

**MOLECULAR CHARACTERIZATION OF CASSAVA MOSAIC
GEMINIVIRUSES IN TANZANIA**

by

Joseph Ndunguru

Submitted in partial fulfillment
of the requirements
for the degree of

Doctor of Philosophy

In the faculty of Natural and Agricultural Sciences
Department of Microbiology and Plant Pathology

University of Pretoria

Pretoria

December 2004

DEDICATION

This PhD thesis is a dedication to my Father Mr Canisius Manyahy Ndunguru and to the Lord for the gift of the Holy Spirit

DECLARATION LETTER

I, the undersigned, declare that the thesis, which I hereby submit for the degree of Doctor of Philosophy at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution

Joseph Ndunguru

November 2004

Molecular characterization of cassava mosaic geminiviruses in Tanzania

By

Joseph Ndunguru

Supervisor: Prof T. A. S. Aveling

Co-Supervisors: Dr G. Thompson (ARC-Institute for Industrial Crops, Private Bag X82075, Rustenburg 0300, South Africa),
Dr James Legg (International Institute of Tropical Agriculture-Eastern and Southern Regional Center and Natural Resource Institute, Box 7878, Kampala, Uganda),
Dr Claude Fauquet (International Laboratory for Tropical Agricultural Biotechnology, Danforth Plant Science Center, St. Louis, MO 63132, USA).

Department: Microbiology and Plant Pathology
University of Pretoria

Degree: Doctor of Philosophy (PhD)

Abstract

Cassava (*Manihot esculenta* Crantz) is a basic staple food crop in Tanzania. Cassava mosaic disease (CMD) caused by cassava mosaic geminiviruses (CMGs) constitutes a major limiting factor to cassava production in the country. This study was undertaken to characterize the CMGs occurring in Tanzania using molecular techniques and to map their geographical distribution to generate information on which the formulation of control measures can be based. Using Polymerase Chain Reaction (PCR) and Restriction Fragment Length Polymorphism (RFLP) for analysis of CMGs DNA-A genomes, different CMGs were found to be associated with CMD. Higher molecular diversity was observed among East African cassava mosaic viruses (EACMVs) than African cassava mosaic viruses (ACMVs), which was confirmed later by complete nucleotide sequence analysis. In addition to EACMV and ACMV isolates, two isolates of EACMV Cameroon virus (EACMCV) were found in Tanzania. These were confirmed to be strains of EACMCV Cameroon, originally described in Cameroon, West Africa and here named EACMCV- [TZ1] and EACMCV-[TZ7]. They had high (92%) overall DNA-A nucleotide sequence identity and EACMCV-[TZ1] was widespread in the southern part of the country. A

subgenomic DNA form of CMG that appeared to be truncated was identified in a CMD-infected cassava plant. It was confirmed upon sequence analysis to be a defect of EACMV DNA-A and had a capacity of attenuating symptoms when coinoculated with wild-type EACMV. In addition, this study revealed for the first time the presence of two novel non-geminivirus single-stranded DNA (ssDNA) sub-genomic molecules associated with CMG infection. They were shown to be dependent on CMG for replication and movement within the plants, confirming their status as satellite molecules named here as satDNA-II and satDNA-III. When present in coinfection with CMGs, they enhance symptoms and can break high levels of resistance in a cassava landrace. Finally a simple, inexpensive technique is described of archiving, transporting and recovering plant DNA for downstream geminivirus characterisation.

Key words: *African cassava mosaic virus*, cassava, cassava mosaic disease, cassava mosaic geminiviruses, *East African cassava mosaic virus*, DNA-A, DNA-B molecular characterisation, Tanzania

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following people for their invaluable contributions toward completion of this study:

Prof T. A. S. Aveling for supervising, directing and guiding this study. Your positive advice and inspiration rendered to me will always be remembered.

Drs James Legg, G. Thompson and C. M. Fauquet for co-supervising this work and for your scientific support.

International Institute of Tropical Agriculture (IITA) and United Kingdom Department for International Development (DFID) for financing this study.

The Lake Zone Agricultural Research and Development Institute (LZARD), Ukiriguru, Tanzania for providing transport during a countrywide collection of samples.

The personnel and students of the Department of Microbiology and Plant Pathology, University of Pretoria, for all their kind support in various ways.

Mr Alloyce Cyprian for assistance during sample collection in Tanzania and encouragement and support.

Mr R. Obonyo and G. Okao-Okuja for their technical support during preliminary molecular analysis of DNA samples in Uganda.

All the International Laboratory for Tropical Agricultural Biotechnology (ILTAB) members in St. Louis, USA especially Ben Fofana and Mrs Maya K. for their good friendship, encouragement and support.

My Father Mr Ndunguru C. for his love. This thesis is his fruit.

My Mother Agnes M. Peace to her soul.

Dr Regina for her support and encouragements.

