch E.F. and Maniatis T. (1989). Molecular cloning. A laboratory Manual 2nd Edition. Cold Spring Laboratory Press pp 1.5 – 1.28]

Reagents

- Solution I – 50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA in sdH_2O.
- Solution II – 0.2N NaOH, 1% SDS in sdH_2O
- Solution III – To get 50 mL use 30 mL 3M potassium acetate and 5.75 mL 5M glacial acetic acid. Make up to 50mL with sdH_2O.

Procedure

1. Pour 1.5–2.0 mL of the overnight bacterial culture into a microfuge tube and spin for 3 minutes at 4°C. Store any remaining suspension at 4°C.
2. Remove supernatant, repeat step 1 if necessary in order to get adequate pellet. Blot dry, leaving the pellet as dry as possible (important step).
3. Resuspend pellet in 250 μL of ice cold Solution I (resuspension buffer) by vigorous vortexing. Leave at room temperature for 5 minutes.
4. Add 250μL of freshly prepared Solution II, mix by inverting gently for 6 times. Do not vortex. Incubate at RT for 5-10 minutes.
5. Add 350 μL of ice cold Solution III (neutralization buffer). Mix by inverting vigorously. Incubate on ice for 5-10 minutes.
6. Centrifuge at high speed for 10 minutes at 4°C. Transfer supernatant to new tube using a pipette avoiding the white precipitate on top as much as possible.
7. Spin again and transfer supernatant to new tube.
8. Add an equal amount of isopropanol (at room temperature). Mix and incubate at RT for 5 minutes.

9. Precipitate DNA by centrifuging for 10 minutes at 4°C. Remove supernatant by using a pipette. Blot dry.
10. Wash pellet in 1 mL 70% ethanol. Spin for 5 minutes at 4°C. Decant ethanol by using pipette. Spin empty microfuge tubes for 5 minutes and decant any ethanol. Invert tubes over blotting paper to dry (or alternatively use heating block at 37°C for about 20min to allow ethanol to evaporate).
11. Resuspend pellet in 50-100 μL TE (or nuclease free H_2O) at 65°C. If not resuspended, put tube in water bath at 65°C.

Appendix 4.5: Preparation of plasmid mixes

Guiding principles

1. There were six plasmid mixes: EACMV-TZ + SatDNA-II, EACMV-TZ + SatDNA-III, EACMV-TZ (control), EACMV-UG2 + SatDNA-II, EACMV-UG2 + SatDNA-III and EACMV-UG2 (control), denoted for simplicity as respectively EII, EIII, E0, UII, UIII, and U0;
2. In each of the six mixes, 50 μg of DNA was required;
3. Where the mix consisted of yet a mix between helper virus and a satellite, then the ratio between the two was always 1:1, i.e. 25 μg virus and 25 μg satellite;
4. There were six samples of clones to be used for the formation of the plasmid mixes and each clone had different DNA concentration, following cell transformation: UA (EACMV-UG genomic component A), UB (EACMV-UG genomic component B), EA (EACMV-TZ genomic component A), EB (EACMV-TZ genomic component B), II (SatDNA-II) and III (SatDNA-III);
5. In the bipartite helper viruses, genomic components (A and B) were also mixed in the ratio 1:1. So, to make treatment EII it required 12.5 μg of EA, 12.5 μg of EB and 25 μg of SatDNA-II;
6. Each of the six plasmid DNA sample was diluted to 1μg/μL, preparing a volume of 60 μL. The amount of each DNA sample required in the experiment was thus calculated:

(a) **UA**: 25 μg in mix E0, 12.5 μg in mix EII, and 12.5 μg in mix EIII. Total 50 μg. To take care of working errors, prepared 60 μg (60 μL): Volume of stock = Desired conc./Available Conc. x Vol. required, this is 1/5.2064 x 60 μL = 11.5 μL. So, you draw 11.5 μL of stock DNA into reaction tube and make up to 60 μL by adding 48.5 μL of sterile distilled water (sdH_2O).

