

https://dspace.org/

Interaction and impact of cassava mosaic begomoviruses and their associated satellites

Mollel, Happyness Gabriel

2014-07-07
http://hdl.handle.net10539/14886

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

INTERACTION AND IMPACT OF CASSAVA MOSAIC BEGOMOVIRUSES AND THEIR ASSOCIATED SATELLITES

By:

HAPPYNESS GABRIEL MOLLEL

Dissertation submitted in partial fulfillment for the degree of Master of Science

The University of Witwatersrand, Johannesburg

Republic of South Africa

Declaration

This thesis is my work and has not been presented for a degree in any other University



Signed…….Date: 13th January 2014

H. G. Mollel

.

Dedication

This work is dedicated to my husband Juma. W. Yabeja, my mother Nay Zakayo and my father Gabriel Jacob for their support and encouragement in the period of my studies. I would like also to dedicate to my lovely son William. J. Yabeja for his patience during this period of studies.

Acknowledgment

I wish to sincerely thank "International Centre for Genetic Engineering and Biotechnology" (ICGEB) for funding this work. I am highly grateful to my supervisors Prof. M.E.C. Rey of the University of Witwatersrand, South Africa, Dr. Joseph Ndunguru, Dr. Fred Tairo, Dr. Peter Sseruwagi of Mikocheni Agricultural Research Institute, Tanzania and Dr. Herve Vanderschuren of ETH, Zurich Switzerland for their timely advice and support in the course of my MSc. work. I am also indebted to Dr. Deusdedith Mbanzibwa and my colleagues Cyprian Rajabu, Doreen Mgonja and Margareth Lupembe of Mikocheni Agricultural Research Institute for their support in laboratory work. Also Ms. Debbie Carmichael a student at the school of Molecular and Cell Biology (MCB), university of the Witwatersrand, South Africa for her commendable assistance in primer designing. Furthermore, I would like to thank Dr. Trino Ascencio- Ibanez, of North Carolina State University, United State for his assistance in optimizing the gene gun. Finally, but not least I thank Ms. Maabo Moralo for her support in the laboratory work in the University of Witwatersrand, Johannesburg, South Africa.

Research outputs

Publications

H.G. Mollel., O.A. Ndomba., F. Mwatui., E. Ateka., P. Sseruwagi., F.Tairo., H. Vanderschuren., M.E.C. Rey and J. Ndunguru., 2013. Genetic diversity and geographic distribution of DNA-II and III molecules associated with cassava and Cassava mosaic disease in Uganda, Kenya, Rwanda and Tanzania. Manuscript submitted to Plant Pathology.

Mollel, G.H., Sseruwagi, P., Tairo, F., Vanderschuren, H., Rey, M.E.C., and Ndunguru, J., 2013. Molecular characterization of integrated DNA molecules associated with cassava mosaic begomovirus. Proceedings of the 12th International Plant Virus Epidemiology Symposium, Arusha, Tanzania, 27th-31st January, 2013.

Table of Contents

Declaration i
Dedication ii
Acknowledgment iii
Research outputs iv
List of tables vii
List of acronyms xi
Abstract 1
Background and Justification for study 5
Overall objectives 7
The specific aims of the study are: 8
CHAPTER ONE 9
LITERATURE REVIEW 9
1.1 Cassava (Manihot esculenta Crantz) 9
1.1.1 Introduction 9
1.1.2 Description of cassava and its importance 11
1.1.3 Production constrains 13
1.1.4 Geminiviruses 14
1.1.5 Cassava mosaic disease 16
1.1.6 Cassava mosaic begomovirus transmission 19
1.2 Satellites 21
1.2.1 Discovery of sub viral agents associated with begomovirus diseases 21
1.2.2 Alphasatellite 22
1.2.3 Betasatellites 23
1.2.4 Replication and compatibility of begomovirus subviral agents 24
1.2.5 Functions of begomovirus subviral agents 26
1.2.6 Common features between begomovirus components and satellites 26
1.2.7 Cassava begomovirus-associated satellites 28
CHAPTER TWO 31
A STUDY OF THE IMPACT OF DNA-II AND DNA-III ON CASSAVA MOSAIC DISEASE, AND INVESTIGATION INTO TRANSCRIPTION OF PUTATIVE EPISOMAL ORFs 31
2.1 Abstract 31
2.2 Introduction 32
2.3 Materials and Methods 34
2.3.1 Micropropagation and acclimatization of cassava cultivars TME3 and cv. 60444 34
2.3.2 Biolistic inoculation of virus and DNA-II and III clones 36
2.3.2.1 Preparation of gold particles 36
2.3.2.2 Coating of virus and DNA-II and III infectious clones 36
2.3.2.3 Transcript determination in the geminivirus – DNA-II and DNA-III - cassava host systems 38
2.3.2.3.1 RNA extraction 38

2.3.3.2 Primer design and RT- PCR analysis of ORF C4 for DNA-II and ORF V1 for DNA-III 40
2.3.3.4. Sequencing 44
2.4 Results 44
2.4.1 Symptoms observation 44
2.4.2 Symptoms severity and disease severity progress curves 47
2.4.3. RT-PCR for ORF C4 of DNA-II, ORF V1 and ORF C2 of DNA-III 50
2. 4. 4 Sequence analysis 53
2.4.4.1 Transcript sequences for DNA-II 53
2.4.4.2 Transcript sequences for ORF V1 of DNA-III 55
2.5 Discussion 57
CHAPTER THREE 63
GENETIC DIVERSITY AND GEOGRAPHIC DISTRIBUTION OF DNA-II AND III MOLECULES ASSOCIATED WITH CASSAVA AND CASSAVA MOSAIC DISEASE IN UGANDA, KENYA, RWANDA AND TANZANIA 63
3.1 Abstract 63
3.2 Introduction 65
3.3 Materials and Methods 67
3.3.1 Sample collection 67
3.3.2 DNA extraction 69
3.3.3 Polymerase chain reaction (PCR) 70
3.3.4 Molecular cloning 71
3.3.5 Sequences analysis 71
3.3.6 Statistical data analysis and presentation 72
3.4 Results 72
3.4.1 Cassava mosaic disease symptom expression in the field 72
3.4.2 PCR analysis of integrated DNA-II and DNA-III molecules 73
3.4.3 Occurrence and geographical distribution of integrated DNA molecules 75
3.4.4 Phylogenetic analysis and sequence comparisons 77
CHAPTER FOUR 83
SUMMARY AND RECOMMENDATIONS 83
Recommendations 85
Referrences 86

List of tables

Table 1: Word cassava production in 2010……....10

Table 2: Production of staple tuber crops 2003 to 2008 (million tons)…….....11

Table 3: Number of treatments used in studying putative transcription of DNA-II and DNA-III.....37

Table 4: Reaction of cassava plants to biolistic inoculation with EACMV-UG2 and ACMV, and in combination with DNA-II, DNA-III and DNA-II + DNA-III infectious clones.....39

Table 5: List of primer pairs used for transcript detection of ORF C4 of episomal DNA-II in antisense strand and ORF VI of DNA-III of sense strand.....43

Table 6: Samples collected from screen house for transcript analysis of DNA-II and DNA-III.....52

Table7: Description of amplified RT-PCR samples for DNA-II sent for sequencing 54

Table 8: Description of amplified RT-PCR samples for DNA-III sent for sequencing.....56

Table 9: Cassava major growing areas where survey was conducted Uganda, Kenya, Rwanda and Tanzania in 2009/2010.....68

Table 10: Analysis of Variance for symptom severity on cassava samples collected from Uganda, Kenya, Rwanda and Tanzania in 2009-2010.....73

List of figures

Figure 1: Cassava mosaic begomovirus genome structure composed of twinned particles; DNA-A (left) and DNA-B (right). Functional ORFs on the DNA-A virion-sense (AV1 and AV2) and the complementary-sense strand (C1 to C4). DNA-B has two ORFs: BV1 and BC1.....16

Figure 2: Cassava mosaic disease symptoms in cassava field.....19

Figure 3: Bipartite DNA-A (left) and DNA-B (right) begomovirus genome and satellite DNA-β which is often associated with DNA-A component of monopartite genome..... 28

Figure 4: Genomic structure of cassava mosaic begomovirus associated satellite-like DNA molecules. (A) satDNA-II with eight putative ORFs (V1-V3) on sense strand, and (C1-C5) on antisense strand, and (B) satDNA-III with four putative OFRs (V1-V2) on sense strand and (C1-C2) on antisense strand 30

Figure 5: Stages for acclimatization process: A) TME3 tissue culture cassava plantlets B) cv. 60444 tissue culture cassava plantlets C) Four- week old cassava plants in small pots D) 8 week-old cassava plants in big pots ready for inoculation.....35

Figure 6: Genomic structure of cassava mosaic begomovirus associated satellite-like DNA molecules showing the positions for ORFs.....41

Figure 7: Genome maps of circular DNA-II and III molecules showing primer positions.....42

Figure 8: Cassava mosaic disease symptom development on cv. 60444 inoculated with A) EACMV-UG2 alone, B) EACMV-UG2 + DNA-II, C) EACMV-UG2 + DNA-III, D) EACMV-UG2 + DNA-II & III. CMD symptom development on TME 3 inoculated with E) EACMV-UG2 alone, F) EACMV-UG2 + DNA-II, G) EACMV-UG2 + DNA-III, H) EACMV-UG2 + DNA-II & III.....46

Figure 9: Cassava mosaic disease symptom development on cv. 60444 inoculated with A) ACMV alone, B) ACMV + DNA-II, C) ACMV+ DNA-III, D) ACMV + DNA-II & III. CMD symptom development on TME 3 inoculated with E) ACMV alone, F) ACMV+ DNA-II, G) ACMV + DNA-III, H) ACMV + DNA-II & III.....47

Figure 10: Disease severity progress curves for cassava plants inoculated with EACMV and in combination with DNA-II, DNA-III and DNA-II + DNA-III in A) cv.60444 and B) TME3.....49

Figure 11: Disease severity progress curves for cassava plants inoculated with ACMV and in combination with DNA-II, DNA-III and DNA-II + DNA-III in A) cv. 60444 and B) TME3.....50

Figure 12: Gel picture showing amplification of ORF C4 and ORF V1 transcripts on 1% agarose by RT-PCR from cassava plants inoculated with either EACMV-UG2, ACMV and in combination with DNA-II, DNA-III and DNA-II + DNA-III clones A) DNA-II; and B) DNA-III.....51

Figure 13: Cassava plants showing cassava mosaic symptoms in the fields A: CMD symptoms; B: mosaic severe; C: leaf curling and distortion D: leaf narrows in cassava field.....69

Figure 14: PCR amplification of integrated DNA-II molecules (16A) and DNA-III (16B) on 1% agarose gel. M=Molecular weight marker of 1 kb plus; +C= positive control and -C = negative control. Lanes 1-20 (16A & 16B) represent samples collected from farmers.....74

Figure 15: Occurrence of integrated DNA-II and DNA-III molecules in Uganda, Kenya, Rwanda and Tanzania in 2009/2010.....75

Figure 16: Geographical distribution of DNA-II and DNA-III molecules in Uganda, Kenya, Rwanda and Tanzania in 2009/2010.....76

Figure 17: Phylogenetic trees obtained from alignment of the nucleotide sequences of DNA-II and III with available sequence of satDNA-II and satDNA-III. 17A: DNA-II sequences; 17B: DNA-III sequences obtained from Kenya, Uganda, Rwanda and Tanzania in 2009/2010.....79

List of acronyms

ACMV	*Africa cassava mosaic virus*
AYVD	Ageratum yellow vein disease
ANOVA	Analysis of variance
bp	Base pairs
°C	Degrees Celcius
CMD	Cassava mosaic disease
CP	Coat protein
CMBs	Cassava begomoviruses
CBSD	Cassava brown streak disease
cDNA	Complementary DNA
DNA	Deoxy-ribonucleic acid
DNA-1	Alphasatellites
DNA-β	Betasatellites
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
EACMV-UG	*East Africa cassava mosaic Ugandan variant*
EACMCV	*East Africa cassava mosaic Cameroon virus*
EACMV	*East African cassava mosaic virus*
EACMKV	*East African cassava Kenya mosaic virus*
EACMMV	*East African cassava Malawi virus*
EACMZV	*East African cassava mosaic Zanzibar virus*
EACMZV	*East Africa cassava mosaic Zanzibar virus*
EDTA	Ethylene-diaminetetraacetate

EST	Expressed sequence tag
FAO	Food and Agricultural Organisation
HCl	Hydrochloric acid
ICMV	*Indian cassava mosaic virus*
MYMV	*Mungbean yellow mosaic virus*
MARI	Mikocheni Agricultural Research Institute
$MgCl_2$	Magnesium chloride
MP	Movement protein
MS	Murashige and Skoog media
MED-NA-ME	North African-Mediterranean-Middle East
µl	Microlitre
µg	Microgram
NW	New world
NCBI	National Centre for Biotechnology Information
NSP	Nuclear shuttle protein
OW	Old world
ORF	Open reading frame
PCR	Polymerase chain reaction
RT - PCR	Reverse Transcriptase polymerase chain reaction
RNA	Ribose nucleic acid
Rep	Replication activator protein
Ren	Replication enhancer protein
ssDNA	Single stranded deoxyribonucleic acid
SDS	Sodium dodecyl sulphate

SCR	Satellites conserved region
Trap	Transcription protein
TRV	*Tobacco ringspot virus*
ToLCV	*Tomato leaf curl virus*
TYLCD	Tomato yellow leaf curl disease
TYLCCV	*Tomato yellow leaf curl China virus*
TGMV	*Tomato golden mosaic virus*
TAE	Tris-acetate and EDTA
Tris-HCl	Tris (hydroxymethyl) aminomethane hydrochloride

List of Appendices

Appendix 2.1: Alignment of DNA-II ORF C4 transcripts isolates collected from mock-inoculated cassava plants and cassava plants inoculated with either EACMV-Ug2, ACMV and in combination with DNA-II, DNA-III and DNA-II + DNA-III clones. Highlighted colours are conserved nucleotides in all sequenced samples. Yellow colour is G, Blue colour is C, Green colour is T and Red colour A.

Appendix 2.2: Percentage nucleotide sequence similarity matrix table of the DNA-II ORF C4 transcript sequences.

Appendix 2.3: Alignment of DNA-III ORF V1 and ORFC2 transcripts isolates collected from mock-inoculated cassava plants and cassava plants inoculated with either EACMV-Ug2, ACMV and in combination with DNA-II, DNA-II and DNA-II + DNA-III clones. Highlighted colours are conserved nucleotides in all the sequenced samples. Yellow colour is G, Blue colour is C, Green colour is T however A nucleotide was not conserved.

Appendix 2.4: Percentage nucleotide sequence similarity matrix table of the DNA-III ORF VI transcripts sequence.

Abstract

Cassava (*Manihot esculenta* Crantz) is affected by two major viral diseases, namely Cassava brown streak disease (CBSD) and Cassava mosaic disease (CMD). Two of the most widely distributed begomoviruses in East Africa associated with CMD, are *East African cassava virus*-Uganda2 (EACMV-UG2) and *African cassava mosaic virus* (ACMV). Despite efforts of generating improved Tropical Manihot Series (TMS) by traditional breeding and using highly resistant geminivirus cassava landraces such as Tropical Manihot Esculenta1 (TME1) and Tropical Manihot Esculenta3 (TME3), more recently two circular single stranded (ss) satellite-like DNA molecules (episomal DNA-II and DNA-III) have been found to be associated with CMD and are able to break resistance to EACMV-UG2 and enhance virus symptoms.

The nature of these satellite-like DNA molecules is unknown, and furthermore, the discovery of integration of partial copies of DNA molecules (DNA-II and III fragments), and evidence for transcription from cassava Expressed Sequence Tag (EST) database screening, has led to an even more perplexing disease complex. In the present study, we attempted to further explore the interaction between the satellite-like DNAs and their associated cassava-infecting begomoviruses by investigating the impact of these DNA molecules on disease development in TME3 (tolerant) and cv. 60444 (susceptible) cassava cultivars, and to also gather biological evidence for transcription of integrated genomic and episomal (putative predicted ORFs) sequences in the ACMV and EACMV-UG2-associated DNA-II and DNA-III.

Biolistic inoculation of EACMV-UG2, ACMV, and in co-bombardment with DNA-II, DNA-III, DNA-II + DNA-III was successfully performed. CMD symptoms were developed earlier on

cassava plants inoculated with ACMV + DNA-II, ACMV + DNA-III, ACMV + DNA-II + DNA-III and EACMV-UG2 + DNA-II, EACMV-UG2 + DNA-III, EACMV-UG2 + DNA-II + DNA-III molecules compared with cassava plants inoculated with begomoviruses alone. Additionally, CMD symptoms were more severe in cv.60444 compared to TME3 when inoculated with begomoviruses alone, or in combination with DNA-II, DNA-III and DNA-II + DNA-III molecules. DNA-II and III were able to break resistance to the highly CMD-tolerant cassava landrace, TME3, and enhance virus symptoms.

In order to confirm EST-generated evidence for transcription of DNA-II and III fragments, cDNA was subjected to RT-PCR. RT-PCR of transcripts was successful for only three putative ORFs: ORF C4 of the antisense DNA-II strand, ORF V1 on sense DNA-II strand, and ORF C2 on antisense strand for DNA-III. Primers for transcripts amplified 250 bp and 220 bp for ORF C4 of DNA-II and ORF V1 of DNA-III, respectively. Transcribed ORFs were confirmed by sequencing, and the sequences were similar to the published sequences of Begomovirus associated DNA-II satellite and Begomovirus associated DNA-III satellite, respectively. These results showed that at least two putative ORFs for DNA-II and one (the largest ORF VI) DNA-III can be transcribed.

Furthermore, surveys were undertaken in order to ascertain the distribution of episomal and integrated DNA-II and III in cassava germplasm from several countries, namely Tanzania, Uganda, Kenya and Rwanda. Results from this research successfully established genetic diversity and wide geographical distribution of integrated DNA-II and DNA-III molecules. Two primer pairs were designed from a central conserved sequence found in all the integrated DNA-II

or III fragments identified from the cDNA libraries (EST database). These primers also amplified integrated sequences of expected size in cassava accessions and wild Manihot species which were similar to satellite-like sequence occurrences in the ESTs.

Using designed primers, PCR amplification yielded integrated DNA-II and DNA-III products of ~895 bp and ~306 bp, respectively. Analysis of 363 field leaf samples detected the presence of DNA-II or DNA-III from Kenya (3.3% or 8.3%), Uganda (18% or 2.5%), Rwanda (6.5% or 19.6%) and Tanzania (5.7% or 11.9%) , results which were confirmed by analysis of the sequenced PCR amplicons. Detection of both DNA-II and DNA-III molecules on the samples collected was also found from Kenya (73%), Uganda (69.1%), Rwanda (50%) and Tanzania (69.3%). Interestingly integrated DNA-II and II copies were amplified from healthy, symptomless and infected cassava samples. DNA-II sequences did not vary significantly (93.3% - 99.8%) and were highly similar to the sequences of Begomovirus associated sat DNA-II (AY836366) and 99% with mentha leaf deformity disease associated satellite DNA-II, while DNA-III sequences and Begomovirus associated DNA-II satellite (AY833667).

In conclusion, this study has provided useful information that contributes to a further understanding of the biological function of integrated and episomal DNA-II and III molecules in begomoviruses infected cassava plant. However the relationship, if any between episomal and integrated forms needs to be established in future, and investigation into whether the transcribed ORFs can produce functional proteins, needs to be undertaken. How DNA-II and III interact with EACMV-UG2 and ACMV in disease modulation remains to be explored, and the replication of episomal DNA-II and III by these associated begomoviruses needs to be confirmed

if these DNA molecules are to truly show a satellite-like relationship. Furthermore, the findings in this study that partial and varied-sized integrated DNA-II and III fragments occur widely in healthy and infected cassava germplasm will enable researchers (plant virologists and breeders) working on cassava in Sub Saharan Africa (SSA) to explore this complex more deeply in order to develop durable management strategies for CMD.

Background and Justification for study

Cassava (*Manihot esculenta* Crantz) is one of the leading crops in Sub-Saharan Africa in terms of production and has become an important source of income to households and small-scale farmers (FAOSTAT, 2010). Cassava is considered as a staple root crop for more than 800 million people living in developing countries (Burns *et al*., 2010). About 70% of world cassava root production is used for human consumption either directly after cooking or in processed forms; the remaining 30% is used for animal feed and industrial products, such as starch, glucose and alcohol (El-Sharkawy, 2004). However, the production across the region is greatly affected by cassava mosaic disease (CMD) and cassava brown streak disease (CBSD). Several studies (Gibson., 1996; Ogbe *et al*., 1996; Legg *et al*., 1999; Fondong *et al*., 2000; Berry and Rey, 2001; Ndunguru *et al.,* 2005; Bisimwa *et al*., 2012) have reported the occurrence of CMD in different countries in Sub-Saharan Africa. A comprehensive cassava mosaic begomovirus (CMB) characterization study by Ndunguru *et al.* (2005) in East Africa showed that seven cassava mosaic geminivirus species occur in Tanzania. Mbanzibwa *et al*. (2009a) reported the prevalence of two Ipomovirus species causing CBSD in the Lake Victoria basin and along the coastal belt of Indian Ocean.

Plant viruses are occasionally associated with satellite molecules which contribute to increase virus symptoms on infected plants. Satellites are either viruses or nucleic acids that depend on a helper virus for their replication but lack general nucleotide sequence homology to the helper virus (Mayo *et al*., 2005). Satellites associated with RNA plant viruses were the first satellites to be discovered and described, and are divided into satellite viruses and satellite RNAs (Schneider, 1969). The majority of satellite RNAs do not encode functional proteins but can have a

remarkable effect on the symptoms induced by their helper viruses, ranging from mild symptoms to severe symptom (Roossink *et al.,* 1992). Two satellites associated with monopartite begomovirus diseases; alphasatellites and betasatellites were isolated from different infected crops of Old World (OW) begomovirus diseases such as in Tomato yellow leaf curl disease in China, Cotton leaf curl disease in Pakistan and India, and Ageratum yellow vein in South East Asia (Dry *et al* 1997; Briddon *et al*, 2004). These satellites depend on helper virus for the protein encapsidation and movement within and between plants (Mayo *et al*., 2005).

More recently, two satellite-like molecules, DNA-II and DNA-III associated with CMBs (*African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV), were described from infected cassava plants in Tanzania and were able to enhance CMD symptoms and break resistance to geminiviruses in the highly tolerant landrace TME 3 (Ndunguru *et al.,* 2005; 2008; Ingelbrecht *et al.,* 2008). Primers to amplify DNA-II (895bp) and DNA-III (306 bp) were previously designed based on the sequences from conserved regions across different matching ESTs. Surprisingly, these primers amplified integrated sequences in cassava accessions and wild Manihot species which were similar to satellite-like sequence occurrences in the ESTs when additional sequences from the cassava genome were examined.

The diversity and geographic distribution of DNA-II and III in Uganda, Kenya, Rwanda and Tanzania is not well documented. Further, it is not known if the open reading frames (ORFs) of the DNA-II and III are transcribed; this information would be helpful to establish if the transcribed products potentially produce functional proteins involved during viral infection in the plant host.

This study aimed to generate information on the genetic diversity and geographic distribution of the integrated DNA-II and III molecules in Kenya, Uganda, Tanzania and Rwanda. Furthermore, an investigation of the putative transcripts of genomic and episomal DNA-II and III molecules was also conducted. Findings from this study will add to the body of knowledge on these unusual DNA molecules and will help researchers (plant virologists and breeders) working on cassava in Sub Saharan Africa (SSA) to develop durable management strategies for CMD in the future.

Overall objectives

The aim of this study is to further uncover the nature of these unusual episomal satellite-like DNA molecules associated with putative begomoviruses in East African countries. Furthermore, this study aimed to;

i. Confirm that these putative episomal DNA-II and III molecules influence the symptom phenotype and severity of CMD in order to further understand the nature of resistance, and investigate biological evidence for transcription of putative ORFs on DNA-II and DNA-III.
ii. Screening of cassava germplasm to establish the geographic distribution and diversity of integrated forms in cassava germplasm.

The specific aims of the study are:

i. To screen cassava genome germplasm in major cassava-producing areas in Uganda, Kenya, Rwanda and Tanzania in order to establish genetic diversity and geographic distribution of integrated DNA-II and DNA-III sequences.

ii. To investigate mechanisms of interaction between the satellite-like DNAs and their associated cassava – infecting begomoviruses by investigating the impact of these DNA molecules on disease development in TME3 (tolerant) and cv. 60444 (susceptible) cassava cultivars.

iii. To determine if any putative regions or ORFs on episomal or integrated DNA-II or III are transcribed.