Dr Pita Psseruwagi for encouragement and moral support.

Mrs Anna Ndunguru, my spouse, who so admirably showed patience, love and encouragement.

My children, Ezra Ndunguru and Gloria Ndunguru for their love, endurance and support.

My sisters Imelda, Bernadetha, Maria, Nyapili, and Evangelista for support

Drs E. Kanju, Katherine. K. for their encouragement and support.

Mrs F. Katagira for her encouragement and support.

CONTENT

ABSTRACT	iv
ACKNOWLEDGEMENT	vi
CONTENT	vii
LIST OF ACRONMYS AND ABBREVIATIONS	xiii
LIST OF TABLES	xv
LIST OF FIGURES	xvi
CHAPTER 1 GENERAL INTRODUCTION	1
1. 1 References	4
CHAPTER 2 LITERATURE REVIEW	8
2.1 CASSAVA (<i>MANIHOT ESCULENTA CRANTZ</i>)	8
2.1.1 Introduction	8
2.1.2 Economic importance of cassava	8
2.1.3 Cassava production in Tanzania	9
2.1.4 Uses of cassava in Tanzania	10
2.1.5 Production constraints	10
2.2 CASSAVA MOSAIC GEMINIVIRUSES	11
2.2.1 Introduction	11
2.2.2 Cassava mosaic disease in Tanzania	12
2.2.3 Cassava mosaic geminivirus species	13
2.2.4 Geographical distribution	13
2.2.5 Impact of cassava mosaic geminiviruses on yield	14
2.2.6 Host ranges	15
2.2.7 Transmission and spread	16
2.2.8 Symptom expression	17
2.2.9 Genome organization	18
2.2.10 Cassava mosaic geminivirus gene products and their role in pathogenicity	20

2.2.11	Cassava mosaic geminivirus replication	23
2.2.12	Virus movement within the host plants	26
2.2.13	Recombination, pseudorecombination and its role in viral pathogenicity	26
2.2.14	Molecular identification and characterization	29
2.2.15	Cassava mosaic geminiviruses in East Africa	30
2.2.16	Synergism between virus species	31
2.2.17	Control strategies	32
2.3	References	34

CHAPTER 3 RESTRICTION AND SEQUENCE ANALYSIS OF PCR-AMPLIFIED VIRAL DNAs SUGGESTS THE EXISTANCE OF DIFFERENT CASSAVA MOSAIC GEMINI VIRUSES ASSOCIATED WITH CASSAVA MOSAIC DISEASE IN TANZANIA

		54
3.1	Introduction	55
3.2	Materials and methods	56
3.2.1	Virus collection	56
3.2.2	DNA extraction	57
3.2.3	PCR analysis	57
3.2.4	DNA precipitation and RFLP analysis	58
3.2.5	Molecular cloning of viral DNA	58
3.2.6	Sequence determination and analysis	58
3.3	Results	60
3.3.1	Cassava sample collection	60
3.3.2	PCR and RFLP-based molecular variability of CMGs in Tanzania	60
3.3.3	Phylogenetic analysis of the AC1 gene and comparison with selected viruses	62
3.3.4	Relationship between virus type, symptom and disease severity	62
3.3.5	Geographical distribution of CMGs in Tanzania	64
3.3.6	Detection of coinfection in cassava plants by RFLP analysis	65
3.4	Discussion	66

3.5 Acknowledgements	69
3.6 References	69
CHAPTER 4 CHARACTERIZATION OF A NATURALLY-OCCURRING TRUNCATED CASSAVA MOSAIC GEMINIVIRUS ASSOCIATED WITH CASSAVA MOSAIC DISEASE	83
4.1 Introduction	84
4.2 Materials and methods	85
4.2.1 Source of virus DNA	85
4.2.2 PCR, cloning and sequencing	85
4.2.3 Sequence analysis	86
4.2.4 Construction of an infectious clone	87
4.2.5 Infectivity assay	87
4.2.6 Southern blot analysis	87
4.3 Results	88
4.3.1 Symptoms on cassava plants	88
4.3.2 PCR	88
4.3.3 Identification of the helper virus	89
4.3.4 Sequence analysis and comparison	89
4.3.4.1 Phylogenetic comparisons of df DNA-A 15 with 23 other begomoviruses infecting cassava	89
4.3.4.2 Analysis of df DNA-A ORFs	91
4.3.5 df DNA-A 15 ameliorates symptoms of EACMCV but not of ACMV-[CM]	91
4.3.6 Mediation of df DNA-A systemic movement by cassava mosaic geminiviruses DNA	91
4.4 Discussion	92
4.5 Acknowledgements	94
4.6 References	94

