BIODIVERSITY AND STRAIN

DIFFERENTIATION OF CASSAVA-INFECTING

GEMINIVIRUSES AND *BEMISIA TABACI* IN

SOUTHERN AFRICA

Shaun David Berry

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

Johannesburg, 2001

DECLARATION

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

_____Berry_____

(Shaun David Berry)

_____30th____ day of ____March_____ 2001

"The only truly failed trial is one that fails to produce an interpretable result"

Author unknown

ABSTRACT

Cassava mosaic disease is prevalent in Africa and significantly affects the growth and yield of cassava. This disease is caused by a number of whitefly-transmitted begomoviruses. The aims of this study were to establish the identity of cassava begomoviruses in southern Africa and to develop assays for their differentiation.

Using primers that target the highly conserved core region of the coat protein gene it was possible to identify and establish the geographical distribution and relatedness of cassava begomoviruses in 6 countries within southern Africa. It was found that *African cassava mosaic virus* occurred in five countries (except Angola), *East African cassava mosaic virus* was present in all countries (except Zambia) and *South African cassava mosaic virus* was present in South Africa and Swaziland. In addition, this study reports for the first time in southern Africa, the Ugandan variant virus (UgV) which occurs frequently in mixed infections with other cassava-infecting begomoviruses.

Bemisia tabaci (Gennadius) is the vector of begomoviruses that cause cassava mosaic disease (CMD) in Africa and India. The taxonomy of the *B. tabaci* complex is problematic, making it unclear whether more than one variant or 'type' of the vector is involved in the transmission of cassava-infecting begomoviruses. Phylogenetic analysis of mitochondrial COI gene sequences revealed that *B. tabaci* colonising cassava in Africa form 3 distinct clades (clade 1: Mozambique, South Africa, Swaziland, Zambia; clade 2:

Cameroon; clade 3: Zimbabwe). These results indicate that topotypes within *B. tabaci* vector populations from cassava exist and suggest a geographic basis for the separation of cassava-colonising *B. tabaci* in Africa.

New cassava viruses and viral strains continue to be discovered and simple, rapid and sensitive techniques are needed for screening of cassava plantations. Here we report on the development of a heteroduplex mobility assay (HMA) for differentiating cassava-infecting begomoviruses. The HMA profiles were able to differentiate four different viral species and eleven different virus strains, and showed a good correlation with sequencing results and phylogenetic comparisons with other sequenced cassava viruses. This technique was found to be sensitive and rapid and had the added advantage of being able to detect mixtures of viruses in field-grown cassava.

Current serological methods and antibodies are limited in their usefulness and specificity and new antibodies need to be developed to detect all the possible viral species. The viability of using phage antibodies to detect begomoviruses proved promising as a number of phage clones were isolated and characterised. These clones, when used in combination, were able to differentiate between several cassava-infecting begomoviruses. However a number of improvements on this technique would need to be implemented before it became an acceptable method for producing antibodies to identify and distinguish between cassava begomovirus species and strains.

ACKNOWLEDGEMENTS

CONTENTS

CHAPTER ONE- LITERATURE REVIEW

CHAPTER TWO- MOLECULAR EVIDENCE FOR DIVERSE POPULATIONS OF CASSAVA-INFECTING BEGOMOVIRUSES IN SOUTHERN AFRICA

CHAPTER THREE- MOLECULAR EVIDENCE FOR DISTINCT *BEMISIA TABACI* GEOGRAPHIC GENOTYPES FROM CASSAVA IN AFRICA

CHAPTER FOUR- DIFFERENTIATION OF CASSAVA-INFECTING BEGOMOVIRUSES USING HETERODUPLEX MOBILITY ASSAYS

CHAPTER FIVE- PHAGE ANTIBODIES TO SOUTH AFRICAN CASSAVA MOSAIC VIRUS

CHAPTER SIX- SUMMARY

CHAPTER SEVEN- REFERENCES

LIST OF FIGURES

page

LIST OF TABLES

LIST OF ABBREVIATIONS

AbMV	abutilon mosaic virus
ACMD	African cassava mosaic disease
ACMV	African cassava mosaic virus
a.s.l.	above sea level
BCTV	beet curly top virus
CBVs	cassava-infecting begomoviruses
CCP	core coat protein
CLV	cassava latent virus
CMD	cassava mosaic disease
CP	coat protein
CR	common region
DNA	deoxyribonucleic acid
DGGE	denaturing gradient gel electrophoresis
dsDNA	double stranded DNA
EACMV	East African cassava mosaic virus
EDTA	ethylenediamine tetra-acetic acid
ELISA	enzyme-linked immunosorbent assay
HIV	Human immunodeficiency virus
HMA	heteroduplex mobility assay
ICMV	Indian cassava mosaic virus
IITA	International Institute of Tropical Agriculture
kB	kilobase
M	molar

MAb	monoclonal antibody
MDE™	mutation detection enhancement™
mtCOI	mitochondrial cytochrome oxidase I
nt	nucleotide
ORF	open reading frame
PAGE	polyacrylamide gel elecrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PNRS	prunus necrotic ringspot virus
PVY	potato virus Y
RAPD	random amplified polymorphic DNA
Rep	replication-associated protein
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rDNA	ribosomal DNA
SA	South Africa
SACMV	South African cassava mosaic virus
SCR	Scottish Crop Research
SSCP	single stranded conformational polymorphism
ssDNA	single stranded DNA
scFv	single chain Fv antibody
SDS	sodium dodecyl sulphate
SiGMV	Sida golden mosaic virus
TGMV	tomato golden mosaic virus
TYLCV	tomato yellow leaf curl virus

TNA   total nucleic acid

UgV   Ugandan variant virus

UK   United Kingdom

USA   United States of America

WTGs   whitefly-transmitted geminiviruses

LIST OF CONFERENCES ATTENDED

INTERNATIONAL CONFERNCES:

1. " Biodiversity of cassava-infecting begomoviruses in South Africa"

   S. Berry, C. Jericho, G. Thompson and M.E.C. Rey

   International Union of Microbiological Societies (IUMS)

   11th International Congress of Virology

   9-13 August 1999

   Sydney, Australia

2. "A study of whitefly populations on cassava in southern Africa"

   M.E.C. Rey, **S. Berry,** G.K. Banks, P. Markham and J.K. Brown

   International Symposium on Tropical Root and Tuber Crops

   19-22 January 2000

   Thiruvanathapuram, India

LOCAL CONFERENCES:

1. "Diversity of cassava-infecting begomoviruses in southern Africa"

   S. Berry and M.E.C. Rey

   BIOY2K Combined Millennium Meeting: South African Society of Plant

   Pathology (SASPP) Congress

   23-28 January 2000

   University of Rhodes, Grahamstown, South Africa

2. "Biodiversity of cassava begomoviruses and their whitefly vectors"

M.E.C Rey, J.K. Brown and **S. Berry**

39th South African Society for Plant Pathology (SASPP) Congress

21-24 January 2001

Greenway Woods, Nelspruit, South Africa

CHAPTER ONE

INTRODUCTION

1.1 Cassava: History, Geography and Uses

Cassava (*Manihot esculenta* Crantz) is a shrubby, perennial plant grown in the tropical and subtropical regions of the world, particularly Africa, Asia and Latin America. Cassava is known by a number of different names depending on where it is grown: ubi kettella or kaspe (Indonesia), tapioca (India, Malaysia), manioca, rumu or yucca (Latin America), mandioca or aipim (Brazil) and ndumbula (South Africa) (Daphne 1990, Essiedu 1995).

Recent studies on the origins of cassava suggest that cassava originated from a wild subspecies, *Manihot flabeliifolia*, found in the southern border of the Amazon River Basin in Brazil (Olsen and Shaal 1999). Cassava has been grown for thousands of years in Central and South America but was unknown in the Old World (Europe, Africa and Asia) until fairly recently i.e. until the arrival of the Portuguese settlers. In 1494 Pierre Martyr noted that the indigenous tribes of Brazil were using the "poisonous roots" of the yucca plant for making bread (FAO website: http://www.fao.org). The Portuguese developed a liking for this 'new' plant and began using cassava flour as a provision on their ships, particularly those using the trade routes between Africa, Brazil and Europe (Mpoko Bokanga, IITA website: http://www.iita. org).

Cassava is believed to have been introduced to West Africa in the 16th Century, East Africa and their surrounding islands in the 18th Century and India and Sri Lanka 50 years later (Fauquet and Fargette 1990). The diffusion of cassava from the coastal regions of West Africa to the inland was initially very slow. Large scale spread and dissemination of cassava only occurred in

the late 19th and 20th centuries as a result of the role of colonial governments in encouraging cassava cultivation (http://www.iita.org/West Africa).

In Central Africa, cassava was first introduced to the Congo region. It was readily adopted as an alternative food source and soon replaced millet, yams and plantain as the principal staple food. Trade in cassava and cassava-products increased in the 19th century particularly along the Alima river in the Congo Republic. At the same time cassava was introduced in the Central Africa Republic, Cameroon, Equatorial Guinea and Gabon. The dissemination of cassava in Central Africa was most significantly influenced by river transport and tribal movements (http://www.iita.org/Central Africa).

In East Africa, cassava was introduced to all the major Portuguese trading posts such as Mozambique, Benguela, Sofala, Zanzibar and Mombasa during the 17th and 18th centuries. Environmental and cultural barriers were the main obstacles in slowing cassava movement from these initial points. Cassava was spread mainly as a result of the movements of Europeans and Arabs migrating towards the interior who saw the value of cassava as a safeguard against famine. Cassava has since spread to many east African countries including Malawi, Zimbabwe, Zambia and South Africa.

Since it's introduction to Africa many years ago, cassava has changed it's status from one of a 'part-time' or secondary crop to that of an extremely important and essential crop. Nowadays cassava is the most important staple food crop grown in Africa, it's production exceeds that of other traditional African crops such as maize and yams. It provides food for over 200 million people and more cassava is now grown in Africa than anywhere else in the world (Dahniya 1994). Today, cassava production in Africa represents 45%

of the total world production. In 1985, 57 million tons of cassava roots were harvested from 7.5 million hectares of land (Fauquet and Fargette 1990), 6 years later, 68.9 million tons of roots was produced on 8.9 million hectares (Dahniya 1994).

Cassava is a short-lived perennial, woody shrub that can grow up to 5m tall and produces between 5-10 enlarged tuberous roots (Nweke *et al.* 1995). The average yield of cassava roots in Africa is 10-15 tons/hectare but yields of up to 70 tons/ha have been reported at some agricultural research stations. The tubers can be left in the ground for up to 3 years and thus make an excellent food security source, particularly in areas of erratic rainfall. The tubers are an important source of carbohydrates, however, processing of the roots depends on the cyanide content. 'Bitter' tasting roots which are usually high in cyanide have to be dried, cut into thin slices, crushed and seaved into flour whereas 'sweeter' varieties (lower in cyanide content) only have to be peeled and boiled. The leaves of cassava are high in protein and vitamins A and C and are consumed as a vegetable (Enete *et al.* 1995).

Cassava is a typical tropical plant, requiring a warm and humid climate for growth. It is usually grown in tropical areas between 30°S and 30°N where temperatures range from 25-35°C and little or no frost occurs. Optimal growth occurs in areas where the rainfall varies from 500-5000 mm/year and at low altitude (up to 1500 m a.s.l.) (Alleman *et al.* 1996). Cassava is ideally suited for growth in Africa due to it's ability to grow in a wide variety of agro-ecological conditions, adaptability to a wide variety of soils, drought tolerance and relative resistance to weeds and insect pests and ease of cultivation (Dahniya 1994, Christiaensen *et al.* 1995).

In contrast to Africa, where the majority of the cassava is grown by subsistence farmers for use as a food source, other tropical countries of Asia and Latin America grow cassava for industrial or export purposes. In Brazil (which is the single largest cassava-growing country in the world) cassava is grown commercially as a fuel source alternative. Cassava roots are rich in starch and a ton of cassava roots can yield 165-180 litres of alcohol (Nestel 1993). In Thailand almost all the cassava grown is used for export. Europe is their largest market where cassava is imported as dried roots and used primarily for compound animal feed (CGIAR website: http://www.warda.org /areas/cassava). The USA and Japan also import a significant amount of cassava in the form of starch for use in their food, textile and paper industries. The by-products of cassava are used in the chemical industry for making products such as monosodium glutamate, citric acid, mannitol, sorbitol and syrup.

