# Epidemiology of Cassava Mosaic Disease in Mozambique

**Nurbibi Saifodine Cossa**

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

Johannesburg, 25th February 2011

# Declaration

I declare that this dissertation is my own, 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 at any other University.

______________________________________

Nurbibi Saifodine Cossa

25th February, 2011

## ABSTRACT

Cassava *(Manihot esculenta* Crantz*)* belongs to the family *Euphorbiaceae*. The crop was introduced to África, including Mozambique, in the 17th century by Portuguese traders. Cassava is one of the most important crops in Africa. It is widely grown in many countries of sub-Saharan Africa and Madagascar, and plays a major role in food security and poverty alleviation.

In Mozambique, cassava is the second most important root crop and used exclusively as human consumption: it is estimated that 70% of the total cassava production is utilized as human food. Cassava is chiefly vegetatively propagated from stem cuttings. The roots and leaves provide a major source of carbohydrate and vitamins. In areas prone to draught and floods, cassava is the main crop.

Cassava mosaic disease (CMD) was reported for the first time in Tanzania in 1894. CMD has since been observed throughout all cassava growing areas of sub-Saharan Africa and its adjacent islands, and occurs throughout Mozambique. The disease is caused by a single-stranded circular bipartite DNA geminivirus belonging to the Family *Geminiviridae* and Genus: *Begomovirus*. It is transmitted by the whitefly vector *Bemisia tabaci* Gennadius. Cassava begomoviruses are highly diverse and currently nine species have been identified. Limited surveys of CMD have been performed in Mozambique, and the widespread loss of crop yield threatens farmers in the region.

This study therefore performed epidemiological studies on CMD in six major cassava-growing provinces to evaluate the incidence and severity of disease. Transmission of CMD by either cuttings or whitefly was estimated, and the genetic identities of the *B. tabaci* vector and begomoviruses were established.

Two surveys were conducted in the 2005/2006 and 2006/2007 growing seasons in six provinces namely, Maputo, Gaza, Inhambane, Nampula, Zambézia and Cabo Delgado. The study showed that CMD incidence recorded was highly variable within districts and between provinces, in both surveys, witth mean incidence of 6% in Nampula and 75% in the Gaza Province.

The whitefly infection was very low (0 to 3%) compared to cutting infection (22 to 60%). Mean whitefly numbers per plant were variable but also low (0-15.8) in all the provinces in

the two surveys. The mean CMD severity score ranged from 2 to 3.3 in the six provinces in the two surveys, which indicated mild to severe symptoms in the cassava-growing regions in Mozambique. A score of 4-5 usually indicates very severe symptoms typical of mixed infections or several species or isolates of cassava begomoviruses. CMD was found to be more prevalent in the southern (Maputo, Zambezia and Gaza) than the northern provinces. These results indicated that cutting infection is more important than the whitefly infection because the farmers plant infected cuttings repeatedly over the years, causing spread of the disease.

In this study, the *B. tabaci* from the six provinces in Mozambique formed a single unified group (with little sequence divergence) within the Sub-Saharan African clade. Specifically the Mozambique haplotypes clustered with the southern African haplotypes from South Africa, Swaziland and Zambia.

In order to determine the biodiversity of CMGs in Mozambique, 285 infected cassava leaf samples (showing symptoms) were collected throughout the six provinces in Mozambique, and from sixty selected CCP- positive samples, full-length DNA A amplification was performed (either by Rolling Circle Amplification (RCA)-PCR or PCR from Total Nucleic Acid) from 55 of the leaf samples, and Restriction Length Fragment Polymorphism (RLFP) analysis was undertaken using the enzymes *EcoRV, DraI* and *MluI.* The results showed that 63.4% (35 of the 55 cassava leaf samples) were EACMV species; 12.7% (6/55) were ACMV species; 10.9% (7/55) were mixed infections of EACMV and ACMV; and 6 samples were unidentified by RFLP.

Six full-length DNA A clones (4 each), one from each province, showing unusual RFLP patterns, were cloned and sequenced. Consensus sequences of six full-length DNA A clones (4 each), one from each province, were aligned with other cassava begomoviruses and selected begomoviruses from southern Africa. Phylogenetic analysis (parsimony) revealed that virus isolates from Maputo, Inhambane and Nampula Provinces exhibited 95-97% nucleotide sequence divergence/similarity to *African cassava mosaic virus*-[Nigeria]; the virus isolate from Gaza Province was 99% similar to *South African cassava mosaic virus* - [South Africa]; while the Zambezia Province virus was most closely aligned (94%) with EACMMV (*East African cassava mosaic Malawi virus*- [Malawi: MH]), and the isolate from the Cabo Delgado Province aligned most closely (96%) with *East African cassava mosaic Cameroon virus*- Cameroon and less closely (87%) to EACMMV.

This study reports diversity for cassava begomovirus species in Mozambique for the first time. It is similar to previous studies in RFLP analysis that indicated that the cassava geminiviruses in Mozambique are mainly genetic variants of ACMV and EACMV, although full length DNA A sequences indicated that EACMMV and EACMCV are present in Mozambique. These results, as in South Africa, demonstrate the mixture of geminivirus species from east and west Africa. This is the first report of SACMV in Mozambique.

# Dedication

I dedicate this thesis to my deceased husband Boaventura Celestino Langa Cossa, who was a very kind and helpful man. He is my inspiration and I loved him very much. This thesis was produced in his honour.

# Acknowlegments

First of all I would like to say thank you to God who gave me strength  and light to do this work and second, to my supervisors,  Prof. Chrissie Rey and Ana Mondjana,  who gave me this opportunity, and the patience she  demonstrate along the way.

I would also like to thanks my kids, Rui Jorge and Patricia, for their patience and the support they gave to me, during the stressed times when I was working on this dissertation. A word of thanks to my mother, brothers and sisters, specially to mana Zauria Saifodine," muito obrigada pela noites perdidas a ajudar-me na formatação da tese".

Thanks to CDR USAID, and the Rockefeller Foundation, for the financial support given to the project which on the other hand was used to cover some of the costs incurred during my research.

Thanks to collegues from the Wits MCB staff especially Maabo, Natasha Imah and Farhana, I have no words to describe my gratitude for your support, God bless you.

Thanks to collegues from the national and provincial Departments of Agriculture located in the Maputo Province, Gaza, Inhambane, Nampula, Zambezia and Cabo Delgado Provinces. Thank you very much for your endeless support.

# Table of contents

Declaration................................................................................................................ II

Dedication................................................................................................................VI

Acknowlegments ......................................................................................................VII

List of Figures...........................................................................................................XI

List of Tables .......................................................................................................... XIII

List of Abreviations ................................................................................................XIV

Rationale for study and objectves..........................................................................XVII

**CHAPTER ONE: LITERATURE REVIEW .......................................................... 1**

1.1 General Introduction ................................................................................ 1

1.2 Economic Importance of Cassava in Africa ............................................... 2

1.3 Cassava Mosaic Disease .......................................................................... 3

1.3.1 Transmission of CMD .............................................................................. 4

1.3.2 Symptoms CMD ...................................................................................... 4

1.3.3 Epidemiology of CMD .............................................................................. 6

1.3.4 Geminiviruses .......................................................................................... 7

1.3.5 Begomovirus Genome Organisation.......................................................... 8

1.3.6 Replication ............................................................................................... 9

1.3.7 Recombination and Pseudorecombination .............................................. 10

1.3.8 Satellite DNAs ........................................................................................ 11

1.3.9 Defective interfering (DI) molecules ........................................................ 12

1.3.10 Cassava Geminivirus Diversity and Nomenclature .................................. 13

1.4 Identification and Diagnostic Methods.................................................... 15

1.5 The Whitefly Vector (Bemisia tabaci)...................................................... 16

1.5.1 *Bemisia tabaci* Host Range......................................................................... 17

1.5.2 Whitefly Biotypes ........ 18

**CHAPTER TWO: EPIDEMIOLOGY OF CASSAVA MOSAIC DISEASE AND THE WHITEFLY VECTOR *BEMISIA TABACI* (GENNADIUS) ........ 21**

2.1 Abstract ........ 21

2.2 Introduction ........ 22

2.3 Methodology ........ 24

2.3.1 Geographical location of surveys ........ 24

2.4 Data collection ........ 25

2.4.1 Whitefly collection ........ 26

2.4.2 Statistical Analyses ........ 27

2.4.3 Molecular Characterization of Whitefly Samples ........ 28

2.4.4 Phylogenetic analysis ........ 28

2.5 Results ........ 29

2.5.1 CMD Incidence ........ 29

2.5.2 CMD severity and whitefly abundance ........ 34

2.5.3 Phylogenetic analysis of whiteflies ........ 38

2.6 Discussion and Conclusions ........ 40

**CHAPTER THREE: IDENTIFICATION OF BEGOMOVIRUSES IN CASSAVA IN THE SIX PROVINCES IN MOZAMBIQUE ........ 46**

3.1 Abstract ........ 46

3.2 Introduction ........ 47

3.3 Materials and Methods ........ 51

3.3.1 Sampling using FTA cards and leaves ........ 51

3.3.2 Preparation of cards for PCR ........ 51

3.3.3 Total nucleic Acid (TNA) Extraction ........ 52

3.3.4 PCR screening of leaf samples for the core coat protein (CCP) ........................... 52
3.3.5 Identification of CMGs using PCR- RFLPs ........................................................ 53
3.3.6 Phylogenetic analysis of selected geminiviruses ................................................. 54
3.4 Results ....................................................................................................................... 55
3.4.1 Core Coat Protein (CCP) - PCR ........................................................................... 55
3.4.2 Near-full length DNA-A Amplification .............................................................. 56
3.4.3 Restriction Length Fragment Polymorphism (RLFP) analyses ...................................... 57
3.4.4 Phylogenetic Analysis of Sequenced DNAs From six Geminivirus Samples ....... 60
3.5 Discussion .................................................................................................................. 63
**SUMMARY AND CONCLUSIONS ........................................................................ 68**
**REFERENCES ......................................................................................................... 72**
**Appendices .............................................................................................................. 87**
Appendix 1: Incidence, severity and whitefly numbers, and statistical (anova) data from all the 6 provinces surveyed. ............................................................ 87
Appendix 2: Nucleotide sequences of full length DNA A from 6 geminivirus isolates from Mozambique (one from each province) ........................................................... 101
Appendix 3: Percentage nucleotide sequences identities of geminivirus isolates from Mozambique compared to selected virus isolates from different countries ........ 108
Appendix 4: Table 4.1: Percentage nucleotide sequence identities of Bemisia tabaci isolates from Mozambique compared to selected isolates from different countries ........................................................................................... 132

# List of Figures


**Figure 1.1:** Cassava Mosaic Disease Symptoms 5

**Figure 1.2:** Organization of African cassava mosaic virus (ACMV) (*Begomovirus*) 8

**Figure 2.1:** Main cassava growing areas where the survey was conducted 29

**Figure 2.2:** Histogram showing the incidence of CMD in six Provinces in the 2005 Survey 31

**Figure 2.3:** Histogram showing the incidence of CMD in six provinces in the 2006 survey 33

**Figure 2.4:** CMD severity and mean whitefly numbers per plant in the six provinces from the 2005 survey 35

**Figure 2.5:** CMD severity and mean whitefly numbers per plant in the six provinces from the 2006 survey 37

**Figure 3.2:** 1% agarose gel of near full-length virus DNA-A amplification by RCA from 16 samples amplified from RCA products using Uni primers. M: 1kb DNA molecular weight marker (0□ GeneRuler Ladder Plus- Fermentas) 56

**Figure 3.1:** PCR amplification of a 550bp fragment amplified from tna extracted from infected cassava leaf samples. Amplicons were amplified using the CCP primers (Results shown above are for only 19 of the 60 samples). M: 1kb DNA molecular weight marker (0□ Gene Ruler LadderPlus Fermantas) 56

**Figure 3.4:** Characterisation of RFLP patterns of 18 CMG DNA-A samples restricted with *EcoR*V. Lane 6: No template control. M: 1kp DNA marker ladder (Fermentas) 57

**Figure 3.3:** Full-length DNA A amplified by PCR (Uni primers) from TNA. M: 1kb DNA molecular weight marker (0□ Gene Ruler Ladder Plus- Fermentas); Lanes 17-21 and 23-26 : ±2.8bp DNA A amplified bands; Lane 22 : Negative control; Lane 23: positive control. 57

**Figure 3.5:** Characterisation of RFLP patterns of 15 CMG DNA-A samples (lanes 1-15) restricted with *Mlu*I. M: 1kp DNA marker ladder (Fermentas) 58

**Figure 3.7:** *Pst*I digestion of the amplified CMGs to screen for the presence of SACMV. Lanes 1-4: Uncut DNA A from 4 samples. 59

**Figure 3.6:** Characterisation of RFLP patterns of 16 CMG DNA-A samples (lanes 1-15) restricted with *Dra*I. M: 1kp DNA marker ladder (Fermentas) 58

**Figure 3.8:** Most parsimonious tree (CLUSTAL V) derived from full-length DNA A comparisons between geminivirus samples from six provinces in Mozambique (C7Maputo; C45Nampula; C23C Inhambane; C16Zambezia; C54Gaza and C50Cabo Delgado) and selected cassava geminiviruses. Accession numbers and virus abbreviations are listed in Appendix 3. *Tomato curly stunt virus* from South Africa was used to root the tree. Bootstrap values are generated from 1000 replicates 61

**Figure 3.9**: Maximum likelihood tree constructed in PAUP, showing the relationship between Mozambique cassava isolates and cassava geminivirus species (ACMV and EACMV) and selected geminiviruses from the African continent. Accession numbers and virus abbreviations are listed in Appendix 3. *Tomato leaf curl Madagascar virus* was used to root the tree. Bootstrap values are generated from 1000 replicates. 62

# List of Tables

**Table 2.1:** Provinces and Districts where the surveys were conducted .............................. 25

**Table 2.2:** Whitefly sample collections from districts in the six provinces chosen for mtCO1 sequencing ........................................................................................ 27

**Table 2.3:** Incidence (percentage) of cassava mosaic disease in six provinces in 2005 ..... 31

**Table 2. 4:** Incidence of CMD in six provinces in the 2006 survey .................................. 33

**Table 2. 5:** Disease severity and whitefly abundance in the 2005 survey .......................... 35

**Table 2.6:** CMD Severity and whitefly numbers in the 2006 survey ................................ 37

**Table 3.1:** Number of geminivirus species identified in each province using RLFPs ....... 59

# List of Abreviations

| | |
|---|---|
| **ACMV** | African cassava mosaic virus |
| **CBSD** | Cassava brown streak disease |
| **CBSV** | Cassava brown steak virus |
| **CCP** | Core coat protein |
| **CLV** | Cassava latent virus |
| **CMD** | Cassava mosaic Disease |
| **CMG's** | Cassava mosaic geminiviruses |
| **CMV** | Cassava mosaic virus |
| **CR** | Common region |
| **CSM** | Cassava starch manufacturing |
| **DAS (ELISA)** | Double antibody sandwich ELISA |
| **DNA** | Deoxyribonucleic acid |
| **DI** | Defective interfering DNA |
| **DS** | Disease severity |
| **dsDNA** | Double stranded DNA |
| **EACMV** | East African cassava mosaic virus |
| **EACMCV** | East African cassava mosaic Cameroon virus |
| **EACMMV** | East African cassava mosaic Malawi virus |
| **EACMZV** | East African cassava mosaic Zanzibar virus |
| **EDTA** | Ethylenediamine tetra-acetic acid |
| **ELISA** | Enzyme-linked immuno-absorbent assay |
| **FAO** | Food and Agriculture Organization |
| **ICMV** | Indian cassava mosaic virus |

| | |
|---|---|
| **IR** | Intergenic region |
| **ITS1** | Internal transcriber sequence |
| **KE** | Kenya |
| **MAbs** | Monoclonal antibodies |
| **mtCOI** | Mitochondrial cytochrome oxidase I |
| **MW** | Malawi |
| **ORF** | Origin of replication |
| **PCR** | Polymerase chain reaction |
| **RAPD** | Random amplified polymorphic DNA |
| **RCA** | Rolling circle amplification |
| **RDP** | Recombination detection program |
| **Rep** | Replicase |
| **RFLP** | Restriction fragment length polymorphism |
| **RNA** | Ribonucleic acid |
| **rDNA** | Ribosomal DNA |
| **SLCMV** | Sri-Lankan cassava mosaic virus |
| **ssDNA** | Single stranded DNA |
| **TAS - ELISA** | Triple antibody sandwich |
| **TGMV** | Tomato golden mosaic virus |
| **TNA** | Total nucleic acids |
| **TYLCV** | Tomato yellow leaf curl virus |
| **TZ** | Tanzania |
| **UgV** | Ugandan varian |

# Rationale for Study and Objectives

The main constraints of cassava production in Mozambique are soil fertility, rain distribution, quality of the planting material, and pests and diseases. The main pests affecting cassava in Mozambique are: Cassava mealybug *(Phenacoccus manihotti*), cassava green mites (*Mononnychellus spp*), cassava red mites (*Tetranychus spp*), cassava nematodes (*Meloidogyne spp*), cassava scale (*Aonidomytilus albus*) and the elegant grasshoppers (*Zonocerus spp*). With regard to diseases the most prevalent are: Cassava bacterial blight (*Xanthomonas campestris*), angular leaf spot *(Xanthomonas campestris*), cassava brown leafspot (*Cercosporidium henningssi*), cassava leaf blight (*Cercospora visosae*), Cassava tuber rot (*Sclerotium rolfsii*), Cassava soft root rot (*Phytophthora spp., Pythium spp. and Fusarium spp*), cassava brown streak and cassava mosaic disease (Théberge, 1985).

The most devastating and most economically important diseases in the country are Cassava brown streak disease (CBSD) and Cassava mosaic disease (CMD). Both diseases are of viral origin and cause enormous losses to cassava production (Thresh *et al*, 1999). CBSD was reported for the first time in 1999, when farmers were reporting major losses of cassava caused by root rotting which by 1998 were considerable, and had infected large areas of the coastal Nampula province in the northern part of the country.

As a result many farmers started to turn to alternative crops. Farmers in Zambézia Province, immediately to the south of Nampula, were finding the same problem of cassava root rotting. While CBSD is still confined to the north of the country, mainly in the coastal regions, CMD occurs all over the country where cassava is grown and causes enormous losses (Segeren and Van den Oever, 1993). Some of the local cultivars, widely grown by farmers, like Munhaça, Fernando Po, and Gangasol, are severely attacked by CMD (Cossa and Jonasse, 2002).

Cassava (*Manihot esulenta Crantz*) is one of most important source of carbohydrates in Mozambique as its tuber is rich in starch. It is produced in the whole country, with more emphasis in the northern Mozambique. It is one of the potential crops in biofuel production.The cassava crop in Mozambique has low yields due to many factors, some of which are, low quality of plant material, low soil fertility, bad crop management by the farmer, and pest and diseases.

One of the main objectives of the National Cassava Program at the Research Institute (IIAM), is to enhance and develop new germplasm, to produce technologies for cassava production and processing, and to use Integrated Pest Management as a control strategy. Not much research has been done on the epidemiology of CMD in Mozambique.

Surveys were carried out between 1999 and 2002 in Zambezia, Cabo Delgado, Nampula and Inhambane Provinces (Thresh, 1999-2001, Cossa and Mangana, 2002, Cossa and Jonasse, 2002), and by Toko *et al.* (2003). The first survey was only on the incidence of CMD and CBSD, followed by the second which was incidence and severity of CMD in Inhambane province and lastly incidence and severity in Cabo Delgado. The 2003 survey was on all cassava pests and diseases, not only on viral pathogens. These surveys did not provide detailed information on the virus diversity/identification (except for a few samples) and whitefly numbers, and the 2001 study only reported incidence of CBSD and CMD. Therefore it was important that a more detailed epidemiological study be undertaken to establish the CMD situation in Mozambique. Since the last surveys of CMD were done in 2003, and not all the provinces were covered, it was important that a more detailed epidemiological study be undertaken to establish the CMD situation in Mozambique.

The overall objective of this research was to study the epidemiology of Cassava Mosaic Disease (CMD) and to undertake molecular studies to establish the geminivirus (es) identities. Furthermore, the study aimed to undertake a study of the vector, *Bemisia tabaci,* of CMGs as this has had not been achieved in previous surveys.

## Specific Objectives

The specific objectives were as follows:

1. To evaluate the incidence and severity of CMD in six provinces.
2. To identify the virus species/isolates associated with CMD, using molecular techniques.
3. To determine mode of transmission (whitefly vs vegetative) and whitefly abundance and distribution.
4. To determine the identity of whitefly genotypes in the six provinces.

# CHAPTER ONE: LITERATURE REVIEW

## *1.1 General Introduction*

Cassava *(Manihot esculenta* Crantz*)* belongs to the family *Euphorbiaceae,* a shrubby tropical plant that grows tall, sometimes reaching 4.6m. The edibles parts are tuberous root and leaves. The tuber is dark brown in color and grows up to 600mm long; the leaves vary in shape and size. It is believed that the plant originated from South America and was introduced to Africa in the 16th Century and to Asia in the late 17th century by Portuguese traders (FAO, 2003).

Currently, cassava is grown in more than 80 countries between 30° South and 30° North of the equator (Fauquet and Fargette, 1990). It is suited to warm, humid lowland tropics. Cassava can grow in poorer soils compared to any other major food plant, so fertilization is rarely necessary if yields are acceptable. The crop normally requires a minimum temperature of 27°C to grow, and can survive during the dry season when the soil moisture is low but humidity is high. Although cassava tolerates drought, it grows best at rainfalls exceeding 1200 mm on many soil types and requires limited agronomic and pest management practices (Pounti-Kaerlas, 1998). Stem cutting are used for propagation, which usually starts at the beginning of the rainy season. Cuttings are selected from mature parts of the stem of parent plants.

Cassava is one of the most important crops in Africa, and it is widely grown in many countries of sub-Saharan Africa and Madagascar. It plays a major role in food security and poverty alleviation (Legg and Thresh, 2003). Cassava is grown on an estimated 80 million hectares, in 34 African countries; 84 million tones of cassava are produced per year in this region (FAO, 2003)

People in many ways consume more than 70% of the total production in Africa (FAO, 2005). It is an important crop in subsistence farming as it requires few production inputs, is drought tolerant, can grow in marginal soils under low rainfall conditions and produces reasonable yields under adverse conditions. The tubers are rich in carbohydrates and provide a good source of cheap calories, while the leaves are used as vegetable and are source of vitamins, minerals and proteins. It can also be used as source of ethanol for fuel, as animal feed and starch for industry (Hillocks, 2002).

## *1.2 Economic Importance of Cassava in Africa*

The World consumption of cassava for food (fresh or processed) is concentrated in developing countries. In Africa, about 70 percent of cassava production is used for consumption (FAO, 2005). It is a major staple food that is consumed by urban and rural people. In the drier regions of Sub-Saharan Africa cassava serves as a household security crop. A recent study indicated that in Africa cassava provides a larger income than any other crop in cassava production areas. The crop is sold in the markets and used for animal feed. The leaves of cassava are utilized as a vegetable and the roots eaten fresh and can be processed into over one hundred different food types. Certain cassava varieties are mainly grown for industrial uses (FAO, 2005).

Cassava was introduced into Mozambique by the Portuguese in the 17th century and was adopted as a food crop by Tsonga tribesman in the north. Later it spread to East Africa through Madagascar. Cassava is also extensively grown in Tanzania, Zambia, Malawi and Zimbabwe. It was probably introduced into Zambia via the Congo basin (where it was well established by the early 1650s), and in Zimbabwe and Malawi via Portuguese trading routes from Mozambique on the east coast of Africa (Haggblade and Zulu, 2003).

In Mozambique, cassava is the second most important root crop and is used exclusively (100%) for human consumption (Zacarias, 2006). Roots and leaves provide a major source of carbohydrate and vitamins. In areas prone to draught and floods, cassava is the main crop. It occupies the first place in terms of volume of production. It is the basic food crop for more than 50% of the population in the northern and central regions of the country (Cassava Strategy Document, 2003).

Cassava is cultivated throughout the country. However in terms of quantity, and production, it is mainly concentrated in four provinces: Cabo Delgado, Nampula, Zambezia and Inhambane. These provinces contribute about 93% of the national production (Mader, 2005). Cassava is the main staple food in Mozambique. It is generally considered a food security crop for most of the population in rural areas, particularly because of its adaptability to cultivation under harsh conditions. Cassava is being introduced along with sweet potato under a government initiative in drought prone areas throughout Mozambique (Mozambique Equator Initiative, 2002).

Large scale production of cassava was previously 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, often at sites where cassava would perform better in terms of reliable yields. However this situation is changing due to drought and unsustainably high maize subsidies. For example, efforts to promote the 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).

Cassava roots are mainly composed of starch (carbohydrate) which serves as a source for energy and some minerals and fiber. Because the root consists largely of starch, cassava is not a balanced food and the protein content is very low. However the protein content of cassava leaves is higher than that found in cassava root (Cassava Strategy Document Mozambique, 2005).

## *1.3 Cassava Mosaic Disease*

Cassava mosaic disease was reported for the first time in Tanzania in 1894. CMD has since been observed throughout all cassava growing areas of Africa and its adjacent islands (Fauquet and Fargette, 1990). The first work on CMD confirmed that the causal agent is transmitted by whiteflies (Storey and Nichols, 1938) and it was later shown to be a geminivirus (Bock *et al*. 1978).

Originally one geminivirus was regarded as the causal agent of CMD, but subsequent studies have shown that several similar, but distinct whitefly-transmitted geminiviruses, cause CMD in Africa and they can occur singly or in combinations (Hong *et al.*,1993; Zhou *et al.* 1997; Rey and Thompson, 1998; Berry and Rey 2001; Fauquet *et al*., 2008). The most important of these virus species are *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV), which belong to the genus *Begomovirus* of the family *Geminiviridae* (Fauquet and Stanley, 2003).

From 1988 to the present, a major pandemic of an unusually severe form of CMD has been spreading throughout East and Central Africa, causing massive losses, which is considered as a distinct strain of EACMV, previously called the Uganda variant (Deng *et al*.,1997), but now named *East African cassava mosaic virus-Uganda* (Fauquet *et al*., 2008). This

strain has serious consequences for cassava production in Africa. In addition to ACMV and EACMV-Uganda, several other cassava begomovirus species have been identified in sub-Saharan Africa (Harrison *et al,*1977., Thresh *et al*., 1998a; Berrie and Rey, 2001; Fondong *et al.,* 2000).

### 1.3.1 Transmission of CMD

Geminiviruses are not transmitted through cassava seeds (Storey and Nichols, 1938), but some species, such as ACMV, can be transmitted mechanically to many *Nicotiana* species (Bock *et al*. 1978; Rossel and Thottapilly, 1984 and 1985). The only insect known to transmit ACMV and other cassava geminiviruses is *Bemisia tabaci* Genn. (Hemiptera: Aleyrodidae). *B.tabaci* is thought to be of Indian/Middle East origin and was probably introduced into Africa, while *B. afer*, which also occurs on cassava, is suspected to be of African origin (Fishpool and Burban, 1994).

The virus is transmitted in a persistent manner and is retained when the vector moults. It is not transmitted congenitally to the progeny of the vector (Dubern, 1979). Virus dissemination between fields and over long distances is primarily through the use of infected stem cuttings as planting material (Fauquet *et al*.1988), and secondarily by the whitefly vector, *B. tabaci*. The viruses can also be transmitted by grafting.

### 1.3.2 Symptoms CMD

The symptoms of CMD are expressed either as a green mosaic or a yellow mosaic (Figure 1.1 below). Leaves affected by green mosaic have contrasting sectors of normal and light green tissue. Leaves affected by yellow mosaic are more obvious, with areas of normal green and yellow tissues. Severe chlorosis is often associated with premature leaf abscission and consequently decreases in growth and yield (Hillocks, 2002).

**Figure 1.1: Cassava Mosaic Disease Symptoms**

Cassava varieties differ in the type, and severity of the symptoms caused by CMD (Gibson *et al*., 1996). Varieties with some degree of tolerance or resistance show mild to moderate symptoms during the late stages of crop growth, when recovery occurs. The severity of symptoms is influenced by environmental factors (Gibson *et a*). Symptoms of CMD are sometimes confused with injury due to pests or nutrient deficiency. Normally the cassava green mite (*Mononychellus tanajoa*) and zinc deficiency cause particular problems in diagnosis. Single ACMV infections typically give rise to mild symptoms, while EACMV give moderate to severe symptoms. Synergism between CMG's has been reported and it influences disease severity, (Pita *et al*., 2001). Generally mixed infections produce severe symptoms and consequently increased yield losses (Legg, 1999).

Crop losses data have been obtained for CMD in many African countries. The crop losses are influenced by virus strains present, sensitivity of the varieties grown and environmental conditions (Thresh and Fargette, 2003). Yield losses with individual cassava cultivars have been reported from different countries to range from 20 to 95% (Thresh and Fargette, 2003).

The virus can infect virtually all cassava cultivars grown in Africa and the continent losses are estimated to be 12-23 million tones of fresh tuberous roots per year, worth about US $1200-2300 million. The mode and time of infection affect the magnitude of yield reduction, and plants grown from infected cuttings develop greater yield losses than those infected subsequently by vectors (Hillocks, 2002).

### 1.3.3 Epidemiology of CMD

The first epidemiological evidence on CMD was provided by Storey and Nichols (1938). The dissemination of the disease through infected cassava cuttings, together with transmission by whiteflies, determines its incidence. The secondary spread is made by viruliferous whiteflies moving between or within cassava plantings. Studies in Côte d'Ivoire, Kenya and Uganda, showed that the spread of CMD into and within experimental plantings is related to the number of adult whiteflies recorded, combined with the overall incidence of infection in the area (Legg *et al*., 1998; Otim-Nape, *et al*., 1993).
The amount of whitefly populations varies with agro-systems; normally when there is a low intensity of whiteflies, the incidence of CMD is expected to be low in the early stages of the growth. The disease incidence is also related to variations in climatic factors, such as temperature, rainfall and prevailing wind. The incidence of CMD tends to be high at the margins of plantings, especially the ones facing the direction of the prevailing wind (Fauquet and Fargette, 1990).

Researchers (Fauquet and Fargette, 1990; Fishpool and Burban, 1994) have found that the adult whiteflies are found more frequently on the top five leaves of the shoot, where feeding and egg laying occur. Normally the populations of whiteflies increase 4 to 5 months after planting while the virus concentration decreases with the plant age (Fargette *et al*., 1988, 1993).

The interactions between the cassava genotype, climate and the development of CMD have significant effects on populations of *B. tabaci* (Abdullahi *et al*., 2003).Previous studies have found, positive correlation between the incidence and severity of CMD, but no correlation of severity and whitefly population density. High temperatures of 25° to 30°C can favor an increase in CMD, because of an increase in whitefly population which transmits the virus.

The density of whitefly populations is also influenced by rain. It was observed that heavy rains caused a decline in adult populations of whitefly (Fauquet and Fargette, 1990). The type of cultivar (resistant or susceptible) can also dictate virus and vector spread, where spread is more rapid in the susceptible than the resistant genotypes (Legg and Thresh, 2003).

Many studies on the epidemiology of cassava mosaic disease have been done in the past 15 years in various countries of Africa such as, Tanzania, Malawi, Kenya, South Africa, Uganda, Congo, Madagascar, Angola, and Rwanda. These studies have shown the importance of spread cassava mosaic disease, and 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 all of these studies, the disease was spread in all countries with various levels of incidence and severity. Cutting infection was the more common means of spread in ll of the previously mentioned studies. The occurrence of the different CMD species differs depending on the country. For example, in a survey of CMD and CBSD, conducted in Tanzania by Legg and Raya (1998), it was reported that CMD occurred throughout the country at low to moderate incidences (1-64%) in the fifteen different regions. The incidence was generally higher along the coastal plain than in higher altitude areas in the interior. There was a significant correlation between numbers of adults of the whitefly vector, *Bemisia tabaci*, and incidence of recently infected plants, although most infection was attributable to the use of infected cuttings.

In Rwanda, Cassava mosaic disease was investigated in a survey conducted in six cassava-producing areas in 2001. CMD was shown to occur throughout the country with a mean incidence of 30%. Cutting infection resulting from use of CMD- affected planting material was the main type of infection. The disease symptoms were generally severe, with little difference between cultivars or locations (Sseruwagi *et al*., 2001).

In a survey, in ten districts, conducted at Eastern Democratic Republic of Congo, with the objective to monitor and diagnose CMD (Obonyo, 2007), CMD was the most important constraint affecting cassava production in Congo. Incidence ranged between (0 to 95%). Moderate to severe symptoms occurred virtually throughout the sampled areas. The trend in the surveys mentioned in the previous paragraph above, regarding cutting infection being high compared with whitefly infection, was observed.

### 1.3.4 Geminiviruses

Geminiviruses have paired icosahedral capsids containing single stranded (ss) genomes of 2.5-2.9 kilobases (kb) and are either monopartite or bipartite. The name Geminivirus was derived from the Latin word Gemini, meaning twins and was proposed by Harrison *et al*.

(1977). The family *Geminiviridae* has been classified into 4 genera; *Curtovirus, Mastrevirus, Topocuvirus* and *Begomovirus*. Classification was based on the virus genome organisation, host range and, most importantly, the vector transmission (Fauquet and Stanley, 2003). Curtoviruses have a 2.5-2.9 kb monopartite genome and are transmitted by leafhoppers and generally infect dicotyledonous plants.

Mastreviruses have monopartite genomes of 2.6-2.9 kb. They are also transmitted by leafhoppers but generally infect monocotyledonous plants. Topocuviruses have monopartite genomes and are treehopper transmitted and infect monocotyledons. Begomoviruses have a monopartite or bipartite genome and are transmitted by whitefly.They have a narrow host range, infecting dicotyledonous plants. Most geminiviruses (80%) belong to the genus *Begomovirus* (Varma and Malathi, 2003).

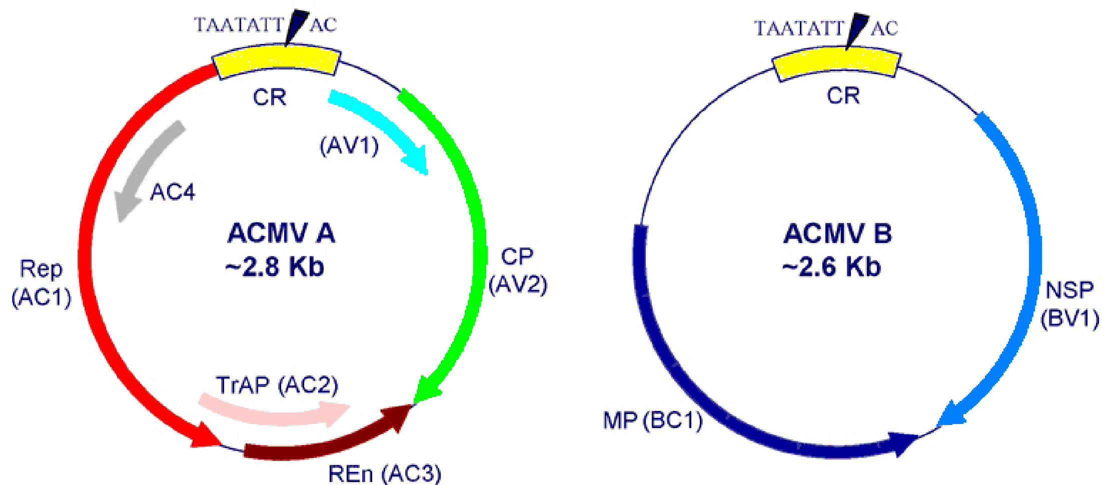


**Figure 1.2: Organization of African cassava mosaic virus (ACMV) (*Begomovirus*)**

### 1.3.5 Begomovirus Genome Organisation

Most begomoviruses are bipartite, consisting of two DNA components: DNA- A and DNA-B, each having a 2.6-2.8 kb genome. They are bidirectionally transcribed and replicate in the nuclei of infected plant cells (Dry *et al*., 1993). DNA-A replicates autonomously, while DNA-B replication is dependent on DNA-A. DNA-A encodes 2 genes on its virion-sense strand: AV1 encoding the coat protein; and AV2, responsible for virus accumulation and symptom development.

The complementary-sense strand of DNA A encodes 4 genes: AC1-AC3 genes encode the replication-protein (Rep), transcriptional activator protein (TrAP) and the replication

enhancer protein (REn), respectively. The AC4 gene is involved in virus movement, symptom severity and host range determination (Jupin *et al*., 1994).

The DNA-B component is required for inter and intra-cellular movement and encodes for 2 genes. BV1 which encodes for a nuclear shuttle protein while BC1 encodes for proteins required for cell-to-cell movement of the virus (Hanley-Bowdoin *et al*., 1999). But some begomoviruses with bipartite genomes are able to cause systemic virus infection in the host plant with only DNA-A (Briddon and Markham, 1994). Previous studies have demonstrated that DNA-A can spread in the absence of DNA-B. However, virus infection and accumulation occurs at low efficiency and without the induction of host symptoms.

DNA-A and DNA-B share a common region (CR) of ~200 base pairs (bp) with a high sequence identity (90-100%) (Harrison and Robinson, 1999). The CR is located within the intergenic region (IR) which lies between the virion and the complementary-sense strands of both DNA-A and DNA-B. The IR contains sequence motifs that are necessary for gene replication and control of gene expression (Eagle *et al*., 1994). Also found in the IR, is a nonanucleotide sequence, TAATATT↓AC, which is conserved across all geminiviruses. Rolling circle replication occurs at this site ↓ which is known as the as the initiation site (Stanley,1995).

### 1.3.6 Replication

The virus replicates by a rolling-circle replication (RCR) mechanism in a plant cell. RCR is initiated within the TAATATTAC sequence at the initiation site. Required for RCR is the Rep protein which recognizes a sequence on the ssDNA and causes a break within the conserved nonanucleotide sequence of the virion sense strand. Rep also binds to plant retinoblastoma protein which then alters the plant cell from a terminally differentiated state to an actively replicating state (Stanley, 1995). This activates the host cell machinery as geminiviruses do not have genes that encode for DNA polymerase.

The ssDNA is then converted to double stranded (ds) DNA which is then used as template for bidirectional transcription of viral genes in the virion and complementary-sense strand. Damaged DNA may occur as a result of the template being used for replication and transcription simultaneously (Brewer, 1988; Hanley-Bowdoin *et al*., 1999).

However investigations undertaken by Jeske *et al*. (2001) revealed through two-dimensional gel electrophoresis and electron microscopy that replication can also occur through a recombination-dependent replication (RDR) mechanism.

### 1.3.7 Recombination and Pseudorecombination

Recombination occurs when a fragment of one nucleotide strand becomes incorporated into that of another individual strand during replication. Research shows that recombination provides these viruses with a selective advantage, allowing evolution and increased diversity. The emergence of new variants and species through recombination has been shown to occur across genera and family (Paidam *et al*., 1999). The frequency of recombination is partly due to RDR which can occur between begomoviruses and their associated satellites. RDR is a method that allows rescuing of damaged or incomplete viral DNA that may occur as a result of replication and transcription occurring simultaneously (Jeske *et al*., 2001).

Damaged or incomplete DNA can also occur from viral replication being hindered for example from digestion of viral DNA by host enzyme or shortage of nucleotides. Recombination of damaged or incomplete DNA occurs at a homologous site and results in extension and synthesis of heterogeneous ssDNA which can then converted to dsDNA and be used for replication resulting in populations of heterogeneous ssDNA which can then be converted to dsDNA and be used for replicating resulting in populations of heterogeneous linear viral DNA (Jeske *et al*., 2001).

In 1997, Zhou *et al*. reported on the occurrence of a geminivirus that had a recombinant DNA-A that was responsible for severe CMD in Uganda. This geminivirus was isolated only from severely affected plants. Analysis of this geminivirus revealed that is a novel geminivirus that is arisen from recombination in the DNA-A component of *African cassava mosaic virus* (ACMV) and *East African cassava mosaic Ugandan virus* (EACMV-UG). This recombinant geminivirus was isolated in all locations in Uganda where the severe CMD epidemic was observed, and in areas where there was no epidemic, ACMV and EACMV isolates were only detected.