CHAPTER 5 TWO NOVEL SATELLITE DNA MOLECULES ASSOCIATED WITH BIPARTITE CASSAVA MOSAIC BEGOMOVIRUSES ENHANCE SYMPTOMS AND BREAK RESISTANCE IN CASSAVA GERMPLASM	106
5.1 Introduction	107
5.2 Materials and methods	109
5.2.1 Virus sources and DNA extraction	109
5.2.2 Polymerase chain reaction for amplification of viral DNA	109
5.2.3 Cloning and sequencing	110
5.2.4 Genome sequence analysis	110
5.2.5 Construction of infectious clones	110
5.2.6 Investigation of a possible biological role of satDNA-II and III	111
5.2.7 Southern blot analysis	111
5.3 Results	111
5.3.1 Disease symptoms in cassava plants containing satellite molecules	111
5.3.2 Isolation of satDNA-II and satDNA-III	112
5.3.3 Sequence analysis	113
5.3.4 Structure features of satDNA-II	114
5.3.5 Structure features of satDNA-III	115
5.3.6 Effect of satDNA-II on CMG symptom expression	115
5.3.7 Enhancement of CMG disease symptoms by satDNA-III	116
5.3.8 Breaking of TME3 natural resistance by satDNA-II and satDNA-III	117
5.3.9 Effect of satDNA-II and satDNA-III on the accumulation of CMG DNA	118
5.4 Discussion	119
5.5 Acknowledgements	123
5.6 References	123

CHAPTER 6 MOLECULAR BIODIVERSITY OF CASSAVA

BEGOMOVIRUSES IN TANZANIA: EVIDENCE FOR THE PRESENCE OF STRAINS OF *EAST AFRICAN CASSAVA MOSAIC CAMEROON VIRUS* AND RECOMBINATION IN CASSAVA GEMINIVIRUSES

	140
6.1 Introduction	141
6.2 Materials and methods	143
6.2.1 Collection of plant samples	143
6.2.2 Symptom reproduction in the growth chamber	143
6.2.3 DNA preparation	143
6.2.4 Polymerase chain reaction, cloning and sequencing	143
6.2.5 Computer analysis of CMG sequences	145
6.2.6 Recombination analysis for cassava geminiviruses	145
6.3 Results	146
6.3.1 Assessment of CMD symptoms	146
6.3.2 Detection of viral genomic components	147
6.3.3 Complete nucleotide sequence characteristics of CMGs from Tanzania	147
6.3.4 Coat protein (CP) gene sequence analysis and comparison with selected viruses	149
6.3.5 The common regions (CRs) of the Tanzanian CMGs	150
6.3.6 Comparisons of the East African and West African isolates of EACMCV	151
6.3.6.1 Comparisons of the A components of EACMCV-[TZ]	151
6.3.6.2 Comparisons of the B components of EACMCV	151
6.3.6.3 Comparisons of the common regions (CRs) of EACMCV from Cameroon and Tanzania	152
6.3.7 Determination of genetic diversity of EACMV DNA-B using partial sequence analysis	152
6.3.8 Recombination analysis	153
6.3.8.1 Pairwise analysis of the A components	153
6.3.8.2 Pairwise analysis of B components	154

6.4 Discussion	155
6.5 Acknowledgements	158
6.6 References	159
CHAPTER 7 USE OF PLANT DNA STORED ON FTA CARDS FOR RECOVERY AND MOLECULAR CHARACTERIZATION OF BIPARTITE AND MONOPARTITE GEMINIVIRUSES	181
7.1 Introduction	182
7.2 Materials and methods	184
7.2.1 Sampling of plants	184
7.2.2 Sample processing from FTA cards for PCR	184
7.2.3 PCR analysis for cassava mosaic geminiviruses	184
7.2.4 Detection of dual infection in infected cassava	185
7.2.5 Cloning, sequencing and sequence analysis	185
7.2.6 Quantification of DNA and viral DNA eluted from FTA card	186
7.3 Results	186
7.3.1 Polymerase chain reaction for detection of viral DNA	186
7.3.2 Detection of CMG species	187
7.3.3 Clones and sequence comparisons of viral DNA from phenol and FTA- processed DNA	188
7.3.4 Quantification of DNA and viral DNA eluted from FTA cards	188
7.4 Discussion	189
7.5 Acknowledgements	191
7.6 References	191
CHAPTER 8 GENERAL DISCUSSION	204
8.1 References	208
SUMMARY	211