CHAPTER ONE

LITERATURE REVIEW

1.1Cassava (*Manihot esculenta* Crantz)

1.1.1 Introduction

Cassava (*Manihot esculenta* Crantz) is extensively cultivated as an annual crop in tropical and subtropical regions of Africa, Asia and Latin America between 30°N and 30°S (El-Sharkawy, 1993). Cassava is considered as a staple root crop for more than 800 million people living in developing tropical countries (Burns *et al.,* 2010). The total worldwide cassava production in 2009 was about 241 million tonnes (Bull *et al.,* 2011) with Africa being the leading producer (FAOSTAT, 2010). About 70% of world cassava root production is used for human consumption either directly after cooking or in processed forms; the remaining 30% is used for animal feed and other industrial products, such as starch, glucose and alcohol (El-Sharkawy, 2004). Cassava is being increasingly used for bio-ethanol production in Brazil and China (Narina and Odeny, 2011) while in Nigeria it is a major cash crop earning the country about US$ 3 billion annually through export of the crop and related products (FAO, 2004). Of the estimated 18 million ha of cassava cultivated worldwide, approximately two thirds are in Africa, producing 110 million tonnes (mt) of tuberous roots annually (FAO, 2006). Worldwide average yield of cassava production in 2010 was estimated to be 12.40 tonnes/ha (FAOstat, 2012) (**Table 1).**

Table 1: World cassava production in 2010

Location	Area harvested (ha)	Yield (tons/ha)	Production (tons)
World	18,568,788	12.4	230,265,639
Africa	11,969,784	10.16	121,661,234
S. America	2,400,720	13.2	31,686,404
Asia	3,901,877	19.26	75,148,313
Tanzania	798,000	5.5	4,392,170
Kenya	61,573	5.25	323,389
Uganda	415,000	12.73	5,282,000
Malawi	195,828	20.43	4,000,990
Mozambique	950,000	6	5,700,000
Rwanda	197,394	12.04	2,377,210
Zambia	198,000	5.82	1,151,700

Source: FAOstat, 2012

Although cassava requires optimal conditions to reach high growth rates, it performs well in drought areas and on poor soils and thus is considered one of the most productive tropical crops on marginal lands (Zhang *et al.,* 2010). Under prolonged water shortages in seasonally dry and semi-arid environments with less than 700 mm of annual rain, improved cultivars can give dry root yields of over 3 t ha-1 (El-Sharkawy, 2006). 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**).

Table 2: Production of staple tuber crops 2003 to 2008 (million tons)

Tuber Crop	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**Ave.**
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

1.1.2 Description of cassava and its importance

Cassava is a shrubby perennial plant that typically grows from one to three meters (3-10 feet) in height. It is grown mainly for its carbohydrate rich tuberous roots (Katz and Weaver, 2003). It belongs to the family *Euphorbiceae* that also includes other commercially important plants like castor bean (*Ricinus communis* L.) and rubber (*Havea bransiliensis* L.). Cassava is believed to have originated in South and Central America and was introduced into West Coast Africa in 16th century and later to East Africa through Madagascar and Zanzibar, and later into Asia in the late 17th century by Portuguese traders. Africa has become the largest producer of cassava, constituting 54% of world production. Most of the spread of cassava in Africa away from the

coastal areas and riverside trading posts took place during the 20th century due to colonial powers encouraging its cultivation as a reserve against famine and the ability of the crop to withstand locust attack (Hillocks, 2002). However cultivation of cassava started to decline in the 1960s due to post colonial governments turning their attention to maize in terms of funding and research efforts, as well as taste preference for maize (Haggblade and Zulu, 2003). Today, cassava is cultivated in more than 80 countries mainly between 30º South and 30º North of the equator (Fauquet and Fargette, 1990).

Cassava is suited to warm humid lowland tropics and can be grown in areas where the mean annual temperature exceeds 20ºC with annual rainfall that varies between 5000 mm and 8000 mm (Pounti-Kearlas, 1998). Cassava is a very hardy plant, tolerating drought better than most other crops, and growing well in very poor and acidic soils (Katz and Weaver, 2003). Cassava crop can yield up to 13 millions kcal/acre (Bender and Bender, 2005).

Cassava naturally is grown by small-scale farmers using traditional methods, and often on land not suitable for other crops (Katz and Weaver, 2003). Cassava is propagated by cutting a mature stem into sections of approximately 15 cm and planting prior to the wet season. Most of the harvest from cassava is used for human consumption, either fresh or in processed forms. Cassava can be processed into a wide variety of products for food and industrial uses, such as starch, flour, alcohol, glucose and others. The leaves, which are rich in proteins, vitamin C and other nutrients, are consumed in some communities to supplement the low protein content of the roots. Cassava is consumed by an estimated 600 million people (FAO, 2006). Cassava flour is used in the production of biscuits, sausage rolls, meat pies and bread (Ogbe, 2001). 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). In most east African countries, though maize is the dominant staple food, cassava is very important in Mozambique, Tanzania, Uganda, and Burundi as a reserve against famine.

1.1.3 Production constraints

Cassava production in East and Southern Africa is particularly exposed to numerous biotic stresses. Common constrains include pests and diseases, poor agronomic practices, low yielding varieties, high cyanide levels, lack of clean planting materials and long maturity periods (Thresh *et al*. 1994). Pests and diseases are the most economically important constrains (Hillocks, 1994) to the cassava production. Pests infesting cassava include cassava green mite (*Mononychellus tanajoa*), green spider mite (*Mononychellus tanajoa*), mealy bugs (*Phenacoccus manihot*), cassava hornworm (*Erinnyis ello*), scales, thrips and whitely (*Bemisia tabaci) (*Montero, 2003*).* Diseases of cassava include cassava mosaic disease (CMD), cassava brown streak disease (CBSD), cassava bacterial blight (CBB), cassava anthracnose disease (CAD), cassava bud necrosis, and root rots (Calvert and Thresh, 2002). Virus diseases, particularly CMD and CBSD, are the most important economical constraints of cassava production in east and southern Africa (Taylor and Fauquet 1997; Thresh *et al*. 1994; Thresh *et al*. 1997). The two diseases are thought to have a risen from infection of cassava by viruses already present in the indigenous African flora (Legg and Hillocks, 2003).

The perpetuation of these diseases in cassava production is influenced by the abundance of efficient insect vectors for transmission and continuous use of unclean planting materials normally selected from the previous seasons. Despite success with control of cassava mealy bug and cassava green mite using biological control programs, CMD and CBSD have remained a challenge to control using sustainable approaches though there are adequate information on the pathogen and efficient diagnostic tools such as PCR and RT-PCR tools being well established. Thus, the two diseases are now key research priorities of many root and tuber programs in many African countries (Legg and Thresh, 2003).

1.1.4 Geminiviruses

Geminiviruses are viruses which infect plants and are characterized by circular single-stranded DNA genomes encapsidated into small twinned icosahedra virions. The family *Geminiviridae* is divided into four genera (*Mastrevirus, Curtovirus, Topocuvirus and Begomovirus*), which are differentiated based on their genome structure; the host plants they infect and the type of insect vector (Fauquet *et al.,* 2008).

The genus *Begomovirus* consists of either two DNA molecules, DNA-A and DNA-B, each of about 2.8 kb which are responsible for different functions in the infection process (Stanley *et a*l., 2005) (**Fig. 1**), or a single DNA component (monopartite). Begomoviruses are the largest group of plant viruses and cause economically important diseases of many vegetable and fibre crops (Rojas *et al*., 2005; Varma and Malathi, 2003). DNA-A contains six partially overlapping open reading frames (ORFs) organized in two opposite transcriptional directions separated by an

intergenic region (IR). On the virion-sense strand, DNA-A component contains AV1 and AV2 ORFs, and AC1–AC4 are on the complementary-sense strand. The DNA-A encoded gene products are replication-associated protein AC1 (Rep); AV1 coat protein (CP); and proteins that participate in the control of replication (AC3) and gene expression (AC2) (TrAP). DNA-B encodes proteins required for nuclear trafficking (BV1) and cell-to-cell movement (BC1) of the viral DNA (Hamilton *et al*., 1984; Hanley-Bowdoin *et al*., 1999). Both DNA components (DNA-A and DNA-B) share a high nucleotide identity in the IR (*c*. 200 nt) called the common region (CR), which contains promoter and sequence elements required for DNA replication and transcription (Lazarowitz, 1992; Eagle *et al.*, 1994; Chatterji *et al.*, 1999). Begomoviruses are transmitted by the whitefly *Bemisia tabaci* (*Bemisia tabaci* Genn.), and are also spread through infected cuttings, which are the usual mode of cassava propagation. CMBs have been reported from all cassava-growing countries in Africa and CMD constitutes a threat to cassava production.

In addition to sub-viral components, geminiviruses have also been shown to be accompanied by smaller-sized DNA molecules, called defective-DNA molecules, that are derived from the helper virus genomes, but are not satellite molecules (Patil and Dasgupta, 2006; Simon *et al.*, 2004). Monopartite begomoviruses are accompanied by circular single stranded DNAs (ssDNA) which are alphasatellites and betasatellites. Most Old World monopartite begomoviruses are associated with one or more betasatellite DNA(s), which are required for induction of typical disease symptoms and depend on the helper begomovirus for replication and movement (reviewed by Mansoor *et al*., 2006; Briddon and Stanley, 2006).

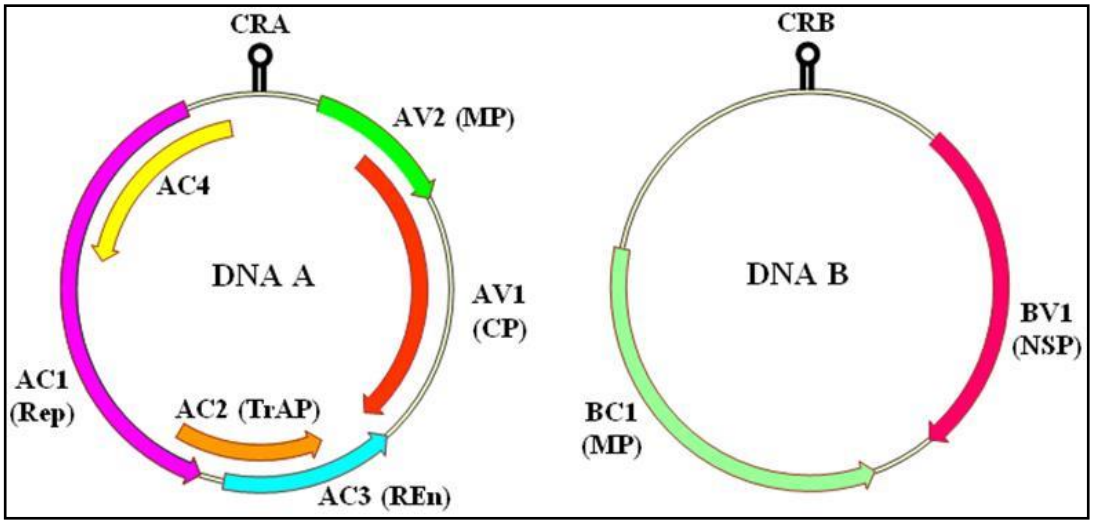


Figure 1: Cassava mosaic begomovirus genome structure composed of twinned particles; DNA-A (left) and DNA-B (right). Functional ORFs on the DNA-A virion-sense (AV1 and AV2) and the complementary-sense strand (C1 to C4). DNA-B has two ORFs: BV1 and BC1. In both DNA-A and DNA-B component there is a non-coding intergenic region referred to as the common region (CR) (Adapted from Alabi, *et al.*, 2011).

1.1.5 Cassava mosaic disease

Cassava mosaic disease (CMD) occurs in all cassava- producing regions in Africa and Indian subcontinents (Legg and Fauquet, 2004; Legg *et al.*, 2006). Cassava mosaic disease is the major constraint to cassava production in Africa (Thresh and Cooter, 2005). CMD symptoms are variable according to season and variety, but always include a chlorotic mosaic on infected

leaves and the colour of chlorosis varies from pale green to whitish yellow. In moderate to severe infections, leaves also show crumpling, the laminae are distorted, and the size is reduced (**Fig. 2**) leading to the stunting of the growth of the plant. There are no symptoms on the stem or roots (Legg and Thresh, 2000).

Cassava mosaic disease in Africa is caused by nine species of begomoviruses (Fauquet *et al.,* 2008). Cassava mosaic begomoviruses (CMBs) were originally placed into three groups based on their reaction to monoclonal antibodies (Swanson and Harrison, 1994). Group A was limited to West Africa, Burundi, Chad, Uganda and Western part of Kenya. Group B occurred in Malawi, Madagascar, Zimbabwe and Eastern part of Kenya and Tanzania; and Group C was restricted to India and Sri Lanka. However due to the movement of infected stakes, these geographic boundaries are no longer defined. Due to the differences in their nucleotide sequences, these viruses were later identified as different virus species (Hong *et al.,* 1993) and named *African cassava mosaic virus* (ACMV), *East Africa cassava mosaic virus* (EACMV) and *Indian cassava mosaic virus* (ICMV) respectively.

More recently, nine species of CMBs have been classified (Fauquet *et al.,* 2008): These viruses are *African cassava mosaic virus (*ACMV*); East African cassava mosaic virus* (EACMV), *East African cassava mosaic Cameroon virus* (EACMCV) (Fondong *et al.,* 2000), *East African cassava mosaic Kenya virus* (EACMKV) (Bull *et al.,* 2006), *East African cassava mosaic Malawi virus* (EACMMV) (Zhou *et al.,* 1998), *East African cassava mosaic Zanzibar virus* (EACMZV) (Maruthi *et al.,* 2004)), *African cassava mosaic Bukina Faso virus* **(**ACMBFV)

(Tiendrébéogo *et al*., 2012), *South African cassava mosaic* (SACMV) (Berrie *et al.*,1998) and *Cassava mosaic Madagascar virus* **(**CMMGV) (Harimalala *et al.,* 2012).

Furthermore, cassava is often infected with more than one begomoviruses that can enhance symptom severity due to synergistic effect and because of recombination (Pita *et al.,* 2001). For instance, ACMV-EACMV recombinant component A, previously designated EACMV-UG2, and a pseudo-recombinant component B, designated EACMV-UG3 (Pita *et al.,* 2001), caused the pandemic of severe CMD currently devastating cassava much of East and Central Africa (Legg *et al.,* 2004). In 1997, only ACMV and EACMV were known to occur in Tanzania with the former occurring only in the western part of the country (Ogbe *et al.,* 1997). The discovery of EACMZV on the island of Zanzibar (Maruthi *et al.,* 2002) together with the recent spread into Tanzania of the EACMV-UG2 associated pandemic of severe CMD (Legg *et al.,* 2004, Ogbe *et al.,* 1997) has exacerbated the CMD situation.

Cassava mosaic disease has been reported only in Africa and Indian subcontinent despite the large scale cultivation of cassava in Latin America and South East Asian countries (Fargette *et al.,* 2006). This is due to the inability of the polyphagous *Bemisia tabaci* B biotype to colonize cassava effectively (Carabali *et al.,* 2005), and recent studies have demonstrated a Southern African clade composed of cassava whitefly haplotypes that are not part of the B group (Berry *et al.,* 2004; Esterhuizen *et al.,* 2012).

In the 1920s, the search for virus-resistant genotypes started, and breeding for resistance has been the major control strategy for CMD (Otim-Nape *et al.,* 1998). Since that time, trials of

cultivars and selections have been done in several countries including Kenya, Tanzania, Uganda, Madagascar, Democratic Republic of Congo and Nigeria (Jennings, 1994). The use of resistant cassava landraces such as TME3 as has remained the most economical and ecologically sustainable control measure.



Figure 2: Cassava mosaic disease symptoms in cassava field.

1.1.6 Cassava mosaic begomovirus transmission

Cassava mosaic begomoviruses are transmitted by the whitefly vector, *Bemisia*. *tabacci* (*B. tabaci*). CMBs were first thought to be transmitted by the B biotype in Africa, but more studies have shown that the whitefly associated with cassava in sub-Saharan Africa is a different haplotype belonging to a unique clade and recently the Q-biotype has been reported in South

Africa (Berry *et al.,* 2004; Esterhuizen *et al.,* 2012; Mugerwa *et al*., 2012). The coat protein (CP) of begomoviruses is adapted for transmission by the local whitefly population, which explains the antigenic similarity of the CPs of the begomoviruses from the same area (Harrison and Robson, 1999) and the proposed theory of co- adaptation between the CMBs and their local whitefly population (Maruthi *et al*., 2002). The whitefly B biotype is more productive and has an extreme broad host range (Colvin *et al.,* 2004) which might have contributed to the transmission of the new viruses from the weed hosts to cultivated crop plants hence the emergence of a number of begomoviruses disease in the world.

Close relatives (haplotype variants) of both biotypes are endemic to adjacent locales, and in some habitats B-like and Q-like haplotypes overlap with respect to geographic distribution and host association. The prototype B and Q biotypes are highly productive, polyphagous, have a propensity to develop insecticide resistance, and cause damage to plants directly through feeding and honeydew excretion, and indirectly, by virus transmission (Dennehy *et al.,* 2005; Horowitz *et al.,* 2005; Prabhaker *et al.,* 2005; Chu *et al.,* 2006; Dennehy *et al.,* 2006; Brown, 2010; Dennehy *et al.,* 2010). The unrestricted movement of whitefly-infested and virus-infected plants has resulted in the widespread distribution of these two aggressive biotypes beyond their respective geographical range, and nearly worldwide (Chu *et al.,* 2006; De La Rua *et al*., 2006; Dennehy *et al*., 2006; Brown, 2007; Bethke *et al.,* 2009; Mckenzie *et al.,* 2009; Dennehy *et al*., 2010). The onset of CMD epidemic in Uganda has led to an increase of whitefly populations comprising two distinct clusters of *B. tabaci* (Ug1 and Ug2) (Legg *et al*., 2002, Sseruwagi *et al*., 2005).

1.2 Satellites

Satellites are defined as viruses or nucleic acids that depend on a helper virus for their replication but lack general nucleotide sequence homology to the helper virus (Mayo *et al*., 2005). Satellites associated with RNA plant viruses were the first satellites to be discovered and described, and are divided into satellite viruses and satellite RNAs (Schneider, 1969). Satellite viruses encode a structural protein that encapsidates its own nucleic acid while satellite RNAs rely on the helper virus structural protein for encapsidation and do not necessarily encode additional non-structural proteins. Satellite RNAs also depends on the helper virus for their replication, but provides a function that is necessary for the biological success of the helper virus. The first satellite RNA was identified in 1969 in association with the nepovirus *Tobacco ringspot virus* (Schneider, 1969), and since that time, a large number of satellite RNAs, associated with several groups of plant viruses, have been reported (Mayo *et al*., 2005). The majority of these satellite RNAs do not encode functional proteins but can have a dramatic effect on the symptoms induced by their helper viruses, ranging from less severe symptoms to an increase in symptom severity (Roossinck *et al*., 1992).

1.2.1 Discovery of sub viral agents associated with begomovirus diseases

The first begomovirus satellite to be discovered, was ToLCV-sat isolated from tomato plants infected with the monopartite begomovirus *Tomato leaf curl virus* (ToLCV) in Australia (Dry *et al.*, 1997). The circular satellite is small (682 nucleotides), has no extensive open reading frames and has little sequence similarity to its helper virus with the exception of sequences within the

apex of two stem-loop structures, one containing the ubiquitous geminivirus TAATATTAC motif and the other containing a putative ToLCV Rep binding motif (Behjatnia *et al.*, 1998). ToLCV-sat is not required for ToLCV infectivity and has no effect on the symptoms induced by the helper virus but is dependent on the helper begomovirus for its replication and encapsidation and hence has a typical feature of satellite DNA.

1.2.2 Alphasatellite

Alphasatellites (DNA-1) are self- replicating satellite-like molecules, depend on the helper virus for movement, encapsidation and vector transmission. The alphasatellites do not have known specific functions that can be attributed to the begomovirus though they are associated with monopartite Old World begomovirus diseases such as Cotton leaf curl disease in Pakistan and India, Tomato yellow leaf curl disease in China and Ageratum yellow vein disease from South East Asia (Briddon *et al.,* 2004). DNA-1 components cannot strictly be defined as satellite DNAs according to current guidelines because they do not rely on helper virus for replication (Mayo *et al.,* 2005).

Alphasatellites have a highly conserved genome organization, encompassing a replication-associated protein 36kDa, adenine- rich region of 200 nts and origin of replication including a conserved nanonucteotide TAGTATT/AC, similar to nanoviruses (Briddon *et al.,* 2004). According to the study conducted on the evolution of geminiviruses and satellites, it was reported that there were no alphasatellites associated with New World begomovirus (Briddon and Stanley 2006, Nawaz-ul-rehman and Fauquet 2009). However, recently, in a study by

Paprotka et al. (2010), the first two alphasatellites were discovered in the New World - one associated with a Brazilian strain of *Euphorbia mosaic virus* (EuMV) and another with a newly discovered *Begomovirus* species, *Cleome leaf crumple virus* **(**ClLCrV). Sequence analysis indicates that these alphasatellites are different from all other known OW satellites.

1.2.3 Betasatellites

A novel ssDNA component of approximately half the size of the helper begomovirus was isolated and shown to induce the yellow vein phenotype when re-introduced with AYVV into *Ageratum* (Saunders *et al.*, 2000). The component was named DNA-β because, in many respects, it functionally resembled the DNA-B component of bipartite begomoviruses. Betasatellites (DNA-β) are completely depending on the helper virus for replication, encapsidation and movement within and between plants (Mayo *et al.*, 2005). Betasatellites are associated with their specific helper component, irrespective of host and geographical distribution. Betasatellites encode a single gene, betaC1 (13kDa protein) in the complementary strand, and contain an adenine-rich region 240 nt, and a 220 nt satellite-conserved region (SCR) which is highly conserved among all betasatellites known (Briddon *et al.*, 2004). While DNA-1 component was derived from nanoviruses, the evolutionary origin of DNA-β remains unclear. DNA-β components contain an A-rich region suggesting that, like DNA-1, they may have originated as a bona fide component of another pathogenic agent prior to being captured by the begomovirus, necessitating a slight increase in size.

1.2.4 Replication and compatibility of begomovirus subviral agents

Comparison of the growing number of DNA-β components has indicated that they have a highly conserved structure. In addition to an A-rich region, DNA-β components encode a single gene (βC1) and contain a highly conserved sequence of approximately 80 nucleotides, referred to as the satellite-conserved region (SCR). PCR primers based on this sequence have provided a simple and robust method for the detection and isolation of DNA-β components (Briddon *et al.*, 2002). The SCR is located adjacent to the putative stem-loop structure, which contains the TAA/GTATTAC motif, which, by analogy to geminiviruses, is the site where Rep introduces a nick during the initiation of virion-sense DNA replication. The fact that naturally occurring mutants lacking the βC1 coding region (Briddon *et al.*, 2003) and mutants in which the βC1 coding region and A-rich region have been deleted *in vitro* (Tao *et al.*, 2004, Qian and Zhou, 2005) are maintained by the helper begomovirus is consistent with the involvement of the SCR and stem-loop in replication, although the identification of precise DNA-β sequences that contribute to replication awaits fine mapping studies. Two-dimensional gel electrophoresis analysis of ToLCV and CLCuD DNA-β replication intermediates suggests that the satellite replicates use similar rolling circle and recombination-dependent replication mechanisms (Alberter *et al.*, 2005). The observation that ToLCV-sat proliferated in the presence of ACMV and BCTV (Dry *et al.*, 1997) was unexpected because neither virus has the precise ToLCV and ToLCV-sat high-affinity Rep binding motifs that should be necessary for trans-replication according to the generally accepted model for the initiation of replication. The ToLCV high-affinity Rep binding site is located adjacent to the stem-loop containing the nick site, as has been found for other begomoviruses.

Comparison with DNA-β components shows that ToLCV-sat contains a slightly modified SCR at this position, but the binding site is located outside of this region, within stem-loop II. In contrast to all other reports for begomoviruses, it has been demonstrated that the binding sites are not essential for ToLCV and ToLCV-sat proliferation in tomato (Lin *et al.,* 2003), indicating that high-affinity binding is not a prerequisite for replication for this monopartite begomovirus and its satellite. However, it is possible that a more transient interaction involving the binding site, a cryptic version of the motif or an entirely unrelated sequence, is necessary to correctly position Rep within the origin. The apparently indispensable nature of the high-affinity binding site for replication of bipartite begomoviruses (Fontes *et al.,* 1994a and Fontes *et al.,* 1994b) suggests that it may have evolved to ensure that both genomic components are stably maintained.

DNA-1 components have been found in association with several distinct monopartite begomoviruses that have been isolated from a range of plant species growing in different regions throughout Africa and Asia (Mansoor *et al.,* 1999, Mansoor *et al.,* 2000a, Mansoor *et al.,* 2000b, Mansoor *et al.,* 2001, Saunders and Stanley, 1999, Briddon *et al.,* 2004 and Wu and Zhou, 2005). Although they are frequently found in association with monopartite begomovirus/DNA-β complexes, there have been no reports to suggest that DNA-1 components occur naturally in association with bipartite begomoviruses.

1.2.5 Functions of begomovirus subviral agents

Begomovirus DNA-β complexes induce a variety of host-specific symptoms ranging from a vein yellowing phenotype, in *Ageratum, Eupatorium* and *Honeysuckle*, that has no obvious adverse effect on plant survival (Wong *et al.*, 1993, Saunders *et al.*, 2003 and Were *et al.*, 2005a), to severe leaf curl, chlorosis and stunting in crops such as cotton, tobacco and mungbean, that can have a significant effect on productivity (Harrison *et al.*, 1997, Briddon and Markham, 2000, Cui *et al.*, 2004b and Rouhibakhsh and Malathi, 2005). Initial studies to resolve the aetiology of AYVD and CLCuD demonstrated that the DNA-β components made an important contribution to the disease phenotype (Saunders *et al.*, 2000 and Briddon *et al.*, 2001). More recently, βC1 has been shown to be a repressor of host immunity i.e. RNA silencing (Cui, *et al*., 2005). AYVD DNA-β does not significantly affect AYVV replication in an *N. benthamiana* leaf disk assay although it does enhance the systemic accumulation of the helper begomovirus in its natural host ageratum (Saunders *et al.*, 2000).