The factors that contribute to recombination are mixed infections which are pivotal for recombination to occur, high virus replication levels and increase in host range of the virus vector. Following the first report of recombination by Zhou *et al*. (1997), numerous reports

of recombination have been reported. In a study by Fondong *et al*. (2000), cassava stem cuttings from Cameroon were observed to contain ACMV and EACMV mixed infections. The DNA-A and DNA-B of these viruses was then determined.

DNA-A of EACMV virus showed evidence of recombination in the BCI region but both DNA-A and DNA-B maintained the normal genome arrangement associated with bipartite begomoviruses. It was also observed that the double infections are associated with the occurrence of more severe symptoms unlike single infections which can result in incomplete systematic infections with leaves remaining symptomless (Fondong *et al*., 2000). It is observed that frequent recombination occur in the AC1 Rep gene and it is believed that recombination at this location allows geminiviruses to increase specificity. AC2-AC4 regions have also been found to be involved in recombination and its suggested that it is because these are the genes that are mostly involved with host factors (Xie and Zhou, 2003).

Genetic variation can also occur as a result of pseudorecombination which occurs you have exchange and re-assortment of DNA between DNA-A and DNA-B of two viruses (Pita *et al*., 2001). However the production of viable infectious pseudorecombinants is generally limited to strains of a particular virus and this is due to the highly specific interaction observed between the Rep protein and the origin of replication. Another limitation is that the pseudorecombinants may possess limited host range as compared to their parental viruses (Unseld, 2000).

The first case of pseudorecombination occurring between geminiviruses that infect different hosts was observed by Gilbertson *et al*.,(1993). Pseudorecombination was observed between *Tomato mottle virus* (ToMoV) and *Bean dwarf mosaic virus* (BDMV). However, the pseudorecombinant clones produced from genetic exchange of the viruses bipartite components resulted in symptoms that were less severe than those produced by parental viruses. It is also believed that some monopartite begomoviruses may have permanently acquired DNA-B components as a result of certain field conditions allowing then to change to bipartite viruses (Seal, 2006).

### 1.3.8 Satellite DNAs

Satellites and satellite RNAs were previously isolated with RNA helper viruses, but later were identified in association with DNA viruses (Briddon *et al*., 2003). Satellites isolated

from geminiviruses depend on co-infection with a helper virus for replication, movement and encapsidation but they do not share any sequence similarity to their helper virus making them easily distinguishable from their helper virus (Dry *et al*., 1993). Satellites can also exacerbate viral symptoms caused by helper virus or produce different symptoms in conjunction with those produced by the helper virus (Dry *et al*., 1993).

Dry *et al.* (1993) reported the first satellite to be associated with a DNA virus, Tomato leaf curl virus (TLCV). The satellite was found to be a small and circular subviral DNA of 682 nucleotides (nt) but showed no evidence of being derived from the helper virus or any other geminiviruses. It did not contain an IR region similar to the one found in the CR of all geminiviruses but it did however contain between nt 10 and 20 the nonanucleotide sequence TAATATT↓AC which is conserved in all geminiviruses. This indicated that isolated DNA fragment was indeed a geminivirus associated molecule.

The subviral DNA was found to be present in both ss and dsDNA form but was transciptionally inactive and therefore was dependent on the helper virus for its replication, spreading and encapsidation. The satellite was found to have no effect on TLCV replication or symptom production in infected plants.

Recently 2 novel subviral DNA's referred to as satDNA-II and satDNA-III were isolated in Tanzania. They were isolated from cassava infected with bipartite begomoviruses (EACMCV and EACMV-UG). These satellites were slightly smaller (1-1.2kb) than the known characterised satellites. However, there is no knowledge about the satellite's replication and gene expression strategies (Ndunguru, 2005). Infectivity studies of the recently characterised satellites demonstrated that co-inoculation of the satellite with ACMV, EACMV-UG and *East African cassava mosaic Cameroon virus* (EACMCV) results in intensified symptoms observed in cassava. Co-infection also resulted in symptoms being present in resistant cassava cultivars (TME3).

### 1.3.9 Defective interfering (DI) molecules

These are small subgenomic DNA (549-1555 nt) referred to as DI-DNA and are found associated with a few full length geminiviruses. They contain a CR region similar to that of their host virus and are derived from DNA-A or DNA-B and contain an IR region similar to the one that is conserved across all geminiviruses (Patil *et al*., 2007). DI- DNAs lack a complete set of genes required to complete an infectious cycle and therefore also require

their helper virus to complete replication (Patil and Dasgupta, 2006). They interfere with helper virus replication and production and delay modulation of symptoms in infected plants and hence referred to as DI molecules (Stanley *et al.*, 1983).

DI-DNAs are about half the size of a full length geminivirus genome but the method by which DI-DNA molecules form is still unclear. However it is thought that they could have occurred as result of viral sequence rearrangement, deletion, or duplication and insertion in viral genomes as result of the DNA polymerase jumping from ssDNA to dsDNA or during ssDNA synthesis during rolling circle replication. The deletions are thought to arise during replication but the exact mechanism is unknown. However there are repeat sequences present at the deletion points indicating that recombination events might also be involved.

The first DI-DNA was characterized by Stanley and Townsend (1986) and it was associated with ACMV and was called "cassava latent virus" and found to be derived from deletions of DNA-B. MacDowell *et al.* (1986) also isolated a similar subgenomic molecule associated with Tomato golden mosaic virus (TGMV) and it was reported to be 1.2kb and also derived from DNA-B through deletion of the NSP and the C-terminal of the MP. Stanely *et al.*, 2003 reported another DI-DNA however it was derived from monopartite begomovirus Ageratum yellow vein virus (AYVV). It contained the IR segments of the 5' end of Rep, C4 and V2 ORFs and some sequences that were not of AYVV origin.

The majority of DI-DNAs are DNA-B derived (Stanley and Townsend, 1985), however recently a DI-DNA derived from DNA-A of East African cassava mosaic virus (EACMV) was isolated by Ndunguru *et al.* (2006). More recently, a sub-genomic DNA derived from South African cassava mosaic virus DNA-B was also isolated (unpublished). Patil *et al.* (2007) isolated for the first time two DI-DNAs derived from recombination of DNA-A and DNA-B. It was suggested that the DI-DNAs occurred from RDR as this replication method results in the occurrence of a lot of intracellular and intermolecular recombination.

### 1.3.10 Cassava Geminivirus Diversity and Nomenclature

Firstly, CMD was assumed to be caused by a single whitefly-borne geminivirus, although later on four types of isolates were recognized: West Africa, Kenya coastal, India and Angola (Bock and Harrison, 1985). These different isolates were grouped into tree clusters: group A included those from Angola, Côte d Ivoire, Nigeria and western Kenya; group B from coastal Kenya, Madagascar and Malawi; group C from India and Sri Lanka (Harrison

*et al*., 1991). Hong *et al*. (1993) analyzed nucleotide sequences from these strains and regarded the virus isolates from different geographic origins as distinct geminiviruses.

Group A and B were renamed *African cassava mosaic virus* (ACMV) and *East African cassava mosaic virus* (EACMV), respectively, and the Indian isolate was named *Indian cassava mosaic virus* (ICMV). Swanson and Harrison (1994) confirmed the distinct nature of the virus using monoclonal antibodies, Swanson and demonstrated that ACMV occurred in West and Central African countries up to the west of the Great Rift Valley and also in South Africa; EACMV in countries east of the Rift Valley including coastal Kenya, coastal Tanzania, Malawi, Zimbabwe and Madagascar; and ICMV in India and Sri Lanka.

In later studies, improved diagnostic techniques resulted in identification of other CGVs. In Uganda, a new virus variant was detected and was named (EACMV-UG) (Deng *et al*.1997) or the Uganda variant cassava mosaic virus (UgV) (Zhou *et al*.1997): in South Africa, *South African cassava mosaic virus* (SACMV), (Berrie and Rey 2001). Later studies in Africa showed that EACMV and ACMV occur over a much wider area. Both ACMV and EACMV occur in single and mixed infections in cassava causing mild and severe symptoms (Legg and Fauquet, 2004).

Fauquet *et al*. (2003) revised the taxonomic criteria for species demarcation and proposed six CGV species demarcation and proposed six CGV species in Africa, of which ACMV, EACMV and *South African cassava mosaic virus* (SACMV) are well characterized, while *East African cassava mosaic Cameroon virus* (EACMVV), *East African cassava mosaic Malawi virus* (EACMMV), and *East African cassava mosaic Zanzibar Virus* (EACMZV) were all regarded as distinct from typical EACMV. The recombinant virus referred to as a Uganda variant (EACMV-UG2) was considered a distinctive strain of EACMV.

More recently, Fauquet *et al*., (2008) proposed seven Africa CGV species, namely *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), *East African cassava mosaic virus Cameroon* (EACMCV), *East African cassava mosaic Kenya virus* (EACMKV), *East African cassava mosaic Malawi virus* (EACMMV), *East African cassava mosaic Zanzibar* virus (EACMZV) and *South African mosaic virus* (SACMV), and two Indian, which are *Indian cassava mosaic virus* (ICMV), and *Sri Lankan cassava mosaic virus* (SLCMV), identified from Indian subcontinent.

Demarcation of species is based on a nucleotide sequence threshold of 89%. In addition to these species, several strains/isolates have been recognized, including: East African cassava mosaic virus-Uganda (EACMV-UG), -Kenya (EACMV-KE), -Tanzania (EACMV-TZ); *South African cassava mosaic virus-South Africa* (SACMV-ZA), -Madagascar (SACMV-MG); *East African cassava mosaic Cameroon virus-Cameroon* (EACMCV-CM),- Tanzania (EACMV-TZ); *Indian cassava mosaic virus-India* (ICMV-IN), Kerala (ICMV-Ker); Sri Lankan cassava mosaic virus-India (SLCMV-IN), -Sri Lanka (SLCMV-LK). Virus isolates are grouped into strains (85-93% nt sequence identity), while variants of strains or species are demarcated at 94-100% (Fauquet *et al*, 2008).

## *1.4 Identification and Diagnostic Methods*

Cassava plants respond to attack by various viruses in a limited number of ways, which can make viral diagnoses difficult. ACMV, EACMV, SACMV, ICMV and others cause similar leaf mosaic symptoms in cassava and they cannot be distinguished by symptoms

The serological techniques of DAS-Elisa (Sequeira and Harrison 1982) and TAS-Elisa (Thomas *et al*., 1986) have been used successfully to detect and distinguish between CMGs, and also panels of monoclonal antibodies were subsequently developed to facilitate the discrimination between ACMV and EACMV (Swanson and Harrison, 1994). This technique was used to produce the first CMGs distribution map for Africa (Swanson and Harrison 1994). To detect the recombinant CMGs, DNA-based diagnostic techniques are now widely used. Much of this diagnostic and detection work makes use of specific PCR primers designed from full length sequences of DNA-A (Zhou *et al*., 1997).

An alternative molecular approach for CMGs diagnostic and variability studies is PCR with RLFP analysis. In this method, universal and abutting CMG DNA-A primers (Biddon *et al*., 1993) are used to amplify near full-length DNA-A fragments from whole plant DNA. Those products are then digested with restriction enzymes (commonly *ECOR*V and *Mlu*I), and the digests run on an agarose gel.

The development and use of rolling circle amplification (RCA) along with PCR and RLFP has helped to provide a solution for molecular characterization of geminiviruses. This method of geminiviruses characterization has recently been developed and studied for geminiviruses such as *Tobacco yellow dwarf virus* and *Abutilon mosaic virus* and is however relatively novel for CMGs (Haible *et al*., 2006).

RCA is a method that mimics the replication mechanism RCR (Rector *et al*., 2004). This uses oligonucleotide primers which anneal to the circular DNA template and result in exponential amplification of DNA in 4-8 hours producing single -stranded complementary concatemers. The hexamers then bind to the ss RCA products and elongation occurs at the new recognition sites resulting in the release of ds RCA concameters (TempliPhI Kit manual) and all this requires no thermal cycling and can be performed on the bench at a wide range of temperatures (Rector *et al*., 2004).

The RCA has the advantage to be used and implemented in resource-poor laboratories as this method is cheaper than PCR and antibody detection (Haible *et al*., 2006). This method is useful when you do not know the sequence of the template DNA, as it uses random primers. It generates approximately 800 copies per hour and with a low error rate due to the enzyme having high proof reading ability.

The enzyme is also stable and does not require thermal cycling (Rector *et al*., 2004). RCA products can be used directly for RLFP for virus identification, even up to species level, without any sequencing information. They can be sent off for sequencing directly without the requirement of any PCR and cloning steps making it very useful and attractive especially in laboratories with limited resources (Haible *et al*., 2006).

## *1.5 The Whitefly Vector (Bemisia tabaci)*

Whiteflies belong to the order: Homoptera, suborder: Sternorrhyncha, Superfamily: A*leyrodidae* (Thompson, 2002). Over 1200 whitefly species are known worldwide, although only a limited number have been closely studied on key herbaceous hosts (Mound and Halsey, 1978; Byene *et al*., 1990a). The species *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) is the most important whitefly on cultivated agricultural crops in tropical and subtropical areas around the world, and is the most widely distributed and economically important species (Brown *et al*., 1995). It is believed to have originated in South East Asia/Indian sub-continent (Gill, 1990; Mound and Halsey, 1978 or possibly in northern Africa (Gill, 1990).

This whitefly gained increased importance during the past thirty years as a pest and vector of plant viruses, particularly whitefly-transmitted geminiviruses in the genus: *Begomovirus* (family: *Geminiviridae*) in tropical and subtropical regions (Muniyappa, 1980; Duffus, 1987; Poulston and Anderson, 1997).

The insect causes significant damage to crops primarily through phloem feeding, phytotoxic disorders and 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 *et al.,* 1983; Bedford *et al* 1994).

*Bemisia tabaci* also causes damage to plants through direct feeding, such as chlorosis of leaves (Maynard and Cantliff, 1989). This insect is very small in size less than 2mm long, the nymphs are sessile 'scales" and the adults are white winged and they can fly up to five to ten kilometers. The adults normally are morphologically indistinguishable. They feed under the surfaces of young apical leaves, where they lay eggs and the immature stages develop. The eggs are oval and elongate, and are attached to the leaf surface by a narrow stalk or pedicel (Avidov, 1956). *Bemisia tabaci* undergoes incomplete metamorphosis and therefore its development is divided into four nymphal instars (Lopez-Avila, 1986). The adults differ in size, and females are larger (~ 1mm) than males (~8 mm) (Gill, 1990). The males are thinner and have more tapered abdomens with a pair of claspers at the tip (Fishpool and Burban, 1994). An average of twelve generations is attained annually under field conditions (Fishpool and Burban, 1994).

Generations ranges from 18.6 days on sweetpotato to 21 days (dry season) and 28 days (rainy season) on cassava in Ivory Coast (Fishpool *et al*., 1995) and in Uganda (Legg, 2001). Temperature is a key determining factor, and higher temperatures (30-33°C) result in faster development times (Gerling *et al*., 1986).

### 1.5.1 Bemisia tabaci Host Range

*Bemisia tabaci* is a polyphagus species (Greathead, 1986), and some biotypes and genotypes are extremely polyphagus (Brown *et al*., 1995). It colonizes mainly annual, herbaceous plants including over 500 species from 74 families (Mound and Halsey, 1978).

Adult whiteflies occur on cassava throughout the crops growing period (Fishpool and Burban, 1994), although their population differ with the stage of the plant growth (Fishpool *et al*., 1995; Otim-Nape *et al*., 2001). The adults invade and establish within the

crop as soon as the plants are sufficiently grown. Small numbers of adults may occur 3 weeks after the initial colonization, which is followed by a rapid increase of population 3 to 4 months after planting. (Fishpool and Burban, 1994).

The population dynamics and activity of *B.tabaci* are believed to depend on changes in the nutritional quality of the host-plant, natural enemies, and climatic factors such as temperature, rain, wind and relative humidity (Legg, 1995). The adults feed by inserting their stylets between host-plant cells and penetrating the phloem of the host plant (Pollard, 1955; Janssen *et al*., 1989). There is usually a greater whitefly population during the first 3 months than later when the plants are more mature (Silvestre and Arraudeau, 1983).

According to Fargette *et al*., (1993) and Otim Nape (1993), increases in whitefly populations are favored by high temperatures and radiation and low rainfall and relative humidity. The adults disperse mainly by the aid of wind and can move short or long distances, but are also disseminated by humans who transport the immature and adult stages on plant material (Joyce, 1981; Mound, 1983; Byrne and Bellows, 1991). Some of the cropping practices such as planting date, crop disposition and intercropping influence whitefly population dynamics and hence the spread of whitefly- transmitted geminiviruses.

### 1.5.2 Whitefly Biotypes

Outbreaks of *B tabaci*, particularly in areas where it was previously unimportant, are linked to the appearance of new biotypes (Simon *et al*., 1994). A biotype can be a group of organisms having the same or nearly the same genotype, such as a particular strain of an insect species. The taxonomic status of *B.tabaci* has been subject to much debate since the emergence of different geographical biotypes, in particular biotype B, and the general consensus is that *B. tabaci* is a species complex (Brown *et al*., 1995).

The B biotype increased in distribution in the Southwestern US and ultimately was shown to have displaced the local A biotype (Costa *et al*., 1993). Most recently outbreaks of the B biotype have occurred in many countries including Australia, China and elsewhere in South East Asia (De Barro *et al*., 2000; Coombs *et al*., 2003). The evolution of agriculture leading to irrigated monocultures, the requirement for intensive agriculture enabling two cropping seasons per year and the use of fertilizers and pesticides are key factors in biotype emergence (Brown *et al*., 1995a). The growing economic importance of *B. tabaci*

generated interest in the development of techniques for systematic and evolutionary studies (Nei, 1987; Doolitle, 1990).

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 designated A-S (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 RLFP□ s.

The application of the sequences of the mitocondrial cytochrome oxydase I (mt COI) and the ribosomal internal transcribed spacer (ITS1) genes provided 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; Abdullahi *et al*., 2003; De Barro *et al*., 2000).

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 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).

# CHAPTER TWO: EPIDEMIOLOGY OF CASSAVA MOSAIC DISEASE AND THE WHITEFLY VECTOR *BEMISIA TABACI* (GENNADIUS)

## *2.1 Abstract*

Cassava mosaic disease contributes to devastating cassava crop losses in Africa, India and Sri Lanka (Thresh *et al*, 1998), and is caused by several species of begomoviruses. The first survey of CMD in Mozambique was carried out between 1999 and 2001 in the three Northern provinces (Zambezia, Cabo Delgado and Nampula), where incidence and severity was recorded. The aims of this study were to investigate the incidence and severity of CMD in 6 provinces in the north and south of Mozambique, and to evaluate the contribution of the whitefly vector, *B. tabaci,* to the disease epidemiology. Characterization of the whitefly populations using molecular markers using (the mitochondrial CO1 gene) was also undertaken.

Two surveys were conducted in the 2005/2006 and 2006/2007 growing seasons in six provinces, namely, Maputo, Gaza, Inhambane, Nampula, Zambézia and Cabo Delgado. From each province 5 districts were selected and 20 fields per district, with a distance from fields of ± 5km, were targeted. In each field, 30 plants were assessed along a 'Z' configuration. CMD incidence, severity, and source of infection (whitefly or from infected cuttings) were assessed and average numbers of whitefly per district in each province enumerated. Results were analyzed using SPSS 11.5 for Windows, and One Way ANOVA was performed.

The study showed that CMD incidence was highly variable within districts and between provinces, in both surveys. In the 2005 survey, CMD incidence ranged from 22% (lowest) in Cabo Delgado to 76% (highest) in Gaza Province, while in 2006 the average mean incidence was 6% in Nampula and 75% in Gaza. In both surveys there were significant differences ($P \leq 0.005$) in incidence between some of the provinces. The mean CMD severity score ranged from (2 to 3.3) in all provinces in the two surveys, which indicated mild to severe symptoms in Mozambique in the cassava-growing regions. A score of 4-5 usually indicates very severe symptoms typical of mixed infections or severe species or isolates of cassava begomoviruses.

The whitefly infection was very low (0 to 3%), compared to cutting infection (22 to 60%) in all of the six provinces in the two surveys. Zambezia had the highest mean whitefly numbers (15.8) across the province in the two surveys compared to Maputo and Inhambane, which had the lowest mean whitefly numbers (0.1) recorded. These results indicate that cutting infection is more important than the whitefly infection, because farmers plant infected cuttings repeatedly over the years. This is the main vehicle of spreading the disease.

In this study, the *B. tabaci* from the six provinces in Mozambique formed a single unified group (with little sequence divergence) within the Sub-Saharan African clade. Specifically the Mozambique haplotypes clustered with the southern African haplotypes from South Africa, Swaziland and Zambia (Berry *et al*., 2004).

## *2.2 Introduction*

In Mozambique, CMD was first reported in 1958, (Orlando Mendes, 1958) and at that time the disease was considered to be spread all over the country.

In 1999-2001, Dr Mike Thresh, from the Natural Resources Institute of UK, conducted surveys in Nampula and Zambezia to investigate the incidence of CBSD, and CMD. The incidence of CBSD was as high as 63% in Nampula, whereas the highest incidence of CMD was 51%. Incidence was even higher in Zambezia Province (75% and 83% for CMD and CBSD, respectively).

In 2000, a survey on cassava pests and diseases was conducted in 6 districts of Inhambane Province, which is situated in south of Nampula and Zambezia in Mozambique, and is one of the main cassava production provinces. The results of the survey showed that the mean CMD incidence, in the six districts, of a popular susceptible cassava variety called Munhaca was 60%, while the mean severity score was 3, on a scale of 1 to 5. Another variety (Chinhembwe) is also widely grown, and a mean CMD incidence of 55% was recorded in the six districts of Inhambane province, while the mean severity was 3. In Inhambane most of the varieties grown (Munhaca, Fernando Po, Gangasol, Chinhembwe and Precoce de Angola) were severely attacked, with a mean severity score of 4.

The concept of host races has been proposed due to occurrence of *B. tabaci* populations with no clear-cut diagnostic differences in morphology but which exhibit differences in

host range, dispersal behauvior, fecundity and competency for begomovirus transmission (Brown and Bird, 1992; Brown *et al*., 1995). Despite the usefulness of genetic markers, such as mtCOI, there is still insufficient data to raise races to a species status. De Barro *et al*., (2005) provided evidence to disqualify the separation of *B. tabaci* and *B. argentifolii* and suggested that the later be considered within the *B. tabaci* species complex. However these markers gave evidence of six well-supported phylogenetic clades or races and an unresolved core of ungrouped genotypes with clear geographic distribution at a continental level that fall within the *B. tabaci* species complex (De Barro *et al*., 2005).

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). Brown (2005) defined 5 major clades for selected haplotypes and biotypes worldwide, based on mtCOI sequences (780bp): Eastern and southern Africa; West Africa; Far East, Near East and Pacific; America and Tropics; and Northern African/Middle Eastern and Mediterranean. More recently, with the exception of recently introduced biotypes, members of the *B. tabaci* species now have been grouped into seven major geographic clades based on mtCOI sequence analysis: (I and II) two sub-Saharan Africa (SSAF I and II) clades that contains a large number of divergent haplotypes, (Berry *et al.*, 2004; Brown 2000; Legg *et al.,* 2002; Sseruwagi *et al.,* 2005, 2006) (III) Mediterranean-North Africa-Middle Eastern clade (MED-NA-ME) that contains the well known B, Q, and MS biotypes and closest relatives; (IV and V) Southeast Asian-Australian clade (SA-AUS) comprising of two sister clades namely Asia I and Australia-Pacific and Asia II found only in China; (VI) two clades within the American Tropics namely north and central Caribbean and; (VII) the south American clade of the American Tropics (Brown, 2010). As may be expected, different kinds of analyses may reveal different demarcations for sub- and sisterclades, although this could also be due to the alignment of sets of sequences employed and use of different outgroups (Boykin *et al*., 2007; De Barro *et al.,* 2005; De la Rua *et al.,* 2006, Dinsdale *et al*., 2010). However, this nomenclature system that group haplotypes within their respective phylogeographic clade still facilitate rapid identification and reveal the origin or possible introduced status of previously unstudied populations.

The phylogeny given above revealed geographic groups that diverged at about 2-26% in terms of nucleotide sequence. Boykin *et al*. (2007) attempted to improve the understanding of whitefly phylogeny and used a Bayesian phylogenetic technique to analyse 366 whitefly mtCOI DNA sequences. They show 12 major well-resolved groups/clades: Mediteranean/Asia Minor/Africa; Mediterranean; Indian Ocean; sub-Saharan Africa silverleafing; Asia I; Australia; China; Asia II; Italy; New World; sub-Saharan Africa non-silverleafing; and Uganda sweetpotato. These studies can provide the basis for analysis of whitefly genetic diversity.

Despite the fact that CMD is spread all over the country, little attention has been given to this disease. No comprehensive and systematic studies have been done in any of the provinces and no studies have been performed since 1999-2001. There is limited information in the distribution of the disease; what species/strains of begomovirus occur; the effects of the disease on cassava growth or yield; or whitefly numbers, biotype and distribution. These are serious limitations that needed to be addressed and in this chapter we present results of a survey in six major cassava-growing provinces in 2005 and 2006, where we measured incidence and severity of CMD, and recorded the type of transmission by either whitefly or cutting infection. Whiteflies were also genetically characterized using the mitochondrial CO1 marker.

## *2.3 Methodology*

### 2.3.1 Geographical location of surveys

Surveys were conducted in 6 provinces of Mozambique, namely Nampula, Zambezia, Cabo Delgado (northen Mozambique), and Maputo, Gaza and Inhambane (southern Mozambique), in June 2005 and May 2006. Figure 1 shows the cassava growing regions in Mozambique. In each province five districts were surveyed and in each district 20 fields were randomly selected within proximity of 5 kilometers. Thirty plants were assessed per field along a Z configuration. A table of provinces and districts is shown below in Table 2.1:

**Table 2.1: Provinces and Districts where the surveys were conducted**

| | Provinces | | | | | |
|---|---|---|---|---|---|---|
| | Maputo | Gaza | Inhambane | Zambezia | Nampula | Cabo Delgado |
| **Districts** | Boane | Guijá | Zavala | Namacurra | Murrupula | Chiúre |
| | Matutuine | Manjacaze | Inharrime | Nicoadala | Mongovolas | Macomia |
| | Marracuene | Chibuto | Homoine | Mocuba | Nacala | Mpuez |
| | Moamba | Xai-Xai | Morrumbene | Maganja | Meconta | Mocimboa |
| | Manhica | Macia | Massinga | Ile | Nacaroa | Palma |

## *2.4 Data collection*

In each field, parameters such as disease incidence, disease severity, type of infection (whitefly vs cutting), and number of adult whiteflies were assessed.

**Disease incidence**: number of diseased plants relative to the total number of plants assessed (Incidence = number of plants with symptoms/total number of plants assessed). The mean incidence was calculated per field, per district and then per province.

**Diseased 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 (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 takes into account the degree of symptom development in diseased plants.

**Type of infection**: For each infected plant assessed, it was indicated whether transmission was by cutting or whitefly infection. For cutting infection (CINC) symptoms appear on lowest first-formed leaves, while for whitefly infection (WINC) symptoms appear only on upper-most leaves.

**Adult whitefly population:** The whitefly population was assessed on each sample plant by counting the average number of adult whitefly on the five top leaves of the plant. The mean numbers of whitefly per plant were calculated. Following this calculation, the mean number was calculated for all the fields per district and then averaged for the province.

**Detection and characterization of cassava mosaic begomoviruses**:

The uppermost leaves with CMD symptoms were collected and kept in a plastic bag on ice for DNA extraction and molecular characterization in the laboratory at the University of the Witwatersrand, Johannesburg, SA. These samples are discussed in Chapter 3.

### 2.4.1 Whitefly collection

*Bemisia tabaci* adults and nymphs (if present) were collected of the ventral surfaces of cassava leaves, in all the sites surveyed. Both adult whiteflies and nymphs were stored in 70% ethanol at -20°C until analysis. DNA extraction and mtCO1 amplification for whitefly identification was carried out by Prof. J Brown, University of Arizona, United States of America. A list of the whitefly codes and names of locations are in Table 2.2.

**Table 2.2: Whitefly sample collections from districts in the six provinces chosen for mtCO1 sequencing**

| Whitefly code | District | Province |
|---|---|---|
| M107 | Ile district | Zambezia |
| M111 | Ile district | Zambezia |
| M119 | Macomia | Cabo Delgado |
| M123 | Palma | Cabo Delgado |
| M134 | Montepuez | Cabo Delgado |
| M142 | Namacurra | Zambezia |
| M155 | Murrupula | Nampula |
| M16 | Zavala | Inhambane |
| M166 | Guija | Gaza |
| M169 | Boane | Maputo |
| M17 | Mogovolas | Nampula |
| M25 | Mocuba | Zambezia |
| M33 | Mocimboa | Cabo Delgado |
| M36 | Nacala | Nampula |
| M41 | Inharrime | Inhambane |
| M45 | Manjacaze | Gaza |
| M5 | Matutuine | Maputo |
| M53 | Mueda | Cabo Delgado |
| M56 | Mogovolas | Nampula |
| M61 | Maganja | Zambezia |
| M64 | Morrumbene | Inhambane |
| M74 | Xai-Xai | Gaza |
| M85 | Moamba | Maputo |
| M89 | Homoine | Inhambane |
| M9 | Marracuene | Maputo |
| M94 | Chibuto | Gaza |
| M95 | Inharrime | Inhambane |

### 2.4.2 Statistical Analyses

Comparisons were first made between mean incidences, whitefly numbers and severity scores for different districts within each province using a One-way ANOVA (data not shown). One-way ANOVA was performed in order to test for overall significant differences in combined disease parameters (mean total disease incidence, severity, cutting incidence, whitefly incidence and whitefly numbers) between the provinces. Multiple comparisons using the post-ANOVA LSD test were carried out to establish if significant differences existed in each of the disease parameters measured between the individual provinces. The analysis was done using the program SPSS version 11.3. ANOVA results are presented in Tables A1 to A8 (appendix 1).

### 2.4.3 Molecular Characterization of Whitefly Samples

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 cytochorme oxidase I gene (mtCOI) fragment were C1-J-2195 (TTGATTTTTTGGTCATCCAGAA GT) and L2-N-3014 (TCCAATGCACTAATCTGCCATATTA) selected from 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 Laboraty for molecular Systematic and Evolution, University of Arizona, Tucson, AZ 85721, USA.

### 2.4.4 Phylogenetic analysis

Sequences were aligned using the Clustal algorithm (ClustalW 1.7) (MegAlign, DNA STAR, Madison, Wiscosin, USA), and aligned sequences were evaluated for genetic relatedness by Parsimony using PAUP (Swofford *et al*., 2002). Bootstrapping was performed with PAUP using the heuristic option for 1000 replications. Whitefly mtCO1 sequences for comparison were obtained from Genbank or unpublished data (J.Brown, University of Arizona) (Refer to Appendix 4). Sequences of the whiteflies *Trialeuroides vaporariorum* (Westwood) and (Inhambane) (Preisener and Hosney) were used as the soutgroups.

**Figure 2.1: Main cassava growing areas where the survey was conducted**

## *2.5 Results*

### 2.5.1 CMD Incidence

#### 2.5.1.1 Survey in 2005

Results for total incidence (with separate cutting and whitefly-borne components), recorded in 2005 are presented below in Table 2.3 and Figure 2.2. Mean incidence figures used for statistical analyses (0-1) were converted to percentages in table 2.3 (and table 2.4) for ease of discussion since incidence is usually presented as percentage of diseased vs total number of plants sampled. Cutting incidence was significantly higher in all cases in the six provinces compared with whitefly-borne incidence.

Mean transmission percentages by whitefly in 3 provinces were low (1-3%), while no transmission by whitefly was recorded in Maputo, Nampula and Cabo Delgado. Gaza was the province with highest CMD incidence (76%), while Cabo Delgado had a lower CMD

incidence (22%) recorded. Maputo province had the second highest (60%) mean CMD incidence recorded amongst the 5 provinces (Table 2.3 Figure 2.2).

There were statistically significant differences ($p \leq 0.05$) in the overall total incidence (combined cutting and whitefly incidence) between provinces (Table A1 appendix).

Gaza and Maputo provinces had the highest total CMD incidence (INC) (76% and 60%, respectively) compared with the other 4 provinces (Table 2.3). Inhambane, Zambezia and Cabo Delgado had significantly lower CMD incidences (27%, 36% and 22%, respectively) compared with all provinces except with Cabo Delgado. Nampula had a similar mean CMD incidence (56%) recorded compared to Maputo (60%). Whitefly transmission (WINC) was lower than cutting transmission (CINC) in all 5 provinces, with the highest scores values in Inhambane (3) and Gaza (2).

Multple comparisons (Table A2) of CINC in each province, indicated that there were significant differences ($p \leq 0.05$) between all of the 6 provinces, except for between Maputo and Nampula, and Cabo Delgado and Inhambane. Mean whitefly transmission (WINC) was less significantly variable ($p \leq 0.05$) compared to cuttings; between some provinces here were significant differences in whitefly transmission, while in others not (see Table A2). The pattern was identical for total incidence (INC) where there were significant differences ($p \leq 0.05$) between all of the 6 provinces, except for between Maputo and Nampula, and Cabo Delgado and Inhambane.

**Table 2.3: Incidence (percentage) of cassava mosaic disease in six provinces in 2005**

| Variable | Province | Number of fields | Mean ± std. error | Mean (%) |
|---|---|---|---|---|
| CINC | Maputo | 100 | 0.60± 0.4 | 60 |
| | Gaza | 100 | 0.74±0.2 | 74 |
| | Inhambane | 100 | 0.22±0.4 | 22 |
| | Zambezia | 100 | 0.36±0.5 | 36 |
| | Nampula | 95 | 0.56±0.4 | 56 |
| | Cabo Delgado | 105 | 0.22±0.4 | 22 |
| WINC | Maputo | 100 | 0.0±0.0 | 0 |
| | Gaza | 100 | 0.2±0.1 | 2 |
| | Inhambane | 100 | 0.3±0.1 | 3 |
| | Zambezia | 100 | 0.1±0.1 | 1 |
| | Nampula | 95 | 0.0±0.0 | 0 |
| | Cabo Delgado | 105 | 0.0±0.0 | 0 |
| INC | Maputo | 100 | 0.60±0.4 | 60 |
| | Gaza | 100 | 0.76±0.2 | 76 |
| | Inhambane | 100 | 0.27±0.6 | 27 |
| | Zambezia | 100 | 0.36±0.5 | 36 |
| | Nampula | 95 | 0.56±0.4 | 56 |
| | Cabo Delgado | 105 | 0.22±0.4 | 22 |

N = number of fields; CINC- Cutting incidence, WINC-Whitefly incidence; INC-Total incidence

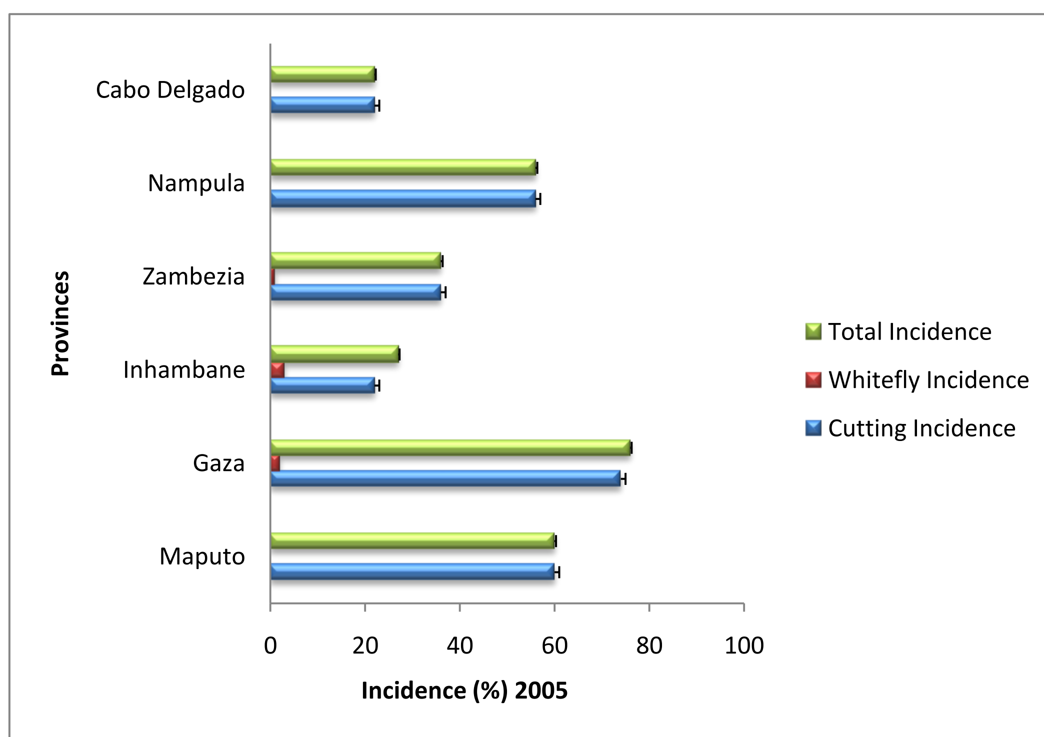


**Figure 2.2: Histogram showing the incidence of CMD in six Provinces in the 2005 Survey**

The pattern of Cassava mosaic disease incidence by cutting or whitefly, and numbers of whitefly recorded can be observed in figure 2.2. Gaza, Nampula and Maputo provinces had higher cutting (symptoms appear on the lower leaves) and total incidences compared with Zambezia, Cabo Delgado and Inhambane provinces. Transmission by whitefly was significantly lower than by cutting, and differences between provinces, as discussed earlier, could be clearly observed. Low whitefly transmission was observed in Zambezia, Gaza and Inhambane provinces, while the other 4 provinces had no to negligible whitefly transmission, as evidenced by no whitefly infection symptoms which appear only on upper-most leaves.

### 2.5.1.2 Survey in 2006

Results for total incidence (combined cutting and whitefly-borne components) recorded in 2006 are presented in Table 2.4 and Figure 2.3. Cutting incidence was significantly higher in all cases in the six provinces compared with whitefly-borne incidence, which was similar to that reported in 2005.

Mean transmission percentages of CMD by whitefly in 3 provinces was low (2-3%), as observed in the 2005 survey, while no whitefly transmission was recorded in Zambezia, Nampula and Cabo Delgado. In the 2006 survey, Gaza was again the province with the highest mean CMD incidence (76%), as was observed in 2005. However, compared to 2005 (56%), Nampula was the province with the lowest INC (7%), while in 2005 Cabo Delgado had the lowest overall mean CMD incidence (22%). Maputo had the second highest recorded mean INC (55%) compard with all provinces. Cutting transmission (CINC) showed a similar pattern in terms of mean percentages as INC.

A similar pattern as in 2005 was observed in the 2006 survey, where there were significant differences in overall mean disease parameters (INC; CINC and WINC) within and between the provinces ($p \leq 0.05$) (Table A3 appendix). Multiple comparsions of disease parameters (ANOVA Table A4) indicated that significant differences in CINC and INC between the provinces were noted, except between Inhambane, and Zambezia and Cabo Delgado; and between Zambezia and Cabo Delgado where there were no significant differences ($p \leq 0.05$). Mutliple whitefly transmission comparisons were significantly different between some provinces, but not others: Maputo and Inhambane; Inhambane and

Gaza; Zambezia and Inhmabane, and Nampula and Zambezia; and Zambezia and Cabo Delgado (Table A4).

