

**EPIDEMIOLOGY OF CASSAVA MOSAIC DISEASE AND  
MOLECULAR CHARACTERIZATION OF CASSAVA MOSAIC  
VIRUSES AND THEIR ASSOCIATED WHITEFLY (BEMISIA  
TABACI) VECTOR IN SOUTH AFRICA**

Kenneth Gaza Mabasa

A dissertation submitted to the school of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment for the degree of Master of Science.

Johannesburg, 2007

## **DECLARATION**

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

---

(Kenneth Gaza Mabasa)

\_\_\_\_\_ Day of \_\_\_\_\_ 2007

## ABSTRACT

Cassava mosaic disease (CMD) is caused by whitefly-transmitted geminiviruses and is a major constraint to cassava production in Africa. Field surveys were conducted in three (Bushbuckridge, Mariti and Tonga) cassava growing areas of Limpopo and Mpumalanga provinces in South Africa during two seasons (2004/2005 and 2005/2006). Results showed that a higher percentage (27.1%) of CMD infection was due to the use of infected planting materials compared to whitefly borne-infections (10.4%). Disease symptoms were generally mild. There was no change in disease incidence over the survey period. Molecular characterization of cassava mosaic geminiviruses (CMG's), using differential primer PCR, restriction fragment length polymorphisms (RFLP's), phylogenetic and recombination analysis and screening for satellite DNA's. Differential primer PCR and RFLP's showed that *African cassava mosaic virus* (ACMV) was the most prevalent virus in South Africa and that mixed infections were a common occurrence. Phylogenetic analysis and RFLP's showed the presence of a 'new' strain of ACMV in South Africa. EACMV isolates from this study showed more frequent recombination compared to ACMV isolates. None of the samples tested positive for satellite DNA's. Phylogenetic analysis of *Bemisia tabaci* using the mitochondrial cytochrome oxidase gene sequences revealed a 'new' sister clade of *B. tabaci* that is closely related to the previously identified southern African clade and the presence of the Q biotype that groups with Q biotypes of North African/Mediterranean origin. Good cultural practices, introduction of resistant cultivars and continuous monitoring are required to reduce the impact of CMD in South Africa.

## ACKNOWLEDGEMENTS

Thanks to my supervisor Prof. Chrissie Rey for giving me this opportunity, patience and allowing me space to exercise free thought.

To my family thanks for your support, *mi ve khokholo leri ndzi nga tiya hi rona ku kondza ndzi ta fika laha ndzi nga kona. Ndza khensa wena manana Tsatsawani N'wa Hasani Chauke, va makwavo, Mthavini, Mamayila, Mphephu, Flora, Selinah na Giyani. Eka tatana John "Vhavhu" Mabasa ndza tshembha a wu ta ti nyungubyisa namunthla.*

Thanks to the USAID, NRF, NSFAS and the Mellon post graduate fund for financial support for the project and personal finance.

Officers of the Department of Agriculture in Thulamahashe, farmers in Tonga, Mariti and Bushbuckridge and Jim of the cassava farm in Barberton thank you.

Thanks to Darren Martin of the University of Cape Town for assistance on recombination analysis

James Legg and staff at the IITA in Uganda for field epidemiological training.

MCB staff and colleagues of the Cassava Biotechnology program at the University of the Witwatersrand, Johannesburg, and thanks for your support.

## CONTENTS

|                            |      |
|----------------------------|------|
| DECLARATION .....          | II   |
| ABSTRACT.....              | III  |
| CONTENTS .....             | V    |
| LIST OF FIGURES .....      | VIII |
| LIST OF TABLES.....        | X    |
| LIST OF ABBREVIATIONS..... | XI   |

## CHAPTER 1: INTRODUCTION

|                                                                                         |    |
|-----------------------------------------------------------------------------------------|----|
| 1.1. THE CASSAVA CROP .....                                                             | 2  |
| 1.2. ECONOMIC IMPORTANCE OF CASSAVA IN SOUTHERN AFRICA.....                             | 4  |
| 1.3. CONSTRAINTS TO CASSAVA PRODUCTION .....                                            | 5  |
| 1.4. CASSAVA MOSAIC DISEASE.....                                                        | 6  |
| 1.4.1. The Causal Agent of CMD .....                                                    | 7  |
| 1.4.2. General Characteristics and Structure of Geminiviruses.....                      | 8  |
| 1.4.3. Cassava Infecting Begomovirus Diversity.....                                     | 9  |
| 1.4.4. Recombination and Psuedorecombination of Cassava Infecting<br>Begomoviruses..... | 11 |
| 1.4.5. Satellite DNA's.....                                                             | 13 |
| 1.4.6. Symptoms .....                                                                   | 13 |
| 1.5. DIAGNOSTIC METHODS .....                                                           | 14 |
| 1.5.1. Serological Tests.....                                                           | 14 |
| 1.5.2. Nucleic Acid Based Tests .....                                                   | 14 |
| 1.6. THE WHITEFLY VECTOR (BEMISIA TABACI: GENNADIUS) .....                              | 16 |

|                                                   |    |
|---------------------------------------------------|----|
| 1.6.1. CMG transmission by <i>B. tabaci</i> ..... | 18 |
| 1.6.2. Taxonomy of the whitefly .....             | 18 |
| 1.6.3. Whitefly Biotypes and Host races.....      | 19 |
| 1.7. OBJECTIVES .....                             | 21 |
| 1.7.1. General Objectives.....                    | 21 |
| 1.7.2. Specific Aims.....                         | 22 |

## **CHAPTER 2: THE EPIDEMIOLOGY OF CASSAVA MOSAIC DISEASE**

|                                      |    |
|--------------------------------------|----|
| 2.1. ABSTRACT.....                   | 24 |
| 2.2. INTRODUCTION .....              | 25 |
| 2.3. Materials and Methods.....      | 28 |
| 2.3.1. Geographical Location.....    | 28 |
| 2.3.2. Survey records .....          | 30 |
| 2.3.3. Data analysis .....           | 31 |
| 2.4. RESULTS .....                   | 31 |
| 2.4.1. CMD incidence .....           | 31 |
| 2.4.2. Disease Severity.....         | 33 |
| 2.4.3. Whitefly Abundance .....      | 33 |
| 2.5. DISCUSSION AND CONCLUSIONS..... | 35 |

## **CHAPTER 3: BIODIVERSITY OF CASSAVA MOSAIC GEMINIVIRUSES IN SOUTH AFRICA**

|                                                 |    |
|-------------------------------------------------|----|
| 3.1. ABSTRACT.....                              | 40 |
| 3.2. INTRODUCTION .....                         | 41 |
| 3.3. MATERIALS AND METHODS .....                | 45 |
| 3.3.1. Virus sources and PCR amplification..... | 45 |

|                                                                      |           |
|----------------------------------------------------------------------|-----------|
| 3.3.2. Restriction Fragment Length Polymorphisms (RFLP's) .....      | 47        |
| 3.3.3. Cloning, Sequencing and Phylogenetic Analysis .....           | 47        |
| 3.3.4. Recombination Analysis .....                                  | 48        |
| 3.4. RESULTS .....                                                   | 49        |
| 3.4.1. PCR Amplification and RFLP's .....                            | 49        |
| 3.4.2. Phylogenetic relationships and recombination analysis .....   | 53        |
| 3.5. DISCUSSION AND CONCLUSIONS.....                                 | 57        |
| <br><b>CHAPTER 4: PHYLOGENETIC ANALYSIS OF <i>BEMISIA TABACI</i></b> |           |
| <b>COLONIZING CASSAVA IN SOUTH AFRICA USING THE</b>                  |           |
| <b>MITOCHONDRIAL CYTOCHROME OXIDASE I (MTCOI) GENE</b>               |           |
| 4.1. ABSTRACT.....                                                   | 63        |
| 4.2. INTRODUCTION .....                                              | 64        |
| 4.3. MATERIALS AND METHODS .....                                     | 67        |
| 4.3.1. Whitefly samples .....                                        | 67        |
| 4.3.2. DNA extraction, PCR and Cloning.....                          | 67        |
| 4.3.3. Phylogenetic analysis.....                                    | 69        |
| 4.4. RESULTS .....                                                   | 69        |
| 4.5. DISCUSSION AND CONCLUSIONS.....                                 | 73        |
| <br><b>CHAPTER 5: SUMMARY</b>                                        |           |
| 5.1. EPIDEMIOLOGY OF CASSAVA MOSAIC DISEASE.....                     | 78        |
| 5.2. GENETIC DIVERSITY OF CASSAVA MOSAIC GEMINIVIRUSES.....          | 79        |
| 5.3 GENETIC DIVERSITY OF <i>BEMISIA TABACI</i> .....                 | 81        |
| 5.4. CONCLUSIONS.....                                                | 82        |
| <b>CHAPTER 6: REFERENCES.....</b>                                    | <b>84</b> |

## **LIST OF FIGURES**

**FIGURE 2.1** Cassava growing areas in Mpumalanga and Limpopo in which surveys were conducted (inserts represent position of respective areas in South Africa).

**FIGURE 2.2** Cassava mosaic disease incidences during 2004/2005 and 2005/2006 growing seasons in Mpumalanga and Limpopo provinces of South Africa.

**FIGURE 2.3** Percentage cutting (C) and whitefly (W) CMD incidence for 2004/2005 and 2005/2006 growing seasons in Mpumalanga (Tonga and Mariti) and Limpopo (Bushbuckridge) Provinces of South Africa.

**FIGURE 2.4.** Average whitefly numbers observed during surveys in 2004/2005 and 2005/2006 growing seasons in the regions of Mpumalanga and Limpopo provinces in South Africa.

**FIGURE 2.5** Disease severity and disease index observed during surveys in Mpumalanga and Limpopo provinces of South Africa.

**FIGURE 3.1** PCR using core coat protein (CCP) primers of some of the samples collected during January 2005 in Limpopo and Mpumalanga provinces of South Africa.

**FIGURE 3.2** Near Full length DNA-A amplified with universal PCR primers UN/F and UN/R. Lane 2 is the positive control.

**FIGURE 3.3** A and B. PCR products obtained using differential primers JSP001, JSP002 (top row) and JS001, JSP003 (bottom row) to distinguish between ACMV and EACMV isolates respectively.

**FIGURE 3.4** (A, B, and C) PCR-RFLP patterns using *EcoRV* (top rows) and *MluI* (bottom rows) restriction enzymes.

**FIGURE 3.5** phylogenetic trees showing near full length DNA-A sequence relationships between four South African cassava mosaic geminiviruses isolate (SA-Barb, SA-Mariti, SA-Tonga and SA-Bush) and other cassava mosaic geminiviruses.

**FIGURE 3.6** Recombination map of putative recombinant fragments for DNA-A of four isolates of cassava mosaic geminiviruses (CMG's) from South Africa together with selected previously identified CMG's.

**FIGURE 4.1** MTCOI tree showing relationships & identification of Southern African and reference whitefly populations

**Figure 4.2** Bootstrap MTCOI tree showing relationships and identification of Southern African and reference whitefly populations

## LIST OF TABLES

**TABLE 2.1** Summary of survey data collected during 2005 and 2006. The following parameters were recorded: disease incidence (cutting and whitefly-borne), disease severity and index, and whitefly numbers.

**TABLE 3.1** Predicted RFLP's for published cassava mosaic geminivirus sequences following computer-based digestion of full-length DNA-A with the restriction enzymes *EcoRV*, *MluI* and *PstI*

**TABLE 3.2** Percentage of viruses identified in Bushbuckridge, Mariti and Tonga region using RFLP's.

**TABLE 4.1** Percentage nucleotide sequence identities between selected South African whiteflies from cassava and other host plants from various African Countries

## LIST OF ABBREVIATIONS

|                    |                                            |
|--------------------|--------------------------------------------|
| <b>ACMV</b>        | African cassava mosaic virus               |
| <b>Barb</b>        | Barberton                                  |
| <b>Bush</b>        | Bushbuckridge                              |
| <b>CBSD</b>        | Cassava brown streak disease               |
| <b>CBSV</b>        | Cassava brown steak virus                  |
| <b>CCP</b>         | Core coat protein                          |
| <b>CLV</b>         | Cassava latent virus                       |
| <b>CMD</b>         | Cassava mosaic Disease                     |
| <b>CMG's</b>       | Cassava mosaic geminiviruses               |
| <b>CMV</b>         | Cassava mosaic virus                       |
| <b>CR</b>          | Common region                              |
| <b>CSM</b>         | Cassava starch manufacturing               |
| <b>DAS (ELISA)</b> | Double antibody sandwich ELISA             |
| <b>DNA</b>         | Deoxyribonucleic acid                      |
| <b>DI</b>          | Disease index                              |
| <b>DS</b>          | Disease severity                           |
| <b>dsDNA</b>       | Double stranded DNA                        |
| <b>EACMV</b>       | East African cassava mosaic virus          |
| <b>EACMCV</b>      | East African cassava mosaic Cameroon virus |
| <b>EACMMV</b>      | East African cassava mosaic Malawi virus   |
| <b>EACMZV</b>      | East African cassava mosaic Zanzibar virus |
| <b>EDTA</b>        | ethylenediamine tetra-acetic acid          |
| <b>ELISA</b>       | enzyme-linked immuno-absorbent assay       |

|                  |                                                    |
|------------------|----------------------------------------------------|
| <b>FAO</b>       | Food and Agriculture Organization                  |
| <b>HMA</b> s     | Heteroduplex mobility assays                       |
| <b>ICMV</b>      | Indian cassava mosaic virus                        |
| <b>ICTV</b>      | International Committee of the Taxonomy of Viruses |
| <b>IR</b>        | Intergenic region                                  |
| <b>ITS1</b>      | Internal transcriber sequence                      |
| <b>KE</b>        | Kenya                                              |
| <b>MAbs</b>      | Monoclonal antibodies                              |
| <b>Mld</b>       | Mild                                               |
| <b>mtCOI</b>     | Mitochondrial cytochrome oxidase I                 |
| <b>MW</b>        | Malawi                                             |
| <b>ORF</b>       | Open reading frame                                 |
| <b>PCR</b>       | Polymerase chain reaction                          |
| <b>RAPD</b>      | Random amplified polymorphic DNA                   |
| <b>RCA</b>       | Rolling circle amplification                       |
| <b>RDP</b>       | Recombination detection program                    |
| <b>Rep</b>       | Replicase                                          |
| <b>RFLP</b>      | Restriction fragment length polymorphism           |
| <b>RNA</b>       | Ribonucleic acid                                   |
| <b>rDNA</b>      | Ribosomal DNA                                      |
| <b>SA</b>        | South Africa                                       |
| <b>Svr</b>       | Severe                                             |
| <b>SLCMV</b>     | Sri-Lankan cassava mosaic virus                    |
| <b>ssDNA</b>     | Single stranded DNA                                |
| <b>TAS ELISA</b> | Triple antibody sandwich ELISA                     |

|              |                                    |
|--------------|------------------------------------|
| <b>TGMV</b>  | Tomato golden mosaic virus         |
| <b>TNA</b>   | Total nucleic acids                |
| <b>TYLCV</b> | Tomato yellow leaf curl virus      |
| <b>TZ</b>    | Tanzania                           |
| <b>UGV</b>   | Uganda variant                     |
| <b>WTGs</b>  | Whitefly transmitted geminiviruses |

## CHAPTER 1: INTRODUCTION

### 1.1. The Cassava Crop

Cassava (*Manihot esculenta* Crantz), is a shrubby perennial plant grown mainly for its carbohydrate rich tuberous roots. It belongs to the family *Euphorbiaceae* that also includes other commercially important plants like castor bean (*Ricinus communis* L.) and rubber (*Hevea brasiliensis* L.). Cassava and some 90 other species make up the genus *Manihot*, and it is the only widely cultivated member of this genus.

Cassava is believed to have originated in South America and was introduced into Africa in the 16<sup>th</sup> century, and later into Asia in the late 17<sup>th</sup> century by Portuguese traders. Today, cassava is cultivated in more than 80 countries mainly between 30° south and 30° north of the equator (Fauquet and Fargette, 1990). Cassava is suited to warm humid lowland tropics and can be cultivated in most areas where the mean annual temperature exceeds 20°C with annual rainfall that varies between 500mm and 8000 mm (Pounti-Kearlas, 1998). Although cassava tolerates drought, it grows best at rainfalls exceeding 1200mm on many soil types and requires only limited agronomic and pest management practices. Furthermore the roots can be left in the ground for a long time before harvesting, thus giving poor farmers a useful security against famine. These characteristics make cassava the most cheaply cultivated crop as compared to other major staple crops such as rice, maize, wheat, and sugarcane, thus making cassava convenient for small-scale farmers in many tropical countries who have limited access to expensive agricultural inputs.

Most of the harvest from cassava is used for human consumption, either fresh or in various processed forms. The rest is processed to animal feed and industrial products. Cassava is consumed by an estimated 600 million people (FAO, 2006). As a raw material, 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.

Although cassava has a relatively recent history in Africa, it has become one of the most important crops grown on the continent. It was introduced from Brazil into the west coast of Africa and later to East Africa through Madagascar and Zanzibar. Africa has become the largest producer of cassava, constituting 54% of world production. Cassava is also widely grown in Latin America and Asia. Today, cassava is grown in all African countries south of the Sahara and has since become the dominant staple food. In Central Africa for example, cassava constitutes over 50% of the average staple food consumption in the Democratic Republic of Congo, Congo (Brazzaville), and the Central African Republic (Nweke and Lynam, 1997). 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.

Most of the spread of cassava in Africa away from the coastal areas and riverside trading posts took place during the 20<sup>th</sup> 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).

In Africa cassava is primarily produced for human consumption, where it is consumed in various processed forms. Food products include *gari*, *fufu*, and tapioca. The use of cassava flour is gaining ground in the production of biscuits, sausage rolls, meat pies and bread (Ogbe, 2001).

## **1.2. Economic Importance of Cassava in southern Africa**

Cassava was introduced into Mozambique by the Portuguese in the 17<sup>th</sup> century and was adopted as a food crop by Tsonga tribesman, who later spread westwards into the old Eastern Transvaal (Now Mpumalanga) and Swaziland, and south into Northern KwaZulu-Natal (Woodward *et al.*, 1997). Cultivation took hold only gradually and it appears that plantings in South Africa came mainly with the major tribal movements of the 1830s and 1860s (Trench and Martin, 1984). Cassava is also extensively grown in Zambia, Malawi and Zimbabwe and was probably introduced into Zambia via the Congo basin (where it was well established by the early 1650s), in Zimbabwe and Malawi via Portuguese trading routes from Mozambique on the east coast of Africa (Haggblade and Zulu, 2003).

In South Africa, cassava is grown as a secondary staple food mainly in the provinces of Mpumalanga, Kwazulu-Natal and Limpopo (Woodward *et al.*, 1997). One of the most important uses of cassava in South Africa is the production of industrial starch products by a private company called Cassava Starch Manufacturing Company (CSM) with a factory at Dendron in Limpopo province. Commercial cassava farms have recently been established in Mpumalanga (Barberton) and plans for a second factory there are in progress. The main consumers of their starch are the food, textile, paper, corrugated cardboard and the mining industries. The company owns about 2000 hectares under cassava and they get additional raw materials by contracting small-scale farmers to produce cassava for their factory.

In Mozambique, food security is a big problem, and one third of the total population is considered chronically food insecure, and subsistence agriculture is the main form of livelihood, providing more than 80% of basic food needs (Equator Initiative, 2002). Although maize is the major staple crop grown in all regions of the

country, other crops include rice, sorghum, millet, Irish potatoes, cassava and a wide variety of vegetables. Cassava is grown mainly in northern Mozambique where it is the main staple food in the provinces of Nampula, Zambezia and Cabo Delgado (Thresh, 2001). Cassava is being introduced along with sweet potato under government initiative in drought prone areas throughout Mozambique (Equator Initiative, 2003).

Large-scale production of cassava was constrained by a taste preference for maize and post colonial government policies in most southern African countries, which favored maize over cassava. As a result maize is grown in areas environmentally not suited to it and where cassava would perform far better in terms of reliable yields. However, since the early 1990s this situation is changing due to droughts and unsustainably high maize subsidies. For example efforts to promote cultivation of cassava in Zambia and Malawi have paid off and cassava production in these countries has grown by between 6% and 8% per year respectively, which is among the fastest growth rates in Africa and the world (Haggblade and Zulu, 2003)

Therefore such initiatives are needed in Mozambique as well as in South Africa where cassava could replace maize in drought-prone areas and marginal soils without interfering with land most suited to maize production. Cassava can play a useful role in the development of subsistence farming in southern Africa, by providing an opportunity for rural farmers to develop a cash crop.

### **1.3. Constraints to Cassava Production**

Although Africa is the largest producer of cassava, yields in Africa are very low (estimated at 8.9 tons/ha) as compared to other cassava producing regions (Asia and Latin America) despite the fact that, under optimal conditions cassava can produce up

to 80 tons/ha of tubers in a 12 months culture period (Legg and Thresh, 2003). A number of factors are responsible for the severely reduced yields in Africa, and the most important constraints are virus diseases, particularly cassava mosaic disease (CMD), cassava brown streak disease (CBSD) as well as bacterial blight (caused by *Xanthomonas axonopodis* pv. *manihotis*). Other constraints are poor agricultural practices and various other diseases caused by bacteria, fungi and nematodes (Hillocks and Wydra, 2002), most of which are considered of minor importance. Cassava mosaic disease is now considered to be the most damaging pest or disease constraint to cassava production in Africa and Asia, while CBSD is most prevalent in the coastal regions of East Africa with the greatest effects in northern Mozambique (Thresh, 2001).

#### **1.4. Cassava Mosaic Disease**

Symptoms of CMD were first reported by Warburg in 1894 in what is now Tanzania. The disease was later reported in many other countries in east, west and central Africa and it is now known to occur in all the cassava-growing countries of Africa and the adjacent islands, India and Sri Lanka. In East Africa, the disease was not reported to cause serious damage until the 1920s whereas in West Africa, CMD was first recorded in the coastal areas of Nigeria, Sierra Leone and Ghana in 1929 and had spread northward by 1945 (Fauquet and Fargette, 1990).

Available data on the incidence of CMD is highly variable and various sources of data indicate the incidence to range from 15-50% (Thresh *et al.*, 1997). However, Thresh *et al.*, (1997) suggested that the overall incidence of CMD is currently 50-60% and diseased plants sustain losses of up to 40%. On these assumptions losses in Africa

are estimated at 15-20%, equivalent to 12-23million tons compared with actual production estimates of 73million tons. Due to limited information available, these figures cannot be taken as definitive estimates of the losses caused by CMD. Recent estimates in 2003 puts losses in Africa at 19-27t (Legg and Thresh, 2004). However losses of up to 82% in CMD pandemic-affected areas compared to areas not yet affected (Legg *et al.*, 2006).

#### **1.4.1. The Causal Agent of CMD**

Cassava is a vegetatively propagated crop and virus diseases cause particular problems as they are carried from one crop cycle to next through stem cuttings that are used as planting material. At least sixteen different viruses have been isolated from cassava and these belong to at least four families and genera, namely; *Comoviridae: Nepovirus*, *Geminiviridae: Begomovirus*, *Potyviriidae: Ipomovirus*, and *Caulimoviridae: Caulimovirus* (Legg and Thresh, 2003). However there is limited information on the properties, distribution, effects and economic importance of most of these viruses. Only two genera are of economic importance in Africa with regard to cassava, namely *Ipomovirus: cassava brown streak virus* (CBSV) and *Begomovirus: cassava mosaic geminiviruses* (CMG's) of the family *Geminiviridae*. Although CBSV is important in some parts of Africa (i.e. Mozambique), CMG's are the most economically important viruses of cassava in Africa as a whole.

Geminiviruses are currently divided into four genera on the basis of their genome organizations, biological properties and their vector (Fauquet and Stanley, 2003). Those that have monopartite genomes and are transmitted by leafhopper vectors, primarily to monocotyledonous plants, are included in the genus *Mastrevirus* (Group I) of which *Maize streak virus* is the type species. Viruses that have monopartite genomes distinct from those of the mastreviruses and that are transmitted by

leafhopper vectors to dicotyledonous plants are included in the genus *Curtovirus* (Group II) with *Beet curly top virus* as the type species. The genus *Topocuvirus*, recently recognized by the International Committee on Taxonomy of Viruses (ICTV) (Pringle, 1999), has only one member (also the type species), *Tomato pseudo-curly top virus*, which has a monopartite genome and is transmitted by a treehopper vector to dicotyledonous plants. The genus *Begomovirus* (group III) contains viruses that are transmitted by the whitefly *Bemisia tabaci* (Gennadius) to dicotyledonous plants, with *Bean golden yellow mosaic virus* (originally *Bean golden mosaic virus* – Puerto Rico) as the type species (Legg and Thresh, 2003). Cassava mosaic viruses belongs to this group (i.e. group III).