1.2.6 Common features between begomovirus components and satellites

The unique common feature for all begomoviruses and their satellites (**Fig. 3**) is the rolling circle replication (RCR) mechanism and the presence of a stem loop with the nanonucleotide origin of replication (TAATATT/AC for betasatellites or TAGTATT/AC for alphasatellites). RCR is a basic replication mechanism, for the replication of bacterial and archaeal plasmids as well as circular human, animal and plant viruses, which depends on a replication-associated protein, possessing nicking and ligation functions of double stranded DNA replicative forms (Ilyina and

Koonin, 1992). The begomovirus-encoded Rep protein performs the same functions as prokaryotic plasmid replication associated proteins during the RCR (Jeske *et al.*, 2001).

Although begomoviruses and their satellites are highly diverse in nucleotide sequence of their genomes and ORFs (Briddon *et al.*, 2008), the genetic architecture, localization, length of individual genes at specific locus and the function of individual proteins are highly conserved among the members of each genus (**Fig. 3**). For example, members of the genus *Mastrevirus* have four ORFs in their genomes, LIR and SIR and a Rep protein interrupted by an intron, irrespective of the species considered, monocot or dicot hosts and geographical distribution (Stanley *et al.*, 2005). The *TrAP* proteins of the OW *Tomato yellow leaf curl China virus* (TYLCCNV-C2), *African cassava mosaic virus* (ACMV-AC2), *Mungbean yellow mosaic virus* (MYMV-AC2), and the NW *Tomato golden mosaic virus*, (TGMV-AC2) have all been associated with a transcription activator function of late viral genes (Cui *et al.*, 2005). Although small differences are present in their mode of action and regulation of genes, all TrAPs of these diverse viruses share the trans-activation of viral genes and nuclear localization. Similar conclusions can be drawn for all geminivirus and satellite genes.

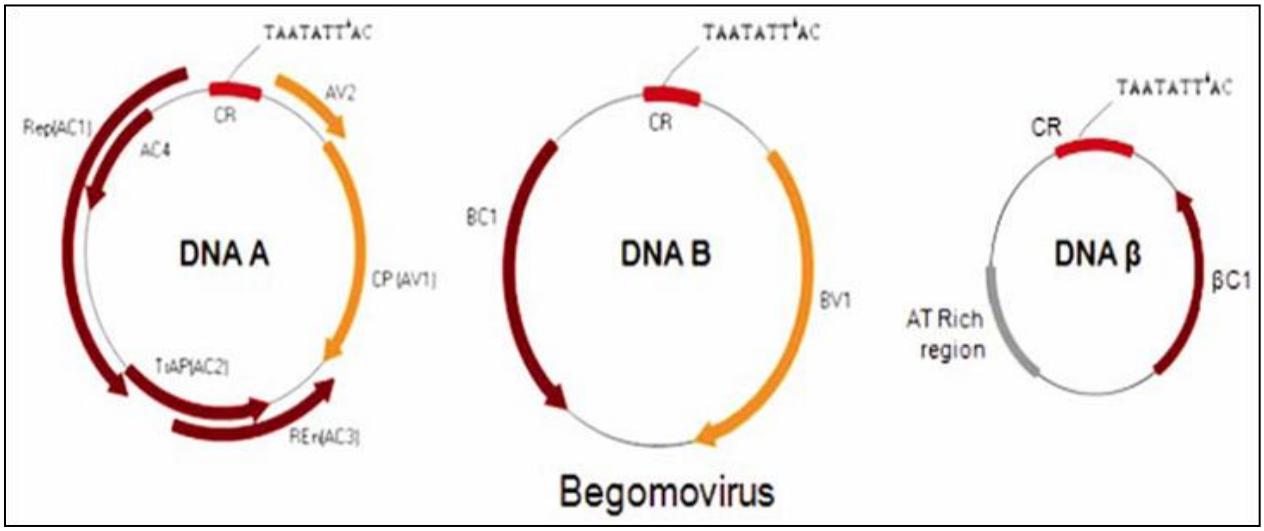


Figure 3: Cassava mosaic begomovirus genome structure composed of twinned particles; DNA-A (left) and DNA-B (right) represent bipartite genome and satellite DNA-β is often associated with DNA-A component of monopartite genome. BC1 protein of DNA β performs the role of RNAi suppression in monopartite begomoviruses. The location of the conserved region of all geminiviruses, namely, TAATATT/AC has been indicated (Briddon and Stanley, 2006).

1.2.7 Cassava begomovirus-associated satellites

Two novel subviral DNAs (DNA-II and DNA-III) have recently been isolated from cassava infected with bipartite begomoviruses (Ndunguru *et al.,* 2005) in Tanzania. DNA-II and DNA-III (AY836366 and AY836367, respectively) were discovered in cassava-growing areas in the coastal region and Lake Victoria basin in north-western of Tanzania respectively, and due to their ability to modify symptoms in cassava when co-inoculated with their cognate begomovirus, were thought to be satellite-like. Cassava plants containing DNA-II displayed very severe

symptoms, characterised mainly by leaf distortion, narrowing, yellowing, and mosaic. Cassava plants from which DNA-III was detected expressed very unique symptoms, characterized by a severe leaf narrowing due to rapid loss of leaf blades (filiform) and very prominent leaf yellowing (Ndunguru *et al.*, 2005). These circular single stranded DNA-II and III molecules are relatively small, approximately 1032 and 1209 nucleotides, respectively. They contain a GC-rich region, and yet are distinct from each other (23% nucleotide identity) and from all begomoviruses and other subviral components. As they have been isolated only from symptomatic plants, they may represent a novel type of satellite DNA that is adapted to bipartite begomoviruses.

The genome structure of cassava mosaic begomovirus-associated satellite-like DNA molecules consist of eight putative ORFs and four ORFs for DNA-II and DNA-III, respectively (**Fig. 4**). Whether these ORFs of DNA-II and DNA-III are actually expressed is still unknown. Although little is yet known about their replication and gene expression strategies, infectivity studies have demonstrated that DNA-II and DNA-III enhance symptoms in cassava caused by *African cassava mosaic virus*, *East African cassava mosaic virus* and *East African cassava mosaic Cameroon virus*, and allow these begomoviruses to produce symptomatic infections in resistant cassava landrace (TME3) (Ndunguru *et al.*, 2008 unpublished data). These novel satellites need a more detailed analysis and assessment of their contribution to the cassava mosaic disease pandemic caused by begomoviruses that is currently affecting many central and east African countries (reviewed by Legg and Fauquet, 2004).

An earlier study in Tanzania also showed a high occurrence of DNA-II and III in a large proportion of cassava leaf samples tested (Ndomba, 2012). The aim of this study was to establish genetic diversity and geographic distribution of DNA-II and DNA-III in Kenya, Uganda, Rwanda and Tanzania. Moreover the study aimed to explore the interaction of these unusual sub-genomic DNA-II and DNA-III with their associated begomovirus such as viruses EACMV-UG2 and ACMV in a screen-house experiment.

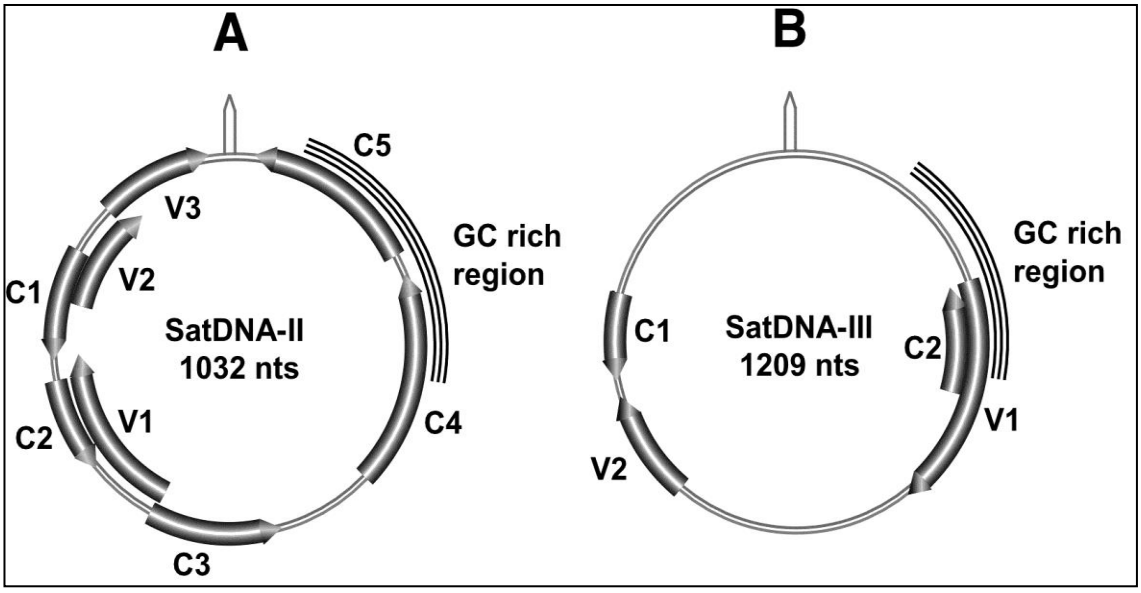


Figure 4: Genomic structure of cassava mosaic begomovirus associated satellite-like DNA molecules. (A) satDNA-II with eight putative ORFs (V1toV3) on sense strand and (C1toC5) on antisense strand, and (B) satDNA-III with four putative OFRs (V1toV2) on sense strand and (C1to C2) on antisense strand (Adapted from Ndunguru *et al*., unpublished)

CHAPTER TWO

A STUDY OF THE IMPACT OF DNA-II AND DNA-III ON CASSAVA MOSAIC DISEASE, AND INVESTIGATION INTO TRANSCRIPTION OF PUTATIVE EPISOMAL ORFs

2.1 Abstract

A study was conducted to investigate the transcription of putative ORFs of DNA-II and DNA-III molecules from cassava plants inoculated with either EACMV-Ug2, or ACMV, and co-inoculated with DNA-II, DNA-III or both DNA-II and III infectious clones. The experiment was conducted at Mikocheni Agricultural Research Institute greenhouse in Tanzania in 2012. Young cassava mosaic disease symptomatic and symptomless cassava leaves from inoculated cassava plants were collected for the analysis. Using primers to amplify selected regions (putative ORFs) of the DNA-II and III genome, RT-PCR yielded transcripts of 250 bp of ORF C4 for DNA-II and 220 bp of ORF V1 for DNA-III, and these were confirmed by sequencing. Transcripts of ORF C4 for DNA-II had >94% nt sequence similarity to cassava begomovirus associated satellite DNA-II deposited in the Genebank (AY836366), while ORF V1 of DNA-III had >94% nt similarity to cassava begomovirus associated satellite DNA-III (AY836367). This study reports two ORFs, ORFs C4 for DNA-II in antisense strand and ORF V1 of DNA-III in sense strand which could be transcribed in all experimental cassava treatments (healthy non-infected and infected) of cv.60444 and TME3 cassava cultivars, confirming that these ORFs are also part of expressed cassava genome-derived transcripts.

2.2 Introduction

Plant viruses are often associated with satellite DNA/RNA molecules, which modulate replication and enhance symptom expression of their cognate helper virus (Roossinck *et al*., 1992; Simon *et al*., 2004). These satellite molecules completely lack sequence identity to their helper viruses and depend on the helper virus for replication, encapsidation, movement, and transmission (Mayo *et al*., 2005). Satellites were initially reported to be associated with RNA viruses and are very well characterized (Simon *et al*., 2004). In the last decade, more than 500 satellite sequences associated with DNA begomoviruses (family *Geminiviridae*) have been isolated from a diverse range of cultivated crops and weeds (Briddon and Stanley, 2006; Briddon *et al*.,2008). Two classes of DNA satellites are associated with several monopartite begomoviruses, namely alphasatellites (formerly called DNA-1) and betasatellites (formerly called DNA-β) (Briddon *et al*., 2008). In addition to these sub-viral components, geminiviruses are accompanied by smaller-sized DNA molecules, called defective DNA molecules, that are derived from the helper virus genomes, but are not satellite molecules (Patil and Dasgupta, 2006; Simon *et al*., 2004). Alphasatellites are nanovirus-like components approximately 1.3 kb that, in some cases, suppress viral disease symptoms. Although alphasatellites encode a replication-associated protein, they depend on the helper virus encoded proteins for movement and encapsidation (Briddon and Stanley, 2006; Nawaz-ul-Rehman and Fauquet, 2009). The first betastellite components identified were associated with cotton leaf curl disease (CLCuD) and Ageratum yellow vein disease (AYVD) complexes. Since then, many betasatellites have been reported to be associated with a diverse range of begomovirus disease complexes (Briddon *et al*., 2004; Mansoor *et al*., 2001; Saunders and Stanley, 1999).

Two novel satellite-like DNA molecules (episomal DNA-II and DNA-III) originating in Tanzania have recently been amplified, using universal betasatellite primers (Briddon *et al*., 2001), from cassava infected with bipartite begomoviruses in cassava cultivation areas in the Coastal and Lake Victoria regions (Ndunguru, 2005). Cassava plants containing episomal DNA-II displayed very severe symptoms, characterized mainly by leaf distortion, yellowing, and mosaic. Cassava plants from which episomal DNA-III was detected expressed very unique symptoms, characterized by a severe leaf narrowing due to rapid loss of leaf blades (filiform) and very prominent leaf yellowing (Ndunguru *et al.,* 2005). These DNA molecules are relatively small (~1032 bp for episomal DNA-II and ~1209 bp for episomal DNA-III) and contain a GC-rich region, yet are distinct from each other (23% nucleotide identity), from all known geminiviruses and other sub-viral components.

A previous study conducted in 2005 showed that these DNA molecules were isolated only from cassava plants showing symptoms for cassava mosaic disease (CMD). A more recent study (unpublished) showed that even in healthy cassava plants, these DNA-II and III molecules could be detected by PCR in a partial form which resembled the episomal sequences deposited in Genbank (AY836366 and AY836367), suggesting that fragments of DNA-II and DNA-III are integrated into the cassava genome. Although little is yet known about their function, putative mode of replication and gene expression strategies, infectivity studies have demonstrated that episomal DNA-II and DNA-III enhance symptoms in cassava caused by *African cassava mosaic virus*, *East Africa cassava mosaic virus*-UG2 (ECMV-UG2) *and East African cassava mosaic Cameroon virus* (EACMCV), and allow these begomoviruses to produce symptomatic infections in resistant cassava landrace (TME3) (Ndunguru *et al.,* 2008 unpublished data). More recently,

portions (nearly full-length to smaller fragments) of DNA-II and DNA-III have been confirmed to be integrated in genes at different locations in the cassava genome. Although both DNA-II and DNA-III have several putative ORFs, it is not known if they are transcribed or not.

In this study we have established that transcripts of ORF C4 in the antisense strand of DNA-II, ORF V1 in the sense strand, ORF C2 in the antisense strand of DNA-III molecule and part of GC-rich region were expressed from mock- inoculated cassava cultivars TME3 and cv. 60444 (cassava plants inoculated with gold particles via biolistics) and diseased cassava plants (cassava plants inoculated with EACMV-Ug2 alone, ACMV alone and co-inoculated with episomal DNA-II or III or both DNA-II + DNA-III clones). Using RT-PCR, transcription from ORF C4 of DNA-II and ORF VI for DNA-III was demonstrated using designed primers (**Table 5**).

2.3 Materials and Methods

2.3.1 Micropropagation and acclimatization of cassava cultivars TME3 and cv. 60444

The experiment was conducted at Mikocheni Agricultural Research Institute (MARI), in Tanzania in which two cassava cultivars, TME3 resistant to cassava mosaic disease (CMD) and cv. 60444 a susceptible (control) were used. Plantlets of cassava cultivars were obtained from tissue culture using nodal cuttings. Prior to biolistic inoculation, TME3 and cv. 60444 were screened for the presence of CMD, associated episomal DNA-II and DNA-III and integrated DNA molecules (DNA-II and DNA-III) using specific primers. CMD- free (but not integrated

DNA-II or DNA-III molecules) cassava plants were micro propagated in Murashige and Skoog media (MS + Vitamin, sucrose and phytogel) in tissue culture laboratory. Cassava nodes were cut from the base of plants and subculutred in a sterile bottle containing MS media asceptically in lamina floor hood. Cassava plantlets were kept in growth room of 28°C for 16 hours light and 8 hours dark for shoot and root development until ready for acclimatization. Then cassava plantlets were transferred into pots and kept in the greenhouse free from insects for 8 weeks until use (**Fig. 5**).



Figure 5: Stages for acclimatization process; A) TME3 tissue culture cassava plantlets B) cv. 60444 tissue culture cassava plantlets C) Four- week old cassava plants in small pots D) 8 week-old cassava plants in big pots ready for inoculation.

2.3.2 Biolistic inoculation of virus and DNA-II and III clones

2.3.2.1 Preparation of gold particles

Sixty micrograms of gold particles were sterilized in 800 µl absolute ethanol, vortexed vigorously for 5 min and spun down at 13,000 rpm for 2 min. The supernatant was removed, 500 µl of absolute ethanol was added, vortexed for 2 min and centrifuged at 13,000 rpm for 2min. The supernatant was removed; the pellet was washed with 500 µl of sterile double distilled water, vortexed for 1 min and centrifuged at 13,000 rpm for 1 min. After washing gold particles, the supernatant was removed and then 500 µl of sterile double distilled water was added, vortexed and split into 10 tubes each with 50 µl.

2.3.2.2 Coating of virus and DNA-II and III infectious clones

Infectious clones of helper viruses, EACMV-UG2 and ACMV DNA components A and B were obtained from MARI in pBlueScript vector while DNA-II and DNA-III infectious clones were in pGEM-T^{R} Easy vector. One microgram of each DNA component, DNA-II and DNA-III clones was coated with 50 µl of gold particles as indicated in Table 3. The gold particles were vortexed, and while vortexing, 1µg of DNA A, DNA B of EACMV-UG2 and ACMV components, DNA-II and DNA-III clones were added. This was followed by the addition of 20µl of 0.1M spermidine and 50 µl of 2.5M calcium chloride. The mixture was left on the bench for 10 min. The mixture was then spun at 13,000 rpm for 2 min and the resultant supernatant was removed

and 500 µl of absolute ethanol was added. The mixture was vortexed for 2 min, centrifuged at 13,000 rpm for 1 min and the supernatant was removed. Lastly, 50 µl of absolute ethanol was added to the coated DNA, vortexed and used to shoot 8-week old cassava plants in pots at a pressure of 90 psi using a handheld DNA gene gun (Biorad). Coated DNA was inoculated biolistically into TME3 or cv. 60444 cassava plants (**Table 3**). Mock plants were used as negative controls. After bombardment, cassava plants were monitored visually in the greenhouse for symptom development over 50 days and symptoms severity was scored using a scale of 1-5 (1 represents symptomless cassava plants and 5 represents most severe symptoms in cassava plants).

Table 3: Number of treatments used in studying putative transcription of DNA-II and DNA-III

Treatments	Virus specie/strain	DNA molecules	No. of plants
1	EACMV-UG2	-	4
2	EACMV-UG2	II	4
3	EACMV-UG2	III	4
4	EACMV-UG2	II + III	4
5	No virus	II	4
6	No virus	III	4
7	No virus	II + III	4
8	Mock	-	4
9	ACMV	-	4
10	ACMV	II	4
11	ACMV	III	4
12	ACMV	II + III	4
13	No virus	II	4
14	No virus	III	4
15	No virus	II + III	4
16	Mock	-	4

2.3.2.3 Transcript determination in the geminivirus – DNA-II and DNA-III - cassava host systems

2.3.2.3.1 RNA extraction

Young apical symptomatic and symptomless cassava leaves were collected from biolistic-inoculated cassava plants in greenhouse (**Table 4**). Total RNA was extracted using modified CTAB protocol (Allen *et al*., 2006). Leaf samples were ground with pestle in a mortar containing warmed 700 µl of 2% CTAB extraction buffer with 98% β-mercaptoethanol (10% CTAB, 1M Tris-HCL, 0.5M EDTA, 5M NaCL), and the homogeneous mixture was transferred to clean 1.5 ml eppendorf tube. The mixture was incubated at 65°C in a water bath for 30 minutes, mixed by inverting the tubes every 10 minutes and left at room temperature for 10 minutes. An equal volume (700 µl) of chloroform: isoamyl (24:1) was added to the mixture and inverted for 10 minutes and centrifuged at 13,000 rpm for 15 min. The supernatant (500 µl) was transferred to a new 1.5 ml tube; an equal volume of cold isopropanol was added to each tube and centrifuged at 13,000 rpm for 10 min to precipitate RNA. The RNA pellet was washed with 500 µl of 70% ethanol treated with DEPC (diethylpyrocarbonate), air dried for 40-60 minutes and re-suspended in 50 µl of DEPC treated water followed by further DNAse1 treatment.

Table 4*:* Reaction of cassava plants to biolistic inoculation with EACMV-UG2, ACMV, and in combination with DNA-II, DNA-III and DNA-II + DNA-III infectious clones

Cassava variety	Treatments	No. of plants inoculated	Symptomatic plants
	EACMV-UG2	4	3
	EACMV-UG2 + II	4	4
	EACMV-UG2 + III	4	4
	EACMV-UG2 + II + III	4	4
	II	4	0
	III	4	0
	II + III	4	0
	Mock	4	0
	ACMV	4	4
	ACMV + II	4	4
	ACMV + III	4	4
	ACMV + II + III	4	4
	II	4	0
	III	4	0
	II + III	4	0
TMS 60444	Mock	4	0
	EACMV-UG2	4	0
	EACMV-UG2 + II	4	1
	EACMV-UG2 + III	4	0
	EACMV-UG2 + II +III	4	2
	II	4	0
	III	4	0
	II + III	4	0
	Mock	4	0
	ACMV	4	3
	ACMV + II	4	4
	ACMV + III	4	4
	ACMV + II + III	4	4
	II	4	0
	III	4	0
	II + III	4	0
TME3	Mock	4	0

2.3.3.2 Primer design and RT- PCR analysis of ORF C4 for DNA-II and ORF V1 for DNA-III

Primers to amplify putative ORFs for DNA-II and DNA-III, were designed from the genome map of DNA-II (AY836366) and DNA-III (AY836367) respectively (**Fig. 6**). Primer pairs SatII-2F/3R and SatIII-7F/9R were used to detect transcripts for DNA-II ORF C4 in the antisense strand and ORF V1 in the sense strand, ORF C2 in the antisense strand and part of GC-rich region for DNA-III, respectively (**Table 5**). A two step Reverse Transcriptase-Polymerase chain reaction (RT- PCR) procedure was performed. Total genomic RNA was first transcribed into cDNA using oligo dTs primer according to manufactures instructions. cDNA synthesis involved 2 steps; Master mix I involved sterile double distilled water (sdd H_2O) (9 µl), oligo dT (1 µl) and RNA as a template (2 µl). Then the mixture was incubated for 10min at 70°C, cooled on ice for 5 min and master mix II was added. Master II contained sdd H_20 (6.5 µl), 5 X RT buffer (4.0 µl), dNTPs (1 µl), RNAsin (0.5 µl) and RT enzyme (1 µl). The mixture was heated at 42°C for 55 min, 70°C for 10 min and then stored at 4°C for further use. The resulting cDNA was used as a template in PCR of 30 cycles each consisting of 94°C, 1 min; 94°C, 1 min; 56°C, 1 min; 72°C, 2 min; 72°C; 10 min using designed specific primers from ORF C4 and ORF V1 for DNA-II and DNA-III respectively (Table 5). PCR was performed in a 50 µl reaction volume, which contained sterile distilled water (41.25 µl) 10x Dream Taq buffers (3.5 µl), 2.5 mM dNTPs (1.0 µl), 10 mM each amplification primer (1.0 µl), Dream Taq (0.25 µl) and cDNA template (2.0 µl). After completion of PCR, samples (25.0µl) were loaded onto a 1% agarose gel in 1xTAE (Tris-acetate- ethylenediaminetetraacetic acid) buffer and the ethidium stained DNA-bands visualized using ultraviolet light.