**Table 2. 4: Incidence of CMD in six provinces in the 2006 survey**

| Variable | Province | Number of Fields | Mean ±Std. error | Mean % |
|---|---|---|---|---|
| CINC | Maputo | 100 | 0.52±0.4 | 52 |
| | Gaza | 100 | 0.73±0.2 | 73 |
| | Inhambane | 100 | 0.26±0.6 | 26 |
| | Zambezia | 100 | 0.27±0.4 | 27 |
| | Nampula | 95 | 0.7±0.1 | 7 |
| | C.Delgado | 105 | 0.31±0.4 | 31 |
| WINC | Maputo | 100 | 0.3±0.1 | 3 |
| | Gaza | 100 | 0.3±0.1 | 3 |
| | Inhambane | 100 | 0.2±0.1 | 2 |
| | Zambezia | 100 | 0.0±0.0 | 0 |
| | Nampula | 95 | 0.0±0.0 | 0 |
| | C.Delgado | 105 | 0.0±0.0 | 0 |
| INC | Maputo | 100 | 0.55±0.4 | 55 |
| | Gaza | 100 | 0.76±0.2 | 76 |
| | Inhambane | 100 | 0.28±0.6 | 28 |
| | Zambezia | 100 | 0.27±0.4 | 27 |
| | Nampula | 95 | 0.7±0.1 | 7 |
| | C.Delgado | 105 | 0.32±0.4 | 32 |

CINC-Cutting incidence, WINC- Whitefly incidence, INC- Total incidence

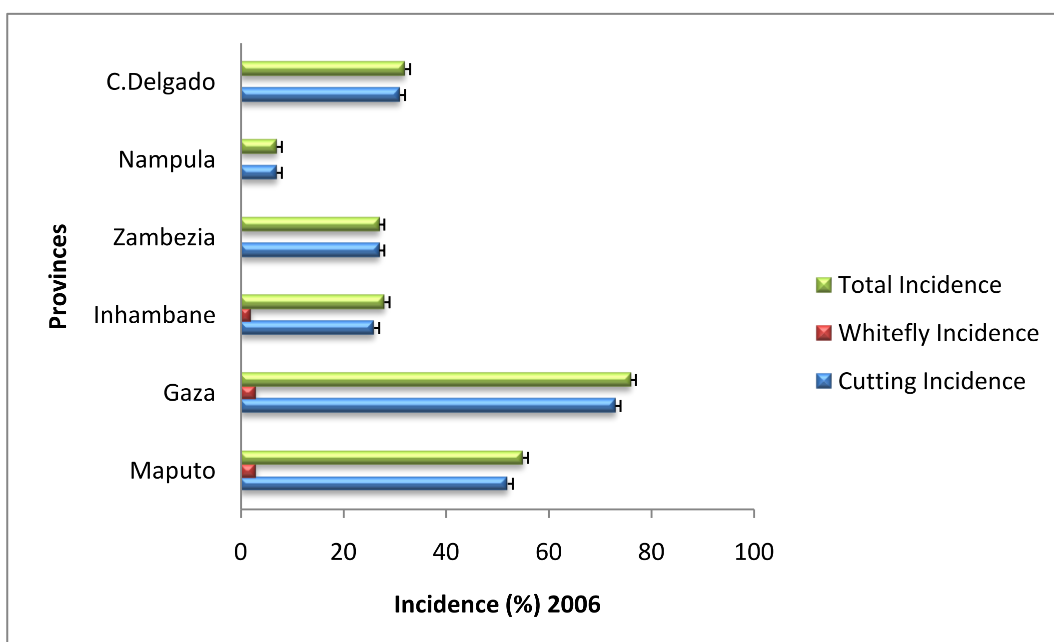


**Figure 2.3: Histogram showing the incidence of CMD in six provinces in the 2006 survey**

The histogram (Figure 2.3) clearly shows that infection by whitefly does not have a significant impact on CMD incidence and a similar pattern was observed for total incidence and incidence by vegetative cuttings as seen in the 2005 survey. Gaza clearly had the highest mean CINC and WINC compared to the other provinces, followed by Maputo and Cabo Delgado.

### 2.5.2 CMD severity and whitefly abundance

#### *2.5.2.1 Survey in 2005*

In 2005, the mean severity score (scale 0-5) was low (2.1) to moderate (3.3) in all provinces, with Nampula having the lowest symptom severity (2.1) and Maputo the highest (3.3) (Table 2.5, Figure 2.4). Although the mean scores were low to moderate, some plants did show more severe symptoms (score >3.3). There were overall significant differences within and between provinces ($p \leq 0.05$) (Table A5). Multiple comparisons of CMD severity demonstrated that these were significantly different, except between Maputo and Cabo Delgado and between Cabo Delgado and Zambezia The southern provinces Maputo and Ihhambane showed a similar severity score (3.3 and 3.2, respectively)(Table 2.4 and Table A6 appendix), and had higher severity than the northern provinces Nampula and Cabo Delgado.

The mean number of whitefly per plant in each province ranged between 1 in Maputo to 15.1 in Zambezia (Table 2.5, Figure 2.4). While Maputo and Inhmabane had the highest severity score, they showed the lowest mean whitefly numbers per plant (0.1- 1.0). The overall mean whitefly numbers per plant were significantly different within and between the six provinces (Tables A5 and A6 appendix).

**Table 2. 5:** Disease severity and whitefly abundance in the 2005 survey

| Variable | Province | N | NP | Mean score ± SE |
|---|---|---|---|---|
| **CMDSEV** | Maputo | 100 | 1775 | 3.3±04 |
| | Gaza | 100 | 3000 | 2.5±0.9 |
| | Inhambane | 100 | 829 | 3.2±0.5 |
| | Zambezia | 100 | 814 | 2.9±0.5 |
| | Nampula | 95 | 295 | 2.1±1.2 |
| | C.Delgado | 105 | 977 | 2.9±0.1 |
| | | | | **Mean no. per plant ± SE** |
| **WFNO** | Maputo | 100 | 3000 | 1.0±0.1 |
| | Gaza | 95 | 3000 | 2.5±0.7 |
| | Inhambane | 105 | 2999 | 0.1±0.0 |
| | Zambezia | 100 | 3000 | 15.1±1.3 |
| | Nampula | 100 | 2998 | 3.4±0.2 |
| | C.Delgado | 100 | 2997 | 3.9±0.4 |

**CMDSEV: Cassava mosaic disease severity; Scores (1-5); WFNO: mean numbers per plant per province; N= number of fields per province; NP = number of plants scored per province**

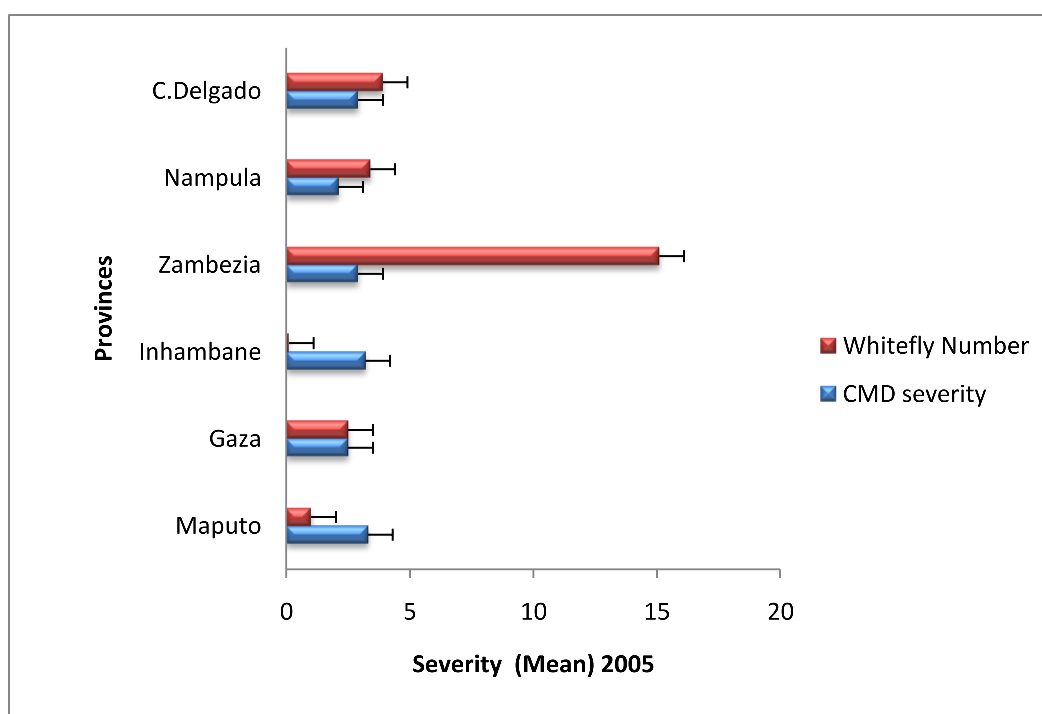


**Figure 2.4: CMD severity and mean whitefly numbers per plant in the six provinces from the 2005 survey**

### 2.5.2 Survey in 2006

In 2006, the severity score (scale 0-5) was moderate (2.7) to high (3.3) in all provinces, with Zambezia having the lower severity (2.7) and Gaza the highest (3.3) (Table 2.6). These figures were similar as observed in 2005. The southern Mozambique provinces, Maputo, Gaza and Inhambane, had higher severity than the northern provinces, Zambezia, Nampula and Cabo Delgado, which was a similar trend observed in 2005. The mean severity score ranges (2.7 to 3.3) for all the provinces were the same in the two growing seasons when the surveys were done.

There were overall significant differences in CMD severity within and between provinces ($p \leq 0.05$) (Table A7 appendix). Multiple comparisons showed that there were significant differences in CMD severity between provinces ($p \leq 0.05$), except between Maputo and Inhambane and between Cabo Delgado and Nampula, which showed similar mean severity scores (Table A8 appendix 1).

The mean number of whitefly per plant in each province ranged between 0.1 in Maputo to 15.8 in Zambezia. No whitefly in Maputo were observed (Figure 2.5). The mean whitefly numbers per plant were significantly different within and between between the six provinces (Tables A7 and A8 appendix 1). The mean number of whitefly recorded per plant in the three northern provinces (Cabo Delgado, Nampula and Zambezia) was higher (5.8 – 7.8) than the southern provinces in the 2006 survey (Figure 2.5), with a similar pattern as 2005.

As in 2005, Zambezia had the mean highest number of whitefly per plant (15.8) as observed in 2005 (15.1). One notable difference was in Nampula where there was a significant increase in the mean number of whitefly per plant from 2005 (3.4) to 2006 (7.8). The other five provinces exhibited a similar pattern in 2005 and 2006.

**Table 2.6:** CMD Severity and whitefly numbers in the 2006 survey

| Var | Provinces | N | NP | Mean score ± SE |
|---|---|---|---|---|
| CMDSEV | Maputo | 100 | 1657 | 3.2±0.4 |
| | Gaza | 100 | 2290 | 3.3±0.3 |
| | Inhambane | 100 | 877 | 3.2±0.5 |
| | Zambezia | 100 | 1080 | 2.7±0.5 |
| | Nampula | 95 | 1759 | 2.9±0.7 |
| | Cabo Delgado | 105 | 566 | 2.8±0.6 |
| | | | | **Mean no. per plant ± SE** |
| WFNO | Maputo | 100 | 3000 | 0.1±0.4 |
| | Gaza | 95 | 3000 | 2.7±0.9 |
| | Inhambane | 105 | 3000 | 0.8±0.5 |
| | Zambezia | 100 | 3000 | 15.8±0.1 |
| | Nampula | 100 | 3000 | 7.8±0.8 |
| | Cabo Delgado | 100 | 2399 | 5.8±0.4 |

**CMDSEV: Cassava mosaic disease severity; WFNO: mean whitefly number per plant per province; mean score = mean score between range of 1-5; N= number of fields per province; NP = number of plants scored per province**

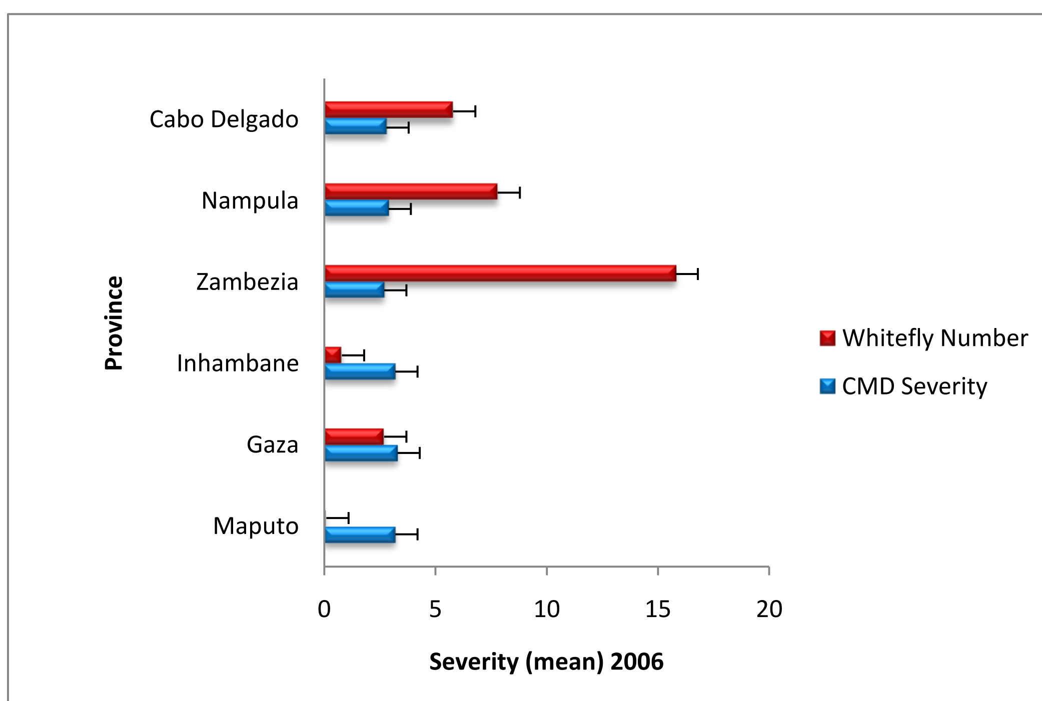


**Figure 2.5: CMD severity and mean whitefly numbers per plant in the six provinces from the 2006 survey**

### 2.5.3 Phylogenetic analysis of whiteflies

A list of all the reference whitefly sequences, and nt % sequence identities, used to compare with the Mozambique collection, are presented in Appendix 2. Two hundred and three whitefly collections were made but only 63 adult instars were sequenced from 5 districts in each of the six provinces. From the table in appendix 2, it can be observed that the mtCO1 gene of *B. tabaci* from Mozambique samples show little nucleotide sequence variability (<5%) and form a distinct phylogenetic group within the southern African clade (Figure 2.6).

Not surprisingly, *B. tabaci* from Mozambique, showed the highest % nt sequence identity (>89%) with those from South Africa, and other southern African cassava haplotypes. Figure 2.6 further demonstrates that the cassava *B. tabaci* haplotypes in east and southern Africa form a tight sub-clade (sub-group).

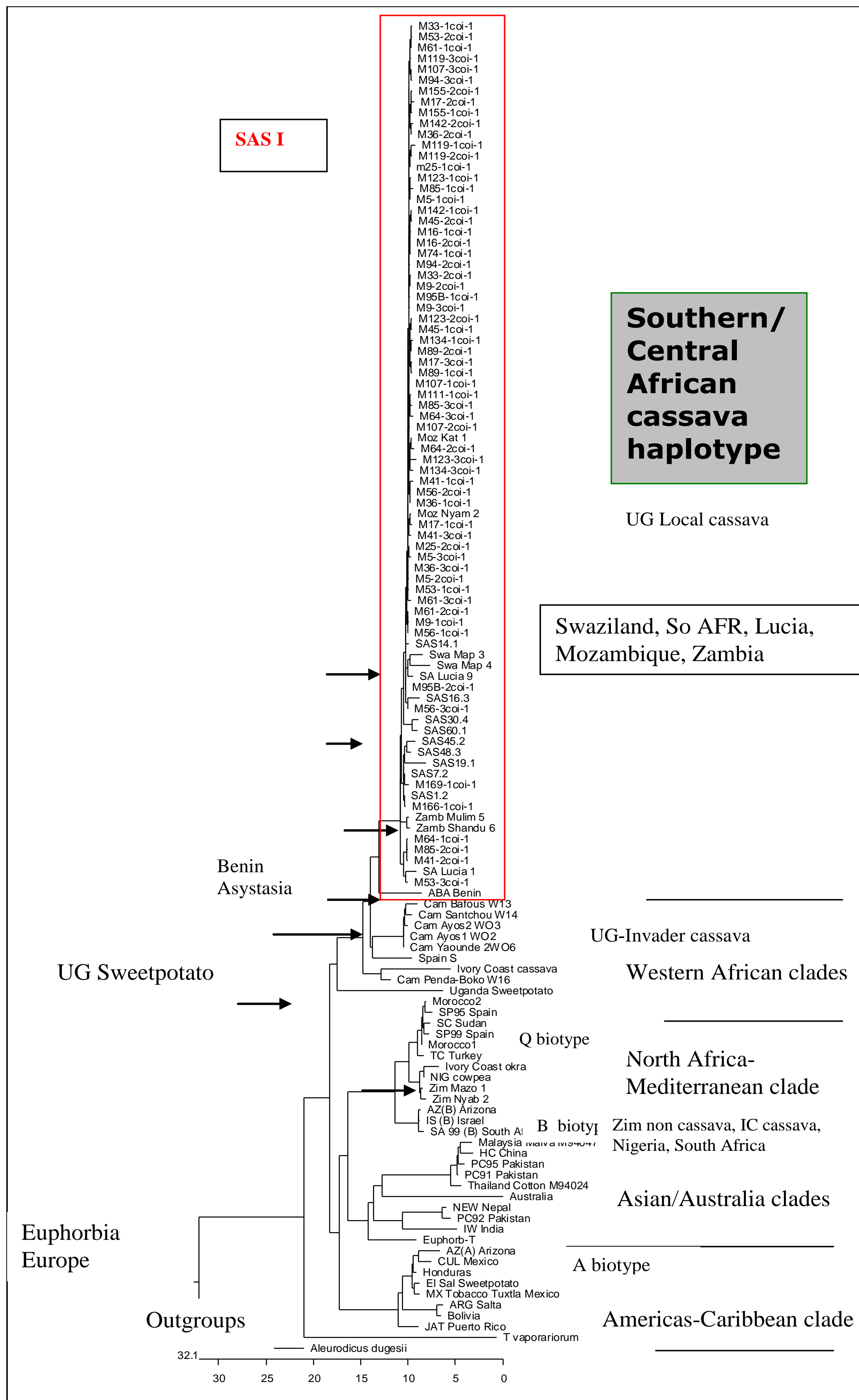


**Figure 2.6: Phylogenetic relationships of *Bemisia tabaci* from Mozambique with selected haplotypes and biotypes based on the mtCOI sequence (780bp)**

## 2.6 Discussion and Conclusions

Cassava mosaic disease (CMD) was found throughout the six provinces surveyed. This survey showed that CMD is one of the most important constraints affecting cassava production in Mozambique, and confirms earlier studies in 1999-2001, that the disease is having a wide impact on cassava production.

The general pattern for both surveys in 2005 and 2006 was that CMD was significantly higher in the southern provinces. It is interesting to note that Cassava mosaic disease (CMD) is now reported as being more prevalent in the northern provinces and is spreading south (unpublished data). Some cassava plants showed both CMD and CBSD and the occurrence of these two viral pathogens in Mozambique is of huge concern. Cassava is grown extensively in both the northern and southern provinces.

The high variability in CMD incidence in the districts (data not shown) and variability in disease incidence between provinces (appendix 1) could have been for several reasons, one being the variety of cassava grown. Some cassava cultivars, such as Munhaca, are highly susceptible to CMD. High CMD incidences (55-76%) were found in the Maputo and Gaza provinces where the farmers widely grow the Munhaca variety. The variability of different cassava cultivars and landraces to CMD has been described by Zacarias in Mozambique (unpublished data).

In Inhambane, Zambezia, Nampula and Cabo Delgado the incidence was moderate to low, while in Nampula the incidence was high in 2005 (56%), but in 2006 CMD incidence was drastically lowered to 7% in this province. This may probably be due to the agronomic station in Nampula multiplying and distributing clean (CMD and CBSD) cassava cuttings to the farmers.

Results of this study are consistent with those of Thresh *et al.* (1999-2001) where huge differences in incidence between provinces and districts were observed. For example, in the 1999 survey in the provinces of Nampula and Zambezia, the overall incidence of CMD was 13% in Nampula and 29% in Zambezia. In both provinces, there were substantial differences between districts. Mean incidence ranged from 0 to 34% in Nampula and from 4 to 75% in Zambezia, while in the 2003 survey assessing cassava and sweet potato pests and diseases in Mozambique (Toko, 2003), CMD incidence ranged between 0% to 100% in the seven provinces surveyed. This indicates that the situation had not changed over the

three years between surveys probably because farmers were planting the same varieties and new or virus-free varieties were not available for distribution.

While significant differences in CMD incidence were observed between provinces (P□ 0.05) in both 2005 and 2006 seasons, we cannot say if there were significant differences between seasons as a Two-Way ANOVA was not perfomed (raw data lost from computer hard drive before this was done). However we can certainly point out that the patterns observed between the two years appeared to be similar.

Cutting infection was significantly higher (max. 73 %) than whitefly infection, which was low (max. 21%) in all survey areas in 2005 and 2006 (Tables 2.3. and 2.6). While there was variability or differences in CMD incidence between the districts (data not shown), whitefly numbers and transmissions by whitefly were consistently low. This emphasizes the fact that the primary source of infection is through cuttings. The above observations are consistent with various studies conducted elsewhere in Africa continent indicating that cutting infection is the primary source of the virus (Trench and Martin, 1985; Bock, 1994; Legg and Ogwal, 1998; Legg and Raya, 1998). An exception would be the epidemic in Uganda in the 1990s where transmission of a new recombinant virus (EACMV-UG) was shown to be driven by a novel and invading *B. tabaci* genotype Ug2 (Legg *et al*., 2002).

In South Africa, a study by Mabasa in 2004 -2006 (unpublished) also demonstrated that dissemination of infected cuttings (stakes) is the main cause of CMD. Therefore the control of CMD through whitefly control and propagation of clean material can assist in keeping infection lower. The problem, however, is the widespread illegal movement of cassava across borders, which is difficult to control.

The whitefly borne disease incidence was very low in all surveyed areas in the two seasons (see fig. 2.4 and 2.5). The above 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 found the same results were conducted in Rwanda, Kenya and the Democratic Republic of Congo (DRC) (Legg, *et al.,* 2001; Thresh, 2001; Munga and Thresh, 2002; Okao-Okuja, 2004).

Numbers of whitefly were moderate to low in all provinces except in Zambezia where the mean numbers per plant per province were high (15.1 and 15.8 in 2005 and 2006, respectively). The high numbers of whitefly in Zambezia did not correspond to high %

CMD incidence caused by whitefly, and overall incidence in Zambezia was lower (27 to 36%) than some of the other provinces. CMD incidence in Zambezia was mainly by infected cuttings.

The disease incidence refered above suggests that for some reason the whitefly may not be transmitting disease and that they arrived later in the season. It is possible, but highly unlikely, that the whitefly in Zambezia is not the *B. tabaci* species but another whitefly species, for example *B. afer*, which is known to feed on cassava (Brown *et al*., 1995). MtCO1 sequencing did not reveal any non-*B.tabaci* species, but this does not rule out the possible presence of sporadic mixtures of other whitefly species. However, no prior information on whitefly in Mozambique was available. These results concur with findings of other surveys (Thresh, 2001; Toko, 2003) where higher whitefly numbers were observed in the Zambezia Province. Reasons for higher whitefly numbers between 1999 and 2006 in Zambezia are unknown at this stage. Information on whitefly preference for different cassava varieties is not known, but may be a contributing factor. The results, obtained in this study, could further be explained by the fact that Zambezia has higher rainfall compared to the provinces in the south, and that resulted in vigorous plant growth that supported higher whitefly populations. It has been observed that cassava growth in Zambezia is noticeably higher than the other 5 provinces from this survey, but reasons for this, possibly soil fertility, need to be established in further studies. Differences in rainfall alone would not satisfactorily explain the reasons for lower CMD in the northern provinces of Nampula and Cabo Delgado, but the trend appears to be a decrease in CMD as one moves north. The higher prevalence of CBSD in the north may play a role in shifting the importance of CMD as both viruses are reported to be transmitted by *B. tabaci*. We speculate that some displacement of begomoviruses by CBSV, through possible vector preference for virus transmission. The fact that wetter weather supports higher CMD spread compared with drier weather has been established by Storey (1936).

Disease severity scores were low to moderate (Tables 2.5 and 2.6) in all provinces in the two seasons, although in the southern provinces, infected cassava appeared to have higher severity scores (>3) compared to the northern provinces (<3). This could be explained by the use of the Munhaca variety in the south which is highly susceptible to CMD. During this study the predominant species found was EACMV-UG which generally induces more severe symptoms than ACMV. A larger number of samples analysed for virus identity

were from the southern provinces (34) compared to the three northern provinces (21) (due to leaf/ DNA degradation) (Table 3.1; Chapter 3). Out of 55 samples, 7 plants had mixed infections, while 13 and 5 plants were infected with EACMV-UG in the three southern and three northern provinces, respectively. EACMV-UG is known to cause more severe symptoms that ACMV, and may have contributed to higher symptom severity in the southern provinces. Mixed infections may also increase severity (Legg and Fauquet, 2004), although in this study there were not enough samples with mixed infections (5 from the northern provinces and 2 from the southern provinces) to infer a significant correlation. In no cases was the mean severity score higher than 4, although many individual samples showed above average severe symptoms (>3.3). Previous surveys in Mozambique showed average scores to be as high as 4 (pers. observation; data unpublished), and it is thought that the symptom scoring may have been slightly underetsmated. In Tanzania recently, severity of CMD has been recorded as >4 and has thought to be due to mixed infections with the presence of a putative satellite (Ndunguru *et al*., 2005).

While this study did not screen specifically for satellites, a preliminary study screening cassava from Inhambane Province by RCA did not detect any satellites (unpublished data). Should satellites spread from Tanzania south into Mozambique this could result in exacerbation of the disease and a disaster for the farmers. In summary, it would appear that CMD is important throughout Mozambique, but is more prevalent in the southern regions.

Molecular phylogenetic analysis of *Bemisia tabaci* (Gennadius) mitochondrial cytochrome oxidase I (mtCO1) has shown populations grouping into four major phylogenetic lineages (clades) represented by the Sub-Saharan Africa Group; Asia-Australia group; American-Caribbean Group and Mediterranean-North Africa-Middle East (Brown and Idris, 2005).

Genetic comparisons based on the mtCOI gene for the *B. tabaci* species complex has revealed as high as 26% variance (Brown and Idris, 2005) between groups. It is well documented that that *B. tabaci* native to Africa colonized cassava after its introduction into Africa from South America hundreds of years ago (Story and Nichols, 1938). In this study, the *B. tabaci* from the six provinces in Mozambique form a single unified group (with little to no sequence divergence) within the Sub-Saharan African clade.

Specifically the Mozambique haplotypes clustered with the southern African haplotypes from South Africa, Swaziland and Zambia (Berry *et al*., 2004). The nucleotide sequence

percentage similarity between samples within the Mozambique cluster was 5 % while the nt sequence divergence from the other southern African haplotypes was more variable (89-97%). This demonstrates that further divergence between *B. tabaci* haplotypes in the sub-Saharan Africa clade.

These above results confirm the previous studies (Frohlich *et al.,* 1999; Brown, 2000) that have demonstrated that sister clades within the major phylogenetic groups occur with a basis in phylogeography. In a previous study by Berry *et al.* (2004), the % nt sequence divergence of *B. tabaci* from Zambia, South Africa, Mozambique and Swaziland, within the Sub-Saharan Africa clade, was shown to be 0.4-16.5%.

The Mozambique unified group showed a maximum of 5% nt sequence divergence amongst the isolates sequenced. While the majority of haplotypes in a defined geographic region, such as southern Africa, would fall within a single sister clade, certain haplotypes may fall outside the group, but these generally are suspected to be "strays" that have moved via human traffic (Viscarret *et al*., 2003; Berry *et al*., 2004).

Within the Sub-Saharan African major clade, there appears to be an Eastern-southern African sub-group of early monophagous cassava-restricted whitefly haplotypes. Other studies have demonstrated several African indigenous *B. tabaci* haplotypes (Brown, 2000), including the Ivory Coast (western Africa) biotype which also has a limited host range of cassava and wild eggplant and transmits ACMV (Burban *et al*., 1992).

The Mozambique-South African cluster clusters separately from the distinct cluster associated with the epidemic of severe cassava mosaic disease in Uganda (Legg *et al*., 2002). Two *B. tabaci* genotype clusters were identified in Uganda, which diverged by 8%. Neither the Ugandan genotypes, nor the southern African cluster, were identified with the widely distributed polyphagous B biotype of Old World origin (also identified in South Africa) (Berry *et al*., 2004). It is hypothesised that the occurrence of two major cassava mosaic virus lineages (West and East Africa; ACMV and EACMV) may be linked to the divergence of two major *B. tabaci* sister clades in the same geographical locations. This would be in line with the observation that begomovirus transmission determinants of the whitefly are linked to the coat protein of the virus suggesting that virus and vector have co-evolved (Harrison and Robinson 1999; Brown and Czosnek 2002). It is likely that the near-monophagous cassava whitefly group in southern Africa (including Mozambique and

South Africa) has evolved in part as the result of geographical isolation and host range restriction. This hypothesis is supported by the limited hosts that *B. tabaci* has been found on, from studies in SA (unpublished data). In conclusion, the mitochondrial CO1 sequence for *B. tabaci* has shown to be informative for predicting phylogenetic relationships between clades (genetic groups) within the *B. tabaci* species complex and was able to group haplotypes of whitefly of cassava in Moambique into the sub-Saharan lineage. Further analysis of polymorphic sites in consensus amino acids for the mtCO1 gene would be useful in tracking selection of haplotypes in relation to cassava begomoviruses and their recombinants. Analysis of genome changes in the *B. tabaci* mtCO1 coding region (Brown and Idris, 2005) has suggested that this gene has evolved under positive selection leading to genetic differentiation in the four major extant phylogenetic lineages, one of which is the sub-Saharan clade.

# CHAPTER THREE: IDENTIFICATION OF BEGOMOVIRUSES IN CASSAVA IN THE SIX PROVINCES IN MOZAMBIQUE

## *3.1 Abstract*

Cassava mosaic geminiviruses (CMGs) show large diversity because their genomes have a high plasticity, and variations can occur naturally due to inter- and intra-species recombination. Geminiviruses can recombine and exchange sequences between genomic components. Intra- and interspecies recombination events have contributed to begomovirus diversity and evolution. In sub-Saharan Africa, six distinct species of CMGs have been sequenced and many genetic variants or isolates have also been reported.

In order to determine the biodiversity of CMGs in Mozambique, 285 infected cassava leaf samples were collected throughout the six provinces in Mozambique. Of these 109 were screened by PCR using core coat protein (CCP) universal primers for begomoviruses. Sixty samples revealed positive CCP amplification. Many of the samples were degraded due to lack of cooling or freezing facilities in some of the remote districts. From the 60 CCP- positive samples, full-length DNA A amplification was performed (either by Rolling Circle Amplification (RCA)-PCR or PCR from Total Nucleic Acid) from 55 of the leaf samples. Restriction Length Fragment Polymorphism (RLFP) analysis was undertaken using the enzymes *EcoR*V, *Dra*I and *Mlu*I. The results showed that 63.4% (35 of the 55 cassava leaf samples) were EACMV, 12.7% (6/55) were ACMV species; 12.7 (7/55) were mixed infections of EACMV and ACMV; and 6 samples were not unidentified by RFLP. However, RFLP was not able to distinguish between EACMV and EACMMV and EACMCV.

Six full-length DNA A clones, one from each province, showing unusual RFLP patterns, were cloned and sequenced. Consensus sequences were aligned with other cassava begomoviruses and selected begomoviruses from southern Africa. Phylogenetic analysis (parsimony) revealed that virus isolates from Maputo, Inhambane and Nampula Provinces exhibited 95-97% nucleotide sequence divergence/similarity to *African cassava mosaic virus*-[Nigeria]; the virus isolate from Gaza Province was 99% similar to *South African cassava mosaic virus* - [South Africa]; while the Zambezia Province virus was most closely aligned (94%) with EACMMV (*East African cassava mosaic Malawi* virus-

[Malawi: MH]).The isolate from Cabo Delgado aligned most closely (96%) with the *East African cassava mosaic Cameroon virus*- Cameroon, and less closely (87%) to EACMMV. This study reports, for the first time, diversity for cassava begomovirvus species in Mozambique, similar to previous studies in South Africa, and demonstrates the mixture of geminivirus species from east and west Africa. Sequence variation of > 92% indicates that the cassava geminiviruses in Mozambique are genetic variants ACMV-[NG], EACMMV and EACMCV. This is the first report of SACMV in Mozambique.

## *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 (CMGs) are characterized by a circular single stranded DNA (ssDNA) genome encapsidated in a twinned (germinate) particle of approximately 20 x 30nm (Zhang *et al*., 2001; Bottcher *et al*., 2004) and transmission by *Bemisia tabaci*. CMGs have genomes that consist of two components, termed DNA-A and DNA-B (Stanely and Gay, 1983; Stanley *et al*., 2005). Together these components contain six open reading frames s) four on DNA-A and two on DNA-B (Stanley and Thwnsend, 1986).

The two components share a conserved intergenic "common region" (IR or CR) of approximately 200bp in size and have 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) (Hanley-Bowdoin *et al*., 2004; Vanitharani *et al*., 2004). DNA-B encodes for the nuclear shuttle protein (NSP, BC1) and the movement protein (MP, BC2) responsible for virus movement within and between cells (Hanley-Bowdoin *et al*., 2004).

Initially 3 distinct groups of whitefly transmitted CMGs were distinguished, based on nucleotide sequences by Hong *et al*. (1993). Group A, *African cassava mosaic virus* (ACMV) was isolated in West Africa, Burundi, Chad, Uganda, Cameroon and western parts of Kenya. Group B, *East African cassava mosaic virus* (EACMV) was isolated in the eastern parts of Kenya, Malawi, Madagascar, Zimbabwe and Tanzania and later detected in Uganda and Zambia in 1997 and in Cameroon and Nigeria in 1998 and 1999, respectively. Group C consisting of *Indian cassava mosaic virus* (ICMV) was discovered in India and

Sri Lanka. Berrie *et al*. (2001) isolated and characterized from South Africa the 4th distinct group of CMGs and this distinct group was named *South African cassava mosaic virus* (SACMV). SACMV was identified by coat protein (CP) and (CR) nucleotide comparison with other CBVs and host range studies. SACMV was found to be closely related to EACMV in both its DNA-A (85%) and DNA-B (90%). However, it was considered to be a distinct virus as a result of its DNA-A containing less than 90% similarity to DNA-A of EACMV Berrie and Rey. (2001). Natural recombination was associated with the occurrence of evolution of a new distinct virus Berrie and Rey. (2001).

More recently, geminivirus taxonomy and nomenclature (Fauquet *et al*., 2008) has been defined by a system of demarcation criteria, refined from the ICTV Geminiviridae Study Group which had proposed an 89% nucleotide (nt) identity threshold between full-length DNA-A component nucleotide sequences for defining begomovirus species (Fauquet *et al*., 2003). The present classification by Fauquet, (2009) proposed guidelines to classify and name geminiviruses below species level.

Based on pairwise sequence comparisons (Clustal V algorithm in DNAStar MegAlign software), viruses at the 85-94% nt identity level are proposed to be strains, while viruses between 92 and 100% nt identity constitute variants. Eight distinct species of cassava – infecting begomoviruses have been identified, of which seven infect cassava in Africa: *African cassava mosaic virus*; *East African cassava mosaic virus* (EACMV), East *African cassava mosaic Cameroon virus* (EACMCV), *East African cassava mosaic Kenya virus* (EACMKV), *East African cassava mosaic Malawi virus* (EACMMV), *East African cassava mosaic Zanzibar virus* (EACMZV) and *South African cassava mosaic virus* (SACMV) (Fauquet *et al*., 2008).

With the availability of improved techniques such as PCR and Rolling Circle Amplification (RCA) a better understanding and appreciation of the complexity and distribution of CMGs has been achieved. Over the past thirteen years several cassava geminiviral species have been identified in different regions of the African continent. For example, several studies have shown the presence of ACMV in all parts of the continent where cassava is grown. EACMV is now found in West Africa. Fondong *et al*. (2000) isolated an EACMV-like virus (EACMCV) in Cameroon and in Ivory Coast. Pita *et al*. (1999) isolated a similar virus designated EACMCV-IC.

In southern Africa, Berrie *et al*., (1997) isolated SACMV for the first time in South Africa. It 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 found CMG species (i.e ACMV, ACMV-[UG], EACMV and SACMV) in six southern African countries (South Africa, Swaziland, Mozambique, Angola, Zambia and Zimbabwe).

The identification of the Ugandan variant (EACMV-UG [UG:Severe]), which was responsible for the CMD epidemic that almost wiped out an entire crop in Uganda during the early 1990s (Zhou *et al*., 1997), was one of the milestones in CMGs diagnostics. It was in this study that the first evidence of recombination and/or pseudorecombination in geminiviruses were presented. Recombination is now known to be a driving force in the evolution of ssDNA geminiviruses (Padidam *et al*., 1999; Ndunguru *et al*., 2005; Rojas *et al*., 2005; Shepherd *et al*., 2008) and is thought to contribute to adaptation to host. In order to test for recombination, various recombination detection methods have been used (Martin *et al*., 2005). The RDP2 method by Martin *et al*. (2005) uses ten published recombination methods, and has been rigorously applied to the study of evolution of geminviruses.

In order to establish the genetic relationship between virus isolates, phylogenetic analyses need to be carried out on the nucleotide or amino acid sequences of the viruses. A phylogeny is the evolutionary history of a group of entities, where evolutionary relationships are represented by branching diagrams or trees, with branches joined by nodes (Xiong, 2006). Bioinformatic analysis requires sequence comparisons or alignments that can be carried out on databases such as Genbank, the primary nucleotide sequence database in NCBI (www.ncbi.nlm.nih.gov/Genbank).

The degree of sequence conservation in the alignment of geminiviruses can therefore provide information on relatedness between species, strains and isolates. Sequence similarity between virus nucleotide (or amino acid) sequences can be quantified using percentages (Xiong, 2006). Pairwise sequence alignment of an unknown sequence (virus) with sequences in a database requires algorithms for searching and retrieving sequences from a database (Harrison and Langdale, 2006).

One type of algorithm is the heuristic type, which is a computational strategy to essentially take shortcuts to search large databases more reliably and in a short space of time (Xiong,

2006). Heuristic algorithms, such as BLAST (Basic Local Alignment Search Tool), are 50-100 times faster than dynamic programming, and use a heuristic word method for fast pairwise sequence alignment (Altschul *et al*., 1997). When sequences are retrieved from BLAST searches, they are allocated an e-score, which is an indication of similarity between two sequences. The closer the e-value is to zero, the higher degree of similarity between two sequences. Once data are aligned there are three major types of phylogenetic techniques that can be used, namely parsimony, maximum likelihood (ML) and neighbour joining (NJ). All of these analyses can be performed using PAUP software (Swofford, 2003).

Neighbour joining methods calculate pairwise distances between sequences and group sequences that are most similar. Parsimony, on the other hand, assumes that shared nucleotides in different groups result from a common ancestry. Groups would contain entities (viruses) that share nucleotide (aa) sequences and the simplest explanation for the evolution of these nucleotides is taken to be the most parsimonious one.

A consensus tree is derived from multiple trees and includes non-contradictory topologies. Maximum likelihood analysis computes the probability that a data set fits a tree derived from the data set, given a specified model of sequence evolution. The maximum likelihood score (best fit) is a function of both the branch lengths (the longer the branches, the more the nucleotides have changed and the less the similarity between two sequences are) and topology.