LIST OF ACRONYMS AND ABBREVIATIONS

ACMV	African cassava mosaic virus
AYVV	Ageratum yellow vein virus
CBSD	cassava brown streak virus disease
CBSV	cassava brown streak virus
CMD	cassava mosaic disease
CMGs	cassava mosaic geminiviruses
CP	coat protein
CLV	cassava latent virus
CR	common region
CM	Cameroon
COSCA	collaborative study of cassava in Africa
ClcuMV	cotton leaf curl mosaic virus
°C	degree Celsius
df	defective
DNA	deoxyribonucleic acid
DIFID	Department for International Development, UK
DI	defective interfering
DRC	Democratic Republic of Congo
EACMV	East African cassava mosaic virus
EACMMV	East African cassava mosaic Malawi virus
EACMZV	East African cassava mosaic Zanzibar virus
FAO	Food and Agricultural Organization of the United Nations
Fig.	figure
ICMV	Indian cassava mosaic virus
IITA	International Institute of Tropical Agriculture
ILTAB	International Laboratory for Tropical Agricultural Biotechnology
IR	intergenic region
kbp	kilobase pair
KE	Kenya
L	liters
MAFS	Ministry of Agriculture and Food Security

min	minutes
MAS	marker-assisted selection
MP	movement protein
NG	Nigeria
NSP	nuclear shuttle protein
ORF	open reading frame
PAUP	phylogenetic analysis using parsimony
PCR	polymerase chain reaction
PCNA	proliferating cell nuclear antigen
PNACL	protein and nucleic acid chemistry laboratory
REn	replication enhancer
RFLP	restriction fragment length polymorphism
satDNA	satellite deoxyribonucleic acid
SACMV	South African cassava mosaic virus
siRNA	short interfering ribonucleic acid
SLCMV	Sri-Lankan cassava mosaic virus
ssDNA	single-stranded deoxyribonucleic acid
SqLCV	squash leaf curl virus
TGMV	tomato golden mosaic virus
ToLCV	tomato leaf curl virus
TrAP	transcription activating protein
TZ	Tanzania
UG	Uganda
μl	microlitter
μg	microgram
WTGS	whitefly-transmitted geminiviruses

LIST OF TABLES

Table 3.1 Extent of CMG coinfection in cassava plants with CMD symptoms	74
Table 4.1 Cassava mosaic geminiviruses used in comparative sequence analysis and their respective GenBank accession numbers	98
Table 6.1 List of the oligonucleotide primers used in this study for amplification of cassava mosaic geminiviruses from Tanzania (^a nfl = near-full length, ps = partial sequence)	163
Table 6.2 Nucleotide sequence identities (percentages) of the DNA-A full length of cassava mosaic geminiviruses from Tanzania and other geminiviruses from the African and Indian Sub-continent. Values above 89% are in bold in grey boxes and isolates from Tanzania are in blue	164
Table 6.3 CP gene nucleotide sequence identity (%) between the selected cassava and other begomoviruses. Values above 80% are in bold	165
Table 6.4 Percent similarity (in the upper triangle) of nucleotide sequence of common region of East and West African isolates of EACMCV	166
Table 7.1 Oligonucleotides used for PCR amplification of CMGs and MSV where K = G+T, R = A + G, S = G + C (Invitrogen, Life Science, USA) (^a nfl = near-full length)	194

LIST OF FIGURES

- Figure 2.1** Bipartite genome organization of cassava mosaic geminiviruses. The part of intergenic region whose sequence is identical in both begomovirus components is called the common region (CR) and includes a sequence capable of forming a stable hairpin loop structure. CP, coat protein; MP, movement protein; Rep, replication initiator protein; TrAP, transcription activator protein; Ren, replication enhancer protein. The proteins encoded by the BV1 and BC1 ORFs are both movement proteins. MP, movement protein; NSP, nuclear shuttle protein. 19
- Figure 2.2** Begomovirus replication in the nucleus of the plant cell. On the infection of plant and following uncoating of genomic component, the encapsidated ssDNA is converted into transcriptionally active dsDNA replicative form prior to gene expression by the synthesis of complementary-sense DNA. The RF dsDNA serve as a template for virus-sense strand synthesis to generate free ssDNA, which then (i) re-enter the DNA replication pool, (ii) associate with CP or (iii) be transported outside the nucleus and to the neighbouring cell, most probably through plasmodesmata, with the help of viral MPs. 25
- Figure 3.1** Restriction fragment length polymorphism patterns of 2.8 Kb DNA fragments of selected putative cassava mosaic geminiviruses. EACMV-like virus pattern of DNA fragments digested with *MluI* (Top a): (a) EACMV-[TZ], (b) EACMV-[TZ1], (c) EACMV-[TZ2], (d) EACMV-[TZ3], (e) EACMV-[TZ5], (f) EACMV-[TZ6], (g) EACMV-[TZ8], (h) EACMV-[TZ9], (i) EACMV-[TZ10], and (j) EACMV-[TZ11]. (Bottom b): Banding patterns showing detection of virus dual infection of EACMV and ACMV by restriction analysis with *EcoRV* endonuclease. DNA size markers are indicated in base pairs in the left margins. 75
- Figure 3.2** Restriction fragment length polymorphism patterns of 2.8 Kb DNA fragment of cassava mosaic geminiviruses, EACMV-[TZ Mara], EACMV-[UG2], ACMV-[NG] and ACMV-[TZ]. DNA fragments were separately digested with *EcoRV* and *MluI*. 76