In South Africa, over 60% of the cassava cultivated is processed into flour, vegetable or crisps for human consumption, a large percentage of the rest is used for animal feed and the remainder is transformed into secondary products such as starch and alcohol (Daphne 1980). Although the majority of cassava is grown by subsistence farmers in the rural areas of Venda, Gazankulu, KaNgwane and Kwazulu-Natal, this may change as recent attempts have been made in the growth and utilisation of cassava as a commercial crop for starch production (Sparks 1996).

1.2 African Cassava Mosaic Disease

Typical yields of cassava plantations in Africa are generally lower than that in the more developed countries. This is further exacerbated by other factors (abiotic and biotic) which have been found to affect cassava productivity and yield. Approximately 18 different viruses have been found that infect cassava, 8 of these occur in Africa, 8 in South America and 2 in Asia (Thresh *et al.* 1994a). The most significant disease to affect cassava in Africa is African Cassava Mosaic Disease (ACMD) caused by a group of whitefly-transmitted begomoviruses. Other diseases that affect cassava to a lesser extent are vascular bacteriosis, anthracnoses and damage due to pests such as mealybug, green spider mite and the variegated grasshopper (Fauquet and Fargette 1990).

The most characteristic feature of ACMD are the symptoms which result in mosaic colouring of leaves and leaf curling. The symptoms can range from mild (slight mosaic with little/no stunting of leaves) to severe (severe mosaic and stunting of plant and leaves) depending on the cultivar and age of the plant. Similar mosaic diseases of cassava also occur in India and South America although in these areas the disease is caused by *Indian cassava mosaic virus* (geminivirus) and *Cassava common mosaic virus* (potexvirus) respectively. ACMD was first reported in Tanzania in 1894 and later also reported in countries in East Africa (1920's) and West Africa (1930's).

Today ACMD is found in all cassava-growing countries in Africa. The disease was assumed to be caused by a virus since it could be transmitted by grafting and a whitefly vector (*Bemisia tabaci* Gennadius) although this was

not confirmed until fairly recently. In 1975 Bock managed to inoculate cassava and *Nicotiana clevelandii* with sap from infected cassava and viral particles were obtained this way. The exact role of this virus in causing disease was uncertain as it could be obtained from infected cassava plants from western Kenya, Tanzania and Uganda but **not** from infected cassava from coastal Kenya. Hence the virus was initially called cassava latent virus (CLV) since it was not assumed to be the causal agent of ACMD. This was later resolved when successful transmissions were made to *Nicotiana benthamiana* from cassava samples from **both** western and coastal Kenya. This virus causing ACMD was later characterised and renamed *African cassava mosaic virus* (ACMV) (Bock and Woods 1983).

The increase in the incidence of ACMD in Africa is thought to be as a result of the rapid expansion and intensity of cassava cultivation in many countries since the beginning of the 20th century. Reliable estimates of the exact losses attributable to ACMD are difficult to calculate but it has been suggested that losses due to ACMD currently range from 15-24% which in 1994 was equivalent to a loss of 73 million tons adding up to a loss of $1.3-2.3 billion per annum (Otim-Nape *et al.* 1998b). The effect of ACMD on the yield of a plant is complex and depends on a number of factors such as: resistance/susceptibility of cultivar, age of plant when infected, mode of infection (infected stem cuttings vs. whitefly transmission), climatic conditions and the virulence of the virus. Results of recent surveys of the incidence of ACMD (Table 1.1) suggest that ACMD occurs in all areas where cassava is grown although the incidences and severity of symptoms varies within and between regions.

Table 1.1: Surveys of the incidence of African cassava mosaic disease in certain African countries (Thresh *et al.* 1998)

Ref.	Country (organisation)	Year	No. Fields assessed	Cassava area (mill. ha.)	Ha./field surveyed	CMD Incidence
1.	India (UAS, Bangalore)	1988	180	0.23	1,305	19
	Andhra Pradesh	1988	10	0	0	<1
	Karnataka	1988	40	0	0	5
2.	Uganda (NARO)	1990-92	1,350	0.36*	267	57
3.	Chad (AID)	1992	48	0.07	1,458	40
4.	Malawi (National)	1993	450	0.07	156	21
5.	Tanzania (National)	1993-94	325	0.69	2,123	26
6.	Ghana (ESCaPP)	1994	40	0.61	15,250	72
6.	Benin (ESCaPP)	1994	31	0.14	4,516	53
6.	Cameroon (ESCaPP)	1994	61	0.08	1,311	67
7.	Nigeria (IITA)	1994	93	2.00*	21,505	55
6.	Nigeria (ESCaPP)	1994	111	2.00*	18,018	82
8.	Uganda (NARO)	1994	1,200	0.38*	317	65
	All African countries: total (mean)	1990-1994	3,709	4.03*	(1,087)	(54)+

* Assuming single values for Uganda and Nigeria.
+ In calculating the overall mean incidence the results of the first Uganda survey were omitted, the results of the two Nigerian surveys in 1994 were combined and the country values were weighted according to the area of cassava grown.
Sources:1. Mathew (1989), 2. Otim-Nape *et al.* (1998), 3. Johnson (1992), 4. Nyirenda *et al.* (1993), 5. Legg & Raya (1998), 6. ESCaPP (Yaninek *et al.*, 1994; Wydra & Msikita, 1998 and unpublished), 7. L.C. Dempster, unpublished, 8. G.W. Otim-Nape *et al.*, unpublished.

Figures available for Mozambique suggest that cassava mosaic disease occurs in all areas where cassava is grown although the incidence is generally low (21%) and the symptoms are generally mild (M. Thresh, personal communication). Studies on the distribution of ACMD in South Africa (Trench and Hemmes 1985) found a high incidence of disease in the central to north-eastern parts of the country with the incidence decreasing until relatively absent further south.

It has been difficult to correlate the results of various epidemiological studies on ACMD since the studies were carried out at different times, on different cultivars, with different environmental conditions and different experimental conditions. Cassava is grown in a wide range of different, diverse agro-ecological conditions hence significant relationships between climatic factors and cassava growth and infection are more complex than

initially thought. Particularly since the environmental conditions affect, to different degrees, the plant, virus and vector.

Results from studies in the Ivory Coast suggest an annual periodicity of disease with large year-to-year variations suggesting that one or more climatic factors are involved. At Adiopodoume where the climate is hot and humid for most of the year, temperature was found to be the chief climatic factor, other parameters such as rainfall and humidity were important but less significant. High temperatures positively impacted on whitefly population numbers and activity as well as influencing virus multiplication in the plant (Fargette *et al.* 1993). Similar results were obtained at Kiwanda in Tanzania where spread was greatest at high temperatures and least at low temperatures (Fargette and Thresh 1994). Adiopodoume and Kiwanda have humid conditions all year round thus it is not possible to assess the effects of rainfall on ACMD spread at these areas. Other studies at Toumodi (Ivory Coast) and Ibadan (Nigeria) where rainfall is limited found that most rapid spread occurred during the rainy seasons (when temperature and rainfall were high), signifying that rainfall-associated factors also influence disease spread (Fargette and Thresh 1994). In Uganda where 2 main rainy seasons occur, it was found that the peaks of ACMD incidence coincided with peaks in rainfall (Adipala *et al.* 1998). Other factors such as planting date, age of crop, cropping intensity, wind speed and direction and altitude also play a role in disease incidence.

It is difficult to establish any valid generalisations regarding climatic factors and disease spread although the evidence suggests that spread is most rapid in humid low-lying areas where temperature and soil moisture

conditions favour crop growth and whitefly reproduction. There tends to be less spread in areas with cooler temperatures, less rain and higher in altitude (>1000 m a.s.l.) (Thresh *et al.* 1994b).

In some parts of Africa, mosaic disease occurs so regularly and infects so many plants that some subsistence farmers in Africa understandably assume that ACMD is inevitable and unavoidable. Since cassava is a vegetatively propagated crop it is natural that without some form of control the disease incidence will ultimately reach 100% with successive plantings. Since the discovery of ACMD in 1894 much work has been done on investigating various control measures in preventing/slowing disease spread. The 2 most important forms of control involve (a) planting of resistant varieties and (b) phytosanitation.

(a) Use of resistant cassava varieties

Increasing reports of ACMD in various parts of Africa in the 1930s led Storey and Nichols to begin breeding cassava plants for resistance. Initially they interbred cassava plants but the levels of resistance obtained were not sufficient for areas of high innoculum pressure. Higher levels of resistance were obtained when cassava varieties were bred with *Manihot glaziovii* (rubber plant) and then backcrossed to cassava and inter-crossed to restore the original cassava qualities (Fargette and Fauquet 1990). This process has formed the basis of the cassava breeding programme at IITA, Ibadan, Nigeria (Herren 1994). There are several features that contribute towards resistance:

- plants are resistant but not immune to infection, this is important if the spread of infection is primarily by whitefly vectors. Of little use if the crop is vegetatively propagated as all plants will eventually become diseased.

- symptoms are slow to develop and generally inconspicuous.

- virus content of plants tends to be low, thus they become poor source of innoculum for further spread by vectors.

- virus does not become fully systemic in the plant. This form of non-systemicity or 'reversion' is important in preventing the gradual build-up of infection over successive generations and provides a source of 'clean' planting material for future use.

Resistant cultivars exhibit mild symptoms and low virus titres. The correlation between virus content, symptom intensity, disease incidence and recovery suggest that different genetic mechanisms apply for resistance to vector and virus (Thresh *et al.* 1994b, Fargette *et al.* 1996, Gibson and Otim-Nape 1997).

(b) Phytosanitation

The aim of phytosanitation is to remove the source/s of innoculum before and after planting to prevent any further spread of disease particularly by whitefly vectors.

There are 3 aspects to phytosanitation:

- Removal of all diseased plants around and within the area before planting

- Planting of virus-free planting material

- Removal (roguing) of diseased plants within crop stands

The success of phytosanitation depends on a number of factors. In some areas it is difficult or sometimes impossible to obtain virus-free planting material. Removal of infected planting material is particularly difficult to implement and encourage since farmers are naturally reluctant to remove any sources of potential food. However phytosanitation has proved to be highly effective particularly in areas of high infection pressure where spread of disease is primarily by whiteflies and where less resistant cultivars are grown (Otim-Nape *et al.* 1994, Thresh *et al.* 1994b).

The effect of using a combination of the above 2 control methods can be seen by results of modelling the temporal spread of ACMD in Adiopodoume (Fargette *et al.* 1994). With no reversion or cutting selection the disease incidence progressively increased over successive generations till it reached 100%. However with the adoption of reversion and/or cutting selection, the disease incidence reaches an equilibrium level below 100% at which the effects of reversion/and or cutting selection balanced new viral infections by whiteflies (Fargette and Vie 1995).

1.3 Begomoviruses: Taxonomy, Gene Organisation and Function

The viruses that cause ACMD in cassava are members of the *Geminiviridae* family of plant viruses. Geminiviruses are recognised by their characteristic twinned, isometric capsid morphology and their circular ssDNA genomes (Goodman 1977, Harrison 1977). Geminiviruses are economically important

plant pathogens that are responsible for causing a wide range of diseases in a diverse array of crops in many tropical regions of the world. Studies on geminiviruses over the last 2 decades have revealed that, although these viruses have many features in common, there are also discernible differences with regard to insect vector, host range and genome structure. This has resulted in the formation of 4 subgroups within the *Geminiviridae* family (Francki *et al.* 1991, Padidam *et al.* 1995, Fauquet *et al.* 2000):

	Mastrevirus	*Curtovirus*	*Topocuvirus*	*Begomovirus*
Type species	*Maize streak virus*	*Beet curly top virus*	*Tomato pseudo-curly top virus*	*Bean golden mosaic virus*
Insect vector	Leafhopper	Leafhopper	Treehopper	Whitefly
Host range	Narrow host range, mainly infect *Graminae*	Wide host range, 44 plant families	Narrow host range, infect dicotyledons	Narrow host range, infect dicotyledons
Transmission	Not mechanically transmitted	Possible by mechanical transmission	Not known	Mechanically transmissible
Genome	~2.7 kB, monopartite, encodes 4 genes (2 viral and 2 complementary)	~3 kB, monopartite, has 6 ORFs (2 viral and 4 complementary)	~2.8 kB, monopartite, has 6 ORFs (2 viral and 4 complementary)	~2.8 kB, majority are bipartite (have DNA-A and DNA-B), some monopartite

A number of investigations examining relationships between geminiviruses using phylogenetic analyses have been conducted. Alignments of the coat protein (CP) and replication-associated proteins (Rep) of 16 geminiviruses

revealed that phylogenetic trees constructed from CP alignments correlated with vector specificity i.e. had 2 separate branches for whitefly transmitted and leafhopper transmitted geminiviruses. Alignment of Rep genes also had 2 branches which correlated with host specificity i.e. viruses separated according to whether they infected monocotyledonous or dicotyledonous plants, with a further division of dicotyledonous-infecting viruses into those that are found in the New World and those in the Old World (Howarth and Vandemark 1989). More recent and more extensive phylogenetic analyses comparing the whole genomes of even more geminiviruses have confirmed these results and have shown that geminiviruses form 2 distinct clusters, those that are leafhopper-transmitted and infect monocots and those that are whitefly-transmitted and infect dicots (Rybicki 1994, Padidam *et al.* 1995).