Parsimony and ML trees are preferred over NJ as they are more rigorous in exploring the relationship between the tree and the groups included (Harrison and Langdale, 2006). Parsimony and ML trees both rearrange the branches of the starting or initial trees to choose the best trees, but parsimony selects the tree that minimizes the number of character (nucleotide or aa) state changes, whereas ML selects for the tree that best fits the data (likelihood). Molecular trees should be rooted to provide some evidence for ancestor-descendant relationships, and one of the most common methods is to use out-group rooting (Maddison *et al*., 1984).

An ideal out-group, for example, would be a virus sequence that is closely related to, but does not fall within, the in-group that is being compared. To statistically assess how accurately the phylogenetic trees reflect true biological events, adding bootstrap values can

test the robustness of the data used to generate the trees. The bootstrap value shows the percentage of times that a clade appears when individual nucleotides in the data set are randomly swopped with other nts from the set over a number of specified replications (Felsenstein, 1985). Most bootstrap values are generated from 1000 replicates.

Previous surveys of cassava mosaic geminiviruses (CMGs) in Mozambique were undertaken from 1999-2003 in only three provinces (Nampua, Inhambane and Zambezia), and the viruses were never identified. In this study, RFLP and phylogenetic sequence analyses were applied to identify CMGs in six provinces in Mozambique, and to establish their relationship to other geminiviruses.

## *3.3 Materials and Methods*

### 3.3.1 Sampling using FTA cards and leaves

A total of 285 young symptomatic leaves were removed from infected plants and placed on FTA classic cards. An eppendorf tube was used to apply pressure with a slight twisting until sap penetrated the reverse side of the FTA paper. Cards were dried at room temperature until requied for PCR.

### 3.3.2 Preparation of cards for PCR

One 1 mm diameter punches were removed from each cholorophyll-stained region using a harris punch, and placed in a sterile 1,5 ml eppendorf tube.The paper discs were washed in 500µl of TE buffer for five minutes, followed by sequential five minutes washes into 500µl of 70% ethanol and FTA Purification Reagent (Whatman). Sample punches were then transferred to a fresh 1.5 ml eppendorf tube and allowed to dry for two hours at room temperature. The punched discs were then used for the PCR reaction. The same procedure for PCR from leaf samples was used (see below). PCR was performed twice on the same leaf samples. PCR with FTA cards was not sucessful. The same number of symptomatic leaf samples (287), were collected from six provinces, placed in small plastic bottles, and stored on ice until they were brought to the laboratory where they were stored at 20°C.

### 3.3.3 Total nucleic Acid (TNA) Extraction

Total nucleic acid (TNA) extraction was performed on 109 of the leaf samples using the CTAB method as described by Doyle and Doyle (1987).

An extraction buffer comprising of 20g CTAB/L, 1.4 M NaCl, 20 mM EDTA, 0.1M Tris, pH 8.0 was preheated at 65°C and 500ul added along with 1ul ß- mercaptoetanol (0.2%v/v) to 50mg ot ground leaf samples and this incubated at 65°C for 60 min. TNA was extracted by adding 500 ul chloroform:isoamyl alcohol (24:1) the mixture was then inverted and centrifuged at maximum speed (13 rpm) at 4°C for 10min. The aqueous layer was transferred to new sterile tube and the step repeated.

To precipitate the TNA, 500ul of isopropanol was added and tube centrifuged at maximum speed at 4°C for 10 min and the supernatant then poured off. The pellet was washed in 500ul ice-cold 70% ethanol and spun at maximum speed at 4°C for 10min and this step repeated. The pellet was air dried, then resuspended in 50ul 1X TE buffer (10mM Tris-Cl, 1mM EDTA pH 8.0) and 1 ul RNase A (10mg/ml). After complete resuspension, TNA was quantified using a ND-1000 Nanodrop (Nanodrop ) and quality was determined on 1% agarose gel in 1X TAE and visualized by ethidium bromide staining DNA (10ug/ml).

### 3.3.4 PCR screening of leaf samples for the core coat protein (CCP)

Core coat protein (CCP) polymerase chain reaction (PCR) was carried out in the Biorad ThermoCycler (Biorad, US). This was performed to screen the 109 TNA leaf samples for the presence of CMGs. Samples were screened for whitefly-transmitted begomoviruses using degenerate CCP primers that amplify ACMV, EACMV and/or SACMV.

CCP primers used were AV514 GCCCWTGTAGAGRAAGCCMAGRA and AC1048 GRTTDGARGCATGHGTCANGCC (Wyatt and Brown, 1996). These anneal to a highly conserved region within the core region of the coat protein amplifying a 550 bp region. Each CCP reaction consisted of 5µl of 10X $(NH_4)2SO_4$ buffer, 1 µl dNTP's (10 mM), 7 µl $MgCl_2$ (25mM), 2.5 µl of each of the CCP primers (10 µM), 1U Taq polymerase, 1µg TNA (2 µl) and nuclease free water to make up a 50 µl volume reaction.

An ACMV-infected TNA sample was used as a positive control while no template was placed in the negative control reaction mixture. CCP PCR cycling conditions were as

follows: initial denaturation at 95 ºC 5 min, 35 cycles at 95 ºC for 45 sec, 55 ºC for 45 sec and 72 ºC for 45 sec and a final extension at 72 ºC for 7 min. Amplification products were analysed in 1% agarose in 1X TAE and visualised by ethidium bromide staining (10ug/ml). Of the 109 samples 60 were CCP positive.

Near full-length virus DNA-A (c. 2760-2780 bp of) from 60 selected infected cassava samples (selection based on different districts and provinces) were PCR amplified using the universal primers UNI/F (5' KSGGGTCGACGTCAAGACTTRTAC 3') and UNI/R (5' AARGAATTCATKGGGGCCCARRGACTGGC3 ) (Briddon and Markham, 1994). These primers anneal to the AC1 region of DNA A.

Each PCR reaction mixture consisted of 5µl of 10X (NH4)2SO4 buffer, 1 µl dNTP's (10 mM), 4 µl $MgCl_2$ (25mM), 1µl of each of the Uni primers (10 µM), 1U Taq polymerase (Accuzyme; Bioline), 1 µl of 10 fold diluted RCA product and nuclease free water to make up a 50 µl volume reaction. As with the CCP PCR, the same experimental controls were included. The amplification conditions were as follows: 95˚C for 2min, 30 cycles of 95˚C for 1min, 56˚C for 90 sec and 72˚C for 4min followed by a final extension of 10 min at 72˚C. Amplified products were analyzed in 1% agarose gel in 1X TAE and visualized by ethidium bromide staining (10ug/ml).

Near full-length DNA A amplification was difficult for some TNA samples as some of the DNA had degraded. Therefore RCA, using the TempliPhi Kit (Amersham Bioscience), was used to amplify these samples. The kit- provided protocol was followed: 5 ul of TempliPhi sample buffer was added to 1ug of TNA and heated at 95 ºC for 3 min, then cooled down to room temperature. Once cool, 5 µl of reaction buffer (salts and deoxynucleotides) and 0.2 µl of enzyme mix (Phi 29 DNA polymerase and random hexamers in 50% glycerol) were added and incubated for 18 hrs at 30˚C followed by inactivation of the enzyme at 65˚C for 10 min. A positive control using 5ng pTZ plasmid DNA and negative control (no template) experiment were also performed.

### 3.3.5 Identification of CMGs using PCR- RFLPs

Random Fragment Length Polymorphisms (RFLPs) were performed on the near-full length DNA-A PCR products (amplified from PCR or RCA products using the Uni primers). The near-full length DNA-A PCR products were first quantified using a Nanodrop ND-1000 Spectrophotometer (Nanodrop Manufacturers). RFLP patterns were generated from the

amplified near-full length DNA-A PCR products using enzymes *Mlu*I (Fermentas) and *EcoR*V (Fermentas), *DRa*1(Fermentas), and *PST*1 . Each restriction reaction consisted of 2 µl of 10X appropriate enzyme buffer, 1U appropriate enzyme, 1 µg of the near-full length PCR product and nuclease free water to make up a total volume of 20 µl.

The digestion reactions were carried out at 37°C for 3h. Post incubation all (20 µl) of the restriction digest reactions were loaded and analyzed in 1% agarose gel in 1X TAE and visualized by ethidium bromide staining (10ug/ml). Patterns obtained were analyzed and compared to available and established CBV RFLP patterns. For samples that were preliminarily identified as EACMV, additional restriction digests using *Pst*I (Fermentas) were performed (as described above) (Mabasa, 2007). This was to distinguish between EACMV and SACMV because *Pst*I has a restriction site on SACMV and not EACMV DNA-A. Once again samples were run in analyzed in 1% agarose gel in 1X TAE and visualized by ethidium bromide staining (10ug/ml).

### 3.3.6 Phylogenetic analysis of selected geminiviruses

Six geminivirus isolates (Clone 50 Cabo Delgado; Clone 54 Gaza; Clone 16 Zambezia; Clone C23C Inhambane; Clone 7 Maputo; and Clone 45 Nampula) from cassava were selected, one from each province, based on unusual RFLP pattern results. Clones were sequenced using M13/pUC F primer. However for clones that were bigger (700bp to 1500bp), both the M13/pUC F and M13/pU R primers were used. For the near-full length virus clones both the M13/pUC F and M13/pUC R were used, and internal primers designed (Inqaba, SA) for sequences that were longer than 1.4 kb.

Sequences were edited using Chromas (Version 1.45 Conor McCarthy School of Health Science, Griffith University, Australia). Sequence data of clones obtained was identified using BLAST at ncbi:nlm:nih:gov/blast site. Pairwise sequence alignment was generated using Clustal X to generate maps of the clones using Vector NTI Advance Suite (Invitrogen). Multiple sequence alignment was then performed using Clustal W in DNAMAN alignment program and the multiple sequence default parameters were used. Following this, a neighbourjoining phylogenetic tree was constructed, and a value of 1000 bootstraping applied.

The consensus sequences of each of the clones were were aligned in BLAST for nucleotide sequence similarity searches and the five most similar sequences to each were selected for

phylogeny analyses. Multiple sequence alignments (pairwise sequence comparisons with cassava geminiviruses) were performed using the ClustalW algorithm (contained in the MEGA 4.0 Software). Phylogenetic trees were constructed using the parsimony algorithm available in MEGA 4.0 * version 4.0b10 (Tamura *et al*., 2007). The bootstrap value was set at 1000 replicates, using the ≥70% confidence limits and placed at nodes. The trees were rooted using *Tomato curly stunt virus* (AF 350330) and *Tobacco leaf curl Zimbabwe virus* (AF 3500330), two begomoviruses from southern Africa (Zimbabwe and South Africa). A second tree was constructed in PAUP (Swofford, 2003) using the maximum likelihood algorithm and bootstrap value set at 1000 replicates, with the Mozambique cassava viruses compared to a selected number of geminiviruses in southern Africa. This tree was rooted using *Tomato leaf curl Madagascar virus* (ToLCMGV). A list of geminiviruses used in the analysis (with Genbank accession numbers) is presented in Appendix 3.

## *3.4 Results*

### 3.4.1 Core Coat Protein (CCP) - PCR

From the 109 cassava leaf samples analyzed, 60 tested positive for CMGs. The total nucleic acid (TNA) from the remainder of the samples were degraded due to melting of the ice in the cooler boxes when they were collected, and also as some of the districts did not have electricity so the samples could not be stored at 4°C. Some of the leaf samples were brownish in colour due to oxidation of plant components.

PCR amplifications from infected cassava leaf samples, stored on FTA cards, were unsuccessful. Since the same samples collected, and used for TNA extractions, yielded PCR products, we concluded that the DNA from the squashed leaves had not degraded. The PCR reaction was carried out in the PCR tubes with the Whatman filter paper containing the DNA in the tube. It could be that this may have affected the PCR reactions.

Figure 3.1 represents the amplified CCP PCR bands (550 bp) from 19 of the 60 infected cassava samples.

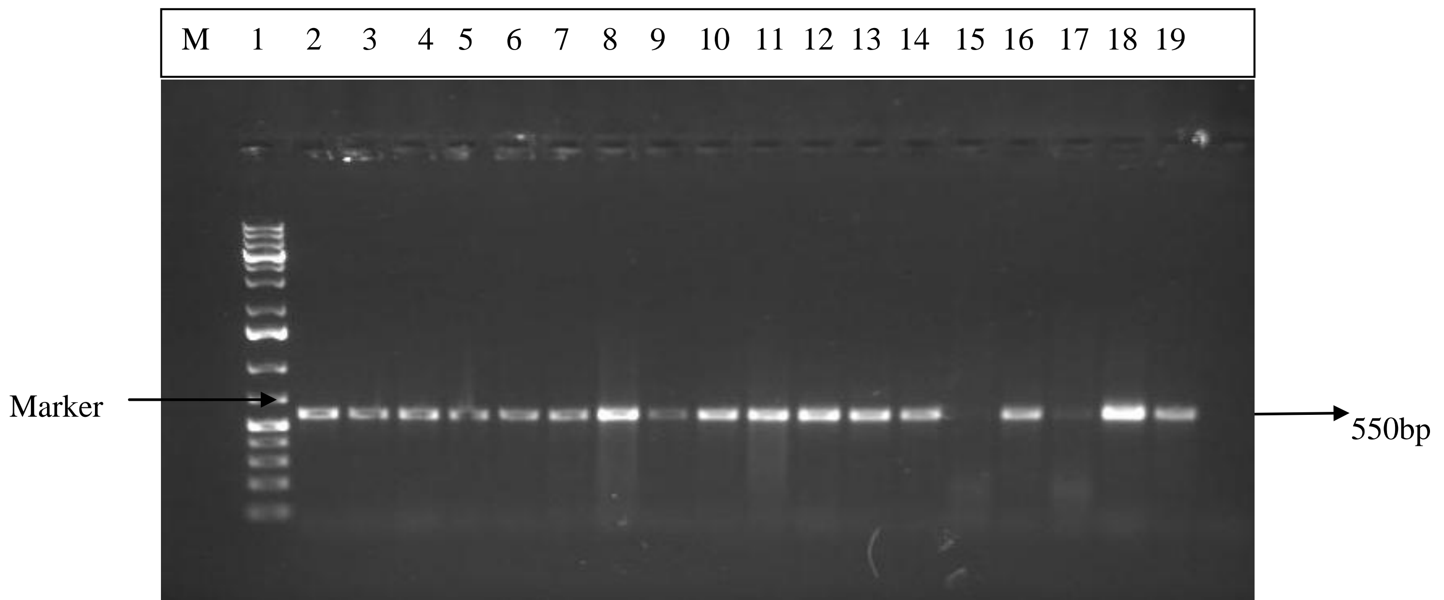


**Figure 3.1:** PCR amplification of a 550bp fragment amplified from tna extracted from infected cassava leaf samples. Amplicons were amplified using the CCP primers (Results shown above are for only 19 of the 60 samples). M: 1kb DNA molecular weight marker (0 Gene Ruler LadderPlus Fermantas)

### 3.4.2 Near-full length DNA-A Amplification

Near full length virus DNA-A were amplified in 55 out of the 60 samples samples using the Uni primers either using PCR-RCA (Figure 3.2) or from PCR amplification from TNA (Figure 3.3) Smaller fragments (1.5kp) were also amplified in some lanes [Lanes 3, 15 (Fig. 3.2) and 19 Fig. 3.3)].

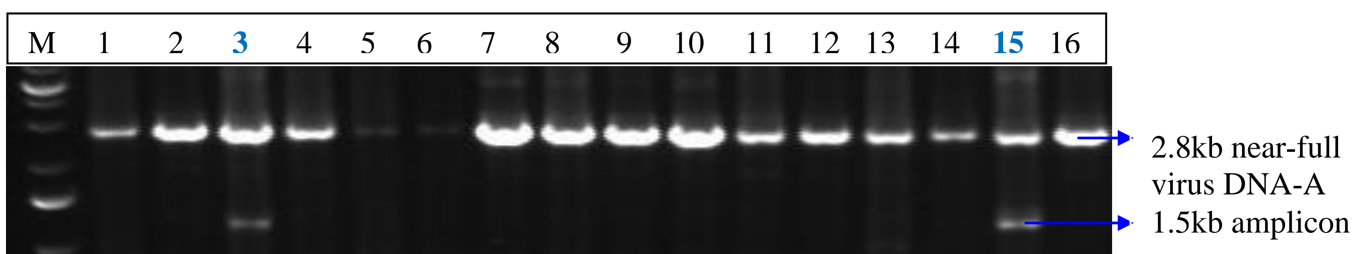


**Figure 3.2:** 1% agarose gel of near full-length virus DNA-A amplification by RCA from 16 samples amplified from RCA products using Uni primers. M: 1kb DNA molecular weight marker (0 Gene Ruler Ladder Plus Fermentas)

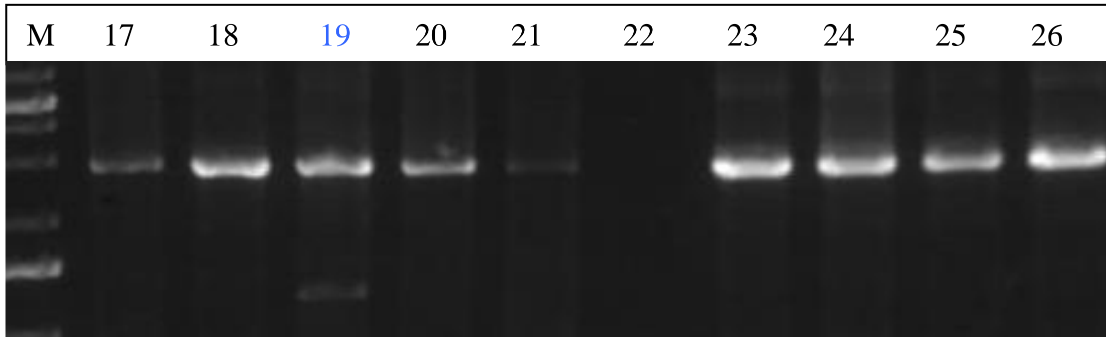


**Figure 3.3:** Full-length DNA A amplified by PCR (Uni primers) from TNA. M: 1kb DNA molecular weight marker (0□ Gene Ruler Ladder Plus-Fermentas); Lanes 17-21 and 23-26 : ±2.8bp DNA A amplified bands; Lane 22 : Negative control; Lane 23: positive control.

### 3.4.3 Restriction Length Fragment Polymorphism (RLFP) analyses

RLFP was carried out on the 55 near full-length DNA-As derived from selected samples in Mozambique (Table 3.1). The amplified near-full length DNA-As were restricted using restriction enzymes *Mlu*1, *Eco*RV, *Dra*1 and *Pst*1 (Figure 3.4). RFLP restriction enzyme (RE) patterns for the different cassava begomoviruses have been established in previous studies (Ndunguru *et al*., 2005; Sseruwagi, *et al*., 2004; and Sserubombwe *et al*., 2008), and the identity of each of the Mozambique samples were determined (Figures 3.4 – 3.7; Table 3.1). Figures 3.4 to 3.7 below illustrate some of the DNA A-restricted samples.

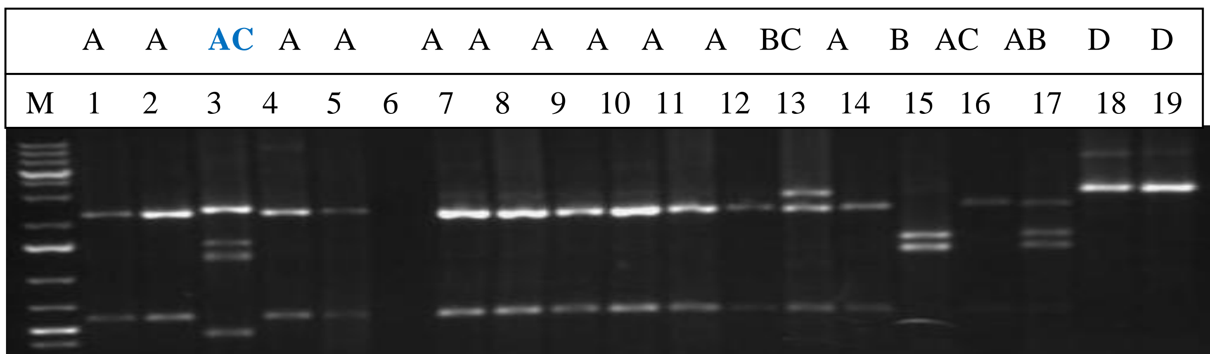


**Figure 3.4:** Characterisation of RFLP patterns of 18 CMG DNA-A samples restricted with *Eco*RV. Lane 6: No template control. M: 1kp DNA marker ladder (Fermentas)

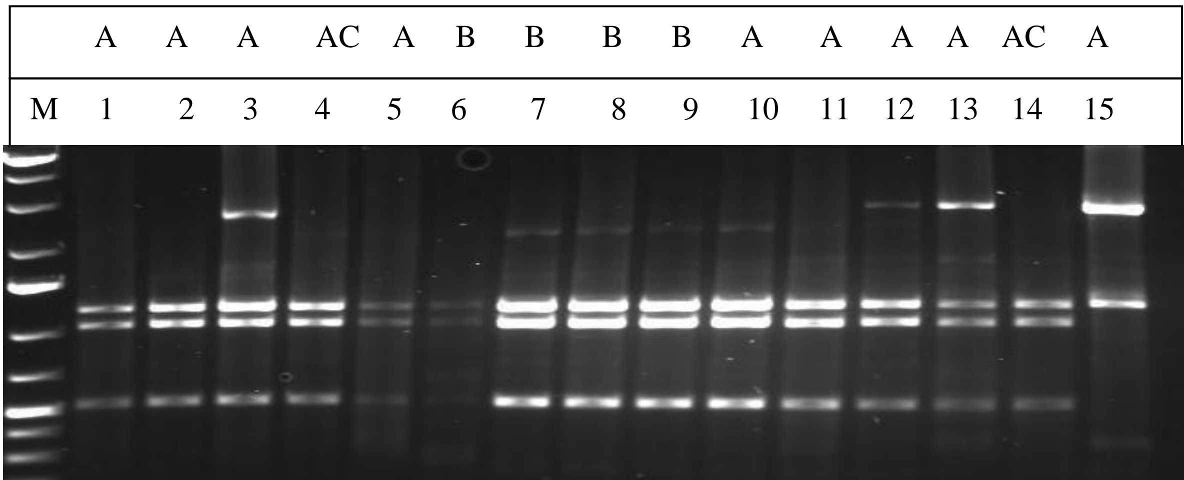


**Figure 3.5:** Characterisation of RFLP patterns of 15 CMG DNA-A samples (lanes 1-15) restricted with *Mlu*I. M: 1kp DNA marker ladder (Fermentas)

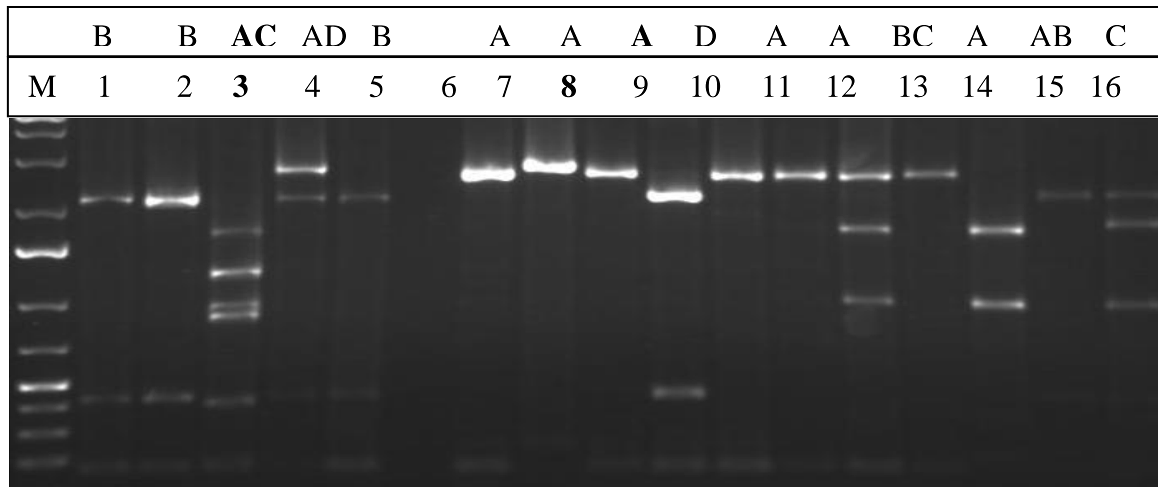


**Figure 3.6:** Characterisation of RFLP patterns of 16 CMG DNA-A samples (lanes 1-15) restricted with *Dra*I. M: 1kp DNA marker ladder (Fermentas)

**Abbreviations for Figures 3.4 to 3.6:**

A = EACMV – [KE,MW,TZ] B = EACMV – UG

C = ACMV – [UG] D = Unknown

Near-full length DNA-A amplified from all the 55 samples was further restricted with *Pst*I (Mabasa, 2007) to screen for the presence of SACMV since *EcoR*V, *Mlu*I and *Dra*I cannot distinguish between some EACMV species and SACMV. None of the amplified CMGs in the 55 samples produced RFLP patterns consistent with SACMV restricted with *Pst*I. However, later sequencing of one of the DNA A samples from Gaza Province did show that indeed one of the CMGs is SACMV, suggesting a nt difference in the *Pst*I cutting site.

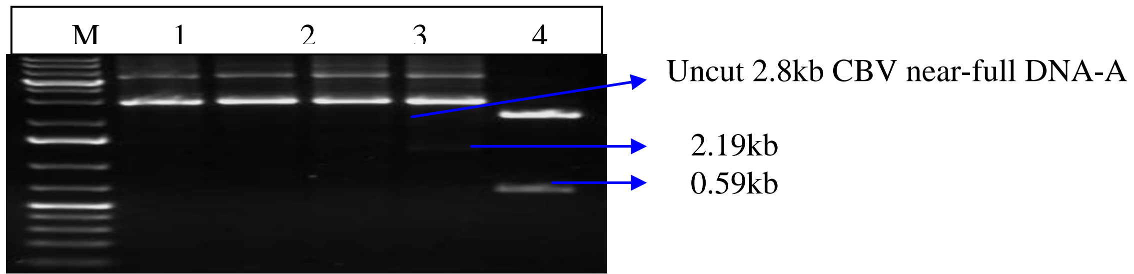


**Figure 3.7:** *Pst*I digestion of the amplified CMGs to screen for the presence of SACMV. Lanes 1-4: Uncut DNA A from 4 samples.

Of the 55 samples that were RLFP analyzed (Table 3.2), 14 came from Maputo, 8 from Gaza, 9 from Inhambane, 8 from Zambezia, 6 from Nampula and 7 from Cabo Delgado. In total, 34 samples were from the three southern provinces and 21 were from the three northern provinces. Results showed that 12.7% (7/55) were mixed infections, 63.6% (35/55) were EACMV (30% were EACMV- [KE, MW and TZ] while 32% were EACMV-UG isolates), and 12.7% (7/55) were ACMV-[UG]. Six of the 55 samples could not be distinguished unequivocally by RLFPs. This could be due to point mutations at restriction sites or mixtures of genetic variants. It is noted that because the sampling was biased to the southern three provinces (due to DNA degradation), no statistical analyses were performed. Results present only the virus identities based on the actual 55 samples and only comment on a pattern of virus distribution rather than any significant differences between the two regions.

**Table 3.1: Number of geminivirus species identified in each province using RLFPs**

| **Province** | **ACMV-[UG]** | **EACMV-[KE, MW, TZ,]** | **EACMV/ ACMV** | **EACMV-UG** | **Unknown** |
|---|---|---|---|---|---|
| Maputo | **2** | **3** | **1** | **7** | **1** |
| Gaza | **2** | **3** | **0** | **2** | **1** |
| Inhambane | **2** | **4** | **1** | **4** | **1** |
| Zambezia | **0** | **4** | **2** | **1** | **1** |
| Nampula | **0** | **1** | **1** | **3** | **1** |
| C.Delgado | **1** | **2** | **2** | **1** | **1** |
| **Total** | **7** | **17** | **7** | **18** | **6** |

### 3.4.4 Phylogenetic Analysis of Sequenced DNAs From six Geminivirus Samples

The near full-length sequences of six representative CMGs, one from each province, were determined. The full sequences are presented in Appendix 4. The DNA-A components ranged in size, from 2795nt (C16Zambezia) and 2819nt (C23C Inhambane), to 2833nt (C7 Maputo), 2861nt (C50Cabo Delgado), 2852 (C45 Nampula), and 2866nt (C54Gaza). The DNA A components contained all six ORFs, as expected in Old World bipartite geminiviruses (Rojas *et al*., 2005).

Three of the Mozambique isolates (from the provinces of Maputo, Nampula and Inhambane) grouped together in a 1000 bootstrap-supported branch with ACMV-Nigeria (Figure 3.8), sharing a nt sequence identity of 95-97%. Clone C54 from Gaza Province had 99% nt sequence similarity with SACMV-South Africa, while the virus isolates from Zambezia and Cabo Delgado provinces clustered with the EACMV clade, showing higher nt sequence divergence of 87 – 94% compared to EACMCV and EACMMV. Table of virus percentage nucleotide identities are in Appendix 3.

Phylogenetic comparisons of the ACMV and EACMV major clades or groups containing the geminiviruses from Mozambique, were made with cassava geminiviruses from southern Africa using ML algorithms in PAUP, and similar results were obtained as the CLUSTAL trees (Figure 3.9). It is interesting to note that other tobacco geminiviruses, for example *Tobacco leaf curl Zimbabwe virus* (TbLCZV), and tomato leaf curl viruses from the region, for example, *Tomato curly stunt virus* (ToCSV), a new geminivirus recently isolated in South Africa, formed a distinct clade from the cassava geminiviruses. ToCSV showed an average of 74% nt similarity with the cassava begomoviruses.

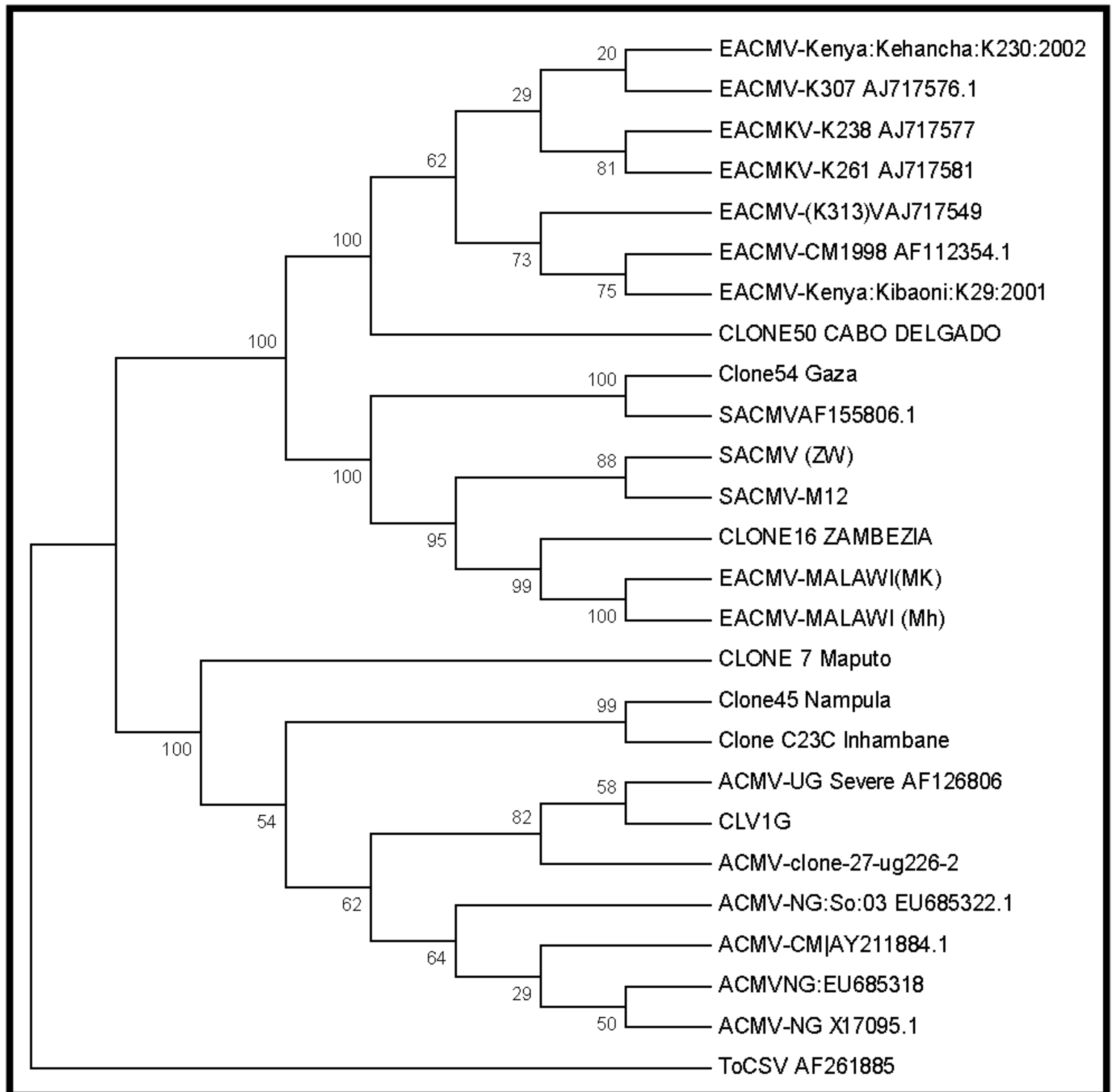


**Figure 3.8:** Most parsimonious tree (CLUSTAL V) derived from full-length DNA A comparisons between geminivirus samples from six provinces in Mozambique (C7Maputo; C45Nampula; C23C Inhambane; C16Zambezia; C54Gaza and C50Cabo Delgado) and selected cassava geminiviruses. Accession numbers and virus abbreviations are listed in Appendix 3. *Tomato curly stunt virus* from South Africa was used to root the tree. Bootstrap values are generated from 1000 replicates

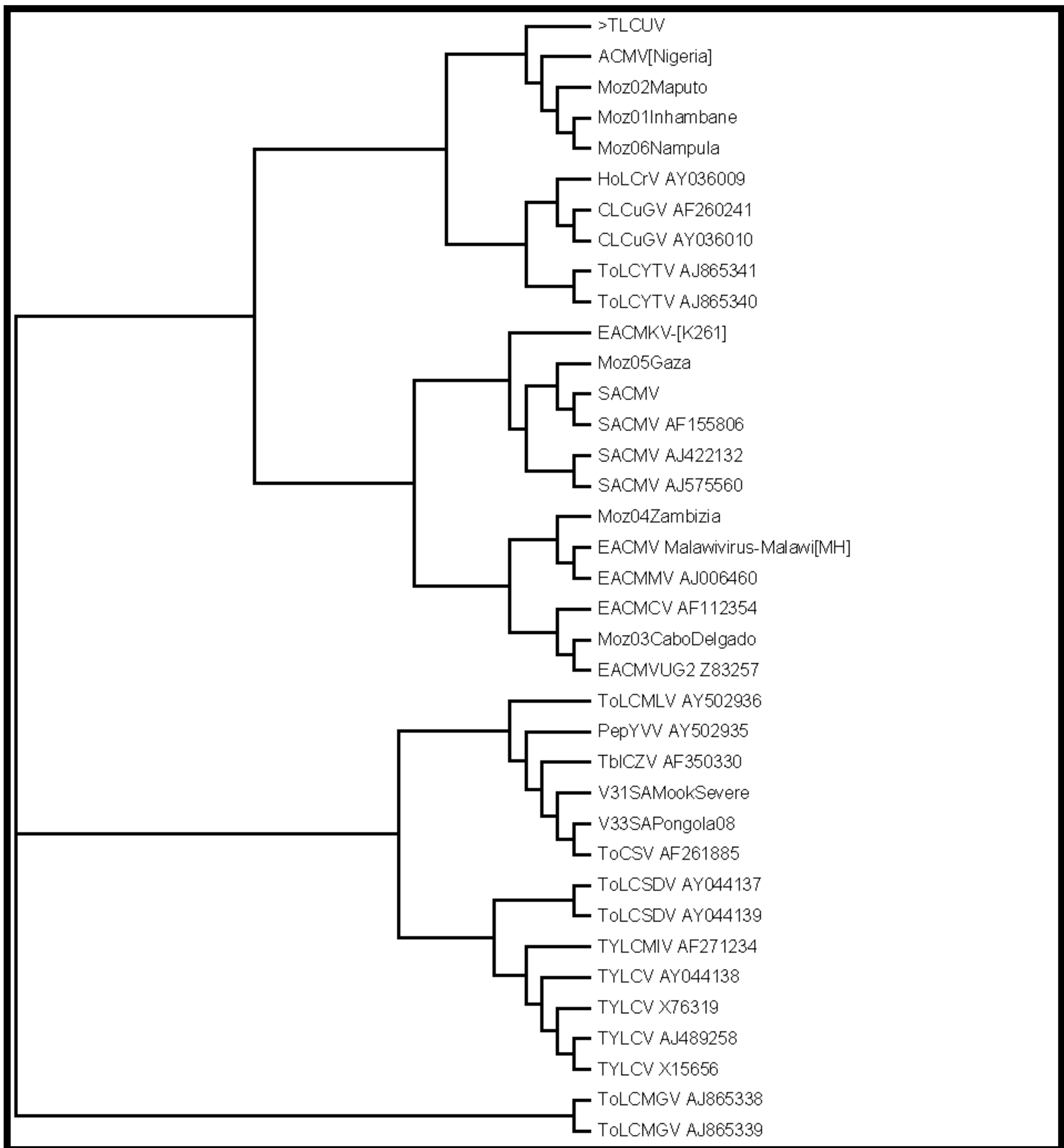


**Figure 3.9**: Maximum likelihood tree constructed in PAUP, showing the relationship between Mozambique cassava isolates and cassava geminivirus species (ACMV and EACMV) and selected geminiviruses from the African continent. Accession numbers and virus abbreviations are listed in Appendix 3. *Tomato leaf curl Madagascar virus* was used to root the tree. Bootstrap values are generated from 1000 replicates.

## *3.5 Discussion*

Six geminivirus species, and many strains and variants of these species, have been reported from cassava in Africa (Legg and Fauquet, 2004; Fauquet *et al*., 2007). In the past 15 years many studies have revealed the diversity of cassava geminiviruses, where once ACMV was thought to be the only virus species causing CMD.

Several epidemiological studies have been conducted in southern, east and central African countries (South Africa; Zimbabwe; Rwanda; Kenya; Malawi; Tanzania; Uganda; Madagascar) following the availability of molecular techniques such as PCR for virus identification, and the pandemic in Uganda which revealed the first evidence for recombination among geminiviruses (Zhou *et al*., 1997). Studies have revealed the enormous diversity of geminiviruses, including those of cassava, and have led to speculations as to the rapid emergence and diversification of this group of DNA viruses, whose strategies for evolving would need to differ from the abundant RNA genome plant-infecting viruses (Rojas *et al*., 2005).

In this study, CCP primers were successful in confirming that cassava mosaic symptoms in the field were due to geminivirus infection. Degenerate CCP primers are a reliable preliminary method for screening of large numbers of samples, and have been used extensively to detect Old World and New World begomoviruses (Rojas *et al*., 1993; Deng *et al*., 1994; Wyatt and Brown, 1996).

The CCP primers were tested on 109 samples and 60 samples tested positive. Higher leaf sample numbers would have amplified the CCP, but poor sampling storage methods in the remote areas in many of the provinces led to material deterioration, disappointingly even those on FTA cards (Whatman) (discussed in results). The CCP PCR was only carried out to confirm symptoms, which are usually reliable and distinct in the case of CMD. However, our epidemiological data (which was subjected to ANOVA) is only based on viral incidence and severity (plants actually showing distinct CMD symptom phenotypes), and does not attempt to distinguish between virsues. We cannot therefore make any statistical conclusions that there are significant differences in the virus species or isolate numbers in each province. We can therefore only comment on the general pattern/distribution and numbers out of the numbers of positive PCR samples (which were 55 in this study). This study is the first comprehensive molecular screening of CMGs in

Mozambique. Prior to this study, little information was available on the occurrence and distribution of CMD and identification of CMGs in the country.