(b) **UB**: 25 μg in mix U0, 12.5 μg in mix UII and 12.5 μg in mix UIII, total 50 μg (as usual, prepare 60 μL). Vol. of stock = 1/5.1968 x 60μL = 11.5 μL. This is dissolved in 48.5 μL of sdH_2O.

(c) **EA**: 25 μg in mix E0, 12.5 μg in mix EII and 12.5 μg in mix EIII. Total 50 μg (but prepare 60 μL). Vol. of stock = 1/5.196 x 60 μL = 11.5 μL, to be dissolved in 48.5 μL of sdH_2O.

(d) **EB**: 25 μg in mix E0, 12.5 μg in mix EII and 12.5 μg in mix EIII. Total 50 μg. Vol. of stock = 1/5.1527 x 60 μL = 11.6 μL, to be dissolved in 48.4 μL of sdH_2O.

(e) **II**: 25 μg in mix EII and 25 μg in mix UII. Total 50 μg. Vol. of stock = 1/5.2488 x 60 μL = 11.4 μL, to be dissolved in 48.6 μL of sdH_2O.

(f) **III**: 25 μg in mix EIII and 25 μg in mix UIII. Total 50 μg. Vol. of stock required = 1/5.203 x 60 μL = 11.5 μL, to be dissolved in 48.5 μL of sdH_2O.

Appendix 4.6: Protocol for particle bombardment

(Based on User Manual provided by the Helios Gen Gun manufacturers)

Preparing gold particles

1. Prepare stock solution of 20 mg/mL PVP in ethanol. Dilute to a working concentration of 0.05 mg/mL with more ethanol.
2. Weigh 25 mg gold into 1.5 mL microfuge tube.
3. To the gold, add 100 μL of 0.05 M spermidine (volume of plasmid DNA in step 5 below should not be greater than 100 μL).
4. Vortex gold and spermidine for few seconds then sonicate for 3-5sec to break up gold clumps.

5. To the gold and spermidine mixture, add the required amount of plasmid (should be 1μg/μL) to achieve the desired DLR. For co-transfection of multiple plasmids, add each of the plasmids at this step.
6. Mix DNA, spermidine and gold by vortexing for approx 5sec.
7. While vortexing the mixture at moderate rate on a variable speed vortexer, add 100 mL of 1 M $CaCl_2$ dropwise to the mixture. The volume added should equal that of the spermidine in step 3.
8. Allow the mixture to precipitate at room temperature for 10 minute, after which the supernatant should be relatively clear.
9. Spin the microcarrier solution in a microfuge tube for about 15 sec to pellet the gold. Remove supernatant and discard.
10. Resuspend pellet in the remaining supernantant by vortexing briefly. Wash pellet 3 times with 1mL of fresh absolute ethanol each time; spin for approx 5 sec in the microfuge between each wash. Discard supernatants.
11. Resuspend pellet in 200μL of the ethanol solution containing PVP at 0.05 mg/mL. Transfer the suspension to a 15 mL disposable centrifuge tube with screw cap. Rinse the microfuge tube with 200 μL with the same ethanol/PVP solution and add to the centrifuge tube. Make the volume to 3 mL with more ethanol/PVP solution.
12. The suspension is ready for tube preparation. It can be stored for up to 2 month at -20°C. Prior to freezing, tighten the cap and put parafilm. After storage at -20°C, allow particle suspension to come to room temperature before breaking the parafilm seal.

Preparing gold coat tubing and cartridges

1. Cut 75 cm of tube, wash with absolute ethanol using syringe to suck.
2. Insert tube into Tubing Prep Station, dry with N_2 at 50 kPa for 15 minute.
3. Vortex suspension (then sonicate) and draw into tube using syringe, approximately 58 cm (about 17 cm) from the end. Avoid drawing bubbles into the tube.
4. Remove tube from the suspension and continue drawing another approximately 6 cm to leave some space at each end.