- Figure 3.3** Neighbor-joining phylogenetic dendrogram based upon alignments of partial AC1(360 nts) sequences of CMG DNA-A (Tanzanian isolates) and selected other begomoviruses. The sequences were either generated by forward (UniF) (a) or reverse (UniR) (b) primers which anneal at position 1700-1834 of the AC1 gene on the complementary sense of the DNA-A component. Numbers at nodes indicate percentage bootstrap scores (1000 replicates). 77
- Figure 3.4** Cassava mosaic disease symptoms caused by selected cassava mosaic geminiviruses in farmers' fields (a-h). Symptoms are on cassava cultivars; (a) Lyongo (caused by ACMV-[NG]), (b) Nakalai (EACMV-[TZ1]), (c) Jawa (EACMV-[TZ6]), (d) Gajaigaja (EACMV-[UG2]), (e) Unnamed local (mixed infection of EACMV-[UG2] and ACMV-[NG]), (f) Msufi (EACMV-[TZ2]), (g) Jawa (EACMV-[TZ3]), (h) Kigoma (EACMV-[TZ Mara]). 81
- Figure 3.5** Map of Tanzania illustrating the distribution of the most prevalent cassava mosaic geminiviruses in the major cassava-growing zones. Only the widespread cassava mosaic geminiviruses are listed and the rest are shown as "others". 82
- Figure 4.1** Disease symptoms of CMG (a) on cassava field plant containing the df DNA-A 15 and (b) symptoms reproduced in the growth chamber. *Nicotiana benthamiana* plants inoculated with (c) ACMV-CM, (d) ACMV-CM + df DNA-A15, (e) EACMCV, (f) EACMCV + df DNA-A 15 and (d) df DNA-A 15 alone 14 days after inoculation. 99
- Figure 4.2.** Isolation and characterization of a circular 1.5 kbp subgenomic defective DNA molecule associated with CMG natural infection. (a) PCR-amplification of df DNA-A 15 from field sample number 15 and a near-full length 2.8 kbp of the helper virus (EACMV-TZ15) from the same plant. Lane -C was loaded with DNA extracted from a healthy cassava plant to serve as a negative control. (b) The characteristic banding patterns obtained after treatment of the PCR amplification products (2.8 kbp) of the helper virus with the restriction endonucleases *EcoRV* and *MluI* are indicated. DNA products were electrophoresed through an ethidium bromide-stained 1%

agarose gel in 1% TAE buffer. Sizes of the DNA marker (M) are shown.

100

Figure 4.3a Phylogenetic tree obtained from the alignment of complete nucleotide sequences of DNA-A of 22 cassava mosaic geminiviruses with df DNA-A 15. The tree was generated using the Neighbor-joining method with the Mac Vector 7.2 (Accelry, San Diego, Inc. USA) computer program. The numbers at each branch indicate the percentage of 1000 bootstraps, which support the grouping at that node.

101

Figure 4.3b Consensus phylogenetic tree (1000 bootstrap replication) obtained from comparison of the CR sequences of selected cassava mosaic geminiviruses with the df DNA-A 15. Numbers indicate the percentage of boot strap for each branch.

102

Figure 4.4 Alignment of the nucleotide sequences of the CR of df DNA-A 15 with similar sequences of the EACMV isolates selected from GenBank. The iterated sequences (iterons), TATA box, the variable region and the invariant nononucleotides (TAATATTAC) containing the origin of replication are indicated.

103

Figure 4.5 Genome organization of the subgenomic df DNA-A 15. Predicted ORFs in both orientations ('C' for complementary and 'V' for virion sense) are shown. The missing part of the genome based on the known genome of a full-length DNA-A component of EACMV is shown (dashed).

104

Figure 4.6 Replication of df DNA A 15 in *Nicotiana benthamiana* inoculated biolistically with clones of (a-b) ACMV-[CM] (DNA-A and B), (c) EACMCV (DNA-A and DNA-B). A blot of the DNA from original cassava obtained from the infected field plant is included in panel d in duplicate wells (e) accumulation of EACMCV DNA-B when inoculated singly or with df DNA-A 15. Position of the replicative forms linear (Lin), open circular (OC), single stranded (SS), and super coiled (SC) are indicated.

105

Figure 5.1 Symptoms caused by the association of DNA satellites to cassava mosaic geminiviruses. Pictures (a-c) represent close-up leaf symptoms from

plants imported from a Tanzanian fields, while (g-l) represent symptoms of plants inoculated in the laboratory with DNA infectious clones using particle bombardment.

(a) Healthy plant from a Tanzanian cassava cv.; (b) leaf from a Tanzanian cassava cv. in which satDNA-II was detected; (c) leaf from a Tanzanian cassava cv. in which satDNA-III was detected.