In a previous study by Berry and Rey (2001), 16 cassava samples were collected during 1998 – 2000 in the Mozambique provinces of Sofala, Manica, Maputo and Zambezia. In this survey, using differential PCR primers (Zhou *et al*., 1998), 63% of the samples revealed the presence of either ACMV or EACMV, with no occurrence of the severe Ugandan variant (UgV), now named *East African cassava mosaic virus*-Uganda, which caused a huge pandemic in 1996/1997 in Uganda. In this study, seven out of the 55 samples contained EACMV-UG in mixed infections with other CMGs.

In another survey conducted in Mozambique by Toko *et al*. (2003), the objective was to assess cassava and sweet potato pests, diseases and yields in Mozambique. Some cassava-infected leaves were collected from individual fields for virus identification and about 80.5% of them were infected with EACMV in the six provinces surveyed. There were eleven fields infected with mixed virus species, namely ACMV and EACMV, while seven fields were infected with EACMV- [UG2] and six fields were infected with ACMV alone. This study screened a much larger number of samples (55) and districts in a systematic manner in six provinces, and found EACMV species to be predominant (64%).

The enzymes *Dra*I, *Mlu*I and *EcoR*V used in the RFLPs can distinguish between EACMV-UG and ACMV species patterns (figures 3.4 – 3.7), but are not be sufficient to distinguish SACMV, EACMMV, EACMCV and EACMV-TZ or detect all sequence variations. Sequencing of full-length DNA-A, however did reveal that genetic variants (94% nt sequence divergence) of EACMMV and EACMCV were present in Mozambique. More intense sequencing of full-length DNA As would be required to ascertain the extent and distribution of these geminivirus species and their variants.

Mixed geminivirus infections are common, and often lead to recombination (Berry and Rey, 2001; Pita *et al*., 2001). The pandemic of CMD in Uganda in the mid- to late 1990s is a good example of synergism and recombination playing a critical role in new and severe epidemics (Harrison *et al*., 1997; Pita *et al*., 2001).

In a study by Mabasa (2006) in South Africa between 2004 and 2006, 14.8% of cassava infections were found by RFLP to be mixtures of ACMV and EACMV. In this survey, 12.7% of infections were detected using RFLP. In another study by Sserubombe *et al*.

(2008), six RFLP profiles of EACMV (EA1-EA6) and two of ACMV were detected with *DraI*, MluI and EcoRV in a study in Uganda.

In a similar study in Kenya, Bull *et al*. (2006), using identical RFLP methods as Sserubombe *et al*. (2008), identified EACMV-UG at high incidence in the west of the country at the border with Uganda and found CMD-associated begomovirus diversity to be higher in Kenya and Tanzania compared to Uganda. However, RFLP analysis is not as sensitive as differential primers (Zhou *et al*., 1998) in picking up mixed virus infections, and some of the samples cannot be identified (in this study 7/55 samples could be identified). In a study by Zhou *et al*. (1998) differential primers were successful in differentiating variation in DNA A among isolates of EACMV from Kenya, Malawi and Tanzania.

RFLPs may be able to detect nucleotide (nt) changes at the restriction sites (RE) targeted, but may not pick up any potential recombination domains or other nt substitutions in areas outside of the RE sites. In this study for example, SACMV was not detected by RFLP in any of the samples tested. However, sequencing of DNA A of one isolate from Gaza showed the presence of SACMV. Single nucleotide polymorphisms do occur in geminivirus species populations, and therefore SACMV, and other virsues, may be present in higher numbers in mixed infections in reality. This is not critical for a general survey where patterns and general observation are made, but it would be necessary to sample and sequence a much greater number of viruses if more accurate and significant numerical comparisons were to be performed.

While RFLPs are reliable in initial screening of large sample numbers and are able to differentiate between the two characterized major species ACMV and EACMV, other methods such as differential primers and amplification of full-length DNA A by universal primers (Briddon and Markham, 1994) and sequencing, may be more reliable in characterization of new genetic variants and recombination events.

Rolling circle amplification (RCA) has recently proved to be useful in the diagnostic and genomic studies of geminviruses (Haible *et al*., 2006). In addition to amplifying circular full-length DNAs of geminiviruses from plants, it also can amplify any circular satellites or defective DNA molecules often found complexed with geminiviruses (Ndunguru et. al., 2006). In this study, RCA was used, where PCR from TNA failed, to amplify circular

geminiviral DNA concatemers, which were then used as templates for near full-length DNA A amplification by universal primers (Briddon and Markham, 1994).

It would appear that bacteriophage phi 29 DNA polymerase was far more robust and less sensitive than long-template Taq polymerase (Accuzyme) in our cassava TNA samples, where some degradation had taken place. A limitation of this study was that at least half of the 285 infected leaf samples degraded because of lack of cold conservation facilities during collection, even those on the FTA® cards.

In many of the remote districts there was no electricity or ice for the cooler box. Even though FTA® Classic Cards (Whatman) have been shown to preserve DNA integrity up to six months, or longer, in our hands DNA extractions and PCR from plant leaf material squashes on the cards was not successful. Reasons were discussed in the results section.

RCA from DNA extracts from dried plant tissue has been successfully carried out on a range of plants including *Euphorbia heterophylla*, which belongs to the same family as cassava (Shepherd *et al*., 2008). This finding would support our observation that phi 29 DNA polymerase was able to perform RCA on partially degraded leaf samples. It would be useful to determine in future studies whether dried cassava leaf material would be suitable for RCA. Cassava is a difficult plant to extract nucleic acids due to polyphenols, latex and other carbohydrates.

The results of this study demonstrate, for the first time, the diversity of cassava begomoviruses in six provinces in Mozambique, where cassava is intensively grown. As in other countries such as South Africa, Kenya, Tanzania and Uganda, ACMV species and EACMV variants occur, both singly, and in mixed infections, with EACMV/ KE AND UG being most predominant (64%). RFLP patterns could not differentiate between ACMV –[NG], ACMV –[CM] and ACMV-[CI] variants, but sequencing demonstrated that ACMV-[NG] was present in Maputo.

ACMV-[UG] was also detected by RFLP. It may be expected that ACMV-[UG] would be found in Mozambique, as Tanzania and Uganda are directly north of the country and provide routes for cassava root material trade. However it is interesting to note the presence of ACMV-[NG] from Nigeria. A number of EACMV variants were detected using RFLP and DNA A sequencing, namely EACMV-KE and EACMV-UG, and two other viral species, *East African cassava mosaic Malawi vírus* (EACM-[MV]) in Zambezia

Province and *East African cassava mosaic Cameroon* virus (EACMCV) in Cabo Delgado, were identified.

SACMV-South Africa was reported also for the first time in Mozambique, in the Gaza Province. Such high diversity of begomovirsues has implications for disease control strategies and poses further threats in that recombination may occur more frequently. Recombination was not tested in this study, and larger number of samples need to be collected and sequenced in future studies, to screen for recombination events, and to accurately establish quantitative data on the identities of cassava geminiviruses in Mozambique.

# SUMMARY AND CONCLUSIONS

Cassava mosaic disease (CMD) caused by cassava mosaic geminivirus CMGs) is undoubtedly the most important constraint to the production of cassava in Africa at the outset of the 21st century. Although the disease was recorded for the first time in the later part of the 19th century, for much of the intervening period, it was relatively benign in most areas, and was considered to be of minor economic significance. Towards the end of 20th century, however, the dynamics of the epidemic changed, with the discovery that the cassava begomovirsues were evolving, and genetic diversity was demonstrated.

Furthermore, recombination between various cassava begomovirsues was demonstrated, such as as the recombinant hybrid of the two principal species (ACMV and ECAMV) that was identified, initially in Uganda, but has now been shown to be associated with an unusually severe and rapidly spreading epidemic of CMD.

The movement of infected germplasm throughout sub-Saharan Africa has led to the spread of all the seven major cassava begomovirus species, including to Mozambique. This has led to mixed infections and recombinations in cassava and alternate hosts, thereby exacerbating the disease. In the past, ACMV was restricted to the west of Africa and EACMV mainly to the east, but both these virus species, and their strains, are now found across southern and eastern Africa.

Subsequent spread throughout East and Central Africa has led to the consequent devastation of production of the cassava crop, a staple food in this region, and to widespread hunger. The observation of similar virus recombination events elsewhere, highlights the inherent danger posed to man by the capacity of these viruses to adapt to their environment and optimally exploit their relationships with the whitefly vector, plant host and human cultivator.

In Mozambique, where cassava is widely grown as a subsistence and food security crop, CMD, as well as Cassava brown streak disease (CBSD), threaten the livelihoods of many people. The Mozambique climate is ideal for cassava cultivation, and recently the interest in cassava starch for biofuels has intensified the interest in commercialization of the crop. Therefore it is essential that the incidence and severity, and genetic diversity of cassava begomoviruses, be established so that control measures can be implemented to protect the crop cultivation. This study undertook to study the epidemiology of CMD and to

undertake molecular studies to establish the geminivirus(es) identities. Furthermore, the study aimed to undertake a study of the vector, *Bemisia tabaci*, of CMGs as this has had not been achieved in previous surveys. It was important to observe whether there had any changes or shifts in the CMG populations since the last surveys in 1999-2001, although only some provinces were surveyed previously and the identity of the viruses were never established.

The epidemiological study showed that CMD incidences recorded were highly variable within districts and between provinces in both surveys., In the 2005 survey, CMD incidence ranged from 22% in Cabo Delgado to 76% in Gaza Province, while in 2006 the incidence was 6% in Nampula and 75% in Gaza. In both surveys there were significant differences ($P \leq 0.005$) between some of the provinces in CMD. However CMD incidence did not appear to change noticeably over the two years (2005 and 2006).

Infections caused by whitefly were very low (0 to 3%), compared to cutting infection (22 to 60%) in the two surveys. Cutting infection was significantly higher compared to whitefly transmission in the six provinces in the two surveys. Zambezia had the highest (15.8) mean whitefly numbers per province in the two surveys compared to Maputo and Inhambane, which had the lowest mean whitefly numbers (0.1) recorded.

The mean CMD severity score ranged from 2.0 to 3.3 in all provinces in the two surveys, which demonstrates moderate symptoms, compared to higher severity (scores >4) in some parts of Tanzania and Uganda. Results indicated that cutting infection was more important than whitefly infection, because normally the farmers plant infected cuttings repeatedly over years and this is the main source of spreading the disease. A rigorous culling of infected cassava stakes or cuttings and the planting of healthy material could reduce CMD incidence in Mozambique. There is an urgent need to identify and assess the susceptibility to CMD of varieties preferred by farmers and multiply and distribute them to farmers. Improvement in CMD situation was observed during these two surveys, were the incidence in Nampula province reduced from 56% in 2005 to 7%. This could be a result of a huge program at the IIAM station with the collaboration of NGOS to identify, multiply and distribute virus-free (CMD/CBSD) cuttings to farmers.

Restriction Length Fragment Polymorphism (RLFP) analysis was undertaken on fifty five infected cassava samples for the six provinces in order to establish virus diversity. The

results showed that 63.4% (35 of the 55 cassava leaf samples) were EACMV, 12.7% (6/55) were ACMV species; 12.7 (7/55) were mixed infections of EACMV and ACMV; and six samples were not identified by RFLP. Six full-length DNA A clones, one from each province, showing unusual RFLP patterns, were cloned and sequenced.

Consensus sequences were aligned with other cassava begomoviruses and selected begomoviruses from southern Africa, and phylogenetic analysis (parsimony) revealed that virus isolates from the Maputo, Inhambane and Nampula Provinces exhibited 95-97% nucleotide sequence divergence/similarity to *African cassava mosaic virus*-[Nigeria]; the virus isolate from Gaza Province was 99% similar to *South African cassava mosaic virus* -[South Africa]; while the Zambezia Province virus was most closely aligned (94%) with EACM-[MV] (*East African cassava mosaic Malawi virus*- [Malawi: MH]), and the isolate from Cabo Delgado aligned most closely (96%) with *East African cassava mosaic Cameroon virus*- Cameroon and less closely (87%) to EACM-[MV].

For the first time, this study reports diversity of cassava begomovirus species in Mozambique, similar to previous studies in South Africa, Tanzania and Uganda, and demonstrates the mixture of geminivirus species from east and west Africa. Sequence variation of > 92% indicates that the cassava geminiviruses in Mozambique are genetic variants of ACMV-[NG], EACMV-[MW] and EACMCV. SACMV was reported for the first time in Mozambique, which is not suprising as SACMV, although first discovered in South Africa, has now been found in Madagascar and Zimbabwe. Further studies are likely to show that SACMV may be more widely found in eastern Africa, than is currently known.

A molecular study based on the mtCOI gene from *B. tabaci*, showed that this whitefly vector from Mozambique was closely related to the South African haplotypes, and formed a unified group within the Eastern/southern African clade. These results were not suprising based on previous *B. tabaci* studies, which have shown a phylogeographic origin. South Africa and Mozambique border on each other and are close to the countries north and east of them (Zimbabwe, Tanzania and Uganda).

CMD occurred in all the areas surveyed, with various levels of incidence and with some provinces such as Gaza having a disease incidence as high as 76%. The CMD symptoms seen were mild to moderate with effects on leaf size and overall plant vigour also observed.

Moreover, there was little evidence of current season spread by whitefly vector. This suggests that much of the disease recorded was due to the use of infected cuttings as planting material.

The existing pockets of high severity coupled with high incidence in some areas could be precursors of more serious attacks in years to come. In either situation CMD can be an avoidable problem if the use of healthy planting material is adopted routinely. However, due to the time it takes to micro-propagate and bulk up virus-free material, and to distribute it to the farmers, this process is prohibitive and farmers cannot often access clean planting material.

It is clear from experience in Uganda and elsewhere that there are big differences between cassava varieties in their reaction to CMD. Some varieties are severely affected and yield is reduced in early stages of growth. There is an urgent need to evaluate the main varieties being grown by farmers in the country. They should be categorized according to their main features, farmer preference, response to CMD and suitability within the cropping system. It will then be possible to advise farmers on the risk of losses due to the disease. In the long term the planting of resistant CMD and CBSD varieties is needed as mentioned before CBSD is devastating in the north of Mozambique and now is probably spread to south (unpublished results).

The results of the DNA analysis revealed the existence of several strains of CMGs of which EACMV is the most common. EACMV-[UG] was found in 18 samples of the 55, and the presence of mixed infections, is an indicator that intervention stategies are required to reduce spread of CMD in the country. It is therefore important to determine the source of infection of the disease (whether by cutting or else) in order to predict how fast the disease would spread within and outside the affected provinces, develop quarantine measures to slow down/stop the spread of the disease to non-affected provinces, and ultimately screen varieties with levels of resistance or tolerance, for multiplication and distribution to farmers.

## 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.

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J. (1997) Gapped BLAST and PSI-BLAT: a new generation of protein database search programs. Nucleic Acid Research **25**, 3389-3402.

Avidov Z. (1956) Bionomics of the tobacco whitefly (*Bemisia tabaci* Genn.) in Israel. Ktavin **7**, 25-41.

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**, 311-32.

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.

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.

Bock K.R., Guthrie E.J. and Meredit 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 Harrison, B.D. (1985) *African cassava mosaic virus*. *CMI/AAB Descr. Pl. Viruses* No. 297 pp.6.

Basavaprabu L.P. and Dsagupta I. (2006) Defective interfering DNAs of plant viruses. *Plant Science*, **25**, pp 47-64.

Basavaprabhu L. Patil., Nitin Dutt ., Rob W. Briddon ., Simon E. Bull., Dirk Rothenstein., Basanta K. Borah., Indranil Dasgupta., John Stanley and Holger Jeske· (2007). Deletion and recombination events between the DNA-A and DNA-B components of Indian cassava infecting geminiviruses generate defective molecules in *Nicotiana benthamiana*.

Bird, J. (1957) A whitefly-transmitted mosaic of *Jatropha gossypifolia*. *Agric*. *Exp*. *Stn*. *Univ*. *PR,* v.22, p.1-35.

Basavaprabu L.P and Fauquet C. (2009) Cassava mosaic geminivirus:actual knowelege and perpectives. *Molecular Plant Pathology*10 (5), 685-701.

Brewer B.J and Fangman W.L (1988) A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell*, **55**, 637–643.

Briddon R.W.,Prescott A.G.,Lunness P.,Chamberlin L.C.L.,Markham P. (1993) Rapid production of full- length, infectious geminivirus clones by abutting primer PCR. Journal of Virology Methods, **43**,7-20.

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/].

Bottcher B.,Unseld S.,Ceulemans H.,Russel R.B.,Jeske H. (2004) Geminate structures of *African cassava mosaic virus*. *Journal of Virology* **78**, 6758-6765.

Boykin L.M.,Shatters R.G., Rossell R.C., Mckenzie, C.L., Bagnall R.A., de Barro, P. & Frohlich, D.R. (2007) Global relationships of *Bemisia tabaci* (Hemiptera:Aleyrodidae) revealed using Bayesian analysis of mitocondrial COI DNA sequences. *Molecular Phylogenetics and Evolution* **44,** 1306-1319.

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. (1995b) 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. & Idris A.M. (2005) Genetic differentiation of whitefly (*Bemisia tabaci*) mitochondrial cytochrome oxidase I, and phylogeographic concordance with the coat protein of the plant virus genus begomovirus. *Annals of Entomology society of America* **95,** 827-883.

Bull R. A., Tu E.T., McIver, C. J., Rawlinson, W. D. & White, P. A. (2006) Emergence of a new norovirus genotype II.4 variant associated with global outbreaks of gastroenteritis. *J Clin Microbiol* **44**, 327–333.

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.

Byrne D. N. and Bellows, T.S.Jr. (1991). Whitefly biology. *Annual Review of Entomology* **36**, 431-458.

Byrne D.N. and Bellows T.S.Jr., Parella M.P., (1990a). Whiteflies in agricultural systems. In: Gerling, D. (Ed), Whiteflies: Their Bionomics, Pest Status and Management. Intercept, Andover, pp.227-261.

Cassava Strategy document,(2003), Document submitted do Ministry of Agriculture in Mozambique.

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.

Costa H.S., Brown J.K.., Sivasupramaniam S. & Bird J. (1993) Regional distribution, insecticide resistance and reciprocal crosses between the A and B biotypes of *Bemisia tabaci*. *Insect Science and its Application* **14,** 127-138.

Cossa N. and Mangana S. (2002). Relatório levantamento sobre o listrado castanho da mandioca na provincial de Cabo Delgado. (Unplublished Results).

Cossa N. and Jonasse F. (2002). Relatório levantamento sobre do mosaico africano da mandioca na provincial de Inhambane. (Unplublished Results).

De Barro P.J., Driver F., Trueman J.W.H. and Curran J. (2000) Phylogenetic relashinships 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.

Deng D., Otim-Nape G. W., Sangarem A., Ogwal S., Beachy R. N. and Fauquet C.M. (1997) Presence of a new virus closely related to East African cassava mosaic viru asssociated with the cassava mosaic outbreak in Uganda. *African Journal of Root and Tuber Crops* **2**, 23-28.

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.

Dry I.B., Rigden J.E., Krake L.R. Mullineaux P.M. and Rezaian M.A. (1993) Nucleotide sequence and genome organisation of tomato leaf curl geminivirus. *J. Gen. Virol.* **74**, 147-151.

Duffus, J.E. (1987) Whitefly transmission of plant viruses. In: HARRIS, K.F., ed. *Current topics in vector research.New York, Springer-Verlag*. v.4, pp.73-91.

Dubern J. (1979) Transmission of African cassava mosaic geminivirus by the whitefly (*Bemisia tabaci*). *Tropical Science* **34**, 82-91.

Eagle P.A., Orozco B.M. & Hanley-Bowdoin L. (1994) A DNA sequence required for geminivirus replication also mediates transcriptional regulation. *Plant Cell* **6**, 1157–1170

Jupin I., De Kouchkovsky F., Jouanneau F., and Gronenborn B (1994) Movement of *Tomato yellow leaf curl virus* (TYLCV): Involvement of the protein encoded by ORF C4. *Virology* **204**, 82–90.

FAO (2003) Cassava production statistics. http:/www.fao.org.

FAO (2005) Special Report FAO/WFP. Crop and food supply Assessment Mission to Mozambique, 20 June 2005.

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., Fargette, D. & Thouvenel, J.C. (1988) Some aspects of the epidemiology of African Cassava Mosaic in Ivory Coast. *Tropical Pest Management* **340**, 92-96.

Fauquet C.M. and Stanley J. (2003) Geminivirus classification and nomenclature: progress and problems. *Annals of Applied Biology.* **142**, 165–189.

Fauquet C.M., Briddon R.W., Brown J.K., Moriones, E., Stanley J., Zerbini M. and Zhou X. (2008) Geminivirus strain demarcation and nomenclature. *Arch Virol* **142**, 165-189.

Felsenstein J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**, 783-791.

Fishpool L.D.C. and C. Burban. (1994) *Bemisia tabaci*: The whitefly vector of African cassava mosaic geminivirus. *Trop. Sci*. **34**, 55-72.

Fishpool, L.D.C.; Fauquet C.; Fargette D., Thouvenel J.C., Burban C.; Colvin, J. (1995) The phenology of *Bemisia tabaci* (homoptera: aleyrodidae) populations on cassava in southern Côte D'ivoire. *Bulletin of Entomological Research*, v. **85**, p.197-207.

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 Africa cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *J. Gen. Virol*. **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.

Gerling D.(1986) Natural enemies of Bemisia tabaci, biological characteristics and potential as biological control agents: *A review. Agriculture, Ecosystems and Environment* **17**,99-110.

Greathead A.H.,(1986) Host Plants. In: *Bemisia tabaci* A Literature Survey on the Cotton Whitefly with an Annotated Bibliography, Cook, M.J.W. (Ed.), *CAB. International, London*, pp,17-26.

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.

Gilbertson R.L.; Faria J.C.; Ahlquist P.; Maxwell D.P. (1993) Genetic diversity in geminiviruses causing bean golden mosaic disease: the nucleotide sequence of the infectious cloned DNA components of a Brazilian isolate of bean golden mosaic geminivirus. *Phytopathology* **83**, 709-715.

Gill R.J. (1990)**.** The morphology of whiteflies: Their bionomics, pest status management. *Newcastle, Intercept*., **348**, 13-46

Haible D., Kober S. and Jeske H. (2006). Rolling circle amplification revolutionizes diagnosis and genomics of geminiviruses.*Journal of Virology Methods* **135**, 9-16.

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

Hanley-Bowdoin L., Settlage S.B., Orozoo B.M., Nagar S., and Robertson D. (1999) Geminiviruses: Models for Plant DNA Replication, Transcription, and Cell Cycle Regulation. *Crit. Rev. Plant Sci*. **18,** 71-106.

Hanley-Bowdoin, L., Settlage, S. B. & Robertson D. (2004) Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. *Molecular Plant Pathology* **5**,149–156.

Harmse J. (2007) Engineering virus resistant transgenic cassava: the design of long hairpin RNA constructs against *South African cassava mosaic virus*. MSc Dissertation.

Harrison S.F., Baker H., Guthrie E.J., Meredith G. and Atkinson M (1977). Plant viruses with circular single stranded DNA. *Nature.* **270,** 760-762.

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

Harrison C.J. and Langdale J.A. (2006) A step by step guide to phylogeny reconstruction. *The Plant Journal* **45**, 561-572.

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.

Hillocks R.J., Tresh J.M., Tomas J., Botao M., Macia R., Zavier R. (2002). Cassava brown streak in northen Mozambique. *International Journal of Pest Management* **48**, 179-182.

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.

Janssen J.A. M., W.F. Tjallingii, and J.C. Van Lenteren (1989) Electrical recording and ultrastructure of stylet penetration by the greenhouse whitefly. *Entomol. Exp. Appl* **52**, 69–81..

Jeske H., Lutgemeier M., and Preiss W. (2001) Distinct DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic geminivirus. *EMBO J.* **8**, 6158–6167.

Jupin I., De Kouchkovsky F., Jouanneau F., and Gronenborn B. (1994) Movement of *tomato yellow leaf curl* geminivirus (TYLCV): Involvement of the protein encoded by ORF C4. *Virology* **204**, 82–90.

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 transacts 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 and C.M. Fauquet, (2004) Cassava mosaic geminiviruses in Africa. *Plant Molecular Biology* **56**, 585–599.

Lopez-Avila, A. (1986) Taxonomy and biology. In: Bemisia tabaci - A Literature Survey on the Cotton Whitefly with an Annotated Bibliography, M.J.W. Cock (ed.). C.A.B. *International Institute of Biological Control, Silwood Park, United Kingdom*, pp. 3-11.

Mabasa K.G. (2007) Epidemiology of Cassava Mosaic Disease and Molecular Characterization of Cassava Mosaic Viruses and their associated Whitefly *Bemisia tabaci* vector in South Africa. Msc Thesis. University of the Witwatersrand.105 pp.

MacDowell S.W., R.H. Coutts and K.W. Buck. (1986) Molecular characterisation of subgenomic single-stranded and double-stranded DNA forms isolated from plants infected with tomato golden mosaic virus, *Nucleic Acids Res.* **14** pp. 7967–7984.

Mader, (2005) A report on cassava production in Mozambique,Ministry of Agriculture.

Maddison W.P., Donoghue M.J. and Maddison D.R. (1984) Out-group analysis and parsimony. *Systemic Zoology* **33**, 83-103.

Martin D.P., Van der Walt E., Posada D. and Rybicki E.P. (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* **21**, 260-262.

Mozambique Equator Initiative, (2002) Documento Publicado pelo Ministerio da Agricultura.

Mound L.A. Halsey S.H.Eds. (1978) Whitefly of the world:a systermatic catalogue of the Aleyrodidae (Homoptera) with host plant and natural enemie data.Britsh Museum (Natural History), London UK.Wiley Chchester.

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, pp. 305-313.

Munga.T. and J.M. Tresh, (2002) Incidence of cassava mosaic and cassava brown streak virus disease in coastal Kenya, Roots **8(1)**, 12-14.

Muniyapa V., (1980) Whiteflies, In: Harris, K.F., Maramorosch, K. (Eds.), Vectors of Plant Pathogens. *Academic Press, New York*, pp. 39-85.

Nei M., Saitou N. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

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., Traore P., and Jorge A. M. ( 2004). Viruses Associated with Cassava Mosaic Disease in Senegal and Guinea Conakry. *Journal of Phytopathology* **152**:2, 69–76.

Orlando M, (1958) Manual sobre a mandioca,pp 297-314,Editado em Portugal.

Otim-Nape G.W. (1993). The epidemiology of the African cassava mosaic geminivirus disease in Uganda. PhD Thesis. University of Redding 252 pp.

Otim-Nape G.W., Alicai T. and Thresh J.M. (2001). Changes in the incidence and severity of *Cassava mosaic virus disease*, varietal diversity and cassava production Uganda. *Annals of Applied Biology* **138**, 312-327.

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., 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.

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.

Pollard D.G. (1995) Feeding habitats of the cotton whitefly,Bemisia tabaci Genn.(Homoptera:Aleyrodidae). *Annals of Aplied Biology*,**43**, 664-671.

Polston, J. E., and Anderson, P. K. (1997) The emergence of whitefly transmitte geminiviruses in tomato in the Western Hemisphere. *Plant Disease.* **81,** 1358–1369.

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

Ranomenjanahary S., Rabindran R. and Robinson D. J. (2002) Occurrence of three distinct begomoviruses in cassava in Madagascar. *Annals Applied Biology* **140**, 315–318.

Rector A., Tachezy R. and Van Ranst M. (2004) A sequence- independent strategy for detection and cloning of circular DNA virus genomes by using multiply primed rolling-circle amplification. *Journal of Virology* 4993-4998.

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.

Rey M.E.C. and Thompson G. (1998) Cassava mosaic virus disease in South Africa. Roots **5**, 3-5.

Rojas M. R. (1993) Use of degenerate primers in the polymerase chain reaction to detect whitefly-transmitted geminiviruses. *Plant Disease*, **v. 77,** 340-347.

Rojas M.R., Hagen C., Lucas W.J., Gibertson R.L., (2005). Exploiting chinks in the plant's armor: evolution and emergenceof geminiviruses. *Annual Review of Phytopathology* **43**, 361-394.

Seal.S.E (2006) Factors influencing begomovirus evolution and their increasing global significance: Implications for sustainable control.

Segeren P. e Vanden O.R. (1993) Pragas, doencas e ervas daninhas nas raizes e tuberculos.Centro de Formacao Agraria e de Documentacao Rural.**55** pp.18-19.

Sequeira J.C. and Harrison, B.D (1982). Serological studies on cassava latent virus. *Annals of Applied Biology* **101**, 33-42.

Simon C. Frati F., Beckenbach.(1994). Evolution, weighting, and phylogenetic utitlity of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America,***87**, 651-701.

Silvestre P. & Arradeau M.(1983) Le manioc techniques agricoles et productions tropicales 23.editions maisonneuve+larose, pp 262.

Sserubombwe W.S.., R.W. Briddon.,Y.K. Baguma., G.N. Ssemakula., S.E. Bull.,A. Bua., T. Alicai., C. Omongo., G.W. Otim-Nape and J. Stanley. (2008) Diversity of begomoviruses associated with mosaic disease of cultivated cassava (Manihot esculenta Crantz) and its wild relative (Manihot glaziovii Mu¨ ll.Arg.) in Uganda. *Journal of General Virology,* **89**, 1759–1769

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. (1983) Infectivity of cloned geminivirus genome requires sequences from both DNA's. *Nature, London* **301**, 643-645.

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 and Fauquet C.M. (2003) Geminivirus classification and nomenclature: progress and problems. *Annals of Applied Biology* **142**,165–189.

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.

Storey H. and Nichols R.F.W. (1938) Studies of the mosaic disease of cassava. *Annals of Applied Biology* **25**, 790-806

Swanson M.M. and Harrison B.D. (1994) Properties, relationships and distribution of cassava mosaic geminiviruses. *Tropical Science* **34**, 15-25.

Swofford D.L. (2003). PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods), Version 4.0b. Sunderland, Massachusetts: Sinauer Associates.

Tamura K, Dudley J, Nei M & Kumar S. (2000). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596-1599. (Publication PDF at http://www.kumarlab.net/publications)]

Thomas J.E, Massalski P.R, Harrison BD. (1986) Production of monoclonal antibodies to African cassava mosaic virus and differences in their reactivities with other whitefly-transmitted gemini viruses. *Journal of General Virology* **67,** 2739–48.

Thresh J.M. and Mbwana M.W. (1998) Cassava mosaic and cassava brown streak virus diseases in Zanzibar. *Roots* **5**, 6-9.

Thresh J.M. (1999-2001) Reports on visits to Mozambique. National Resources Institute. http://www.cpp.uk.com/downloads/cbsd/btor1.pdf.

Thresh J.M., Otim-Nape G.W., Legg J.P. and Fargette D. (1997) African cassava mosaic virus disease: the magnitude of the problem. *African Journal of Root and Tuber Crops* **2**, 13-19.

Thresh J.M., Otim-Nape G.W., Thankappan M. and Muniyappa V. (1998) The mosaic disease of cassava in Africa and India caused by whitefly-borne geminiviruses. *Plant Pathology*. **77**, 935-945.

Thompson W.M.O. (2002) Comparison of *Bemisia tabaci* (Homoptera: Aleyrodidae) development on uninfected cassava plants and cassava plants infected with *East* 104 *African Mosaic Virus*. *Annals of the Entomological Society of America* **95**, 387-394.

Trench T.N. & Martin M.M. (1985) An assessment of African cassava mosaic disease in South Africa and Swaziland. *South African Journal of Plant and Soil* **2**, 169-170.

Toko M.., Andrade M., Danny C., Rachid H., Legg J.., Abu J., Elsa M., and Afonso S. (2003). Report on assessement of cassava and sweet potato pests and diseases in Mozambique. (Unpublished Results).

Unseld S., Ringel M., Höfer P., Höhnle M., Jeske H., Bedford I. D., Markham P. G. & Frischmuth, T. (2000) Host range and symptom variation of pseudorecombinant virus produced by two distinct bipartite geminiviruses. *Arch Virol* **145**, 1449–1454

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

Varma A. and Malathi,V. (2003) Emerging geminivirus problems. A serious threat to crop production. *Ann. Appl. Biol*. **142**, 145-164.

Viscarret M.M., I. Torres-Jerez, E. Agostini De Manero, S. N. López, E. E. Botto, and J.K. Brown (2003) Mitochondrial DNA evidence for a distinct New World group of *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae) indigenous to Argentina and Bolivia, and presence of the Old World B biotype in Argentina. *Ann. Entomol. Soc. Am.* **96**, 65–72.

Wyatt S.D. and Brown J.K. (1996) Detection of subgroup III geminivirus isolated in leaf extracts by degenerate primers and polymerase chain reaction. *Phytopathol.* **86**, 1288-1293.

Y. Xie and X. P. Zhou (2003) Molecular characterization of squash leaf curl Yunnan virus,a new begomovirus and evidence for recombination. *Arch Virology* **148,** 2047-2054.

Xiong J. (2006) Essential Bioinformatics. Cambridge University Press.

Zhang S.C., Wege, C. & Jeske H. (2001) Movement proteins (BC1 and BV1) of Abutilon mosaic geminivirus are cotransported in and between cells of sink but not of source leaves as detected by green fluorescent protein tagging. *Virology* **290**, 249–260.

Zhou X., Liu Y., Calvert L., Munoz D., Otim-Nape G. W., Robinson D. J. and Harrison B. D. (1997) Evidence that DNA-A of a geminivirus associated with severe cassava mosaic disease in Uganda has arisen by interspecific recombination. *Journal of General Virology* **78**, 2101-2111.

# Appendices

## *Appendix 1: Incidence, severity and whitefly numbers, and statistical (anova) data from all the 6 provinces surveyed.*

**Table A1: ANOVA of cutting incidence, whitefly incidence, and total incidence in the 2005 survey**

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| CINC | Between Groups | 23.363 | 5 | 4.673 | 45.859 | .000 |
| | Within Groups | 60.523 | 594 | .102 | | |
| | Total | 83.885 | 599 | | | |
| WINC | Between Groups | .099 | 5 | .020 | 13.323 | .000 |
| | Within Groups | .880 | 594 | .001 | | |
| | Total | .979 | 599 | | | |
| INC | Between Groups | 22.683 | 5 | 4.537 | 44.009 | .000 |
| | Within Groups | 61.232 | 594 | .103 | | |
| | Total | 83.915 | 599 | | | |

Value of 0.000 indicates overall significance difference between provinces; $p \leq 0.05$
Group = Province
CINC: cutting incidence; WINC: whitefly incidence
INC: total incidence (WINC + CINC)

**Table A2: Multiple comparisons of cutting incidence, whitefly incidence and total incidence between six provinces in the 2005 survey**

| Dependent Variable CINC | | | | |
|---|---|---|---|---|
| Province (I) | Province (J) | Mean Difference (I-J) | Std. Error | Sig. |
| Maputo | Gaza | -0.14* | 0.05 | 0.00 |
| | Inhambane | 0.38* | 0.05 | 0.00 |
| | Zambezia | 0.24* | 0.05 | 0.00 |
| | Nampula | 0.04 | 0.05 | 0.44 |
| | C. Delgado | 0.38* | 0.04 | 0.00 |
| Gaza | Maputo | 0.14* | 0.05 | 0.00 |
| | Inhambane | 0.52* | 0.05 | 0.00 |
| | Zambezia | 0.38* | 0.05 | 0.00 |
| | Nampula | 0.18* | 0.05 | 0.00 |
| | C. Delgado | 0.52* | 0.04 | 0.00 |
| Inhambane | Maputo | -0.38* | 0.05 | 0.00 |
| | Gaza | -0.52* | 0.05 | 0.00 |
| | Zambezia | -0.14* | 0.05 | 0.00 |
| | Nampula | -0.34* | 0.05 | 0.00 |
| | C.Delgado | 0.00 | 0.04 | 0.99 |
| Zambezia | Maputo | -0.24* | 0.05 | 0.00 |
| | Gaza | -0.38* | 0.05 | 0.00 |
| | Inhambane | 0.14* | 0.05 | 0.00 |
| | Nampula | -0.20* | 0.05 | 0.00 |
| | C. Delgado | 0.14* | 0.04 | 0.00 |
| Nampula | Maputo | -0.04 | 0.05 | 0.44 |
| | Gaza | -0.18* | 0.05 | 0.00 |
| | Inhambane | 0.34* | 0.05 | 0.00 |
| | Zambezia | 0.20* | 0.05 | 0.00 |
| | C. Delgado | 0.34* | 0.05 | 0.00 |
| C.Delgado | Maputo | -0.38* | 0.04 | 0.00 |
| | Gaza | -0.52* | 0.04 | 0.00 |
| | Inhambane | 0.00 | 0.04 | 0.99 |
| | Zambezia | -0.14* | 0.04 | 0.00 |
| | Nampula | -0.34* | 0.05 | 0.00 |

| Dependent Variable WINC | | | | |
|---|---|---|---|---|
| Province (I) | Province (J) | Mean Difference (I-J) | Std. Error | Sig. |
| Maputo | Gaza | -0.02* | 0.01 | 0.00 |
| | Inhambane | -0.03* | 0.01 | 0.00 |
| | Zambezia | 0.00 | 0.01 | 0.37 |
| | Nampula | 0.00 | 0.01 | 0.96 |
| | C.Delgado | 0.00 | 0.01 | 0.83 |
| Gaza | Maputo | 0.02* | 0.01 | 0.00 |
| | Inhambane | -0.01* | 0.01 | 0.01 |
| | Zambezia | 0.02* | 0.01 | 0.00 |
| | Nampula | 0.02* | 0.01 | 0.00 |
| | C. Delgado | 0.02* | 0.01 | 0.00 |
| Inhambane | Maputo | 0.03* | 0.01 | 0.00 |
| | Gaza | 0.01* | 0.01 | 0.01 |
| | Zambezia | 0.03* | 0.01 | 0.00 |
| | Nampula | 0.03* | 0.01 | 0.00 |
| | C.Delgado | 0.03* | 0.01 | 0.00 |
| Zambezia | Maputo | 0.00 | 0.01 | 0.37 |
| | Gaza | -0.02* | 0.01 | 0.00 |
| | Inhambane | -0.03* | 0.01 | 0.00 |
| | Nampula | 0.01 | 0.01 | 0.35 |
| | C.Delgado | 0.00 | 0.01 | 0.48 |
| Nampula | Maputo | 0.00 | 0.01 | 0.96 |
| | Gaza | -0.02* | 0.01 | 0.00 |
| | Inhambane | -0.03* | 0.01 | 0.00 |
| | Zambezia | -0.01 | 0.01 | 0.35 |
| | C.Delgado | 0.00 | 0.01 | 0.79 |
| C.Delgado | Maputo | 0.00 | 0.01 | 0.83 |
| | Gaza | -0.02* | 0.01 | 0.00 |
| | Inhambane | -0.03* | 0.01 | 0.00 |
| | Zambezia | 0.00 | 0.01 | 0.48 |
| | Nampula | 0.00 | 0.01 | 0.79 |

| Dependent Variable INC | | | | |
|---|---|---|---|---|
| Province (I) | Province (J) | Mean Difference (I-J) | Std. Error | Sig. |
| Maputo | Gaza | -0.16* | 0.05 | 0.00 |
| | Inhambane | 0.34* | 0.05 | 0.00 |
| | Zambezia | 0.24* | 0.05 | 0.00 |
| | Nampula | 0.04 | 0.05 | 0.40 |
| | C.Delgado | 0.38* | 0.04 | 0.00 |
| Gaza | Maputo | 0.16* | 0.05 | 0.00 |
| | Inhambane | 0.50* | 0.05 | 0.00 |
| | Zambezia | 0.40* | 0.05 | 0.00 |
| | Nampula | 0.20* | 0.05 | 0.00 |
| | C.Delgado | 0.54* | 0.04 | 0.00 |
| Inhambane | Maputo | -0.34* | 0.05 | 0.00 |
| | Gaza | -0.50* | 0.05 | 0.00 |
| | Zambezia | -0.10* | 0.05 | 0.04 |
| | Nampula | -0.30* | 0.05 | 0.00 |
| | C.Delgado | 0.04* | 0.04 | 0.35 |
| Zambezia | Maputo | -0.24* | 0.05 | 0.00 |
| | Gaza | -0.40* | 0.05 | 0.00 |
| | Inhambane | 0.10* | 0.05 | 0.04 |
| | Nampula | -0.20* | 0.05 | 0.00 |
| | C.Delgado | 0.14* | 0.04 | 0.00 |
| Nampula | Maputo | -0.04 | 0.05 | 0.40 |
| | Gaza | -0.20* | 0.05 | 0.00 |
| | Inhambane | 0.30* | 0.05 | 0.00 |
| | Zambezia | 0.20* | 0.05 | 0.00 |
| | C.Delgado | 0.34* | 0.05 | 0.00 |
| C.Delgado | Maputo | -0.38* | 0.04 | 0.00 |
| | Gaza | -0.54* | 0.04 | 0.00 |
| | Inhambane | -0.04 | 0.04 | 0.35 |
| | Zambezia | -0.14* | 0.04 | 0.00 |
| | Nampula | -0.34* | 0.05 | 0.00 |