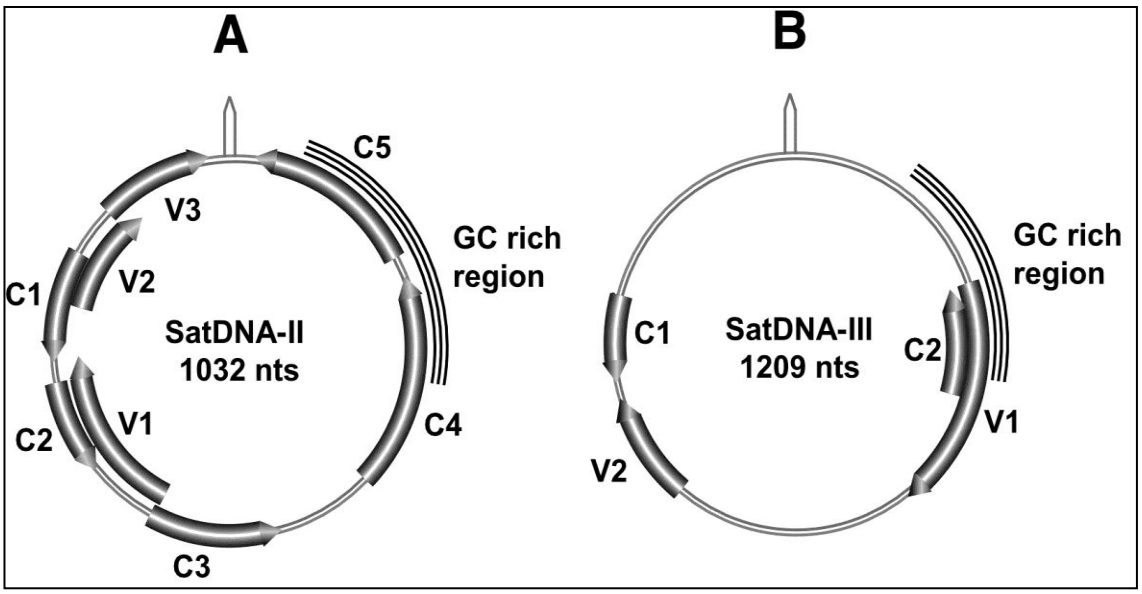


Figure 6: Genomic structure of cassava mosaic begomovirus associated satellite-like DNA molecules. (A) satDNA-II with eight putative ORFs (V1-V3) on sense strand, and (C1-C5) on antisense strand, and (B) satDNA-III with four putative OFRs (V1-V2) on sense strand and (C1-C2) on antisense strand (Adapted from Ndunguru *et al*., unpublished)

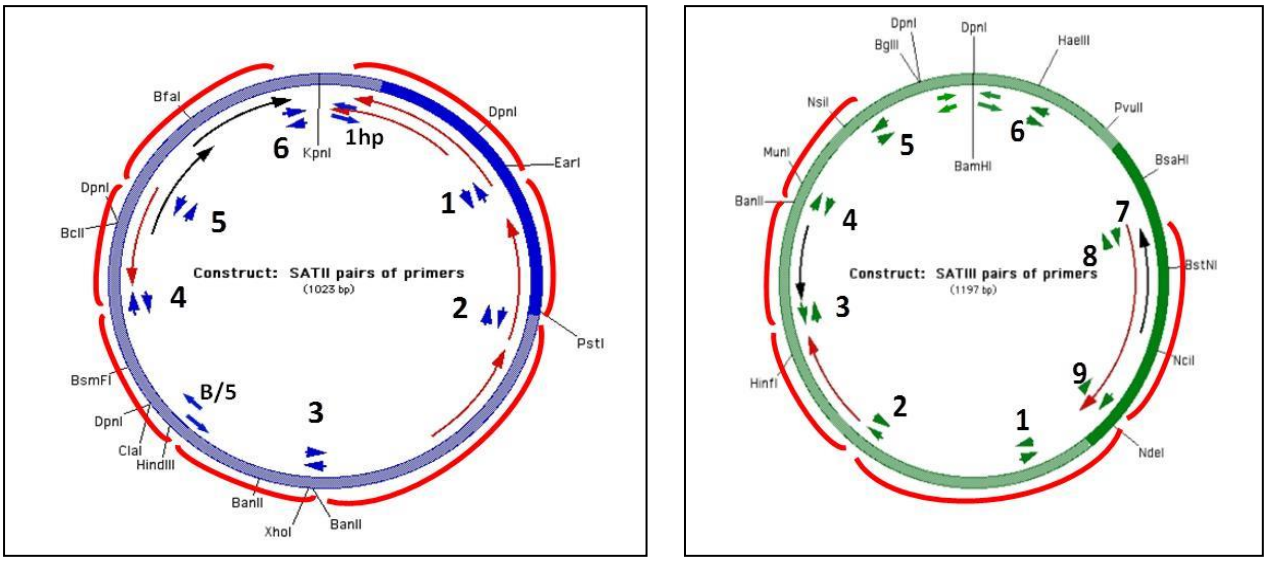


Figure 7; Genome maps of circular DNA-II and III molecules showing primer positions

Table 5: List of primer pairs used for transcript detection of ORF C4 of DNA-II in antisense strand, ORF C2 of the antisense strand and ORF VI of the sense strand of DNA-III.

S/N	Primer name	Primer sequence 5' - 3'	Remark
1	IIIV1-F	TACGAAAGATGAGGAGAAG	Detect transcript for ORF V1 for DNA-III in sense strand
2	IIIV1-R	AAAAGCACACATATGAAACCA	Detect transcript for ORF V1 of DNA-III in sense strand
3	SatII-2F	GCAGTTCAGCAGTTCAACTG	Detect transcript for ORF C4 of DNA-II in antisense strand
4	SatII-3R	TCGAGCTCCAAAGGTCCT	Detect transcript for ORF C4 of DNA-II in antisense strand
5	SatIII-7F	CATGCTGTCAACGCCATTGCTG	Detect transcript for ORF V1 in sense strand, ORF C2 in the antisense strand and part of GC-rich region of DNA-III
6	SatIII-9R	AACCACAGTGATTGACGTAGC	Detect transcript for ORF VI in sense strand, ORF C2 in the antisense strand and part of GC-rich region of DNA-III

2.3.3.4. Sequencing

RT-PCR products were precipitated with the addition of 10 µl of 3M sodium acetate and 300 µl absolute ethanol, and then centrifuged at 13,000 rpm for 15 min. The pellet was washed with 400 µl 70% ethanol and air dried at room temperature for 1 hour. Precipitated RT-PCR products were sent to North Carolina State University for sequencing using primers used for RT-PCR. Sequences were aligned with MUSCLE alignment in Geneious Computer Software. All sequence editing was performed using BioEdit computer software. Three sequences for each primer set were done and the consensus sequence was used for analysis.

2.4 Results

2.4.1 Symptoms observation

Cassava mosaic disease (CMD) symptoms were observed visually on cv. 60444 and TME3 cassava plants inoculated with EACMV-UG2 alone, EACMV-UG2 + DNA-II, EACMV-UG2 + DNA-III, EACMV-UG2 + DNA-II +DNA-III, ACMV alone, ACMV + DNA-II, ACMV + DNA-III and ACMV + II + III 7days post inoculation (dpi) except for TME3 cassava plants inoculated with EACMV-UG2 alone (Table 4). Cassava plants of cv. 60444 and TME3 inoculated with DNA-II alone, DNA-III alone, both DNA-II + DNA-III and mock plants (cassava plants inoculated with gold particles) did not develop CMD symptoms as expected (Table 4).

Cassava plants inoculated with either EACMV-UG2, ACMV, and in combination with DNA-II, DNA-III and both DNA-II + DNA-III clones developed several CMD symptom phenotypes (**Fig. 8 & 9**) except for Figure 7E which did not show symptoms. CMD symptoms observed were mosaic, yellowing and leaf deformation and symptoms started as mild symptoms and then developed into severe symptoms over time. Fifteen out of sixteen plants of cv. 60444 inoculated with EACMV-UG2 alone and in combination with DNA-II, DNA-III and DNA-II + DNA-III clones developed cassava mosaic symptoms. Symptoms were observed on young emerging leaves on plants inoculated with EACMV-UG2 alone and in combination with DNA-II or / and III. TME3 (tolerant cassava cultivar) showed cassava mosaic symptoms on plants inoculated with EACMV-UG2 in combination with DNA-II, DNA-III and both DNA-II + DNA-III clones. Three plants out of sixteen TME3plants inoculated with EACMV-UG2 alone, and in combination with DNA-II, DNA-III and both DNA-II + DNA-III clones developed CMD symptoms (**Table. 4**). In addition, unusual malformation of cassava leaves (**Fig. 8 F& G**) was observed in TME3 plants inoculated with EACMV-UG2 in combination with DNA-II and DNA-III clones.

Cassava mosaic disease symptoms were also observed on TME3 and cv. 60444 plants biolistically inoculated with another different virus ACMV alone and in combination with DNA-II, DNA- III and both DNA-II + DNA-III clones (**Fig. 9**). CMD symptoms observation differed between inoculated cassava cultivars cv. 60444 and TME3 infected with ACMV alone or co-inoculated with DNA-II, DNA-III and both DNA-II + DNA-III clones. Symptoms observed in the susceptible cassava cultivar, cv. 60444, were much more severe (3 scale) compared with resistant cassava cultivar, TME3 (2 scale), when inoculated with ACMV in combination with

DNA-II and III clones (**Fig. 9 D& H**). All 16 plants of cv. 60444 inoculated with ACMV alone and in combination with DNA-II, DNA-III and DNA-II + DNA-III clones developed CMD symptoms while 15 plants out of 16 TME3 plants inoculated with ACMV alone, and in combination with DNA-II, DNA-III and both DNA-II + DNA-III clones developed CMD symptoms as shown previously (**Table. 4**).

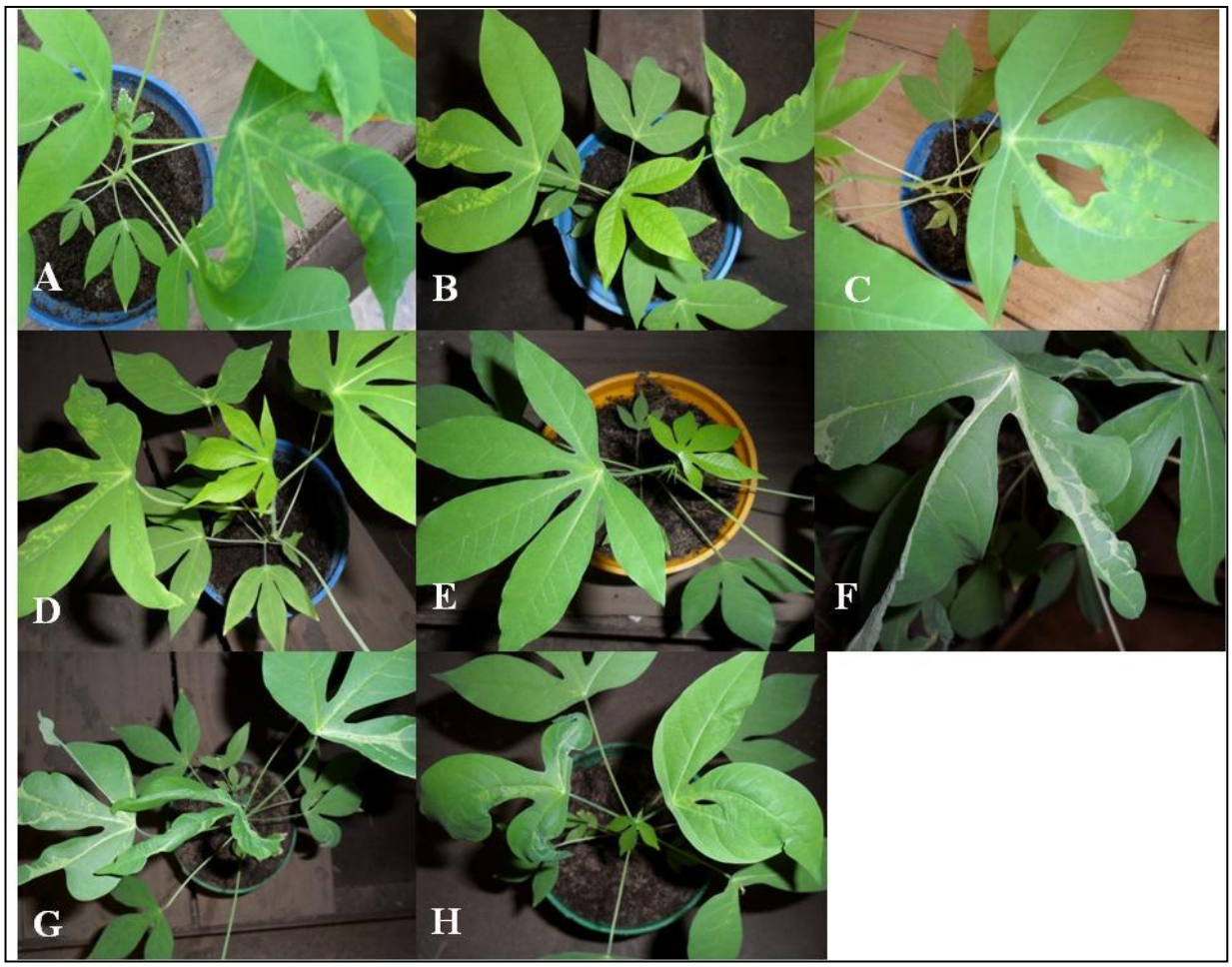


Figure 8: Cassava mosaic disease symptom development on cv. 60444 plants inoculated with A) EACMV-UG2 alone, B) EACMV-UG2 + DNA-II, C) EACMV-UG2 + DNA-III, D) EACMV + DNA-II & III. CMD symptom development on TME3 plants inoculated with E) EACMV-UG2 alone, F) EACMV-UG2 + DNA-II, G) EACMV-UG2 + DNA-III, H) EACMV-UG2 + DNA-II & III.

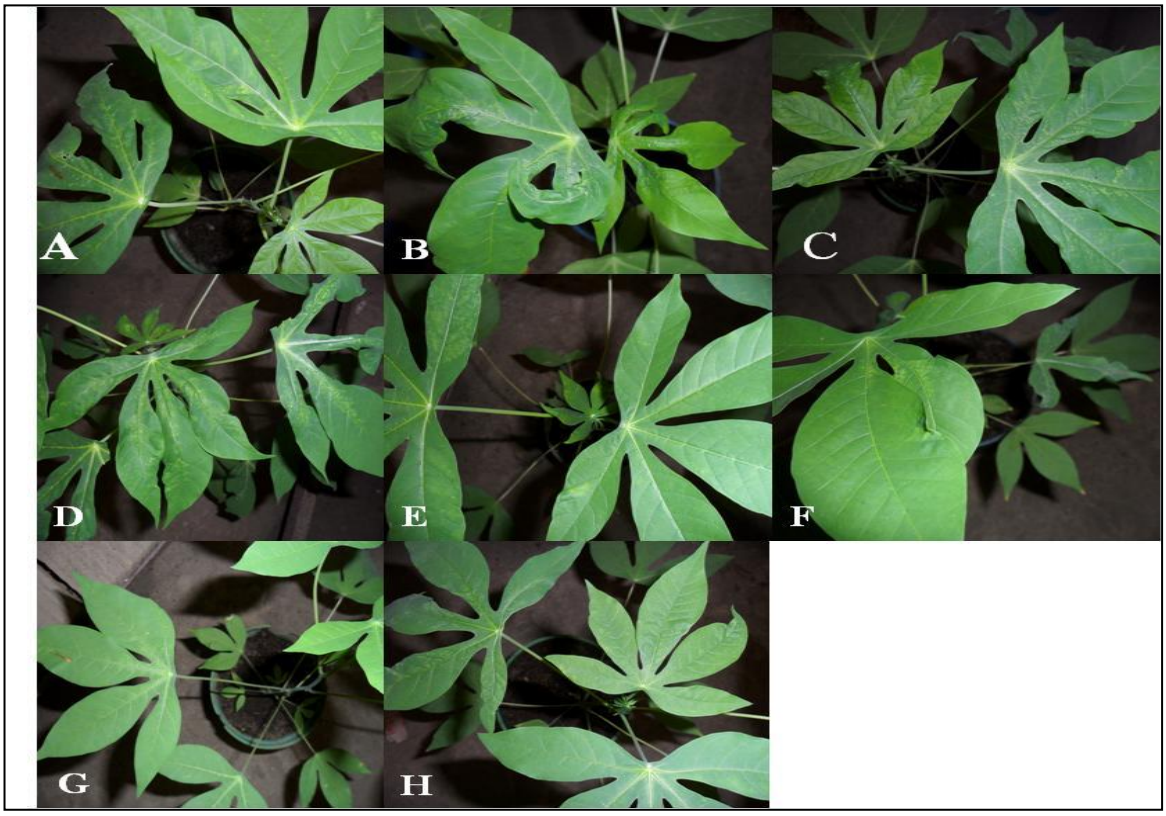


Figure 9: Cassava mosaic disease symptom development on cv. 60444 plants inoculated with A) ACMV alone, B) ACMV + DNA-II, C) ACMV+ DNA-III, D) ACMV + DNA-II & III.

CMD symptom development on TME3 plants inoculated with E) ACMV alone, F) ACMV+ DNA-II, G) ACMV + DNA-III, H) ACMV + DNA-II & III.

2.4.2 Symptoms severity and disease severity progress curves

Cassava mosaic disease symptom severity on inoculated cassava plants was recorded every 7 days over 7 week (49 days) period. From the disease progress curves (**Fig. 10**), generally CMD symptom severity score ranged from 1-3 with either of the four treatments in cv.60444 and TME3. The highest score was achieved after 14 days post inoculation in cv.60444 when inoculated with EACMV-UG2 + DNA-II, 21 day post inoculation in cv. 60444 and TME3

inoculated with EACMV-UG2 alone and EACMV-UG2 + DNA-II + DNA-III respectively, 28 day post inoculation in cv. 60444 inoculated with EACMV-UG2 + DNA-II + DNA-III and lastly 35 day post inoculation in cv. 60444 inoculated with EACMV-UG2 + DNA-III. However cassava plants inoculated with EACMV-UG2 alone and in combination with DNA-III had a severity score of 1 in TME3 throughout the 7 weeks monitoring period. Infection rates for EACMV-UG2-inoculated cassava plants were observed to be 93% and 12.5% for cv. 60444 and TME3, respectively and cassava plants inoculated with DNA-II, DNA-III , both DNA-II + DNA-III and gold particle only did not show symptoms up to 49 days monitoring, as expected.

From the disease progress curve (**Fig. 11**), generally CMD symptom severity scores ranged from 2-3 with either of the four treatments in cv.60444 and TME3. The highest score was achieved after 14 days post inoculation in cv.60444 and TME3 when inoculated with ACMV + DNA-II, ACMV-DNA-II + DNA-III and ACMV + DNA-III, ACMV + DNA-II + DNA-III respectively. In cv. 60444 inoculated with ACMV, ACMV + DNA-III, the highest severity score was achieved after 21 day post inoculation. After 28 and 35 day post inoculation in TME3, the highest severity score was achieved in cassava plants inoculated with ACMV + DNA-II and ACMV alone respectively.

The observation made in this experiment is that, cassava plants cv. 60444 and TME3 inoculated with either of EACMV-UG2 or ACMV and in combination with DNA-II, DNA-III or both DNA-II+ DNA-III developed CMD symptoms earlier that cassava plants inoculation with EACMV-UG2 or ACMV alone. However CMD symptoms never got severe with DNA-II, DNA-III and both DNA-II + DNA-III when compared to cv.60444 or TME3 inoculated with EACMV-

UG2 or ACMV alone. Infection rates for ACMV- inoculated plants were 100% and 93% for cv. 60444 and TME3, respectively and cassava plants inoculated with DNA-II, DNA-III , both DNA-II + DNA-III and gold particle only did not show symptoms up to 49 days monitoring, as expected.

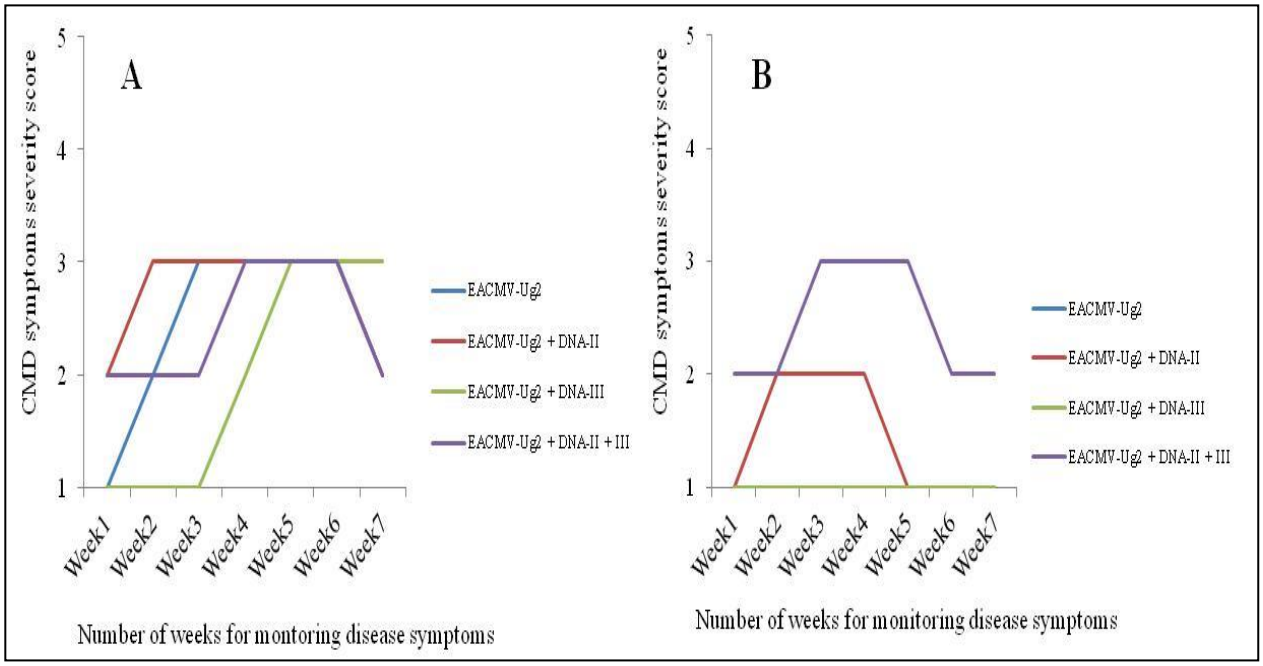


Figure 10: Disease severity progress curves for cassava plants inoculated with EACMV-UG2 and in combination with DNA-II, DNA-III and DNA-II + DNA-III in A) cv.60444 and B) TME3

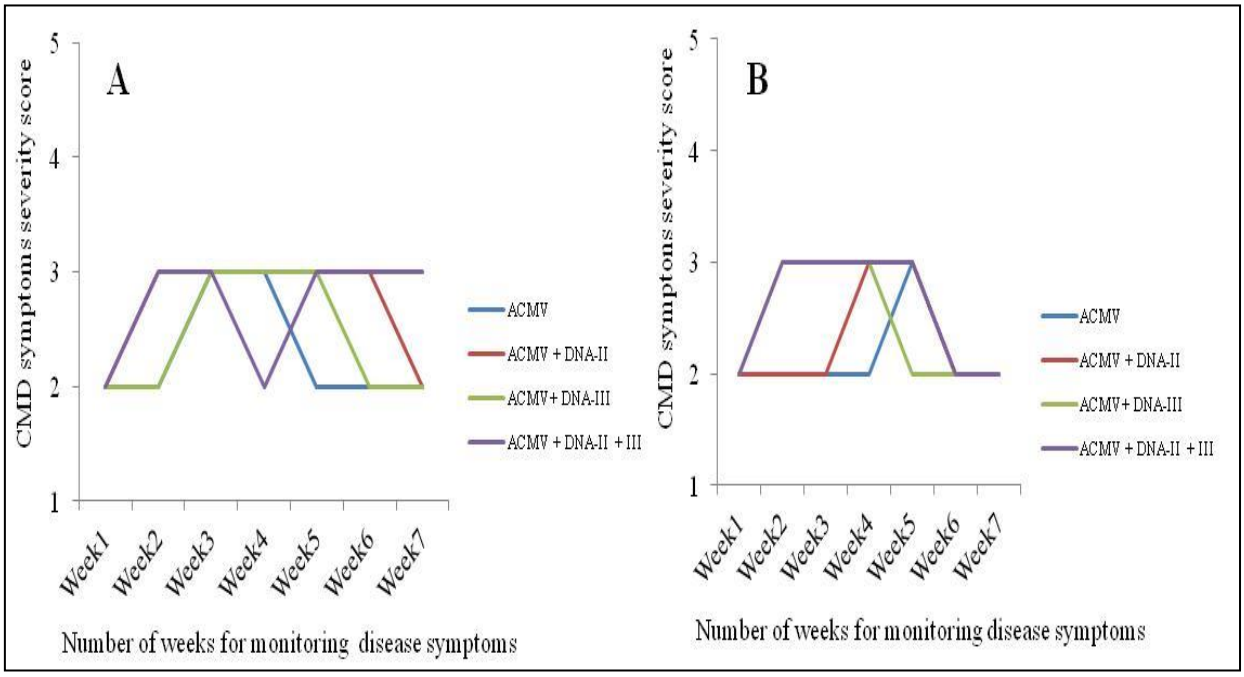


Figure 11: Disease severity progress curves for cassava plants inoculated with ACMV and in combination with DNA-II, DNA-III and DNA-II + DNA-III in A) cv. 60444 and B) TME3.

2.4.3. RT-PCR for ORF C4 of DNA-II, ORF V1 and ORF C2 of DNA-III

The amplification of transcripts for ORF C4 of DNA-II and ORF V1, ORF C2 of DNA-III were successful from cv.60444 and TME3 inoculated with EACMV-UG2 alone, ACMV alone, begomovirus co-inoculated with DNA-II, DNA-III, DNA-II + DNA-III and mock -inoculated cassava plants with gold particles. Primer pairs SatII-2F/SatII-3R and SatIII-7F/SatIII-9R amplified a DNA fragment band size of ~250bp for DNA-II and ~220bp for DNA-III, respectively (**Fig. 12A&B**) and primer pair IIIV1F/R yielded a band size of ~228 bp for DNA-III.