#### **1.4.2. General Characteristics and Structure of Geminiviruses**

The family *Geminiviridae* is a unique group of viruses, characterized by their twinned icosahedral particle morphology. The twinned particles together are approximately 15-20nm by 25-35nm in size (Harrison, 1985). The viruses contain protein subunits arranged in an icosahedral array of 22 pentameres of approximately 30kDa (Bock and Guthrie, 1978), which enclose a genome of single stranded circular DNA. Geminate virions may contain one (monopartite) or two (bipartite) distinct DNA molecules depending on the genus, for a total genome size ranging from 2.5 to 5.2 kb.

The *Geminivirus* genomes replicate via double stranded DNA intermediates and are believed to replicate through a rolling cycle mechanism. For example, ACMV DNA forms that are consistent with rolling cycle replication of virion sense DNA have been identified (Stanley, 1995). Furthermore, Koonin and Ilyina (1992) have shown that *Geminivirus*-encoded replicase associated proteins (Reps) are distantly

related to proteins involved in the initiation of rolling circle replication of certain prokaryotic plasmids.

Bipartite geminiviruses such as ACMV have a sequence of approximately 200 nucleotides that is conserved between the genomic components and located primarily within an intergenic region (referred to as the common region). In the common region there is a nano-nucleotide motif TAATATTAC, found in all geminiviruses at the apex of a stem-loop structure. This motif is cleaved during initiation and termination of rolling circle replication by *Geminivirus* Rep proteins (Laufs *et al.*, 1995).

Geminiviruses have an inherent capacity to recombine between and among themselves, and are thereby constantly evolving to generate new biodiversity, combined with their ability to act in a synergistic manner; this makes them highly opportunistic and capable of generating dramatic new epidemics (Zhou *et al.*, 1997) such as those significantly impacting cassava farmers in Uganda and Central Africa. Characterization and classification of geminiviruses, therefore is very critical in combating such epidemics and can result in rapid diagnostic methods for the viruses.

#### **1.4.3. Cassava Infecting Begomovirus Diversity**

For many years CMD has been assumed to be caused by a virus (Zimmermann, 1906). The first epidemiological information came from a study by Storey and Nichols (1938) who described virus strains based on severity of disease and thus divided them into mild and severe strains. Storey and Nichols (1938) further described the mechanism of transmission and concluded that the whitefly *B. tabaci* was probably the vector. However its etiology was not clear until in the late 1970s when Bock and Guthrie (1978) described a virus that could be transmitted by mechanical inoculation of sap from mosaic-infected cassava to *Nicotiana clevelandi* Grey and

they named the causal agent of CMD; *cassava latent virus*. The etiology of the virus was finally determined in 1983 by Bock and Woods (1983), who proved Koch's postulates for the virus and named it *African cassava mosaic virus* (ACMV).

As mentioned above the ACMV genome consists of two DNA molecules of similar size but different nucleotide sequence, known as DNA-A (or DNA-1) and DNA-B (or DNA-2) (Stanley and Gay, 1983), which are in the size range of 2500-2900 nucleotides. Both DNA components contain protein coding nucleotide sequences in the virus strand and in the complementary strand. DNA-A generally contains two genes (AV1 and AV2) in the virus strand and four genes (AC1, AC2, AC3 and AC4) in the complementary strand. AV1 encodes for the virus coat protein and AC1 codes for a replication associated protein (Rep). The protein product of AC2 is a transcriptional activator (TrAP) for virus-sense genes. The virus-sense and complementary strands of DNA-B each contain one gene (BV1 and BC1, respectively). Both gene products are involved in virus movement within the host (Sanderfoot and Lazarowitz, 1996). Between the initiation codons of AV2 and AC1 in DNA-A lies the intergenic region (IR) and in DNA-B there is an equivalent IR between the initiation codons of BV1 and BC1.

Comparisons of nucleotide sequences among bipartite begomoviruses showed that the sequences of DNA-B are more diverse than those of DNA-A. It is relatively easy to align the DNA-A sequences of diverse begomoviruses, but alignments of DNA-B sequences are less straightforward because there are fewer conserved elements. However, the part of the genome that shows the greatest variation is the IR (Rybicki, 1994; Harrison and Robinson, 1999).

#### 1.4.4. Recombination and Psuedorecombination of Cassava Infecting

##### Begomoviruses

Early classification of geminiviruses was based on serological methods and as a result three distinct groups of cassava infecting whitefly-transmitted *Geminiviruses* have been identified based on their reactions to a panel of antibodies (Hong *et al.*, 1993). When a panel of 17 antibodies (MAbs) to ACMV were used to determine the epitope profiles of a number of geminivirus strains from cassava, considerable differences were found. Group A, *African cassava mosaic virus* (ACMV) reacted with at least 15 monoclonal antibodies and was found in West Africa, including Burundi, western parts of Kenya, Chad, Uganda, Cameroon and South Africa (Harrison *et al.*, 1991, Swanson and Harrison, 1994). Group B isolates known as *East African cassava mosaic virus* (EACMV), reacted with five to nine monoclonal antibodies and were found in Malawi, Madagascar, Zimbabwe and the eastern parts of Kenya and Tanzania (Swanson and Harrison, 1994). A third group (group C) known as the *Indian cassava mosaic virus* occurred in India and Sri Lanka, and reacted with only two or three monoclonal antibodies.

The trafficking of cassava stakes resulted in many of the geographical boundaries for virus species being crossed. For example, EACMV was reported in Cameroon (Fondong *et al.*, 2000) as well as other West African countries where it was previously thought not to occur. Such crossing of geographical boundaries can result in mixed infections, thus creating opportunities for recombination and pseudo-recombination. The introduction of ‘new’ viruses into an area and the presence of whiteflies can further exacerbated CMD spread by creating opportunities for mixed infections with resulting recombination/pseudo-recombination and synergism of the virus species involved. Pseudo-recombination occurs when DNA-A of one virus

trans-replicates with DNA-B of another (Legg and Fauquet, 2004). Recombination and pseudo-recombination seem to be the major driving force in the evolution of cassava infecting geminiviruses and such recombination events could result in the evolution of new chimeric viruses having greater virulence and disease capabilities. A very good example is presented by the recombinant virus EACMV-UG (Uganda variant), which caused a severe form of CMD in Uganda that has caused serious reduction in yields (Zhou *et al.*, 1997).

The advent of improved molecular techniques resulted in identification of more viruses such as *South African cassava mosaic virus*-[South Africa] (SACMV-[ZA]), (Berrie *et al.*, 1997; Berrie *et al.*, 1998), the Uganda variant of EACMV known as EACMV-UG has also been reported (Zhou *et al.*, 1997), which appears to be a recombinant virus with most of the coat protein gene of ACMV inserted in an EACMV-like DNA-A component. Variants of EACMV-UG have been isolated in Uganda and were designated EACMV-UG1, EACMV-UG2 and EACMV-UG (Pita *et al.*, 2001). In the Indian sub-continent another isolate was reported in Sri Lanka and was designated *Sri Lankan cassava mosaic virus* (SLCMV), which appears to be possibly a recombinant of a monopartite begomovirus and ICMV through acquisition of a DNA-B component from ICMV (Saunders *et al.*, 2002). Other examples of recombination in CMG's include, *East African cassava mosaic Zanzibar virus* (EACMZV) (Maruthi *et al.*, 2002), SACMV (Berrie *et al.*, 2001) and *East African cassava mosaic Malawi virus* (EACMMV) (Zhou *et al.*, 1998).

Continuous efforts to elucidate the biodiversity of CMG's reveal a complex picture of recombination and pseudo-recombination. Sequence comparisons of many more begomoviruses revealed that recombination is very common and has great implications for evolution of viruses. It has been shown that ACMV isolates show

greater nucleotide sequence similarity while EACMV isolates show considerable variation and recombination (Pita *et al.*, 2001, Fondong *et al.*, 2000, Berrie *et al.*, 2001, Zhou *et al.*, 1998).

#### **1.4.5. Satellite DNA's**

Subviral agents, viroids, and satellites are well known in infectious RNA systems (Mayo, *et al.*, 1995). The production of satellite DNA's is often associated with geminivirus infection (Stenger *et al.*, 1992, Frischmuth and Stanley, 1992). Satellites show little or no sequence similarity to viral or host genome and are completely dependent on a helper virus for their replication. Some satellites have been found to exacerbate viral symptoms or induce symptoms distinct from those induced by the helper virus alone (Ndunguru, *et al.*, 2006). However, a number of them interfere with helper virus replication and ameliorate disease expression, which has led to considerable interest into the investigation of their potential as sources of viral resistance (Dry, *et al.*, 1997).

#### **1.4.6. Symptoms**

Cassava plants infected with CMG's express a range of symptoms which depend on the virus species/strain, environmental conditions, and the sensitivity of the cassava host. The most typical symptoms consist of a yellow or pale green chlorotic mosaic of leaves, commonly accompanied by distortion and crumpling. Symptoms are readily distinguished from those of mineral deficiency or cassava green mite damage as the virus induced chlorosis and malformation of leaflets is asymmetrical about the midrib. Where symptoms are severe, the plant becomes generally stunted and petioles immediately below the shoot tip may be angled downwards. Occasionally, leaves may

become necrotic, shrivel, and absciss. Where the virus or virus strain is mild or the cassava variety is tolerant, leaf chlorosis may be patchy and absent on some leaves, and there is little or no leaf distortion or malformation and little effect on overall plant vigour.

## **1.5. Diagnostic methods**

### **1.5.1. Serological Tests**

Serological tests provided the first evidence of the diversity of cassava-infecting geminiviruses through use triple antibody sandwich (TAS) ELISA (Swanson and Harrison, 1994) as well as double antibody sandwich (DAS) ELISA (Sequeira and Harrison, 1982). These tests are conducted through the use of a small panel of monoclonal antibodies (MAbs). The antibodies include at least one that reacts with all viruses, one that reacts only with either ACMV or ICMV and one that reacts with ACMV and EACMV but not ICMV, and the three viruses can be distinguished reliably (Harrison *et al.*, 1997). Although these tests have been successfully used to distinguish between ACMV, EACMV and ICMV they failed to detect the causal agent of a severe epidemic in Uganda, which was later detected by use of DNA-based techniques (Zhou *et al.*, 1997). These monoclonal antibodies also fail to distinguish for example between SACMV and EACMV (Berrie *et al.*, 2001).

### **1.5.2. Nucleic Acid Based Tests**

More recently, DNA-based tests for CMG's have received increased attention. Although nucleic acid hybridisation tests are feasible (not for differentiation though), the increased specificity and greatly increased sensitivity of the polymerase chain

reaction (PCR) make it the preferred test in most instances. Evidence for the occurrence of recombinant CMG's (Zhou *et al.*, 1997), made it clear that DNA-based diagnostic techniques were required if reliable diagnosis is to be made. When devising DNA-based tests, one must bear in mind that different geminiviruses tend to share nucleotide sequences in various parts of their DNA-A molecules and that this tendency is much less in DNA-B (Harrison *et al.*, 1997). For differentiation of CMG's, PCR analysis can be based on shared or unique sequences.

Two PCR-based approaches have proved valuable in differentiating a wide range of begomoviruses. One approach relies on selecting primers based on nucleotide sequences that do not occur in other CMG's; thus only the target virus is detected (Deng *et al.*, 1994). For example, primers ACMV-AL1/F and ACMV-CP/R3 detect ACMV only, UV-AL1/F1 and ACMV-CP/R3 detect ACMV-UG only, UV-AL1/F1 and EACMV-CP/R detect EACMV only and UV-AL1/F1 and UV-AL1/R1 detect either EACMV-UG or EACMV (Zhou *et al.*, 1997).

In another approach, degenerate primers based on sequences occurring in several CMG's are used in PCR and the viruses are distinguished by the pattern of fragments obtained by restriction endonuclease treatment (restriction fragment length polymorphism) (RFLP) of the amplified DNA (Rojas *et al.*, 1993). RFLP involves digesting DNA with restriction enzymes, separating the resultant fragments by gel electrophoresis, blotting the fragments to a filter and hybridizing probes to the separated fragments (Lukman, 2003) or separating the fragments by agarose gel electrophoresis. Variation in restriction sites can therefore be detected as restriction fragment length polymorphisms.

Heteroduplex mobility assays (HMA's) have also been used for differentiating CMG's (Berry and Rey, 2001). This technique is sensitive and rapid and can detect

mixtures of viruses in field-grown cassava. However this technique has not been used extensively for CMG's and is limited by use of a single pair of primers that may not detect recombination of large DNA fragments in other regions of the genome. Therefore it may require the use a large range of reference samples for quicker and detailed identification of unknown isolates.

Of particular use for expanding diversity studies are group or virus-specific sequences that can be targeted by utilizing core coat protein CCP sequences, or additional conserved regions around which broad spectrum primers can be designed to target variable sequences in key open reading frames (ORF's) or non-coding regions. Prospective markers under exploration were selected with a basis in the most highly conserved viral ORF's, CP (AV1) and a portion of replication-associated protein (Rep) (AC1), and a key non-coding sequence that contains sufficient variability and/or virus-specific sequences which are consequently of potential epidemiological relevance (Brown, 2000). However the only begomovirus sequence formally approved by the International Committee on Taxonomy of Viruses (ICTV) for indicating begomovirus identity is the highly conserved CCP gene and its translation product (Fauquet and Stanley, 2003).

#### **1.6. The Whitefly Vector (*Bemisia tabaci*: Gennadius)**

The whitefly *B. tabaci* was first described as *Aleyorodes tabaci* Gennadius in 1889 from whiteflies collected from tobacco in Greece. Eleven years later another whitefly was described in the USA and was designated *Aleyorodes inconspicua* Quaintance (Quaintance, 1900). In 1914 *A. inconspicua* was moved to the genus *Bemisia*, giving rise to *Bemisia inconspicua* (Quaintance and Baker, 1914). Subsequently, over the

following fifty years or so numerous *Bemisia* isolates were described under various names from various countries and hosts. The overlap of the morphological characters that were used to classify whiteflies led to confusion in naming species of this genus, hence the subsequent synonymization of the various species, including *A. tabaci* into the *B. tabaci* species complex (Corbett, 1935; Russell, 1957; Danzig, 1966; Mound and Halsey, 1978). Today more than 1100 *Bemisia* species have been identified, however only three are known to be vectors of plant viruses, the most important being *B. tabaci*.

*Bemisia tabaci* belongs to the order Homoptera, family Aleyrodidae and is primarily a polyphagous insect that primarily colonizes annual herbaceous plants (Brown *et al.*, 1995). These whiteflies are about 2-3 mm in length and wings are present in the adult stages of both sexes. The abdomen lacks cornicles and the hind wings are nearly as long as the forewings (Bellows *et al.*, 1994). Most homopterans undergo gradual metamorphosis however the metamorphosis of whiteflies is different, showing a pattern more towards complete metamorphosis (Borror *et al.*, 1989). According to Borror *et al.* (1989) there are five instars in the development cycle of *B. tabaci* including the adult.

*Bemisia tabaci* has become one of the most important agricultural pests and virus vectors of agricultural and ornamental crops in all tropical, subtropical and some temperate areas. *Bemisia tabaci* causes significant damage to crops primarily through phloem feeding, phytotoxic disorders and the transmission of plant viruses. The development of insecticide resistance, reduction in natural enemies and monocultural practices have been considered as the main drivers in the emergence of *B. tabaci* as the primary agricultural pest in tropical and subtropical agricultural systems (Brown *et al.*, 1995).

### **1.6.1. CMG transmission by *B. tabaci***

*Bemisia tabaci* is the only known whitefly vector of cassava mosaic geminiviruses (CMG's) (Harrison, 1985). Since the late 1980's and early 1990's a number of CMD epidemics have been reported in sub-Saharan Africa, the most prominent being the East African epidemic that started in northern Uganda in the late 1980's and still spreading to surrounding countries (Otim-Nape *et al.*, 1997; Legg, 1999). The epidemic was caused by a recombinant virus known as the Uganda variant of *East African cassava mosaic virus* (UG) involving a recombination of DNA-A between *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV) (Zhou *et al.*, 1997). A similar form of severe CMD caused by a recombinant virus has been reported in Cameroon (Fondong *et al.*, 2000). One of the most prominent features about the Ugandan epidemic was the association with high population density of the whitefly vector, *B. tabaci* (Gibson *et al.*, 1996; Legg and Owai, 1998; Colvin *et al.*, 2004). Positive correlations have been observed between *B. tabaci* populations and CMD spread into initially healthy cassava plantings (Fargette *et al.*, 1993; Legg and Raya, 1998). The size of the whitefly populations has also been positively correlated with virus spread about one month after invasion, which corresponds with to the time necessary for symptom development (Fauquet and Fargette, 1990). These examples are therefore an indication that disease spread might be facilitated when a high population density of *B. tabaci*, feeds on plants containing a high virus titre and subsequently infects disease-free plants over a large area.

### **1.6.2. Taxonomy of the whitefly**

The taxonomy of the whitefly family, *Aleyrodidae*, has long been known to be problematic because morphological traits of adults do not readily permit

differentiation. However, some morphological features of immature forms, specifically those of the fourth instar or pupae have been used and as a result several species have been synonymized into a single taxonomic unit *Bemisia tabaci*, which included several of the so-called host races that exhibited obvious biological differences (Russell, 1958). Paramount to decisions to synonymize was the fact that the morphology of the *B. tabaci* fourth nymphal instar varied with host plant. Other biological differences include the ability to utilize host plants for feeding and reproductive purposes, viral transmission and the ability to induce phytotoxic disorders. In most cases where these differences occurred the morphological characteristics were indistinguishable (Russell, 1958).

### **1.6.3. Whitefly Biotypes and Host races**

The current international importance of *B. tabaci* has been attributed to the appearance and dispersal of a new biotype designated biotype B. This biotype has been considered a new species by some (Perring *et al.*, 1993) and was described as *B. argentifolii* (Bellows *et al.*, 1994). However its taxonomic status was subject to much debate and the general consensus is that *B. tabaci* is a species complex (Brown *et al.*, 1995). Due to the complexity of classifying whiteflies, the application of molecular techniques such as PCR and DNA probes as well as biochemical tools for determining banding patterns have been important in the taxonomic studies of *B. tabaci* (Perring, 2001).

Biochemical studies on whitefly protein polymorphisms were undertaken by several workers and resulted in distinguishing three whitefly species through esterase patterns (Prabhaker *et al.*, 1987), followed by Costa and Brown (1991) who also used esterases to distinguish host-associated populations derived from cotton. Perring *et al.*

(1993) examined allelic polymorphism among 14 enzymes from 17 colonies of either A or B biotypes. Based on esterase profiles, new biotypes were proposed and at present 20 of them, designated A-S are recognised (Brown *et al.*, 1995; Banks *et al.*, 1999). Although biochemical studies were useful in distinguishing whitefly populations, new PCR-based techniques are increasingly being used, providing better resolution of differences between whitefly populations and revealing polymorphisms.

The first DNA marker to be used to identify biotypes was random amplified polymorphic DNA PCR (RAPD-PCR) (Reiter *et al.*, 1992). This marker corroborated the esterase studies, yet simplified the experimental process for biotype identification. The RAPD-PCR technique was embraced due to the relatively high levels of polymorphism it reveals and its low cost compared to other techniques, such as allozymes and RFLP's.

The application of the sequences of the mitochondrial cytochrome oxidase I (mt COI) and the ribosomal internal transcribed spacer (ITS1) genes provided an entirely new perspective of *B. tabaci* phylogeny. The use of such sequence-based molecular informative markers that can be linked to geographic genotypes and/or biotic phenotypes constitute the most reliable approach for tracking the distribution and dispersion of *B. tabaci* (Brown, 2000; De Barro *et al.*, 2000; Abdullahi *et al.*, 2003). The use of mtCOI has been shown to reveal groups or clusters of *B. tabaci* with an overriding basis in geographic origin (Frohlich *et al.*, 1999; Brown, 2000; De Barro *et al.*, 2005; De La Rau *et al.*, 2006). Further support for geographic delineation of genotype clusters was provided by a number of other studies using the ITS1 gene sequence sometimes together with the mtCOI gene sequence (Frohlich *et al.*, 1999; Brown, 2000; De Barro *et al.*, 2000; De Barro *et al.*, 2005, De La Rau *et al.*, 2006). Despite the usefulness of these genetic markers there is still insufficient data to raise

racas to a species status. In fact De Barro *et al.*, (2005) provided evidence to disqualify the separation of *B. tabaci* and *B. argentifolii* and suggested the latter be considered within the *B. tabaci* species complex. However these markers give evidence of six well-supported phylogenetic clades or races and an unresolved core of ungrouped genotypes, with a clear geographic distribution at a continental level that fall within the *B. tabaci* species complex (De Barro *et al.*, 2005). Therefore to clarify the identity of the race to which the *B. tabaci* under investigation is known, the following nomenclature was suggested, *B. tabaci* (Asia), *B. tabaci* (Bali), *B. tabaci* (Australia), *B. tabaci* (sub-Saharan Africa), *B. tabaci* (Mediterranean/Asia Minor/Africa), and *B. tabaci* (New world) (De Barro *et al.*, 2005).

## **1.7. Objectives**

### **1.7.1. General Objectives**

The emergence of “new” begomoviruses, high frequency of virus genome recombination and the existence of whitefly genotypes contribute to epidemics in Africa. Some studies have been carried out in southern Africa to establish the diversity and extent of cassava begomoviruses and their whitefly vectors (Berry *et al.*, 2004). This study is an extension of such studies with particular emphasis in South Africa and is part of a larger study that aims to eventually screen for and select and/or develop cassava cultivars that are resistant to infection by cassava mosaic viruses and distribute them to resource poor farmers in South Africa. The aim of this project was to undertake a comprehensive epidemiology study of CMD in cassava cultivation areas in South Africa and to establish the genetic diversity of the viruses and vectors present on cassava in South Africa, thus identifying the problem with regards to CMD and its vectors in South Africa. The results of this study will impact the type of

genetic-based strategies that are needed in order to develop CMD-resistant cassava cultivars.

### **1.7.2. Specific Aims**

1. Epidemiological studies of cassava mosaic disease (CMD) in South Africa.
  - a) Monitoring the incidence and severity of CMD
  - b) Collect whitefly abundance data
2.
  - a) To investigate the genotype and distribution of cassava-infecting begomoviruses in South Africa using CCP-PCR, differential primers and RFLP's
  - b) To investigate possible recombination and occurrence of satellite DNA molecules associated with cassava begomoviruses
3. To investigate the genotype and distribution of the whitefly *B. tabaci* in South Africa using the mitochondrial cytochrome oxydase I (mtCOI) gene as a molecular marker

## **CHAPTER 2: THE EPIDEMIOLOGY OF CASSAVA MOSAIC DISEASE**

## 2.1. ABSTRACT

Cassava mosaic disease (CMD) is caused by whitefly-transmitted geminiviruses and poses a serious threat to cassava production in Africa. The disease is transmitted by *Bemisia tabaci* and disseminated in stem cuttings. To investigate the incidence and severity of CMD and its whitefly vector, surveys were conducted in 2004/2005 and 2005/2006 seasons in Bushbuckridge, Mariti and Tonga (Mpumalanga). Plants were assessed for CMD incidence, severity, infection type (cutting/whitefly-borne) and whiteflies numbers. Results were analyzed by one-way ANOVA t-test. The results indicated that cutting-borne CMD was a major source of infections compared to whitefly-Borne CMD. Average whitefly numbers were lowest in Bushbuckridge (2.1) compared to Mariti (2.7) and Tonga (4.3) and may have contributed to higher disease incidences there. There was little variation in disease severity (ranging from 2.4 to 2.8) for all regions and all surveys. The results showed that the CMD was stable over the two seasons. These results therefore indicate that cutting-borne infections should be of greatest concern in all regions and cultural practices such as phytosanitation could help a great deal. Whitefly-borne infection is nonetheless very important in disease spread more so if cassava cultivation is intensified.