(d-i) CMD symptoms on cassava cv. 60444 7dpi inoculated with different combinations of geminiviruses and satellites: (d)- ACMV-[CM]; (e)- ACMV-[CM] + satDNA-II; (f)- EACMCV-[CM]; (g)- EACMCV-[CM] + satDNA-III; (h)- EACMV-UG2, (i)- EACMV-UG2 + satDNA-II.

(j) *Nicotiana benthamiana* plants inoculated with the combination EACMCV-[CM] DNA-A and DNA-B (left) and the combination EACMCV-[CM] DNA-A and DNA-B plus satDNA-III (right), 35 dpi;

(k-l) virus resistant cassava landrace TME3 inoculated with the combination: (k)- EACMV-UG2 DNA-A and EACMV-UG3 DNA-B; (l)- EACMV-UG2 DNA-A and EACMV-UG3 DNA-B plus satDNA-II, 60 dpi.

129

Figure 5.2 PCR detection of cassava mosaic virus-associated satellite DNA

molecules. Total DNA was extracted from leaves of field-collected cassava where mosaic symptoms were observed. (a) satDNA-II; (b) satDNA-III. SatDNA-II and satDNA-III were amplified using the primer pairs Beta01/Beta02 and DNA1-F/DNA1-R respectively; lanes 1-4 represent a samples for infected symptomatic cassava plants; (-c) represents sample from healthy control cassava; M: DNA marker; Arrows on the right indicate the expected fragment for satDNA-II (1.03 kbp) and satDNA-III (1.2 kbp); Lane b1 represents a CMG-infected plant but which is PCR-negative for satDNA-III. kbp: kilo base pair

131

Figure 5.3 Complete nucleotide sequence of cassava mosaic virus associated

satellites DNA molecules. (a) satDNA-II and (b) satDNA-III. The region of high GC composition in both satellites is boxed. Direct repeat sequences are shaded black. A consensus putative transcript polyadenylation signal (AATAAA) is indicated by gray shade with an underline. The putative TATA boxes in the two satellites are shaded gray. The putative TTGTA motif frequently associated with functional polyadenylation signals (Rothnie *et al.*, 1994) is positioned upstream of the polyadenylation signal

shown in bold and boxed. For sat DNA-III, the putative ACGTCA motif believed to be associated with strong promoter activity in undifferentiated actively dividing cell types (Nakayama *et al.*, 1992) is indicated in bold and underlined. The predicted AP-1 binding motifs are underlined.

The position of primers Beta01 and DNA1-F used for amplification of satDNA-II and satDNA-III, respectively, are indicated. 133

Figure 5.4 Genome organization of cassava mosaic geminivirus-associated satellites DNA molecules. (a) satDNA-II; (b) satDNA-III. The putative genes encoded on either virion-sense (V) or complementary sense (C) are indicated by dark arrows and their direction of transcription is shown. The name of each open reading frame is shown in bold type. The hatched regions represent the GC-rich region present in both satDNA-II and satDNA-III. 135

Figure 5.5 Accumulation of viral DNA in systemic leaves of *Nicotiana benthamiana* and cassava plants of landrace TME3. Southern blots show detection of EACMCV-[CM], EACMV-UG2, satDNA-II and satDNA-III using specific probes. (a) *N. benthamiana* plants inoculated with the combinations (EACMCV-[CM] DNA-A and DNA-B), (EACMV-UG2 DNA-A and EACMV-UG3 DNA-B), satDNA-II and satDNA-III alone or in combination; (b) TME3 plant were inoculated with the combination (EACMV-UG2 DNA-A and EACMV-UG3 DNA-B) and SatDNA-II alone or in combination. Samples were collected at 14 days post-inoculation (dpi) for *N. benthamiana* and 21 dpi for TME3, inoculated with viral DNA infectious clones. DNA accumulation was quantified and the values indicated under each lane for each blot refer to the percentage of viral DNA amount. 136

Figure 5.6 Accumulation of viral DNA in systemic leaves of *Nicotiana benthamiana* and cassava plants. Southern blots show detection of EACMCV-[CM], EACMV-UG2, satDNA-II and satDNA-III using specific probes. (a) accumulation of satDNA-II and satDNA-III in cassava field samples. -control corresponds to DNA samples from CMD-affected cassava plants that tested negative for satDNA-II and satDNA-III (b) accumulation of satDNA-II and satDNA-III in *N. benthamiana* and TME3 inoculated with viral DNA infectious clones. Only one sample corresponds to DNA from TME3 and is indicated

on the figure. + control corresponds to full-length satDNA-II and satDNA-III digested from plasmid infectious clone. 138