Viruses belonging to the genus, *Begomovirus*, are responsible for cassava mosaic disease in Africa. The majority of the begomoviruses are bipartite (have 2 genomic components, DNA-A and DNA-B), although quite a number of monopartite viruses have been found, particularly those causing tomato leaf curl disease (Czosnek *et al.* 1988, Navot *et al.* 1991, Dry *et al.* 1993).

In 1983 Stanley and Gay discovered that ACMV had 2 similar sized circular DNA components which they named DNA-1 (2779 nt) and DNA-2 (2724 nt) respectively. Later it was found that **both** genomic components were required for systemic infection within the plant (Stanley 1983) and for transmission by the whitefly vector (Liu *et al.* 1997). DNA-A (new name for DNA-1) encodes all the gene products necessary for viral replication and

plants and produce viral particles in the absence of DNA-B, while DNA-B can replicate only in the presence of DNA-A (Klinkenberg and Stanley 1990).

Table 1.2: Geminivirus Gene Products and their Functions (Hanley-Bowdoin *et al.* 1999)

	Subgroup I	Subgroup II	Subgroup III
Replication			
• Plus-strand origin	Long intergenic region (LIR)	5′ intergenic region (IR)	5′ intergenic region (IR) or common region (CR)
• Minus-strand origin	Short intergenic region (SIR)	Unknown	Unknown but may be in common region
• Essential factor	C1:2 (Rep)	C1 (Rep)	AL1 (C1, Rep)
• Accessory factor	C1′ (RepA)	C3 and V3 (?)	AL3 (C3)
Transcription			
• Promoters	Long intergenic region (LIR)	5′ intergenic region (IR)	5′ intergenic region (IR) or common region (CR)
• Activators	C1:2 (Rep)	C2 (TrAP)	AL2 (C2, TrAP)
• Repressors			AL1 (C1, Rep) and AL4 (C4)
Host activation	C1′ (RepA)	C4	AL1 (C1, Rep)
Encapsidation	V2	V2	Double component: AR1 (V1)
			Single component: V2
Movement	V1 and V2	V1 and V2	Double component: BL1 (BC1) and BR1 (BV1)
			Single component: V1 and V2
Symptoms		C4	BL1 (BC1)

The 2 genomic components are similar in size but differ in their sequences and gene products except for a 200 bp region (known as the common region/intergenic region) which is identical for the 2 components. The DNA-A of begomovirus*s* has 6 ORFs, 2 overlapping sequences on the viral strand and 4 on the complementary strand. The most conserved region of the genome is that which codes for the coat protein (CP) gene. The CP is responsible for encapsidating the virus genome but is not required for systemic spread or symptom development in bipartite begomoviruses (Stanley and Townsend 1986). ACMV CP deletion mutants are still infectious after agroinoculation although the onset of symptoms is slightly delayed and symptoms attenuated (Klinkenberg *et al.* 1989). In contrast, the CP is essential for viral spread and symptom development for monopartite

infectious after agroinoculation although the onset of symptoms is slightly delayed and symptoms attenuated (Klinkenberg *et al.* 1989). In contrast, the CP is essential for viral spread and symptom development for monopartite begomoviruses (Padidam *et al.* 1996). The CP also plays an important role in vector transmissibility of the virus. Replacing the CP of ACMV (whitefly transmitted) with that of BCTV (leafhopper transmitted) resulted in transmission of a chimeric clone of ACMV by a leafhopper vector (Briddon *et al.* 1990). Similarly replacing the CP of AbMV (non-transmissable virus) with that of SiGMV (transmissable virus) resulted in transmission of an insect non-transmissable virus by *B. tabaci* (Hofer *et al.* 1997).

Mutational analysis of the complementary genes on DNA-A of ACMV revealed that 3 ORFs (AC1, AC2 and AC3) are highly conserved among dicot-infecting geminiviruses. AC1 is a prerequisite for replication since disruption of this ORF results in the DNA being non-infectious and no replication occurring. Disruption of AC2 prevents infection of the plant, suggesting that this product is involved in virus spread within the plant. AC3 mutants are still able to replicate and infect plants but the onset of symptoms is delayed and attenuated, thus AC3 is most probably responsible for modulating the levels of virus in infected cells (Etessami *et al.* 1991, Morris *et al.* 1991). Thus AC1 is the replication-associated protein gene essential for viral DNA replication, AC2 is required for virus spread by transactivating the expression of the CP and DNA-B genes and AC3 modulates the levels of viral DNA replication (Haley *et al.* 1992, Hong and Stanley 1995).

DNA-B encodes 2 gene products, BV1 and BC1. These gene products are believed to be movement proteins that are essential for systemic infection of

the plant, symptom production and host range of the virus (von Arnim *et al.* 1993, Haley *et al.* 1995, Schaffer *et al.* 1995). BV1 is involved in localised cell-to-cell spread, it is believed to bind ssDNA and shuttle the viral genome into and out of the nucleus (Pascal *et al.* 1994). BC1 is found at the periphery of cells and interacts with BV1 to mediate the long distance vascular spread of the virus (Sanderfoot and Lazarowitz 1995).

Due to the bipartite nature of geminiviruses it has been possible to investigate the relationships and specificity's of various geminiviral genes by co-inoculating mixtures of DNA-A and -B components of different viruses in common hosts. Replacing the CP gene of ACMV with the movement genes of DNA-B of TGMV inhibited systemic infection suggesting lack of complementation (von Arnim and Stanley 1992).

Co-agroinoculation of mixtures of Old World and New World viruses showed that Old World viruses could mediate the systemic movement of most New World viruses but not *vica versa*, suggesting an evolutionary divergence between the 2 virus types (Frischmuth *et al.* 1993). Other studies found that viable pseudorecombinants between viruses can only be produced if the 2 viruses were strains of the same species and that the inability of pseudorecombination was not due to A- and B-component differences but more likely due to inability of AC1 (AL1) replication protein to initiate replication of heterologous -B components (Sunter *et al.* 1994).

Once inside a host plant cell, a geminivirus needs to replicate and produce viral progeny for systemic infection of the plant. Geminiviruses have small, circular ssDNA genomes which are replicated by the rolling-circle mechanism of replication (Saunders *et al.* 1991). This is a two step process

where leading and lagging strand synthesis are two separate events. The first step is where ssDNA is replicated to form a double stranded replicative form which is then used as a template for producing free ssDNA (Hanley-Bowdoin *et al.* 1999). Initiation of replication occurs at the origin of replication which is found within the 200 bp highly conserved common region (CR).

Geminivirus replication origins are composed of at least 3 domains: a putative stem-loop structure, a specific high affinity binding site for Rep protein, and at least 1 additional element that contributes to specific origin recognition (Fontes *et al.* 1994, Arguello-Astorga 1994a, Orozco *et al.* 1998). The organisation of iterated elements (iterons) in terms of number, orientation and spacing is specific for each geminivirus and is highly conserved within the major lineages of dicot-infecting geminiviruses (Arguello-Astorga 1994b). Within the CR of all geminiviruses is a stem-loop structure containing the characteristic nonanucleotide motif, TAATATTAC. Geminivirus-encoded Rep proteins bind specifically to a region adjacent to the stem-loop region and introduce a nick in the viral DNA at the nonanucleotide motif which initiates replication (Stanley 1995).

1.4 Cassava-infecting Begomoviruses

Studies carried out using monoclonal antibodies against a range of geminiviruses have shown that all whitefly-transmitted geminiviruses (WTGs) are serologically related to each other. This is true even between Old World and New World WTGs irrespective of host or geographical origin (Roberts *et al.* 1984). Initially viral isolates infecting cassava were thought of

as strains of the same virus but as more viruses were found questions began to be raised as to whether these were strains of the same virus or actually new viruses.

Serological testing of a number of isolates from cassava from different geographical regions within Africa and India showed that CBVs could, on the basis of their serology, be placed into 3 groups, designated Group A, B and C. Group A included the type strain from western Kenya (ACMV-T), as well as isolates from Angola (ACMV-A), Nigeria (ACMV-N) and Ghana (ACMV-G). Group B comprised isolates from coastal Kenya (ACMV-C) whereas Group C consisted of isolates from India (ACMV-I) (Harrison *et al.* 1986). This separation of isolates was further supported by evident biological differences between the groups. Group A isolates were readily mechanically transmitted to *N. benthamiana* and *N. clevelandii,* accumulated to high concentrations and had an optimum temperature of 23°C; Group B isolates were less readily mechanically transmitted to *N. benthamiana* and *N. clevelandii,* occurred in low concentrations and required a higher optimum temperature (30°C); Group C isolates could be transmitted to and maintained in *N. benthamiana* and multiplied best at 30°C (Robinson *et al.* 1984, Harrison *et al.* 1986).

Genetic differences between the groups was shown by the ability of a Group A ACMV-T DNA-A probe to bind with different affinities to isolates within groups B and C whereas a full-length probe to DNA-B reacted only weakly, often not at all, with isolates from groups B and C (Robinson *et al.* 1984). These differences between the groups with regard to gene sequence was further investigated once the full-length sequences of ACMV-M (group

B isolate) and ICMV (formerly ACMV-I of group C) were available. Nucleotide sequence comparisons between these isolates and that of ACMV-K (formerly ACMV-C from group A) revealed that these 3 isolates are significantly different enough to justify their inclusion as **distinct** viruses designated as *African cassava mosaic virus, East african cassava mosaic virus* and *Indian cassava mosaic virus* (Hong *et al.* 1993).

Cassava was introduced from South America to Africa and India. In South America where cassava originated, a mosaic disease of cassava does occur but it is caused by a potexvirus. Cassava was distributed to 3 major points, the west coast of Africa, East Africa and India and Sri Lanka. It is thought that indigenous geminiviruses present in natural hosts within these 3 areas infected cassava which resulted in the appearance of cassava mosaic disease and the occurrence of 3 distinct virus species. The exact origin of CBVs is still not known. Geminiviruses are found in a wide range of plant species including many weed species which grow near infected cassava plantations, but the exact role of these plants with respect to cassava infection is unclear.

There have not been any large-scale, co-ordinated, systematic surveys of the distributions of ACMV, EACMV and ICMV. The most comprehensive survey to-date has been a large scale serological survey by Harrison (1994) which suggested that each of these 3 viruses had separate, non-overlapping geographic distributions. ICMV only occurs in India and Sri Lanka whereas ACMV and EACMV only occur in Africa. ACMV has been found mainly in West Africa countries, particularly Benin, Burkina Faso, Ghana, Ivory Coast, Nigeria, Senegal, as well as Angola, Burundi, Cameroon, Chad, Congo, western Kenya, Mozambique, Tanzania, Uganda and Zambia. EACMV is

mainly confined to East Africa, in Malawi, Madagascar, Zimbabwe, coastal regions of Kenya and Tanzania (Harrison 1995, Swanson and Harrison 1994).

However more recent findings seem to contradict this simplistic geographical separation of viruses:

- Certain areas of central Kenya, parts of Kenya and Tanzania (near Lake Victoria) and Ukerewe island have ACMV and EACMV occurring together in the same area (Harrison *et al.* 1997).

- EACMV which was originally thought only to occur in East Africa, east of the Rift Valley, has since been found in Ghana (Offei *et al.* 1999), Cameroon (Fondong *et al.* 1998) and Nigeria (Ogbe *et al.* 1999).

- Severe mosaic disease in Cameroon has been found to be as a result of mixed infections of ACMV and EACMV (Fondong *et al.* 2000).