5. Immediately bring the tube to horizontal position, slide the loaded tube with syringe attached to the support cylinder until the tube passes through the O-ring.
6. Allow the microcarriers to settle for 3-5 minutes. Using the syringe, remove ethanol at the rate of 0.5-1"/sec (this should require 30-45sec).
7. Detach the syringe and immediately rotate tube 180° while in the groove and allow the gold to begin coating the inside surface of the tubing for 3-4sec.
8. Switch ON the tubing prep station to start rotating the Tubing prep station. Allow the gold to smear in the tube for 20 to 30 sec.
9. While rotating, open the valve of the flow meter to allow 0.35-0.4L/min of nitrogen to dry the gold coat tubing. Continue drying the gold tubing while turning for 3-5 minutes.
10. Switch OFF the motor. Turn off the nitrogen. Remove the tubing from the tubing support cylinder.
11. Examine the coated tube to verify that the microcarriers are evenly distributed over the length of the tubing. Cut off and discard sparsely and unevenly coated tubing from one of the ends. Mark with a marker pen any sections of tubing to be discarded.
12. Cut tube into 0.5 in-length cartridges using tubing cutter. These can be wrapped with parafilm and stored at 4°C for at least 8months.

DNA delivery to tissues

1. Insert cartridges in cartridges holder of the gun as numbered from 1 to 12 and load it onto the gun.
2. Set the required helium pressure (240 psi = 1654.7 kPa).
3. Apply 3 shots per plant at a distance of about 4.5 cm from the target.

Appendix 4.7: CTAB total nucleic acid extraction protocol

[Adapted and Optimized from Colosi and Schaal, 1993. *Nucleic Acids Res*. 21:1051-1052]. Note: The protocol is designed for 1.5 ml microfuge tubes.

Reagents

- **DNA extraction buffer (CTAB buffer) consisting of:**
  2% CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl (pH 8.0), add 0.2% beta-mercaptoethanol just prior to use.
- **TE buffer consisting of:**
  10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0).
- **CIA consisting of:**
  24 parts chloroform: 1 part octan-1-ol (isoamyl alcohol).
- **Isopropanol**, cold, stored at -20°C
- **70% ethanol**, cold, stored at -20°C
- **100% ethanol**, cold, stored at -20°C

Before You Start

- Place mini-spin at 4°C, at least an hour before you start.
- Switch ON water bath and set it at 65°C.
- Place extraction buffer in the water bath (65°C) so it can heat up while the samples are being crushed.
- Place 70% molecular grade ethanol on ice.

Procedure

1. Aliquot enough CTAB buffer for your samples (500 microliters each plus 10% extra) into a disposable 50 mL Falcon tube and incubate in 65°C water bath.
2. Harvest 0.050 to 0.100 g of succulent young leaf tissue and place it into a 1.5 mL eppendorf tube. [The leaf samples can be frozen and stored at -70°C].
3. Carefully pour liquid nitrogen into the tube until fizzing stops.
4. Grind tissue to a fine powder using a drill. [You can use a 0.25 inch drill bit fitted snugly with a 1 mL-pipette tip, with the very end of the tip bent over]. Change the tip between samples. Keep ground samples on ice.
5. Before adding CTAB buffer to your samples, add 0.2% beta-mercaptoethanol to (equivalent to 1 μL to 500 μL CTAB buffer) final volume of CTAB. Work under the fume hood.

6. Add 500 mL of warm CTAB buffer to the ground leaf tissue, vortex, and incubate in a water bath at 65°C for 30 to 60 minutes.
7. Working under the fume hood, add 500 mL of CIA to the warm mixture (to extract the top aqueous layer containing the total nucleic acids) and invert 50 times. Do NOT vortex to prevent shearing of the DNA. Also, be careful not to erase labels by spilling drops of CIA!
8. Centrifuge at maximum speed for 10 minutes at 4°C.
9. While in the hood, transfer as much of the supernatant as possible to a clean tube without disturbing the interface. Repeat steps 7 to 9 and discard the bottom layer into the appropriate hazardous waste container.
10. Add an equal volume of ice cold isopropanol to the supernatant and mix thoroughly to precipitate the DNA.
11. Centrifuge at maximum speed for 10 minutes and decant supernatant to obtain the pellet.
12. Wash the pellet with 500 mL of ice cold 70% molecular grade ethanol.
13. Centrifuge at max speed for 10 minutes.
14. Repeat the washing step with 500 mL cold 70% ethanol and centrifuge at maximum speed for 5-10 minutes.
15. Decant the ethanol and allow the pellet to air dry or place in heating block at 37°C for more rapid drying.
16. Resuspend the pellet in 20 to 100 mL of 1x TE buffer containing 20µg/mL RNase A. The TE buffer can be pre-heated at 65°C for a few minutes to allow for easy resuspension. Store total nucleic acids at 4°C.