Figure 6.1 (a) Pairwise analysis of begomoviruses in the old world that do not exhibit putative recombinant fragments, at the species level (green curve) and at the strain level (blue curve). (b) Pairwise analysis of EACMCV-[TZ1] A component, paired with the sequence of the A component of other cassava geminiviruses like EACMCV-[TZ7] (blue line), ACMV-[TZ] (brown line), EACMV-KE (red line) and EACMZV-ZB (green line), showing the recombinant fragment of this virus (1200 – 2000 nts) as well as the one from EACMZV-ZB (2000 – 2900 nts). The linearized genome organization of these geminiviruses is depicted at the bottom of the graph. 167

Figure 6.2 CMD symptoms on naturally infected cassava plants (a,c, e and g) in the field with their corresponding plants raised from field collected cuttings maintained in the growth chamber (b, d, f and h). Only plants containing single virus infection are shown. Plants (a) and (b) contained a single infection of EACMV-KE[TZM] virus strain, (c) and (d) contained ACMV-[TZ], (e) and (f) were infected by EACMCV-[TZ1] and (g) and H by EACMV-UG2[TZ10]. 169

Figure 6.3 Phylogenetic tree (100 boot strap replications) showing the DNA-A complete nucleotide sequence relationships between the seven Tanzanian cassava mosaic geminiviruses strains and other selected begomoviruses. *Tomato golden mosaic virus* (TGMV) (K02029) was used as the out group. Abbreviations and accession numbers: EACMCV-[CM], *East African cassava mosaic virus*-Cameroon (AF112354); EACMCV-CM[IC], *East African cassava mosaic virus*-Cameroon Ivory Coast (AF259896); EACMV-KE-K2B, *East African cassava mosaic virus* (Isolate K2B) (Z83258); EACMV-[TZ], *East African cassava mosaic virus*-Tanzania (Z53256); EACMV-UG2Svr, *East African cassava mosaic virus*-Uganda2 severe (AF126806); EACMV-UG2Mld, *East African cassava mosaic virus*-Uganda2 mild (AF126804); EACMV-UG2, *East Africa cassava mosaic virus*-Uganda2 (Uganda variant) (Z83257); EACMMV-MW[MH], *East African cassava mosaic virus*-Malawi [MH] (AJ006459); EACMMV-WM[K], *East African cassava mosaic Malawi virus*-[K] (AJ006460); EACMZV, *East African*

cassava mosaic Zanzibar Virus (AF422174); EACMZV-Kekil, *East African cassava mosaic Zanzibar virus-Kenya* [Kil] (AJ516003); SACMV, *South African cassava mosaic virus* (AF155807); SACMV-[M12], *South African cassava mosaic virus*-(Isolate M12) (AJ422132); ACMV-[IC], *African cassava mosaic virus*-[Ivory Coast] (AF259894); ACMV-[Nig-Ogo], *African cassava mosaic virus*-[Nigeria-Ogo] (AJ427910); ACMV-[NG], *African cassava mosaic virus*-[Nigeria] (X17095); ACMV-[CM], *African cassava mosaic virus*-[Cameroon] (AF112352); ACMV-[CM/D02], *African cassava mosaic virus*-[Cameroon D02] (AF366902); ACMV-UGMld, *African cassava mosaic virus*-Uganda mild (AF126800); ACMV-UGSvr, *African cassava mosaic virus*-Uganda severe (AF126802); ACMV-[KE], *African cassava mosaic virus*-[Kenya] (J02057); SLCMV-[Col], *Sri-Lanka cassava mosaic virus*-[Colombo] (AF314737). 171

Figure 6.4 Consensus phylogenetic tree (1000 bootstrap replications) obtained from comparison of the complete nucleotide sequence of EACMCV-[TZ1] DNA-B and selected cassava mosaic geminiviruses DNA-B components (Abbreviations and GenBank accession numbers are indicated in the Phylogenetic tree). 173

Figure 6.5 Relationship dendrogram of the coat protein gene (CP) nucleotide sequences of the cassava mosaic geminiviruses strains from Tanzania and other selected begomoviruses. The tree was constructed using PAUP (1000 bootstrap replications) and reconstructed using Discovery Studio (DS) Gene software for Windows 1.5 (Accelrys Inc, USA). Sequence of *tomato golden mosaic virus* (TGMV) was used as the outgroup. For more abbreviations and accession numbers for the published sequences see figure 6.3. Bootstrap percent values more than 50 are numbered along branches. 174

Figure 6.6a Alignment of common region nucleotide sequences of DNA-A of the EACMV strains from Tanzania with the closely related isolates of EACMV from the Database sequences. The TATA box for AC1 is boxed and indicated. The putative CR iterative sequences (iterons) are boxed and indicated with arrows. The conserved nonanucleotide sequences TAATATTAC together with its stem loop are boxed and shown. Mismatched nucleotides are highlighted in white. 175

Figure 6.6b Alignment of common region nucleotide sequences of DNA-A of the ACMV-[TZ] strain from Tanzania with its closely related selected isolates of ACMV from Africa derived from the Database sequences. The TATA box for AC1 is boxed and indicated. The putative common region iterative sequences (iterons) are boxed and indicated with arrows. The conserved nonanucleotide sequences TAATATTAC together with its stem loop are boxed and shown. Mismatched nucleotides are highlighted in white. 176