In recent years geminiviral taxonomy and classification has undergone a number of revisions and adjustments (Rochester *et al.* 1993, Rybicki 1994, Padidam *et al.* 1995, Padidam *et al.* 1997, Rybicki 1998, Fauquet *et al.* 2000). This is made particularly difficult by the continual discovery of 'new' geminiviruses. For example:

(A) Since the 1980s Uganda has been experiencing a severe pandemic of ACMD which has totally destroyed cassava cultivation in many areas and left millions without food (Gibson *et al.* 1996). The epidemic has been caused by a number of factors. A new virus, the Uganda variant, UgV, has

arisen by recombination between ACMV and EACMV (Zhou *et al.* 1997,
Deng *et al.* 1997). Compared to ACMV and EACMV, UgV was more
virulent and multiplied to higher concentrations in plants (Harrison *et al.*
1997). The CP of UgV consisted of 3 parts, the 5' end was identical to
EACMV, the middle portion to ACMV and 3' end was similar to EACMV,
yet the CP of UgV was indistinguishable from ACMV in serological tests
(Zhou *et al.* 1997). Increased populations of whitefly (*Bemisia tabaci*
Gennadius) vector which may have contributed to the spread of disease.
Planting of extremely susceptible cassava varieties unable to withstand even
slight viral infections may have exacerbated spread (Otim-Nape *et al.*
1998a).

(B) In South Africa a novel type of geminivirus, *South African cassava*
mosaic virus (SACMV) that infected cassava was discovered (Berrie *et al.*
1998). This virus was a bipartite virus that clustered together with other Old
World geminiviruses. SACMV is a recombinant virus sharing significant
sequence homology with *Tomato yellow leaf curl virus* (TYLCV) in the
intergenic region and EACMV in the coat protein region (Berrie *et al.* 2001).

(C) Full length DNA-A sequence analysis of EACMV isolates from Malawi
revealed that the viral sense genes differed from other EACMV isolates and
was more like those of *Tomato yellow leaf curl virus-Israel*, implying that a
recombination event must have occurred in the past (Zhou *et al.* 1998). This
further complicates the taxonomic status of EACMV since it appears that
perhaps 2 distinct types of EACMVs now occur, those that cluster more

closely to other Old World geminiviruses and those that cluster with EACMV isolates from Kenya and Tanzania.

(D) A recently discovered isolate from Cameroon, EACMV/CM, was found to have undergone recombination in the AC2-AC3 **and** BC1 coding regions. This is the first example of a geminivirus undergoing recombination in the complementary sense genes of both DNA-A and -B suggesting that this virus originated from 2 different viral species (Fondong *et al.* 2000).

1.5 Reasons for the Emergence of New Geminiviruses

Recombination appears to be a powerful means by which geminiviruses have evolved over the years. Evidence seems to suggest that *Curtovirus*es resulted from an ancient recombination event between *Begomovirus*es and *Mastrevirus*es (Rybicki 1994). Recombination between closely related geminiviruses has been shown to occur fairly frequently (Padidam *et al.* 1999). Two isolates considered strains of the same virus that share AC1 and CR/IR sequence similarities can complement each other's function when co-inoculated into common hosts. The fact that geminiviruses within geographical areas are closely related in terms of their CP phylogeny and can be transmitted by the same whitefly vectors suggests that opportunities do exist for recombination to occur.

Genomic variation within geminiviruses can be caused by a number of mechanisms:

- *Mutation*: work on maize streak virus (*Mastrevirus*) has shown that isolates of the same virus can differ by 1 or more point mutations which affects symptom production, host range and whitefly transmission.

- *Acquisition of extra DNA components*: most begomoviruses, with the exception of those causing *Tomato yellow leaf curl* diseases, are bipartite. However some viruses have been found that have additional genomic components, in addition to DNA-A and -B. Certain populations of geminiviruses have also been found that have additional circular DNA molecules half the size of normal genomic components. These molecules, called defective interfering DNA, occur in naturally infected plants, their role is to decrease the accumulation of viral DNA and ameliorate symptoms (Stanley *et al.* 1990). *Ageratum yellow vein virus* in Pakistan has been found to be associated with a begomovirus **and** a nanovirus-like molecule half the size of normal begomovirus genome and with similarity to *Cotton leaf curl virus* (Mansoor *et al.* 2000).

- *Pseudorecombination*

- *Recombination*: Recombination of geminiviruses seems to be on the increase. Recombinant geminiviruses have been found that infect cassava in Africa (Zhou *et al.* 1997, Deng *et al.* 1997), cotton in Pakistan (Sanz *et al.* 2000), pepper in Mexico and tomato in Central America. Padidam *et al.* (1999) used a gene recombination detection program (GENECONV) to detect possible sites of recombination within geminiviral genomes. They found that recombinations appeared to occur **throughout** the geminiviral genomes and more recombination events were recorded for WTGs from Asia compared to those from Africa or America. In contrast

recombinations between viruses from the Old World and New World and begomoviruses and curtoviruses were concentrated at the N-terminal region of the Rep protein. They suggest that 3 factors could contribute to this apparently high level of recombination i.e. mixed infections (which is common in geminiviral diseases, particularly since geminiviruses share many hosts in common), high levels of viral replication (which is true for geminiviruses which replicate by the rolling circle mechanism, generating high copy numbers) and increased host range of the vector (the appearance of a new biotype (B-biotype) *B. tabaci* which has increased fecundity and host range has increased the opportunity for increased viral spread and recombination) (Harrison *et al.* 1999, Padidam *et al.* 1999).

In conclusion, the propensity of geminiviruses to mutate by any of the abovementioned mechanisms, along with the emergence of new more devastating biotypes of *B. tabaci*, has led to the emergence of new begomovirus variants and concomitantly devastating epidemics such as that affecting cassava in Uganda and cotton in Pakistan.

1.6 The Whitefly Vector (*Bemisia tabaci* Gennadius)

Bemisia tabaci or whitefly was first identified in Greece in 1889 and was considered a pest of tobacco. Since then many additional whitefly species have been found. To date there are more than 1100 identified species of *Bemisia,* yet only 3 are vectors of plant viruses, the most important species

being *Bemisia tabaci*. This whitefly belongs to the order Homoptera, family Aleyrodidae and is primarily a polyphagous insect that colonises annual herbaceous plants (Brown *et al.* 1995[a]). Whiteflies are small insects, measuring 0.6-2.0 mm in length, have piercing-sucking mouthparts and feed on plant sap. Whiteflies undergo 5 developmental stages from egg to adult, with only the first and last stages allowing mobility (Muniyappa 1980).

Since first being described as a pest of tobacco many years ago, the distribution and host range of *Bemisia tabaci* has increased. Today they occur in almost all tropical and sub-tropical regions of the world, and are appearing in some temperate regions as well, and affect the majority of food, fibre and ornamental crops. Whiteflies were also initially only considered as sporadic pests but recent increases in whitefly-associated diseases worldwide have elevated their status to that of serious plant pests and pathogens (Brown 1994).

The effect of whiteflies on plant growth and yield can be divided into direct and indirect damage. Direct damage involves the removal of large amounts of phloem sap which can severely debilitate the plant, as well as the induction of certain physiological disorders such as silvering of leaves, blanching of stems, vein clearing and irregular ripening. Indirect damage is due to virus transmission and honeydew production. To date approximately 60 members of the geminivirus family are known to be transmitted by *B. tabaci*. These viruses occur throughout the tropics and sub-tropics and are responsible for significant yield losses in many important crop species, such as cassava, cotton, tomato, melons and tobacco. One of the by-products of whitefly feeding on plants is the production of honeydew on the leaf surfaces.

This serves as a substrate for sooty moulds and powdery mildew which reduces overall plant yields and marketability (Schuster *et al.* 1995).

Whitefly-transmitted geminiviruses (WTGs) are transmitted in a non-circulative, persistent manner. The viruses are acquired from an infected plant host, remain in the vector for a certain period, without affecting the vector, and are then transmitted into new hosts. The viruses have fairly similar transmission characteristics, acquisition and inoculation can occur in a minimum of 10 minutes but usually feeding times of up to 24 h are optimal, latent periods of 4-21 h are necessary but this is usually variable depending on the virus type (Duffus 1995).

It has been shown that efficient transmission of geminiviruses relies on specific vector-virus interactions where the coat protein of geminiviruses plays a predominant role. This is shown by the closer serological similarities between WTGs in different hosts in the same area than that of WTGs infecting the same host in different areas (McGrath and Harrison 1995). Another factor also influencing viral transmission is vector-host interactions, incompatible whitefly-host interactions prevent the acquisition and inoculation of geminiviruses.

1.6.1 *Bemisia tabaci* and Cassava Mosaic Disease

ACMV was one of the first viruses to be shown to be experimentally transmitted by whiteflies. It has a minimum (and optimum) acquisition access period of 3 h (5h), latent period of 3-4 h (6h) and inoculation access period of 10 min and the virus can persist in the vector for up to 9 days. Evidence from Cote d'Ivoire suggests that only 1.7% of adult whiteflies

become viruliferous after feeding on infected cassava whereas evidence from Uganda shows 13% of whiteflies are infective (Otim-Nape *et al.* 1995). Two other whitefly species, *Bemisia afer* (Priesner and Hosny) and *Aleurodicus dispersus* (Russel) colonise cassava in Africa. *Bemisia afer* is a vector for cassava brown streak virus and *Aleurodicus dispersus* causes feeding damage but neither of these species transmits ACMV (Legg 1994).

Bemisia tabaci affects cassava in Africa yet in South and Central America (from where cassava originated), *B. tabaci* affects a wide range of other plant species but not cassava. In contrast in Africa a number of different host-associated biotypes exist. Studies in Cote d'Ivoire have revealed that 2 biotypes of *B. tabaci* occur, one biotype (cassava biotype) colonises cassava and egg plant and transmits ACMV whereas the other biotype (okra biotype) is more polyphagous but does not occur on cassava and does not transmit ACMV (Burban *et al.* 1992).

Transmission and spread of ACMV is complex and affected by a range of different factors. These factors influence both virus-vector and host-vector interactions. The virus-vector relationship has been difficult to fully investigate. Some studies have investigated the specificity of interaction between virus and vector while other studies have examined the ecology of these interactions i.e. associating whitefly numbers with disease incidence.

Generally the appearance of large numbers of whiteflies a few weeks prior to disease appearance results in rapid spread of ACMV. Host-vector relationships are influenced by both those factors relating to the host and those to the vector. Host-associated factors include: host chemistry (resistance and cyanogenicity of plant), host architecture (width of leaves,

height of plant), host age (younger plants more vulnerable to infection than older ones) and host growth (rapid rate of growth of plant associated with rapid increase in whitefly numbers) (Legg 1994). Examples of vector-associated factors are: temperature (the higher the temperature the quicker the development rate of whitefly), rainfall (positively correlates with whitefly population dynamics), wind speed, wind direction, planting date, plant density, intercropping and crop deployment (planting upwind compared to downwind) (Fishpool *et al.* 1995, Otim-Nape *et al.* 1995).

1.6.2 Appearance of new *Bemisia tabaci* biotypes

Bemisia tabaci is thought to have originated somewhere on the Indian subcontinent and then have spread to Africa, Europe and the Americas by the transport of plant material. *Bemisia tabaci* has been reported to occur on more than 500 plant species and it is this polyphagous nature that has enabled *B. tabaci* to successfully survive in new areas and transmit indigenous viruses.

In the 1950s it became evident that some populations of *B. tabaci* that were morphologically indistinguishable exhibited different biological traits. These traits included host range, host plant adaptability and plant virus transmission capabilities. Hence the separation of *B. tabaci* species into biotypes or host races was proposed. These biotypes are now separated on the basis of: host plant affiliations, degrees of phytotoxic symptom induction, insecticide resistance, morphology and behaviour (Brown *et al.* 1995[a]). This is supported by evidence that populations of *B. tabaci* that are geographically

(and reproductively) isolated differ in their ability to utilise particular host plants and in their virus transmission characteristics (Brown 1994).

Whitefly-transmitted geminiviral diseases had until a few years ago in most parts of the world been sporadic and controllable. But in the 1980s in parts of the southwestern USA a number of severe epidemics started to affect many food and fibre crops. Studies on whitefly populations associated with these epidemics suggested that a new strain of *B. tabaci* was occurring. This newly introduced species resulted in the loss of more than $500 million in 1991 in the USA alone (Perring *et al.* 1993). This whitefly was morphologically similar to the sweet potato (A-biotype) whitefly which had been known in the USA for more than 100 years but exhibited a number of different biological characteristics. This species of whitefly, which became known as the "superbug", "type B" or "pointsettia" strain was more mobile, able to spread further distances, had increased fecundity, a significantly larger host range and caused many types of phytotoxic disorders (Markham *et al.* 1995).

Since it's 'discovery' in the early 1990's the B-biotype has spread and now occurs in many tropical and sub-tropical regions throughout the world, in the process replacing many of the indigenous whitefly populations. This B-biotype is thought to have originated from the Old World and was disseminated throughout the world by the trade and traffic in ornamentals, particularly poinsettias (Bedford *et al.* 1992). This has been confirmed by phylogenetic studies of 16s rDNA markers which show that the B-biotype whitefly clusters with Old World whiteflies suggesting that the B-biotype originated in the Old World and has recently been introduced to the New World (Frohlich *et al.* 1999).