Appendix 4.8: Protocol for Southern blot analysis

[According to Instruction Manual by DIG High Prime DNA Labeling and Detection Kit I. Roche Applied Science, Mannheim Germany and Sambrook J, E.F. Fritsch, and T. Maniatis (1989). Molecular Cloning: A Laboratory manual 2nd Ed. Cold Spring Harbor Laboratory Press Pg 9.34].

Solutions needed to make up buffers

- 2M Tris-HCl (pH 7.2) 500 mL: 121.1g Tris base, 300 mL dH_2O. Adjust pH with Conc. HCl, make to 500mL. Autoclave 20 minutes.

- 5M NaCl 500 mL: 146.1g NaCl. Autoclave 20 minutes.
- 10N NaOH 100 mL: 40g in 100 mL dH_2O.
- 0.5M EDTA (pH 8.0) 250 mL: 46.52 g EDTA, 150 mL dH_2O, pH using 10N NaOH. Make up to 250 mL with dH_2O.

Buffers for blotting

- 0.25M HCl 250 mL: 62.5 mL 1M HCl, 187.5 mL dH_2O.
- 20X SSC 1L: 175.3g NaCl, 88.0g Na_3Citrate (Trisodium citrate), 800 mL dH_2O, pH to 7.2 with 10N HCl, make up to 1L.
- Neutralization buffer 500 mL: 125 mL 2M Tris-HCl pH 7.2, 150 mL 5M NaCl, 1 mL 0.5M EDTA pH 8.0. Autoclave 20 minutes, make up 500 mL.
- Denaturation buffer 500 mL: 150 mL 5M NaCl, 25 mL 10N NaOH, 325 mL dH_2O. Autoclave 20 minutes.

Solutions for hybridization and washing

- Hybridization working Solution, 64 mL: Add 64 mL sdH_2O in two portions to the DIG Easy Hyb granules (bottle 7), first add 32 mL sdH_2O warm 5 minutes at 37^oC mix gently then add another 32 mL sdH_2O.
- 2X SSC 250 mL: 25 mL 20X SSC, 0.25g SDS, 225 mL dH_2O, autoclave.
- 0.1X SSC 250 mL: 1.25 mL 20X SSC, 0.25g SDS, 248.75 mL dH_2O, Autoclave.

Additional reagents for immunological detection:

(i) Washing buffer (for washing membrane), 500 mL: 0.1M Maleic acid, 0.15M NaCl; pH 7.5 (20^oC); 0.3% (v/v) Tween 20. Store at 15-25^oC.

(ii) Maleic acid buffer (for dilution of blocking solution), 250 mL: 0.1M Maleic acid, 0.15M NaCl; adjust with NaOH (solid) to pH 7.5 (20^oC). Store at 15-25^oC.

(iii) Detection buffer (Alkaline phosphatase buffer), 20 mL: 0.1M Tris-HCl, 0.1M NaCl, pH 9.5 (20^oC). Store at 15-25^oC.

(iv) TE Buffer (for stopping colour reaction): 10 mM Tris-HCl, 1 mM EDTA, pH 8.0. Store at 15-25^oC.

DNA extraction

1. 100 mg of leaf samples from the plants will be taken, placed into a plastic bag and frozen in liquid nitrogen.
2. Total DNA from the tissues extracted by CTAB protocol.