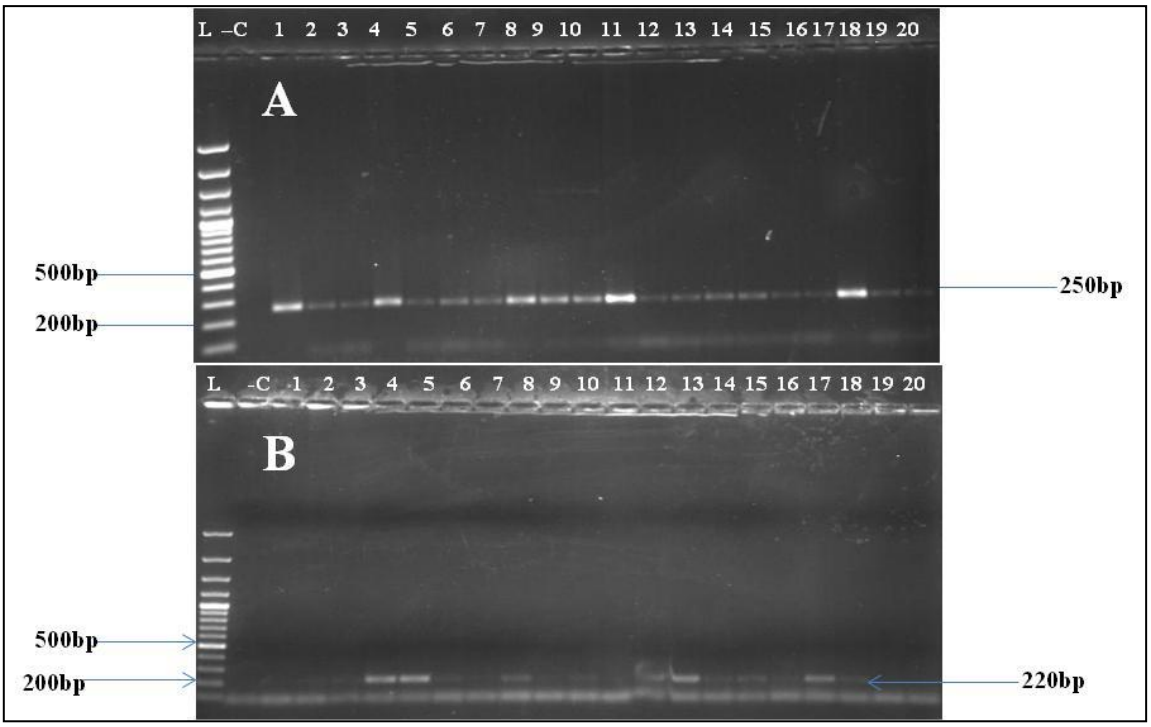


Figure 12: Gel picture showing amplification of ORF C4 and ORF V1 transcripts on 1% agarose by RT-PCR from cassava plants inoculated with either EACMV-UG2, ACMV and in combination with DNA-II and DNA-III clones A) DNA-II; and B) DNA-III. L= Molecular marker of 1kb plus, C= Negative control and, Lanes 1-20 are described in table 6 below.

Table 6: Samples collected from screen house for transcript analysis of DNA-II and DNA-III

Sample No	Sample description (Treatment)
1	EACMV-UG2 alone in cv. 60444
2	EACMV-UG2 + DNA-II in cv. 60444
3	EACMV-UG2 + DNA-III in cv. 60444
4	EACMV-UG2 + DNA-II +DNA-III in cv. 60444
5	Mock in cv. 60444
6	EACMV-UG2 alone in TME3
7	EACMV-UG2 + DNA-II in TME3
8	EACMV-UG2 + DNA-III in TME3
9	EACMV-UG2 + DNA-II + DNA-III in TME3
10	Mock in TME3
11	ACMV alone in cv. 60444
12	ACMV + DNA-II in cv. 60444
13	ACMV + DNA-III in cv. 60444
14	ACMV + DNA-II + DNA-III in cv. 60444
15	Mock in cv. 60444
16	ACMV alone in TME3
17	ACMV+ DNA-II in TME3
18	ACMV+ DNA-III in TME3
19	ACMV +DNA-II +DNA-III
20	Mock in TME3

2. 4. 4 Sequence analysis

2.4.4.1 Transcript sequences for DNA-II

From NCBI BLAST search, it showed that transcript sequences for DNA-II ORF C4 had highly (>94%) nt sequence similarity to the cassava begomovirus-associated satellite-like DNA-II (AY836366) in Genbank. Alignment results of ORF C4 transcript sequences from 19 DNA-II isolates illustrates that several sequences were highly conserved; for example nucleotide sequences AGGCATC at position 76-82, GGTAA at position 37-41 and GAGAA at position 117-121 (**Appendix** 2.1). Phylogenetic tree and similarity distance for transcripts from different samples are highly related up to 100% nucleotide sequence similarity (**Appendix** 2.2).

Table 7: Description of amplified RT- PCR samples for transcripts DNA-II sent for sequencing.

Code number	Sample number	Sample Description (Treatments)
1.12	12	ACMV + DNA-II in cv. 60444
2.14	14	ACMV + DNA-II + DNA-III in cv.60444
3. 2	2	EACMV-UG2 + DNA-II in cv.60444
4.11	11	ACMV alone in cv.60444
5.20	20	Mock in TME3
6.8	8	EACMV-UG2 + DNA-III in TME3
7.3	3	EACMV-UG2 + DNA-III in cv.60444
8.1	1	EACMV-UG2 alone in cv.60444
9.5	5	Mock in cv.60444
10.15	15	Mock in cv.60444
11.18	18	ACMV+ DNA-III in TME3
12.4	4	EACMV-UG2 + DNA-II + DNA-III in cv.60444
13.7	7	EACMV-UG2 + DNA-II in TME3
14.16	16	ACMV alone in TME3
15.17	17	ACMV+ DNA-II in TME3
16.13	13	ACMV + DNA-III in cv.60444
17.9	9	EACMV-UG2 + DNA-II + DNA-III in TME3
18.6	6	EACMV-UG2 alone in TME3
19.10	10	Mock in TME3

2.4.4.2 Transcript sequences for ORF V1 of DNA-III

BLAST search shows that transcript sequences for DNA-III had high (>94%) nt sequence similarity to the cassava begomovirus-associated satellite-like DNA-III (AY836367). Alignment results for transcript DNA-III illustrated that several nucleotide sequences from several plant sample isolates are highly conserved; for example nucleotide sequences TCGTCGCCTTTGGTCCTCAATCACA at position 94-124, nucleotide C at position 70, 73, 89, 124 and similarly nucleotide G is conserved at several positions for example at position 40 (**Appendix** 2.3). Phylogenetic tree and nucleotide similarity distances for transcripts from different isolates are highly related (up to 100% nt sequence similarity) (**Appendix** 2.4).

Table 8: Description of amplified RT- PCR samples for transcripts DNA-III sent for sequencing.

Code number	Sample number	Sample Description (Treatment)
1.27	27	EACMV-UG2+ DNA-II in TME3
2.37	37	ACMV + DNA-II inTME3
3.24	24	EACMV-UG2 + DNA-II + DNA-III in cv. 60444
4.25	25	Mock in cv. 60444
5.45	45	EACMV-UG2 alone in TME3
6.56	56	ACMV + DNA-II + DNA-III in TME3
7.52	52	ACMV + DNA-III in cv. 60444
8.4	40	Mock in TME3
9.29	29	EACMV-UG2 + DNA-II + DNA-III in TME3
10.39	39	ACMV + DNA-II + DNA-III inTME3
11.28	28	EACMV-UG2 + DNA-III in TME3
12.31	31	ACMV alone in cv. 60444
13.32	32	ACMV + DNA-II in cv. 60444
15.33	33	ACMV + DNA-III in cv. 60444
14.34	34	ACMV + DNA-II + DNA-III in cv. 60444
16.36	36	ACMV alone in TME3
17.22	22	EACMV-UG2 + DNA-II in cv. 60444
18.26	26	EACMV-UG2 alone in TME3
19.51	51	ACMV + DNA-II in cv. 60444
20.21	21	EACMV-UG2 alone in cv. 60444
21.35	35	Mock in cv. 60444
22.38	38	ACMV+ DNA-III inTME3
23.3	3	Mock in TME3

2.5 Discussion

From the biolistic co-inoculation experiment conducted (**Fig. 8**), it was observed that cassava mosaic disease (CMD) symptoms developed earlier on cv. 60444 compared to TME3 plants when inoculated with EACMV-UG2, or EACMV-UG2 co-inoculated with DNA-II or DNA-III, and both DNA-II + DNA-III infectious clones, and this may be due to the susceptible nature of cv. 60444 compared with the tolerant TME3 cultivar. It was difficult to differentiate between CMD symptom severity on cv. 60444 when plants were inoculated with EACMV-UG2 alone and in combination with DNA-II, DNA-III or both DNA-II + DNA-III, as the highest severity scores were the same. Probably the effect of DNA-II and DNA-III molecules on symptom severity could have been discerned if the experiment had been extended. TME3 plants tolerant to CMD developed cassava mosaic symptoms when EACMV-UG2 was co-inoculated with either DNA-II or both DNA-II + DNA-III clones.

Failure of EACMV-UG2 alone to infect TME3 but in the presence of either of the satellites or both indicated that EACMV-UG2 requires the episomal DNA-II and III molecules to induce infection. In addition, DNA-II and DNA-III molecules appear to work with EACMV-UG2 to break CMD tolerance in TME3, and enhance symptoms as previously reported (Ndunguru, 2008 unpublished data). It is postulated that these DNA-II and III sequences may participate in assisting EACMV-UG2 to overcome antiviral defense mechanisms (either post transcriptional gene silencing (PTGS) of virus derived transcripts, or transcriptional gene silencing (TGS) through methylation of viral genome). Several studies support a possible role of methylation as an important plant defense strategy against geminivirus. For example, *in vitro* methylation of

geminiviral DNA drastically reduced its infectivity (Brough., 1992; Ermak., *et al* 1993; Bisaro, 2006).

Another biolistic infection experiment (**Fig. 9**) was performed with a different begomovirus species, ACMV, and DNA-II, DNA-III and both DNA-II + DNA-III in cv. 60444 and TME3. Cassava mosaic disease symptoms severity ranged from 2-3 based on the CMD severity scale of 1-5 (Sseruwagi *et al*., 2004). The highest severity score was 3 in both cv. 60444 and TME3 however cv. 60444 plants developed symptoms earlier than TME3 plants. CMD symptoms were developed earlier on cv.60444 than TME3 inoculated with ACMV alone and in combination with DNA-II, DNA-III or both DNA-II + DNA-III. In cv. 60444, in contrast to EACMV-UG2, cassava mosaic symptoms was more severe in plants inoculated with ACMV in combination with DNA-II and III compared with those inoculated with ACMV alone, suggesting the influence of DNA-II and DNA-III on symptom enhancement and may also be due to susceptibility nature between ACMV and EACMV-UG2 in the cv. 60444. A research study conducted by (Ayeh and Ramsell., 2008) showed differences in susceptibility between begomoviruses in the same cassava cultivar. For example, a similar pattern of resistance/susceptibility has also been observed in some of the cultivars on cassava fields in Ghana. A study conducted in Uganda showed that different cassava varieties had different responses of susceptibility to begomovirus infection in the field as well as in the laboratory (Sserubombwe *et al*., 2008).

Tolerant TME3 cassava plants showed cassava mosaic symptoms when inoculated with ACMV alone on young emerging leaves and recovered 42 days post inoculation. The same phenomenon

was observed in the field trials when cassava plants inoculated with ACMV did not develop CMD symptoms. TME3 was reported to be tolerant to CMD because it has a dominant gene (CMD2) important to cassava production (Akano *et al*., 2002) and which was able to tolerate viral infection but recent studies showed that it is no longer tolerant (Ndunguru *et al*., 2008 unpublished). Recovery is a well-known phenomenon in many plants in addition to cassava (Hagen, *et al*., 2008; Chellappan *et al.,* 2005). Symptom recovery in the infected plants over time was shown to correlate with the accumulation of short interfering RNAs (siRNA) targeted against specific genes of the virus (Rodriguez-Negrete *et al*., 2008).

In the two cassava cultivars used in this study (i.e. cv. 60444 and TME3) differences in symptom expression were observed. In general CMD symptoms were observed earlier, and were more severe, in cv.60444 than TME3 plants, and this is due to the susceptibility of cv. 60444 to ACMV infection. Differences in symptom severity between the same cassava mosaic begomovirus in the different cultivars have been also reported by Ndomba (2012). For example a study conducted in Ghana revealed that there is a difference in symptom severity of cassava breeding lines infected by the same infectious clones. Variation in cassava cultivars genetic ability to resist CMD has also been reported by other studies (Lapidot and Friedmann, 2002; Ogbe *et al*., 2002).

Recovery in TME3 could also be due to the fact that some leaves exclude the spread of virus which may be related to expression of certain host defense mechanisms later on, leading to virus-free leaves. Viruses may not be evenly distributed between leaves from inoculated plants and some leaves may have lower virus titer due to patterns in virus movement.

Symptoms recovery from begomoviral infection of inoculated TME3 cassava plants in this study may be due to presence of major dominant gene (CMD2) that controls resistance to CMBs reported by Akano *et al*. (2002). But recent evidence showed that recovery to plant viruses in many cases, may be due to innate basal immunity known as RNA silencing (Fregene *et al*., 2001a; Dixon *et al*., 2001; Fraser, 2006, Patil and Fauquet, 2009). This occurs either in PTGS or TGS silencing leading to a decrease in virus replication. RNA silencing involves the production of small interfering RNAs (siRNAs) by the enzyme DICER, which splices viral mRNAs (Rodriguez-Negrete *et al*., 2008). The production of siRNAs in plants challenged or containing various geminiviruses can however be affected by temperature (Nelson and Citovsky, 2005). RNA silencing was reported to increase as the temperature was raised from 25°C to 30°C, but in this study this was not the case. Changes in symptom phenotype in this study may be related to the presence of partially characterized DNA-II and DNA-III molecules. This observation was also made from another study conducted by Ndunguru et al., (2013) (manuscript in preparation) who observed enhanced CMD symptom on TME3 cassava land race due to the presence of DNA-II and DNA-III in combination with the wild type EACMV-Ug2.

Putative ORFs were predicted for episomal DNA-II and DNA-III forms (Ndunguru *et al*., 2005). RT-PCR transcripts of ORF C4 for DNA-II on the antisense strand and ORF V1 on the sense strand, ORF C2 on the antisense strand and part of GC-rich region for DNA-III were transcribed in all the experimental cassava plants. ORF C4 for DNA-II out of eight ORFs, ORF V1 and ORF C2 for DNA-III out of four ORFs present in the genome structure of episomal DNA-II and DNA-III respectively were amplified. Transcripts obtained were from inoculated cassava plants as indicated in Table 6. Amplification of transcripts from mock-inoculated and begomovirus-

only inoculated cassava plants demonstrates that episomal DNA-II and DNA-III are integrated into the cassava genome, at least in partial form, and additionally that transcription of DNA-II and DNA-III molecules are independent from their associated virus. The transcriptions of the remaining ORFs were not studied. In addition, RT-PCR was performed using cassava germplasm samples from farmer's field and transcripts of the same ORFs were obtained (result not shown here). The results from this study provide evidence that transcription of ORFs is independent of helper virus and environmental conditions.

In order to confirm transcription of episomal ORFs C4 and C2 from DNA-II, and VI from DNA-III, the same experiments need to be performed in another host, such as *N. benthamiana* where no integrated sequences occur. Furthermore, the roles/functions of the transcripts are yet to be determined. However it is hypothesised, based on the ability of episomal DNA-II and III to overcome tolerance/resistance in TME3 when co-inoculated with EACMV-UG2 (which does not infect TME3 alone), and transcribed integrated DNA-II and III sequences, that silencing of homologous cognate transcripts leads to silencing of critical genes involved in begomovirus defence. Transcript sequencing results showed that nucleotide sequences are highly conserved amongst DNA-II and III in inoculated and non-inoculated cassava plants, suggesting that these sequences of DNA-II and DNA-III may have a genomic function in cassava, or begomovirus infection in cassava plants.

In conclusion, this experiment has shown that ORF C4 of DNA-II on the antisense strand and ORF V1 on the sense strand and ORF C2 on the antisense strand of DNA-III, as well as part of

GC-rich region, could be transcribed, and there is a need to conduct more research to understand the biological functions of these conserved sequences.

CHAPTER THREE

GENETIC DIVERSITY AND GEOGRAPHIC DISTRIBUTION OF DNA-II AND III MOLECULES ASSOCIATED WITH CASSAVA AND CASSAVA MOSAIC DISEASE IN UGANDA, KENYA, RWANDA AND TANZANIA

Manuscript submitted to Plant Pathology (modified with additional data from Dr. Ndomba's PhD).

3.1 Abstract

DNA-II and III are small fragments, first detected as episomal ssDNA form associated with cassava mosaic begomoviruses (*East African cassava mosaic virus*-Uganda2 and *African cassava mosaic virus*), and later found during screening of cassava EST databases to be integrated into the cassava genome. Co-inoculation of episomal ssDNA-II and DNA-III with ACMV and EACMV-UG2 led to enhanced symptom severity in cassava cultivar TME3, leading to speculation of the function of these DNA molecules. A study was conducted to determine the genetic diversity and geographic distribution of integrated DNA-II and III molecules in Kenya, Rwanda, Uganda and Tanzania in 2009 and 2010. Young CMD-symptomatic and symptomless cassava plants (3-6 months) were collected in cassava fields across major cassava growing areas in each country. The samples were subjected to DNA extraction and specific PCR primer amplification in order to determine the frequency and genetic diversity of DNA-II/III in different cassava germplasm. Sequence analysis showed that DNA-II and III molecules were integrated in 75.7% and 77.7% of the cassava genotypes, respectively. The distribution of both DNA-II and DNA-III molecules in collected samples was 69.1 % in Uganda, 73.3% in Kenya, 50% in

Rwanda and 69.3% in Tanzania. DNA-II occurred at a frequency of 18.5% in Uganda, 3.3% in Kenya, 6.5% in Rwanda and 5.7% in Tanzania, while DNA-III was 2.5% in Uganda, 8.3% in Kenya, 19.6% in Rwanda and 11.9% in Tanzania. CMD symptoms were more severe (3.2 mean symptom severity score based on 1-5 scoring scale), in Uganda followed by Rwanda (3.0) and Tanzania (2.9) and least severe in Kenya (2.5). DNA-II and DNA-III were detected from cassava plants displaying CMD symptoms of narrow and deformed leaves, and also non-symptomatic cassava plants, and is suspected to play a role in disease phenotype modulation during viral infection. There was little nucleotide sequence divergence (88% - 100%) between the DNA-II and DNA-III sequences: DNA-II sequences had 95-99% nt sequence similarity to the sequences of EACMV-associated episomal satDNA-II (AY836366) and 99% with mentha leaf deformity disease associated satellite DNA II (EU862815), while DNA-III shared sequence similarity of 92-99% with EACMV-associated episomal satDNA-III (AY836367). This study revealed that both DNA-II and DNA-III are widely distributed in the cassava genotypes of almost all the cassava genotypes collected from surveyed areas in Uganda, Kenya, Rwanda and Tanzania.

3.2 Introduction

Cassava (*Manihot esculenta* Crantz) is an important staple food for over 700 million people in Sub Saharan Africa (FAOSTAT, 2011). However, the production of the crop is hugely limited by a number of pests and diseases, among which cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) are currently the most important (Legg *et al.,* 2011).

Cassava mosaic disease is caused by cassava mosaic begomoviruses (CMBs) which belong to the genus *Begomovirus* of the family *Geminiviridae*. The viruses are transmitted by whitefly *Bemisia tabaci* Genn and spread by humans through planting of infected cuttings, which is the usual mode of cassava propagation (Katz *et al.,* 2003; Thresh *et al*., 1994). *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV) are the most commonly occurring CMBs on the sub-Saharan African continent (Pita *et al*., 2001). Cassava mosaic begomovirus (CMB) genome consists of two circular molecules of single-stranded DNA (DNA-A and DNA-B), each *c.* 2.8 kbp (Lazarowitz, 1992). Currently, there are nine reported species of cassava begomoviruses (CBVs) infecting cassava in sub-Saharan Africa countries. These viruses are *African cassava mosaic virus (*ACMV*); East African cassava mosaic virus (*EACMV*), East African cassava mosaic Cameroon virus* (EACMCV) (Fondong *et al.,* 2000), *East African cassava mosaic Kenya virus* (EACMKV) (Bull *et al.,* 2006), *East African cassava mosaic Malawi virus* (EACMMV) (Zhou *et al.,* 1998), *East African cassava mosaic Zanzibar virus* (EACMZV) (Maruthi *et al.,* 2004)), *African cassava mosaic Bukina Faso virus* (ACMBFV) (Tiendrébéogo *et al*., 2012), *South African cassava mosaic* (SACMV) (Berrie *et al.,*1998) and *Cassava mosaic Madagascar virus* **(**CMMGV) (Harimalala *et al.,* 2012). The control of CMD

in East Africa has involved the introduction of improved Tropical Manihot Series (TMS) from the International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria, initially into Uganda and later in Tanzania and other countries.

Despite the achievements of developing CMD resistant cassava varieties through traditional breeding (Hahn *et al*., 1980) and using of several tolerant landraces, such as TME1 and TME3 (Akano *et al*., 2002), these cassava varieties are no longer resistant/tolerant to CMBs. More recently, two episomal circular satellite-like molecules, DNA-II and DNA-III associated with ACMV and EACMV-UG2, were detected in infected cassava plants using universal betasatellite primers (Briddon *et al*., 2002), and were able to break resistance to EACMV-UG2 in the highly resistant landrace TME3 (Ndunguru *et al.,* 2005; 2008; Ingelbrecht *et al.*, 2008). Additionally, integrated DNA-II and III molecules were identified during screening of cassava EST databases (unpublished). Plant virus sequences (19,831) were downloaded from the DPV web plant virus database (Adams and Antoniw, 2006) and aligned to the set of 10577 transcripts in the CIAT-RIKEN cassava (MATI16) full-length cDNA library using BLAST. Two viral sequences sharing sequence similarity with cassava transcripts were detected, namely ACMV-associated DNA-II and DNA-III (unpublished). The range of alignments for DNA-II and DNA-III was 112-346 bp and 65-226 bp, respectively.

It is not known if or how these genomic DNA-II and III molecules in cassava interact with cassava CMBs and cassava brown streak viruses (CBSVs), or whether the episomal forms are transmitted by the whitefly (*Bemisia tabaci*) vector. The exact impact of DNA-II and III on cassava production and their putative role in the current pandemic of cassava mosaic disease are

key priorities, and require further research. This study aimed to establish the genetic diversity and geographic distribution of DNA molecules in Kenya, Rwanda, Uganda and Tanzania.

3.3 Materials and Methods

3.3.1 Sample collection

Samples were collected from the major cassava growing areas in Uganda, Kenya, Rwanda and Tanzania (**Table 9**). Leaf samples were collected from CMD-symptomatic and symptomless cassava plants. In particular, leaves that displayed an unusual narrow (filform) phenotype were collected (**Fig.13A**). The samples were placed in sampling bags, and then stored in cool boxes containing ice cubes until they were delivered to the laboratory for molecular analysis. CMD symptom severity was assessed using a 1 to 5 scoring scale as described by Sseruwagi *et al*. (2004); where 1= symptomless and 5= very severely distorted leaf symptoms. In each field, the geographical location and geo-coordinates (latitude, longitude and altitude) were recorded using a GPS.