Studies using PCR-based DNA differentiation tests, allozymic frequency analyses, crossing experiments and mating behavioural studies found that this new whitefly species was actually a **distinct** species. It was proposed that this whitefly formerly termed type-B be called "silverleaf whitefly" and later a new species, *Bemisia argentifolii* (Perring *et al.* 1993).

1.6.3 Taxonomy of *Bemisia* species

The taxonomic status of *B. tabaci* was, until very recently, based on morphological characteristics of the fourth instar. The majority of the B-biotypes lack a 4th setae in the anterior submarginal region and have narrow and delicate anterior wax fringes. In contrast the A-biotype has a 4th setae and the wax fringes are more robust and wide (Brown *et al.* 1995[a]). But it has been shown that the morphological characteristics of the nymphs vary in response to differences in leaf surface topology and environmental and physical factors such as temperature and humidity. Thus morphological characteristics of pupae are not on their own sufficient for classifying individuals as to being *B. tabaci* or *B. argentifolii* (Rosell *et al.* 1997).

In addition to biological evidence that supports the existence of 2 or more species of *Bemisia* species there is also molecular and biochemical evidence to corroborate this. Protein polymorphisms in the form of non-specific esterase profiles of whiteflies from the Old and New World have clearly shown that the A- and B-biotypes have distinct esterase patterns (Coats *et al.* 1994, Brown *et al.* 1995[b], Coats *et al.* 1995, Byrne *et al.* 1995). This has been very useful in tracking the movements of whiteflies in many parts of the world. Isozyme polymorphisms between A- and B-biotypes separate

populations on the basis of geographical origin i.e. those occurring in the New World and those in the Old World. Genetic polymorphism studies have investigated 16s rDNA, RAPDs, RFLPs and COI markers. Grouping *Bemisia* species on the basis of 16s rDNA sequences is useful at the subspecies level and results in 4 groups: group 1- individuals from India, group 2- New World individuals, group 3- individuals from Sudan and group 4- other individuals from the Old World as well as the B-biotype (Frohlich *et al*. 1996).

Differences in RAPD banding patterns proved that the A- and B-biotypes were indeed separate species but the RAPD technique is more useful in distinguishing closely-related organisms than for determining species natures (Brown *et al.* 1995[a]). The mitochondrial cytochrome oxidase I (COI) gene is more variable than 16s sequence and thus is thought to be more informative. Results of COI groupings are similar to those of 16s rDNA groupings and further narrows down the origin of the B-biotype to somewhere within the dry zones of North-East Africa, Middle East or Arabian Peninsula. In addition the Old World whitefly species are divided into 3 groups: Africa; India; and Sudan, Yemen, Israel (Frohlich *et al.* 1999).

The evident biological, phenotypic and genotypic differences between species of *Bemisia* have complicated the taxonomy of this genus and whether *Bemisia* consists of races or biotypes or is a species complex is currently under much scrutiny and debate. It is hoped that soon sufficient evidence will be available to improve and entrench the taxonomic status of this very important plant pest and pathogen.

1.7 Objectives

1.7.1 Background

This project was funded in part by the British Council and was performed in collaboration with overseas researchers, Dr. Peter Markham and colleagues, John Innes Centre, Norwich, UK. The theme of the collaborative research project being "Biodiversity and strain differentiation of ACMV and *B. tabaci*"

Our component of the project involved examining the genetic diversity of cassava begomovirus isolates and whitefly vectors from several southern African countries as well as developing new sensitive and reliable techniques for differentiating cassava viral isolates. The other component of the project which was undertaken by the overseas collaborators was the investigation of the relationships between DNA-A and -B genomic components with respect to virus movement in the plant.

1.7.2 Specific Aims

- Examine cassava begomovirus diversity within southern Africa by PCR amplification of the highly conserved core coat protein region of DNA-A followed by cloning and sequencing of amplification products and phylogenetic analyses. Investigate the occurrence of recombinant UgV-like viruses within southern Africa using specific, discriminatory PCR.

- Investigate the diversity of *Bemisia tabaci* populations in southern Africa by examining the genetic relatedness of whitefly isolates using the mitochondrial cytochrome oxidase I gene.

- Develop new sensitive, specific and reliable diagnostic techniques for identifying and characterising cassava-infecting begomovirus isolates. In particular the Heteroduplex Mobility Assay which is a mutation detection technique enabling rapid characterisation of isolates without resorting to traditional methods of cloning and sequencing. As well as utilising differential phage display technology to produce monoclonal antibodies to SACMV.

CHAPTER TWO

MOLECULAR EVIDENCE FOR DIVERSE

POPULATIONS OF CASSAVA-INFECTING

BEGOMOVIRUSES IN SOUTHERN AFRICA

2.1 ABSTRACT

Cassava mosaic disease is prevalent in Africa and significantly affects the growth and yield of cassava. This disease is caused by a number of whitefly-transmitted begomoviruses. In this paper we describe results obtained from screening field populations of begomoviruses from infected cassava from six countries in southern Africa. Using primers that target the highly conserved core region of the coat protein gene it was possible to identify and establish the geographical distribution and relatedness of cassava begomoviruses. It was found that *African cassava mosaic virus* occurs in five of the six countries (except Angola), *East African cassava mosaic virus* is present in five countries (except Zambia) and *South African cassava mosaic virus* was detected only in South Africa and Swaziland. In addition, we report for the first time in southern Africa, the appearance of the Ugandan variant virus (UgV) which was found frequently in mixed infections with other cassava-infecting begomoviruses.

2.2 INTRODUCTION

The most important disease affecting cassava (*Manihot esculenta* Crantz) production in Africa is cassava mosaic disease (CMD). This disease is caused by a group of begomoviruses which are members of the *Geminiviridae* family of plant viruses and are characterised by their ssDNA-containing genomes, their twinned, isometric particle morphology (Harrison 1985) and their transmission by *Bemisia tabaci* Gennadius (Brown *et al.* 1995). Currently four different cassava-infecting begomoviruses (CBVs) have been found associated with CMD in Africa namely *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV), Ugandan variant virus (UgV) which is an ACMV-EACMV recombinant virus (Zhou *et al.* 1997) and *South African cassava mosaic virus* (SACMV) (Berrie *et al.* 1998), and in India and Sri Lanka, *Indian cassava mosaic virus* (ICMV) (Thresh *et al.* 1998).

Serology, nucleic acid hybridisation and sequence comparison analyses have shown that these viruses are sufficiently different to be considered distinct virus species (Robinson *et al.* 1984, Thomas *et al.* 1986, Hong *et al.* 1993). Until recently, extensive serological surveys suggested that CBVs in Africa had separate, non-overlapping distributions: ACMV isolates were found mainly in West Africa; EACMV occurred mainly in East Africa (Harrison and Swanson 1994); SACMV was only reported in South Africa and UgV was mainly confined to Uganda. However, more recent findings contradict a non-overlapping geographical distribution of viruses: EACMV has been found in parts of Ghana (Offei *et al.* 1999), Cameroon (Fondong *et*

al. 1998) and Nigeria (Ogbe 1999) and in regions of Kenya and Tanzania (Harrison *et al.* 1997), and Cameroon (Fondong *et al.* 2000) mixtures of both ACMV and EACMV have been reported.

Although the incidence and distribution of CBVs in many countries in West, Central and East Africa have been extensively examined, little is known about cassava begomovirus populations in southern Africa. In this paper we describe the results obtained from a study in six countries within southern Africa. Molecular evidence demonstrates that four of the described cassava-infecting begomoviruses are present and that a significant number of plants contain these viruses in combination with UgV. This has implications for pseudo-recombinations, recombinations and viral disease diagnosis and control.

2.3 MATERIALS AND METHODS

Cassava leaf samples displaying CMD symptoms of mosaic, stunting and curling were collected from various regions within Angola, Mozambique, South Africa, Swaziland, Zambia and Zimbabwe (Table 2.1).

Total nucleic acid (TNA) was extracted from the leaf samples according to the method of Doyle and Doyle (1987). The field samples were screened for viruses using two sets of primers, AV514 and AC1048 (Wyatt and Brown 1996) that anneal to highly conserved regions within the core region of the coat protein gene of whitefly-transmitted begomoviruses and primers UV-AL1/F1 and ACMV-CP/R3 (Zhou *et al.* 1997) which specifically amplify ~1600 bp region of recombinant UgV-like virus DNA-A, including parts of the AC1, common/intergenic and coat protein regions. Core coat protein

(CCP) PCR amplification was performed as described by Paximadis *et al.* 1997. For PCR of UgV, amplification was carried out in the same thermal cycler using a modified 'hot start' and cycling conditions as reported in Zhou *et al.* 1997.

Table 2.1: Distribution of the cassava collection sites within southern Africa

Country	Area	Province/Region	Accession numbers
Angola	Okavango	Southern Angola	AF329228
Mozambique	Beira	Sofala	AF329246
	Chimoio	Manica	AF329245
	Nyamundinn	Manica	AF329247
	Maputo	Maputo	AF329237-AF329239, AF329243
	Zambezia	Zambezia	AF329240-AF329242, AF329244
South Africa	Hoxane	Mpumalanga	AF329216, AF329220, AF329223
	Tonga	Mpumalanga	AF329225-AF329227
	St. Lucia	Kwazulu-Natal	AF329217-AF329219, AF329222
	Eastern Natal	Kwazulu-Natal	AF329215, AF329221, AF329224
Swaziland	Maphiveni	Eastern Swaziland	AF329234-AF329236
Zambia	Unknown	Unknown	AF329232-AF329233
Zimbabwe	Harare	North-eastern Zimbabwe	AF329229-AF329231

The resulting amplification products were electrophoresed on 1% agarose gels and viewed with UV illumination. The CCP amplification products were cloned into pBluescript vectors and sequenced by automated sequencing (ABI Prism 310). CCP clones were sequenced in both directions using the universal KS and M13 primers to confirm the duplicity of sequences. The CCP sequences were then aligned and phylogenetic trees constructed using 5000 bootstraps and the Multiple Sequence Alignment and Phylogenetic Tree Construction functions of the DNAMAN software program (Lynnon Biosoft ©, Quebec, Canada). Nucleotide sequence comparisons were made with other cassava and Old World Begomoviruses (names, abbreviations and accession numbers in Figure 2.2 legend). Although UgV primers are specific discriminatory primers that amplify only UgV-like viruses, to confirm that the PCR amplification products were indeed viral, one isolate from Mozambique was cloned and sequenced.

2.4 RESULTS

The CCP PCR primers are degenerate primers that amplify ACMV, EACMV and SACMV thus the occurrence of ~580 bp amplification products suggested that either of these viruses or a mixture of viruses could be present. The PCR screening results (Table 2.2) showed that samples from Angola and Zambia contained either ACMV, EACMV or SACMV, whereas for Mozambique, South Africa, Swaziland and Zimbabwe, 63%, 24%, 25% and 40% of the samples respectively, were positive for ACMV, EACMV or SACMV with no occurrence of UgV. From the results of the UgV PCR (Table 2.2) it can be noted that a number of the samples collected from

Mozambique, South Africa, Swaziland and Zimbabwe contained UgV-like viruses whereas samples from Zambia and Angola did not appear to be infected with UgV.

Table 2.2: Results of the survey of cassava begomoviruses in southern Africa using CCP and UgV primers

PCR   **Begomoviruses**

	CCP PRIMERS	UGV PRIMERS	ACMV/EACMV /SACMV	ACMV/ UGV + EACMV/ SACMV	UGV ALONE
Angola (n=1)	1/1 (100%)	0/1 (0%)	1/1 (100%)	0/1 (0%)	0/1 (0%)
Mozambique (n=16)	14/16 (88%)	4/16 (25%)	10/16 (63%)	4/16 (25%)	0/16 (0%)
South Africa (n=42)	28/42 (67%)	22/42 (52%)	10/42 (24%)	18/42 (43%)	4/42 (9%)
Swaziland (n= 4)	4/4 (100%)	3/4 (75%)	1/4 (25%)	3/4 (75%)	0/4 (0%)
Zambia (n=3)	3/3 (100%)	0/3 (0%)	3/3 (100%)	0/3 (0%)	0/3 (0%)
Zimbabwe (n=5)	3/5 (60%)	2/5 (40%)	2/5 (40%)	1/5 (20%)	1/5 (20%)

From Mozambique, 4 out of the 16 samples tested were positive for UgV and all the samples contained UgV in mixed infections with one or other CBV. In South Africa, 52% of the samples tested were positive for UgV and of these 43% were mixed infections of UgV and other CBVs whereas the

remaining 9% contained UgV alone. Four samples were tested from Swaziland and, of these, 3 (75%) were positive, with UgV only occurring in mixed infections with other CBVs. In Zimbabwe, 40% of the samples were UgV positive with half exhibiting UgV in mixed infections and the other half exhibiting single UgV infections.