PCR and electrophoresis

1. Amplify *AC1* gene fragments (1077bp) from DNA-A by PCR using primers EAC1F 5'-GGA ATT CCA TAT GCC GAG AGC CGG TCG TTT TCA AAT A-3' and EAC1R 5'-CCG CCT CGA GGC TTG CCT GTG CCT GGC T-3'. Also, amplify SatII (895bp) and SatIII (306bp) fragments using primers: SatII-F 5'-GCC GCA CCA CTG GAT CTC-3', SatII-R 5'-CGT TTA CAG CCC ACC TCT GT-3', SatIII-F 5'-AGG CCT CGT TAC TAA AAG TGC-3', SatIII-R 5'-ACC TGA CGG CAA GGA AT-3'.
2. Detect positive result by visualizing the PCR product on 1% agarose gel in 1x TAE buffer. Electrophorese the gel at 80V for 50min. Confirm identity of the PCR product by Southern blot hybridization as follows.

Preparations of probes and labeling

1. Amplify *AC1, SatII* and *SatIII* fragments by PCR from cassava total DNAs to obtain PCR products.
2. Precipitate the PCR products using ice-cold isopropanol and resuspend pellet in 15uL TE buffer. Use these as probes.
3. Add 1μg (about 8uL) template DNA (linear or supercoiled) to autoclaved, double distilled water to a final volume of 16μL to a reaction vial.
4. Denature the DNA by heating in a boiling water bath at 69°C for 10min and quickly chilling in an ice/water bath. **NB**: Complete denaturation is essential for efficient labeling.
5. Mix DIG-High Prime (vial 1) thoroughly and add 4μL to the denatured DNA, mix and centrifuge briefly. Incubate O/N at 37°C. **NB**: Longer incubations (up to 20h) will increase the yield of DIG-labeled DNA.
6. Stop the labeling reaction by adding 2μL of freshly prepared 0.2M EDTA (pH 8.0) and by heating to 65°C for 10min.

Denature DNA fragments, transfer to membrane and cross-linking

1. Place the gel (containing electrophoresed DNA) to depurination treatment (0.25M HCl) for 20 minutes. Depurination breaks DNA into smaller fragments, thus facilitating transfer to membrane. [**NB**: Not a necessary step if the fragment is shorter than 3kb].
2. Rinse the gel in distilled water, followed by incubation in denaturation buffer for 15 minutes at RT with shaking.
3. Replace denaturation buffer with fresh buffer and leave for 15 minutes.
4. Rinse with sdH_20, transfer to neutralization buffer, leave for 15 minutes at room temperature, shaking.
5. Replace neutralization buffer with fresh buffer leave for 15 minutes. Equilibrate gel in 20X SSC for 10 minutes.
6. Place denatured gel onto sheets of moist filter paper (the wells facing down) and dip the filter papers into buffer reservoir containing 20x SSC.
7. A nylon membrane (positively charged) is laid over the gel; a number of dry filter papers are placed on top of the membrane; a weight about 500 g is placed on top and left overnight (18-22h).
8. Rinse membrane in 6X SSC buffer for 1 minute and dry it.
9. Treat membrane with UV light in UV transilluminator (or in UV cross-linker) at 120milliJoules for 3-5 minute to irreversibly bind DNA to the membrane. To do this, place the membrane on Whatman paper soaked in 10X SSC. UV-crosslink the wet membrane without prior washing. Then wash the membrane briefly in dH_2O and allow to air dry.

Hybridization

1. Prepare DIG Easy Hyb working solution by carefully adding 64 mL sdH_2O in two portions to the DIG Easy Hyb granules (bottle 7), first add 32 mL sdH_2O warm 5 minutes at 37^oC mix gently then add another 32 mL sdH_2O.
2. Pre-heat DIG easy hybridization solution (10 mL/100 cm^2 filter) to hybridization temperature (37^oC–42^oC) for about 1 h.
3. Pre-hybridize filter for 30 minutes with gentle agitation in an appropriate container (membrane should move freely during agitation).

4. Denature the DIG-labeled DNA probe (about 25 ng/mL) by boiling (68°C) for 5min and rapidly cooling in ice/water.
5. Add denatured DIG-labeled DNA probe to the preheated DIG Easy Hyb (3.5m L/100 cm^2 membrane) and mix well but avoid foaming (bubbles may lead to background problems).
6. Pour off prehybridization solution and add probe/hybridization mixture to membrane. Incubate overnight (± 18h) with gentle agitation (22 rpm at 35°C).