## 2.2. INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is one of the most important food crops in Africa. It is grown in most countries of sub-Saharan Africa including Madagascar. Since its introduction from Latin America into Africa by Portuguese traders in the late 16<sup>th</sup> century (Fauquet and Fargette, 1990), cassava has become very important as a food security crop and for poverty alleviation in Africa. Today, cassava is grown on an estimated 17 million hectares in 34 African countries (FAO, 2006). The bulk of cassava produced in Africa is for human consumption and is consumed fresh or in various processed forms.

Cassava was introduced from Mozambique into South Africa during the major tribal movements of the mid-1800s subsequently spreading into Mpumalanga (formerly Eastern Transvaal), Swaziland and Northern KwaZulu-Natal (Daphne 1980). Due to low winter temperatures in most parts of South Africa, cassava cultivation is confined to these areas, where temperatures are not as low as they are in other parts of the country. This together with taste preference for maize means that cassava has remained a minor crop in South Africa, grown predominantly by small scale farmers on average field sizes of 0.05-0.25ha and is grown as a secondary staple food and/or for sale locally. However there is growing interest in South Africa to produce cassava for industrial purposes. Already there are commercial cassava farms in Barberton in Mpumalanga and Dendron in the Limpopo province that feed a starch processing factory originally situated in Dendron and now moved to Barberton.

Cassava yields in South Africa, like many other countries in the African continent, are consistently low, ranging from 7.8t/ha to 15t/ha (Mathews, 2000), this is mainly due to cassava pests and diseases. Cassava mosaic disease (CMD) is the

most important disease affecting cassava production in Africa. The disease is caused by a group of begomoviruses, which belong to the family *Geminiviridae* (Harrison, 1985). These viruses are characterized by a bi-partite single stranded DNA (ssDNA) genome, transmission by *B. tabaci* and twinned isometric particle morphology (Harrison 1985; Harrison *et al.* 1977). CMD symptoms are easily recognized by the appearance of characteristic leaf mosaic. The most severe symptoms result in stunting of the plant and extreme reduction of the leaf surface area with consequent reduction in root yield.

Cassava mosaic geminiviruses (CMG's) are transmitted in a consistent manner by the whitefly *B. tabaci* (Storey and Nichols, 1938; Dubern, 1994; Brown *et al.*, 1995). However the primary source of CMD dissemination is through the stem cuttings used as planting material. Most of the planting materials in South Africa are obtained locally or are bought from Mozambican migrant workers. The repeated use of local planting material, most of which is already infected, could result in the deterioration of crop quality. Furthermore the supply of planting materials from neighboring countries could lead to introduction of new virus stains, which may possibly be more virulent.

Previous studies have shown the importance of various factors that influence the pattern of spread of virus disease within and between fields and the factors that inhibit or favor such spread. In 1990, Fauquet and Fargette (1990) reported that disease incidence largely reflects fluctuations in whitefly populations which partly depend on climatic factors, including temperature, rainfall and wind. A number of studies have suggested high temperature as the primary factor driving the increase in whitefly populations (Fauquet *et al.*, 1985, Fargette *et al.*, 1993), however this may not be the case where drought limits plant growth or in the event of an epidemic such as the one

that occurred in Uganda in the early 1990s where the epidemic was observed to be spreading rapidly into somewhat cooler areas (Legg and Ogwal, 1998). Higher rainfall and humidity have also been positively correlated with higher disease incidence that results from higher whitefly populations that are supported by vigorous plant growth (Dengel, 1981; Robertson, 1985; Fargette and Thresh, 1994). Legg and Raya (1998) showed that in Tanzania, regions with the highest incidence were hot, wet coastal areas as well as drier inland areas moderated by neighboring lakes. Wind speed and direction influence the distribution of the whitefly population in a field and it has been shown that the incidence of the disease was higher on the upwind edges than on the downwind edges of the field (Fauquet and Fargette, 1990).

Synergism between CMG's has been reported by several workers and is particularly important in influencing disease severity (Harrison *et al.*, 1997; Fondong *et al.*, 2000; Pita *et al.*, 2001). These studies have shown that mixed infections (eg. ACMV/EACMV-UG) result in more severe symptoms than single infections. This phenomenon is of primary importance for the emergence of new geminivirus diseases and has been shown to be a key factor in the genesis and spread of the CMD epidemic in East Africa that started in Uganda (Harrison *et al.*, 1997; Legg, 1999). Other factors that influence disease spread are cassava varieties used, proximity of other fields or source of inoculum, crop density and virus strains present. It is important to note though, that the interactions between these factors are complex and should be treated in a manner that takes local conditions into consideration.

There have been three studies conducted in South Africa on the epidemiology of CMD (Berry and Rey, 2001; Jericho *et al.*, 1999; Trench and Martin, 1985). Trench and Martin (1985) only went as far as to record the presence of CMD in various areas. The study was limited to irrigation schemes and did not give a comprehensive

overview of the disease situation. The significance and relevance of that study however, was that it confirmed the presence of CMD in South Africa then, and had shown that the principal mode of CMG transmission in South Africa was through stem cuttings. A second study conducted by Berry and Rey (2001) was essentially a survey of the viruses present and their genetic variation, but did not include any epidemiological data. A third study by Jericho *et al.* (1999) provided the first quantitative record of CMD in South Africa in terms of incidence and severity and also went further to assess other cassava pests. This study was also limited by having a small sample size, being done in a single season and not including whitefly data.

The growing realization of the commercial value of cassava in South Africa could lead to intensification of cassava cultivation. This will require a thorough understanding of the disease situation in order to guard against possible epidemics and other problems that may arise from such intensification. The aim of this study therefore was to thoroughly investigate the epidemiology of CMD by monitoring incidence and severity and collecting whitefly abundance data.

## **2.3 MATERIALS AND METHODS**

### **2.3.1. Geographical Location**

During the November to March 2004/2005 and 2005/2006 growing seasons, surveys were conducted in Mpumalanga and Limpopo. The surveys were conducted in the areas of Bushbuckridge, Mariti and Tonga (figure 2.1). Bushbuckridge and Mariti are close to one another separated by a low mountain escarpment and together they share a radius of approximately 50km. Tonga is over 200km away by road from Bushbuckridge and Mariti and is further separated from these areas by the Kruger

National Park. These three areas are situated in a subtropical climate zone with an annual rainfall of approximately 800mm with the rainy season starting in September and continuing to May. The agro-ecology of the three areas varies, with the Bushbuckridge area characterized by an annual rainfall between 500 and 800mm with a landscape of slightly undulating plains, grayish and excessively drained soil that contain less than 15% clay. In contrast the Mariti area which lies adjacent to Bushbuckridge is characterized by low mountain ranges, a higher annual rainfall of between 800 mm and 1300 mm and the soil is red or yellowish, massive or weak structured with more than 35% clay content. Whereas the landscape and rainfall pattern in Tonga are similar to that of Bushbuckridge, the soil type is different and is characterized by marked clay accumulation, strongly structured and reddish in color with 15-35% clay content. The Tonga area lies in the Nkomati river valley, which contributes to higher humidity (<http://www.environment.gov.za/>)

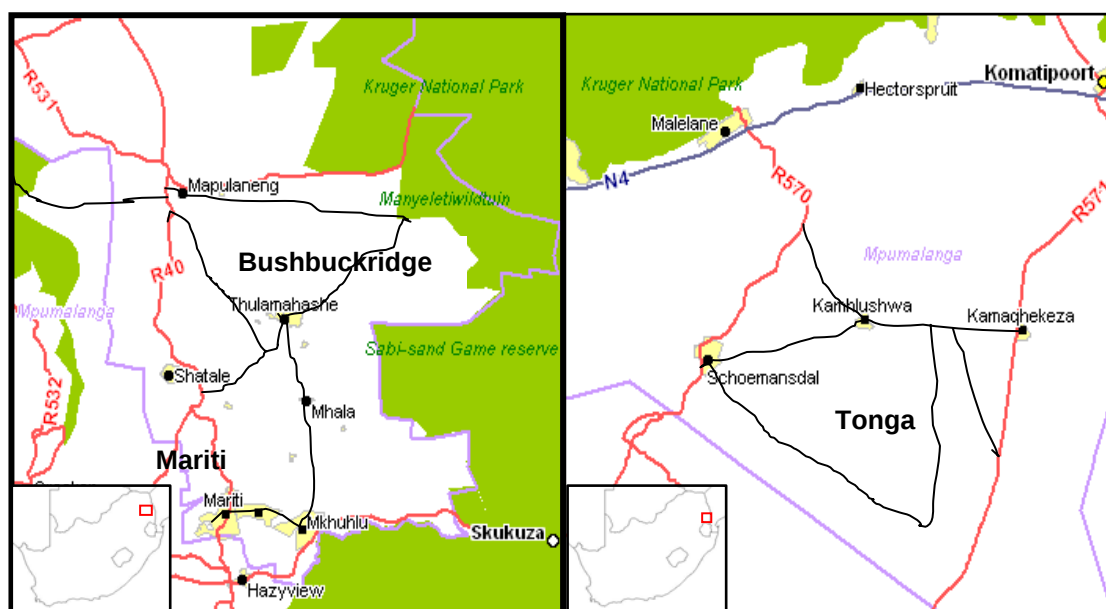


Figure 2.1 Cassava growing areas in Mpumalanga and Limpopo in which surveys were conducted (inserts represent position of respective areas in South Africa ([http://www.southafrica.info/ess\\_info/sa\\_glance/geography/](http://www.southafrica.info/ess_info/sa_glance/geography/))).

### 2.3.2. Survey records

The surveys were conducted over two growing seasons, although the initial plan was to conduct three surveys per growing season (October, January and March) coinciding with the beginning to the end of the rainy season. Due to unpredictable weather patterns and drought, most farmers had not planted in October, hence a small sample size was available for the October survey. Only two surveys (January and March 2005) in the 2004/2005 season had a satisfactory sample size to be included. In the 2005/2006 season, reasonable rains only fell in January, thus only one survey was conducted in February 2006 where a reasonable sample size was obtained. For statistical purposes only the January 2005 and February 2006 seasons were compared since cassava plants were more or less of the same age (i.e. 3-4 months old).

In each region, 10-15 fields were randomly selected within a minimum proximity of approximately 5km. Thirty plants were assessed along a diagonal or 'Z' configuration in each field. Fields that were too narrow were assessed along a diagonal and in fields that had fewer than 30 plants all plants were assessed. The following parameters were measured during observations: a) **Disease incidence**: number of plants diseased relative to the total number of plants assessed, [i.e. Incidence % = [(number of plants with symptoms/total number of plants assessed) x 100]]. b) **Disease severity**: Area or volume of plant tissue that is diseased relative to the total area or volume. It is normally expressed using a scale that indicates the extent of symptom development. The scale of 1 to 5 was used, 1 indicating no symptoms and 5 the most severe symptoms with leaf distortion and stunting of plants (Sseruwagi *et al.*, 2004). Disease severity (DS) takes into account the degree of symptom development in diseased plants only (DS = Average disease severity score/number of symptomatic plants). c) **Type of infection**: for each infected plant

assessed, it was also indicated whether it was cutting-borne (C) (older bottom leaves are symptomatic) or whitefly-borne (W) infection (new topmost leaves show symptoms). d) **Adult whitefly population**: the whitefly population was assessed on each sampled plant by counting the average number of adult whitefly on the five topmost leaves.

### **2.3.3. Data analysis**

Comparisons were made between mean CMD incidences, whitefly numbers and severity scores for different regions and seasons using One-way ANOVA (Non-parametric tests). All percentage data (total incidence, Cutting-borne and whitefly-borne incidences) were arcsine-transformed prior to analysis. Disease severity data was analyzed using the chi-squared test. Whitefly numbers were log-transformed before analysis. The analysis was done using Graphpad InStat 5® (version 3.06).

## **2.4. RESULTS**

### **2.4.2. CMD incidence**

Results for total CMD incidence(C + W) recorded during the two surveys are presented in figure 2.2. Average CMD incidences recorded during 2004/2005 season were 27.4% in Bushbuckridge, 39.8% in Mariti and 52.5% in Tonga and not significantly different from one another ( $P>0.05$ ). In the 2005/2006 season total CMD incidence was not significantly different between Tonga (45.1%) and Mariti (35.9%;  $P>0.05$ ,  $t=0.838$ ) and between Mariti and Bushbuckridge (24.1%;  $P>0.05$ ,  $t=1.959$ ). It was however significantly different between Tonga and Bushbuckridge ( $P<0.05$ ,  $t=2.505$ ).

Cutting-borne and whitefly-borne CMD incidences are shown in figure 2.3. Cutting-borne CMD incidences in 2004/2005 and 2005/2006 seasons were not significantly different from one another (Tonga, 25.8%, 23.2%,  $t = 1.559$ ; Mariti, 27.5%, 25.2%,  $t = 0.2766$ ; Bushbuckridge, 26.6%, 34.0,  $t = 0.05977$  respectively). Whitefly-borne disease incidences were not significantly different between Tonga (12.3%) and Mariti (26.7%,  $P > 0.05$ ,  $t = 0.7310$ ). However, in Bushbuckridge (0.9%) whitefly-borne CMD incidence was significantly lower than incidences in both Tonga ( $P < 0.05$ ,  $t = 3.359$ ) and Mariti ( $P < 0.05$ ,  $t = 3.503$ ) during 2004/2005. The same trend was observed in 2005/2006 where there was no significant difference in cutting-borne incidence between Tonga (11.1%) and Mariti (10.7%,  $P > 0.05$ ,  $t = 0.7195$ ). Again incidence in Bushbuckridge was significantly lower than those of Tonga ( $P < 0.05$ ,  $t = 2.251$ ) and Mariti ( $P < 0.05$ ,  $t = 3.667$ ) during 2005/2006.

Comparison of cutting-borne and whitefly-borne disease incidence revealed that in Tonga there was no significant difference between the two (25.8% and 26.7% respectively,  $P > 0.05$ ,  $t = 1.008$ ) during 2004/2005. Significant differences in whitefly-borne and cutting-borne incidences were observed in Mariti (C=27.5%, W=12.3%;  $P < 0.05$ ,  $t = 2.829$ ) and Bushbuckridge (C=26.5%, W=0.9%;  $P < 0.05$ ,  $t = 3.580$ ). In the 2005/2006 season whitefly –borne incidences were significantly lower than cutting-borne incidences all the surveyed areas (Tonga,  $P < 0.05$ ,  $t = 4.256$ ; Mariti,  $P < 0.05$ ,  $t = 3.119$ ; Bushbuckridge,  $P < 0.05$ ,  $t = 5.025$ ).

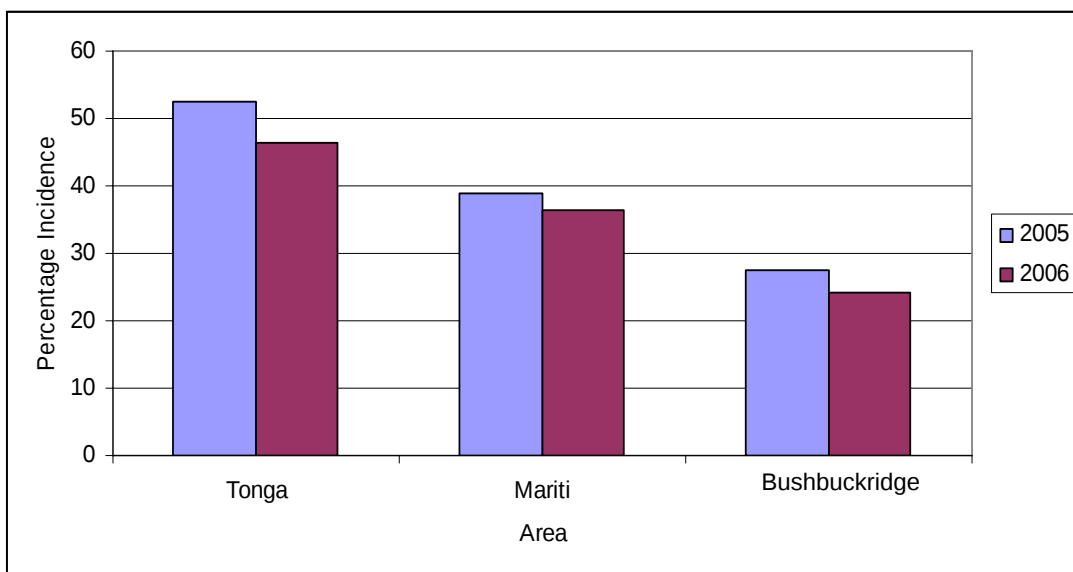


Figure. 2.2. Cassava mosaic disease incidence during 2004/2005 and 2005/2006 growing seasons in Mpumalanga (Tonga and Mariti) and Limpopo (Bushbuckridge) Provinces of South Africa

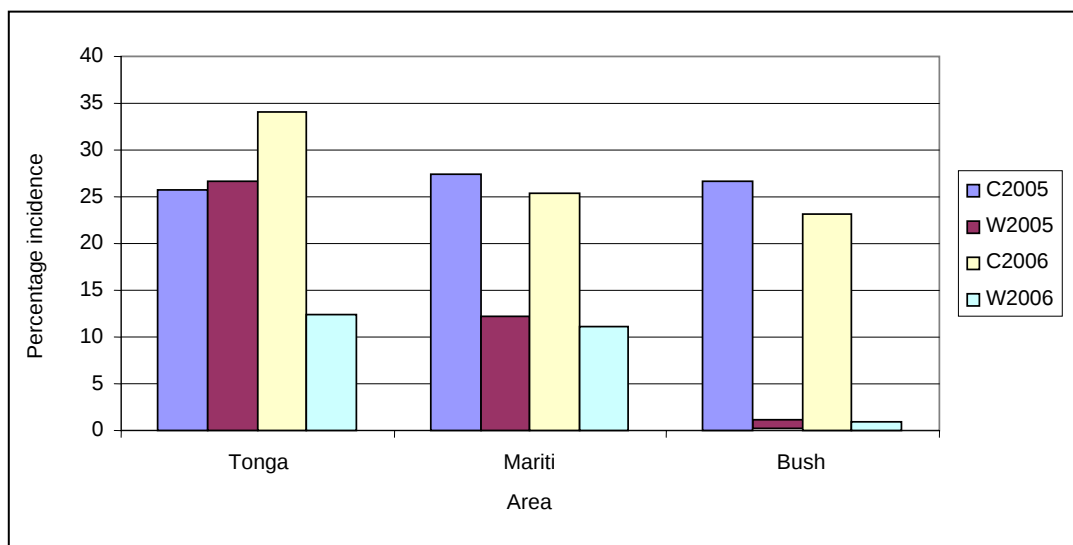


Figure 2.3 Percentage cutting (C) and whitefly (W) CMD incidence for 2004/2005 and 2005/2006 growing seasons in Mpumalanga (Tonga and Mariti) and Limpopo (Bushbuckridge) Provinces of South Africa

#### 2.4.3. Disease severity

Variations in disease severity amongst regions remained insignificant throughout the surveys, ranging from 2.37 in Bushbuckridge to 2.61 and 2.78 in Mariti and Tonga respectively. However, slightly lower but insignificant disease severity was observed

during the March 2005 surveys in Bushbuckridge (2.25), Mariti (2.51) and Tonga (2.61) (figure 2.3).

Table 2.1 Summary of survey data collected during 2004/2005 and 2005/2006. The following parameters were recorded: disease incidence (cutting and whitefly-borne), disease severity and whitefly numbers.

|                     | Bushbuckridge    |                 | Mariti          |                 | Tonga           |                   |
|---------------------|------------------|-----------------|-----------------|-----------------|-----------------|-------------------|
|                     | Jan-05           | Feb-06          | Jan-05          | Feb-06          | Jan-05          | Feb-06            |
| <b>Total Inc. %</b> | 27.4a $\pm$ 10.9 | 24.1a $\pm$ 7.7 | 39.8a $\pm$ 4.8 | 35.9a $\pm$ 4.8 | 52.5a $\pm$ 9.8 | 45.1a $\pm$ 7.0   |
| <b>C. Inf. %</b>    | 26.6a $\pm$ 11.1 | 23.2a $\pm$ 7.9 | 27.5a $\pm$ 3.9 | 25.2a $\pm$ 3.5 | 25.8a $\pm$ 4.2 | 34.0a $\pm$ 5.4   |
| <b>Wf. Inf. %</b>   | 0.9a $\pm$ 0.6   | 0.9a $\pm$ 0.5  | 12.3b $\pm$ 3.4 | 10.7b $\pm$ 3.0 | 26.7b $\pm$ 4.6 | 11.1b( $\pm$ 3.5) |
| <b>Sev. (ave)</b>   | 2.4a $\pm$ 0.2   | 2.5a $\pm$ 0.2  | 2.6a $\pm$ 0.3  | 2.7a $\pm$ 0.2  | 2.8a $\pm$ 0.2  | 2.5a $\pm$ 0.3    |
| <b>Wf. No.(ave)</b> | 0a               | 4.1b $\pm$ 1.2  | 2.3ab $\pm$ 1.1 | 3.6b $\pm$ 0.8  | 1.2a $\pm$ 0.3  | 7.3b $\pm$ 1.2    |

a, b represent values that are not significantly different from one another in each row, +/- # = SEM.

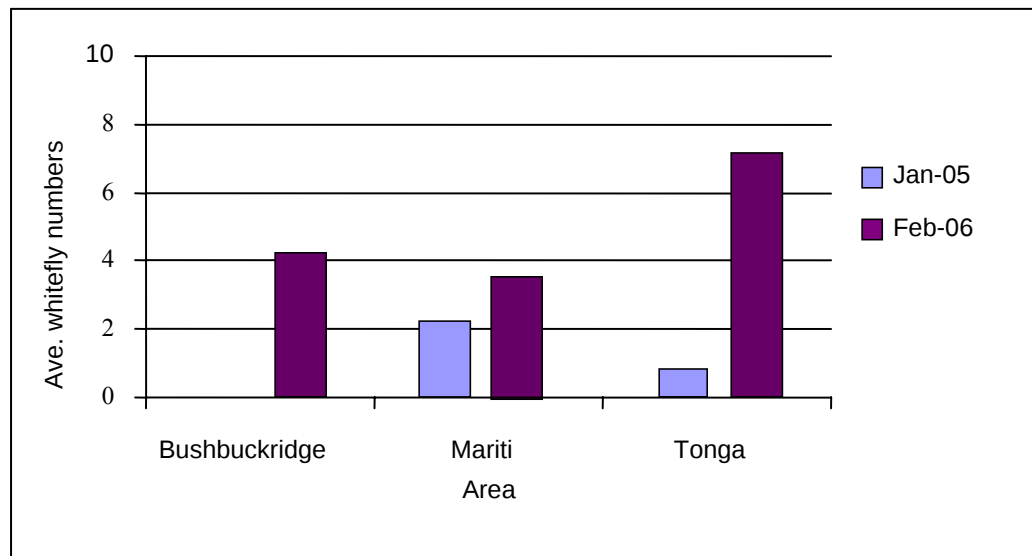


Figure 2.4 Average whitefly numbers observed during surveys in 2004/2005 and 2005/2006 growing seasons in the regions of Mpumalanga and Limpopo provinces in South Africa.

### 2.3.3. Whitefly Abundance

Whitefly abundance data is shown in figure 2.4. Average numbers of adult whitefly on the five topmost leaves ranged from 0 to 7.3. No whiteflies were observed in Bushbuckridge during 2004/2005. In Tonga, there was a significant difference in

whitefly numbers between 2004/2005 and 2005/2006 (1.9, 7.3 respectively,  $P>0.05$ ,  $t=3.594$ ). No significant difference in whitefly numbers was observed in Mariti between 2004/2005 (2.3) and 2005/2006 (3.6,  $P>0.05$ ,  $t=0.7346$ ). Whitefly numbers were significantly higher in Tonga (7.3) during 2005/2006 than in Mariti (3.6,  $P>0.05$ ,  $t=2.54$ ), but not significantly different than numbers in Bushbuckridge ( $P>0.05$ ,  $t=0.4158$ ).