Sequencing of a 776 bp region of DNA-A, encompassing the CR/IR, AV1 and AV2 genes, of the UgV-like begomovirus from Mozambique revealed a virus which had nucleotide sequence similarities of 95% with ACMV-UGMld (accession number AF126800) and ACMV-UGSvr (accession number AF126802) from Uganda. These results demonstrate for the first time that recombinant UgV-like viruses causing the epidemic in Uganda are also now found in a number of other cassava-growing countries within southern Africa and occur frequently in mixed infections.

In order to establish the identities of the core coat protein PCR products and to determine their relatedness and phylogenetic relationships, a large number of the CCP amplification products were cloned and sequenced. The percentage nucleotide similarity in the coat protein between ACMV and EACMV is in the region of 80-85%. It has been proposed that any new virus isolate having more than 90% sequence similarity to a previously characterised virus be considered a strain of an already described virus species, whereas an isolate with less than 90% similarity be considered a new strain (Padidam *et al.* 1995). The core coat protein gene is generally accepted as a good indicator for assigning provisional identities to begomoviruses (Brown *et al.* 2001). Figure 2.1 illustrates a phylogenetic tree depicting the relationships between the southern African CBV isolates.
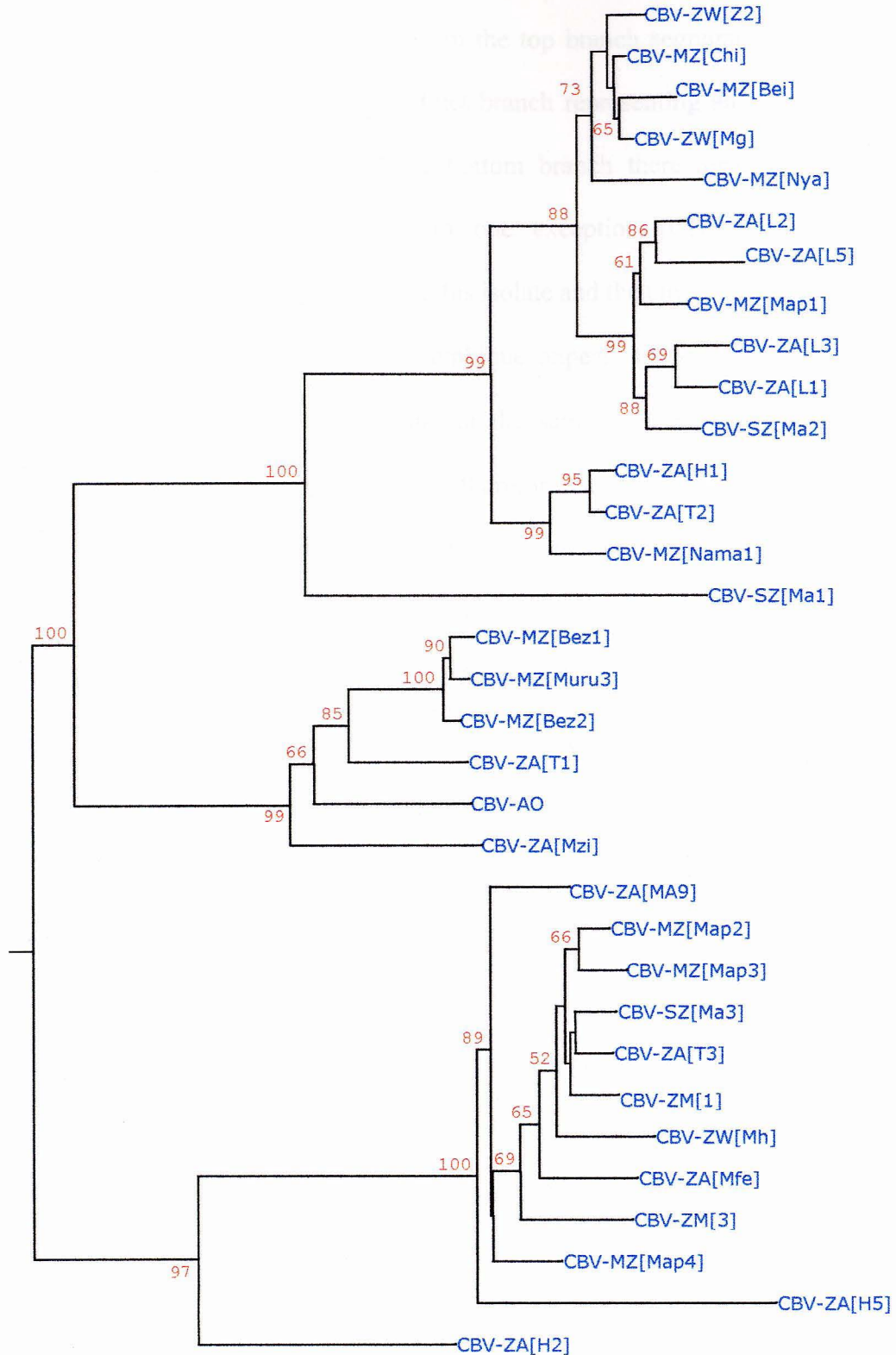
Figure 2.1: Phylogenetic Tree showing the relationships between 33 cassava virus isolates from six southern African countries by aligning the core coat protein sequences. Abbreviations used: CBV=cassava infecting begomovirus, ZW=Zimbabwe, MZ=Mozambique, ZA=South Africa, SZ=Swaziland, ZM=Zambia and AO=Angola.

The tree splits into two main branches (sequence similarity between the two branches of 75%) with isolates in the top branch segregating into two main clusters and one separate distinct branch representing an isolate from Swaziland (CBV-SZ[Ma1]). In the bottom branch there appears to be a single closely related cluster, with one exception (CBV-ZA[H2]) (nt sequence similarity of 86% between this isolate and the cluster).

The separation of CBVs in Mozambique appears to be geographically-based, with those viruses occurring in the same provinces being closely related to each other (94-98%). Furthermore, a number of the isolates collected from Mozambique are closely related to isolates collected from neighbouring countries. An isolate from Maputo (CBV-MZ[Map1]) is a strain (>96% homology) of an isolate from St. Lucia, South Africa; another isolate from Maputo (CBV-MZ[Map4]) is closely related to isolates from Zambia (95%) and Hoxane, South Africa (92%); an isolate from Beira is 98% similar to an isolate from Zimbabwe (CBV-ZW[Mg]). Thus it appears that there may have been multiple introductions of CBVs both into and out of Mozambique, particularly from neighbouring countries and little movement of infected planting material between provinces.

In terms of geographical separation and relatedness of isolates within SA, the three Hoxane isolates were sufficiently different to each other (nt similarities ranged from 73% to 83%) to be considered different viruses; the three isolates from Tonga were even more distantly related to one another (relatedness of 75% to 81%); the four St. Lucia isolates were almost identical (similarities of 97-98%); and the remaining three isolates from Natal had nucleotide sequence similarities ranging from 77%-96%. It is apparent from

these results, that in South Africa there exists several begomovirus strains and species, and these viruses are not as strictly geographically limited as those observed, for example, in Mozambique. From the limited survey in Swaziland it appears that the CBVs there are highly diverse and probably arose from the extensive trafficking of cassava stakes from the neighbouring countries of South Africa and Mozambique.

The second phylogenetic tree (Figure 2.2) establishes the identities of the isolates and depicts the relationships between these same CBV isolates with that of several characterised Old World begomoviruses. The bipartite, Old World Asian begomoviruses ToLCVIn1 and -In2, TYLCVTha, CLCuV, AYVV and ICMV infecting tomato, cotton, *Ageratum* and cassava, respectively, formed a separate cluster distinct from the African cassava group of viruses.

The cassava-infecting begomovirus field isolates fell within the four established branches of EACMV1, EACMV2, SACMV and ACMV. EACMV isolates from Malawi, EACMV-MW[MH] and EACMV-MW[K] (EACMV2) are more closely related in the CP region to TYCLV-Is than EACMV isolates from Kenya and Tanzania (EACMV1) (Zhou *et al.* 1998). The southern Africa isolates that clustered with the EACMV2 viruses from Malawi shared nt sequence similarities of 92% to 96% in the CCP with these viruses. The Swaziland isolate (CBV-SZ[Mal]) was the only isolate to share significant homology with SACMV which in turn shared 82% homology with the monopartite TYLCV viruses from Israel, Sardinia and Spain.
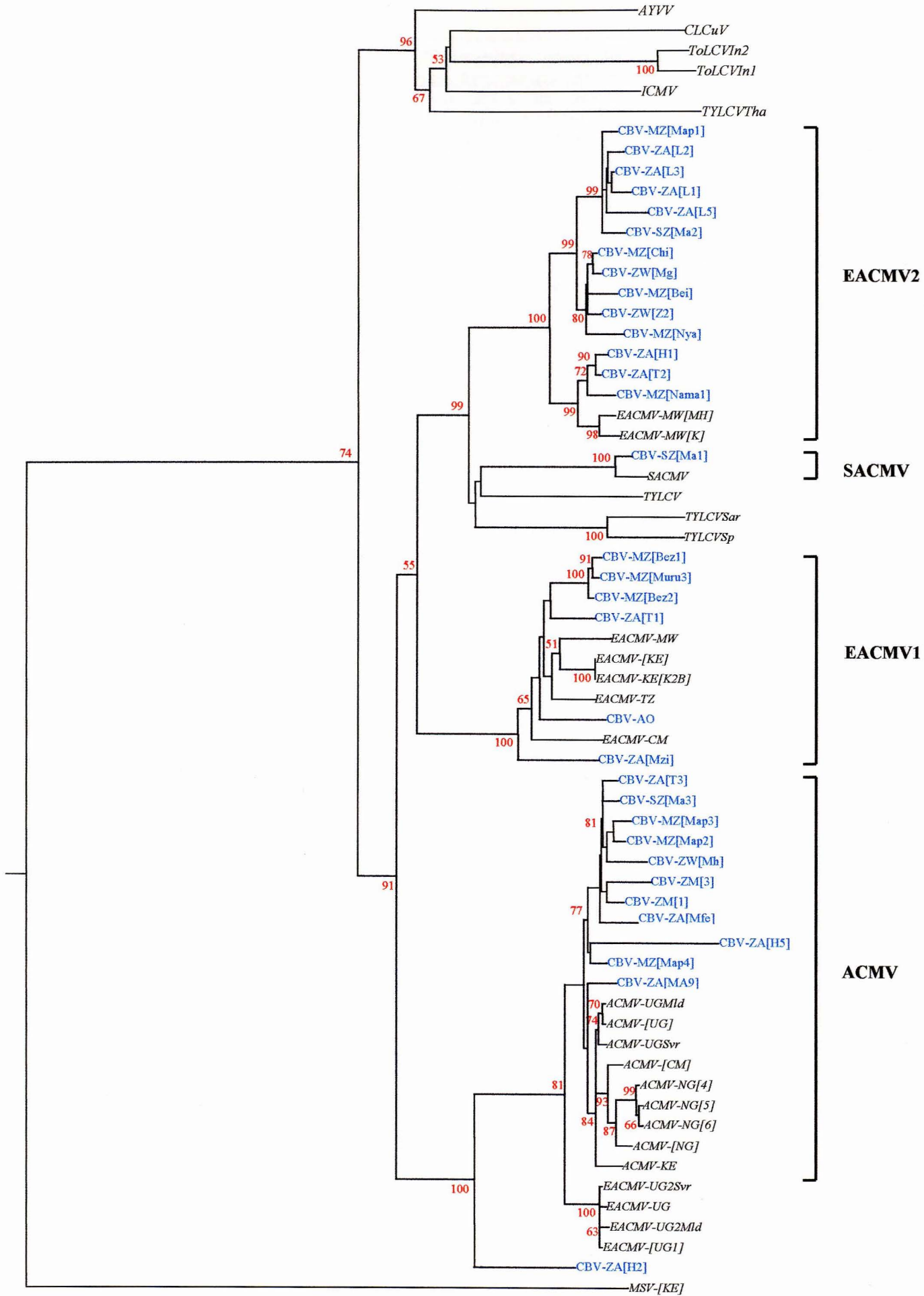
Figure 2.2: Phylogenetic Tree showing the CCP relationships between the 33 field cassava virus isolates and bipartite Old World Begomoviruses from cassava and other plant hosts. Maize Streak Virus-Kenya, MSV-[KE], (X01089) was used as the outgroup. Abbreviations and accession numbers: *AYVV*, Ageratum yellow vein virus (X74516); *CLCuV*, Cotton leaf curl virus-Pakistan1 (AJ132430); *ToLCVIn1*, Tomato leaf curl New Delhi virus-[Severe] (U15015); *ToLCVIn2*, Tomato leaf curl New Delhi virus-[Mild] (U15016); *ICMV*, Indian cassava mosaic virus (Z24758); *TYLCVTha*, Tomato yellow leaf curl Thailand virus (X63015); *EACMV-MW[MH]*, East African cassava mosaic virus-Malawi [MH] (AJ006459); *EACMV-MW[K]*, East African cassava mosaic virus-Malawi [K] (AJ006460); *SACMV*, South African cassava mosaic virus (AF155807); *TYLCV*, Tomato yellow leaf curl virus- Israel (X15656); *TYLCVSar*, Tomato yellow leaf curl virus-Sardinia (X61153), *TYLCVSp*, Tomato yellow leaf curl Sardinia virus-Spain [1] (Z257510); *EACMV-MW*, East African cassava mosaic virus- Malawi (AJ006461); *EACMV-[KE]*, East African cassava mosaic virus- [Kenya] (Z83258); *EACMV-KE[K2B]*, East African cassava mosaic virus (isolate K2B) (AJ006458); *EACMV-TZ*, East African cassava mosaic virus- Tanzania (Z83256); *EACMV-CM*, East African cassava mosaic virus- Cameroon (AF112354); *ACMV-UGMld*, African cassava mosaic virus- Uganda Mild (AF126800); *ACMV-[UG]*, African cassava mosaic virus-Uganda (X83252); *ACMV-UGSvr*, African cassava mosaic virus- Uganda severe (AF126802); *ACMV-[CM]*, African cassava mosaic virus- [Cameroon] (AF112352); *ACMV-NG[4]*, African cassava mosaic virus- Nigeria (clone pPCR4) (X68320); *ACMV-NG[5]*, African cassava mosaic virus- Nigeria (clone pPCR5) (X68319); *ACMV-NG[6]*, African cassava mosaic virus- Nigeria (clone pPCR6) (X68318); *ACMV-[NG]*, African cassava mosaic virus- [Nigeria] (X17095); *ACMV-KE*, African cassava mosaic virus- Kenya (J02057); *EACMV-UG2Svr*, East African cassava mosaic virus- Uganda2 Severe (AF126806); *EACMV-UG2Mld*, East African cassava mosaic virus- Uganda2 Mild (AF126804); *EACMV-[UG1]*, East African cassava mosaic virus- [Uganda1] (AF230375).