Table 9: Cassava major growing areas where survey was conducted in Uganda, Kenya, Rwanda and Tanzania in 2009/2010

Country	District/ Region	Latitude	Longitude	Country	District/ Region	Latitude	Longitude
Uganda	Wakiso	00°10'00"N	32°30'00"E	Rwanda	Kamonye	02°00'00"S	29°54'00"E
	Luwero	01°00'00"N	32°20'00"E		Muhanga	02°05'00"S	29°45'00"E
	Mpigi	00°50'00"N	32°00'00"E		Ruhango	02°13'57"S	29°46'49"E
	Mukono	00°15'00"N	32°55'00"E		Nyanza	02°35'00"S	29°43'00"E
	Kayunga	01°00'00"N	32°52'00"E		Huye	02°35'00"S	29°41'20"E
	Jinja	00°26'20"N	33°12'11"E		Gakenke	01°42'00"S	29°47'00"E
Kenya	Kisumu	00°15'00"S	34°55'00"E	Tanzania	Morogoro	08°00'00"N	37°00'00"E
	Busia	00°25'00"N	34°15'00"E		Dodoma	06°00'00"N	36°00'00"E
	Siaya	00°50'00"S	34°15'00"E		Tabora	05°30'00"N	32°49'00"E
	Nyondo	00°45'00"S	38°12'11"E		Shinyanga	03°45'00"N	33°00'00"E
					Mwanza	02°45'00"N	32°45'00"E
					Kagera	02°00'00"N	31°30'00"E
					Tanga	05°00'00"N	38°15'00"E
					Coastal	07°00'00"S	39°00'00"E
					Lindi	09°30'00"N	38°30'00"E
					Mtwara	10°40'00"N	39°00'00"E
					Songea	11°00'00"N	36°00'00"E
					Iringa	09°30'00"N	35°00'00"E

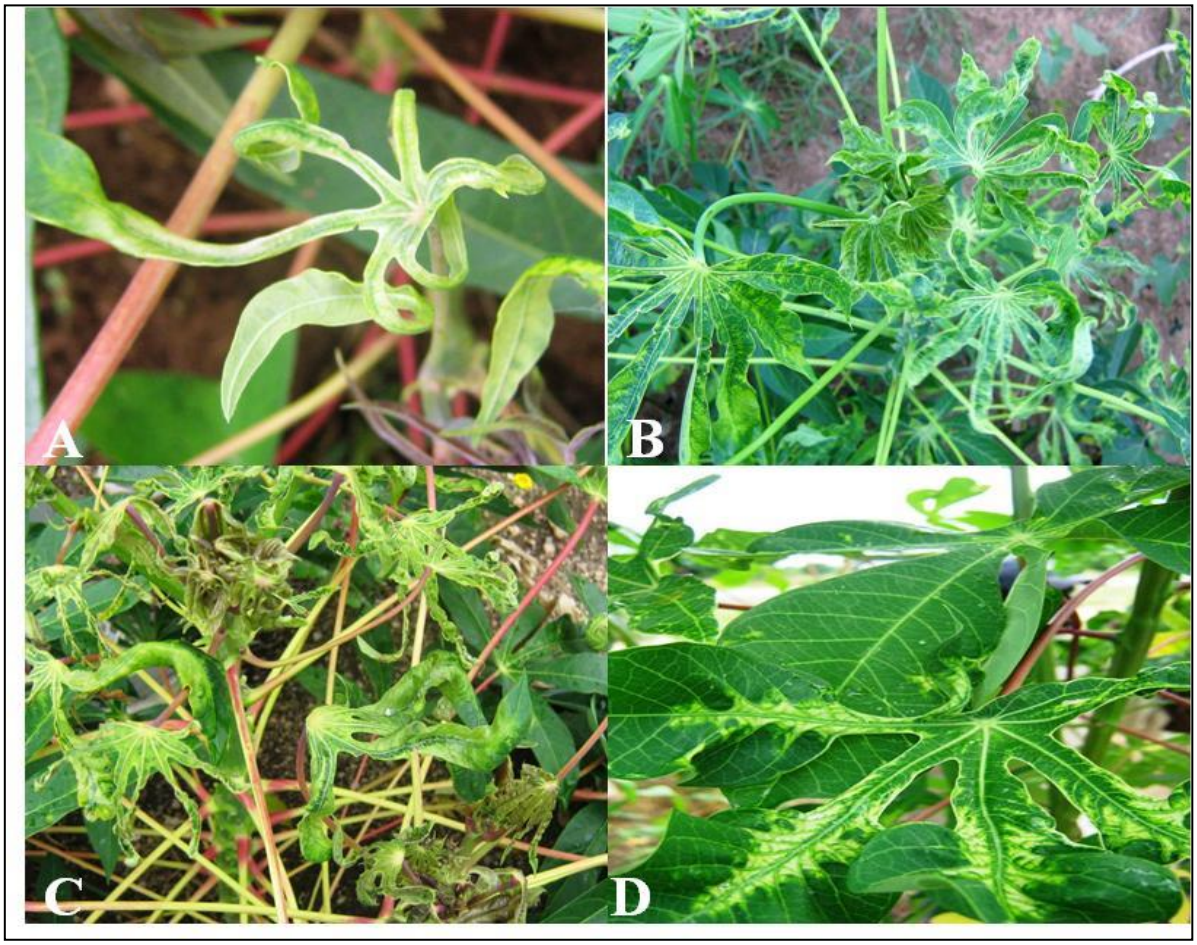


Figure 13: Cassava plants showing cassava mosaic symptoms in cassava fields A; CMD symptoms; B; mosaic severe; C; leaf curling and distortion D; leaf narrowing.

3.3.2 DNA extraction

Total genomic DNA was extracted from young CMD-symptomatic and asymptomatic cassava leaves using a modified Dellaporta *et al* (1983) method. For each leaf sample, 700 mg were ground with a pestle in a mortar containing 500 µl extraction buffer (1M Tris-HCl, 0.5M EDTA and 5M NaCl) with 98% β-mercaptoethanol and the homogeneous mixture was transferred to

clean 1.5 ml eppendorf tubes. In each tube, 40 µl of sodium dodecyl sulphate (SDS) solution was added, mixed and the contents incubated in a 65°C water bath for 30 min. The content was mixed by inverting the tubes after every 10 min. Then 200 µl of 3M sodium acetate solution was added to each tube, mixed, incubated in ice for 20 min and centrifuged at 14,000 rpm for 10 min at room temperature. The supernatant (500 µl) was transferred to a new 1.5 ml tube to which 2 µl of 10 mg/ml RNase A enzyme was added and incubated in water bath at 37°C for 30 min. To precipitate the DNA, 500 µl of cold isopropanol was added to the contents, stored at -20°C for 30 min, followed by a 10 min centrifugation to precipitate the DNA. After washing with 70% ethanol, the DNA were air-dried for 30 min, re-suspended in 100 µl of sterile double distilled water and stored at-20°C.

3.3.3 Polymerase chain reaction (PCR)

PCR amplification for integrated DNA sequences was performed using specific primers: satIIF 5'-GCCGCACCACTGGATCTC-3' forward primer, satIIR 5'-CAGCAGCCAGTCAGGAAGTT-3' reverse primer for DNA-II and satIIIF 5'-AGGCCTCGTTACTAAAAGTGC-3' forward primer, satIIIR 5' ACCTGACGGCAGAAGGAAT-3' reverse primer for DNA-III (German Plata CIAT, Colombia). PCR was performed using a Thermorcycler (GeneAmp^{R} PCR System 9700) with the first cycle at 94°C for 1 min followed by 30 cycles at 94°C for 1 min, 55°C for 1 min, 72°C for 2 min and a final extension at 72°C for 10 min. A total reaction mixture of 50 µl made up of 41.75 µl of sterile distilled water, 2.5 µl PCR buffer (10x), 1.5 µl MgCl2 (25 mM), 0.5 µl dNTPs (2.5 mM), 1 µl each of the forward (10 mM) and reverse primer (10mM), 0.25 µl Taq

polymerase and 2.0 µl of the DNA template was used. To detect the PCR products, electrophoresis was performed at 100 volts in a 1% agarose gel buffered with 1X TAE (Savant HG330 electrophoretic gel system), and stained with 6x Orange DNA loading dye. The gels were run for 1 hour, and visualized using the bench top UV transilluminator, and photographed with a gel documentation system (BioDoc-It^{TM} Imaging System).

3.3.4 Molecular cloning

Amplicons of integrated DNA-II and DNA-III were purified from the agarose gels using a QIAgen Gel Extraction Kit (Invitro, Life Science, USA) as per the manufacturer's protocol. In each country, at least five positive amplicons each of DNA-II and DNA-III were ligated into the pGEM^{R}-T Easy Vector using a TA Cloning Kit (Promega, USA) following the manufacturer's guidelines, and transformed into chemically highly competent cell *Escherichia coli* strain JM109 (Promega, USA). Transformed bacteria colonies were cultured in Luria Bertani broth containing appropriate antibiotics and incubated at 37°C overnight while shaking at 200 rpm. Plasmids were purified by using QIAgen Minprep Purification Kit (Promega, USA) and the purified plasmids were sequenced using universal primers (M13F/R) at BeCA-ILRI- Nairobi, Kenya.

3.3.5 Sequences analysis

DNA sequences of integrated DNA-II and DNA-III molecules were assembled from contiguous sequences with SEQMAN program (DNASTAR). Putative Open Reading Frames (ORFs) were identified with a FramePlot software program. To determine whether the molecules were similar

to published DNAs sequences, they were initially aligned, using BLASTN, to known sequences in the National Center for Biotechnology Information (NCBI). ORFs were translated to protein and BLASTed (http://www.ncbi.nlm.nih.gov/BLAST/) to sequences in Genbank to determine whether the proteins were similar to known satellite proteins, for example CAA65841, CBJ19308, CAJ85966 and YP-184756. Sequences were aligned with the MEGA software, (ClustalW) (Thompson *et al*., 1997) and phylogenetic trees constructed using the neighbour-joining algorithm.

3.3.6 Statistical data analysis and presentation

Geo-coordinates were entered into excel and used to generate disease distribution maps for the data collected in the surveyed areas. Significance of the DNA-II and III occurrence and distribution data was analysed using a one-way analysis of variance (ANOVA) and means separated using the least significance differences (LSD) at 5%.

3.4 Results

3.4.1 Cassava mosaic disease symptom expression in the field

Cassava mosaic disease symptoms varied widely among the infected plants in cassava fields. The symptoms observed were mosaic, severe mosaic, leaf narrowing, leaf curling, leaf distortion and leaf deformation (**Fig.13A, B, and C &D**).

Analysis of variance (ANOVA) indicated that, symptom severity scores differed significantly ($P<0.011$, $df = 3$, $F = 3.76$) between Tanzania, Uganda, Rwanda and Kenya where the surveys were conducted. CMD severity averaged 2.9 and was highest in Uganda (3.2) and lowest in Kenya (2.5) (**Table 10**).

Table 10: Analysis of Variance for symptom severity on cassava samples collected from Uganda, Kenya, Rwanda and Tanzania in 2009-2010

Country	Mean symptom severity (scale 1-5)
Uganda	3.2
Kenya	2.5
Rwanda	3.0
Tanzania	2.9
Mean	2.9
P-value (5%)	0.011
F-statistics	3.76

3.4.2 PCR analysis of integrated DNA-II and DNA-III molecules

Partial fragments of DNA-II (895bp) and DNA-III (306bp) molecules were amplified (**Fig. 14A & B**) in 275 (75.8%) and 282 (77.7%), respectively, of the 363 total samples collected. Of The 81 samples in Uganda, 15 (18%) were positive for DNA-II only, 2 (2.5%) for DNA-III only, and 56 (69.1%) for both DNA-II and DNA-III. In Kenya, of 60 samples collected, 2 (3.3%) had

DNA-II only, 5 (8.3%) had DNA-III only, and 44 (73.3%) were positive for both DNA-II and DNA-III. Of the 46 samples collected from Rwanda, 3 (6.5%) had DNA-II, 9 (19.6%) had DNA-III, and 23 (50%) were positive for both DNA-II and DNA-III. Of the 176 samples from Tanzania, 10 (5.7%) samples had DNA-II only, 21 (11.9%) had DNA-III only, and 122 (69.3%) confirmed the presence of both DNA-II and DNA-III molecules. Combined PCR results for all samples collected in the surveys showed that 67% (245/363) of the samples had with both integrated DNA-II and DNA-III molecules. Only 8% (30/363) and 10% (37/363) had either integrated DNA-II or DNA-III, respectively, and 14% (51/363) did not yield any amplification.

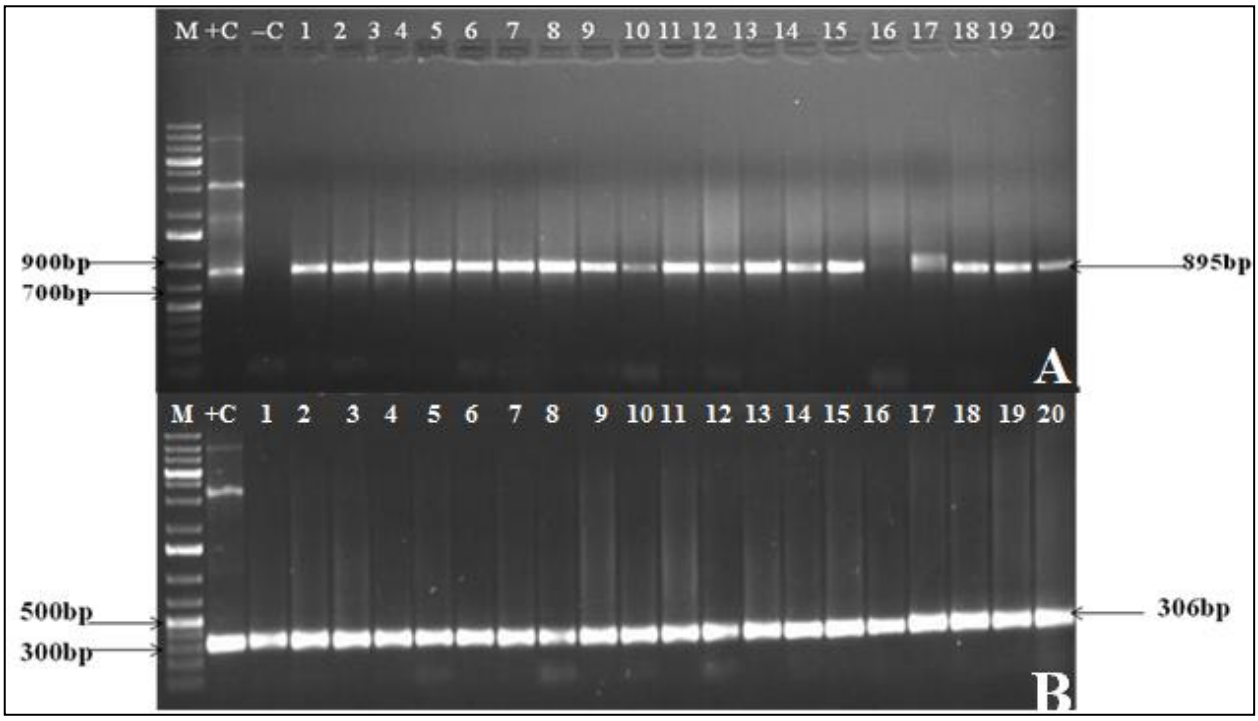


Figure 14: PCR amplification of integrated DNA-II molecules (16A) and DNA-III (16B) on 1% agarose gel. M=Molecular weight marker of 1 kb plus; +C= positive control and -C = negative control. Lanes 1-20 (16A &16B) represent samples collected from farmers' field.

3.4.3 Occurrence and geographical distribution of integrated DNA molecules

Generally, DNA-II and III molecules occurred at a high frequency in major cassava growing areas from surveyed regions (**Table 9**). Percentage frequency of DNA molecules in the surveyed areas was 73% in Uganda, 85% in Kenya, 76% in Rwanda and 86% in Tanzania, based on the total number of samples collected in each country. Dual infection of DNA-II and DNA-III was detected more frequently in the major cassava growing areas compared with single DNA-II or DNA-III (**Fig 15**).

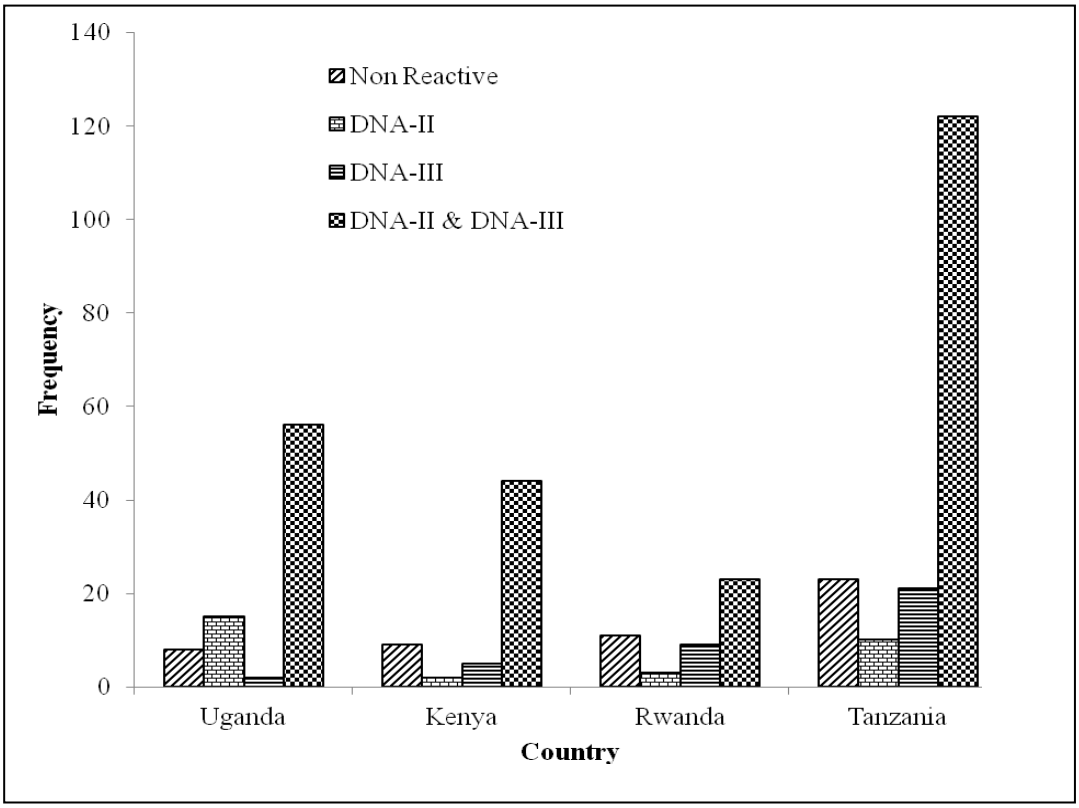


Figure 15: Occurrence of integrated DNA molecules in Uganda, Kenya, Rwanda and Tanzania in 2009/2010.

There was no apparent defined pattern in terms of geographical distribution of the integrated DNA molecules among the countries and within the country, since the integrated DNA molecules were frequently distributed in all the surveyed areas, either singly or as dual infection (**Fig. 16**). The DNA molecules were distributed throughout the surveyed areas in major cassava growing areas in Lake Victoria, Southern zone and Coastal zone in Tanzania, Central areas in Uganda, Northern and Southern provinces in Rwanda and Western province in Kenya.

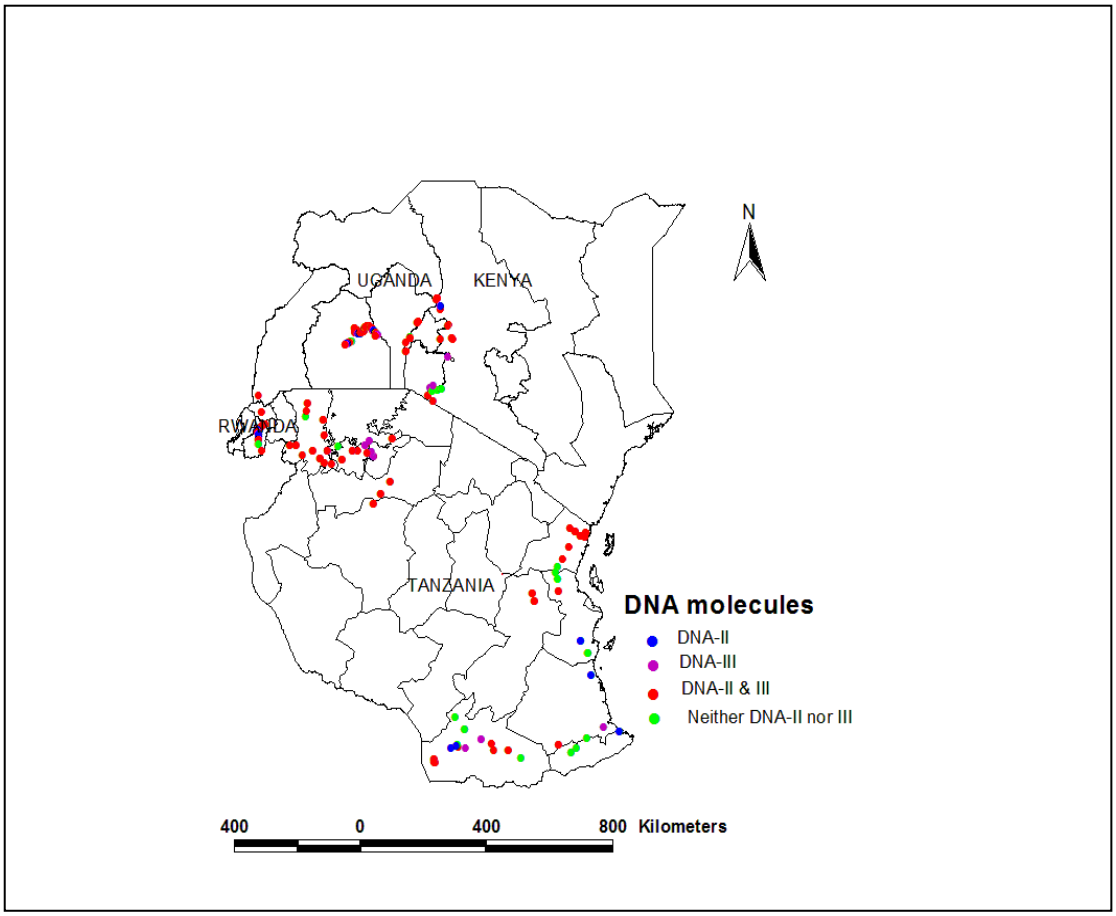


Figure 16: Geographical distribution of DNA-II and DNA-III molecules in Uganda, Kenya, Rwanda and Tanzania in 2009/2010

3.4.4 Phylogenetic analysis and sequence comparisons

DNA-II and DNA-III sequences were aligned and gaps were removed. After alignment, phylogenetic trees were generated in which DNA-II sequences grouped into a single cluster (**Fig. 17A**), while DNA-III sequences formed two clusters (**Fig. 17 B**). Grouping of the isolates was random within the two clusters indicating that there was no specific definite geographical distribution of genotypes. Analyzed sequences for DNA-II molecules obtained from Uganda, Kenya, Rwanda and Tanzania revealed an average integrated sequence length of 841 nucleotides (nt). The variation in length within the sequences obtained was due to deletion of 54 nt in expected sequences of 895 bp within DNA-II molecules (unpublished). Pair-wise comparison within the analysed sequences showed 88% - 100% sequence identity to each other. All DNA-II sequences analysed had GC region typical of episomal DNA-II in the database (AY836366). Nucleotide sequences derived from these sequences were compared with GeneBank sequences, and were highly similar (95-99%) to the published episomal DNA sequence accession number (AY836366) (Ndunguru *et al*, 2008) and 99% to Mentha leaf deformity associated DNA-II (Accession No. EU862815) (Borah *et al*, 2010) associated with Menthal leaf deformity virus.

Analyzed sequences for DNA-III revealed a sequence length that ranged from 290-314 nt. Within analysed sequences for DNA-III, sequences obtained from Rwanda had an average length of 290 nt, while those of Kenya, Uganda and Tanzania ranged between 306 and 314 nt. The variation in length within DNA-III sequences analysed was due to a deletion of 16 nt (CTCAACTCTCCATGCC), in the Rwandan DNA-III isolates, in the expected sequence (306 bp) amplified in earlier studies (unpublished). In contrast the Kenyan, Ugandan and Tanzanian

DNA-III sequences had a longer length of 314 nt due to insertion of 8 nt. All analysed DNA-III sequences had GC region typical of episomal DNA-III in the published sequence (AY836367). The sequences were blasted against the GeneBank sequence and showed to be highly similar (92-99%) to episomal DNA-III (Ndunguru *et al*, 2008).