## 2.5. DISCUSSION AND CONCLUSIONS

The most significant limitation of the current study was the drought that prevailed in the survey areas during the 2005/2006 season; hence variations in the disease situation within a season at different growth stages may have been missed. These variations may include phenomena such as symptom reversion and change in whitefly populations and the effect these have on CMD spread.

The results obtained from the current study showed no significant variation in disease incidence between the two seasons, thus indicating a stable disease situation over 2004/2005 and 2005/2006. These results are comparable with those of Jericho *et al.* (1999) who showed the lowest CMD incidence as 13% (Table 3 in Jericho *et al.*, 1999) in Bushbuckridge (Northern Province/Limpopo) in 1998; in comparison to incidences in Tonga (Mpumalanga) where CMD incidence (63%) was significantly higher.

The principal mode of infection was shown to be through cuttings since the incidence of cutting infection was much higher than whitefly-borne infections in Mariti and Bushbuckridge except for Tonga (25.8%, 26.7% respectively) in 2004/2005. Furthermore, the results of this study indicate that there were no

significant differences in cutting-borne infections between all the surveyed areas as well as between the two seasons in which the surveys were conducted. The fact that cutting-borne disease incidence was not significantly different between the survey areas and seasons, and was higher than whitefly-borne incidence, further emphasizes the fact that the primary source of infection is through cuttings. This observation is consistent with various studies conducted elsewhere in the African continent indicating that cutting infection is the primary source of the virus (Trench and Martin, 1985; Legg and Ogwal, 1998; Legg and Raya, 1998; Jericho *et al.*, 1999).

Average whitefly-borne disease incidences were very low (10.4%) as compared to cutting incidences (27.1%) amongst all survey areas and between the seasons in South Africa (see figure 2.2 and table 2.1). This observation is consistent with other studies in sub-Saharan Africa, for example in Tanzania a countrywide survey indicated whitefly borne incidence of 3.6% compared to 27% due to cuttings (Legg and Raya, 1998). Other studies that show a similar trend were conducted in Mozambique, Rwanda, Kenya and the Democratic Republic of Congo (DRC) (Legg, *et al.*, 2001; Thresh, 2001; Munga and Thresh, 2002; Okao-Okuja *et al.*, 2004). The exception to this trend was in the epidemic areas of Uganda and neighboring countries such as Rwanda and the DRC where the epidemic is believed to be spreading; up to 80% whitefly-borne disease incidence was reported at the epidemic front in Uganda (Legg and Ogwal, 1998) and up to 83% in some districts in Rwanda (Legg, *et al.*, 2001). The epidemic front in Uganda and elsewhere in East Africa was characterized by very high whitefly numbers and the occurrence of the Ugandan strain of EACMV (EACMV-UG) which is associated with the CMD epidemic. This emphasizes the importance of whiteflies as vectors of CMD especially in areas of intense cultivation with little pest control interventions. In Mariti and Tonga whitefly numbers were

significantly higher than in Bushbuckridge during the 2005 survey, however numbers improved in 2006 probably due to better rainfall. Therefore it was not surprising to see a relatively higher CMD incidence in these areas. These variations in whitefly numbers can be explained by differences in climatic conditions. Climatic data indicates that the areas of Tonga and Mariti have a relatively higher rainfall that results in vigorous plant growth, thus supporting higher whitefly populations. The fact that wetter weather supports CMD spread compared to drier weather has been established as early as 1936 by Storey (1936) as well as Doughty (1958) who indicated that CMD was more prevalent in coastal areas.

Disease severity has been shown to be generally low (2.4 to 2.8) in South Africa compared to other parts of Africa, especially in the CMD epidemic areas of East Africa where severity scores of over 3.0 are frequent. The highest scores were recorded in epidemic areas where in most cases disease severity was above 4.0 (Legg *et al.*, 2001). A number of factors play a role in disease severity; these include cultivar, virus strain/species, rainfall and the quality of soil. In South Africa the low disease severity can be mainly attributed to virus species found. During this study it has been shown that the predominant virus species is ACMV (chapter 3), which is known to induce milder symptoms than EACMV (Fondong *et al.*, 2000; Maruthi *et al.*, 2005). Symptoms are even more severe when double infections occur (Harrison *et al.*, 1997; Fondong *et al.*, 2000; Pita *et al.*, 2001). Therefore the degree of symptom severity shown here is to be expected considering that ACMV is the predominant species in South Africa. Another factor in disease severity is occurrence of mixed infections, which were shown to be common in South Africa (Chapter 3, Berry and Rey, 2001). It was therefore surprising that these mixed infections were not accompanied by higher disease severity. A possible explanation is that even though

some fields showed very high disease severity there were many that showed mild symptoms thus the average was lower. Furthermore researcher bias could have played a role since there could be a tendency to collect leaves with the most visible symptoms which have a better chance of having mixed infections.

However higher disease incidence in Tonga and Mariti can be attributed to higher whitefly numbers in these areas as compared to Bushbuckridge. The low whitefly numbers in Bushbuckridge (2004/2005) can be attributed to poor agro-ecological conditions such as low rainfall, air humidity and excessively drained soils and dry conditions that prevailed in the area in 2004/2005.

The fact that in these areas there is only one predominant cultivar (locally known as Munyaca), the continuous use of this cultivar could pose a threat to the crop should a more virulent strain or species emerge due to recombination or introduced into the area unless interventions in the form of introduction of resistant cultivars and phytosanitation are practiced. Phytosanitation would be effective in controlling CMD, but it may become ineffective if cassava cultivation is to be intensified leading to abundance of whiteflies. This was evident in the commercial farm in Barberton where some fields were entirely infected and whiteflies were in large numbers (personal observation). There will be a need to encourage the use of resistant cultivars.

**CHAPTER 3: BIODIVERSITY OF CASSAVA MOSAIC  
GEMINIVIRUSES IN SOUTH AFRICA**

### 3.1. ABSTRACT

In the past decade a growing number of cassava mosaic geminiviruses (CMG) have been described and shown to be more diverse than previously thought, mainly due to recombination. Field-collected cassava leaves were analyzed by restriction length polymorphisms (RFLP) (using *EcoRV* *MluI* and *PstI*) and differential primers. Both RFLP's and differential primers indicated that ACMV was the predominant and mixed infections were common (43%). No SACMV was detected in any of the samples. Four isolates representing four different RFLP patterns were sequenced and analyzed for phylogenetic relationships and recombination. Three of the four isolates were ACMV. One of the ACMV isolated was shown to be a 'new' strain of ACMV by phylogenetic analysis. The ACMV isolates had a recombination pattern similar to previously identified ACMV isolates. The fourth CMG isolate was closely related to two isolates (EACMMV-[MW:MH:96] with over 95% homology and had similar recombination patterns. The occurrence of a unique strain of ACMV raises concerns for the diversity of CMG's, availability of a larger gene pool than previously thought with implications for recombination and synergistic interactions and the emergence of new strains and species.

### 3.2. INTRODUCTION

Cassava mosaic disease is considered the most important disease affecting cassava in Africa (Fargette *et al.*, 1988). The disease is caused by a group of begomoviruses that belong to the family *Geminiviridae*. Cassava mosaic geminiviruses (CMG's) are characterized by bi-partite circular single-stranded DNA (ssDNA) genomes encapsidated in a twinned (germinate) particle of approximately  $20 \times 30\text{nm}$  (Zhang *et al.*, 2001; Bottcher *et al.*, 2004) and transmission by *Bemisia tabaci*. CMG's have genomes that consist of two components, DNA-A and DNA-B (Stanley and Gay, 1983; Stanley *et al.*, 2005). Together these components contain 8 open reading frames (ORF's), six on DNA-A and two on DNA-B (Stanley and Townsend, 1986). The two components share a conserved intergenic 'common region' (IR or CR) of approximately 200bp in size and has about 90-100% sequence similarity between DNA-A and DNA-B. DNA-A encodes the coat protein (AV1, CP), replication associated protein (AC1, Rep) and proteins associated with movement (AV2), transactivation of AV1 and BV1 (AC2, TrAP) and replication enhancement (AC3, REn) and the transcriptional activator (AC4) (Hanley-Bowdoin *et al.*, 2004; Vanitharani *et al.*, 2004). DNA-B encodes for the nuclear shuttle protein (NSP, BV1) and the Movement protein (MP, BC1) responsible for virus movement within and between cells (Hanley-Bowdoin *et al.*, 2004).

Six distinct CMG species have been identified and associated with cassava in Africa so far, these are; *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *East African cassava mosaic Cameroon virus*-Cameron (EACMCV-CM[CM:98]), *East African cassava mosaic Malawi virus*-Malawi (EACMMV-[MW/MH:96]), *East African cassava mosaic Zanzibar virus*-[Tanzania:Uguja:1998] (EACMZV-[TZ:Ugu:98]) and *South African cassava mosaic*

*virus* (SACMV-[ZA]) (Fauquet and Stanley, 2003). Prior to 1994 only ACMV and EACMV were known to infect cassava in Africa and were thought to be limited to specific geographical areas, whereby ACMV was thought to occur only in West Africa and EACMV in East Africa (Swanson and Harrison, 1994). However improved diagnostic techniques and the advent of PCR have resulted in better understanding and appreciation of the complexity of their distribution. Over the past thirteen years several strains and/or species have been identified in different regions of the African continent. To date several studies have shown the presence of ACMV in all parts of the continent where cassava is grown and EACMV is now found in West Africa as well. For example Fondong *et al.* (2000) isolated EACMCV-CM[CM:98] in Cameroon. Pita *et al.* (1999) did the same in Ivory Coast. In southern Africa Berrie *et al.* (1997) isolated SACMV in South Africa, which shares a high nucleotide sequence similarity with EACMV. SACMV has also been found in Madagascar and Zimbabwe (Ranomenjanahary *et al.*, 2002; Briddon *et al.*, 2003). Berry and Rey (2001) have shown evidence of the occurrence of several CMG species (i.e. ACMV, ACMV-UG, EACMV and SACMV) in six Southern African countries (i.e. South Africa, Swaziland, Mozambique, Angola, Zambia and Zimbabwe).

One of the most important milestones in CMG's diagnostics was the identification of the Ugandan variant of EACMV-UG which was responsible for the CMD epidemic that almost wiped out an entire crop in Uganda during the early 1990's (Zhou *et al.*, 1997). It was in this study that evidence of recombination was first presented. A number of subsequent studies have since established the importance of recombination and/or pseudorecombination as the driving force behind *Geminivirus* evolution and biodiversity (Padidam *et al.*, 1999; Fondong *et al.*, 2000; Pita *et al.*, 2001; Ndunguru *et al.*, 2005; Bull *et al.*, 2006).

The use of PCR has made it possible to use other diagnostic techniques such as restriction fragment length polymorphisms (RFLP's) which allowed for rapid preliminary identification of CMG's. The RFLP technique has been used successfully in a number of countries in Africa (Okao-Okuja *et al.*, 2004; Sseruwagi *et al.*, 2004; Ndunguru *et al.*, 2005; Rothenstein *et al.*, 2006). However, high levels of diversity in a virus population could present problems as shown by Bull *et al.* (2006), where a single nucleotide gain or loss in a restriction site could result in misidentification. RFLP's therefore can be useful for broader screening of viruses for preliminary identification; hence it is always desirable to use RFLP's together with sequencing. The use of PCR has also allowed identification of satellite DNA's which may be important in symptom development (Ndunguru *et al.*, 2005).

Recently a new technique called rolling circle amplification (RCA) has been shown to be easier and cheaper than PCR and antibody detection, and allows reliable diagnosis of geminiviruses (Haible *et al.*, 2006). Used together with Restriction Fragment Length Polymorphism (RFLP) this technique is able to distinguish virus species up to strain level without any sequence information. However to obtain further information such as possible recombination and phylogenetic relatedness of the viruses it is necessary to obtain virus sequences. Because RCA products can be sequenced directly, RCA does not require the use of a thermocycler and circumvents cloning making the procedure much cheaper thus making it accessible for resource poor laboratories and it could be very useful in future studies. Sequence information thus obtained can be used to elucidate the genetic structure of populations and change thereof providing information to breeding and genetic engineering programs. Sequencing of *Geminivirus* genomes has allowed comparisons of large numbers of geminivirus sequences and the revelation of a complex picture of recombination and

psuedorecombination. It was also very clear from such comparisons that EACMV and EACMV-like viruses show a high degree of variation and recombination, meanwhile ACMV isolates showed a greater degree of nucleotide sequence similarity regardless of where they came from (Zhou et al. 1998; Fondong *et al.*, 2000; Berrie *et al.*, 2001; Pita *et al.*, 2001).

The geographical overlap of *Geminivirus* distribution throughout cassava-growing areas, aided by trafficking of planting materials across borders resulting from population movements and trade, provides opportunities for synergism between CMG's. A number of synergistic interactions have been reported amongst CMG's. Such synergistic interactions were first reported in Uganda and neighboring countries where mixed infection of ACMV and EACMV resulted in a severe form of CMD (Harrison *et al.*, 1997; Legg, 1999; Pita *et al.*, 2001). Fondong *et al.* (2000) observed a similar interaction between ACMV and EACMCV in Cameroon. Interestingly this is the only case known for geminiviruses and for plant viruses belonging to the same family and was shown to be a key factor in the genesis and spread of the CMD epidemic in Uganda and surrounding East and Central African countries (Harrison *et al.*, 1997; Legg, 1999). The presence of satellite DNA's in a virus population together with frequent recombination and synergism could further complicate the epidemiology of geminiviruses resulting in epidemics with severe symptom development. Geminiviruses are often associated with satellite DNA molecules, which play a major role in symptom development (Stenger *et al.*, 1992, Frischmuth and Stanley, 1992). Recently, some monopartite begomoviruses, such as *Ageratum yellow vein virus* (AYVV) and *Cotton leaf curl Multan virus* (CLCuMV), have been shown to be associated with a satellite-like molecule, referred to as DNA- $\beta$  (Briddon

*et al.*, 2003; Saunders *et al.*, 2004). Ndunguru *et al.* (2006) reported satellite DNA's associated with EACMV in Tanzania.

Jericho *et al.* (1999) and Berry and Rey (2001) reported the first studies on the diversity of CMG's in South Africa. It was clear from the studies that CMG's were diverse in South Africa, with a total of four different species found and mixed infections a common occurrence. The study reported the occurrence of EACMV-UG in 52% of the samples tested whereas 43% of those were mixed infections of EACMV-UG and either SACMV, ACMV or EACMV, the remaining 9% contained EACMV-UG alone. This situation presents an opportunity for recombination and synergistic interactions among viruses, therefore the virus population could be more diverse than is realized. If this is the case there could be a potentially epidemic situation, especially if the current interest in cassava in South Africa leads to intensification and extensive cultivation of cassava there is therefore a need to intensify research efforts in order to establish the extent of the problem. This requires that virus identities and diversities are established in order to devise appropriate virus control strategies. The aim of the current study therefore was to investigate the genotype and distribution of CMG's in South Africa using CCP-PCR, differential primers PCR and RFLP's; to investigate possible recombination on selected field isolates and to look at the possible occurrence of satellite DNA's in CMG infected field samples.

### **3.3. MATERIALS AND METHODS**

#### **3.3.1. Virus sources and PCR amplification**

Symptomatic cassava leaf samples were collected from all the fields visited in Bushbuckridge, Tonga and Mariti during the 2004 and 2006 surveys (chapter 2).

Although the commercial farm (CSM) in Barberton and Makhathini flats in Kwazulu-Natal were not part of the survey, samples were also collected there. A total of 157 samples were collected and stored on ice at  $\pm 4^{\circ}\text{C}$  until they were brought to the laboratory and stored at  $-20^{\circ}\text{C}$ . The same samples were also smeared on FTA® cards (Whatman). Total nucleic acid (TNA) was extracted from the leaf samples according to the method of Doyle and Doyle (1987).

All the samples were screened for whitefly-transmitted geminiviruses using the CCP primers AV514 and AC1048 (Wyatt and Brown, 1996). Taq DNA polymerase was used to amplify the  $\pm 550\text{bp}$  fragment using a BioPad® thermocycler. Cycling conditions were initiated at  $95^{\circ}\text{C}$  for 5 min, followed by 35 cycles at  $95^{\circ}\text{C}$  for 1 min,  $55^{\circ}\text{C}$  for 1 min,  $72^{\circ}\text{C}$  for 1 min and a final extension at  $72^{\circ}\text{C}$  for 10 min. Both TNA's and FTA card samples were used for PCR amplification. Preliminary identity of the viruses was determined by use of the differential primers JSP001, JSP002 and JSP003. Primer pair JSP001 and JSP002 amplifies the coat protein gene of ACMV, while JSP001 and JSP003 amplify that of EACMV (Fondong *et al.*, 2000).

Near full length virus DNA-A's (c. 2760-2780 bp) of all the samples were also PCR-amplified using the universal primers Uni/F (5'KSGGGTCGACGTCATCAAGACTTTRTAC 3') and Uni/R (5'AARGAATTCATKGGGGCCCARRGACTGGC 3') (Bridson and Markham, 1994). These primers anneal to the common region of geminivirus DNA-A. PCR was performed with Accuzyme® polymerase (Bioline) and the cycling conditions were; first cycle at  $94^{\circ}\text{C}$  for 2 min, 30 cycles at  $94^{\circ}\text{C}$  for 40s,  $56^{\circ}\text{C}$  for 1 min,  $72^{\circ}\text{C}$  for 4 min and a final extension at  $72^{\circ}\text{C}$  for 10 min.

### **3.3.2. Restriction Fragment Length Polymorphisms (RFLP's)**

Restriction Fragment Length Polymorphisms (RFLP's) was performed on a total of 49 samples. The 49 samples were those from which full length DNA-A was obtained. Thirty-five of those are shown in figure 3. RFLP analysis was carried out using the restriction enzymes *EcoRV* and *MluI* (Promega) at 37<sup>0</sup>C for 3 hours. Because these enzymes cannot distinguish between EACMV and SACMV, the enzyme *PstI* (37<sup>0</sup>C for 3 hours) was also used to distinguish these two viruses.

### **3.3.3. Cloning, Sequencing and Phylogenetic Analysis**

The RLFP analysis resulted in four different restriction patterns. To ascertain the results obtained with the RFLP technique, one representative isolate per restriction pattern was picked for cloning and subsequent sequence analysis. The isolates (figure 3.5) were Number 19 (SA Mariti), 21 (SA Barb), 24 (SA Tonga) and 29 (SA Bush). Near full-length DNA was obtained using the universal primers Uni/F and Uni/R. The PCR products were cloned into a pcrSMART™ (Lucigen) vector using the pcrSMART™ cloning kit and transformed into E. cloni® 10G (*E. coli* strain) chemically competent cells as per manufactures instructions. The clones were sent to Inqaba Biotech (Ltd.) for sequencing. Consensus sequences were obtained using the computer program ChromasPro V1.33. A BLAST of the sequences was done to compare with published CMG sequences using the basic alignment search tools (BLASTN) in the NCBI website. Sequences were aligned with the ClustalW option in the MEGA 3.1 computer program. The aligned sequences were used to generate a phylogenetic tree using the neighbor-joining and bootstrap options in MEGA 3.1. (Kumar *et al.* 2004)

### **3.3.4. Recombination Analysis**

The four cloned sequences were analyzed for recombination using the recombination detection program (RDP2) (Martin and Rybicki, 2000). The RDP2 program examines nucleotide sequence alignments and attempts to identify recombinant sequences and recombination breakpoints using 10 published recombination detection methods, including geneconv, bootscan, maximum  $\chi^2$ , chimaera and sister scanning. It allows fast automated analysis of large alignments and interactive exploration, management and verification of results with different recombination detection and tree drawing methods. The four sequences and over 90 other geminivirus sequences obtained from Genebank were aligned in DNAMAN demo version 6 (Lynnon Corporation) and loaded into the RDP program for analysis.

### **3.3.5. Satellite DNA's**

To determine the presence of satellite DNA's in field samples TNAs were extracted by the method of Doyle and Doyle (1987). These were subjected to PCR using abutting primer pair B01 (5'-GGTACCACTACGCAGCAGCAGCC-3') and B02 (5'-GGTACCTACCCTCCCAGGGGTACACAC-3') which amplify a 1.03kb product (Ndunguru *et al.*, 2006). From the 49 samples that were analyzed by RFLP's, all those that tested positive for EACMV were checked for the presence of satellite DNA's. These were selected because so far satellite DNA has only been shown to be associated with EACMV (Ndunguru *et al.*, 2006).

### 3.4. RESULTS

#### 3.4.1 PCR Amplification and RFLP's

Over 90% of the 157 samples analyzed by CCP-PCR tested positive for the whitefly-transmitted geminiviruses. An electrophoresis gel picture of ten of the samples is shown in figure 3.1, with the expected PCR product of approximately 550bp. For preliminary identification of the CMG's differential primers were used to distinguish between EACMV and ACMV and the amplification products are shown on figure 3.3. Ninety three percent of the samples tested positive for either ACMV or EACMV. Of these, 43% were mixed infections of ACMV and EACMV meanwhile 36% contained ACMV alone and 14% had EACMV only.

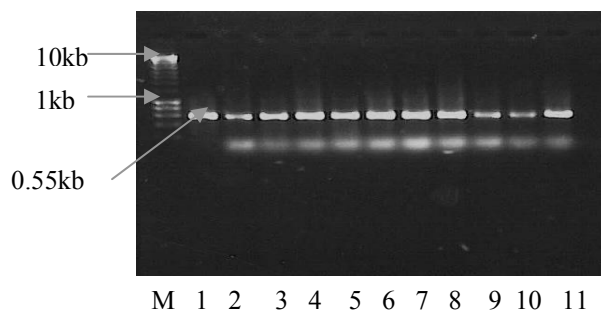


Figure 3.1. PCR using core coat protein (CCP) primers of some of the samples collected during January 2005 in Mpumalanga

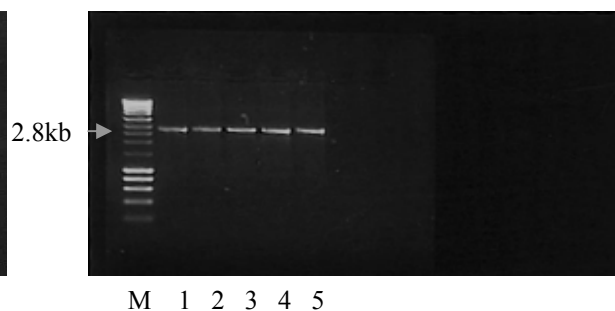


Figure 3.2. Near Full length DNA-A amplified with universal PCR primers UNI/F and UNI/R. Lane 2 is the positive control

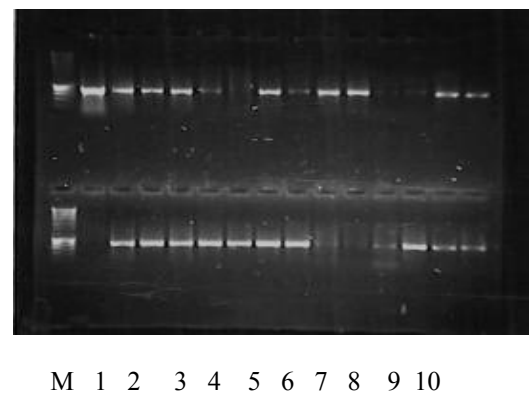
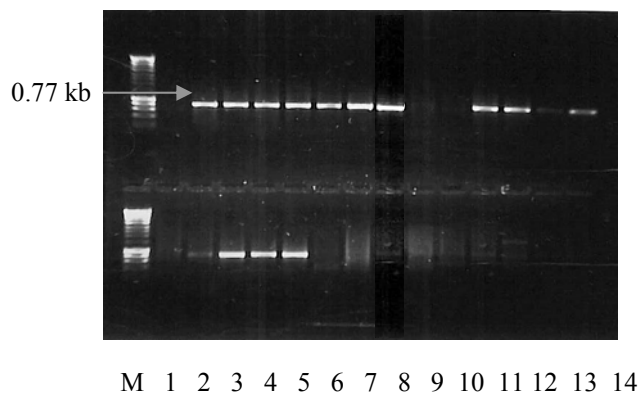


Figure 3.3 A and B. PCR products obtained using differential primers JSP001, JSP002 (top row) and JS001, JSP003 (bottom row) to distinguish between ACMV and EACMV isolates respectively.