* The mean difference is significant at $p \leq 0.05$
CINC: cutting incidence; WINC: whitefly incidence
INC: total incidence (WINC + CINC)

**Table A3: ANOVA of cutting incidence, whitefly incidence and total incidence in the 2006** survey

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| CINC | Between Groups | 26.478 | 5 | 5.296 | 67.513 | .000 |
| | Within Groups | 46.592 | 594 | .078 | | |
| | Total | 73.071 | 599 | | | |
| WINC | Between Groups | .108 | 5 | .022 | 17.784 | .000 |
| | Within Groups | .722 | 594 | .001 | | |
| | Total | .830 | 599 | | | |
| INC | Between Groups | 28.737 | 5 | 5.747 | 71.919 | .000 |
| | Within Groups | 47.469 | 594 | .080 | | |
| | Total | 76.205 | 599 | | | |

Value of 0.000 indicates overall significance difference between provinces; $p \leq 0.05$
Group = Province
CINC: cutting incidence; WINC: whitefly incidence
INC: total incidence (WINC + CINC)

**Table A4: Multiple comparisons of cutting incidence, whitefly incidence and total CMD incidence between six provinces for the 2006 survey**

| Dependent Variable CINC | | | | |
|---|---|---|---|---|
| Province (I) | Province (J) | Mean Diff. (I-J) | Std. Error | Sig. |
| Maputo | Gaza | -0.21* | 0.04 | 0.00 |
| | Inhambane | 0.26* | 0.04 | 0.00 |
| | Zambezia | 0.25* | 0.04 | 0.00 |
| | Nampula | 0.45* | 0.04 | 0.00 |
| | C.Delgado | 0.21* | 0.04 | 0.00 |
| Gaza | Maputo | 0.21* | 0.04 | 0.00 |
| | Inhambane | 0.47* | 0.04 | 0.00 |
| | Zambezia | 0.46* | 0.04 | 0.00 |
| | Nampula | 0.66* | 0.04 | 0.00 |
| | C.Delgado | 0.42* | 0.04 | 0.00 |
| Inhambane | Maputo | -0.26* | 0.04 | 0.00 |
| | Gaza | -0.47* | 0.04 | 0.00 |
| | Zambezia | -0.01 | 0.04 | 0.84 |
| | Nampula | 0.19* | 0.04 | 0.00 |
| | C.Delgado | -0.05 | 0.04 | 0.18 |
| Zambezia | Maputo | -0.25* | 0.04 | 0.00 |
| | Gaza | -0.46* | 0.04 | 0.00 |
| | Inhambane | 0.01 | 0.04 | 0.84 |
| | Nampula | 0.20* | 0.04 | 0.00 |
| | C.Delgado | -0.04 | 0.04 | 0.25 |
| Nampula | Maputo | -0.45* | 0.04 | 0.00 |
| | Gaza | -0.66* | 0.04 | 0.00 |
| | Inhambane | -0.19* | 0.04 | 0.00 |
| | Zambezia | -0.20* | 0.04 | 0.00 |
| | C.Delgado | -0.24* | 0.04 | 0.00 |
| C.Delgado | Maputo | -0.21* | 0.04 | 0.00 |
| | Gaza | -0.42* | 0.04 | 0.00 |
| | Inhambane | 0.05 | 0.04 | 0.18 |
| | Zambezia | 0.04 | 0.04 | 0.25 |
| | Nampula | 0.24* | 0.04 | 0.00 |

| Dependent Variable WINC | | Mean Diff.(I-J) | Std. Error | Sig. |
|---|---|---|---|---|
| Province (I) | Province (J) | | | |
| Maputo | | 0.00 | 0.00 | 0.61 |
| | Inhambane | 0.01 | 0.00 | 0.13 |
| | Zambezia | 0.03* | 0.00 | 0.00 |
| | Nampula | 0.03* | 0.00 | 0.00 |
| | C.Delgado | 0.03* | 0.00 | 0.00 |
| Gaza | Maputo | 0.00 | 0.00 | 0.61 |
| | Inhambane | 0.01* | 0.00 | 0.04 |
| | Zambezia | 0.03* | 0.00 | 0.00 |
| | Nampula | 0.03* | 0.00 | 0.00 |
| | C.Delgado | 0.03* | 0.00 | 0.00 |
| Inhambane | Maputo | -0.01 | 0.00 | 0.13 |
| | Gaza | -0.01* | 0.00 | 0.04 |
| | Zambezia | 0.02* | 0.00 | 0.00 |
| | Nampula | 0.02* | 0.00 | 0.00 |
| | C.Delgado | 0.02* | 0.00 | 0.00 |
| Zambezia | Maputo | -0.03* | 0.00 | 0.00 |
| | Gaza | -0.03* | 0.00 | 0.00 |
| | Inhambane | -0.02* | 0.00 | 0.00 |
| | Nampula | 0.00 | 0.00 | 0.95 |
| | C.Delgado | 0.00 | 0.00 | 0.49 |
| Nampula | Maputo | -0.038* | 0.00 | 0.00 |
| | Gaza | -0.03* | 0.00 | 0.00 |
| | Inhambane | -0.02* | 0.00 | 0.00 |
| | Zambezia | 0.00 | 0.00 | 0.95 |
| | C.Delgado | 0.00 | 0.00 | 0.54 |
| C.Delgado | Maputo | -0.03* | 0.00 | 0.00 |
| | Gaza | -0.03* | 0.00 | 0.00 |
| | Inhambane | -0.02* | 0.00 | 0.00 |
| | Zambezia | 0.00 | 0.00 | 0.49 |
| | Nampula | 0.00 | 0.00 | 0.54 |

| Dependent Variable INC | | | | |
|---|---|---|---|---|
| **Province (I)** | **Province (J)** | **Mean Diff.(I-J)** | **Std. Error** | **Sig.** |
| Maputo | Gaza | -0.21* | 0.04 | 0.00 |
| | Inhambane | 0.27* | 0.04 | 0.00 |
| | Zambezia | 0.28* | 0.04 | 0.00 |
| | Nampula | 0.48* | 0.04 | 0.00 |
| | C.Delgado | 0.23* | 0.04 | 0.00 |
| Gaza | Maputo | 0.21* | 0.04 | 0.00 |
| | Inhambane | 0.47* | 0.04 | 0.00 |
| | Zambezia | 0.49* | 0.04 | 0.00 |
| | Nampula | 0.69* | 0.04 | 0.00 |
| | C.Delgado | 0.44* | 0.04 | 0.00 |
| Inhambane | Maputo | -0.27* | 0.04 | 0.00 |
| | Gaza | -0.47* | 0.04 | 0.00 |
| | Zambezia | 0.01 | 0.04 | 0.75 |
| | Nampula | 0.21* | 0.04 | 0.00 |
| | C.Delgado | -0.04 | 0.04 | 0.37 |
| Zambezia | Maputo | -0.28* | 0.04 | 0.00 |
| | Gaza | -0.49* | 0.04 | 0.00 |
| | Inhambane | -0.01 | 0.04 | 0.75 |
| | Nampula | 0.20* | 0.04 | 0.00 |
| | C.Delgado | -0.05 | 0.04 | 0.22 |
| Nampula | Maputo | -0.48* | 0.04 | 0.00 |
| | Gaza | -0.69* | 0.04 | 0.00 |
| | Inhambane | -0.21* | 0.04 | 0.00 |
| | Zambezia | -0.20* | 0.04 | 0.00 |
| | C.Delgado | -0.25* | 0.04 | 0.00 |
| C.Delgado | Maputo | -0.23* | 0.04 | 0.00 |
| | Gaza | -0.44* | 0.04 | 0.00 |
| | Inhambane | 0.04 | 0.04 | 0.37 |
| | Zambezia | 0.05 | 0.04 | 0.22 |
| | Nampula | 0.25* | 0.04 | 0.00 |

*The mean difference is significant at p≤ 0.05

**Table A5: ANOVA of CMD severity and WF abundance in the 2005 survey**

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| CMDSEV | Between Groups | 956.048 | 5 | 191.210 | 66.883 | .000 |
| | Within Groups | 21967.678 | 7684 | 2.859 | | |
| | Total | 22923.725 | 7689 | | | |
| WFNO | Between Groups | 451543.8 | 5 | 90308.762 | 1340.616 | .000 |
| | Within Groups | 1211737 | 17988 | 67.364 | | |
| | Total | 1663281 | 17993 | | | |

CMDSEV: Severity of disease based on score scale of 1-5
Group = Province
WFNO=Whitefly numbers
Sig= actual P value □ 0.05

**Table A6: Multiple comparisons of CMD severity and whitefly abundance for the 6 provinces in the 2005 survey**

| Dependent Variable CMDSEV | | | | |
|---|---|---|---|---|
| **Province (I)** | **Province (J)** | **Mean Diff.(I-J)** | **Std. Error** | **Sig.** |
| Maputo | Gaza | 0.78* | 0.05 | 0.00 |
| | Inhambane | 0.10 | 0.07 | 0.17 |
| | Zambezia | 0.39* | 0.07 | 0.00 |
| | Nampula | 1.18* | 0.11 | 0.00 |
| | C. Delgado | 0.32* | 0.07 | 0.00 |
| Gaza | Maputo | -0.78* | 0.05 | 0.00 |
| | Inhambane | -0.68* | 0.07 | 0.00 |
| | Zambezia | -0.39* | 0.07 | 0.00 |
| | Nampula | 0.40* | 0.10 | 0.00 |
| | C. Delgado | -0.45* | 0.06 | 0.00 |
| Inhambane | Maputo | -0.10 | 0.07 | 0.17 |
| | Gaza | 0.68* | 0.07 | 0.00 |
| | Zambezia | 0.29* | 0.08 | 0.00 |
| | Nampula | 1.08* | 0.11 | 0.00 |
| | C.Delgado | 0.23* | 0.08 | 0.00 |
| Zambezia | Maputo | -0.39* | 0.07 | 0.00 |
| | Gaza | 0.39* | 0.07 | 0.00 |
| | Inhambane | -0.29* | 0.08 | 0.00 |
| | Nampula | 0.79* | 0.11 | 0.00 |
| | C.Delgado | -0.07 | 0.08 | 0.41 |
| Nampula | Maputo | -1.18* | 0.11 | 0.00 |
| | Gaza | -0.40* | 0.10 | 0.00 |
| | Inhambane | -1.08* | 0.11 | 0.00 |
| | Zambezia | -0.79* | 0.11 | 0.00 |
| | C. Delgado | -0.85* | 0.11 | 0.00 |
| C. Delgado | Maputo | -0.32* | 0.07 | 0.00 |
| | Gaza | 0.45* | 0.06 | 0.00 |
| | Inhambane | -0.23* | 0.08 | 0.00 |
| | Zambezia | 0.07 | 0.08 | 0.41 |
| | Nampula | 0.85* | 0.11 | 0.00 |

| Dependent Variable WFNO | | | | |
| --- | --- | --- | --- | --- |
| **Province (I)** | **Province (J)** | **Mean Diff.(I-J)** | **Std. Error** | **Sig.** |
| Maputo | Gaza | -1.51* | 0.21 | 0.00 |
| | Inhambane | 0.84* | 0.21 | 0.00 |
| | Zambezia | -14.18* | 0.21 | 0.00 |
| | Nampula | -2.41* | 0.21 | 0.00 |
| | C.Delgado | -2.97* | 0.21 | 0.00 |
| Gaza | Maputo | 1.51* | 0.21 | 0.00 |
| | Inhambane | 2.35* | 0.21 | 0.00 |
| | Zambezia | -12.67* | 0.21 | 0.00 |
| | Nampula | -0.90* | 0.21 | 0.00 |
| | C.Delgado | -1.46* | 0.21 | 0.00 |
| Inhambane | Maputo | -0.84* | 0.21 | 0.00 |
| | Gaza | -2.35* | 0.21 | 0.00 |
| | Zambezia | -15.02* | 0.21 | 0.00 |
| | Nampula | -3.25* | 0.21 | 0.00 |
| | C.Delgado | -3.82* | 0.21 | 0.00 |
| Zambezia | Maputo | 14.18* | 0.21 | 0.00 |
| | Gaza | 12.67* | 0.21 | 0.00 |
| | Inhambane | 15.02* | 0.21 | 0.00 |
| | Nampula | 11.77* | 0.21 | 0.00 |
| | C.Delgado | 11.21* | 0.21 | 0.00 |
| Nampula | Maputo | 2.41* | 0.21 | 0.00 |
| | Gaza | 0.90* | 0.21 | 0.00 |
| | Inhambane | 3.25* | 0.21 | 0.00 |
| | Zambezia | -11.77* | 0.21 | 0.00 |
| | C.Delgado | -0.56* | 0.21 | 0.01 |
| C.Delgado | Maputo | 2.97* | 0.21 | 0.00 |
| | Gaza | 1.46* | 0.21 | 0.00 |
| | Inhambane | 3.82* | 0.21 | 0.00 |
| | Zambezia | -11.21* | 0.21 | 0.00 |
| | Nampula | 0.56* | 0.21 | 0.01 |

The mean difference is significant at the .05 level.

**Table A7:** **ANOVA of CMD severity and whitefly abundance in the 2006 survey**

| | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| CMDSEV | Between Groups | 385.022 | 5 | 77.004 | 102.865 | .000 |
| | Within Groups | 6155.733 | 8223 | .749 | | |
| | Total | 6540.755 | 8228 | | | |
| WFNO | Between Groups | 353183.3 | 5 | 70636.661 | 896.901 | .000 |
| | Within Groups | 1369810 | 17393 | 78.756 | | |
| | Total | 1722993 | 17398 | | | |

CMDSEV: Severity of disease based on score scale of 1-5
Group = Province
WFNO=Whitefly numbers

**Table A8: Multiple comparisons of symptom severity and whitefly numbers between the provinces in the 2006 survey**

| Dependent Variable CMDSEV | | | | |
|---|---|---|---|---|
| **Province (I)** | **Province(J)** | **Mean Diff.(I-J)** | **Std. Error** | **Sig.** |
| Maputo | Gaza | -0.14* | 0.03 | 0.00 |
| | Inhambane | -0.05 | 0.04 | 0.14 |
| | Zambezia | 0.42* | 0.03 | 0.00 |
| | Nampula | 0.31* | 0.03 | 0.00 |
| | C.Delgado | 0.33* | 0.04 | 0.00 |
| Gaza | Maputo | 0.14* | 0.03 | 0.00 |
| | Inhambane | 0.08* | 0.03 | 0.02 |
| | Zambezia | 0.56* | 0.03 | 0.00 |
| | Nampula | 0.44* | 0.03 | 0.00 |
| | C. Delgado | 0.47* | 0.04 | 0.00 |
| Inhambane | Maputo | 0.05 | 0.04 | 0.14 |
| | Gaza | -0.08* | 0.03 | 0.02 |
| | Zambezia | 0.48* | 0.04 | 0.00 |
| | Nampula | 0.36* | 0.04 | 0.00 |
| | C.Delgado | 0.38* | 0.05 | 0.00 |
| Zambezia | Maputo | -0.42* | 0.03 | 0.00 |
| | Gaza | -0.56* | 0.03 | 0.00 |
| | Inhambane | -0.48* | 0.04 | 0.00 |
| | Nampula | -0.12* | 0.03 | 0.00 |
| | C. Delgado | -0.09* | 0.04 | 0.04 |
| Nampula | Maputo | -0.31* | 0.03 | 0.00 |
| | Gaza | -0.44* | 0.03 | 0.00 |
| | Inhambane | -0.36* | 0.04 | 0.00 |
| | Zambezia | 0.12* | 0.03 | 0.00 |
| | C. Delgado | 0.02 | 0.04 | 0.56 |
| C.Delgado | Maputo | -0.33* | 0.04 | 0.00 |
| | Gaza | -0.47* | 0.04 | 0.00 |
| | Inhambane | -0.38* | 0.05 | 0.00 |
| | Zambezia | 0.09* | 0.04 | 0.04 |
| | Nampula | -0.02 | 0.04 | 0.56 |

| Province (I) | Province (J) | | | |
|---|---|---|---|---|
| Maputo | | -2.62* | 0.23 | 0.00 |
| | Inhambane | -0.71* | 0.23 | 0.00 |
| | Zambezia | -12.79* | 0.23 | 0.00 |
| | Nampula | -7.76* | 0.23 | 0.00 |
| | C. Delgado | -5.74* | 0.24 | 0.00 |
| Gaza | Maputo | 2.62* | 0.23 | 0.00 |
| | Inhambane | 1.91* | 0.23 | 0.00 |
| | Zambezia | -10.17* | 0.23 | 0.00 |
| | Nampula | -5.14* | 0.23 | 0.00 |
| | C. Delgado | -3.12* | 0.24 | 0.00 |
| Inhambane | Maputo | 0.71* | 0.23 | 0.00 |
| | Gaza | -1.91* | 0.23 | 0.00 |
| | Zambezia | -12.08* | 0.23 | 0.00 |
| | Nampula | -7.05* | 0.23 | 0.00 |
| | C. Delgado | -5.03* | 0.24 | 0.00 |
| Zambezia | Maputo | 12.79* | 0.23 | 0.00 |
| | Gaza | 10.17* | 0.23 | 0.00 |
| | Inhambane | 12.08* | 0.23 | 0.00 |
| | Nampula | 5.03* | 0.23 | 0.00 |
| | C. Delgado | 7.05* | 0.24 | 0.00 |
| Nampula | Maputo | 7.76* | 0.23 | 0.00 |
| | Gaza | 5.14* | 0.23 | 0.00 |
| | Inhambane | 7.05* | 0.23 | 0.00 |
| | Zambezia | -5.03* | 0.23 | 0.00 |
| | C. Delgado | 2.02* | 0.24 | 0.00 |
| C. Delgado | Maputo | 5.74* | 0.24 | 0.00 |
| | Gaza | 3.12* | 0.24 | 0.00 |
| | Inhambane | 5.03* | 0.24 | 0.00 |
| | Zambezia | -7.05* | 0.24 | 0.00 |
| | Nampula | -2.02* | 0.24 | 0.00 |

* The mean difference is significant at the .05 level

## *Appendix 2: Nucleotide sequences of full length DNA A from 6 geminivirus isolates from Mozambique (one from each province)*

>Moz1Inhambe
ACCGGTTGGCCCCGCCCCCTTTAAACGTGGTCCCCGCGCACTACGTATGT
CGGCCAATCATGTTGTAGCGTTAAAGGTTATTTATTAGTGGTTTACCACT
ATATACTTGCAGGCGAAGTTGTTGCTAGTGCGCTATGTGGGATCCACTGG
TGAATGAGTTTCCAGACTCGGTGCATGGGCTTAGGTGTATGCTTGCAATT
AAATATTTGCAGGCCTTAGAGGATACATACGAGCCCAGTACGTTGGGCCA
CGATCTGGTTAGGGATCTAGTCTCAGTCATCAGGGCTCGTAATTATGTCG
AAGCGACCAGGAGATATCATCATTTCCACTCCAGGCTCCAAGGTTCGTCG
AAGGCTGAACTTCGACAGCCCATACAGGAACCGTGCTACTGCCCCCACTG
TCCACGTCACAAATCGAAAATGGGCCTGGATGAACAGGCCCATGTACAGA
AAGCCCATGATGTACAGGATGTACAGAAGCCCAGACATACCTAGGGGTTG
TGAAGGCCCATGTAAGGTCCAGTCGTTTGAGCAGAGAGATGATGTGAAGC
ACCTTGGTATCTGTAAGGTGATTAGTGATGTGACTCGTGGGCCTGGGTTG
ACACACAGGGTCGGAAAGAGGTTTTGTATCAAGTCCATTTACATTCTTGG
TAAGATCTGGATGGATGAAAATATTAAGAAGCAGAATCACACTAATAATG
TGATTTTTTACCTGCTTAGGGATAGAAGGCCTTACGGCAATGCGCCCCAA
GACTTTGGGCAGATATTTAACATGTTTGATAATGAGCCCAGTACTGCAAC
AATTAAGAATGATTTGAGGGATAGGTTTCAGGTGTTGAGGAAATTTCATG
CCACTGTTGTTGGTGGTCCATCTGGCATGAAGGAGCAGGCTTTGGTGAAA
AGGTTTTACAGGTTGAATCATCACGTGACATATAATCATCAGGAGGCAGG
GAAGTATGAGAATCACACAGAGAATGCTTTGCTTCTGTACATGGCATGTA
CTCATGCCTCCAATCCTGTATATGCGACTTTGAAAATACGTATATACTTC
TACGACAGTATTGGCAATTAATAAATATTGAATTTTATTTCATGAGTCAA
CTGACATTCAACAGTTTTTTCAATTACATTGAACAAAACATGATCAGCAG
CTCTAATTACATCGTTAATTGAGATAACACCTATATGATCCAAGTATTTA
AGTACTTGGTATCTAAAGACCCTTAAGAAAAGACCAGTCTGAGGCTGTAA
GGTCGTCCAGATCTTGAAGTTGAGAAAACATTTGTGAATCCCCAGCTCCT
TCCTCAGGTTGTGATTGAATCGAACCTGGACAGTTATTATGTCCTGGTTC
ATTAGGAATGGTCGTTGTTGGTGCCTGGTGATTGTGAAATACAGGGGATT
GTTTATTTCCCAGGTATACACGCCATTCATTGCTTGAGGAGCAGTGATGA
GTTCCCCTGTGCGTAAATCCATGATTGGAGCAGTTGATATGGAGGTAATA
TGAACAGCCACAGACAAGATCCACTCTCCTACGCCTGATAGCCCTCTTCT
TGAATTGTCTGTGATTGACTTTGATTGGAACCTGAGTAGAGTGGTTCTGT
GAGGGTGATGAAGATTGCATTCTTTAATGCCCAGGCCTTTAGTGCTGTTT
GCTTTTCCTCGTCTAGGAACTCTTTATAGGACGAGGTAGGTCCTGGATTG
CAGAGGAAGATAGTGGGAATTCCACCTTTAATTTGAACGGGTTTCCCGTA
TTTCGTGTTTGATTGCCAGTCTTTCTGGGCCCCAATGAATCCAATCTAGA
CTATAGGGAAAGCTTGCATGCAGGCCTCTGCAGTCGACGGGCCCGGGATC
CGATTTGGTCGACGTCTCAAGACGTTGTACCAGGCATCATTATTGAAGAC
CTTTGGACTAAGGTCCAGGTGTCCACACAGGTAATTGTGTGGTCCTAAAG
ATCTGGCCCATATCGTCTTCCCTGTTCTGCTATCACCTTCTATGACAATA
CTATTGGGTCTCCATGGCCGCGCAGCGGAATCCCTAACATTATCAGCGAC
CCATTCTTCAAGTTCATCAGGAACTTGGTCAAAGGAAGAACATGAGAAGG
GAGAAACATAAGGAGCTGGCGGCTCCTGGAAAATCCTATCTAAATTACTA
TTTAGATTATGAAACTGTAGTACAAAGTCCTTTGGGACTAATTCCCTAAT
GACATTAAGAGCTTCTGACTTACTGCCGCTGTTAAGCGCTTTGGCGTAAG
CATCATTCGTTGATTGTTGACCTCCTCTGGCAGATCGTCCATCGATCTGA
AATTGTCCCCATTCGATGGTGTCTCCGTCCTTATCCAGATAGGACTTGAC
ATCCGAGCTTGATTTAGCACCTTGAATGTTGGGGTGGAAACTGGTGCTAC
AGCTTGGGTGTACACAATCGAATAGACGATTGTTCGTAATCGTGATTTTT

CCCTCAAATTGAATGAGGGCATGCAAGTGAGGTTCCCCATTCTGATGCAG
CTCTCTACAGATTTTAATGAATTTAGGGTTGGATGGGAGAGAGAGTGTGT
GAATAAATGACAGCAGGTGTTCTTTGGGTATAGAACACTTTGGGTATGTG
AGAAAGACATTCTTGGCTTGAACTCTAAAACGAGGAGTTCTCATGTTGAC
CAAGTCAATTGGAGACACTCAACTAGAGACACTCTTGAGCATCTCCTCCT
GTTAATTGGAGACATTATATAGTTGTCTCTAAATGGCATTTTTGTAATAA
GTGGAACTTTAATTTGAATGAAAAGGCGCAAAAGGCTCATAACACCCAAG
GGGCCAACCGTATAATATT

>Moz2Maputo
ACCGGTTGGCCCCCGCCCCCCTTTAGACGTGGTCCCCGCGCACTATGTAT
GTCGGCCAATCATGTCGTAGCTTTAAAGGTTATTTATTAGTGGTTTACCA
CTATATACTTGCAGGCGAAGTTGTCGCTAGTGCGCTATGTGGGATCCACT
GGTGAATGAGTTGCCAGAGTCGGTGCATGGGCTTAGGTGTATGCTTGCAA
TTAAATATTTGCAGGCCTTAGAGGATACATACGAGCCCATTACTTTGGGC
CACGATTTGGTCAGGGATCTAATCTCAATTATCAGGGCTCGTAATTATGT
CGAAGCGACCAGGAGATATCATAATTTCCACTCCAGGCTCCAAAGTTCGT
CGAAGATTGAACTTCGACAGCCCATACAGGAACCGTGCTACTGCCCCCAC
TGTCCACGTCACAAATCGAAAACGGGCCTGGATGAACAGGCCCATGTACA
GAAAGCCCATGATGTACAGGATGTATAGAAGCCCAGACATACCTAGGGGC
TGTGAAGGCCCATGTAAGGTGCAGTCGTTTGAGCAGAGGGATGATGTGAA
GCACCTTGGTATCTGTAAGGTGATCAGTGATGTGACTCGTGGGCCTGGGC
TGACACACAGGGTCGGAAAGAGGTTTTGTATCAAGTCCATTTACATTCTT
GGTAAGATCTGGATGGATGAAAATATTAAGAAGCAGAATCACACTAATAA
TGTGATTTTTTACCTGCTTAGGGATAGAAGGCCGTATGGCAATGCGCCCC
AAGACTTTGGGCAGATATTTAACATGTTTGATAATGAGCCCAGTACTGCA
ACAATTAAGAACGATTTGAGGGATAGATTTCAGGTGTTGAGGAAATTCCA
TGCCACTGTTGTTGGTGGTCCATCTGGCATGAAGGAGCAGGCTTTGGTGA
AAAGGTTTTACAAGTTGAATCATCACGTGACATATAACCATCAGGAAGCA
GGGAAGTATGAGAATCACACAGAGAATGCTTTGCTTCTGTACATGGCATG
TACGCATGCCTCCAATCCTGTATATGCTACGTTGAAAATACGTATATATT
TCTATGACAGTATTGGCAATTAATAAACATTGAATTTTATTTCATGAGTC
AACTGACACTCAATAGTTTTTTCAATTACATTGAACAAAACATGATCAGC
AGCTCTAATTACATTGTTAATTGAGATAACACCCATATTATCCAAGTATT
TAATTACTTGGTATCTAAAGACTCTTAAGAAAAGACCAGTCTGAGGCTGT
AAGGTCGTCCAGATCTTGAAGTTGAGAAAACATTTGTGAATCCCCAGCTC
CTTCCTCAGGTTGTGATTGAATCGAACCTGGACCGTAATGATGTCCTGGT
TCATTAGGAATGGTCGTTGTTGGTGCCTGGTGATTGTGAAATACAGGGGA
TTGTTTATTTCCCAGGTATACACGCCATTCATTGCTTGAGGAGCAGTGAT
GAGTTCCCCTGTGCGTAAATCCATGATTGGAGCAGTTGATATGGAGGTAA
TACGAACAGCCACAGACAAGATCCACTCGCCTACGCCTGATGGCTCTCTT
CTTGAATTGTCTGTGATTGACTTTGATTGGAACCTGAGTAGAGTGGTTCT
GTGAGGGTGATGAAGATTGCATTCTTTAATGCCCAAGCCTTTAGTGCTTC
TTGCTTTTCCTCGTCTAAGAACTCTTTATAGGACGAGGTAGGTCCTGGAT
TGCATAGGAAGATAGTGGGAATTCCACCTTTAATTTGAACGGGTTTCCCG
TATTTTGTGTTGGACTGCCAGTCCCTCTGGGCCCCCATGAATTCCTTAAT
CGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTTCCCTAT
AGTGAGTCGTATTAGAGCTTGGTAGATTCATCAAGACGTTGTACCAGGCA
TCATTATTGAAGACCTTGGGACTAAGGTCCAGGTGTCCACACAGGTAATT
GTGTGGGCCTAAAGATCTGGCCCATATCGTCTTCCCTGTTCTGCTATCAC
CTTCTATGACAATACTATTGGGTCTCCATGGCCGCGCAGCGGAATCCCTA
ACATTATCAGCGACCCATTCTTCAAGTTCAACAGGAACTTGGTCAAAGGA
AGAACATAGGAAGGGAGAAACATAAGGAGCTGGTGGCTCCTGGAAAATCC
TATCTAAATTACTATTTAGATTATGAAACTGAAGTACAAAGTCTTTTGGG
ACTAATTCCCTAATTACATTCATAGCTTCGGACTTACTGCCGCTGTTAAG
CGCTTTGGCGTAAGCATCATTCGCTGATTGTTGACCTCCTCTAGCAGATC

GTCCATCGATCTGAAATTGTCCCCATTCGACGGTGTCGCCGTCCTTATCC
AGATAGGACTTGACATCTGAGCTTGATTTGGCACCTTGAATGTTGGGGTG
GAAACTGGTGCGACAGCTTGGGTGTACACAATCGAAGAGACGATTGTTCG
TAATCGTGATTTTGCCCTCGAATTGGATGAGGGCATGCAAGTGAGGTTCT
CCATTCTGATGCAGTTCTCTACAGATTTTAATGAACTTAGGGTTTGATGG
GAGAGAGAGTGTTTGAATGAATGACAGCAGGTGTTCTTTGGGTATAGAAC
ACTTTGGGTATGTGAGAAAGACATTCTTGGCTTGAACTCGAAAACGAGGA
GTTCTCATTTTGACCAAGTCAATTGGAGACACTCAACTAGAGACACTCTT
GAGCATCTCCTCCTGTTAATTGGAGACATTATATAGTTGTCTCTAAATGG
CATTCTTGTAATAAGTGGAACTTTAATTTGAATTAAAAGGCTCAAAAGGC
TCAGAACACCCAGGGGGCCAACCGTATAATATT

>Moz3CaboDelgado
ACCGGATGGCCGCGCCCGAAAAAGCAGATGGACCCCACCAGATTGCCGCC
CCCGTGAAAGAAAGTGGTCCCCGCGCACTTGTTTCTGTCGGCCAGTCATA
TGCACGCGTGAAAGTCTAGATATTTGTTGTGTGTCTTTATAGACTTCGTC
GCGAAGTAGGGAAGCGCGTCAACATGTGGGATCCATTGTTGAATGATTTT
CCTGAAACCGTTCACGGTTTCCGTTCTATGCTTGCTGTTAAATACCTGTT
ACATCTGGAACAGGAATACGATCGCGGTACTGTCGGGGCTGAGTATATAC
GGGATCTGATAGGGGTTTTACGCTGTAAGAGTTATGTCGAAGCGACCAGG
AGATATAATAATCTCAACACCCGTATCCAAGGTGCGGAGGAGGCTGAACT
TCGACAGCCCATTCACGAACCGTGTTGTTGCCCCCACTGTCCGCGTCACC
AGAAGCAAACTATGGGCCAACAGGCCCATGTATCGGAAGCCCAGGATGTA
CAGAATGTATCGAAGCCCAGATGTCCCTAAGGGCTGTGAAGGCCCATGTA
AGGTTCAGTCCTATGAACAGAGGGATGATGTTAAGCACACTGGTATGGTC
CGATGTGTCAGTGATGTGACTCGTGGGTCAGGCATTACCCATAGAGTCGG
GAAGATGTTTTGTGTGAAGTCCATATATATATTGGGGAAGATCTGGATGG
ACGAGAATATCAAGAAGCAAAATCATACGAACCATGTTATGTTCTTCCTC
GTTCGAGATAGAAGGCCTTATGGTCCGAGCCCGCAAGATTTCGGACAAGT
GTTCAATATGTTTGATAATGAACCTACTACTGCAACGGTGAAGAATGATC
TCAGGGACCGGTATCAGGTGTTACGTAAATTTTATGCGACTGTTGTTGGT
GGACCCTCCGGGATGAAGGAACAAGCTCTGGTTAAGAGGTTTTTTAGGAT
CAATAACCATGTAGTGTATAATCATCAGGAACAGGCCAAGTATGAGAATC
ATACTGAGAATGCGTTGTTATTGTATATGGCATGTACACATGCCTCAAAT
CCTGTATACGCTACTCTGAAAATACGCATCTATTTCTATGATGCAGTGAC
AAATTAATAAAGGTTGAATTTTATTACATGTTGCTCCGTAACTTGGAGTG
TGTTTAGTAATACATCGTACATAACATGATCAACAGCTTGAAGTACAGTG
TTAATGGAAATAACGCCTATCATATCTAAATACTTGAGCACTTGATATTT
AAATACTCTTAAGAAAAGACCAGTCGGAGGCCGTAAGGTCGTCCAGACCT
TGAAGTTGAGAAAACACTTGTGAATCCCCAATGCCTTCCTGAGGTTGTGG
TTGAACCGTATCTGGAGTGTGATGATGTCGTGGTTCATGTTCCCTGGCCT
CTTGTCGTGGTTGGTGATTTCGAAATAGAGGGGATTTGTTATTTCCCAGG
TAAAAACGCCATTCTTTGCTTGAGGCGCAGTGATGAGTTCCCCTGTGCGA
GAATCCATGGTTAATGCAGTCGATATGGAGATAGAACGAGCAGCCGCATT
CGAGGTCTACCCTCCTACGTCTGAGGGCTCTAGTCTTCGCTGTGCGGTGT
TGGACTTTGATGGGCACTTGAGAACAATGGCTCGTGGAGGGTGATGAAGG
TGGCATTCTTTAAAGCCCAGGCTTTAAGGGACTGGTTCTTTTCCTCGTCC
AGAAACTCTTTATATGATGATGTTGGTCCTGGATTGCAGAGGAAGATAGT
GGGAATGCCGCCTTTAATTTGAATCGGCTTCCCGTACTTTGTATTGCTTT
GCCAGTCTTTCTGGGCCCCAATGAATTCCTTAATCGGATCCCGGGCCCGT
CGACTGCAGAGGCCTGCATGCAAGCTTTCCCTATAGTGAGTCGTATTAGA
GCTTGGGATTGGGGTCGACGTCATCAAGACGTTGTACCATGCGTCGTTTG
AATAGACCTTTGGAGATAGATCCAGGTGTCCACATAGATAATTATGGGGA
CCCAGTGAACGAGCCCACATGGTTTTTCCGGTTCGGCTATCACCTTCTAG
AACAATACTGATCGGTCTCCATGGCCGCGCAGCGGGACTGCATATATTTT
CTGATACCCATACTTCGATTTCGTCTGGGACGTGTGTAAATGAGGATGAT

AAAAATGGACTTACGTAAGTTTGTGGCGGAGTCTGGAAGATTCTATCTGC
GTTAGCAGATATGTTATGGAACTGTAAAAAAAAAGACTTCGGATCTTTTT
CTTTAATAATTTGAAGAGCTTCTGATTTAGAAGAAGCATTCAACGCTTCG
GCATATACCTGAGCTAAATGCTGGCCCTCACCCCGTGCACTTCTGGCATC
GACTTGGAAAACGCCATCGTCAAGAAAGTCCCCTCCCTTTTCAATGTATG
TTTTGACATCGGACGAGGATTTAGCTCCCTGAATGTTCGGATGGAAATGT
GTTGATCTGGATGGGGAAATGAGATCGAAGAATCTCGGGTTGGTACATTG
GAACTTGCCTTCGAATTGAATGAGAACATGGAGATGAGGCACCCCATCTT
GATGTAGTTCTCGGCAAACCCTAATGAATTTGATATTCGTCGGGTAAGAA
AGGGCTTGTAATTGGGAAAGGGCCTCTTCCTTTGTTAATGAGCATCGGGG
ATAGGTTATGAAATAATTTTTGGCATTAATTTGAAAACGACCGGCTCTTG
GCATATTTGCTGTCGTTTTAGATCGGGGGACACTCAAAACTCCAGGAGAA
CGGTGGAATGGGGGGCATTATATAGGATGTCCCCCAATGGCATATGTGTA
AATAGGTAGACTTACATTCAAAATTTGAATTGCGAATAATGGCGGCCATC
CGATTAATATT

>Moz4Zambezia
ACCGGATGGCCGCGCCCGACAAAGCAGATGGACCCCATTGGACGGCCGCG
CCCGTAAAAGAAAGTGGTCCCCGCGCACGTGTTGCTGTCAGCCAGTCATA
TTCACGCGTGAAAGCCTAGATATTTGGTTTATGTCGTTATATACTTCGTC
GCGAAGTAGTGGAGCGCGTCAACATGTGGGATCCATTGTTGAATGAGTTC
CCCGAGTCTGTGCACGGTTTTCGCTGTATGCTTGCTATTAAATATTTGCA
GGCCCTGGAGGAAACATACGAGCCCAATACTTTGGGCCACGATCTAGTCC
GTGATCTCATCTGTGTTATCCGAGCCCGTGATTATGTCGAAGCGACCCGC
CGATATAATCATTTCCACTCCCGCCTCGAAGGTGCGTCGAAGGCTGAACT
TCGACAGCCCGTTCAGCAGCCGTGCTGCTGTCCCCATTGTCCAAGGCACA
AACAAGCGACGAGCATGGACGTTACGGCCCATGTATCGAAAGCCCAGAAT
GTACAGGATGTACAGAAGTCCTGATGTACCTCGAGGATGTGAAGGCCCAT
GTAAGGTACAGTCATATGAACAGAGAGACGATGTTAAGCACACCGGTGCT
GTTCGTTGTGTTAGTGATGTTACGCGTGGTTCGGGTATTACTCATAGGGT
AGGGAAGAGATTTTGTGTTAAGTCAATATATGTTTTAGGAAAGATCTGGA
TGGATGAAAACATCAAGAAGCAGAATCATACTAATCAGGTTATGTTCTTC
CTAGTCCGTGACAGAAGGCCCTATGGCACAAGCCCCATGGACTTTGGACA
GGTTTTTAATATGTTTGATAATGAGCCCAGTACAGCTACTATTAAGAATG
ATTTGCGAGATAGGTTCCAAGTGTTGCGGAAATTCCATGCCACTGTGGTA
GGTGGTCCCTCGGGTATGAAGGAACAGGCGTTGATTAAGAGGTTTTTTAG
GGTTAATAATCATGTTGTGTATAATCACCAGGAGGCAGCGAAGTATGAGA
ATCATACAGAGAATGCGTTGTTGTTGTATATGGCATGTACGCATGCCTCT
AATCCAGTGTATGCTACGCTTAAAATACGCATCTATTTTTATGATGCAGT
AACAAATTAATAAAGGTTGAATTTTATTTCATGTTGCTCCGTAACTTGGA
GTGTGTTTAGTAATACATTGTACAGAACATGATCAACAGCTCTAAGTACA
GTGTTAATGGAAATAACGCCTATCATATTTAAATACTTGAGCACTTGATA
TTTAAATACTTTTAAGAAAAGACCAGTCGGAGGCCGTAAGGTCGTCCAGA
CCTTGAAGTTGAGAAAACACTTGTGAATCCCCAATGCCTTCCTGATGTTG
TGGTTGAATCGTATCTGGAGTGTGATGATATCGTGGTGCATGTTCCCTGG
CCTCTTGTCGTGGTTGGTGATTGCGAAATAGAGGGGATTTGTTATTTCCC
AGGTAAAAACGCCATTCTTTGCTTGAGGCGCAGTGATGAGTTCCCCTGTG
CGAGAATCCATGGTTGATGCAGTCAATATGGAGATAGAACGAGCAGCCGC
ATTCGAGGTCTACCCTCCTACGTCTGATGGCTCTGTTCTTCGCGGTGCGG
TGTTGGACTTTGATAGGCACTTGAGAACAATGGCTCGTGGAGGGTGATGA
ATGTTGCATTTTTTAAAGCCCAAGCTTTAAGGGACTGGTTTTTTTCCTCT
TCCAGAAACTCTTTATATGATGATGTTGGTCCTGGATTGCAGAGGAAGAT
AGTGGGAATGCCGCCTTTAATTTGAATCGGCTTCCCGTACTTTGTATTGC
TTTGCCAGTCCCTTTGGGCCCCAATGAATTCCTTAATCTAGATGCATCGC
TCTATACGACTCACTATAGGGAAAGCTTGCATGCAGGCCTCTGCAGTCGA
CGGGCCCGGGATCCGATTAAGGAATTCATGGGGGCCCAGTGAACGAGCCC