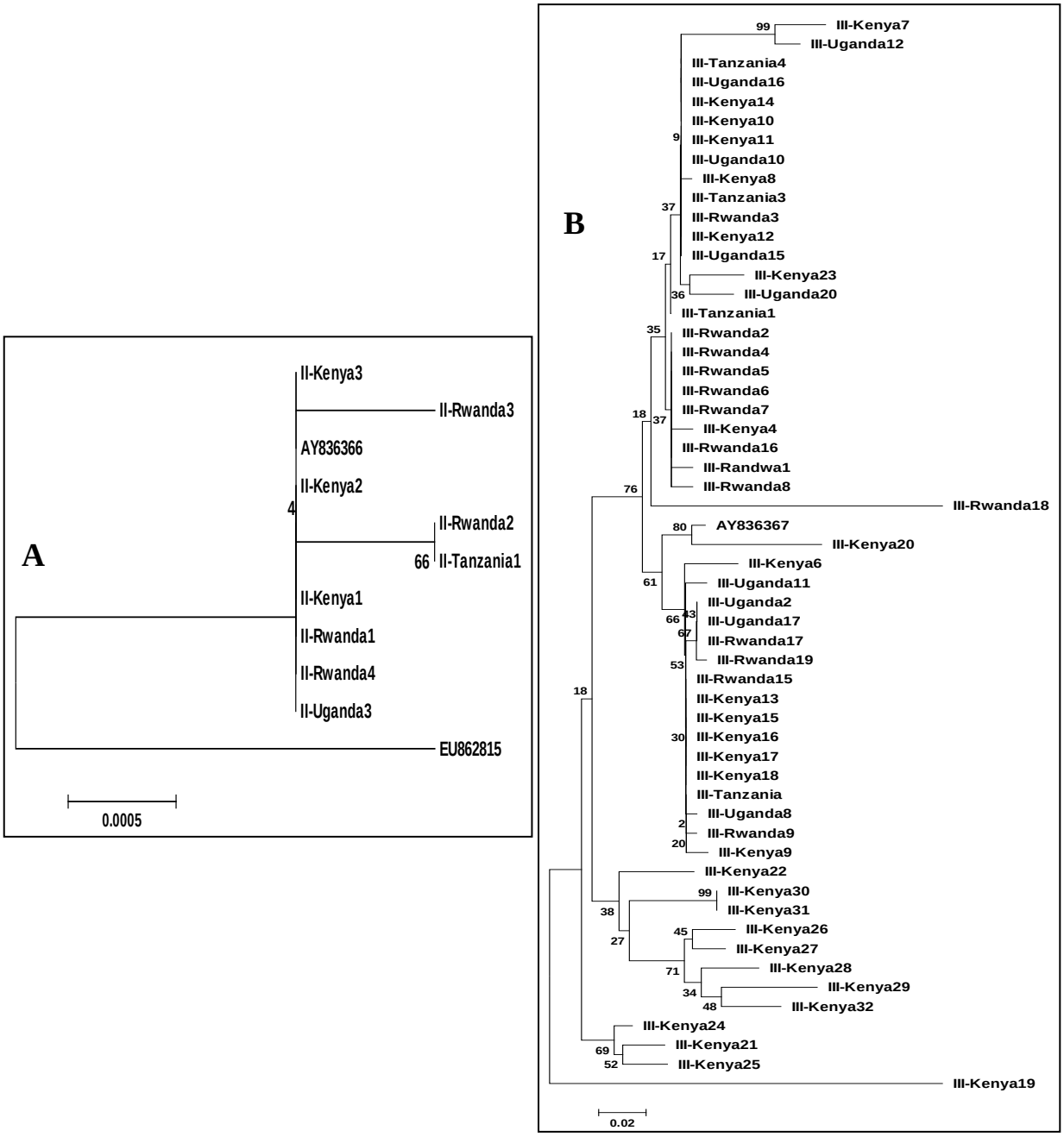


Figure 17: Phylogenetic trees (using the neighbour-joining algorithm) obtained from alignment of the nucleotide sequences of DNA-molecules with available sequence of satDNA-II and satDNA-III. 19A: DNA-II sequences; 19B: DNA-III sequences obtained from Uganda, Kenya,

Rwanda and Tanzania in 2009/2010, obtained from Genbank with accession number AY836366, AY836367 and EU862815.

3.5 Discussion

There is no prior published information from the literature on DNA-II or DNA-III-like molecules (associated with ACMV or EACMV) integrated into the cassava, or any other plant genome. Other DNA molecules associated with symptom modulation of begomovirus infections have been reported in the literature. Amongst these are defective interfering molecules, associated with helper begomoviruses in cassava in South Africa, and Tanzania, and which are derived from helper virus DNA A or B genome (Choge *et al*., 2001; Paximadis *et al*, 2001; Ndunguru *et al*., 2006; Abraham *et al*., 2012). Additionally, betasatellites have been reported in several begomovirus diseases in the Old World, and are known to alter symptom phenotypes in their respective hosts (Briddon and Stanley, 2006). Satellite DNAs or sub-viral agents are dependent on their cognate helper begomovirus for replication, and can exacerbate or reduce symptoms. Some examples include ToLCV-sat components associated with monopartite begomovirus ToLCV in tomato plants (Dry *et al*., 1997), DNA-β associated with AYVV in ageratum weed (Saunder *et al*., 2000), CLCuV in cotton plants (Briddon *et al*., 2000, 2001) and DNA-1 components isolated from a range of plant species growing in different regions throughout Africa and Asia, for example DNA-1 isolated from CLCuD (Briddon *et al*., 2004; Mansoor *et al*., 2001; Mansoor *et al*., 200a; Mansoor *et al*., 200b; Wu and Zhou, 2005).

This study reports for the first time the presence of DNA-II and DNA-III molecules integrated into the cassava genome of CMD-infected and non- infected cassava plants. While the nature of

these integrated DNA fragments are not known, they share high sequence identity with two episomal forms amplified by betasatellite universal primers from ACMV/EACMV-infected cassava in Tanzania (Ndunguru, 2005). These so- called episomal circular ssDNA satellite-like DNA-II and III amplicons were found to exacerbate symptom severity in cassava, and to reduce tolerance in TME3, when co-inoculated with EACMV-UG2 (Ndunguru *et al*., 2005, 2008). Mentha leaf deformity is associated with a geminivirus-satellite complex of cotton leaf curl Multan betasatellite (CLCuMB) (EU862816) and *Tomato leaf curl Karnataka virus* (ToLCKV) infecting mint plants in the Punjab in India (Borah *et al*., 2012).

In addition to the betasatellite, the betasatellite primers amplified a 1019 nucleotide product (EU862815). Surprisingly the Mentha leaf deformity associated DNA-II (MLDA-DNA-II) showed 98% sequence identity with DNA-II (AY836366) reported from cassava associated with EACMV-UG2 from cassava in Tanzania. Neither MLDA-DNA-II or cassava-associated DNA-II appear to contain any known betasatellite ORFs, such as β–C1, although putative ORFs have been postulated. It is interesting to note that both cassava and mentha display severe leaf deformity, and it would appear the episomal DNA-II plays a role in begomovirus symptom phenotype severity. However the exact nature, origin and function of the integrated fragments is not known. As with cassava, MLDA-DNA-II accumulated in the majority but not all symptomatic mentha plants. DNA-III related sequences have not been reported in any other plants in the Genebank database. MLDA-DNA-II could not be transmitted to *Nicotiana benthamiana* by mechanical inoculation (Borah *et al*., 2012), but infectious clones of cassava-associated DNA-II and III were able to infect cassava and *N.benthamiana* by agro-inoculation and biolistic bombardment along with helper CMB (Ndunguru *et al*., 2005; 2008).

Further work is necessary to determine the role of these satellite-like and integrated DNA agents. Since smaller fragments of cassava-associated DNA-II and III were first uncovered from EST databases, and now have been mapped to scaffolds of the cassava genome (unpublished), it is certain that there is some biological function. However what, if any, roles these DNAs may have in gene function/regulation or begomovirus-host interaction, needs to be elucidated.

CHAPTER FOUR

SUMMARY AND RECOMMENDATIONS

Some plant viruses are associated with satellites that are composed of nucleic acid and depend on co-infection of host cell with a helper virus for their multiplication. Many of them modulate disease symptoms and increase viral accumulation (Roossinck *et al*., 1992; Simon *et al*., 2004). The earliest satellites were associated with RNA viruses, but recently satellites associated with DNA viruses, such as geminiviruses, have been discovered. These satellites include TLCV Sat-DNA, betasatellites and alphasatellites which were isolated from monopartite begomoviruses and satellite-like DNA molecules isolated from bipartite begomoviruses. Successful interaction between betasatellites and the New Word (NW) bipartite begomovirus *Cabbage leaf curl virus* (CbLCuV) has been reported by Nawaz-ul-Rehman *et al*. (2009). 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 massive economic losses on crop yields through exacerbated disease symptoms.

Cassava mosaic disease (CMD) is the most important disease reducing cassava productivity in sub-Saharan Africa. The disease is caused by at least nine species of begomovirus that occur singly and in co-infection (Fauquet *et al.,* 2008). Despite the efforts to develop CMD resistant cassava varieties through convectional breeding (Hahn *et al*., 1980), and use of several tolerant cassava landraces such as TME1 and TME3 (Akano *et al*., 2002), these cassava varieties are no longer tolerant/resistant to CMBs. More recently, two circular single stranded DNA molecules,

DNA-II and DNA-III associated with cassava mosaic begomoviruses were detected in infected cassava plants using universal betasatellite primers (Briddon *et al*., 2002), and were able to enhance CMD symptoms and break resistance to EACMV-UG2 in the highly resistant landrace TME3 (Ndunguru *et al.*, 2005; 2008; Ingelbrecht *et al.*, 2008). Additionally, DNA-II and III molecules were found integrated in the cassava genome of both domesticated and wild *Manihot* species during screening of cassava EST databases (unpublished). It has been demonstrated that DNA-II and DNA-III molecules integrate into the host genome while the previously reported satellites were not known to be integrated into genome.

In this study primer pairs satII-F/R and satIII-F/R were able to amplify the expected size for DNA-II and DNA-III molecules integrated in the cassava genome respectively. DNA-II and DNA-III molecules were identified from cassava leaf samples collected from healthy and CMD-infected cassava plants that was confirmed by sequencing. A total of 363 cassava genotype were collected from Kenya, Uganda, Rwanda and Tanzania countries and were screened for integrated DNA-II and DNA-III and were found 86% of the samples .This showed that DNA-II and DNA-III integrated in the cassava genome are widely distributed into the major cassava growing areas.

The other achievement made in this study was the successful biolistic inoculation of EACMV-UG2, ACMV, DNA-II and DNA-III infectious clones to cv.60444 and TME3. Cassava leaf samples collected from inoculated TME3 and cv. 60444 from screen house were used for determination of transcription of ORFs for DNA-II and DNA-III. RT- PCR using primer pairs designed from genome structure of satellite-like DNA-II and DNA-III were performed. SatII-2F/3R and SatIII-7F/9R primers detected the expected sizes for transcript ORF C4 and ORFV1

for DNA-II in antisense strand and DNA-III in sense strand, respectively. This suggests that at least part of DNA-II and DNA-III can be transcribed.

In conclusion this study has generated useful information on the genetic diversity and geographical distribution of the DNA-II and DNA-III molecules from Kenya, Rwanda, Uganda and Tanzania and transcripts analysis for ORF C4 and ORF V1 in antisense strand of DNA-II and sense strand of DNA-III respectively. Information generated can be helpful in understanding the biological function of integrated DNA molecules and if the transcribed ORF can produce functional protein. Furthermore, the study findings will add to the body of knowledge on the DNA II and III and help researchers (plant virologists and breeders) working on cassava in Sub Saharan Africa (SSA) to develop durable management strategies for CMD.

Recommendations

The present study has opened up more areas for further research in order to understand the role of DNA-II and DNA-III molecules in the etiology and pandemic of CMD and which contribute to the loss of cassava productivity in Uganda, Kenya, Rwanda and Tanzania when associated with cassava mosaic begomoviruses. The areas for further research include:

1. The exact impact of DNA-II and III on cassava production and roles in gene function/regulation or begomovirus-host interaction needs to be elucidated.

2. How these genomic DNA-II and III molecules in cassava interact with cassava CMBs or whether the episomal forms are transmitted by the whitefly (*Bemisia tabaci*) vector or co-encapsidated.

3. Further study on the transcribed ORFs if they produce functional proteins, and design primers for the remaining ORFs for DNA-II and DNA-III.

References

Alberter, M., Rezaian, A., Jeske, H., 2005. Replicative intermediates of *Tomato leaf curl virus* and its satellite DNAs. Virol. 331, 441-448.

Alabi, O.J., Kumar, P.L., Naidu, R.A., 2011. Cassava mosaic disease: A curse to food security in Sub-Saharan Africa. Online. APS*net* Features. doi:10.1094/APSnetFeature-0701.

Bender, A., Bender, A.E., 2005. Dictionary of Food and Nutrition, Oxford University Press, 1995, second edition.

Behjatnia, S.A.A., Dry, I.B., Rezaian, M.A., 1998. Identification of the replication-associated protein binding domain within the intergenic region of tomato leaf curl geminivirus. Nucleic Acids Res. 26, 925-931.

Berrie, L.C., Palmer, K.E., Rybicki, E.P., Rey, M.E.C., 1998. Molecular characterisation of a distinct South African cassava infecting geminivirus. Arch. Virol., 143, 2253-2260.

Bethke, J.A., Byrne, F.J., Hodges, G.S., McKenzie, C.L., Shatters, R.G., 2009. First record of the Q biotype of the sweetpotato whitefly, *Bemisia tabaci*, in Guatemala *Phytoparasitica* 37, 61-64.

Bisimwa, E., Walangululu, J., Bragard, C., 2012. Occurrence and Distribution of Cassava Mosaic Begomovirus Related to Agro-ecosystems in the Sud-kivu Province, Democratic Republic of Congo. Asian J. Plant Pathol. 6, 1-12.

Bull, S.E., R.W. Briddon, W.S. Sserubombwe, K. Ngugi, P.G. Markham, Stanley, J., 2006. Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. J. Gen. Virol. 87, 3053-3065.

Burns, A., Gleadow, R., Cliff, J., Zacarias, A., Cavagnaro, T., 2010. Cassava: The drought, war and famine crop in a changing world. Sustain. 2, 3572-3607.

Briddon, R.W., Stanley, J., 2006. Subviral agents associated with plant single-stranded DNA viruses. Virol. 344, 198-210, Oxford University.

Briddon, R.W., Markham, P.G., 2000 .Cotton leaf curl virus disease. Virus Res. 71, 151-159.

Briddon, R.W., Mansoor, S., Bedford, I.D., Pinner. M.S., Saunders, K. Stanley, J.Y. Zafar, Y, Malik, K. A., Markham, P.G., 2001. Identification of DNA components required for induction of cotton leaf curl disease. Virol. 285, 234-243.

Briddon, R.W., Bull, S.E., Mansoor, S., Amin, I., Markham, P.G., 2002.Universal primers for the PCR-mediated amplification of DNA-β: a molecule associated with some monopartite begomoviruses. Mol. Bio-technol. 20, 315-318.

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., Markham, P.G., 2003. Diversity of DNA β: a satellite molecule associated with some monopartite begomoviruses. Virol. 312, 106-121. Article

Briddon, R.W., Bull, S.E., Amin, I., Mansoor, S., Bedford, I.D., Rishi, N., Siwatch, S.S., Zafar, Y., Abedel-Salam, A.M., Markham, P.G., 2004. Diversity of DNA1: a satellite-like molecule associated with monopartite begomovirus-DNAB complexes. Virol. 324, 462-474.

Briddon, R.W., Brown, J., Moriones, E., Stanley, J., Zerbini, M., Zhou X., Fauquet, C., 2008. Recommendations for the classification and nomenclature of the DNA-β satellites of begomoviruses. Arch. Virol.153, 763-781.

Briddon, R.W., Markham, P.G., 1995. Family Geminiviridae; In Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses, 158-165. Edited by Murphy, F.A., Fauquet, C. M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., Summers., M. D. Vienna and New York: Springer-Verlag.

Brown, J.K., Frohlich, D.R., Rosell, R.C., 1995. The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? Annu. Rev. Entomol. 40, 511-534.

Brown, J.K., 2010. Phylogenetic biology of the *Bemisia tabaci* sibling species group. 619, 31-67 in: Bionomics and Management of a Global Pest. P.A. Stansly and S.E. Naranjo (eds.), Springer, Amsterdam. 350.

Bull, S.E., Tsai, W.S., Briddon, R.W., Markham, P.G., Stanley, J., Green, S.K., 2004. Diversity of begomovirus DNA β satellites of Non-malvaceous plants in east and south east Asia, Arch. Virol. 149, 1193-1200.

Carabali, A., Belloti, A.M., Mantoya-Lerma, J., Cuellar, M.E., 2005 Adaptation of *Bemisia tabaci* biotype (Gennadius) to cassava, Manihot esculenta (Crantz).Crop Prot.24, 643-649.

Colvin, J., Omongo, C.A., Maruthi, M.N., Otim-Nape, G.W., Thresh, J.M., 2004. Dual begomovirus infections and high Bemisia tabaci populations: two factors driving the spread of a cassava mosaic disease pandemic. Plant Pathol. 53, 577-584.

Cui, X., Xie, Y., Fauquet, C.M., Zhou, X., 2004b. A DNA β associated with Tomato yellow leaf curl China virus is required for symptom induction, J. Virol. 78, 13966-13974.

Cui, X., Wang, G. Li., Hu, D., Zhou, X., 2005. A begomovirus DNAβ-encoded protein binds DNA, functions as a suppressor of RNA silencing, and targets the cell nucleus, J. Virol.79, 10764-10775.

Chu, D., Zhang, Y.J., Brown, J.K., Cong, B., Xu, B.Y., Wu, Q.J., Zhu, G.R., 2006. The introduction of the exotic Q biotype of *Bemisia tabaci* from the Mediterranean region into China on ornamental crops. Florida Ento. 89, 168f Virol. 73, 5481-9.

Czechowski, T., Bari, R.P., Stitt, M., Scheible, W.R., Udvardi, M.K., 2004. Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: Unprecedented sensitivity reveals novel root- and shoot-specific genes. Plant J. 38, 366–379.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., Scheible, W.R., 2005. Genome wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. 139, 5–17.

De La Rua, P., Simon, B., Cifuentes, D., Martinez-Mora, C., Cenis, J.L., 2006. New insights into the mitochondrial phylogeny of the whitefly *Bemisia tabaci* (Hemiptera Aleyrodidae) in Mediterranean Basin. J. Zool. Syste. Evol. Res. 44, 25-33.

Dennehy, T.J., Degain, B.A, Harpold, V.S, Zaborac, M., Morin, S, Fabrick, J.A, Nichols, R.L., Brown, J.K., Byrne, F.J., Li, X., 2010. Extraordinary resistance to insecticides reveals exotic Q biotype of *Bemisia tabaci* (Gennadius) in the New World. J. Eco Entomol. 103, 2174-2186.

Dennehy, T.J., Degain, B.A., Harpold, V.S., Brown, J.K., Byrne, F.J., Morin, S., Nichols, R.L., 2006. First new world report of Q biotype of *Bemisia tabaci* (Gennadius) reveals high levels of resistance to insecticides. Resistance Pest ManagementNewsletter. 5, 18-19.

Dennehy, T.J., Degain, B.A., Harpold, V.S., Brown, J.K., Morin, S., Fabrick, J.A., Byrne, F.J., Nichols, R.L., 2005. New challenges to management of whitefly resistance to insecticides in Arizona. The University of Arizona-Cooperative Extension Report, Tucson, AZ.

Dry, L.R., Krake, J., Rigden, E., Rezaian, M.A., 1997. A novel subviral agent associated with a geminivirus: the first report of a DNA satellite, Proc. Natl. Acad. Sci. U.S.A. 94, 7088-7093.

Dubern, J., 1994. Transmission of African cassava mosaic geminivirus by the whitefly (*Bemisia tabaci*) Trop. Sci. 34, 82-91.

Elsharkawy, M.A., 2004. Cassava biology and physiology. Plant mol. Biology: 56, 418-501.

Esterhuizen, L.L., Mabasa, K.G., Van Heerden, S.W., Czosnek, H., Brown, J.K., Van Heerden, H., 2012. Genetic identification of members of the *Bemisia tabaci* cryptic species complex from South Africa reveals native and introduced haplotypes. J. Appl. Entomol.136:1-14.

FAO., 2006. Cassava production statistics http:/www.fao.org

Fauquet, C., Fargette, D., 1990. African cassava mosaic virus: Etiology, epidemiology, and control. Plant Dis. 74, 404-11 Retrieved October 23, 2008.

Fauquet, C.M., Briddon, R.W., Brown, J.K., Moriones, E., Stanley, J., Zerbini, M., Zhou, X., 2008. Geminivirus strain demarcation and nomenclature. Arch. Virol. 153, 783-821.

Fauquet, C.M., Sawyer, S., Idris, A.M., Brown, J.K., 2005. Sequence analysis and classification of apparent recombinant begomoviruses infecting tomato in the Nile and Mediterranean basins. Phytopathol. 95, 549-555.

Fontes, E.P.B., Eagle, P.A., Sipe, P.A., Luckow, V.A., Hanley-Bowdoin, L., 1994a. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication, J. Biol. Chem. 269, 8459-8465.

Fontes, E.P.B., Gladfelter, H.J., Schaffer, R.L., Petty, I.T.D., Hanley-Bowdoin, L., 1994b. Geminivirus replication origins have a modular organization, Plant Cell. 6, 405-416.

Haggblade, S., Zulu, B., 2003. The recent cassava surge in Zambia and Malawi. In InWent, IFPRI, NEPAD, CTA Conference Success in African Agriculture. Pretoria, South Africa.

Hahn, S.K., Howland A.K., Terry, E.R., 1980a. Correlated resistance of cassava to mosaic and bacterial blight diseases. Euphytic. 29, 305-311.

Harimalala. M, Lefeuvre, P.A., De Bruyn, F., Tiendrébéogo, M., Hoareau, J., Villemot, S., Ranomenjanahary, A., Andrianjaka, B., Reynaud, Lett, J.M., 2012. A novel cassava-infecting begomovirus from Madagascar: *Cassava mosaic Madagascar virus*. Arch. Virol. 157, 2027-2030.

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

Harrison, Y.L., Liu, S., Khalid, S., Hameed, G.W., Otim-Nape., Robinson, D.J., 1997. Detection and relationships of cotton leaf curl virus and allied whitefly-transmitted geminiviruses occurring in Pakistan. Ann. Appl. Biol. 130, 61-75.

Hillocks, R. J., 2002. Cassava in Africa, In: Hillocks, R.J, Thresh, J.M, Bellolli, A.C, eds. Cassava: Biology, Production and Utilization, CAB International, Wallingford. 41-54.

Hong, Y.G., Robinson, D.J., Harrison, B.D., 1993. Nucleotide sequence evidence for the occurrence of three distinct whitefly- CJG Synergism between ACMV and EACMV transmitted geminiviruses in cassava. J. General. Virol. 74, 2437-2443.

Horowitz, A.R., Kontsedalov, S., Khasdan, V., Ishaaya, I., 2005. Biotypes B and Q of *Bemisia tabaci* and their relevance to neonicotinoid and pyriproxyfen resistance. Arch. Insect Bioch. Phy. 58, 216-225.

Ilyina T.V., Koonin, E.V., 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria, Nucleic Acids Res. 20, 3279-3285.

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

Jeske, H., Lutgemeier, M., Preiss,W., 2001. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus, EMBO J**.** 20, 6158-6167.

Katz, S.H., Weaver, W.W., 2003. Encyclopedia of Food and Culture. New York: Schribner. ISBN 0684805685

Lazarowitz, S.G., 1992. Geminiviruses: genome structure and gene function. Critical Reviews in Plant Scie. 11, 327-49.

Legg, J.P., Fauquet, C.M., 2004. Cassava mosaic geminiviruses in Africa, Plant Mol. Biol. 56, 585-599.

Legg, J.P., French, R., Rogan, D., Okao-Okuja, G., Brown, J.K., 2002. A distinct Bemisia tabaci (Gennadius) (Hemiptera: Sternorrhyncha: Aleyrodidae) genotype cluster is associated with the epidemic of severe cassava mosaic virus disease in Uganda. Mol Ecol. 11, 1219-1229.

Legg, J.P., Hillocks, R.J. Eds., 2003. Cassava brown streak virus disease: past, present and future. Proceedings of an International Workshop, Mombasa, Kenya, October 27-30, 2002. Natural Resources International Limited. Aylesford, UK. 100pp.

Legg, J.P., Ndjelassili, F., Okao-Okuja, G., 2004. First report of cassava mosaic disease and cassava mosaic geminiviruses in Gabon. Plant Pathol. 53, 232.

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

Legg, J.P., Thresh, J.M., 2003. Cassava virus diseases in Africa. In: Proceedings of the First International Conference on Plant Virology in Sub-Saharan Africa (4-8 June 2001, Ibadan, Nigeria), IITA, Ibadan, Nigeria, pp. 517-522.

Lin, A.A., Behjatnia, I.B. Dry, J.W., Randles, M.A. Rezaian., 2003. High-affinity Rep-binding is not required for the replication of a geminivirus DNA and its satellite. Virol. 305, 353-363.

Lisha, V.S., Antony, B., Palaniswami, M.S., Henneberry, T.J., 2003. *Bemisia tabaci* (Homoptera: Aleyrodidae) biotype in India. J. Econo. Entomol. 96, 322-327.

Mansoor, S., Amin, I., Hussain, M., Zafar, Y., Bull, S., Briddon. R.W., Markham, P.G., 2001. Association of a disease complex involving a begomovirus, DNA 1 and a distinct DNA beta with leaf curl disease of okra in Pakistan. Plant Dis. 85, 922.

Mansoor, S., Khan, S.H, M. Hussain, M., Zafar, Y., Pinner, M.S., Briddon, R.W., Stanley. J., Markham, P.M., 2000a. Association of a begomovirus and nanovirus-like molecule with ageratum yellow vein disease in Pakistan. Plant Dis. 84, 101.

Mansoor, S., Mukhtar, S., Hussain, S., Amin, M., Zafar, I., Malik, K., Markham, P.G., 2000b.Widespread occurrence of Cotton leaf curl virus on radish in Pakistan. Plant Dis. 84, 809.

Mansoor, S.H., Khan, A., Bashir, M., Saeed, Y., Zafar, Y., Malik, K.A., Briddon, R.W., Stanley J., Markham, P.G., 1999. Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. Virol. 259, 190-199.

Mayo, M.A., Leibowitz, M.J., Palukaitis, P., Scholthof, K.B.G., Simon, A.E., Stanley, J., Taliansky, M., 2005. Satellites In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), 8th Report of the International Committee on Taxonomy of Viruses. Elsevier/Academic Press, London, 1163-1169.

McKenzie, C.L., Hodges, G., Osborne, L.S., Byrne, F.J., Shatters, Jr. R.G., 2009. Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in Florida investigating the Q invasion. J. Economic Entomol. 102, 670-676.