A total of 49 samples were randomly selected from all the areas for RFLP. Of the 49 samples that were RFLP analyzed 16 came from Tonga, 11 from Bushbuckridge, 12 from Mariti, 6 from Barberton and 4 from Makhathini flats. These samples were subjected to PCR using the universal primers Uni/F and Uni/R to obtain near full-length DNA-A. Figure, 3.2 shows a gel electrophoresis picture of the near full length DNA-A of five of the samples (four were subsequently cloned and sequenced). The near full-length DNA-A's thus obtained were RFLP analyzed, and the restriction patterns segregated into four groups for both enzymes (i.e. *EcoRV* and *MluI*) (figure 3.4). Using predicted restriction patterns (table 3.1) for these enzymes on known CMG's, the four groups were presumed to be A: ACMV-[NG/CI/CM/DO2]; B: EACMV-[KE/MW/TZ,]; C: ACMV-[UGMld]/[UGSvr]/[KE] and D unknown (uncut). The 'uncut' DNA-A was later identified as an ACMV strain after sequencing. Virus mixtures were also observed comprising mixtures of groups AB, BC and BD. Virus mixtures were not as many as revealed by differential primers: 14.8% had virus mixtures when analyzed with RFLP's versus 43% when using differential primers. The percentage distribution of the viruses as revealed by RFLP's is shown on table 3.2. The use of the enzyme *PstI* did not reveal the presence of SACMV (results not shown). ACMV was predominant in Mariti (75%) and EACMV in Bushbuckridge (83.3%). ACMV-UG was only detected in Mariti (occurring in single infections) and Tonga (mixture with ACMV). The four samples obtained from Makhathini flats all contained ACMV.

Table 3.1. Predicted RFLP's for published cassava mosaic geminivirus sequences following computer-based digestion of full-length DNA-A with the restriction enzymes *EcoRV*, *MluI* and *PstI*

| Enzyme       | Fragment | ACMV[NG]/<br>[CI]/[CM]/<br>[CM/DO2] | ACMV-<br>UGMId/UG2S<br>vr | SACMV | EACMV[KE2B]<br>/[MW]/[TZ]/<br>EACMCV/[CI] |
|--------------|----------|-------------------------------------|---------------------------|-------|-------------------------------------------|
| <i>EcoRV</i> | 1        | 1.48                                | 1.48                      | 2.19  | 2.19                                      |
|              | 2        | 1.28                                | 1.28                      | 0.59  | 0.59                                      |
| <i>MluI</i>  | 1        | -                                   | 1.55                      | 1.21  | 1.21                                      |
|              | 2        | -                                   | 1.21                      | 0.67  | 0.67                                      |
|              | 3        | -                                   | -                         | 0.52  | 0.52                                      |
| <i>PstI</i>  | 1        | -                                   | -                         | 2.52  | -                                         |
|              | 2        | -                                   | -                         | 0.28  | -                                         |

Table 3.2. Percentage of viruses identified in each region using RFLP's

|                      | ACMV        | ACMV-UG     | EACMV        | ACMV-<br>UG/<br>EACMV | ACMV/<br>ACMV-UG | ACMV/<br>EACMV |
|----------------------|-------------|-------------|--------------|-----------------------|------------------|----------------|
| Barberton            | 42.80       | 0.00        | 14.30        | 14.30                 | 0.00             | 28.60          |
| Bushbuckridge        | 16.70       | 0.00        | 83.30        | 0.00                  | 0.00             | 0.00           |
| Makhathini           | 100.00      | 0.00        | 0.00         | 0.00                  | 0.00             | 0.00           |
| Mariti               | 75.00       | 12.50       | 12.50        | 0.00                  | 0.00             | 0.00           |
| Tonga                | 50.00       | 6.25        | 12.50        | 0.00                  | 6.25             | 25.00          |
| <b>Total Average</b> | <b>56.9</b> | <b>3.75</b> | <b>24.52</b> | <b>2.86</b>           | <b>1.25</b>      | <b>10.72</b>   |

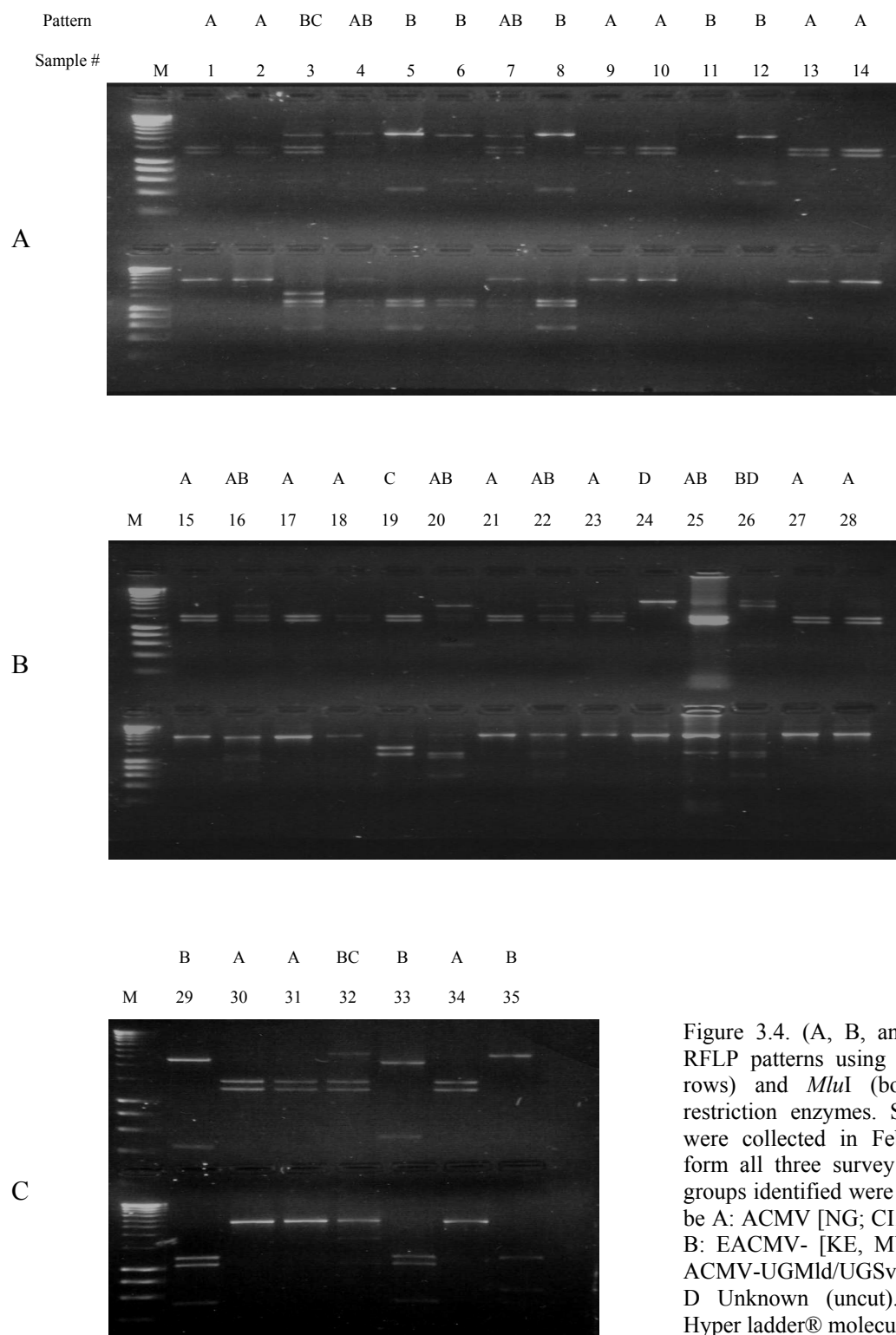


Figure 3.4. (A, B, and C). PCR-RFLP patterns using *EcoRV* (top rows) and *MluI* (bottom rows) restriction enzymes. Sample 1-35 were collected in February 2006 from all three survey areas. Four groups identified were presumed to be A: ACMV [NG; CI; CM; DO2], B: EACMV- [KE, MW, TZ,]; C: ACMV-UGMld/UGSvr/[KE] and D Unknown (uncut). M is the Hyper ladder® molecular marker.

### **3.4.2. Phylogenetic relationships and recombination analysis**

The near full length sequences of four representative CMG's were determined. The DNA-A components ranged in size from 2765-2790 bp and the genetic organization of the ORF's was similar to other begomoviruses. The DNA-A components contained all six open reading frames (ORF's), two in the virion sense (AV1 and AV2) and four in the complementary sense (AC1, AC2, AC3 and AC4).

Three of the four CMG sequences shared a high nucleotide sequence similarity with ACMV isolates compared to EACMV, with nucleotide sequence identity ranging from 92% to 98%. Of the three ACMV-like isolates, one designated SA-Mariti had an RFLP restriction pattern similar to ACMV UG/Mld/Svr with which it shared just over 95% sequence identity. It also shared over 95% sequence identity with a number of West African isolates. However it was most closely related to isolate SA-Barb with 98 % sequence identity. The third ACMV-like isolate designated SA-Tonga was not cut by the two restriction enzymes used, and shared less than 94% sequence identity with the other ACMV isolates identified elsewhere in Africa. However it shared 94.6% and 95.8% with the other two ACMV-like South African isolates SA-Barb and SA-Mariti respectively. As expected the three ACMV isolates grouped together in a 100 bootstrap supported branch with other previously identified ACMV isolates from Africa (figure 3.5). The three South African isolates formed a tight 85 bootstrap supported sub-cluster. The fourth sequenced virus isolate, designated SA-Bush, clustered together with all other EACMV isolates and was shown to be most closely related to EACMMV-[MK/MH] from Malawi sharing a sequence identity of 95%. It shared 78% to 88% sequence identity with the over 50 EACMV isolates analyzed. It also shared quite a high overall sequence identity (86%) with SACMV compared to some EACMV's and other non EACMV's.

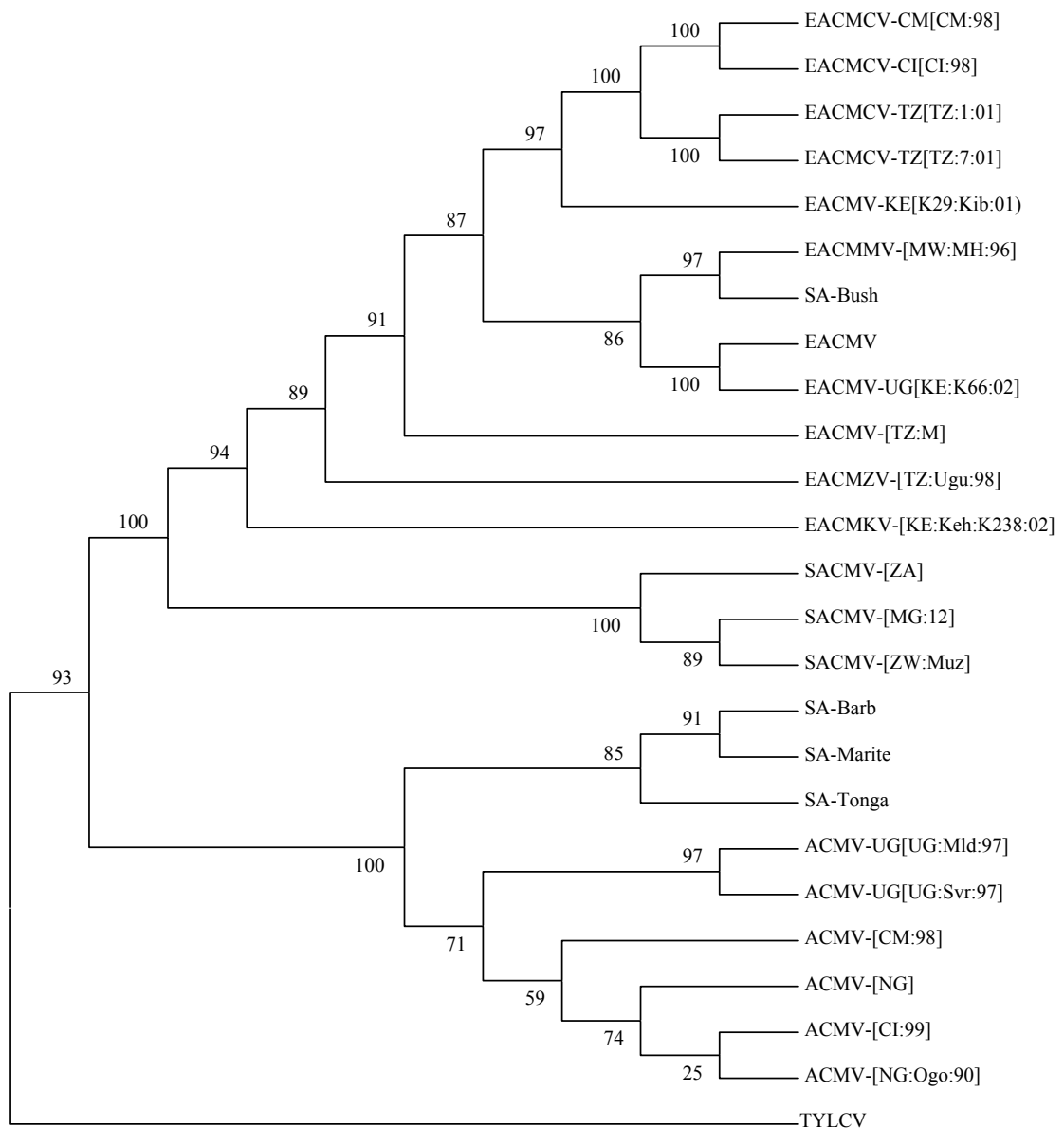


Figure 3.5 Phylogenetic tree showing near full length DNA-A sequence relationships between four cassava mosaic geminivirus isolate from South Africa (SA-Barb, SA-Mariti, SA-Tonga and SA-Bush) and other cassava mosaic geminiviruses. *Tomato yellow leaf curl virus*-TYLCV was used as an out-group. Abbreviations are: (EACMCV-CM[CM:98]) – *East African cassava mosaic Cameroon virus*, EACMCV-CI[CI:98]-*East African cassava mosaic Cameroon virus* [Ivory Coast:98], EACMCV-TZ[TZ:1:01]-*East African cassava mosaic Cameroon virus* [Tanzania:1:2001], EACMCV-TZ[TZ:7:01]-*East African cassava mosaic Cameroon virus* [Tanzania:7:2001], EACMV-KE[K29:Kib:01]-*East African cassava mosaic virus* [Kenya:Kiboni:K29:2001], EACMV-[TZ:M]-*East African cassava mosaic Kenya virus* [Tanzania:M], EACMV-UG[KE:Mum:K66] - *East African cassava mosaic virus*-[Uganda:Kenya:Mumias:K66:2002], EACMZV-[TZ:Ugu:98]-*East African cassava mosaic Zanzibar virus*-[Tanzania:Uguju:1998], EACMKV-[KE:Keh:K238:02]- *East African cassava mosaic Kenya virus*-[Kenya:Kehala:2002], EACMMV [MW:MH:96] - *East African cassava mosaic Malawi virus* – [Malawi:Maha:96], SACMV-[ZA] - *South African cassava mosaic virus*-[South Africa], SACMV [M:12] - *South African cassava mosaic virus* – [Madagascar:12], ACMV-UG[UG:Svr:97] - *African cassava mosaic virus*-Uganda severe, ACMV-UG[UG:Mld:97] - *African cassava mosaic virus*-Uganda mild, ACMV-[CM:98] – *African cassava mosaic virus*-[Cameroon:1998], ACMV-[NG:Ogo:90] – *African cassava mosaic virus*-[Nigeria:Ogo:90]

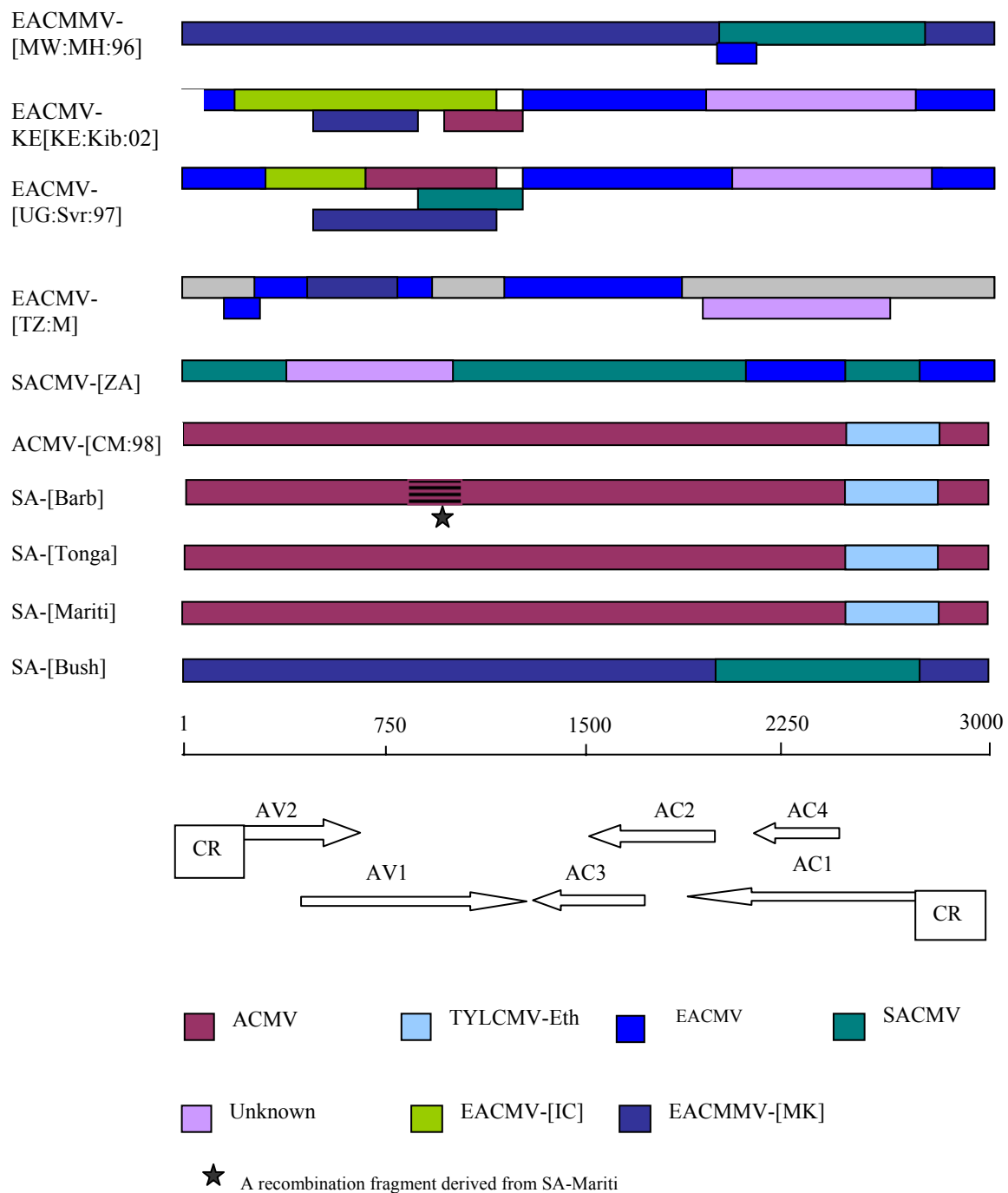


Figure 3.6 Recombination map of putative recombinant fragments for DNA-A of four isolates of cassava mosaic geminiviruses (CMG's) from South Africa together with selected previously identified CMG's. Each horizontal line represents the genotype of each isolate and the colour shades boxes represent the tentative origins of the putative recombinant fragments. Just below the diagram is the length of the genomes and the genome organization. The colour code for the recombinant fragments is indicated at the bottom of the diagram.

Previous studies of ACMV isolates have shown very little or no recombination compared to EACMV and SACMV (Ndunguru *et al.* 2005). In the current analysis

using the RDP program the three ACMV and 10 other previously identified ACMV isolates analyzed had one putative recombination event occurring between nucleotides 2290 and 2600 (figure 3.6). This fragment is of unknown origin but the most probable major parent is *Tomato yellow leaf curl Mali virus*-Ethiopia (TYLCMV-Ethiopia) or a common ancestor. Isolate SA-Barb had a second putative recombination event whose origin is SA-Mariti or a common ancestor. The event occurs between nucleotides 725 and 906 and the fragment is 99% identical to the corresponding fragment in SA-Mariti.

Recombination analysis, as expected, showed a higher degree of recombination with EACMV isolates compared to ACMV isolates with most of the isolates having more than four putative recombination events. The only exceptions were EACMMV-[MW:MH:96] and our isolate SA-Bush which had two and one putative recombination events respectively. The two isolates both had one recombination event spanning nucleotides 1824 to 2540. This fragment is common amongst a number of other EACMV's from East Africa and the tentative major parent is SACMV. EACMMV-[MW:MH:96] had a small additional putative recombination event spanning nucleotides 1836 to 1927.

### **3.4.3. Satellite DNA's**

The use of PCR with primers B01 and B02 to determine the presence of the 1.5kb product showed no amplification product in all the samples.

### 3.5. DISCUSSION AND CONCLUSIONS

Several populations of cassava mosaic viruses have been reported to exist in southern Africa namely; ACMV-UG, ACMV, EACMV and SACMV (Berry and Rey, 2001). These were revealed with CCP sequence comparisons. Although the CCP is accepted by the ICTV for preliminary identification of geminiviruses, complete DNA-A sequence are more appropriate as it has been shown to support well known species demarcation (Pita *et al.*, 2001). In this study CCP-PCR was used to confirm the presence of CMG's in infected field samples. Preliminary identification was achieved by use of differential primers and RFLP's. In addition to these methods, near full length sequences of four South African isolates were determined and compared to other CMG isolates known so far by phylogenetic and recombination analysis.

As expected, CCP-PCR confirmed the presence of CMG's in over 90% of the symptomatic field samples tested. Those that did not show a positive result could be as a result of leaf tissue deterioration over time. Both RFLP's and differential primers showed that ACMV is the predominant virus species in the three sampled areas of Mariti, Bushbuckridge and Tonga in the Mpumalanga province, followed by EACMV. Virus mixtures were a common occurrence with 14.8% shown by RFLP's to be mixtures and even more (43%) by differential primers. There is a very high discrepancy in samples shown to have virus mixtures between the RFLP and differential primer methods. This can be explained by the fact that the PCR product of differential primers is shorter (~750bp) than the universal primer's product (~2790bp), thus the former is more likely to be amplified even at lower concentrations. Virus mixtures were more common in Tonga and Barberton than in Mariti and Bushbuckridge. Tonga (chapter 1) and Barberton (personal observation)

are characterized by higher numbers of whiteflies, thus providing opportunities for super-infections. In Barberton the situation is made worse by the intensity of commercial cultivation, coupled with irrigation, resulting in vigorous plant growth hence providing a good breeding ground for whiteflies. The high incidence of virus mixtures in these areas is a cause for concern since it provides opportunities for synergistic interactions and recombination amongst the different viruses. Berry and Rey (2001) have also shown a high degree of mixed infections (43%) in South Africa including Tonga and Mariti (Hoxane) as well as St Lucia and Makhathini flats in KwaZulu-Natal. This is consistent with the percentage of mixed infections shown in this study by differential primers.

EACMV was the most common virus in Bushbuckridge and the only other virus detected was ACMV. This is contrary to the results shown for the other surveyed areas (Mariti, Tonga and Makhathini flats) where ACMV was predominant possible explanation could be as a result of low samples size in Bushbuckridge. Moreover cassava plantations are much more sparsely distributed in Bushbuckridge than the other areas, particularly Mariti which is close by. It is highly unlikely therefore that the situation in Bushbuckridge could be a true reflection of virus distribution because given the proximity of Bushbuckridge and Mariti exchange of planting materials is common, hence the virus distribution would not be expected to differ significantly. Low sample size in Makhathini flat could also be the reason why only ACMV was detected there since Berry and Rey (2001) has shown the presence of EACMV as well.