The third group of field isolates, from Angola, Mozambique and South Africa, were those related to the EACMV1 viruses from Cameroon, Kenya, Malawi and Tanzania. The separation of EACMV isolates into two groups (EACMV1 and EACMV2) was reiterated in this study where we found that 14/20 of the EACMV-like isolates collected from southern Africa are of the EACMV2 type and the remaining 6/20 are EACMV1. It was interesting to note that the majority of the southern African EACMV isolates were most closely related to the EACMV2 viruses since these viruses are thought to have arisen by recombination with begomoviruses with affinities to TYLCV-Is isolates. Detection of possible recombination within these southern African virus isolates can only be investigated by cloning and sequencing of additional viral genes.

The last group of field isolates clustered together with ACMV strains from the Cameroon, Kenya, Nigeria and Uganda. These isolates, with the exception of CBV-ZA[H2] which was more distantly related at 87% when compared to the ACMV reference viruses, shared relatively high sequence homologies of 94-97%.

2.5 DISCUSSION

Cassava mosaic disease is the most important constraint affecting cassava growth and yield. Surveys on the incidences and severity of CMD have been carried out in a number of countries (Fauquet and Fargette 1990, Thresh et al 1994, Otim-Nape and Thresh 1998). These surveys produced different results depending on the countries examined, with the incidence of CMD

ranging from 21% in Malawi to 82% in Uganda with an average mean incidence of ~54% for Africa (Thresh *et al.* 1998).

Serological surveys of several African countries (Swanson and Harrison 1994, Harrison *et al.* 1997) found EACMV present in Zimbabwe and ACMV in Angola, Mozambique, South Africa and Zambia. Sequencing results from this study now demonstrate that in Angola, Mozambique and South Africa, where previously only ACMV was found, there also exists EACMV species, and in Zimbabwe EACMV and ACMV occur. Although EACMV was not found in Zambia, more field samples would need to be tested. In Swaziland both ACMV and EACMV are present, and SACMV, previously reported only in South Africa, has now also been detected in Swaziland.

In Mozambique approximately 0.99 million hectares of land was used for cassava cultivation in 1996 yielding ~4.7 million metric tons (Thresh, personal communication). CMD occurs in all areas where cassava is grown in Mozambique, the overall incidence is ~21% and the symptoms are generally mild and CMD is largely an avoidable problem if the planting of healthy stem cuttings was enforced (Thresh, personal communication).

Previous studies on the distribution of CMD in South Africa and Swaziland (Trench and Hemmes 1985) found a high incidence of disease in the central to north-eastern parts of the country with the incidence decreasing until absent further south. A high concentration of CMD was also observed on subsistence plantations in central Swaziland. It was proposed that CMD from Mozambique was introduced to South Africa with the movement of infected cassava cuttings southwards and that CMD posed a major threat to commercial growth of cassava in certain areas of the country. More recent

studies have found that the incidence and severity (using a scale of 1-5) of CMD were 61.7% (2.9) and 16.7% (2.3) in Mpumalanga and Kwazulu-Natal provinces respectively (Jericho and Thompson, SASPP conference 1999).

UgV is now reported for the first time to be present in South Africa, Swaziland, Zimbabwe and Mozambique. UgV first detected in severely affected plants within the epidemic area of Uganda and is now believed to be one of the main factors responsible for causing the severe cassava mosaic epidemic in Uganda (Harrison *et al.* 1997, Zhou *et al.* 1997). Monitoring the incidence and distribution of UgV is difficult since it is indistinguishable from ACMV by serological testing and is undetectable using the CCP primers. However using UgV-specific primers (Harrison *et al.* 1997, Zhou *et al.* 1997) it is possible to discriminate between UgV-like viruses and ACMV and EACMV.

It was suggested (Harrison *et al.* 1997) that searches for UgV were best concentrated in areas where ACMV occurs and specifically where the distributions of ACMV and EACMV overlap. UgV had also been found in Sudan (Harrison *et al.* 1997) and Guinae (Winter 1998) and now, from our study, in Mozambique, South Africa, Swaziland and Zimbabwe as well. The exact role of these viruses in causing disease and their interaction with other CBVs would need to be further studied, although it has been shown that UgV was more virulent and accumulated to higher concentrations in cassava (Harrison *et al.* 1997).

In summary, several populations of virus species exist in southern Africa, probably as a result of the movement and trade of infected cassava stakes. This is particularly problematic since it results in the introduction of 'new' viruses into previously uninfected or existing infected areas, and increases the possibilities for mixed infections, pseudo-recombinations, and recombination between viruses. This problem is compounded by the recent finding that mitochondrial COI marker sequences of cassava whitefly populations from southern Africa show a high degree of polymorphism between and within several southern African countries (S. Berry, J.K. Brown and M.E.C. Rey, unpublished results). Thus it is envisaged that the continued movement of infected planting material (both legal and illegal), the continued apparent recombination currently occurring between virus species and strains, and the occurrence of diverse populations of whiteflies could collectively lead to even greater incidences of CMD and increase the diversity and distribution of CBVs leading to greater loss of cassava crops in the future.

2.6 ACKNOWLEDGEMENTS

CHAPTER THREE

MOLECULAR EVIDENCE FOR DISTINCT *BEMISIA TABACI* GEOGRAPHIC GENOTYPES FROM CASSAVA IN AFRICA

This chapter is a modification of the manuscript:

S. Berry, V. N. Fondong, M. E. C. Rey, D. Rogan, C. M. Fauquet and J. K. Brown (2001) Molecular Evidence for Distinct *Bemisia tabaci* Geographic Genotypes from Cassava in Africa. *Molecular Ecology* (submitted and under editorial review)

3.1 Abstract

Bemisia tabaci (Gennadius) is the vector of cassava-infecting begomoviruses that cause cassava mosaic disease (CMD) in Africa and India. Currently at least three distinct begomovirus species have been associated with CMD in Africa, namely *African cassava mosaic virus* (ACMV), *East African cassava mosaic virus* (EACMV) and *South African cassava mosaic virus* (SACMV). The taxonomy of the *B. tabaci* complex is problematic because biological variants are morphologically indistinguishable, making it unclear at this time which genetic and/or biological variants or 'types' of the vector are involved in the transmission of cassava-infecting begomoviruses.

Recent studies, based on polymorphic esterase profiles, RAPD patterns, nuclear ITS, and mitochondria 16S and COI sequences have demonstrated genetic differences within the *B. tabaci* complex, with the latter being particularly useful for resolving higher order differences within the *B. tabaci* complex.

The aim of this study was to examine the relatedness of *B. tabaci* collections from cassava in cassava-growing regions in southern Africa and Cameroon. Adult whiteflies were collected from cropping systems in Cameroon, Zambia, Mozambique, Zimbabwe, Swaziland and South Africa. Mitochondrial COI gene sequences (780 bp) were used as molecular markers to reconstruct phylogenetic trees. Phylogenetic analysis, conducted using maximum parsimony analysis revealed that the *B. tabaci* from Cameroon and southern Africa form a cluster

that is separate from the North African/Mediterranean/Middle East group that also contains African genotypes. Collections from Zambia, Mozambique, South Africa and Swaziland formed a clade of closely related lineages (95-99% COI nt sequence identities), while the Zimbabwean *B. tabaci* grouped most closely with those from okra from Cameroon and Ivory Coast. Results indicate that several distinct genotypes of whitefly vector populations colonise cassava in Africa, and that they group with a strong basis in extant geographic distribution.

3.2 Introduction

Diseases caused by whitefly-transmitted begomovirues, are problematic in many tropical and subtropical regions, and cause heavy losses in production of food and fibre crops, worldwide (Brown 1994). Begomoviruses are ssDNA viruses that belong to family *Geminiviridae* and genus *Begomovirus* (Rybicki 1994; Padidam *et al.* 1995). Current records have shown that the whitefly, *Bemisia tabaci* (Gennadius), can successfully colonise a multitude of plant species worldwide (Cock 1993) and is the only known vector of begomoviruses (Harrison *et al.* 1985). Whereas most *B. tabaci* are polyphagous, some monophagous or nearly monophagous 'biotypes' have been reported (Bird 1957; Bird and Maramorosch 1978).

The occurrence of biotypes or host races were proposed since morphologically indistinguishable variants exhibit differences in host range, dispersal behaviour, fecundity and competency to transmit begomoviruses (Brown and Bird 1992; Brown 1994, Brown *et al.* 1995a). Several distinct 'biotypes' ('A' and 'B') are well studied, including A-like types from the southwest US and northern Mexico (Bedford *et al.* 1994; Brown 1994; Brown *et al.*, 1995a,b; Costa and Brown 1991; Costa *et al.*, 1993). Reproductive incompatibility between a B and a non-B biotype from Sudan was demonstrated experimentally (Byrne *et al.* 1995), and A and B biotypes have also previously been shown not to interbreed (Bedford 1994; Perring *et al.* 1993). In contrast, interbreeding between A and B biotypes (Costa *et al.* 1993; Brown *et al.* 1998; 2001), the B type and two non-B types in

Australia (De Barro and Hart 2000), and between B and Q biotypes from Spain (Ronda *et al.* 2000), have been reported.

The A biotype is indigenous to the south-western US and northwest Mexico (Brown *et al.* 1995a,b; 1999; Costa *et al.* 1993). The B biotype has been shown to be highly fecund, adapted to a wide range of host plants, resistant to pesticides, and is capable of inducing non-viral phytotoxic disorders including silverleaf in curcubits and irregular ripening of tomato (Costa and Brown 1991; Costa *et al.* 1993; Bedford *et al.* 1994; Brown *et al.* 1995a,b).

It has been suggested that the Indian subcontinent is one centre of diversity of *B. tabaci* and that from there this species radiated to other geographical areas, probably millions of years ago (Campbell *et al.* 1996; Cock 1986; Gill 1992) from the introduction of foreign plants, such as cassava from South America to Africa (Brown and Bird 1992; Costa and Russell 1975).