Montero, W.R., 2003. Cassava; Biology, Production and Utilization. Crop. Sci 43,448

Mugerwa, H., Rey, M.E.C., Alicai, T., Ateka, E., Atuncha, H., Ndunguru, J., Sseruwagi, P., 2012. Genetic diversity and geographic distribution of Bemisia tabaci (Gennadius) (Hemiptera: Aleyrodidae) genotypes associated with cassava in East Africa. Ecol. Evol. 2, 2749-2762.

Ndomba, A. O, 2012. Influence of satellite DNA molecules severity of cassava begomoviruses and the breakdown of resistance to cassava mosaic disease in Tanzania. A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy, Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa.

Ndunguru, J., Fofana, B., Legg, J.P., Challappan, P., Taylor, N., Aveling, T., Thomson, G., Fauquet, C., 2008. Two novel satellite DNAs associated with bipartite cassava mosaic begomoviruses enhancing symptoms and capable of breaking high virus resistance in

cassava landraces. In: Proceeding of the First Scientific Meeting of the Global Cassava Partnership (GCP-I) Ghent, Belgium, July 21-25.

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.

Ogbe, F.O., 2001.Survey of cassava begomoviruses in Nigeria and the response of resistant cassava genotypes to African cassava mosaic begomovirus infection PhD thesis, University of Ibadan, Ibadan, Nigeria.

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

Otim-Nape, G.W., Thresh, J.M., Bua, A., Baguma, Y., Shaw, M.W., 1998.Temporal spread of cassava mosaic virus disease in a range of cassava cultivars in different agro-ecological regions of Uganda. Ann. Appl. Biol. 133, 415-430.

Patil, B.L., Dasgupta, I., 2006. Defective interfering DNAs of plant viruses. CRC Crit. Rev. Plant Sci. 25, 47-64.

Patil, B.L., Dutt, N., Briddon, R.W., Bull, S.E., Rothenstein, D., Borah, B.K., Dasgupta, I., Stanley, J., Jeske, H., 2007. Deletion and recombination events between the DNA-A and DNA-B components of Indian cassava-infecting geminiviruses generate defective molecules in *Nicotiana benthamiana*. Virus Res. 124, 59-67.

Pita, J.S., Fondong, V.N., Sangare, 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, 655-665.

Prabhaker, N., Castle, S., Henneberry, T.J., Toscano, N.C., 2005. Assessment of cross-resistance potential to neonicotinoid insecticides in *Bemisia tabaci* (Hemiptera: Aleyrodidae). Bulletin of Entomol. Res. 95, 535-543.

Puonti-Kaerlas, J., 1998. Cassava Biotechnology, Biotechnology and Genetic Engineering Reviews.15, 329-364.

Qian, Y., Zhou, X., 2005. Pathogenicity and stability of a truncated DNA β associated with Tomato yellow leaf curl China virus, Virus Res. 109, 159-163.

Rojas, M.R., Hagen, C., Lucas, W.J., Gilbertson, R.L., 2005. Exploiting chinks in the plant's armor: evolution and emergence of geminiviruses. Annu Rev Phytopathol 43, 361–394.

Roossinck, M.J., Sleat, D., Palukaitis. P., 1992. Satellite RNAs of plant viruses: structures and biological effects. Microbiol Rev. 56, 265-279.

Saunders, K., Bedford. I.D., Briddon, R.W., Markham, P.G., Wong. S.M., Stanley. J., 2000. A novel virus complex causes Ageratum yellow vein disease, Proc. Natl. Acad. Sci. U.S.A. 97, 6890-6895.

Saunders, K., I.D. Bedford, T. Yahara, Stanley. J., 2003.The earliest recorded plant virus disease. Nature, 422, 831.

Saunders, K., Stanley. P., 1999. A nanovirus-like DNA component associated with yellow vein disease of Ageratum conyzoides: evidence for interfamilial recombination between plant DNA viruses, Virol. 264, 142-152.

Schneider, L.R., 1969. Satellite-like particle of tobacco ringspot virus that resembles tobacco ringspot virus. Sci. 166, 1627-1629.

Simon, A.E., Roossinck, M.J., Havelda. Z., 2004. Plant virus satellite and defective interfering RNAs: new paradigms for a new century. Annu. Rev. Phytopathol. 42, 415-437.

Sseruwagi, P., Legg, J.P., Maruthi, M.N., Colvin, J., Rey, M.E.C., Brown, J.K., 2005. Genetic diversity of Bemisia tabaci (Gennadius) Hemiptera: Aleyrodidae) populations and presence of the B biotype and non B biotype that can induce silverleaf symptoms in squach, in Uganda. Ann. Appl. Biol. 147, 253-265.

Sseruwagi, P., Sserumbombwe, W.S., Legg, J.P., Ndunguru, J., Thresh, J.M., 2004. Methods for surveying the incidence and severity of cassava mosaic disease and whitefly vector populations on cassava in Africa: A review. Virus Res. 100, 129-142.

Stanley, J., Bisaro, D. M., Briddon, R.W., Brown, J.K., Fauquet, C.M., Harrison, B.D., Rybicki E.P., Stenger, D.C., 2005. Family Geminiviridae, Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses, Elsevier, London, 301-326.

Swanson, M.M., Harrison, B.D., 1994. Properties, relationships and distribution of cassava mosaic geminiviruses. Trop. Sci. 34, 15-25.

Tao, X., Qing, L., Zhou. X., 2004. Function of A-rich region in DNA β associated with Tomato yellow leaf curl China virus. Chin. Sci. Bull. 49, 1490-1493.

Thresh, J.M., Cooter, R.J., 2005. Strategies for controlling cassava mosaic virus disease in Africa. Plant Pathol. 54, 587-614.

Thresh, J.M., Fargette, D., Otim-Nape, G.W., 1994. Effect of African cassava mosaic virus on the yield of cassava. Trop. Sci. 28, 34-37.

Tiendrébéogo, F., Lefeuvre, P., Hoareau, M., Harimalala, M., De Bruyn, A., Villemot, A., Traoré, J., Konaté, G., Alfred S., Traoré, N., Barro, B., Reynaud, Traoré. O., Lett, J. M.,

2012. Evolution of *African cassava mosaic virus* by recombination between bipartite and monopartite begomoviruses. Virol. J. 9, 67.

Tonukari, N.J, 2004. Cassava and future of starch. Electron. J. Biotechnol. Vol. 7 No 1 2004.

Varma, A., Malathi, V.G. 2003. Emerging geminivirus problems: a serious threat to crop production. Ann. Appl. Biol. 142, 145-164.

Were, H.K., Takeshita, M., Furuya, N., Takanami, Y., 2005a. Molecular characterization of a new begomovirus infecting honeysuckle in Kobe Japan. J. Fac. Agric., Kyushu Univ. 50, 61-71.

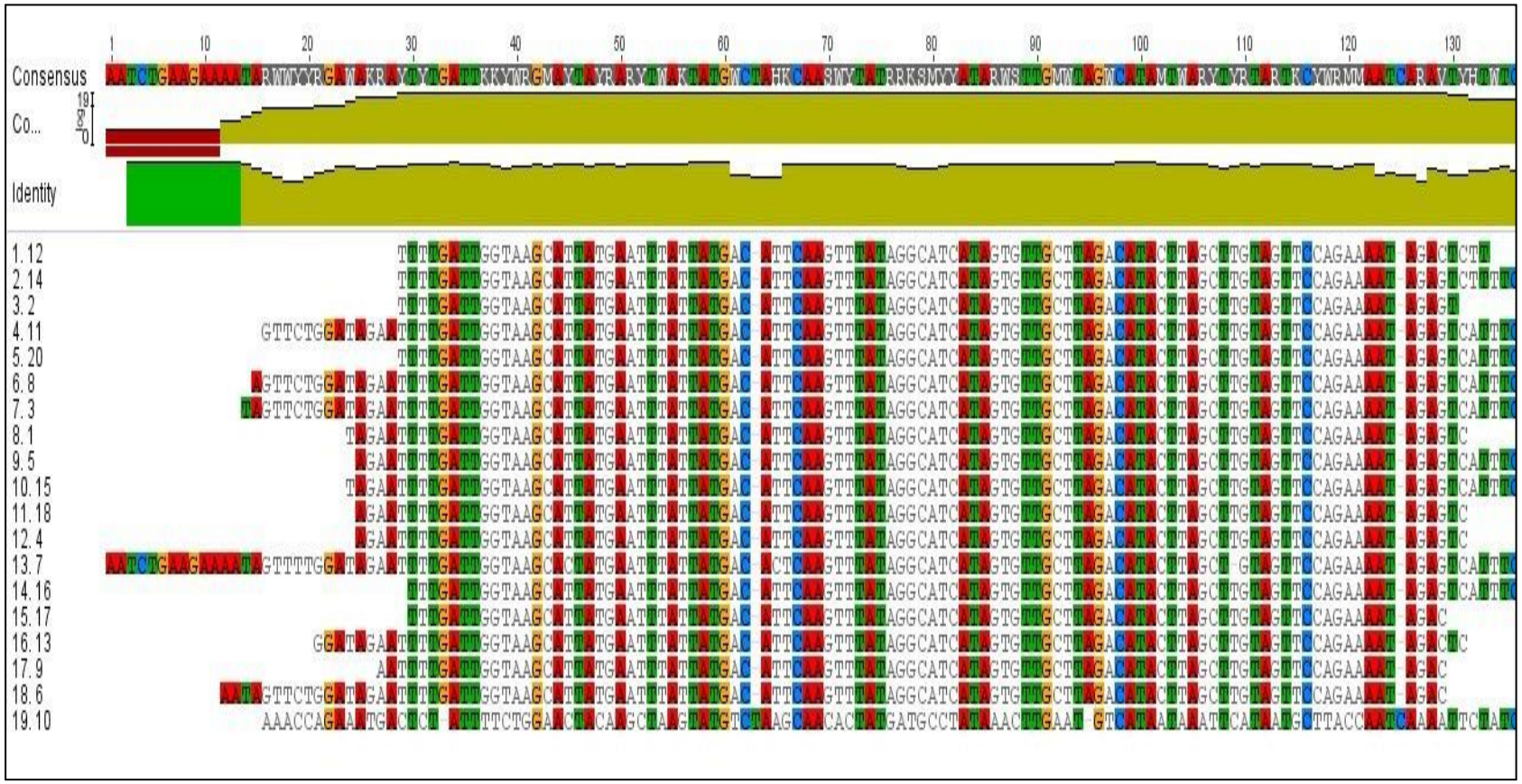
Wong, S.M., Swanson, M.M., Harrison, B.D., 1993. A geminivirus causing vein yellowing of Ageratum conyzoides in Singapore. Plant Pathol. 42, 137-139.

Wu, X., Zhou, P., 2005. Interactions between a nanovirus-like component and the tobacco curly shoot virus/satellite complex. Acta. Biochim. Biophys. Sinica 37, 25-31.

Zhou, X., Robinson, D.J., Harrison, B.D., 1998. Types of variation in DNA A among isolates of *East African cassava mosaic virus* from Kenya, Malawi and Tanzania. J. Gen. Virol. 79, 2835-2840.

List of Appendices

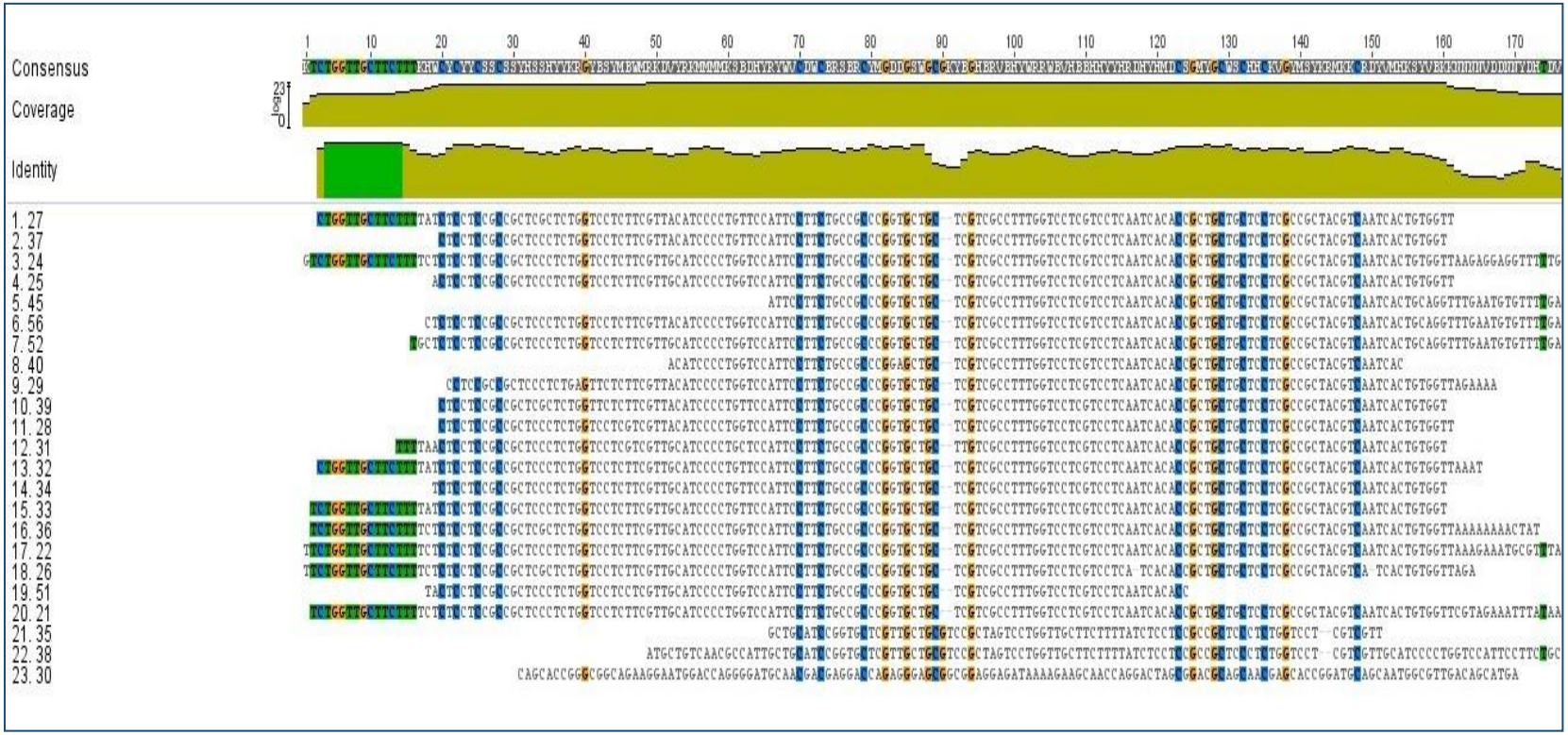
Appendix 2.1**:** Alignment of DNA-II ORF C4 transcripts isolates collected from mock-inoculated cassava plants and cassava plants inoculated with either EACMV-Ug2, ACMV and in combination with DNA-II, DNA-III and DNA-II + DNA-III clones. Highlighted colours are conserved nucleotides in all sequenced samples. Yellow colour is G, Blue colour is C, Green colour is T and Red colour is A



Appendix 2.2**:** Percentage nucleotide sequence similarity matrix table of the DNA-II ORF C4 transcript sequences

	12	14	2	11	20	8	3	1	5	15	18	4	7	16	17	13	9	6	10
12		99.0%	99.0%	98.1%	98.1%	98.1%	98.1%	99.0%	98.1%	98.1%	99.0%	99.0%	95.1%	98.0%	100%	100%	100%	100%	47.6%
14	99.0%		100%	94.3%	96.3%	96.7%	95.0%	100%	94.3%	94.5%	100%	100%	96.2%	99.0%	99.0%	99.0%	99.0%	99.0%	43.5%
2	99.0%	100%		100%	100%	100%	100%	100%	100%	100%	100%	100%	97.0%	100%	99.0%	99.0%	99.0%	99.0%	48.0%
11	98.1%	94.3%	100%		100%	100%	96.2%	100%	95.5%	95.2%	100%	100%	96.6%	100%	99.0%	99.1%	99.0%	99.1%	42.1%
20	98.1%	96.3%	100%	100%		100%	100%	100%	98.9%	99.5%	100%	100%	97.2%	100%	99.0%	99.0%	99.0%	99.0%	42.4%
8	98.1%	96.7%	100%	100%	100%		100%	100%	99.5%	100%	100%	100%	96.7%	100%	99.0%	99.1%	99.0%	99.1%	42.1%
3	98.1%	95.0%	100%	96.2%	100%	100%		100%	97.0%	97.1%	100%	100%	96.7%	100%	99.0%	99.1%	99.0%	99.1%	42.1%
1	99.0%	100%	100%	100%	100%	100%	100%		100%	100%	100%	100%	97.2%	100%	99.0%	99.1%	99.0%	99.0%	47.2%
5	98.1%	94.3%	100%	95.5%	98.9%	99.5%	97.0%	100%		98.9%	100%	100%	97.3%	100%	99.0%	99.0%	99.0%	99.0%	42.5%
15	98.1%	94.5%	100%	95.2%	99.5%	100%	97.1%	100%	98.9%		100%	100%	97.3%	100%	99.0%	99.1%	99.0%	99.0%	42.3%
18	99.0%	100%	100%	100%	100%	100%	100%	100%	100%	100%		100%	97.1%	100%	99.0%	99.0%	99.0%	99.0%	47.7%
4	99.0%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%		97.1%	100%	99.0%	99.0%	99.0%	99.0%	47.7%
7	95.1%	96.2%	97.0%	96.6%	97.2%	96.7%	96.7%	97.2%	97.3%	97.3%	97.1%	97.1%		97.1%	95.9%	96.3%	96.0%	95.7%	47.1%
16	98.0%	99.0%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	97.1%		99.0%	99.0%	99.0%	99.0%	48.6%
17	100%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	95.9%	99.0%		100%	100%	100%	48.0%
13	100%	99.0%	99.0%	99.1%	99.0%	99.1%	99.1%	99.1%	99.0%	99.1%	99.0%	99.0%	96.3%	99.0%	100%		100%	100%	47.7%
9	100%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	99.0%	96.0%	99.0%	100%	100%		100%	47.6%
6	100%	99.0%	99.0%	99.1%	99.0%	99.1%	99.1%	99.0%	99.0%	99.0%	99.0%	99.0%	95.7%	99.0%	100%	100%	100%		46.5%
10	47.6%	43.5%	48.0%	42.1%	42.4%	42.1%	42.1%	47.2%	42.5%	42.3%	47.7%	47.7%	47.1%	48.6%	48.0%	47.7%	47.6%	46.5%	

Appendix 2.3**:** Alignment of DNA-III ORF V1 and ORFC2 transcripts isolates collected from mock-inoculated cassava plants and cassava plants inoculated with either EACMV-Ug2, ACMV and in combination with DNA-II, DNA-II and DNA-II + DNA-III clones. Highlighted colours are conserved nucleotides in all the sequenced samples. Yellow colour is G, Blue colour is C, Green colour is T however A nucleotide was not conserved



Appendix 2.4**:** Percentage nucleotide sequence similarity matrix table of the DNA-III ORF VI transcripts sequence

	27	37	24	25	45	56	52	40	29	39	28	31	32	34	33	36	22	26	51	21	35	38	30
27		99.3%	97.5%	97.2%	96.8%	95.8%	94.4%	98.0%	87.8%	99.3%	97.9%	95.9%	98.7%	98.6%	98.7%	98.1%	97.5%	96.8%	94.3%	97.5%	47.7%	41.6%	29.0%
37	99.3%		98.6%	98.6%	96.8%	97.1%	96.4%	98.0%	88.4%	98.6%	98.6%	97.1%	99.3%	99.3%	99.3%	97.8%	98.6%	96.4%	97.1%	98.6%	47.7%	42.0%	30.0%
24	97.5%	98.6%		99.3%	78.6%	89.2%	90.4%	98.0%	86.2%	97.1%	98.6%	96.6%	97.5%	99.3%	98.7%	95.9%	89.7%	96.3%	97.1%	88.7%	47.7%	35.6%	29.3%
25	97.2%	98.6%	99.3%		96.8%	96.5%	97.2%	98.0%	88.5%	97.1%	98.6%	97.9%	98.6%	98.6%	98.6%	98.6%	99.3%	97.2%	99.0%	99.3%	47.7%	41.6%	29.8%
45	96.8%	96.8%	78.6%	96.8%		100%	99.1%	98.9%	93.0%	96.8%	96.8%	95.7%	92.9%	96.8%	96.8%	89.6%	76.9%	91.8%	100%	77.9%	47.7%	41.1%	26.7%
56	95.8%	97.1%	89.2%	96.5%	100%		98.8%	99.0%	84.8%	95.7%	97.1%	93.6%	93.2%	96.4%	95.7%	91.6%	88.0%	93.1%	96.2%	88.0%	47.7%	39.9%	27.9%
52	94.4%	96.4%	90.4%	97.2%	99.1%	98.8%		98.0%	84.1%	95.0%	96.4%	93.7%	93.2%	97.1%	95.8%	91.7%	88.6%	93.2%	97.1%	88.6%	47.7%	41.2%	27.1%
40	98.0%	98.0%	98.0%	98.0%	98.9%	99.0%	98.0%		99.0%	98.0%	99.0%	96.0%	97.0%	97.0%	97.0%	98.0%	98.0%	96.0%	97.2%	98.0%	46.5%	41.7%	30.1%
29	87.8%	88.4%	86.2%	88.5%	93.0%	84.8%	84.1%	99.0%		88.4%	88.5%	86.2%	86.7%	87.7%	87.7%	87.6%	87.6%	86.6%	83.3%	86.9%	47.7%	40.3%	30.7%
39	99.3%	98.6%	97.1%	97.1%	96.8%	95.7%	95.0%	98.0%	88.4%		97.1%	95.7%	97.8%	97.8%	97.8%	97.8%	97.1%	96.4%	95.1%	97.1%	47.7%	42.0%	29.2%
28	97.9%	98.6%	98.6%	98.6%	96.8%	97.1%	96.4%	99.0%	88.5%	97.1%		97.8%	97.9%	97.8%	97.8%	97.9%	98.6%	96.4%	98.1%	98.6%	47.7%	41.6%	31.3%
31	95.9%	97.1%	96.6%	97.9%	95.7%	93.6%	93.7%	96.0%	86.2%	95.7%	97.8%		97.2%	97.1%	97.2%	95.9%	96.6%	94.5%	96.2%	96.6%	46.5%	42.0%	30.0%
32	98.7%	99.3%	97.5%	98.6%	92.9%	93.2%	93.2%	97.0%	86.7%	97.8%	97.9%	97.2%		100%	100%	97.5%	98.1%	96.3%	96.2%	96.3%	47.7%	40.2%	29.6%
34	98.6%	99.3%	99.3%	98.6%	96.8%	96.4%	97.1%	97.0%	87.7%	97.8%	97.8%	97.1%	100%		100%	98.6%	99.3%	97.1%	97.1%	99.3%	47.7%	42.0%	29.2%
33	98.7%	99.3%	98.7%	98.6%	96.8%	95.7%	95.8%	97.0%	87.7%	97.8%	97.8%	97.2%	100%	100%		98.1%	98.7%	96.8%	96.2%	98.7%	47.7%	42.0%	29.2%
36	98.1%	97.8%	95.9%	98.6%	89.6%	91.6%	91.7%	98.0%	87.6%	97.8%	97.9%	95.9%	97.5%	98.6%	98.1%		96.5%	98.1%	96.2%	95.3%	47.7%	40.0%	29.3%
22	97.5%	98.6%	89.7%	99.3%	76.9%	88.0%	88.6%	98.0%	87.6%	97.1%	98.6%	96.6%	98.1%	99.3%	98.7%	96.5%		97.5%	97.1%	87.1%	47.7%	38.4%	30.7%
26	96.8%	96.4%	96.3%	97.2%	91.8%	93.1%	93.2%	96.0%	86.6%	96.4%	96.4%	94.5%	96.3%	97.1%	96.8%	98.1%	97.5%		95.2%	96.9%	47.7%	40.5%	29.9%
51	94.3%	97.1%	97.1%	99.0%	100%	96.2%	97.1%	97.2%	83.3%	95.1%	98.1%	96.2%	96.2%	97.1%	96.2%	96.2%	97.1%	95.2%		97.1%	44.1%	39.5%	25.5%
21	97.5%	98.6%	88.7%	99.3%	77.9%	88.0%	88.6%	98.0%	86.9%	97.1%	98.6%	96.6%	96.3%	99.3%	98.7%	95.3%	87.1%	96.9%	97.1%		47.7%	36.3%	28.6%
35	47.7%	47.7%	47.7%	47.7%	47.7%	47.7%	47.7%	46.5%	47.7%	47.7%	47.7%	46.5%	47.7%	47.7%	47.7%	47.7%	47.7%	47.7%	44.1%	47.7%		100%	26.7%
38	41.6%	42.0%	35.6%	41.6%	41.1%	39.9%	41.2%	41.7%	40.3%	42.0%	41.6%	42.0%	40.2%	42.0%	42.0%	40.0%	38.4%	40.5%	39.5%	36.3%	100%		25.4%
30	29.0%	30.0%	29.3%	29.8%	26.7%	27.9%	27.1%	30.1%	30.7%	29.2%	31.3%	30.0%	29.6%	29.2%	29.2%	29.3%	30.7%	29.9%	25.5%	28.6%	26.7%	25.4%	