The use of RFLP's has revealed the presence of predominantly three CMG's namely; ACMV, ACMV-UG and EACMV at the sampled locations in South Africa. SACMV was not detected in any of the samples tested. This is consistent with

findings by Berry and Rey, (2001) who have shown the presence of the same CMG species in South Africa, whereas they only detected SACMV once in Swaziland. However SACMV has been reported in Madagascar and Zimbabwe, but the prevalence is not known (Briddon *et al.*, 2003; Ranomenjanahary *et al.*, 2002). It is clear now that SACMV is very rare in South Africa.

Two of the three ACMV-like South African isolates (SA-Barb and SA-Mariti) were more closely related to ACMV isolates from other African countries, sharing just over 95% sequence identity. This is expected of ACMV isolates, which are homogeneous regardless of geographical origin (Pita *et al.*, 2001, Fondong *et al.*, 2000; Ndunguru *et al.*, 2005; Bull *et al.*, 2006). However isolate SA-Tonga shared less than 94% nucleotide sequence similarity with other ACMV isolates except the two South African isolates (SA-Mariti and SA-Barb) with which it shares 94.7% and 95.8% respectively. Furthermore, SA-Tonga was not cut by any of the enzymes used in RFLP analysis. Further restriction analysis of its sequence in DNAMAN® confirmed that it had no restriction sites for both *EcoRV* and *MluI* enzymes. Therefore it can be concluded that SA-Tonga is a distinct strain of ACMV, which clearly showed two polymorphic sites not found in either SA-Barb or SA-Mariti. As expected recombination analysis showed no significant difference in recombination pattern of SA-Tonga with SA-Barb and SA-Mariti. SA-Mariti and SA-Barb virus isolates also showed a similar recombination pattern with other ACMV isolates. SA-Barb however has a unique recombination event in the AV1 region. This event contains a fragment of 181bp in length between nucleotides 725 and 906; it is ACMV-like and is derived from SA-Mariti. This may explain why these two isolates share a much higher nucleotide sequence similarity compared to other ACMV isolates, given that this fragment shares 99% identity with the corresponding fragment

in SA-Mariti. The RDP program uses 10 different methods of recombination analysis and all but two show this event to be significant (Martin and Rybicki, 2000). It is interesting that this is the only other event to occur in ACMV isolates in the current RDP analysis. If this event is indeed a result of recombination and not an artefact, together with the occurrence of a unique recombinant strain (SA-Tonga) of ACMV in South Africa, then contrary to previous studies suggesting that ACMV is a homogeneous group, this is an indication that ACMV could be as capable of recombination as EACMV isolates, but possibly not as frequently. However recombination analysis using other programs did not show any recombination in ACMV, for example Ndunguru *et al.*, (2005) did not detect any recombination in ACMV using the pair-wise comparison sequence analysis (PCSA) method which compares the profile of a pair of sequences to that of an average profile of sequences that are selected *a priori*.

The fourth South African isolate designated SA-Bush was shown by phylogenetic analysis to be most closely related to two EACMV isolates from Malawi, namely EACMV-MH/MK with which it shares 95% nucleotide sequence similarity. This is not surprising since isolates closely related to these Malawian isolates have been identified in South Africa in a previous study (Berry and Rey, 2001). Therefore this study serves to confirm their presence. Recombination analyses of this EACMV isolate shows it has a putative recombination event spanning the AC4 and AC1 region. This event is also found in many other EACMV isolates and is derived from SACMV. SA-Bush shares a similar recombination pattern with the two Malawian isolates and together these three isolates show fewer recombination events as compared to other EACMV isolates, most of which have in excess of four putative recombination events. No satellite DNA's were detected in any of the samples tested,

however their presence cannot be ruled out because there were no positive control samples available at the time of the study. Therefore it is possible that future well controlled studies may show their presence.

In conclusion, the virus diversity in South Africa has not changed since the time of the study by Berry and Rey (2001). The presence of the 'new' ACMV strain warrants continuous investigations to keep in check the virus diversity and possible emergence of new virus strains and species. If SA-Bush is the predominant EACMV genotype in South Africa and showing less recombination activity, the possibility is that South Africa is at the 'bottom' of Africa and more geographically isolated, thus less infection by a variety of strains and opportunity for recombination. However the opening of South Africa's political and economic barriers since the early 1990s with increased movement of people across borders could present an opportunity for introduction of new virus strains and species. This is another reason why continuous monitoring is required.

**CHAPTER 4: PHYLOGENETIC ANALYSIS OF *BEMISIA*  
*TABACI* COLONIZING CASSAVA IN SOUTH AFRICA USING  
THE MITOCHONDRIAL CYTOCHROME OXIDASE I (mtCOI)  
GENE**

#### 4.4. ABSTRACT

*Bemisia tabaci* is the only known vector of begomoviruses. The occurrence of *B. tabaci* populations with no clear-cut diagnostic differences in morphology but exhibiting differences in biological properties makes their classification difficult. The mitochondrial oxidase I (mtCOI) gene sequences have been useful in distinguishing genotype clusters of *B. tabaci* mainly based on geographical boundaries. In this study the mtCOI gene fragment of field-collected *B. tabaci* was cloned and sequenced to determine the phylogenetic relationships of cassava infesting *B. tabaci* populations in the Mpumalanga and Limpopo provinces of South Africa. Phylogenetic analysis indicated that most of the *B. tabaci* populations in South Africa are grouped together with a previously identified southern African clade from cassava samples collected in Zambia, South Africa and Mozambique. Three isolates (SAS 2.5, SAS 24 and SAS 2.6) formed a distinct cluster sharing 97 to 98 % identity with each other and 84% to 94% with the southern African clade forming a 'new' separate cluster of haplotypes within the southern African sister clade. Four isolates from Mariti clustered with the Mediterranean/North African Q types forming a sister clade with previously identified isolates from Zimbabwe. This is the first time this biotype has been reported in South Africa. Four others (SAS27.1-Mariti, SAS37.2-Tonga, SAS41.3-Tonga and SAS60.3-Tonga) were non *Bemisia* species and were probably casual feeders. In conclusion, cassava-infesting whitefly populations in South Africa are monophyletic. The occurrence of haplotypes or sister clades has implications for gene flow between different genotypes, however it remains to be seen if their host is indeed cassava. The occurrence of the Q biotype is cause for concern for other crops in South Africa.

#### 4.5. INTRODUCTION

The whitefly, *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) is one of the most important agricultural pests and virus vectors of agricultural and ornamental crops in all tropical and subtropical areas (Brown *et al.*, 1995). The insect causes significant damage to crops primarily through phloem feeding, phytotoxic disorders and the transmission of plant viruses. The development of insecticide resistance, reduction in natural enemies and monocultural practices have been considered the main drivers in the emergence of *B. tabaci* as the primary agricultural pest in tropical and subtropical agricultural systems (Brown *et al.*, 1995). The majority of *B. tabaci* species are considered polyphagous, colonizing several hundred plant species. However some monophagous or nearly monophagous populations have been reported (Bird, 1957; Mound, 1983; Bedford *et al.*, 1994).

The concept of host races has been proposed due to occurrence of *B. tabaci* populations with no clear-cut diagnostic differences in morphology but exhibiting differences in host range, dispersal behavior, fecundity and competency for begomovirus transmission (Brown and Bird, 1992; Brown *et al.*, 1995). For example, the *Jatropha* race colonizes *Jatropha gossypifolia* L. and *Croton lobatus* L. in Puerto Rico and transmits only *Jatropha mosaic virus* to and from *Jatropha* spp. In contrast another population readily colonizes a wide range of plant species including those in the genus *Sida* and is designated the *Sida* race (Bird, 1957). Other examples are the *Asystasia* spp. restricted *B. tabaci* from Benin (Brown and Bird, 1992; Brown *et al.*, 1995) and the cassava-colonizing *B. tabaci* in sub-Saharan Africa (Storey and Nichols, 1938; Burban *et al.*, 1992) are recognized as monophagous or nearly so on their host plant.

At the same level of distinction as the host race is the *biotype* concept, a term used to designate populations that also lack morphological distinction but possess other characteristics which serve to separate them from other populations (Perring, 2001). Currently at least 20 biotypes have been identified on the basis of electrophoretic pattern of general esterases and were designated A-S with the most widespread being the B-biotype (De Barro *et al.*, 2000; Abdullah *et al.*, 2003). It is the B-biotype that has raised *B. tabaci* to its current international importance as an agricultural pest mainly because it is highly fecund, has a wide host range, is resistant to pesticides and transmits most begomoviruses tested (Brown *et al.*, 1995; Brown, 2001). Not only has the B-biotype different esterase patterns, it also has distinctive genetic and biological characteristics with respect to the A-biotype (another economically important biotype occurring in the Americas or New World). The B biotype is now found all over the world owing to its association with ornamental plants and world trade in this commodity (Perring, 2001; Abdullahi *et al.*, 2003; De Barro *et al.*, 2005).

Although biochemical studies were useful in distinguishing whitefly biotypes, new PCR-based techniques are increasingly being used, providing better resolution of differences between whitefly populations and revealing polymorphisms. The use of RAPD-PCR has been widely reported and has been shown to support findings based on esterase patterns (Reiter *et al.*, 1992; Lima *et al.*, 2002). This technique has been used to distinguish biotype A from biotype B (Perring *et al.*, 1993) and was extended to compare other biotypes from around the world (De Barro and Driver, 1997; De Barro *et al.*, 1998). Other molecular techniques include 16S mitochondrial rDNA sequences (Frohlich *et al.*, 1999), 18S rDNA sequences (Campbell *et al.*, 1994), ITS1 region of ribosomal DNA (De Barro *et al.*, 2000) and more recently the use of

mitochondrial cytochrome oxidase I gene (mtCOI) sequences (Frohlich *et al.*, 1999; Brown *et al.*, 2000).

The mtCOI gene has been widely used in sub-Saharan Africa and other parts of the world; leading to a suggestion that *B. tabaci* is a species complex (Frohlich *et al.*, 1999). This marker has been used in a previous study to genotype *B. tabaci* populations in southern Africa and revealed the existence of five distinct geographic haplotypes associated with cassava, however with the exception of one and two isolates from Cameroon and Zimbabwe respectively (Berry *et al.*, 2004). Specific whitefly phenotypes and the interactions with their host plants directly influence both pest status and the dynamics of virus vector host interactions (Brown *et al.*, 1995). Because management strategies rely on insect biology, behavior, interactions with natural enemies and response to agricultural chemicals, what works for certain populations of *B. tabaci* may not work for others. Therefore understanding of genotype variability in vector populations may help in designing appropriate control strategies for both virus and vector. In this study, the diversity of cassava-colonizing *B. tabaci* populations from cassava growing areas of South Africa was investigated using mtCOI gene sequences.

## **4.6. MATERIALS AND METHODS**

### **4.6.2. Whitefly samples**

*Bemisia tabaci* adults and nymphs (where present) were collected off the ventral surfaces of cassava leaves, in the cassava growing areas of Bushbuckridge (formerly in Limpopo), Mariti and Tonga in the Mpumalanga province of South Africa. A total of 65 samples (table 4.1) were collected during January and March 2005. Both adult whiteflies and nymphs were stored in 70% ethanol at -20 °C until analysis was carried out.

### **4.6.3. DNA extraction, PCR and Cloning**

Total nucleic acids were extracted from individual whiteflies and nymphs according to the method of Frohlich *et al.* (1999). Polymerase chain reaction (PCR) was conducted on all samples collected. PCR primers for amplifying the mitochondrial cytochrome oxidase I gene (mtCOI) fragment were C1-J-2195 and L2-N-3014 selected from the UBC Insect Mitochondrial DNA Primer Oligonucleotide set, with sequences obtained from Simon *et al.* (1999). PCR cycling conditions were conducted as described by Frohlich *et al.* (1999). Mitochondrial cytochrome oxidase I gene fragment (~780bp) amplicons were sequenced in both directions using PCR and sequencing primers and an automated ABI Prism sequencer at the Laboratory for Molecular Systematics and Evolution, University of Arizona, Tucson, AZ 85721, USA.

Table 4.1 Collection sites of whitefly samples

| Collection Site       | Sample Code | Region        |
|-----------------------|-------------|---------------|
| Mandela 1             | SAS1.3      | Bushbuckridge |
| Mandela 2             | SAS5        | Bushbuckridge |
| Green Valley-Brooklyn | SAS6.2      | Bushbuckridge |
| Green Valley-Brooklyn | SAS6.1      | Bushbuckridge |
| Angiencourt           | SAS3.1      | Bushbuckridge |
| Angiencourt           | SAS3.2      | Bushbuckridge |
| Angiencourt           | SAS3.3      | Bushbuckridge |
| New Forest            | SAS14       | Bushbuckridge |
| Dingleydale 1         | SAS15       | Bushbuckridge |
| Dingleydale 4         | SAS16       | Bushbuckridge |
| Mkhuhlu 2             | SAS17.1     | Bushbuckridge |
| Mkhuhlu 2             | SAS17.2     | Bushbuckridge |
| Mkhuhlu 2             | SAS17.3     | Bushbuckridge |
| Jim Brown 1           | SAS10       | Mariti        |
| Jim Brown 2           | SAS11       | Mariti        |
| Nkanini 1             | SAS7        | Mariti        |
| Jim Brown 1           | SAS18       | Mariti        |
| Jim Brown 1           | SAS19       | Mariti        |
| Tsakani 1             | SAS22.1     | Mariti        |
| Tsakani 1             | SAS22.2     | Mariti        |
| Tsakani 1             | SAS22.3     | Mariti        |
| Tsakani 3             | SAS24.3     | Mariti        |
| Tsakani 3             | SAS24.1     | Mariti        |
| Tsakani 3             | SAS24.2     | Mariti        |
| Nkanini 1             | SAS26 (1,2) | Mariti        |
| Nkanini 1             | SAS27.1     | Mariti        |
| Watervaal 3           | SAS28       | Mariti        |
| Watervaal 4           | SAS30       | Mariti        |
| Mandela 1             | SAS31       | Mariti        |
| Goromani 2            | SAS33       | Mariti        |
| Marite 1              | SAS34       | Mariti        |
| Tonga 3               | SAS4        | Tonga         |
| Masibekela 1          | SAS8        | Tonga         |
| Masibekela 2          | SAS9.2      | Tonga         |
| Masibekela 2          | SAS9.3      | Tonga         |
| Hoyi 2                | SAS35       | Tonga         |
| Masibekela 1          | SAS36       | Tonga         |
| Masibekela 1          | SAS37.2     | Tonga         |
| Masibekela 3          | SAS38       | Tonga         |
| Masibekela 3          | SAS39       | Tonga         |
| Mangweni 1            | SAS40       | Tonga         |
| Mangweni 1            | SAS41.3     | Tonga         |
| Mangweni 1            | SAS41.2     | Tonga         |
| Mangweni 2            | SAS42       | Tonga         |
| Mangweni 2            | SAS43       | Tonga         |
| Hoyi 1                | SAS44       | Tonga         |
| Hoyi 1                | SAS45       | Tonga         |
| Tonga 1               | SAS46       | Tonga         |
| Tonga 1               | SAS47       | Tonga         |
| Mzinti 1              | SAS48       | Tonga         |
| Sikhwahlane 1         | SAS50       | Tonga         |
| Sikhwahlane 2         | SAS52       | Tonga         |
| Sihlangu 1            | SAS54       | Tonga         |
| Sihlangu 1            | SAS55       | Tonga         |
| Hlabaville            | SAS56       | Tonga         |
| Hlabaville            | SAS57       | Tonga         |
| Phakama               | SAS58       | Tonga         |
| Laangekop 1           | SAS60       | Tonga         |
| Laangekop 1           | SAS61.3     | Tonga         |
| Laangekop             | SAS61.2     | Tonga         |
| Bekisisa              | SAS63       | Tonga         |
| Ntunda                | SAS64       | Tonga         |
| Ntunda                | SAS65       | Tonga         |

#### 4.6.4. Phylogenetic analysis

Sequences were aligned using the Clustal algorithm (ClustalW 1.7) (MegAlign, DNA'STAR, Madison, Wisconsin, USA), and aligned sequences were evaluated for genetic relatedness by Parsimony using PAUP (Swatford *et al.*, 1991). Bootstrapping was performed with PAUP using the heuristic option for 1000 replications. Sequences of the whiteflies, *Trialeuroides vaporariorum* (Westwood) and *B. afer* (Priesener and Hosny) were used as the out groups.

#### 4.7. RESULTS

A single most parsimonious tree (figure 4.1) was reconstructed from mtCOI sequences. Genetic relatedness (percentage nucleotide sequence identity) of the mtCOI sequences of *Bemisia tabaci* is presented in table 4.1. Figure 4.2 is the bootstrap tree showing the mtCOI gene fragment sequence relationships. The mtCOI sequence was useful in separating all New World *B. tabaci* from Old World *B. tabaci*. *Bemisia tabaci* collected from cassava during this study in South Africa formed a monophyletic group together with the previously identified southern African clade, sharing a sequence identity of 91-99 % (n = 120).

Three isolates collected during this study formed a distinct cluster from the southern African group. The three isolates share a sequence identity of 97-98 % amongst themselves; meanwhile they share 84-94 % (n = 136) sequence identity with the rest of the southern African clade. They also share 88-90 % sequence identity with the non-cassava type ABA Benin. Interestingly, these isolates were collected from all three areas surveyed an indication that they may be common across the survey areas.

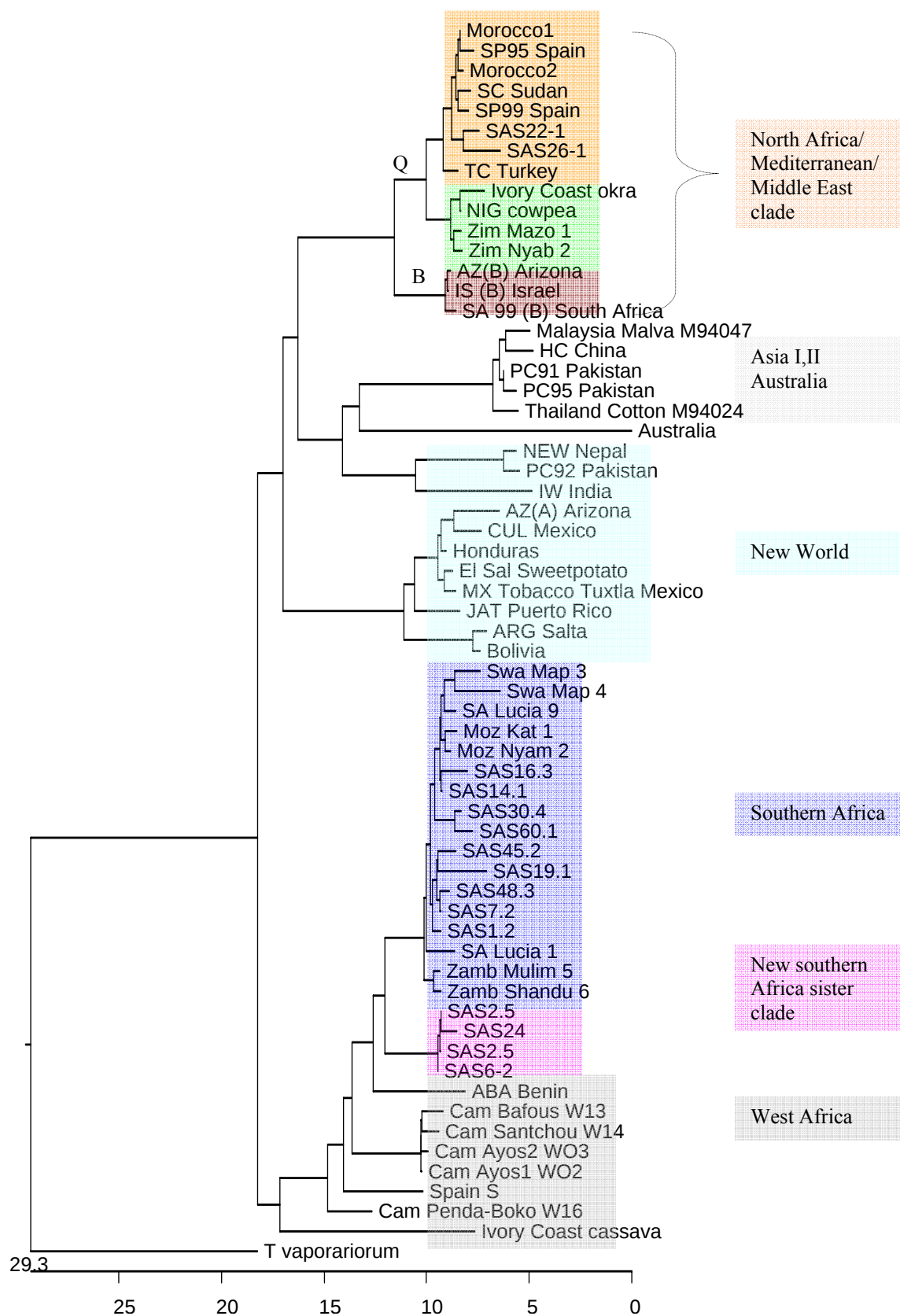


Figure 4.1 MTCOI tree showing relationships and identification of Southern African and reference whitefly populations

Bootstrap

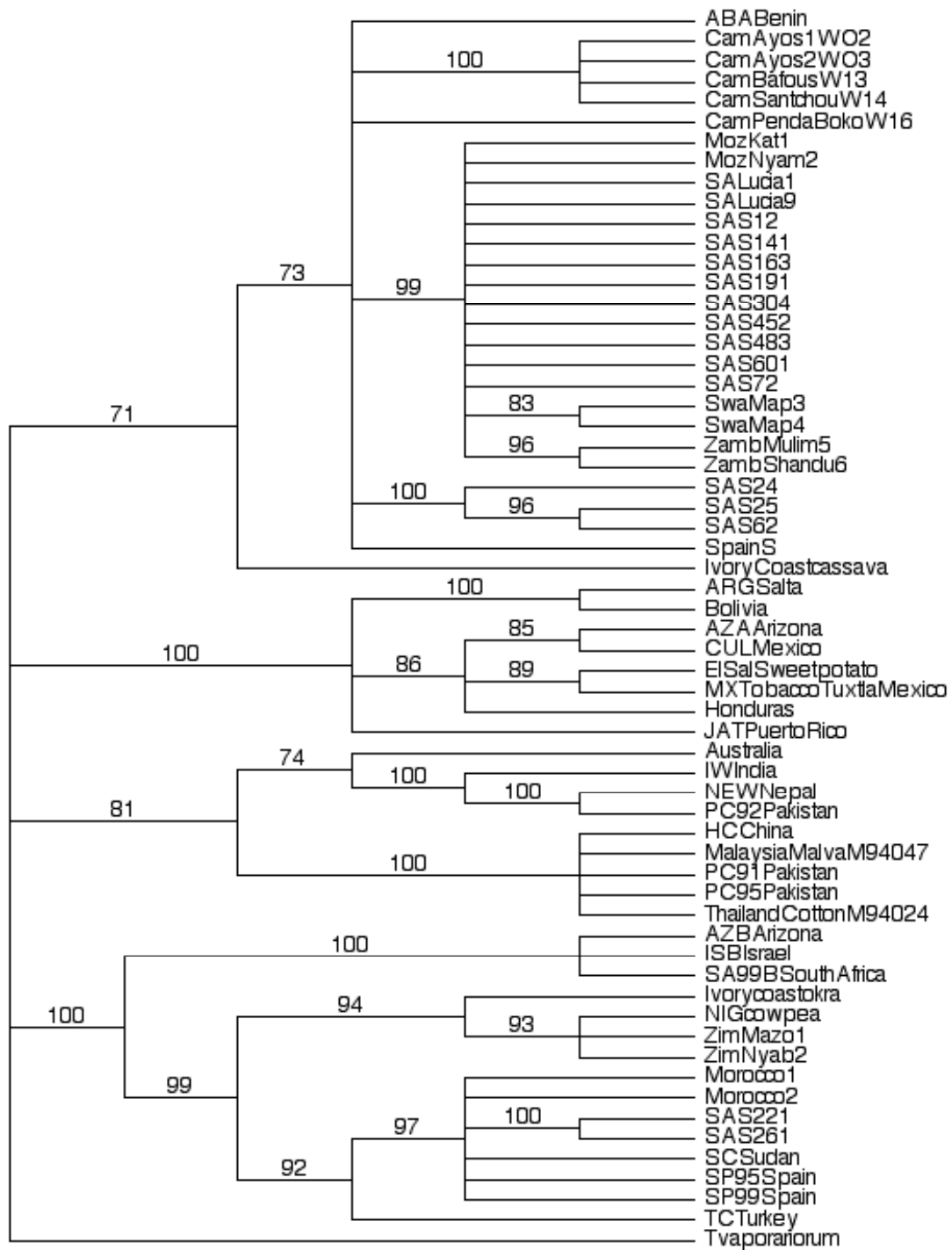


Figure 4.2 Bootstrap MTCOI tree showing relationships and identification of Southern African and reference whitefly populations