ACATGGTTTTACCGGTCCGGCTATCACCTTCTAGAACAATACTGATCGGT
CTCCATGGCCGCGCAGCGGGACTGCATATATTTTCGGATACCCATACTTC
TATGTCGTCTGGGACTTGTGTAAATGAGGATGATAAGAACGGACTTACGT
AAGTTTGTGGCGGAGTCTGGAAGATTTTATCTGCGTTAGCAGATATGTTA
TGGAACTGTAAAAAAAAGGACTTCGGATCCTTTTCTTTAATAATTTGAAG
AGCTTCGGATTTAGAAGAAGCATTCAACGCTTCTGCATATACCTGAGCTA
AATGCTGGCCCTCACCCCTTGCACTTCTGGCATCGACTTGGAAAACTCCA
TCGTCGAGAAATTCCCCTTCCTTTTCAATGTAAGCTTTGACATCAGACGA
TGATTTAGCTCCCTGAATGTTCGGATGGAAATGTGTTGATCTGGATGGGG
AAATAAGATCGAAGAATCTCGGGTTGGTACATTGGAACTTGCCTTCGAAT
TGAATGAGAACATGGAGATGAGGCACCCCATCCTGATGTAGTTCTCTGCA
AACCCTAATGAATTTGATATTCGTCGGGTAAGAAAGGGCTCGTAATTGGG
AAAGTGCCTCTTCCTTTGTTAATGAGCATCGGGGATAGGTTATGAAATAA
TTTTTGGCATTGATTTGAAAACGACCGGCTCTTGGCATATTTGCTGTCGT
TTTGGATCGGGGGACACTCAAAACTCCAGGTAAACGGTGGAATGGGGGGC
ATTATATAGTATGTCCCCCAATGGCATATGTGTAAATAGGTAGACTTACA
TTTAAAATTTGAATTTCGAATAATGGCGGCCATCCGATTAATATT

>Moz5Gaza
ACCGGATGGCCGCGCCCGACAAAGTAGGTGGACCCCATTGGATGACCGCG
CCCGTAAAAGAAAGTGGTCCCCACGCACGTGTTGCTGTCGGCCAGTCATA
TTCACGCGTGAAAGCCTAGATATTTGTTTTATGTCGTTATAGACTTCGTC
GCGAAGTAGTGGAGCGTGTCAACATGTGGGATCCATTGTTGAATGAGTTC
CCCGAGTCTGTGCACGGTTTTCGCTGTATGCTTGCTATTAAATATTTGCA
GGCCGTGGAGGAAACCTACGAGCCCAATACTTTGGGCCACGATCTAGTCC
GTGATCTCATCGGTGTGATCCGAGCCCGTGATTATGTCGAAGCGTCCCGC
CGATATAATCATTTCCACTCCCGTCTCGAAGGTGCGTCGAAGGCTGAACT
TCGACAGCCCGTTCAGCAGCCGTGCTGCTGTCCCCATTGTCCAAGGCACA
AACAAGCGTCGATCATGGACGTTCCGGCCCATGTACCGAAAGCCCAGAAT
GTACAGAATGTTCAAAAGCCCTGATGTTCCGCGTGGCTGTGAAGGCCCAT
GTAAGGTTCAATCTTATGAACAGCGAGATGACGTTAAGCATACTGGCAGT
GTTCGTTGTGTTAGTGATGTCACGCGTGGTTCGGGAATTACACATAGAGT
AGGTAAAAGGTTCTGTATCAAGTCTATATATGTGTTAGGTAAGATATGGA
TGGATGAAAACATCAAGAAGCAGAACCATACAAACCAGGTCATGTTCTTC
TTAGTCCGTGACAGAAGGCCCTATGGCAATAGCCCCATGGACTTTGGACA
GGTTTTTAATATGTTTGATAATGAGCCCAGTACAGCCACTGTGAAGAACG
ATCTTAGGGATAGGTATCGAGTTATGCGGAAGTTTCATGCCACCGTTGTT
GGGGGTCCTTCTGGAATGAAGGAGCAGGCTTTGGTGAGGAGATTTTTTAG
GATAAATAATCATGTTGTGTATAATCACCAGGAGGCAGCTAAGTATGAGA
ATCATACAGAGAATGCGTTATTGTTGTATATGGCATGTACGCATGCCTCT
AATCCAGTGTATGCGACGCTTAAAATACGCATCTATTTTTATGATGCAGT
AACAAATTAATAAAGGTTGAATTTTATTGCATGTTGCTCCGTAACTTGGA
GTGTGTTTAGTAATACATCGTACAGAACATGATCAACAGCTCTGAGTACA
GTGTTAATGGAAATAACGCCTATCATATTTAAATACTTGAGCACTTGATA
TTTAAATACTCTTAAGAAAAGACCAGTCGGAGGCCGTAAGGTCGTCCAGA
CCTTGAAGTTGAGAAAACACTTGTGAATCCCCAATGCCTTCCTGATGTTG
TGGTTGAACCGTATCTGGAGGGTGATGATGTCGTGGTTCATGTTCCCTGG
CCGCTTGTCGTGGTTGGTGATGTCGAAATAGAGGGGATTTGTTATTTCCC
AGGTAAAAACGCCATTCTTTGCTTGAGGCGCAGTGATGAGTTCCCCTGTG
CGAGAATCCATGGTTGATGCAGTCGATATGGAGATAGAACGAGCAGCCGC
ATTCGAGGTCTACCCTCCTACGCCTGAGGGCCCTAGTCTTCGCGGTGCGG
TGTTGGACTTTGATGGGCACTTGAGAACAATGGCTCGTGGAGGGTGATGA
AGGTTGCATTCTTTAAAGCCCAGGCTTTAAGGGACTGGTTCTTTTCCTCG
TCCAGAAACTCTTTATATGATGATGTCGGTCCTGGATTGCATAGGAAGAT
AGTGGGAATGCCGCCTTTAATTTGAATCGGCTTCCCGTACTTGGTATTGC
TTTGCCAGTCTCTCTGGGCCCCCATGAATTCCTTAATTTATTCGCCAGCT

CTATACGACTCACTATAGGGAAAGCTTGCATGCAGGCCTCTGCAGTCGAC
GGGCCCGGGATCCGATTTGGGGTCGACGTCATCAAGACGTTGTACCATGC
GTCGTTGCTGTAAACCTTTGGACTGAGATCCAAATGTCCACATAAGTAGT
TGTGTGGTCCCAGAGATCGGGCCCACATCGTCTTCCCTGTCCTACTATCG
CCCTCGATGACGATACTACTCGGTCTCCATGGCCGCGCAGCGGAACCCAT
CACGTTCTCGGAAACCCAGTCTTCAAGTTCCTCAGGAACATGAGTGAAAG
AAGAAGAAAGAAAGGGAGAAATATAAGGAATCGGAGGCTCCTGAAAAATC
CTATCTAAATTGCTATTTAAATTATGAAACTGTAAAACAAAATCCTTTGG
GGCTAGTTCCCGGATTACATTAAGAGCCTCTGTTTTACTTGCTGCGTTAA
GAGCCTTGGCGTAAGCGTCATTGGCGGATTGTTGTCCGCCGCGAGCAGAT
CGTCCGTCGATCTGAAACTCGCCCCATTGGATGGTGTCTCCGTCCTTGTC
CAAATAGGACTTGACGTCAGAACTGGATTTAGCTCCCTGAATGTTTGGAT
GGAAATGTGTTGACCTGGAAGGGGATATGAGGTCGAAGAATCGTTGGTTG
GTACAATTGTACTTGCCCTCGAACTGAATGAGGGCATGCAAATGAGGTTC
CCCATTTTCATGGAGTTCTCTGCAGATCTTGATGAACAATTTATTTGTTG
GGGTTTGGAGTTGTCGGAGTTGATCTAATGCCGCTTCTTTCGAGAGAGTG
CATTTCGGATACGTGAGGAAATAATTTTTGGCTTTTATGCTAAAACGACC
AGCCCTCGGCATTTTCGCTGTCGTATAGCAATCGGGGGGCACTCAAAGTT
CCTAGCAATCGGGGGAATGGGGGGGCAATTTATATGATGCCCCCCAAATG
GCATATGTGTAATTTTGTGATGAAATTTGAATTTCGAACGTGGAAAGCGG
CCATCCGTCTAATATT

>Moz6Nampula
ACCGGTTGGCCCCGCCCCCCTTTAAACGTGGTCCCCGCGCACTACGTATG
TCGGCCAATCATGTTGTAGCTTTAAATGTTATTTATTAGTGGTTTACCAC
TATATACTTGCAGGCGAAGTTGTTGCTAGTGCGCTATGTGGGATCCACTG
GTGAATGAGTTTCCAGACTCGGTGCATGGGCTTAGGTGTATGCTTGCAAT
TAAATATTTGCAGGCCTTAGAGGATACATACGAGCCCAGTACTTTGGGCC
ACGATCTGGTTAGGGATCTAGTCTCAGTCATCAGGGCTCGTAATTATGTC
GAAGCGAGCAGGAGATATCATCATTTCCACGCCAGGCTCCAAGGTTCGTC
GAAGGCTGAACTTCGACAGCCTATACAGGAACCGTGCTACTGCCCCCACT
GTCCACGTCACAAATCGAAAACGGGCCTGGATGAACAGGCCCATGTACAG
AAAGCCCATGATGTACAGGATGTATAGAAGCCCAGACATACCTAGGGGCT
GTGAAGGCCCATGTAAGGTCCAGTCGTTTGAGCAGAGAGATGATGTGAAG
CACCTTGGTATCTGTAAGGTGATTAGTGATGTGACTCGTGGGCCTGGGTT
GACACACAGGGTCGGAAAGAGGTTTTGTATCAAGTCCATTTACATTCTTG
GTAAGATCTGGATGGATGAAAATATTAAGAAGCAGAATCACACTAATAAT
GTGATTTTTTACCTGCTTAGGGATAGAAGGCCGTATGGCAATGCGCCCCA
AGACTTTGGGCAGATATTTAACATGTGTGATAATGAGCCCAGTACTGCAA
CAATTAAGAACGATTTGAGGGATAGGTTTCAGGTGTTGAGGAAATTTCAT
GCCATTGTTGTGGGTGGTCCATCTGGCATGAAGGAGCAGTCTTTGGTGAA
AAGGTTTTACAGGTTGAATCATCACGTGACATATAATCATCAGGAGGCAG
GGAAGTATGAGAATCACACAGAGAATGCTTTGCTTCTGTACATGGCATGT
ACTCATGCCTCTAATCCTGTATATGCTACTTTGAAAATACGTATATACTT
CTACGACAGTATTGGCAATTAATAAACATTGAATTTTATTTCATGAGTCA
ACTGACATTCAATAGTTTTTTCAATTACATTGAACAAAACATGATCAGCA
GCTCTAATTACATCGTTAATTGAGATAACACCTATATGATCCAAGTATTT
AAGTACTTGGTATCTAAAGACCCTTAAGAAAAGACCAGTCTGAGGCTGTA
AGGTCGTCCAGATCTTGAAGTTGAGAAAACATTTGTGAATCCCCAGCTCC
TTCCTCAGGTTGTGATTGAATCGAACCTGGACTGTTATTATGTCCTGGTT
CATTAGGAATGGTCGTTGTTGGTGCCTGGTGATTGTGAAATACAGGGGAT
TGTTTATTTCCCAGGTATACACGCCATTCATTGCTTGAGGAGCAGTGATG
AGTTCCCCTGTGCGTAAATCCATGATTGTAGCAGTTGATATGGAGGTAAT
ATGAACAGCCACAGACAAGATCCACTCTCCTACGCCTGACAGCTCTCTTC
TTGAATTGTCTGTGATTGACTTTGATTGGAACCTGAGTAGAGTGGTTCTG
TGAGGGTGATGAAGATTGCATTCTTTAATGCCCAGGCCTTTAGTGCTGCT

TGCTTTTCCTCGTCTAGGAACTCTTTATAGGACGAGGTTGGTCCTGGATT
GCAGAGGAAGATAGTGGGAATGCCACCTTTAATTTGAACGGGTTTCCCGT
ATTTCGTGTTTGATTGCCAGTCTCTCTGGGCCCCAATGAATTCCTTAATC
GGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTTCCCTATA
GTGAGTCGTATTAGAGCTGGCGAATGCCTCTAGATTGGGGTCGACGTCAT
CAAGACGTTGTACCAGGCATCATTATTGAAGACCTTTGGACTAAGGTCCA
GGTGGCCACACAGGTAATTGTGTGGGCCTAAAGATCTGGCCCATATCGTC
TTCCCTGTTCTGCTATCACCTTCTATGACAATACTATTGGGTCTCCATGG
CCGCGCAGCGGAATCCCTAACATTATCAGCGACCCATTCTTCAAGTTCAT
CAGGAACTTGGTCAAAGGAAGAACATGAAAAGGGAGAAACATAAGGAGCG
GGCGGCTCCTGGAAAATCCTATCTAAATTACTATTTAGATTATGAAACTG
TAGTACAAAGTCCTTTGGGACTAATTCCCTGATGACATTAAGAGCTTCTG
ACTTACTGCCGCTGTTAAGCGCTTTGGCGTAAGCATCATTCGTTGATTGT
TGACCTCCTCTGGCAGATCGTCCATCGATCTGAAATTGTCCCCATTCGAT
GGTGTCTCCGTCCTTATCCAGATATGACTTGACATCGGAGCTTGATTTAG
CACCTTGAATGTTGGGGTGGAAACTGGTGCTACAGCTTGGGTGTACACAA
TCGAATAGACGATTGTTCGTAATCGTGATTTTTCCCTCGAATTGAATGAG
GGCATGCAAGTGAGGTTCCCCATTCTGATGCAGCTCTCTACATATTTTAA
TGAATTTAGGGTTGGATGGGAGAGAGAGTGTGTGAATAAATGACAGTAGG
TGTTCTTTGGGTATAGAACACTTTGGGTATGTGAGAAAGACATTCTTGGC
TTGAACTCTAAAACGAGGAGTTCTCATGTTGACCAAGTCAATTGGAGACA
CTCAACTAGAGACACTCTTGAGCATCTCCTCCTGTTAATTGGAGACATTA
TATAGTTGTCTCTAAATGGCATTCTTGTAATAAGTGGAACTTTAATTTGA
ATGAAAAGGCGCAAAAGGCTCATAACACCCAAGGGGCCAACCGTATAATA
TT

## *Appendix 3: Percentage nucleotide sequences identities of geminivirus isolates from Mozambique compared to selected virus isolates from different countries*

Alignment Score Sequence Number Sequence Name Sequence Length

| SeqA | Name | Len(nt) | SeqB | Name | Len(nt) | Score |
|---|---|---|---|---|---|---|
| 1 | ACMV[Nigeria] | 2781 | 2 | EACMKV-[K261] | 2797 | 75 |
| 1 | ACMV[Nigeria] | 2781 | 3 | EACMV_Malawivirus-Malawi[MH] | 2804 | 72 |
| 1 | ACMV[Nigeria] | 2781 | 4 | SACMV | 2800 | 75 |
| 1 | ACMV[Nigeria] | 2781 | 5 | >TLCUV | 2747 | 78 |
| 1 | ACMV[Nigeria] | 2781 | 6 | Moz01Inhambane | 2819 | 95 |
| 1 | ACMV[Nigeria] | 2781 | 7 | Moz02Maputo | 2833 | 95 |
| 1 | ACMV[Nigeria] | 2781 | 8 | Moz03CaboDelgado | 2861 | 70 |
| 1 | ACMV[Nigeria] | 2781 | 9 | Moz04Zambizia | 2795 | 70 |
| 1 | ACMV[Nigeria] | 2781 | 10 | Moz05Gaza | 2866 | 75 |
| 1 | ACMV[Nigeria] | 2781 | 11 | Moz06Nampula | 2852 | 95 |
| 1 | ACMV[Nigeria] | 2781 | 12 | V33SAPongola08 | 2768 | 74 |
| 1 | ACMV[Nigeria] | 2781 | 13 | V31SAMookSevere | 2768 | 75 |
| 1 | ACMV[Nigeria] | 2781 | 14 | CLCuGV_AF260241 | 2761 | 68 |
| 1 | ACMV[Nigeria] | 2781 | 15 | CLCuGV_AY036010 | 2764 | 67 |
| 1 | ACMV[Nigeria] | 2781 | 16 | EACMCV_AF112354 | 2802 | 68 |
| 1 | ACMV[Nigeria] | 2781 | 17 | EACMMV_AJ006460 | 2804 | 71 |
| 1 | ACMV[Nigeria] | 2781 | 18 | EACMVUG2_Z83257 | 2799 | 74 |
| 1 | ACMV[Nigeria] | 2781 | 19 | HoLCrV_AY036009 | 2755 | 69 |
| 1 | ACMV[Nigeria] | 2781 | 20 | PepYVV_AY502935 | 2786 | 73 |
| 1 | ACMV[Nigeria] | 2781 | 21 | SACMV_AF155806 | 2800 | 75 |
| 1 | ACMV[Nigeria] | 2781 | 22 | SACMV_AJ422132 | 2800 | 75 |
| 1 | ACMV[Nigeria] | 2781 | 23 | SACMV_AJ575560 | 2800 | 76 |

1 ACMV[Nigeria] 2781 24 TblCZV_AF350330
2767 72
1 ACMV[Nigeria] 2781 25 ToCSV_AF261885
2766 74
1 ACMV[Nigeria] 2781 26 ToLCMLV_AY502936
2773 73
1 ACMV[Nigeria] 2781 27 ToLCMGV_AJ865338
2777 77
1 ACMV[Nigeria] 2781 28 ToLCMGV_AJ865339
2775 77
1 ACMV[Nigeria] 2781 29 ToLCYTV_AJ865341
2765 77
1 ACMV[Nigeria] 2781 30 ToLCYTV_AJ865340
2768 76
1 ACMV[Nigeria] 2781 31 ToLCSDV_AY044137
2779 73
1 ACMV[Nigeria] 2781 32 ToLCSDV_AY044139
2768 74
1 ACMV[Nigeria] 2781 33 TYLCMlV_AF271234
2782 73
1 ACMV[Nigeria] 2781 34 TYLCV_AJ489258
2781 74
1 ACMV[Nigeria] 2781 35 TYLCV_AY044138
2780 75
1 ACMV[Nigeria] 2781 36 TYLCV_X76319
2790 74
1 ACMV[Nigeria] 2781 37 TYLCV_X15656
2787 74
2 EACMKV-[K261] 2797 3 EACMV_Malawivirus-
Malawi[MH] 2804 80
2 EACMKV-[K261] 2797 4 SACMV
2800 88
2 EACMKV-[K261] 2797 5 >TLCUV
2747 79
2 EACMKV-[K261] 2797 6 Moz01Inhambane
2819 74
2 EACMKV-[K261] 2797 7 Moz02Maputo
2833 73
2 EACMKV-[K261] 2797 8 Moz03CaboDelgado
2861 87
2 EACMKV-[K261] 2797 9 Moz04Zambizia
2795 78
2 EACMKV-[K261] 2797 10 Moz05Gaza
2866 87
2 EACMKV-[K261] 2797 11 Moz06Nampula
2852 74
2 EACMKV-[K261] 2797 12 V33SAPongola08
2768 75
2 EACMKV-[K261] 2797 13 V31SAMookSevere
2768 75
2 EACMKV-[K261] 2797 14 CLCuGV_AF260241
2761 66
2 EACMKV-[K261] 2797 15 CLCuGV_AY036010
2764 66
2 EACMKV-[K261] 2797 16 EACMCV_AF112354
2802 79

2 EACMKV-[K261] 2797 17 EACMMV_AJ006460
2804 80
2 EACMKV-[K261] 2797 18 EACMVUG2_Z83257
2799 85
2 EACMKV-[K261] 2797 19 HoLCrV_AY036009
2755 74
2 EACMKV-[K261] 2797 20 PepYVV_AY502935
2786 76
2 EACMKV-[K261] 2797 21 SACMV_AF155806
2800 88
2 EACMKV-[K261] 2797 22 SACMV_AJ422132
2800 88
2 EACMKV-[K261] 2797 23 SACMV_AJ575560
2800 87
2 EACMKV-[K261] 2797 24 TblCZV_AF350330
2767 72
2 EACMKV-[K261] 2797 25 ToCSV_AF261885
2766 75
2 EACMKV-[K261] 2797 26 ToLCMLV_AY502936
2773 74
2 EACMKV-[K261] 2797 27 ToLCMGV_AJ865338
2777 78
2 EACMKV-[K261] 2797 28 ToLCMGV_AJ865339
2775 80
2 EACMKV-[K261] 2797 29 ToLCYTV_AJ865341
2765 80
2 EACMKV-[K261] 2797 30 ToLCYTV_AJ865340
2768 77
2 EACMKV-[K261] 2797 31 ToLCSDV_AY044137
2779 74
2 EACMKV-[K261] 2797 32 ToLCSDV_AY044139
2768 74
2 EACMKV-[K261] 2797 33 TYLCMlV_AF271234
2782 75
2 EACMKV-[K261] 2797 34 TYLCV_AJ489258
2781 74
2 EACMKV-[K261] 2797 35 TYLCV_AY044138
2780 76
2 EACMKV-[K261] 2797 36 TYLCV_X76319
2790 74
2 EACMKV-[K261] 2797 37 TYLCV_X15656
2787 74
3 EACMV_Malawivirus-Malawi[MH] 2804 4 SACMV
2800 85
3 EACMV_Malawivirus-Malawi[MH] 2804 5 >TLCUV
2747 77
3 EACMV_Malawivirus-Malawi[MH] 2804 6 Moz01Inhambane
2819 71
3 EACMV_Malawivirus-Malawi[MH] 2804 7 Moz02Maputo
2833 71
3 EACMV_Malawivirus-Malawi[MH] 2804 8 Moz03CaboDelgado
2861 87
3 EACMV_Malawivirus-Malawi[MH] 2804 9 Moz04Zambizia
2795 94
3 EACMV_Malawivirus-Malawi[MH] 2804 10 Moz05Gaza
2866 85

3 EACMV_Malawivirus-Malawi[MH] 2804 11 Moz06Nampula
2852 71
3 EACMV_Malawivirus-Malawi[MH] 2804 12 V33SAPongola08
2768 76
3 EACMV_Malawivirus-Malawi[MH] 2804 13 V31SAMookSevere
2768 76
3 EACMV_Malawivirus-Malawi[MH] 2804 14 CLCuGV_AF260241
2761 64
3 EACMV_Malawivirus-Malawi[MH] 2804 15 CLCuGV_AY036010
2764 64
3 EACMV_Malawivirus-Malawi[MH] 2804 16 EACMCV_AF112354
2802 80
3 EACMV_Malawivirus-Malawi[MH] 2804 17 EACMMV_AJ006460
2804 98
3 EACMV_Malawivirus-Malawi[MH] 2804 18 EACMVUG2_Z83257
2799 86
3 EACMV_Malawivirus-Malawi[MH] 2804 19 HoLCrV_AY036009
2755 70
3 EACMV_Malawivirus-Malawi[MH] 2804 20 PepYVV_AY502935
2786 76
3 EACMV_Malawivirus-Malawi[MH] 2804 21 SACMV_AF155806
2800 85
3 EACMV_Malawivirus-Malawi[MH] 2804 22 SACMV_AJ422132
2800 86
3 EACMV_Malawivirus-Malawi[MH] 2804 23 SACMV_AJ575560
2800 87
3 EACMV_Malawivirus-Malawi[MH] 2804 24 TblCZV_AF350330
2767 73
3 EACMV_Malawivirus-Malawi[MH] 2804 25 ToCSV_AF261885
2766 75
3 EACMV_Malawivirus-Malawi[MH] 2804 26 ToLCMLV_AY502936
2773 75
3 EACMV_Malawivirus-Malawi[MH] 2804 27 ToLCMGV_AJ865338
2777 77
3 EACMV_Malawivirus-Malawi[MH] 2804 28 ToLCMGV_AJ865339
2775 78
3 EACMV_Malawivirus-Malawi[MH] 2804 29 ToLCYTV_AJ865341
2765 79
3 EACMV_Malawivirus-Malawi[MH] 2804 30 ToLCYTV_AJ865340
2768 77
3 EACMV_Malawivirus-Malawi[MH] 2804 31 ToLCSDV_AY044137
2779 73
3 EACMV_Malawivirus-Malawi[MH] 2804 32 ToLCSDV_AY044139
2768 74
3 EACMV_Malawivirus-Malawi[MH] 2804 33 TYLCMlV_AF271234
2782 75
3 EACMV_Malawivirus-Malawi[MH] 2804 34 TYLCV_AJ489258
2781 75
3 EACMV_Malawivirus-Malawi[MH] 2804 35 TYLCV_AY044138
2780 76
3 EACMV_Malawivirus-Malawi[MH] 2804 36 TYLCV_X76319
2790 75
3 EACMV_Malawivirus-Malawi[MH] 2804 37 TYLCV_X15656
2787 76
4 SACMV 2800 5 >TLCUV
2747 80

4 SACMV 2800 6 Moz01Inhambane
2819 73
4 SACMV 2800 7 Moz02Maputo
2833 73
4 SACMV 2800 8 Moz03CaboDelgado
2861 80
4 SACMV 2800 9 Moz04Zambizia
2795 84
4 SACMV 2800 10 Moz05Gaza
2866 99
4 SACMV 2800 11 Moz06Nampula
2852 73
4 SACMV 2800 12 V33SAPongola08
2768 77
4 SACMV 2800 13 V31SAMookSevere
2768 77
4 SACMV 2800 14 CLCuGV_AF260241
2761 67
4 SACMV 2800 15 CLCuGV_AY036010
2764 67
4 SACMV 2800 16 EACMCV_AF112354
2802 72
4 SACMV 2800 17 EACMMV_AJ006460
2804 85
4 SACMV 2800 18 EACMVUG2_Z83257
2799 80
4 SACMV 2800 19 HoLCrV_AY036009
2755 74
4 SACMV 2800 20 PepYVV_AY502935
2786 78
4 SACMV 2800 21 SACMV_AF155806
2800 100
4 SACMV 2800 22 SACMV_AJ422132
2800 92
4 SACMV 2800 23 SACMV_AJ575560
2800 93
4 SACMV 2800 24 TblCZV_AF350330
2767 75
4 SACMV 2800 25 ToCSV_AF261885
2766 77
4 SACMV 2800 26 ToLCMLV_AY502936
2773 77
4 SACMV 2800 27 ToLCMGV_AJ865338
2777 80
4 SACMV 2800 28 ToLCMGV_AJ865339
2775 82
4 SACMV 2800 29 ToLCYTV_AJ865341
2765 82
4 SACMV 2800 30 ToLCYTV_AJ865340
2768 79
4 SACMV 2800 31 ToLCSDV_AY044137
2779 76
4 SACMV 2800 32 ToLCSDV_AY044139
2768 76
4 SACMV 2800 33 TYLCMlV_AF271234
2782 78

4 SACMV
2781 77
4 SACMV
2780 78
4 SACMV
2790 74
4 SACMV
2787 77
5 >TLCUV
2819 78
5 >TLCUV
2833 78
5 >TLCUV
2861 74
5 >TLCUV
2795 76
5 >TLCUV
2866 80
5 >TLCUV
2852 78
5 >TLCUV
2768 79
5 >TLCUV
2768 79
5 >TLCUV
2761 72
5 >TLCUV
2764 72
5 >TLCUV
2802 71
5 >TLCUV
2804 77
5 >TLCUV
2799 74
5 >TLCUV
2755 75
5 >TLCUV
2786 79
5 >TLCUV
2800 80
5 >TLCUV
2800 81
5 >TLCUV
2800 81
5 >TLCUV
2767 75
5 >TLCUV
2766 78
5 >TLCUV
2773 77
5 >TLCUV
2777 83
5 >TLCUV
2775 83
5 >TLCUV
2765 84

2800 34 TYLCV_AJ489258
2800 35 TYLCV_AY044138
2800 36 TYLCV_X76319
2800 37 TYLCV_X15656
2747 6 Moz01Inhambane
2747 7 Moz02Maputo
2747 8 Moz03CaboDelgado
2747 9 Moz04Zambizia
2747 10 Moz05Gaza
2747 11 Moz06Nampula
2747 12 V33SAPongola08
2747 13 V31SAMookSevere
2747 14 CLCuGV_AF260241
2747 15 CLCuGV_AY036010
2747 16 EACMCV_AF112354
2747 17 EACMMV_AJ006460
2747 18 EACMVUG2_Z83257
2747 19 HoLCrV_AY036009
2747 20 PepYVV_AY502935
2747 21 SACMV_AF155806
2747 22 SACMV_AJ422132
2747 23 SACMV_AJ575560
2747 24 TblCZV_AF350330
2747 25 ToCSV_AF261885
2747 26 ToLCMLV_AY502936
2747 27 ToLCMGV_AJ865338
2747 28 ToLCMGV_AJ865339
2747 29 ToLCYTV_AJ865341

5 >TLCUV 2747 30 ToLCYTV_AJ865340
2768 83
5 >TLCUV 2747 31 ToLCSDV_AY044137
2779 77
5 >TLCUV 2747 32 ToLCSDV_AY044139
2768 78
5 >TLCUV 2747 33 TYLCMlV_AF271234
2782 81
5 >TLCUV 2747 34 TYLCV_AJ489258
2781 79
5 >TLCUV 2747 35 TYLCV_AY044138
2780 81
5 >TLCUV 2747 36 TYLCV_X76319
2790 79
5 >TLCUV 2747 37 TYLCV_X15656
2787 80
6 Moz01Inhambane 2819 7 Moz02Maputo
2833 95
6 Moz01Inhambane 2819 8 Moz03CaboDelgado
2861 70
6 Moz01Inhambane 2819 9 Moz04Zambizia
2795 70
6 Moz01Inhambane 2819 10 Moz05Gaza
2866 72
6 Moz01Inhambane 2819 11 Moz06Nampula
2852 97
6 Moz01Inhambane 2819 12 V33SAPongola08
2768 74
6 Moz01Inhambane 2819 13 V31SAMookSevere
2768 74
6 Moz01Inhambane 2819 14 CLCuGV_AF260241
2761 70
6 Moz01Inhambane 2819 15 CLCuGV_AY036010
2764 70
6 Moz01Inhambane 2819 16 EACMCV_AF112354
2802 67
6 Moz01Inhambane 2819 17 EACMMV_AJ006460
2804 70
6 Moz01Inhambane 2819 18 EACMVUG2_Z83257
2799 73
6 Moz01Inhambane 2819 19 HoLCrV_AY036009
2755 69
6 Moz01Inhambane 2819 20 PepYVV_AY502935
2786 74
6 Moz01Inhambane 2819 21 SACMV_AF155806
2800 73
6 Moz01Inhambane 2819 22 SACMV_AJ422132
2800 74
6 Moz01Inhambane 2819 23 SACMV_AJ575560
2800 75
6 Moz01Inhambane 2819 24 TblCZV_AF350330
2767 72
6 Moz01Inhambane 2819 25 ToCSV_AF261885
2766 74
6 Moz01Inhambane 2819 26 ToLCMLV_AY502936
2773 73

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | Moz01Inhambane | 2819 | 27 | ToLCMGV_AJ865338 | 2777 | 77 |
| 6 | Moz01Inhambane | 2819 | 28 | ToLCMGV_AJ865339 | 2775 | 77 |
| 6 | Moz01Inhambane | 2819 | 29 | ToLCYTV_AJ865341 | 2765 | 77 |
| 6 | Moz01Inhambane | 2819 | 30 | ToLCYTV_AJ865340 | 2768 | 76 |
| 6 | Moz01Inhambane | 2819 | 31 | ToLCSDV_AY044137 | 2779 | 73 |
| 6 | Moz01Inhambane | 2819 | 32 | ToLCSDV_AY044139 | 2768 | 73 |
| 6 | Moz01Inhambane | 2819 | 33 | TYLCMlV_AF271234 | 2782 | 74 |
| 6 | Moz01Inhambane | 2819 | 34 | TYLCV_AJ489258 | 2781 | 74 |
| 6 | Moz01Inhambane | 2819 | 35 | TYLCV_AY044138 | 2780 | 75 |
| 6 | Moz01Inhambane | 2819 | 36 | TYLCV_X76319 | 2790 | 74 |
| 6 | Moz01Inhambane | 2819 | 37 | TYLCV_X15656 | 2787 | 74 |
| 7 | Moz02Maputo | 2833 | 8 | Moz03CaboDelgado | 2861 | 70 |
| 7 | Moz02Maputo | 2833 | 9 | Moz04Zambizia | 2795 | 70 |
| 7 | Moz02Maputo | 2833 | 10 | Moz05Gaza | 2866 | 73 |
| 7 | Moz02Maputo | 2833 | 11 | Moz06Nampula | 2852 | 96 |
| 7 | Moz02Maputo | 2833 | 12 | V33SAPongola08 | 2768 | 74 |
| 7 | Moz02Maputo | 2833 | 13 | V31SAMookSevere | 2768 | 74 |
| 7 | Moz02Maputo | 2833 | 14 | CLCuGV_AF260241 | 2761 | 70 |
| 7 | Moz02Maputo | 2833 | 15 | CLCuGV_AY036010 | 2764 | 70 |
| 7 | Moz02Maputo | 2833 | 16 | EACMCV_AF112354 | 2802 | 67 |
| 7 | Moz02Maputo | 2833 | 17 | EACMMV_AJ006460 | 2804 | 70 |
| 7 | Moz02Maputo | 2833 | 18 | EACMVUG2_Z83257 | 2799 | 73 |
| 7 | Moz02Maputo | 2833 | 19 | HoLCrV_AY036009 | 2755 | 69 |
| 7 | Moz02Maputo | 2833 | 20 | PepYVV_AY502935 | 2786 | 74 |
| 7 | Moz02Maputo | 2833 | 21 | SACMV_AF155806 | 2800 | 73 |
| 7 | Moz02Maputo | 2833 | 22 | SACMV_AJ422132 | 2800 | 74 |
| 7 | Moz02Maputo | 2833 | 23 | SACMV_AJ575560 | 2800 | 73 |
| 7 | Moz02Maputo | 2833 | 24 | TblCZV_AF350330 | 2767 | 72 |

7 Moz02Maputo 2833 25 ToCSV_AF261885
2766 74
7 Moz02Maputo 2833 26 ToLCMLV_AY502936
2773 73
7 Moz02Maputo 2833 27 ToLCMGV_AJ865338
2777 76
7 Moz02Maputo 2833 28 ToLCMGV_AJ865339
2775 76
7 Moz02Maputo 2833 29 ToLCYTV_AJ865341
2765 76
7 Moz02Maputo 2833 30 ToLCYTV_AJ865340
2768 75
7 Moz02Maputo 2833 31 ToLCSDV_AY044137
2779 73
7 Moz02Maputo 2833 32 ToLCSDV_AY044139
2768 74
7 Moz02Maputo 2833 33 TYLCMlV_AF271234
2782 74
7 Moz02Maputo 2833 34 TYLCV_AJ489258
2781 74
7 Moz02Maputo 2833 35 TYLCV_AY044138
2780 75
7 Moz02Maputo 2833 36 TYLCV_X76319
2790 74
7 Moz02Maputo 2833 37 TYLCV_X15656
2787 74
8 Moz03CaboDelgado 2861 9 Moz04Zambizia
2795 88
8 Moz03CaboDelgado 2861 10 Moz05Gaza
2866 80
8 Moz03CaboDelgado 2861 11 Moz06Nampula
2852 71
8 Moz03CaboDelgado 2861 12 V33SAPongola08
2768 73
8 Moz03CaboDelgado 2861 13 V31SAMookSevere
2768 73
8 Moz03CaboDelgado 2861 14 CLCuGV_AF260241
2761 63
8 Moz03CaboDelgado 2861 15 CLCuGV_AY036010
2764 64
8 Moz03CaboDelgado 2861 16 EACMCV_AF112354
2802 86
8 Moz03CaboDelgado 2861 17 EACMMV_AJ006460
2804 87
8 Moz03CaboDelgado 2861 18 EACMVUG2_Z83257
2799 91
8 Moz03CaboDelgado 2861 19 HoLCrV_AY036009
2755 70
8 Moz03CaboDelgado 2861 20 PepYVV_AY502935
2786 74
8 Moz03CaboDelgado 2861 21 SACMV_AF155806
2800 80
8 Moz03CaboDelgado 2861 22 SACMV_AJ422132
2800 80
8 Moz03CaboDelgado 2861 23 SACMV_AJ575560
2800 80

8 Moz03CaboDelgado 2861 24 TblCZV_AF350330
2767 72
8 Moz03CaboDelgado 2861 25 ToCSV_AF261885
2766 73
8 Moz03CaboDelgado 2861 26 ToLCMLV_AY502936
2773 72
8 Moz03CaboDelgado 2861 27 ToLCMGV_AJ865338
2777 75
8 Moz03CaboDelgado 2861 28 ToLCMGV_AJ865339
2775 76
8 Moz03CaboDelgado 2861 29 ToLCYTV_AJ865341
2765 75
8 Moz03CaboDelgado 2861 30 ToLCYTV_AJ865340
2768 74
8 Moz03CaboDelgado 2861 31 ToLCSDV_AY044137
2779 72
8 Moz03CaboDelgado 2861 32 ToLCSDV_AY044139
2768 72
8 Moz03CaboDelgado 2861 33 TYLCMlV_AF271234
2782 73
8 Moz03CaboDelgado 2861 34 TYLCV_AJ489258
2781 73
8 Moz03CaboDelgado 2861 35 TYLCV_AY044138
2780 74
8 Moz03CaboDelgado 2861 36 TYLCV_X76319
2790 73
8 Moz03CaboDelgado 2861 37 TYLCV_X15656
2787 73
9 Moz04Zambizia 2795 10 Moz05Gaza
2866 85
9 Moz04Zambizia 2795 11 Moz06Nampula
2852 70
9 Moz04Zambizia 2795 12 V33SAPongola08
2768 75
9 Moz04Zambizia 2795 13 V31SAMookSevere
2768 75
9 Moz04Zambizia 2795 14 CLCuGV_AF260241
2761 63
9 Moz04Zambizia 2795 15 CLCuGV_AY036010
2764 63
9 Moz04Zambizia 2795 16 EACMCV_AF112354
2802 79
9 Moz04Zambizia 2795 17 EACMMV_AJ006460
2804 94
9 Moz04Zambizia 2795 18 EACMVUG2_Z83257
2799 86
9 Moz04Zambizia 2795 19 HoLCrV_AY036009
2755 69
9 Moz04Zambizia 2795 20 PepYVV_AY502935
2786 75
9 Moz04Zambizia 2795 21 SACMV_AF155806
2800 84
9 Moz04Zambizia 2795 22 SACMV_AJ422132
2800 85
9 Moz04Zambizia 2795 23 SACMV_AJ575560
2800 85