Post-hybridization washes

1. Wash 2x 5 minutes in ample 2X SSC + 0.1% SDS at 15-25°C under constant agitation. (For 50 mL: 500 uL 10% SDS + 5 mL 20X SSC + 44.5 mL sdH_2O).
2. Wash 2x 15 minutes in 0.5X SSC + 0.1% SDS (pre-warmed to wash temperature) at 65°C under constant agitation. (For 50 mL: 500 μL 10% SDS + 1.25 mL 20X SSC + 48.25 mL sdH_2O).

Immunological detection

Prepare kit working solutions:

(i) Blocking solution (for blocking unspecific binding sites), 100 mL: Prepare 1X working solution by diluting 10X blocking solution (vial 6 or 2% skim milk) 1:10 with maleic acid buffer. Always prepare this fresh.

(ii) Antibody solution (for binding to the DIG-labeled probe), 20 mL: Centrifuge Anti-Digoxigenin-AP (vial 4) for 5 minutes at 10,000 rpm in the original vial prior to use, and pipette the necessary amount carefully from the surface. Dilute Anti-Digoxigenin-AP 1:5000 in blocking solution. Store antibody solution at 2-8°C for up to 12h.

(iii) Colour substrate solution (for visualization of antibody binding), 10 mL: Add 200μL of NBT/BCIP stock solution (vial 5) to 10 mL of detection buffer. Always prepare fresh.

Procedure for immunological detection

1. After hybridization and stringent washes, rinse membrane briefly (1-5 minutes) in washing buffer.

2. Incubate membrane for 30 minutes in 100 mL blocking solution (10 mL blocking solution in 90 mL maleic acid buffer).
3. Incubate for 30 minutes in 20 mL antibody solution (2 mL Blocking solution in 18 mL maleic acid buffer + 4 μL Anti-DIG-AP). Spin down the Anti-DIG-AP (vial 4) for 5 minutes at 10,000 rpm and remove 4 μL (for the 20 mL) from the top of solution.
4. Wash 2x 15minutes in 100 mL washing buffer.
5. Equilibrate 2-5 minutes in 20 mL Detection buffer.
6. Change tray for colour detection. Incubate membrane in 10 mL freshly prepared colour substrate solution (10 mL detection buffer + 200 μL colour substrate (NBT/BCIP - vial 5) in an appropriate container in the dark. Do not shake during colour development. **NB**: The colour precipitate starts to form within a few minutes and the reaction is usually complete after 16h. The membrane can be exposed to light for short time periods to monitor colour development.
7. Stop the reaction, when desired spot or band intensities are achieved, by washing the membrane for 5 minutes with 50 mL sdH_2O or TE buffer. Document results by photocopying the wet filter or by photography.

Appendix 6.1: Members of Genus: *Begomovirus-associated Satellite DNA*s

1. *African cassava mosaic virus-associated DNA-II*
2. *African cassava mosaic virus-associated DNA-III*
3. *Ageratum yellow leaf curl betasatellite* (AYLCB)
4. *Ageratum yellow vein betasatellite* (AYVB)
5. *Ageratum yellow vein Sri Lanka betasatellite* (AYVSLB)
6. *Ageratum yellow vein virus-associated DNA-1*
7. *Bean leaf curl China betasatellite* (BLCCNB)
8. *Bhendi yellow vein betasatellite* (BYVB)
9. *Cardiospemum yellow leaf curl betasatellite* (CarYLCB)
10. *Chilli leaf curl Multan virus-associated DNA-1*
11. *Cotton leaf curl Gezira betasatellite* (CLCuGB)
12. *Cotton leaf curl Multan betasatellite* (CLCuMB)