Table 4.2 Percentage nucleotide sequence identities between selected South African whiteflies from cassava and other host plants from various African Countries

|                     | SAS14.1<br>Bush-<br>NF3 | SAS1.2<br>Bush-<br>MAN1 | SAS16.3<br>Bush-<br>DD4 | SAS6.1<br>Bush-<br>GVB | SAS19.1<br>Mariti-<br>JM1 | SAS7.2<br>Mariti-<br>NK1 | SAS30.4<br>Mariti-<br>NK4 | SAS45.2<br>Tonga-<br>H1 | SAS48.3<br>Tonga-<br>MZ1 | SAS60.1<br>Tonga-<br>LAN1 | SAS6.2<br>Bush-<br>GVB | SAS24.3<br>Mariti-<br>TS3 | SAS2.5<br>Tonga-<br>MS3 |
|---------------------|-------------------------|-------------------------|-------------------------|------------------------|---------------------------|--------------------------|---------------------------|-------------------------|--------------------------|---------------------------|------------------------|---------------------------|-------------------------|
| ABA Benin           | 86                      | 90                      | 87                      | 87                     | 87                        | 90                       | 89                        | 88                      | 88                       | 89                        | 88                     | 90                        | 90                      |
| Ivory Coast cassava | 84                      | 84                      | 83                      | 79                     | 83                        | 85                       | 84                        | 84                      | 84                       | 82                        | 82                     | 83                        | 83                      |
| Ivory Coast okra    | 80                      | 81                      | 79                      | 77                     | 81                        | 81                       | 80                        | 79                      | 81                       | 80                        | 78                     | 80                        | 80                      |
| Morocco2            | 81                      | 81                      | 80                      | 78                     | 82                        | 82                       | 81                        | 80                      | 81                       | 80                        | 78                     | 80                        | 80                      |
| Morocco1            | 81                      | 82                      | 80                      | 78                     | 83                        | 83                       | 81                        | 81                      | 82                       | 81                        | 80                     | 82                        | 82                      |
| SA 99 (B) S. Africa | 81                      | 81                      | 80                      | 78                     | 80                        | 82                       | 81                        | 80                      | 81                       | 80                        | 78                     | 80                        | 80                      |
| SC Sudan            | 81                      | 81                      | 79                      | 73                     | 82                        | 81                       | 80                        | 80                      | 80                       | 80                        | 78                     | 80                        | 80                      |
| Uganda Sweetpotato  | 79                      | 77                      | 77                      | 7                      | 79                        | 79                       | 79                        | 80                      | 78                       | 79                        | 78                     | 80                        | 80                      |
| IS (B) Israel       | 82                      | 82                      | 80                      | 79                     | 81                        | 82                       | 82                        | 81                      | 82                       | 81                        | 79                     | 81                        | 81                      |
| NIG cowpea          | 81                      | 82                      | 81                      | 78                     | 82                        | 83                       | 8                         | 81                      | 82                       | 81                        | 80                     | 81                        | 82                      |
| Cam Ayos1 WO2       | 89                      | 89                      | 88                      | 87                     | 87                        | 90                       | 89                        | 89                      | 89                       | 88                        | 89                     | 91                        | 92                      |
| Cam Ayos2 WO3       | 89                      | 89                      | 88                      | 87                     | 87                        | 89                       | 89                        | 88                      | 89                       | 87                        | 89                     | 91                        | 92                      |
| Cam Bafous W13      | 87                      | 87                      | 85                      | 85                     | 85                        | 87                       | 87                        | 86                      | 87                       | 86                        | 87                     | 89                        | 89                      |
| Cam Penda-Boko W16  | 90                      | 90                      | 88                      | 88                     | 88                        | 90                       | 90                        | 89                      | 89                       | 88                        | 89                     | 92                        | 92                      |
| Cam Santchou W14    | 89                      | 89                      | 87                      | 85                     | 87                        | 89                       | 89                        | 88                      | 89                       | 87                        | 87                     | 89                        | 89                      |
| Cam Yaounde 2WO6    | 89                      | 89                      | 88                      | 87                     | 87                        | 90                       | 89                        | 89                      | 89                       | 88                        | 89                     | 91                        | 91                      |
| Zim Mazo 1          | 81                      | 82                      | 79                      | 77                     | 81                        | 82                       | 81                        | 80                      | 81                       | 80                        | 78                     | 80                        | 80                      |
| Zim Nyab 2          | 81                      | 82                      | 79                      | 77                     | 81                        | 82                       | 81                        | 80                      | 80                       | 80                        | 78                     | 80                        | 80                      |
| SA Lucia 1          | 93                      | 94                      | 92                      | 92                     | 92                        | 95                       | 94                        | 94                      | 94                       | 92                        | 90                     | 92                        | 92                      |
| SA Lucia 9          | 95                      | 94                      | 93                      | 94                     | 92                        | 95                       | 94                        | 94                      | 94                       | 93                        | 91                     | 93                        | 93                      |
| Swa Map 3           | 95                      | 95                      | 93                      | 92                     | 92                        | 95                       | 94                        | 94                      | 94                       | 94                        | 88                     | 90                        | 90                      |
| Swa Map 4           | 93                      | 93                      | 92                      | 93                     | 91                        | 93                       | 93                        | 94                      | 93                       | 92                        | 88                     | 90                        | 90                      |
| Moz Kat 1           | 96                      | 96                      | 95                      | 94                     | 93                        | 96                       | 95                        | 95                      | 96                       | 95                        | 90                     | 92                        | 93                      |
| Moz Nyam 2          | 96                      | 96                      | 94                      | 94                     | 93                        | 96                       | 95                        | 95                      | 95                       | 95                        | 91                     | 93                        | 93                      |
| Zamb Mulim 5        | 95                      | 96                      | 94                      | 93                     | 93                        | 96                       | 95                        | 95                      | 95                       | 94                        | 91                     | 93                        | 93                      |
| Zamb Shandu 6       | 95                      | 96                      | 94                      | 93                     | 94                        | 96                       | 95                        | 95                      | 95                       | 94                        | 91                     | 93                        | 93                      |
| SAS14.1 Bush-NF3    | -                       | 97                      | 97                      | 96                     | 94                        | 98                       | 99                        | 88                      | 88                       | 98                        | 92                     | 94                        | 94                      |
| SAS1.2 Bush-MAN1    | 97                      | -                       | 96                      | 94                     | 95                        | 97                       | 98                        | 90                      | 89                       | 96                        | 91                     | 94                        | 94                      |
| SAS16.3 Bush-DD4    | 97                      | 96                      | -                       | 94                     | 93                        | 96                       | 96                        | 86                      | 86                       | 96                        | 89                     | 91                        | 91                      |
| SAS6.1 Bush-GVB     | 96                      | 94                      | 94                      | -                      | 92                        | 96                       | 95                        | 90                      | 90                       | 96                        | 92                     | 90                        | 88                      |
| SAS19.1 Mariti-JM1  | 96                      | 95                      | 93                      | 92                     | -                         | 95                       | 96                        | 90                      | 87                       | 95                        | 89                     | 91                        | 91                      |
| SAS7.2 Mariti-NK1   | 98                      | 97                      | 96                      | 96                     | 95                        | -                        | 99                        | 90                      | 89                       | 98                        | 93                     | 93                        | 91                      |
| SAS30.4 Mariti-NK4  | 99                      | 98                      | 96                      | 95                     | 96                        | 99                       | -                         | 89                      | 94                       | 98                        | 92                     | 94                        | 94                      |
| SAS45.2 Tonga-H1    | 95                      | 94                      | 95                      | 93                     | 94                        | 95                       | 93                        | -                       | 94                       | 89                        | 86                     | 84                        | 84                      |
| SAS48.3 Tonga- MZ1  | 94                      | 95                      | 93                      | 94                     | 93                        | 96                       | 94                        | 94                      | -                        | 89                        | 85                     | 84                        | 84                      |
| SAS60.1 Tonga-LAN1  | 98                      | 96                      | 96                      | 96                     | 95                        | 98                       | 98                        | 91                      | 92                       | -                         | 92                     | 93                        | 93                      |
| SAS6.2 Bush-GVB     | 92                      | 91                      | 89                      | 92                     | 89                        | 93                       | 92                        | 86                      | 85                       | 92                        | -                      | 97                        | 97                      |
| SAS24.3 Mariti-TS3  | 94                      | 94                      | 91                      | 90                     | 91                        | 93                       | 94                        | 84                      | 84                       | 93                        | 97                     | -                         | 99                      |
| SAS2.5 Tonga-MS3    | 94                      | 94                      | 91                      | 88                     | 91                        | 91                       | 94                        | 84                      | 84                       | 93                        | 97                     | 99                        | -                       |
| T vaporariorum      | 69                      | 70                      | 68                      | 63                     | 68                        | 70                       | 69                        | 69                      | 68                       | 67                        | 65                     | 67                        | 67                      |

Four isolates from Mariti (SAS22.1, SAS26.1, SAS26.2 and SAS24.2) clustered with the Mediterranean/North African Q types. Four others (SAS27.1-Mariti, SAS37.2-Tonga, SAS41.3-Tonga and SAS60.3-Tonga) were non-*Bemisia* species and another three (SAS3.2-Bushbuckridge, SAS9.2-Tonga, and SAS17.1/2-Bushbuckridge) were *Bemisia* species. Although these were collected on cassava it is most probable that they were just visitors or transient feeders.

#### **4.8. DISCUSSION AND CONCLUSIONS**

The mtCOI sequences in this study placed most of the cassava whitefly populations collected in South Africa (this study) in the same monophyletic group as previously identified whitefly populations, which consists of five distinct genotype sub-clusters of cassava-colonizing *B. tabaci* in sub-Saharan Africa (Berry *et al.* 2004). These clusters include the southern African (cassava), western African (non-cassava), western African-Cameroon (cassava), western African (cassava) and western African (cassava and wild *Solanum* spp.) clades. As expected most of the collections made during this study belong to the southern African cluster. Berry *et al.* (2004) have shown that this group shares over 96 % identity amongst themselves. Although the majority of the isolates collected in the current study show a similar trend, some of them have sequence identities that go as low as 83%, (sequence identity range from 83 % to 100 %), thus exhibiting a wide sequence diversity within the populations.

It is generally accepted that *B. tabaci* populations are geographically delimited (Frohlich *et al.*, 1999; Brown, 2000; De Barro *et al.*, 2000), however evidence suggests that distinct genotypes are not as tightly delimited by sub-geographies within southern

Africa (Berry *et al.* 2004). The current results agree with this observation in that whiteflies collected from three different areas did not show any geographical affiliation. The Bushbuckridge and Mariti regions are close (see chapter 2, section 2.3.1) to one another and therefore it is expected that isolates from these areas do not show geographic affiliation.

Other than the major southern African group, we report here the occurrence of a ‘new’ southern African sister clade associated with cassava or other plant species close to cassava fields. Isolates of the ‘new’ southern African sister clade were found in each of the three surveyed areas; this could be an indication that this group is common in these areas and there is a possibility that it may be colonizing any of the crops or weed species near or within cassava fields. However, it cannot be ruled out that this group may colonize cassava or readapting to cassava from ornamental or weed plants because it falls within the cassava colonizing group that contains only one non-cassava type namely, ABA Benin. So far it has not been established if ABA Benin cannot colonize cassava at all, and Abdullahi *et al.* (2003) noted that although ABA Benin was not collected from cassava and the fact that cassava *B. tabaci* is a host specialist, does not exclude it from being found on other plants as a visitor. Therefore it is possible that this ‘new’ southern African sister clade may be hosted by cassava. A look at whiteflies colonizing other crops or weeds within or around cassava fields would also give insight as to the host range of cassava biotypes, with implications on sustainable cultural control approaches such as closed season, crop rotation, trap cropping and uniform planting. The significance of this finding is that the occurrence of a biotype that is closely related to the cassava-colonizing biotype could present opportunities for gene flow between cassava and non-cassava

genotypes thus increasing the possibility of the emergence of new biotypes that may potentially alter the dynamics of the virus-vector relationship.

Another interesting finding in the current study is the presence of Q biotype-like isolates for the first time in South Africa. These isolates share a sequence identity of 95% to 98% with the Q biotypes found in the Mediterranean basin and North Africa. However two isolates from Zimbabwe were shown to be closely related to the North African/Mediterranean/B biotype group which is also a close relative of the Q biotype group from the same region (Berry *et al.*, 2001). The Q biotype is one of the most important *B. tabaci* biotypes and it is particularly significant in the spread of geminiviruses with devastating effects on crop production in the Mediterranean basin and North Africa. It is interesting in that contrary to a report by De La Rau *et al.* (2006) suggesting that the Q biotype was only found in the north of the Sahara and a proposition that the separation of biotype J and Q was imposed by the Sahara Desert it is reported this far in the south of the continent. It appears therefore that the Q biotype may be more widespread than previously thought. The four Q biotype-like isolates were all collected from the same area (Mariti) and the collection sites were in very close proximity to one another. This could be an indication that its occurrence is likely a result of a single introduction from elsewhere and that it is not indigenous to the area. It is known that this biotype does not colonize cassava and is therefore not a threat to cassava production however it could have serious implications for other crops in the area and South Africa in general should population numbers grow and spread. Also the possibility of hybridization and cross-breeding could have implications for increased diversity with possible

emergence of more fecund strains. Hybrids from crosses between the B biotype from Sudan and the Q biotype from Spain have been reported (Ronda *et al.*, 1999).

In conclusion, the surveyed cassava-growing areas in South Africa do not appear to have significant genetic variation in both the whitefly and virus populations, whereby only one significant whitefly genotype (clade) occurs and ACMV dominates. Currently, whitefly transmission of geminiviruses is not of great concern and phytosanitation and cultural practices could significantly reduce the incidence of cassava mosaic disease. However there is wider range in whitefly mtCOI gene sequence diversity (91-99%). Cross-breeding and hybridization amongst different biotypes has been reported (Ronda *et al.*, 1999; Brown *et al.*, 2001) and could play a major role in increasing their diversity. In areas where cassava is grown on a larger scale such as the cassava farm in Barberton (Mpumalanga) whiteflies seem to play a significant role and could cause severe crop losses (personal observation). Therefore given the potential of cassava as a commercial crop in South Africa further studies on transmission efficiencies, mating studies, gene flow studies and virus-vector relationships would be very useful in anticipating future control strategies.

## **CHAPTER 5: SUMMARY**

The aim of this study was to undertake a comprehensive epidemiology study of CMD in cassava cultivation areas in South Africa and to establish the genetic diversity of the viruses and vectors present on cassava in South Africa.

### **5.1. Epidemiology of Cassava Mosaic Disease**

Cassava mosaic disease is a major constraint to cassava production in Africa. It is spread by the use of infected stem cuttings as planting material and almost exclusively transmitted by the whitefly vector *Bemisia tabaci*. Several epidemiological studies have been conducted in many cassava-growing regions of Africa. Most of these studies were conducted in East and Central Africa beginning in the early 1990's, mostly as a result of the CMD epidemic that started in Uganda. The epidemic was associated with a recombinant severe strain of EACMV designated EACMV-UG and unusually high populations of the whitefly vector *B. tabaci* (Legg *et al.*, 2002, Gibson *et al.*, 1996).

The aim of this study was to investigate the epidemiology of CMD in South Africa, by monitoring the incidence (cutting and whitefly-borne), severity and whitefly population numbers. The results showed that cutting-borne disease incidence was the major source of infections. Whitefly-borne disease incidence contributes a smaller portion of infections, especially in Bushbuckridge where whitefly populations were much lower, particularly in the beginning of the season. However in Mariti and Tonga whitefly-borne infections were significantly higher than in Bushbuckridge, but also significantly lower than cutting infections. In Mariti and Tonga the observed higher whitefly-borne infections were also accompanied by higher whitefly numbers compared to Bushbuckridge. Therefore whiteflies could play a major role in the epidemiology of

CMD especially if cassava cultivation is intensified, and commercial farms in the surrounding areas reduce spraying.

Disease severity and index was lower (2.37 to 2.78 and 1.25 to 2.05 respectively) in South Africa compared to, for example in the epidemic areas of East and Central Africa ( $DS > 4$ ) (Legg *et al.*, 2001). This is not surprising, given that the predominant CMG species in South Africa is ACMV (Chapter 3), which is known to induce milder CMD symptoms compared to EACMV (Fondong *et al.*, 2000, Maruthi *et al.*, 2005).

Thus far cutting infections are of paramount importance. The significance of cutting infections in South Africa is well supported by two previous studies (Trench and Martin, 1985; Jericho *et al.*, 1999). The current study, together with the two previous studies indicated that the disease situation in South Africa is stable. However the situation could change if the current interest in cassava in South Africa results in the intensification of cultivation. Indeed in South Africa there is a growing industrial interest in cassava for production of starch and bio-fuels. The use of cultural control practices such as phytosanitation, roguing and resistant cultivars could greatly improve cassava production in South Africa.

## **5.2. Genetic Diversity of Cassava Mosaic Geminiviruses**

Investigations into the genetic diversity of CMG's in South Africa have revealed the presence of four distinct CMG's, namely; ACMV, EACMV, ACMV-UGV and SACMV (Berrie *et al.*, 1998; Berry *et al.*, 2001). In the current study the use of differential primers indicated the presence of both ACMV and EACMV mostly in mixed infections. The use of RFLP's has shown four different restriction patterns presumed to represent ACMV,

ACMV-UG, EACMV and an unknown isolate. SACMV was not detected in any of the samples tested.

Phylogenetic analysis of four isolates, representing each RFLP pattern, revealed that three of the isolates were ACMV. One (SA-Tonga) of the three isolates had the least sequence similarity with previously sequenced isolates from sub-Saharan Africa (<93% nt sequence similarity) including the other (this study) two ACMV isolates (>95% nt sequence similarity) and it was concluded that it is a distinct strain of ACMV. The SA-Mariti ACMV isolate demonstrated a similar RFLP pattern to ACMV-UG despite having a high nucleotide sequence similarity to SA-Barb (>98%), while SA-Barb and SA-Tonga isolates shared <95% nucleotide sequence similarity in phylogenetic studies. The fourth isolate was identified as EACMV and was most closely related to EACMMV-MK/MH isolates from Malawi. Therefore these results confirm the presence of ACMV, EACMV and ACMV-UG, as shown in previous studies, in South Africa. In addition a 'new' strain of ACMV was identified. SACMV can therefore be considered a rare species.

Recombination analysis has shown widespread recombination in EACMV isolates. The EACMV isolated in this study has only exhibited a single recombination event compared to most other EACMV's that have at least four recombination events (figure 3.6). All ACMV isolates (including current study) analyzed showed a single recombination event of unknown origin (minor parent TYLCV-Ethiopia). These results indicate that recombination activity is low in South Africa compared to other African countries, despite the fact that mixed infections are a common occurrence. The lower frequencies of recombination detected to date could perhaps be due to the geographical and political (before 1990) isolation of South Africa. However a wider screening of full

length DNA (A/B)'s by RCA's in future for example could pick up further recombination events. The political situation has changed since and increased movement of people and goods into South Africa could present opportunities for introduction of new virus strains and species.

### **5.3. Genetic Diversity of *Bemisia tabaci***

The whitefly *Bemisia tabaci* is the only known insect vector of CMG's. Phylogenetic analysis of the mtCOI gene molecular marker has revealed the presence of distinct groups of cassava infesting whitefly populations in southern Africa: Whiteflies from Mozambique, South Africa, Swaziland and Zambia formed a closely related (95-99%) clade (southern African clade) and those from Zimbabwe formed a separate group closely related to whitefly infecting okra from Cameroon and Ivory Coast (North African/Mediterranean clade) (Berry *et al.*, 2004).

The current study aimed to investigate the genetic diversity of cassava-infesting whiteflies in South Africa using the mtCOI gene fragment. The majority of the whiteflies in South Africa formed a monophyletic group with the southern African clade. In addition, three isolates formed a distinct sister clade to the southern African group. However it remains to be seen if the 'new' southern African sister clade is cassava-colonizing or just a transient feeder on cassava. If it does colonize cassava it could have implications for the diversity of *B. tabaci*, raising the possibility for cross-breeding and hybridization. A further four isolates collected in close proximity to one another in Mariti grouped together with the Q biotype from the North Africa/Mediterranean clade and formed a sister clade with the previously identified Q biotype from Zimbabwe (Berry *et*

*al.*, 2004). However it is unlikely that these are primary colonizers of cassava. Still, their presence is a cause for concern because of the possibility for cross-breeding.

In conclusion, three important findings in this study were:

- 1) The previously identified southern African clade showed greater genetic diversity (sequence identity range of 91-99%) than previously (>96% nucleotide sequence identity) reported (Berry *et al.*, 2004),
- 2) The occurrence of a 'new' southern African sister clade
- 3) The Q biotype of North African/Mediterranean origin was reported for the first time in South Africa.

#### **5.4. Conclusions**

Epidemiological studies have shown that the disease situation in South Africa is stable, remaining with respect to disease incidence and virus genetic diversity unchanged for at least eight years. Over this period the disease incidence remained virtually unchanged with cutting infections remaining the most important mode of CMD transmission. Therefore cultural practices could help maintain the status quo or even reduce CMD incidence. In addition introduction of resistant cultivars could supplement cultural practices in controlling the disease. Virus diversity also did not show significant changes since the last study published in 2001 (Berry and Rey, 2001). The discovery of a new ACMV strain reiterates the concern for the rapid evolution of CMG's in Africa including South Africa, and predicts that wider screening of full length DNA sequences would be likely to reveal a higher degree of genetic diversity and recombination. Another concern is the increased movement of people and goods across borders in the Southern Africa

Development Community (SADC) region increasing the possibility of new virus strains and species being introduced. Whiteflies also showed greater diversity than in previous studies and together with the occurrence of a new southern Africa sister clade and the presence of the Q biotype, raises concern for the emergence of new biotypes through cross-breeding and hybrids. Therefore continuous monitoring of CMD, CMG's and *B. tabaci* populations is necessary in South Africa in order to devise appropriate control strategies, breeding programs and genetic engineering strategies.

## **CHAPTER 6: REFERENCES**

- Abdullahi, I., Winter, S., Atiri, G. I. and Thottappilly, G. 2003. Molecular characterization of whitefly, *B. tabaci* (Hemiptera: Aleyrodidae) populations infesting cassava. *Bulletin of Entomological Research* **93**: 97-106.
- Banks, G. K., Bedford, I. D., Beitia, F. J., Rodriguez-Cerezo, E. and Markham, P.G. 1999. A novel geminivirus of *Ipomoea indica* (Convolvulaceae) from Southern Spain. *Plant Disease* **83**: 486.
- Bedford, I. D., Briddon, R. W., Brown, J. K., Rosell, R. C. and Markham, P. G. 1994. Geminivirus transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. *Annals of Applied Biology* **125**: 311-325
- Bellows, T. S., Perring, T. M., Gill, R. J. and Hendrick, D. H. 1994. Description of a species of *Bemisia* (Homoptera: Aleyrodidae). *Annals of Entomological Society of America* **87**: 195-206.
- Berrie, L. C., Palmer, K., Rybicki, E. P., Hiyadat, S. H., Maxwell, D. P. & Rey, M. E. C. 1997. A new isolate of African cassava mosaic virus in South Africa. *African Journal of Root and Tuber Crops* **2**: 49-52
- Berrie, L. C., Palmer, K. E., Rybicki, E. P. and Rey, M. E. C. 1998. Molecular characterisation of a distinct South African cassava infecting geminivirus. *Archives of Virology* **143**: 2253-2260.
- Berrie, L. C., Rybicki, E. P. and Rey, M. E. C. 2001. Complete nucleotide sequence and host range of *South African cassava mosaic virus*: further evidence for recombination amongst begomoviruses. *Journal of General Virology* **82**: 53-58.