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | Moz04Zambizia | 2795 | 24 | TblCZV_AF350330 | 2767 | 72 |
| 9 | Moz04Zambizia | 2795 | 25 | ToCSV_AF261885 | 2766 | 74 |
| 9 | Moz04Zambizia | 2795 | 26 | ToLCMLV_AY502936 | 2773 | 74 |
| 9 | Moz04Zambizia | 2795 | 27 | ToLCMGV_AJ865338 | 2777 | 76 |
| 9 | Moz04Zambizia | 2795 | 28 | ToLCMGV_AJ865339 | 2775 | 77 |
| 9 | Moz04Zambizia | 2795 | 29 | ToLCYTV_AJ865341 | 2765 | 78 |
| 9 | Moz04Zambizia | 2795 | 30 | ToLCYTV_AJ865340 | 2768 | 76 |
| 9 | Moz04Zambizia | 2795 | 31 | ToLCSDV_AY044137 | 2779 | 73 |
| 9 | Moz04Zambizia | 2795 | 32 | ToLCSDV_AY044139 | 2768 | 73 |
| 9 | Moz04Zambizia | 2795 | 33 | TYLCMlV_AF271234 | 2782 | 74 |
| 9 | Moz04Zambizia | 2795 | 34 | TYLCV_AJ489258 | 2781 | 75 |
| 9 | Moz04Zambizia | 2795 | 35 | TYLCV_AY044138 | 2780 | 75 |
| 9 | Moz04Zambizia | 2795 | 36 | TYLCV_X76319 | 2790 | 74 |
| 9 | Moz04Zambizia | 2795 | 37 | TYLCV_X15656 | 2787 | 75 |
| 10 | Moz05Gaza | 2866 | 11 | Moz06Nampula | 2852 | 70 |
| 10 | Moz05Gaza | 2866 | 12 | V33SAPongola08 | 2768 | 77 |
| 10 | Moz05Gaza | 2866 | 13 | V31SAMookSevere | 2768 | 78 |
| 10 | Moz05Gaza | 2866 | 14 | CLCuGV_AF260241 | 2761 | 71 |
| 10 | Moz05Gaza | 2866 | 15 | CLCuGV_AY036010 | 2764 | 71 |
| 10 | Moz05Gaza | 2866 | 16 | EACMCV_AF112354 | 2802 | 72 |
| 10 | Moz05Gaza | 2866 | 17 | EACMMV_AJ006460 | 2804 | 85 |
| 10 | Moz05Gaza | 2866 | 18 | EACMVUG2_Z83257 | 2799 | 79 |
| 10 | Moz05Gaza | 2866 | 19 | HoLCrV_AY036009 | 2755 | 74 |
| 10 | Moz05Gaza | 2866 | 20 | PepYVV_AY502935 | 2786 | 78 |
| 10 | Moz05Gaza | 2866 | 21 | SACMV_AF155806 | 2800 | 99 |
| 10 | Moz05Gaza | 2866 | 22 | SACMV_AJ422132 | 2800 | 92 |
| 10 | Moz05Gaza | 2866 | 23 | SACMV_AJ575560 | 2800 | 93 |
| 10 | Moz05Gaza | 2866 | 24 | TblCZV_AF350330 | 2767 | 75 |

10 Moz05Gaza 2866 25 ToCSV_AF261885
2766 77
10 Moz05Gaza 2866 26 ToLCMLV_AY502936
2773 77
10 Moz05Gaza 2866 27 ToLCMGV_AJ865338
2777 80
10 Moz05Gaza 2866 28 ToLCMGV_AJ865339
2775 82
10 Moz05Gaza 2866 29 ToLCYTV_AJ865341
2765 82
10 Moz05Gaza 2866 30 ToLCYTV_AJ865340
2768 79
10 Moz05Gaza 2866 31 ToLCSDV_AY044137
2779 76
10 Moz05Gaza 2866 32 ToLCSDV_AY044139
2768 76
10 Moz05Gaza 2866 33 TYLCMlV_AF271234
2782 78
10 Moz05Gaza 2866 34 TYLCV_AJ489258
2781 77
10 Moz05Gaza 2866 35 TYLCV_AY044138
2780 79
10 Moz05Gaza 2866 36 TYLCV_X76319
2790 74
10 Moz05Gaza 2866 37 TYLCV_X15656
2787 78
11 Moz06Nampula 2852 12 V33SAPongola08
2768 74
11 Moz06Nampula 2852 13 V31SAMookSevere
2768 75
11 Moz06Nampula 2852 14 CLCuGV_AF260241
2761 70
11 Moz06Nampula 2852 15 CLCuGV_AY036010
2764 70
11 Moz06Nampula 2852 16 EACMCV_AF112354
2802 67
11 Moz06Nampula 2852 17 EACMMV_AJ006460
2804 70
11 Moz06Nampula 2852 18 EACMVUG2_Z83257
2799 73
11 Moz06Nampula 2852 19 HoLCrV_AY036009
2755 69
11 Moz06Nampula 2852 20 PepYVV_AY502935
2786 74
11 Moz06Nampula 2852 21 SACMV_AF155806
2800 73
11 Moz06Nampula 2852 22 SACMV_AJ422132
2800 74
11 Moz06Nampula 2852 23 SACMV_AJ575560
2800 75
11 Moz06Nampula 2852 24 TblCZV_AF350330
2767 73
11 Moz06Nampula 2852 25 ToCSV_AF261885
2766 74
11 Moz06Nampula 2852 26 ToLCMLV_AY502936
2773 73

11 Moz06Nampula 2852 27 ToLCMGV_AJ865338
2777 77
11 Moz06Nampula 2852 28 ToLCMGV_AJ865339
2775 77
11 Moz06Nampula 2852 29 ToLCYTV_AJ865341
2765 77
11 Moz06Nampula 2852 30 ToLCYTV_AJ865340
2768 77
11 Moz06Nampula 2852 31 ToLCSDV_AY044137
2779 73
11 Moz06Nampula 2852 32 ToLCSDV_AY044139
2768 74
11 Moz06Nampula 2852 33 TYLCMlV_AF271234
2782 74
11 Moz06Nampula 2852 34 TYLCV_AJ489258
2781 74
11 Moz06Nampula 2852 35 TYLCV_AY044138
2780 75
11 Moz06Nampula 2852 36 TYLCV_X76319
2790 74
11 Moz06Nampula 2852 37 TYLCV_X15656
2787 74
12 V33SAPongola08 2768 13 V31SAMookSevere
2768 96
12 V33SAPongola08 2768 14 CLCuGV_AF260241
2761 71
12 V33SAPongola08 2768 15 CLCuGV_AY036010
2764 71
12 V33SAPongola08 2768 16 EACMCV_AF112354
2802 70
12 V33SAPongola08 2768 17 EACMMV_AJ006460
2804 75
12 V33SAPongola08 2768 18 EACMVUG2_Z83257
2799 74
12 V33SAPongola08 2768 19 HoLCrV_AY036009
2755 72
12 V33SAPongola08 2768 20 PepYVV_AY502935
2786 82
12 V33SAPongola08 2768 21 SACMV_AF155806
2800 77
12 V33SAPongola08 2768 22 SACMV_AJ422132
2800 77
12 V33SAPongola08 2768 23 SACMV_AJ575560
2800 77
12 V33SAPongola08 2768 24 TblCZV_AF350330
2767 84
12 V33SAPongola08 2768 25 ToCSV_AF261885
2766 97
12 V33SAPongola08 2768 26 ToLCMLV_AY502936
2773 78
12 V33SAPongola08 2768 27 ToLCMGV_AJ865338
2777 80
12 V33SAPongola08 2768 28 ToLCMGV_AJ865339
2775 79
12 V33SAPongola08 2768 29 ToLCYTV_AJ865341
2765 80

12 V33SAPongola08 2768 30 ToLCYTV_AJ865340
2768 80
12 V33SAPongola08 2768 31 ToLCSDV_AY044137
2779 76
12 V33SAPongola08 2768 32 ToLCSDV_AY044139
2768 77
12 V33SAPongola08 2768 33 TYLCMlV_AF271234
2782 79
12 V33SAPongola08 2768 34 TYLCV_AJ489258
2781 78
12 V33SAPongola08 2768 35 TYLCV_AY044138
2780 79
12 V33SAPongola08 2768 36 TYLCV_X76319
2790 78
12 V33SAPongola08 2768 37 TYLCV_X15656
2787 77
13 V31SAMookSevere 2768 14 CLCuGV_AF260241
2761 71
13 V31SAMookSevere 2768 15 CLCuGV_AY036010
2764 71
13 V31SAMookSevere 2768 16 EACMCV_AF112354
2802 70
13 V31SAMookSevere 2768 17 EACMMV_AJ006460
2804 75
13 V31SAMookSevere 2768 18 EACMVUG2_Z83257
2799 74
13 V31SAMookSevere 2768 19 HoLCrV_AY036009
2755 74
13 V31SAMookSevere 2768 20 PepYVV_AY502935
2786 82
13 V31SAMookSevere 2768 21 SACMV_AF155806
2800 77
13 V31SAMookSevere 2768 22 SACMV_AJ422132
2800 77
13 V31SAMookSevere 2768 23 SACMV_AJ575560
2800 77
13 V31SAMookSevere 2768 24 TblCZV_AF350330
2767 84
13 V31SAMookSevere 2768 25 ToCSV_AF261885
2766 96
13 V31SAMookSevere 2768 26 ToLCMLV_AY502936
2773 78
13 V31SAMookSevere 2768 27 ToLCMGV_AJ865338
2777 80
13 V31SAMookSevere 2768 28 ToLCMGV_AJ865339
2775 79
13 V31SAMookSevere 2768 29 ToLCYTV_AJ865341
2765 79
13 V31SAMookSevere 2768 30 ToLCYTV_AJ865340
2768 79
13 V31SAMookSevere 2768 31 ToLCSDV_AY044137
2779 77
13 V31SAMookSevere 2768 32 ToLCSDV_AY044139
2768 77
13 V31SAMookSevere 2768 33 TYLCMlV_AF271234
2782 79

13 V31SAMookSevere 2768 34 TYLCV_AJ489258
2781 78
13 V31SAMookSevere 2768 35 TYLCV_AY044138
2780 80
13 V31SAMookSevere 2768 36 TYLCV_X76319
2790 79
13 V31SAMookSevere 2768 37 TYLCV_X15656
2787 78
14 CLCuGV_AF260241 2761 15 CLCuGV_AY036010
2764 95
14 CLCuGV_AF260241 2761 16 EACMCV_AF112354
2802 62
14 CLCuGV_AF260241 2761 17 EACMMV_AJ006460
2804 64
14 CLCuGV_AF260241 2761 18 EACMVUG2_Z83257
2799 64
14 CLCuGV_AF260241 2761 19 HoLCrV_AY036009
2755 84
14 CLCuGV_AF260241 2761 20 PepYVV_AY502935
2786 68
14 CLCuGV_AF260241 2761 21 SACMV_AF155806
2800 67
14 CLCuGV_AF260241 2761 22 SACMV_AJ422132
2800 72
14 CLCuGV_AF260241 2761 23 SACMV_AJ575560
2800 67
14 CLCuGV_AF260241 2761 24 TblCZV_AF350330
2767 71
14 CLCuGV_AF260241 2761 25 ToCSV_AF261885
2766 72
14 CLCuGV_AF260241 2761 26 ToLCMLV_AY502936
2773 67
14 CLCuGV_AF260241 2761 27 ToLCMGV_AJ865338
2777 73
14 CLCuGV_AF260241 2761 28 ToLCMGV_AJ865339
2775 69
14 CLCuGV_AF260241 2761 29 ToLCYTV_AJ865341
2765 73
14 CLCuGV_AF260241 2761 30 ToLCYTV_AJ865340
2768 75
14 CLCuGV_AF260241 2761 31 ToLCSDV_AY044137
2779 70
14 CLCuGV_AF260241 2761 32 ToLCSDV_AY044139
2768 71
14 CLCuGV_AF260241 2761 33 TYLCMlV_AF271234
2782 72
14 CLCuGV_AF260241 2761 34 TYLCV_AJ489258
2781 69
14 CLCuGV_AF260241 2761 35 TYLCV_AY044138
2780 71
14 CLCuGV_AF260241 2761 36 TYLCV_X76319
2790 72
14 CLCuGV_AF260241 2761 37 TYLCV_X15656
2787 70
15 CLCuGV_AY036010 2764 16 EACMCV_AF112354
2802 63

15 CLCuGV_AY036010 2764 17 EACMMV_AJ006460
2804 64
15 CLCuGV_AY036010 2764 18 EACMVUG2_Z83257
2799 64
15 CLCuGV_AY036010 2764 19 HoLCrV_AY036009
2755 86
15 CLCuGV_AY036010 2764 20 PepYVV_AY502935
2786 68
15 CLCuGV_AY036010 2764 21 SACMV_AF155806
2800 67
15 CLCuGV_AY036010 2764 22 SACMV_AJ422132
2800 72
15 CLCuGV_AY036010 2764 23 SACMV_AJ575560
2800 67
15 CLCuGV_AY036010 2764 24 TblCZV_AF350330
2767 71
15 CLCuGV_AY036010 2764 25 ToCSV_AF261885
2766 72
15 CLCuGV_AY036010 2764 26 ToLCMLV_AY502936
2773 67
15 CLCuGV_AY036010 2764 27 ToLCMGV_AJ865338
2777 73
15 CLCuGV_AY036010 2764 28 ToLCMGV_AJ865339
2775 70
15 CLCuGV_AY036010 2764 29 ToLCYTV_AJ865341
2765 74
15 CLCuGV_AY036010 2764 30 ToLCYTV_AJ865340
2768 75
15 CLCuGV_AY036010 2764 31 ToLCSDV_AY044137
2779 70
15 CLCuGV_AY036010 2764 32 ToLCSDV_AY044139
2768 72
15 CLCuGV_AY036010 2764 33 TYLCMlV_AF271234
2782 72
15 CLCuGV_AY036010 2764 34 TYLCV_AJ489258
2781 70
15 CLCuGV_AY036010 2764 35 TYLCV_AY044138
2780 71
15 CLCuGV_AY036010 2764 36 TYLCV_X76319
2790 72
15 CLCuGV_AY036010 2764 37 TYLCV_X15656
2787 70
16 EACMCV_AF112354 2802 17 EACMMV_AJ006460
2804 80
16 EACMCV_AF112354 2802 18 EACMVUG2_Z83257
2799 84
16 EACMCV_AF112354 2802 19 HoLCrV_AY036009
2755 69
16 EACMCV_AF112354 2802 20 PepYVV_AY502935
2786 71
16 EACMCV_AF112354 2802 21 SACMV_AF155806
2800 72
16 EACMCV_AF112354 2802 22 SACMV_AJ422132
2800 73
16 EACMCV_AF112354 2802 23 SACMV_AJ575560
2800 72

16 EACMCV_AF112354 2802 24 TblCZV_AF350330
2767 69
16 EACMCV_AF112354 2802 25 ToCSV_AF261885
2766 70
16 EACMCV_AF112354 2802 26 ToLCMLV_AY502936
2773 71
16 EACMCV_AF112354 2802 27 ToLCMGV_AJ865338
2777 71
16 EACMCV_AF112354 2802 28 ToLCMGV_AJ865339
2775 73
16 EACMCV_AF112354 2802 29 ToLCYTV_AJ865341
2765 73
16 EACMCV_AF112354 2802 30 ToLCYTV_AJ865340
2768 71
16 EACMCV_AF112354 2802 31 ToLCSDV_AY044137
2779 68
16 EACMCV_AF112354 2802 32 ToLCSDV_AY044139
2768 68
16 EACMCV_AF112354 2802 33 TYLCMlV_AF271234
2782 70
16 EACMCV_AF112354 2802 34 TYLCV_AJ489258
2781 70
16 EACMCV_AF112354 2802 35 TYLCV_AY044138
2780 71
16 EACMCV_AF112354 2802 36 TYLCV_X76319
2790 70
16 EACMCV_AF112354 2802 37 TYLCV_X15656
2787 70
17 EACMMV_AJ006460 2804 18 EACMVUG2_Z83257
2799 86
17 EACMMV_AJ006460 2804 19 HoLCrV_AY036009
2755 70
17 EACMMV_AJ006460 2804 20 PepYVV_AY502935
2786 76
17 EACMMV_AJ006460 2804 21 SACMV_AF155806
2800 85
17 EACMMV_AJ006460 2804 22 SACMV_AJ422132
2800 86
17 EACMMV_AJ006460 2804 23 SACMV_AJ575560
2800 86
17 EACMMV_AJ006460 2804 24 TblCZV_AF350330
2767 73
17 EACMMV_AJ006460 2804 25 ToCSV_AF261885
2766 75
17 EACMMV_AJ006460 2804 26 ToLCMLV_AY502936
2773 74
17 EACMMV_AJ006460 2804 27 ToLCMGV_AJ865338
2777 77
17 EACMMV_AJ006460 2804 28 ToLCMGV_AJ865339
2775 78
17 EACMMV_AJ006460 2804 29 ToLCYTV_AJ865341
2765 78
17 EACMMV_AJ006460 2804 30 ToLCYTV_AJ865340
2768 76
17 EACMMV_AJ006460 2804 31 ToLCSDV_AY044137
2779 73

17 EACMMV_AJ006460 2804 32 ToLCSDV_AY044139
2768 74
17 EACMMV_AJ006460 2804 33 TYLCMlV_AF271234
2782 75
17 EACMMV_AJ006460 2804 34 TYLCV_AJ489258
2781 75
17 EACMMV_AJ006460 2804 35 TYLCV_AY044138
2780 76
17 EACMMV_AJ006460 2804 36 TYLCV_X76319
2790 75
17 EACMMV_AJ006460 2804 37 TYLCV_X15656
2787 75
18 EACMVUG2_Z83257 2799 19 HoLCrV_AY036009
2755 71
18 EACMVUG2_Z83257 2799 20 PepYVV_AY502935
2786 74
18 EACMVUG2_Z83257 2799 21 SACMV_AF155806
2800 80
18 EACMVUG2_Z83257 2799 22 SACMV_AJ422132
2800 80
18 EACMVUG2_Z83257 2799 23 SACMV_AJ575560
2800 80
18 EACMVUG2_Z83257 2799 24 TblCZV_AF350330
2767 72
18 EACMVUG2_Z83257 2799 25 ToCSV_AF261885
2766 74
18 EACMVUG2_Z83257 2799 26 ToLCMLV_AY502936
2773 73
18 EACMVUG2_Z83257 2799 27 ToLCMGV_AJ865338
2777 75
18 EACMVUG2_Z83257 2799 28 ToLCMGV_AJ865339
2775 76
18 EACMVUG2_Z83257 2799 29 ToLCYTV_AJ865341
2765 76
18 EACMVUG2_Z83257 2799 30 ToLCYTV_AJ865340
2768 74
18 EACMVUG2_Z83257 2799 31 ToLCSDV_AY044137
2779 72
18 EACMVUG2_Z83257 2799 32 ToLCSDV_AY044139
2768 72
18 EACMVUG2_Z83257 2799 33 TYLCMlV_AF271234
2782 72
18 EACMVUG2_Z83257 2799 34 TYLCV_AJ489258
2781 73
18 EACMVUG2_Z83257 2799 35 TYLCV_AY044138
2780 74
18 EACMVUG2_Z83257 2799 36 TYLCV_X76319
2790 72
18 EACMVUG2_Z83257 2799 37 TYLCV_X15656
2787 73
19 HoLCrV_AY036009 2755 20 PepYVV_AY502935
2786 72
19 HoLCrV_AY036009 2755 21 SACMV_AF155806
2800 74
19 HoLCrV_AY036009 2755 22 SACMV_AJ422132
2800 74

19 HoLCrV_AY036009 2755 23 SACMV_AJ575560
2800 74
19 HoLCrV_AY036009 2755 24 TblCZV_AF350330
2767 72
19 HoLCrV_AY036009 2755 25 ToCSV_AF261885
2766 73
19 HoLCrV_AY036009 2755 26 ToLCMLV_AY502936
2773 72
19 HoLCrV_AY036009 2755 27 ToLCMGV_AJ865338
2777 76
19 HoLCrV_AY036009 2755 28 ToLCMGV_AJ865339
2775 74
19 HoLCrV_AY036009 2755 29 ToLCYTV_AJ865341
2765 77
19 HoLCrV_AY036009 2755 30 ToLCYTV_AJ865340
2768 76
19 HoLCrV_AY036009 2755 31 ToLCSDV_AY044137
2779 74
19 HoLCrV_AY036009 2755 32 ToLCSDV_AY044139
2768 72
19 HoLCrV_AY036009 2755 33 TYLCMlV_AF271234
2782 73
19 HoLCrV_AY036009 2755 34 TYLCV_AJ489258
2781 72
19 HoLCrV_AY036009 2755 35 TYLCV_AY044138
2780 75
19 HoLCrV_AY036009 2755 36 TYLCV_X76319
2790 74
19 HoLCrV_AY036009 2755 37 TYLCV_X15656
2787 72
20 PepYVV_AY502935 2786 21 SACMV_AF155806
2800 78
20 PepYVV_AY502935 2786 22 SACMV_AJ422132
2800 78
20 PepYVV_AY502935 2786 23 SACMV_AJ575560
2800 79
20 PepYVV_AY502935 2786 24 TblCZV_AF350330
2767 83
20 PepYVV_AY502935 2786 25 ToCSV_AF261885
2766 82
20 PepYVV_AY502935 2786 26 ToLCMLV_AY502936
2773 79
20 PepYVV_AY502935 2786 27 ToLCMGV_AJ865338
2777 82
20 PepYVV_AY502935 2786 28 ToLCMGV_AJ865339
2775 80
20 PepYVV_AY502935 2786 29 ToLCYTV_AJ865341
2765 80
20 PepYVV_AY502935 2786 30 ToLCYTV_AJ865340
2768 79
20 PepYVV_AY502935 2786 31 ToLCSDV_AY044137
2779 78
20 PepYVV_AY502935 2786 32 ToLCSDV_AY044139
2768 77
20 PepYVV_AY502935 2786 33 TYLCMlV_AF271234
2782 78

20 PepYVV_AY502935 2786 34 TYLCV_AJ489258
2781 79
20 PepYVV_AY502935 2786 35 TYLCV_AY044138
2780 80
20 PepYVV_AY502935 2786 36 TYLCV_X76319
2790 78
20 PepYVV_AY502935 2786 37 TYLCV_X15656
2787 79
21 SACMV_AF155806 2800 22 SACMV_AJ422132
2800 92
21 SACMV_AF155806 2800 23 SACMV_AJ575560
2800 93
21 SACMV_AF155806 2800 24 TblCZV_AF350330
2767 75
21 SACMV_AF155806 2800 25 ToCSV_AF261885
2766 77
21 SACMV_AF155806 2800 26 ToLCMLV_AY502936
2773 77
21 SACMV_AF155806 2800 27 ToLCMGV_AJ865338
2777 80
21 SACMV_AF155806 2800 28 ToLCMGV_AJ865339
2775 82
21 SACMV_AF155806 2800 29 ToLCYTV_AJ865341
2765 82
21 SACMV_AF155806 2800 30 ToLCYTV_AJ865340
2768 79
21 SACMV_AF155806 2800 31 ToLCSDV_AY044137
2779 76
21 SACMV_AF155806 2800 32 ToLCSDV_AY044139
2768 76
21 SACMV_AF155806 2800 33 TYLCMlV_AF271234
2782 78
21 SACMV_AF155806 2800 34 TYLCV_AJ489258
2781 77
21 SACMV_AF155806 2800 35 TYLCV_AY044138
2780 78
21 SACMV_AF155806 2800 36 TYLCV_X76319
2790 74
21 SACMV_AF155806 2800 37 TYLCV_X15656
2787 77
22 SACMV_AJ422132 2800 23 SACMV_AJ575560
2800 95
22 SACMV_AJ422132 2800 24 TblCZV_AF350330
2767 75
22 SACMV_AJ422132 2800 25 ToCSV_AF261885
2766 76
22 SACMV_AJ422132 2800 26 ToLCMLV_AY502936
2773 77
22 SACMV_AJ422132 2800 27 ToLCMGV_AJ865338
2777 80
22 SACMV_AJ422132 2800 28 ToLCMGV_AJ865339
2775 82
22 SACMV_AJ422132 2800 29 ToLCYTV_AJ865341
2765 82
22 SACMV_AJ422132 2800 30 ToLCYTV_AJ865340
2768 79

22 SACMV_AJ422132 2800 31 ToLCSDV_AY044137
2779 76
22 SACMV_AJ422132 2800 32 ToLCSDV_AY044139
2768 76
22 SACMV_AJ422132 2800 33 TYLCMlV_AF271234
2782 78
22 SACMV_AJ422132 2800 34 TYLCV_AJ489258
2781 77
22 SACMV_AJ422132 2800 35 TYLCV_AY044138
2780 78
22 SACMV_AJ422132 2800 36 TYLCV_X76319
2790 76
22 SACMV_AJ422132 2800 37 TYLCV_X15656
2787 77
23 SACMV_AJ575560 2800 24 TblCZV_AF350330
2767 75
23 SACMV_AJ575560 2800 25 ToCSV_AF261885
2766 76
23 SACMV_AJ575560 2800 26 ToLCMLV_AY502936
2773 77
23 SACMV_AJ575560 2800 27 ToLCMGV_AJ865338
2777 80
23 SACMV_AJ575560 2800 28 ToLCMGV_AJ865339
2775 82
23 SACMV_AJ575560 2800 29 ToLCYTV_AJ865341
2765 83
23 SACMV_AJ575560 2800 30 ToLCYTV_AJ865340
2768 79
23 SACMV_AJ575560 2800 31 ToLCSDV_AY044137
2779 76
23 SACMV_AJ575560 2800 32 ToLCSDV_AY044139
2768 76
23 SACMV_AJ575560 2800 33 TYLCMlV_AF271234
2782 78
23 SACMV_AJ575560 2800 34 TYLCV_AJ489258
2781 78
23 SACMV_AJ575560 2800 35 TYLCV_AY044138
2780 79
23 SACMV_AJ575560 2800 36 TYLCV_X76319
2790 74
23 SACMV_AJ575560 2800 37 TYLCV_X15656
2787 78
24 TblCZV_AF350330 2767 25 ToCSV_AF261885
2766 83
24 TblCZV_AF350330 2767 26 ToLCMLV_AY502936
2773 76
24 TblCZV_AF350330 2767 27 ToLCMGV_AJ865338
2777 77
24 TblCZV_AF350330 2767 28 ToLCMGV_AJ865339
2775 76
24 TblCZV_AF350330 2767 29 ToLCYTV_AJ865341
2765 75
24 TblCZV_AF350330 2767 30 ToLCYTV_AJ865340
2768 76
24 TblCZV_AF350330 2767 31 ToLCSDV_AY044137
2779 75

24 TblCZV_AF350330 2767 32 ToLCSDV_AY044139
2768 77
24 TblCZV_AF350330 2767 33 TYLCMlV_AF271234
2782 77
24 TblCZV_AF350330 2767 34 TYLCV_AJ489258
2781 76
24 TblCZV_AF350330 2767 35 TYLCV_AY044138
2780 78
24 TblCZV_AF350330 2767 36 TYLCV_X76319
2790 76
24 TblCZV_AF350330 2767 37 TYLCV_X15656
2787 76
25 ToCSV_AF261885 2766 26 ToLCMLV_AY502936
2773 78
25 ToCSV_AF261885 2766 27 ToLCMGV_AJ865338
2777 79
25 ToCSV_AF261885 2766 28 ToLCMGV_AJ865339
2775 78
25 ToCSV_AF261885 2766 29 ToLCYTV_AJ865341
2765 78
25 ToCSV_AF261885 2766 30 ToLCYTV_AJ865340
2768 79
25 ToCSV_AF261885 2766 31 ToLCSDV_AY044137
2779 76
25 ToCSV_AF261885 2766 32 ToLCSDV_AY044139
2768 77
25 ToCSV_AF261885 2766 33 TYLCMlV_AF271234
2782 78
25 ToCSV_AF261885 2766 34 TYLCV_AJ489258
2781 78
25 ToCSV_AF261885 2766 35 TYLCV_AY044138
2780 79
25 ToCSV_AF261885 2766 36 TYLCV_X76319
2790 78
25 ToCSV_AF261885 2766 37 TYLCV_X15656
2787 77
26 ToLCMLV_AY502936 2773 27 ToLCMGV_AJ865338
2777 78
26 ToLCMLV_AY502936 2773 28 ToLCMGV_AJ865339
2775 79
26 ToLCMLV_AY502936 2773 29 ToLCYTV_AJ865341
2765 79
26 ToLCMLV_AY502936 2773 30 ToLCYTV_AJ865340
2768 76
26 ToLCMLV_AY502936 2773 31 ToLCSDV_AY044137
2779 75
26 ToLCMLV_AY502936 2773 32 ToLCSDV_AY044139
2768 75
26 ToLCMLV_AY502936 2773 33 TYLCMlV_AF271234
2782 76
26 ToLCMLV_AY502936 2773 34 TYLCV_AJ489258
2781 77
26 ToLCMLV_AY502936 2773 35 TYLCV_AY044138
2780 78
26 ToLCMLV_AY502936 2773 36 TYLCV_X76319
2790 77

26 ToLCMLV_AY502936 2773 37 TYLCV_X15656
2787 78
27 ToLCMGV_AJ865338 2777 28 ToLCMGV_AJ865339
2775 93
27 ToLCMGV_AJ865338 2777 29 ToLCYTV_AJ865341
2765 84
27 ToLCMGV_AJ865338 2777 30 ToLCYTV_AJ865340
2768 84
27 ToLCMGV_AJ865338 2777 31 ToLCSDV_AY044137
2779 80
27 ToLCMGV_AJ865338 2777 32 ToLCSDV_AY044139
2768 80
27 ToLCMGV_AJ865338 2777 33 TYLCMlV_AF271234
2782 81
27 ToLCMGV_AJ865338 2777 34 TYLCV_AJ489258
2781 80
27 ToLCMGV_AJ865338 2777 35 TYLCV_AY044138
2780 82
27 ToLCMGV_AJ865338 2777 36 TYLCV_X76319
2790 81
27 ToLCMGV_AJ865338 2777 37 TYLCV_X15656
2787 80
28 ToLCMGV_AJ865339 2775 29 ToLCYTV_AJ865341
2765 85
28 ToLCMGV_AJ865339 2775 30 ToLCYTV_AJ865340
2768 82
28 ToLCMGV_AJ865339 2775 31 ToLCSDV_AY044137
2779 79
28 ToLCMGV_AJ865339 2775 32 ToLCSDV_AY044139
2768 79
28 ToLCMGV_AJ865339 2775 33 TYLCMlV_AF271234
2782 80
28 ToLCMGV_AJ865339 2775 34 TYLCV_AJ489258
2781 80
28 ToLCMGV_AJ865339 2775 35 TYLCV_AY044138
2780 81
28 ToLCMGV_AJ865339 2775 36 TYLCV_X76319
2790 80
28 ToLCMGV_AJ865339 2775 37 TYLCV_X15656
2787 80
29 ToLCYTV_AJ865341 2765 30 ToLCYTV_AJ865340
2768 90
29 ToLCYTV_AJ865341 2765 31 ToLCSDV_AY044137
2779 79
29 ToLCYTV_AJ865341 2765 32 ToLCSDV_AY044139
2768 78
29 ToLCYTV_AJ865341 2765 33 TYLCMlV_AF271234
2782 80
29 ToLCYTV_AJ865341 2765 34 TYLCV_AJ489258
2781 80
29 ToLCYTV_AJ865341 2765 35 TYLCV_AY044138
2780 81
29 ToLCYTV_AJ865341 2765 36 TYLCV_X76319
2790 81
29 ToLCYTV_AJ865341 2765 37 TYLCV_X15656
2787 80

30 ToLCYTV_AJ865340 2768 31 ToLCSDV_AY044137
2779 77
30 ToLCYTV_AJ865340 2768 32 ToLCSDV_AY044139
2768 79
30 ToLCYTV_AJ865340 2768 33 TYLCMlV_AF271234
2782 80
30 ToLCYTV_AJ865340 2768 34 TYLCV_AJ489258
2781 79
30 ToLCYTV_AJ865340 2768 35 TYLCV_AY044138
2780 81
30 ToLCYTV_AJ865340 2768 36 TYLCV_X76319
2790 81
30 ToLCYTV_AJ865340 2768 37 TYLCV_X15656
2787 79
31 ToLCSDV_AY044137 2779 32 ToLCSDV_AY044139
2768 92
31 ToLCSDV_AY044137 2779 33 TYLCMlV_AF271234
2782 83
31 ToLCSDV_AY044137 2779 34 TYLCV_AJ489258
2781 84
31 ToLCSDV_AY044137 2779 35 TYLCV_AY044138
2780 86
31 ToLCSDV_AY044137 2779 36 TYLCV_X76319
2790 85
31 ToLCSDV_AY044137 2779 37 TYLCV_X15656
2787 84
32 ToLCSDV_AY044139 2768 33 TYLCMlV_AF271234
2782 85
32 ToLCSDV_AY044139 2768 34 TYLCV_AJ489258
2781 83
32 ToLCSDV_AY044139 2768 35 TYLCV_AY044138
2780 86
32 ToLCSDV_AY044139 2768 36 TYLCV_X76319
2790 86
32 ToLCSDV_AY044139 2768 37 TYLCV_X15656
2787 83
33 TYLCMlV_AF271234 2782 34 TYLCV_AJ489258
2781 84
33 TYLCMlV_AF271234 2782 35 TYLCV_AY044138
2780 87
33 TYLCMlV_AF271234 2782 36 TYLCV_X76319
2790 89
33 TYLCMlV_AF271234 2782 37 TYLCV_X15656
2787 84
34 TYLCV_AJ489258 2781 35 TYLCV_AY044138
2780 89
34 TYLCV_AJ489258 2781 36 TYLCV_X76319
2790 92
34 TYLCV_AJ489258 2781 37 TYLCV_X15656
2787 98
35 TYLCV_AY044138 2780 36 TYLCV_X76319
2790 92
35 TYLCV_AY044138 2780 37 TYLCV_X15656
2787 89
36 TYLCV_X76319 2790 37 TYLCV_X15656
2787 92

***Appendix 4: Table 4.1: Percentage nucleotide sequence identities of Bemisia tabaci isolates from Mozambique compared to selected isolates from different countries***

| | **77** | **105** | **75** | **103** | **80** | **92** | **99** | **84** | **91** | **110** | **122** | **119** | **61** | **86** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **22** | 18.2 | 18.4 | 18.8 | 17.9 | 18.8 | 18.7 | 18.4 | 18.1 | 19.5 | 17.9 | 17.8 | 17.8 | 18.7 | 18.2 |
| **28** | 18.6 | 18.5 | 18.4 | 18.6 | 18.7 | 18.5 | 18.7 | 18.3 | 18.7 | 18.5 | 18.5 | 18.5 | 18.5 | 18.6 |
| **39** | 7.7 | 7.8 | 7.8 | 7.8 | 7.8 | 7.9 | 7.8 | 7.8 | 8.4 | 7.9 | 7.9 | 7.7 | 7.7 | 7.7 |
| **36** | 7.4 | 7.5 | 7.5 | 7.5 | 7.9 | 7.5 | 7.4 | 8.0 | 7.5 | 7.4 | 7.6 | 7.4 | 7.4 | 7.5 |
| **37** | 7.8 | 8.0 | 8.0 | 8.3 | 9.8 | 8.1 | 7.8 | 7.8 | 8.4 | 7.8 | 8.0 | 8.0 | 7.8 | 8.9 |
| **38** | 8.9 | 8.9 | 9.4 | 9.4 | 7.8 | 9.3 | 9.1 | 9.3 | 9.5 | 9.3 | 9.2 | 9.0 | 9.0 | 9.1 |
| **39** | 7.7 | 7.7 | 7.8 | 7.8 | 2.6 | 8.4 | 8.3 | 7.9 | 8.4 | 7.8 | 7.9 | 7.7 | 7.7 | 7.7 |
| **42** | 2.9 | 2.9 | 2.7 | 2.8 | 1.9 | 1.7 | 2.6 | 2.8 | 2.6 | 2.9 | 2.9 | 2.9 | 2.7 | 2.9 |
| **43** | 1.9 | 1.3 | 1.4 | 1.0 | 2.5 | 2.7 | 1.0 | 1.0 | 1.8 | 2.4 | 0.8 | 1.3 | 1.2 | 1.4 |
| **44** | 2.4 | 2.4 | 2.2 | 2.2 | 3.4 | 3.2 | 2.2 | 2.4 | 2.2 | 2.4 | 2.5 | 2.4 | 2.4 | 2.4 |
| **45** | 3.0 | 3.0 | 2.9 | 2.9 | 1.2 | 4.5 | 2.9 | 3.0 | 3.3 | 3.6 | 2.9 | 3.0 | 3.0 | 3.0 |
| **48** | 2.0 | 2.0 | 1.7 | 2.0 | 1.3 | 2.0 | 1.7 | 1.8 | 1.8 | 1.8 | 2.0 | 1,8 | 2.0 | 1.8 |

- Sequence pair distances of MZtreeKeySASOct06 ClustalV (Weighted)
- Key for table next page

**Table 4.2. : Samples numbers and their correlating locations of Whitefly sequences which were used in determining the percentage nucleotide sequence identities.**

| Sample No | Name (Location) | No | Name (Location) |
|---|---|---|---|
| 77 | M16-1coi-1 (Mozambique) | 28 | Uganda Sweetpotato (Uganda) |
| 105 | M61-1coi-1(Mozambique) | 39 | Cam Penda-Boko W16 (Cameroon) |
| 75 | M155-1coi-1(Mozambique) | 36 | Cam Ayos1 WO2 (Cameroon) |
| 103 | M56-2coi-1(Mozambique) | 37 | Cam Ayos2 WO3 (Cameroon) |
| 80 | M169-1coi-1(Mozambique) | 38 | Cam Bafous W13 (Cameroon) |
| 92 | M41-1coi-1(Mozambique) | 39 | Cam Penda-Boko W16 (Cameroon) |
| 99 | M53-1coi-1(Mozambique) | 42 | SA Lucia 1(St Lucia) |
| 84 | M25-1coi-1(Mozambique) | 43 | SA Lucia 9(St Lucia) |
| 91 | M41-1coi-1(Mozambique) | 44 | Swa Map 3(Swaziland) |
| 110 | M64-3coi-((Mozambique) | 45 | Swa Map 4 Swaziland) |
| 119 | M9-3coi-1(Mozambique) | 48 | Zamb Mulim 5(Zambia) |
| 61 | M107-1coi-1(Mozambique) | 28 | Uganda Sweetpotato(Uganda) |
| 86 | M33-1coi-1(Mozambique) | 39 | Cam Penda-Boko W16(Cameroon) |
| 22 | SA 99 (B) (South Africa) | 36 | Cam Ayos1 WO2( Cameroon) |
| 37 | Cam Ayos2 WO3( Cameroon) | | |
| 38 | Cam Bafous W13( Cameroon) | | |
| 39 | Cam Penda-Boko W16 (Cameroon) | | |
| 42 | SA Lucia 1(St Lucia) | | |
| 43 | SA Lucia 9(St Lucia) | | |
| 44 | Swa Map 3(Swaziland) | | |
| 45 | Swa Map 4 Swaziland) | | |
| 48 | Zamb Mulim 5(Zambia) | | |