13. *Cotton leaf curl Multan virus satellite U36-1*
14. *Cotton leaf curl virus-associated DNA-1*
15. *Croton yellow vein betasatellite* (CroYVB)
16. *Croton yellow vein mosaic betasatellite* (CroYVMB)
17. *Erectites yellow mosaic betasatellite* (ErYMB)
18. *Eupatorium yellow vein betasatellite* (EpYVB)
19. *Hibiscus leaf curl virus-associated DNA-1*
20. *Hollyhock leaf crumple betasatellite* (HoLCB)
21. *Hollyhock leaf crumple virus-associated DNA-1*
22. *Honeysuckle yellow vein betasatellite* (HYVB)
23. *Honeysuckle yellow vein Japan betasatellite* (HYVJB)
24. *Honeysuckle yellow vein mosaic betasatellite* (HYVMB)
25. *Lindernia anagallis yellow vein betasatellite* (LaYVB)
26. *Ludwigia yellow vein betasatellite* (LuYVB)
27. *Luffa leaf distortion betasatellite* (LuLDB)
28. *Malvastrum leaf curl betasatellite* (MaLCuB)
29. *Malvastrum yellow mosaic virus-associated DNA-1*
30. *Malvastrum yellow vein betasatellite* (MaYVB)
31. *Malvastrum yellow vein Yunnan betasatellite* (MaYVYnB)
32. *Mimosa yellow leaf curl virus-associated DNA-1*
33. *Okra leaf curl betasatellite* (OLCuB)
34. *Okra leaf curl virus-associated DNA-1*
35. *Okra yellow vein mosaic virus-associated DNA-1*
36. *Papaya leaf curl betasatellite* (PaLCuB)
37. *Sida leaf curl betasatellite* (SiLCuB)
38. *Sida leaf curl virus-associated DNA-1*
39. *Sida yellow mosaic China betasatellite* (SiYMCNB)
40. *Sida yellow vein betasatellite* (SiYVB)
41. *Sida yellow vein Vietnam betasatellite* (SiYVVNB)
42. *Sida yellow vein Vietnam virus-associated DNA-1*
43. *Siegesbeckia yellow vein betasatellite* (SibYVB)
44. *Siegesbeckia yellow vein Guangxi betasatellite* (SibYVGxB)
45. *Tobacco curly shoot betatsatellite* (TbCSB)

46. *Tobacco curly shoot virus-associated DNA-1*
47. *Tobacco leaf curl betasatellite* (TbLCB)
48. *Tobacco leaf curl virus-associated DNA-1*
49. *Tomato leaf curl Yunnan virus-associated DNA-1*
50. *Tomato leaf curl Bangalore betasatellite* (ToLCBB)
51. *Tomato leaf curl Bangladesh betasatellite* (ToLCBDB)
52. *Tomato leaf curl betasatellite* (ToLCB)
53. *Tomato leaf curl China betasatellite* (ToLCCNB)
54. *Tomato leaf curl Java betasatellite* (ToLCJB)
55. *Tomato leaf curl Joydebpur betasatellite* (ToLCJoB)
56. *Tomato leaf curl Laos betasatellite* (ToLCLB)
57. *Tomato leaf curl Maharastra betasatellite* (ToLCMaB)
58. *Tomato leaf curl Philippines betasatellite* (ToLCPHB)
59. *Tomato leaf curl virus-associated DNA-1*
60. *Tomato yellow leaf curl China betasatellite* (TYLCCNB)
61. *Tomato yellow leaf curl China virus-associated DNA-1*
62. *Tomato yellow leaf curl Thailand betasatellite* (TYLCTHB)
63. *Tomato yellow leaf curl Thailand virus-associated DNA-1*
64. *Tomato yellow leaf curl Vietnam betasatellite* (TYLCVNB)
65. *Zinnia leaf curl betasatellite* (ZLCuB)

Tentative members of the Genus

1. *Ageratum leaf curl Cameroon betasatellite*
2. *Cotton leaf curl Rajasthan virus-associated DNA-1*
3. *Gossypium darwinii symtomless alphasatellite*
4. *Gossypium davidsonii symptomless alphasatellite*
5. *Gossypium mustilinum symptomless alphabetasatellite*
6. *Mentha leaf deformity associated satellite DNA-II*
7. *Okra leaf curl virus associated DNA-1*
8. *Tomato leaf curl Pakistan virus-associated DNA-1*
9. *Tomato leaf curl Patna virus betasatellite*.