- Berry, S. D. and Rey, M. E. C. 2001. Molecular evidence for diverse populations of cassava infecting begomoviruses in southern Africa. *Archives of Virology* **146**, 1795-1802.
- Berry, S. D., Fondong, V. N., Rey, C., Rogan, D., Fauquet, C. M. and Brown, J. K. 2004. Molecular evidence for five distinct *Bemisia tabaci* (Homoptera: Aleyrodidae) geographic genotypes associated with cassava plants in sub-Saharan Africa. *Annals of the Entomological Society of America* **97**: 852–859.
- Bird, J. 1957. A whitefly-transmitted mosaic of *Jatropha gossypifolia*. *Agric. Exp. Stn. Univ. Puerto Rico* **22**:1-35.
- Bock, K. R. and Guthrie, E. J. 1978. Transmission of African cassava mosaic by mechanical inoculation. *Plant Disease Reporter* **62**: 580-581. Bock, K. R., Guthrie, E. J. and Meredith, G. 1978.
- Bock, K. R., Guthrie, E. J. and Meredith, G. 1978. Distribution, host range, properties and purification of cassava latent virus, a geminivirus. *Annals of Applied Biology* **90**: 361-367.
- Bock, K. R. and Woods, R. D. 1983. Etiology of African cassava mosaic disease. *Plant Disease* **67**: 994-995.
- Borror, D. J, Triplehorn, C. H. A. and Johnson, N. F. 1989. *An introduction to study of insects*. 6th edition. Saunders College publishing.
- Böttcher, B., Unseld, S., Ceulemans, H., Russell, R.B and Jeske, H. 2004. Geminata Structures of *African Cassava Mosaic Virus*. *Journal of Virology* **78**: 6758-6765
- Briddon, R. W. and Markham, P. G. 1994. Universal primers for the PCR amplification of dicot-infecting geminiviruses. *Molecular Biotechnology* **1**: 202-205.

- Briddon, R.W., Robertson, I., Markham, P.G. and Stanley, J. 2003. Occurrence of South African cassava mosaic virus (SACMV) in Zimbabwe. *New Disease Reports* 8: Rep. [<http://www.bspp.org.uk/ndr/>]
- Brown, J. K. and Bird, J. 1992. Whitefly transmitted geminiviruses and associated disorders in the Americas and the Caribbean basin. *Plant Disease* **76**: 220-225.
- Brown, J. K., Frohlich, D. R. and Rosell, R. C. 1995. The sweet potato or silver leaf whiteflies: Biotypes of *Bemisia tabaci* or a species complex? *Annual Review of Entomology* **40**: 511-534.
- Brown, J.K. 2000. Molecular markers for the identification and global tracking of whitefly vector-begomovirus complexes. *Virus Research* **71**: 233-260.
- Brown, J.K. 2001. The Molecular Epidemiology of Begomoviruses. Pages **279-316** in: *Trends in Plant Virology* (J. A. Khan and J. Dykstra), The Haworth Press, Inc., NY. 537pp.
- Bull, S. E., Briddon, R. W., Sserubombwe, W. S., Ngugi, K., Peter G. Markham, P. G. and Stanley, J. 2006. Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. *Journal of General Virology* **87**: 3053-3065
- Burban, C., Fishpool, L. D. C., Fauquet, C., Fargette, D. and Thouvenel, J.-C. 1992. Host associated biotypes within West African populations of the whitefly *Bemisia tabaci* (Genn.) (Hom., Aleyrodidae). *Journal of Applied Entomology* **113**: 416-423.
- Campbell, B.C., Steffen-Campbell J. D. and Gill, R. J. 1994. Evolutionary origin of whiteflies (Hemiptera: Sternorrhyncha: Aleyrodidae) inferred from 18S rDNA sequences. *Insect Molecular Biology*, **3**: 73-89.

- Colvin, J., Otim-Nape, G.W., Holt, J., Omongo, C., Seal, S., Stevenson, P., Gibson, G., Cooter, R. J. and Thresh, J. M. 1999. Factors driving the current epidemic of severe cassava mosaic disease in East Africa. In: *VIIIth International Plant Virus Epidemiology Symposium – Plant Virus Epidemiology: Current Status and Future Prospects, 1999.* : International Society of Plant Pathology, 76–7.
- Colvin, J., Omongo, C. A., Maruthi, M. N., Otim-Nape G. W. and Thresh, J. M. 2004. Dual begomovirus infections and high *Bemisia tabaci* populations: two factors driving the spread of a cassava mosaic disease pandemic. *Plant Pathology* **53**: 577–584
- Corbett, G.H. 1935. On new Aleurodidae (Hem.) *Annals and Magazine of Natural History* **16**: 240–252.
- Costa, H. S and Brown, J. K. 1991. Variation in biological characteristics and in esterase patterns among populations of *Bemisia tabaci* (Genn.) and the association of one population with silverleaf symptom development. *Entomologia experimentalia et applicata* **61**:211-219.
- Danzig, E.M. 1966. The whiteflies (Homoptera, Aleyrodoidea) of the southern Primor'ye (Soviet Far East). *Entomologicheskoe Obozrenie* **45**: 364–386. [English translation in *Entomological Review. Washington*, 45, 197–209.
- De Barro, P. J., Liebrechts, W. and Carver, M. 1998. Distribution and identity of biotypes of *B. tabaci* (Gennadius) (Hemiptera: Aleyrodidae) in member countries of the Secretariat of the Pacific Community. *Australian Journal of Entomology* **37**: 214-218.

- De Barro, P. J., Driver, F., Trueman, J. W. H. and Curran, J. 2000. Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. *Molecular Phylogenetics and evolution* **16**: 29-36.
- De Barro, P. J., Trueman, J. W. H. and Frohlich, D. R. 2005. *Bemisia argentifolii* is a race of *B. tabaci* (Hemiptera: Aleyrodidae): the molecular genetic differentiation of *B. tabaci* populations around the world. *Bulletin of Entomological Research* **95**: 193–203
- De La Rau, P., Simon, B., Cifuentes, D., Martinez-Mora, C. and Cenis, J. L. 2006. New insights into the mitochondrial phylogeny of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in the Mediterranean Basin. *JZS* **44**: 25-33.
- Deng, D., McGrath, P. F., Robinson, D. J. and Harrison, B. D. 1994. Detection and differentiation of whitefly-transmitted geminiviruses in plants and vector insects by the polymerase reaction with degenerate primers. *Annals of Applied Biology* **125**: 327-336.
- Dengel, H. J. 1981. Untersuchungen Ueber das Auftreten der Imagines von *Bemisia tabaci* (Genn.) auf verschiedenen Manioksorten. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz* **88**: 355–366.
- Doyle, J. J. and J. L. Doyle. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem Bull.* **19**:11-15.
- Doughty, L. R., 1958. Cassava breeding for resistance to mosaic and brown streak viruses: a review of 21 years' work. *Annual Report, East African Agricultural and Forestry Research Organisation for 1958*. 48–55.

- Dubern, J. 1994. Transmission of African cassava mosaic geminivirus by the whitefly (*Bemisia tabaci*). *Tropical Science* **34**: 82-91.
- FAO. 2006. Cassava production statistics. <http://www.fao.org>
- Fargette, D., Fauquet, C. and Thouvenel, J. C. 1985. Field studies on spread of African cassava mosaic. *Annals of Applied Biology* **106**: 285-294.
- Fargette, D., Fauquet C. and Thouvenel, J. C. 1988. Yield losses induced by African cassava mosaic virus in relation to the mode and the date of infection. *Tropical Pest Managemet* **34**: 89-91.
- Fargette, D., Muniyappa, V., Fauquet, C., N'Guessan, P. and Thouvenel J.C. 1993. Comparative epidemiology of three tropical whitefly-transmitted geminiviruses. *Biochemie* **75**: 547-554.
- Fauquet, C. M and Fargette, D. 1990. African cassava mosaic virus: etiology, epidemiology and control. *Plant Disease* **74**: 404-411.
- Fauquet, C. M. and Stanley, J. 2003. Geminivirus classification and nomenclature: progress and problems. *Annals of Applied Biology*. **142**: 165–189.
- Fondong, V. N., Pita, J. S., Rey, C., Beachy, R. N. and Fauquet, C. M. 1998. First report of the presence of East African cassava mosaic virus in Cameroon. *Plant Disease* **82**: 1172.
- Fondong, V. N., Pita, J. S., Rey, M. E. C., de Kochko, A., Beachy, R. N. and Fauquet, C. M. 2000. Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *Journal of General Virology* **81**: 287-297.

- Frohlich, D. R., Torres-Jerez, I., Bedford, I. D., Markham, P. G. and Brown, J. K. 1999. A phylogeographical analysis of *Bemisia tabaci* species complex based on mitochondrial DNA markers. *Molecular Ecology* **8**: 1683-1691.
- Frischmuth, T. and Stanley J. 1992. Characterization of beet curly top virus subgenomic DNA localizes sequences required for replication. *Virology* **189**: 808-811.
- Gibson, R. W., Legg, J. P. and Otim-Nape, G. W. 1996. Unusually severe symptoms are a characteristic of the current epidemic of mosaic virus disease of cassava in Uganda. *Annals of Applied Biology* **128**: 479-490.
- Hagblade, S. and Zulu, B. 2003. The recent cassava surge in Zambia and Malawi. In InWent, IFPRI, NEPAD, CTA *Conference Success in African Agriculture*. Pretoria, South Africa.
- Haible, D., Kober, S. and Jeske H. 2006. Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses. *Journal of Virological Methods* **135**: 9-16.
- Hanley-Bowdoin, L., Settlege, S. B. & Robertson, D. 2004. Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. *Molecular Plant Pathology* **5**: 149–156.
- Harrison, B. D., Barker, H., Bock, K. R., Guthrie, E. J., Meredith, G. and Atkinson, M. 1977. Plant viruses with circular single-stranded DNA. *Nature* **270**: 760- 762.
- Harrison, B. D. 1985. Advances in geminivirus research. *Annual Review of Phytopathology* **23**: 55-82.

- Harrison, B. D., Swanson, M. M., McGrath, P. F. and Fargette, D. 1991. Patterns of antigenic variation in whitefly transmitted geminiviruses. *Report of the Scottish Crop Research Institute for 1990*. 88-90.
- Harrison, B. D., Zhou, X., Otim-Nape, G. W., Liu, Y. and Robinson, D. J. 1997. Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. *Annals of Applied Biology* **131**: 437-448.
- Harrison, B. D. and Robinson, D. J. 1999. Natural, Genomic and Antigenic Variation in the Whitefly Transmitted Geminiviruses (begomoviruses). *Annual Review of Phytopathology* **37**: 369-398
- 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. and Harrison, B. D. 1993. Nucleotide sequence evidence for the occurrence of three distinct whitefly-transmitted geminiviruses in cassava. *Journal of General Virology* **74**: 2437-2443.
- Jericho, C., G.J. Thompson, U. Gerntholtz, and J.C. Viljoen. 1999. Occurrence and distribution of cassava diseases in South Africa. In: Akoroda, M. O., Teri, J. M. (Eds), *Proceedings of the Scientific Workshop of the Southern African Root Crops Research Network (SARRNET)*, Lusaka, Zambia, 17-19August 1998. 252-262.
- Koonin, E. V., and Ilyina, T. V. 1992. Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA initiation proteins. *Journal of General Virology* **73**: 2763 - 2766.

- Kumar, S., Tamura, K. and Nei, M. 2004. MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment *Briefings in Bioinformatics* **5**:150-163.
- Legg, J. P. and Ogwal, S. 1998. Changes in the incidence of African cassava mosaic virus disease and the abundance of its whitefly vector along south-north transects in Uganda. *Journal of Applied Entomology* **122**: 169-178.
- Legg, J. P. and Raya, M. 1998. A survey of cassava virus diseases in Tanzania. *International Journal of Pest Management* **44**: 17–23.
- Legg J. P., 1999. Emergence, spread and strategies for controlling the pandemic of cassava mosaic virus disease in east and central Africa. *Crop Protection* **18**, 627-237.
- Legg, J. P., Okao-Okuja, G., Mayala, R., Muhinyuza J. B. 2001. Spread into Rwanda of the severe cassava mosaic virus disease pandemic and associated Uganda variant of *East African cassava mosaic virus* (EACMV-Ug). *Plant Pathology* **50**: 796.
- Legg, J. P. and 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.
- Legg, J. P., Owor, B., Sseruwagi, P., and Ndunguru, J. 2006. Cassava mosaic virus disease in East and Central Africa: Epidemiology and management of a regional pandemic. *Advances in Virus Research* **67**: 335-418.
- Laufs, J., Schumacher, S., Geisler, N., Jupin, I., and Gronenborn, B. 1995. Identification of the nicking tyrosine of geminivirus Rep protein. *FEBS letter* **18**: 258 - 2262.

- Lukman, R. 2003. Molecular mapping of major genes influencing flowering time in wheat (*Triticum aestivum* L. em. Thell). Technischen Universität München. PhD thesis: 14-15.
- Martin, D. P. and Rybicki, E. P. 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**: 562-563
- Mayo, M. A., Berns, K. I., Fritsch, C., Kaper, J. M., Jackson, A. O., Leibowitz, M. J. and Taylor, J. M. (1995) in *Virus taxonomy. Classification and nomenclature of viruses. Sixth Report of the International Committee on Taxonomy of Viruses.*, ed. Fauquet, A. M. F. (Springer, New York), pp. 487–492.
- Mayo, M. A. 2002. Virus taxonomy—Houston 2002. *Archives of Virology* **147**:1071-1076.
- Maruthi, M. N, Colvin, J., Seal, S., Thresh, J. M. 2002. First report of a distinct begomovirus infecting cassava from Zanzibar. *Plant Disease* **86**, 187.
- Mayo, M. A., Berns, K. I., Fritsch, C., Kaper, J. M., Jackson, A. O., Leibowitz, M. J. and Taylor, J. M. (1995) in *Virus taxonomy. Classification and nomenclature of viruses. Sixth Report of the International Committee on Taxonomy of Viruses.*, ed. Fauquet, A. M. F. (Springer, New York), pp. 487–492.
- Mound, L.A. and Halsey, S. H. 1978. Whitefly of the world: A systematic catalogue of the Aleyrodidae (Homoptera) with host plant and natural enemy data. *British Museum (Natural History)*: Chichester. 329p.
- Mound, L. A. (1983) Biology and identity of whitefly vectors of plant pathogens. In: *Plant virus epidemiology. The spread and control of insect-borne viruses.* Blackwell Scientific Publications, Oxford, United Kingdom. 305-313.

- Munga, T. and J.M. Thresh. 2002. Incidence of cassava mosaic and cassava brown streak virus diseases in coastal Kenya. *Roots* **8**(1): 12–14.
- Ndunguru, J., Legg, J. P., Aveling, T. A. S., Thompson, G. and 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. *Virology Journal* **2**: 2-21
- Ndunguru, J., Legg, J. P., Fofana, I. B. F., Aveling, T. A. S., Thompson, G. and Fauquet, C. M. 2006. Identification of a defective molecule derived from DNA-A of the bipartite begomovirus of *East African cassava mosaic virus*. *Plant Pathology* **55**: 2–10
- Okao-Okuja, G., Legg, J. P. Traore, L. and Alexandra Jorge M. 2004. Viruses Associated with Cassava Mosaic Disease in Senegal and Guinea Conakry. *Journal of Phytopathology* **152**:2, 69–76
- Otim-Nape, G. W., Bua, A., Thresh, J. M., Baguma, Y, Ogwal, S., Semakula, G. N., Acola, G., Byabakama, B. and Martin, A. 1997. Cassava mosaic viruses in Uganda: The current pandemic and approaches to control. Natural Resources Institute, Chatham, UK. Pp 65.
- 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.
- Padidam, M., Sawyer, S. and Fauquet, C. M. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**: 218–225.
- Perring, T. M. 2001. The *Bemisia tabaci* species complex. *Crop Protection* **20**: 725-737.

- Perring, T. M., Cooper, A. D., Rodriguez, R. J., Farrar, C. A. and Bellows, T. S. 1993. Identification of a whitefly species by genomic and behavioural studies. *Science* **259**: 74-77.
- Pita, J. S., Sangare, A., Beachy, R. N. and Fauquet, C. M. 1998. Cassava Mosaic disease (CMD) in Africa: a parallel study between Uganda (UG) and Ivory Coast (IC). *Plant Diseases* **88**: 571.
- Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G.W., Ogwal, S. and Fauquet C. M. 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology* **82**: 655–665
- Puonti-Kaerlas, J. 1998. Cassava Biotechnology. *Biotechnology and Genetic Engineering Reviews*. **15**: 329-364
- Prabhaker, N. Coudriet, D. L. and Meyerdick, D. E. 1987. Determination of three whitefly species by electrophoresis of non-specific esterases. *Journal of Applied Entomology* **103**: 447-451.
- Pringle, C. R. 1999. Virus taxonomy – 1999 – The universal system of virus taxonomy, updated to include the new proposals ratified by the International Committee on Taxonomy of Viruses during 1998. *Archives of Virology* **144**: 421-429
- Quaintance, A.L. 1900. Contribution towards a monograph of the American Aleurodidae. *Technical Series, United States Department of Agriculture Bureau of Entomology* **8**: 9–64.

- Quaintance, A.L. and Baker, A.C. 1914. Classification of the Aleyrodidae Part II. *Technical Series, United States Department of Agriculture Bureau of Entomology* **27**: 95–109.
- Ranomenjanahary, S., Rabindran, R. and Robinson, D. J. 2002. Occurrence of three distinct begomoviruses in cassava in Madagascar. *Annals Applied Biology* **140**, 315–318.
- Reiter, R. S., Williams, J. G. K., Feldmann, K. A., Rafalski, J. A., Tingey, S. V. and Scolnik, P. A. 1992 Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNA's . *Proc. Natl. Acad. Sci. USA* , **89**: 1477-1481.
- Robertson, I. A. D. 1985. The role of *Bemisia tabaci* (Gennadius) in the epidemiology of ACMV in East Africa. Biology, population dynamics and interaction with cassava varieties. In: *Proceedings of the International Seminar on African Cassava Mosaic Disease and its Control*, pp 57–63
- Rojas, M. R., Gilbertson, R. L., Russell, D. R. and Maxwell, D. P. 1993. Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Disease* **77**: 340-347.
- Ronda, M. A., Adan, A., Cifuentes, D. Cenis, J. L. and F. Beitia.1999. Laboratory evidence of interbreeding between biotypes of *Bemisia tabaci* (Homoptera, Aleyrodidae) present in Spain. In: pp 83–84, VIIth International Plant Virus Epidemiology Symposium-Plant Virus Epidemiology: Current Status and Future Prospects 1999; Aguadulce, Spain.

- Rothenstein, D., Haible, D., Dasgupta, I., Dutt, N., Patil, B, L. and Jeske, H. 2006. Biodiversity and recombination of cassava-infecting begomoviruses from southern India. *Archives of Virology* **151**: 55-69.
- Russell, L.M. 1957. Synonyms of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae). *Bulletin of the Brooklyn Entomological Society* **52**: 122-123.
- Russell, L.M. 1958. *Orchamoplatus*, an Australasian genus (Homoptera: Aleyrodidae). *Proceedings of the Hawaiian Entomological Society* **16**: 389-410.
- Rybicki, E. P. 1994. A phylogenetic and evolutionary justification for three genera of Geminiviridae. *Archives of Virology* **139**: 49-77.
- Sanderfoot, A. A. and Lazarowitz, S. G. 1996. Getting it together in plant virus movement: Cooperative interactions between bipartite geminivirus movement proteins. *Trends in Cell Biol.* **6**: 353-358
- Saunders, K., Salim, N., Mali, V. R., Malathi, V. G., Briddon, S. R., Markham, P. G. and Stanley, J. 2002. Characterisation of *Sri Lankan cassava mosaic virus* and *Indian cassava mosaic virus*: Evidence for acquisition of a DNA B component by a monopartite begomovirus. *Virology* **293**: 63-74.
- Sequeira, J. C. and Harrison, B. D. 1982. Serological studies on cassava latent virus. *Annals of Applied Biology* **101**: 33-42.
- Simon, J. C., Baumann, S. and Sunnucks, P. 1999. Reproductive mode and population genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Molecular Ecology* **8**: 531-545
- Sseruwagi, P., Aritua, V., Otim-Nape, G. W. and Legg, J. P. 1998. Epidemic characteristics of cassava mosaic disease in Uganda: a 1998 update. In

- Proceedings of the Seventh Triennial Symposium of the International Society for Tropical Root Crops - Africa Branch*, pp 11–17. Eds M. O. Akoroda and I. J. Ekanayake.
- Sseruwagi, P., Sserumbombwe, W. S., Legg, J. P., Ndunguru, J. and 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 Research* **100**: 129-142.
- Stanley, J. and Gay, M. R. 1983. Nucleotide sequence of cassava latent virus DNA. *Nature* **301**: 260- 262.
- Stanley, J. and Townsend, R. 1986. Infectious mutants of cassava latent virus generated *in vivo* from intact recombinant DNA clones containing single copies of the genome. *Nucleic Acids Research* **14**: 5981- 5998.
- Stanley, J. 1995. Analysis of African cassava mosaic virus recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. *Virology* **206**: 707-712.
- Stanley, J., Bisaro, D. M., Briddon, R. W., Brown, J. K., Fauquet, C. M., Harrison, B. D., Rybicki, E. P. and Stenger, D. C. 2005. *Geminiviridae*. In *Virus Taxonomy. VIIIth Report of the International Committee on Taxonomy of Viruses*, pp. 301–326. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. London: Elsevier/Academic Press.
- Stenger, D. C., Stevenson, M. C., Hormuzdi, S. G, Bisaro, D. M. 1992. A number of subgenomic DNA's are produced following agroinoculation of plants with beet curly top virus. *Journal of General Virology* **73**: 237-242.

- Storey, H. H. 1936. Virus diseases of East African plants. VI-A progress report on studies of the disease of cassava. *East African Agricultural Journal* **2**: 34-39.
- Storey, H. H. and Nichols, R. F. W. 1938. Studies of the mosaic disease of cassava. *Annals of Applied Biology* **25**: 790-806.
- Swafford, D. L. 1991. PAUP: phylogenetic analysis using parsimony, 3.0 ed. Illinois Natural History Survey, Champaign, Ill.
- Swanson, M. M. and Harrison, B. D. 1994. Properties, relationships and distribution of cassava mosaic geminiviruses. *Tropical Science* **34**: 15-25.
- Thresh, J. M. 2001. Report on visits to Mozambique. National Resources Institute. <http://www.cpp.uk.com/downloads/cbsd/btor1.pdf>
- Trench, T. N. and Martin, M. M. 1985. An assessment of cassava African mosaic disease in South Africa and Swaziland. *South African Journal of Plant and Soil* **2**: 169-170.
- Vanitharani R., Chellappan P., Pita J. S and Fauquet, C. M. 2004. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and posttranscriptional genome silencing suppression. *Journal of Virology* **78**: 9487–9498
- Woodward, B. R.; Allemann, J. and O' Regan, B. P. 1997. Tissue culture of cassava: a South African perspective. *African Journal of Root and Tuber Crops* **2(1/2)**: 243-245

- Wyatt, S. D. and Brown, J. K. 1996. Detection of subgroup III geminivirus isolates in leaf extracts by degenerate primers and polymerase chain reaction. *Phytopathology* **86**: 1288-1293.
- Zhang, W., Olson, N. H., Baker, T. S., Faulkner, L., Agbandje-McKenna, M., Boulton, M., Davies, J. W and McKenna, R. 2001. Structure of the maize streak virus geminate particle. *Virology* **279**: 471-477.
- Zhou, X., Robinson, D. J. and Harrison, B. D. 1998. Types of variation in DNA-A among isolates of East African cassava mosaic virus from Kenya, Malawi and Tanzania. *Journal of General Virology* **79**: 2835-2840.
- Zhou, X., Liu, Y., Calvert, L., Munoz, D., Otim-Nape, G. W., Robinson, D. J. and Harrison, B. D. 1997. Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by inter-specific recombination. *Journal of General Virology* **78**: 2101-2111.
- Zimmermann, A. 1906. Die Krauselerkrankheit des Maniok. *Pflanzer* **2**: 145