Figure 6.7 Alignment of the nucleotide sequences (DNA-A and –B) of the common region of EACMCV isolates from West Africa and Tanzania. Large boxes indicate the positions of iterons, TATA box and the conserved TAATATTAC stem loop. 177

Figure 6.8 Linearized recombination map of putative recombinant fragments for the (a) top and (b) bottom components of cassava geminiviruses. Each horizontal box represents one genotype and the color coded boxes represent the origin of the putative recombinant fragments. The length of the genomes is indicated on the top of each diagram and the genome organization is depicted on the bottom, while the name of the viruses are listed on the left. The color code for the recombinant fragments is indicated in the boxes at the bottom of each diagram. The vertical arrows indicate the position of possible “hot spots” for recombination. For clarity of the figure, we only indicated the name of one isolate for each genotype and those from Tanzania. 178

Figure 6.9 Phylogenetic tree (1000 boot strap replications) showing the relationship between the 13 different Tanzanian EACMV isolates (DNA-B component nucleotide sequences) and selected cassava mosaic geminiviruses. Virus abbreviations follow those indicated in Figure 6.3 legends. The accession numbers for the reference EACMV DNA-B components are shown against each name in this figure. Bootstrap percent values more than 50 are numbered along branches. 180

Figure 7.1 Application of samples collected from infected (a) cassava and (b) maize on FTA card (c). Samples were dried and stored on FTA card at room temperature after which three discs (2 mm) were

punched from each sample area of the FTA card (panel D), pooled and the total plant DNA eluted for use in PCR analysis

of cassava mosaic geminiviruses and *maize streak virus* (MSV). 195

Figure 7.2 PCR amplification of near-full length DNA-A of CMG using universal primers (UNIF and UNIR) either on DNA processed by the Dellaporta method (a) or eluted from FTA card (b). In addition, primers Universal 1 and Universal 2 (Table 7.1) were used for PCR amplification of ~ 1.0 kbp DNA-A fragment from DNA eluted from FTA card (c). Samples were CMD-infected cassava plants held in the greenhouse. 196

Figure 7.3 PCR amplification of EACMV and ACMV species on total nucleic acid from cassava leaves processed either by the Dellaporta (phenol purified) method or FTA elution. Amplifications are: 555 bp (IR/BC1) fragment of EACMV DNA-B component (panels a and b) and 770 bp coat protein gene of ACMV (panels c and d). Negative controls (-C) contained nucleic acid from healthy cassava leaves. M = 1 Kbp plus DNA ladder (for sizes see Figure 7.2). 197

Figure 7.4 PCR amplification of (a) 0.5 kbp Maize streak virus DNA fragment from infected maize plant leaf samples collected from Kenya and Malawi on FTA cards, (b) NPTII transgene in genetically modified cassava plants held in the greenhouse. Samples (lanes 1-5) were collected on FTA cards and the genomic DNA processed for PCR amplification. Negative control samples (-C) consisted of DNA extracted from healthy cassava plants 198

Figure 7.5 Restriction enzyme digestion of (a) near-full length DNA-A of CMG for detection of ACMV/EACMV dual infection. Samples were collected from CMD-infected plants in the fields, archived on FTA cards, DNA eluted and used for PCR analysis. DNA was recovered from the agarose gel, purified and digested with *EcoRV* for 1.5 hrs at 37 °C (b). Recombinant plasmids carrying 556 bp DNA-B viral fragment (IR/BC1) of EACMV. PCR was done on FTA or phenol processed template DNA, cloned in pGEM-T Easy vector (Promega) and recombinant plasmids recovered by miniprep analysis by digestion with *EcoRI* to release the viral inserts. 199

Figure 7.6 Nucleotide sequence comparison of EACMCV-CM clones of viral DNA

fragment obtained from FTA or phenol-purified PCR DNA templates. A 556 bp PCR product was cloned and sequenced and sequences of duplicate clones were compared by Multiple Sequence Alignment using MegAlign option of DNASTAR package.

201

Figure 7.7 Quantification of recombinant plasmids eluted from FTA cards loaded with known amount of recombinant plasmid DNA + 8 µl of sap from healthy cassava leaf extracted in distilled water (a). The eluted DNA was used for PCR amplification of 555 bp viral DNA fragments insert using primers EAB555F/EAB555R (b). Positive control lanes (+C) contained the 555 bp viral DNA-B fragment PCR-amplified from cassava plants infected with EACMV from the growth chamber. To detect the limit of PCR amplification signal, 0.6 µg/µl was serially diluted to 0.6×10^{-6} µg/µl and each dilution used for PCR amplification of the 555 bp viral DNA fragment described above.

203