The taxonomy and classification of whiteflies is complex, confounded by the use of the pupae case of the fourth instar for identification of most species (Gill 1992; Rosell *et al.* 1997). Several molecular based methods differentiate different whitefly biotypes or polymorphs within populations. These include the use of isozymes such as esterase (Costa and Brown 1991; Coats *et al.* 1994), RAPDs (Gawel and Bartlett 1993), 16S mitochondrial rDNA sequences (Frohlich *et al.* 1996), 18S rDNA sequences (Campbell *et al.* 1994), and more recently the use of mitochondrial cytochrome oxidase I gene (mtCOI) sequences (Zhang and Hewitt 1996; Frohlich *et al.* 1999). Mating studies, introduction of phytotoxic disorders, different esterase electromorphs, and phylogenetic

analyses using nuclear and mitochondrial genome markers, suggest that *B. tabaci* (Genn.) is a species complex, or a highly cryptic group of sibling species (Frohlich *et al.* 1999).

In recent studies of African whitefly populations, it has been suggested that several indigenous yet uncharacterised *B. tabaci* populations exist (Brown *et al.* 1996). Indian and African *B. tabaci* from cassava have been shown to be reproductively isolated. The most well-studied *B. tabaci* populations in Africa have been on the cassava (*Manihot esculenta* Crantz), since this host plant is the most important carbohydrate source in much of Africa. Whitefly populations were shown to readily colonise cassava after its introduction into Africa from South America several hundreds of years ago (Storey and Nichols 1938) despite the fact that indigenous *B. tabaci* populations from Brazil could not colonise cassava (Costa and Russell 1975).

Subsequently, in the Ivory Coast in west Africa two races or biotypes of *B. tabaci* were described, one 'cassava biotype' having a limited host range comprising wild eggplant and cassava and transmitting *African cassava mosaic virus* (ACMV) and another biotype which has a wide host range, is more polyphagous and does not transmit ACMV (Burban *et al.* 1992). Two other whitefly species in Africa colonise cassava, *Bemisia afer* (Preisener and Hosney) and *Aleurodicus disperses* (Russell) (Legg 1994), but do not transmit cassava mosaic viruses.

At least four begomovirus species, and several strains, have been associated with African Cassava Mosaic Disease (ACMD), which is the most economically important disease of cassava in Africa (Thresh *et al.* 1998). These viruses include ACMV, *East African cassava mosaic virus* (EACMV), Ugandan variant virus (UgV) and EACMV-CM (strains of EACMV-UG), and *South African cassava mosaic virus* (SACMV) (Berrie *et al.* 1998, 2001; Hong *et al.* 1993; Stanley and Gay 1983; Zhou *et al.* 1997; Fondong *et al.* 2000). Cassava begomoviruses are believed to be transmitted by cassava host-adapted *B. tabaci*, but limited experimental evidence for virus-vector co-adaptation exists (Burban *et al.* 1992; Maruthi *et al.* 2000; Colvin *et al.* 1999). Little is known about genetic polymorphisms within *B. tabaci* populations in southern Africa, or whether distinct whitefly populations or biotypes are associated with the different cassava-infecting begomoviruses. No evidence exists for differences in virus transmission efficiencies within or between populations of *B. tabaci* on cassava (Legg 1994).

Recently the existence of the 'B' biotype in South Africa was reported (J. Brown and P. Markham, pers. comm.). The objectives of this study were to examine the relatedness of *B. tabaci* populations from cassava-growing regions in southern Africa and Cameroon, and to compare them with other *B. tabaci* genotypes and biotypes, worldwide.

3.3 Materials and Methods

3.3.1 Whitefly Samples

Bemisia tabaci adults were collected from okra in Cameroon and from cassava in several geographic locations in Cameroon, South Africa, Swaziland, Zimbabwe, Zambia and Mozambique (Table 3.1). In Cameroon, whitefly adults were collected from 11 regions, the majority from the south- western parts of the country. Whitefly isolates were collected from St. Lucia in Kwazulu-Natal, South Africa, from the Maphiveni district in Swaziland, from the Nyamundinn and Katandica regions in Mozambique, from 2 southern regions (Mazowe and Nyabira) in Zimbabwe, and six regions (Kano, Kawi, Masu, Mulim, Shandu and Kato) in Zambia. Adult whiteflies were stored in 70% ethanol at –20 ^{0}C until analysis was performed.

3.3.2 Whitefly DNA extraction, PCR and cloning

Total nucleic acids were extracted from individual whiteflies according to the method of Frohlich *et al.* (1999). Polymerase chain reaction (PCR) was conducted on samples from all geographical locations (Table 3.1). PCR primers for amplifying the mitochondrial cytochrome oxidase I gene (mtCOI) fragments were C1-J-2195 and L2-N-3014 selected from the UBC Insect Mitochondrial DNA Primer Oligonucleotide Set, with sequences obtained from Simon *et al.*, 1994. PCR assay conditions were conducted as described by Frohlich *et al.* (1999).

Table 3.1: Collection sites and accession numbers of whitefly samples

Sample	Region	Country	Accession number
Cam Bertoua WO1	Bertoua	Cameroon	AF344245
Cam Ayos 1 WO2	Ayos	Cameroon	AF344246
Cam Ayos 2 WO3	Ayos	Cameroon	AF344247
Cam Bertoua 2 WO4	Bertoua	Cameroon	AF344248
Cam Yaounde 1 WO5	Yaounde	Cameroon	AF344249
Cam Yaounde 2 WO6	Yaounde	Cameroon	AF344250
Cam BangaBak 2 WO9	Banga-Bakundu	Cameroon	AF344258
Cam KonyeKonye W10	KonyeKonye	Cameroon	AF344251
Cam Fontem W11	Fontem	Cameroon	AF344252
Cam Ekona W12	Ekona	Cameroon	AF344253
Cam Bafous W13	Bafousam	Cameroon	AF344254
Cam Santchou W14	Santchou	Cameroon	AF344255
Cam Bambui W15	Bambui	Cameroon	AF344256
Cam PendaBoko W16	Penda-Boko	Cameroon	AF344257
SA Lucia 1	St. Lucia	South Africa	AF344259
SA Lucia 2	St. Lucia	South Africa	AF344260
SA Lucia 3	St. Lucia	South Africa	AF344261
SA Lucia 4	St. Lucia	South Africa	AF344262
SA Lucia 5	St. Lucia	South Africa	AF344263
SA Lucia 6	St. Lucia	South Africa	AF344264
SA Lucia 7	St. Lucia	South Africa	AF344265
SA Lucia 8	St. Lucia	South Africa	AF344266
SA Lucia 9	St. Lucia	South Africa	AF344267
SA Lucia 10	St. Lucia	South Africa	AF344268
Swa Map 1	Maphiveni	Swaziland	AF344269
Swa Map 2	Maphiveni	Swaziland	AF344270
Swa Map 3	Maphiveni	Swaziland	AF344271
Swa Map 4	Maphiveni	Swaziland	AF344272
Swa Map 11	Maphiveni	Swaziland	AF344273
Swa Map 12	Maphiveni	Swaziland	AF344274
Swa Map 13	Maphiveni	Swaziland	AF344275
Swa Map 14	Maphiveni	Swaziland	AF344276
Swa Map 15	Maphiveni	Swaziland	AF344277
Zim Mazo	Mazowe	Zimbabwe	AF344285
Zim Nyab	Nyabira	Zimbabwe	AF344286
Zamb Kano	Kano	Zambia	AF344280
Zamb Kawi	Kawi	Zambia	AF344281
Zamb Masu	Masu	Zambia	AF344282
Zamb Mulim	Mulim	Zambia	AF344283
Zamb Shandu	Shandu	Zambia	AF344284
Zamb Kato	Kato	Zambia	AF354462
Moz Kat	Katandica	Mozambique	AF344278
Moz Nyam	Nyamundinn	Mozambique	AF344279

mtCOI (~780bp) amplicons were sequenced in both directions using PCR and sequencing primers and an automatic ABI Prism sequencer at the Laboratory for Molecular Systematics and Evolution, University of Arizona, Tucson, AZ 85721, USA.

3.3.3 Phylogenetic analyses

Sequences were aligned using the Clustal algorithm (ClustalW 1.7) (MegAlign, DNASTAR, Madison, Wisconsin, USA), and aligned sequences were evaluated for genetic relatedness by parsimony using PAUP (Swofford *et al.* 1991). Bootstrapping was performed with PAUP using the heuristic option for 1000 replications. Sequences of the whiteflies, *Trialeuroides vaporariorum* (Westwood) and *Bemisia afer* (Preisener and Hosney) were used as the outgroups.

3.4 Results

A single most parsimonious tree (Figure 3.1) was reconstructed from mtCOI sequences. A table (Table 3.2) indicating genetic relatedness (percentage nucleotide sequence identity) of the COI sequences of *Bemisia tabaci* is also presented. Figure 3.1 indicates that the mtCOI sequence was useful for separating all New World *B. tabaci* from four other groups of Old World *B. tabaci*. From this study, *B. tabaci* from cassava in Cameroon and southern Africa clustered as three distinct lineages: *B. tabaci* from Cameroon; *B. tabaci*

from Zambia, South Africa, Mozambique and Swaziland; and *B. tabaci* from Zimbabwe which clustered with others from the Mediterranean region, northern Africa (Sudan, Morocco), Turkey, Spain, Israel, and a group which also contained the B biotype sub-cluster, including a B biotype from tomato in South Africa. Whitefly from cassava in Ivory Coast grouped independently from the Cameroon group (Figure 3.1). *B. tabaci* from Africa shared less than 10% sequence identity with *B. tabaci* in the Far East (Nepal, Pakistan, India, Malaysia and Thailand) group.

It is interesting to note that *B. tabaci* from cassava from Mazowe and Nyabira regions in Zimbabwe grouped separately from the other southern African *B. tabaci* individuals, in that they were more closely related to the northern African/Mediterranean/B biotype group than the other southern African *B. tabaci*. Furthermore, they were closely related (97% and 98% nucleotide sequence identity) to *B. tabaci* sequences from okra from the Ivory Coast and Cameroon, respectively (Table 3.2). *B. tabaci* from the two regions in Zimbabwe were highly similar (99% nt sequence identity), indicating that they probably radiate from the same geographically related genotype or topotype.

Bootstrap

T vaporariorum
B afer
AZ(A) Arizona
CUL Mexico
Honduras
El Sal Sweetpotato
MX Tobacco Tuxtla Mexico
JAT Puerto Rico
ARG Salta
Bolivia
NEW-Nepal
PC92 Pakistan
IW India
PC91 Pakistan
PC95 Pakistan
Malaysia Malva M94047
Thailand Cotton M94024
Morocco 1
Morocco 2
SC Sudan
SP95 Spain
SP99 Spain
TC Turkey
Ivory Coast Okra
Cam BangaBak 2 Okra
Zim Mazo
Zim Nyab
AZ(B) Arizona
IS (B) Israel
SA99 (B) South Africa
ABA Benin
Tanzania Pepper
Tanzania Cabbage
SA Lucia 10
SA Lucia 2
SA Lucia 3
SA Lucia 5
SA Lucia 6
SA Lucia 8
Swa Map 15
SA Lucia 9
Swa Map 1
Swa Map 11
Swa Map 12
Swa Map 13
Swa Map 14
Swa Map 2
Swa Map 3
Swa Map 4
Moz Kat
Moz Nyam
Zamb Kato
SA Lucia 1
SA Lucia 4
SA Lucia 7
Zamb Kano
Zamb Kawi
Zamb Masu
Zamb Mulim
Zamb Shandu
Cam Ayos1 WO2
Cam Ayos2 WO3
Cam Bafous W13
Cam Bertoua 2WO4
Cam Bertoua WO1
Cam Ekona W12
Cam Fontem W11
Cam Konye W10
Cam Santchou W14
Cam Yaounde 1WO5
Cam Yaounde 2WO6
Cam Bambui W15
Cam Penda Boko W16
Ivory Coast cassava
Uganda SwPot

Bootstrap values: 98, 100, 70, 100, 70, 86, 85, 85, 100, 71, 100, 100, 100, 74, 84, 99, 99, 90, 96, 97, 100, 92, 73, 66, 99, 88, 98, 94, 67, 83, 100, 90, 99

Figure 3.1: Phylogenetic Tree showing the mtCOI relationships between *B. tabaci* isolates from cassava in southern Africa and Cameroon and *B. tabaci* from other countries and plant hosts

Table 3.2: Percentage nucleotide sequence identities between selected southern African whiteflies from cassava and other host plants from various African countries

	SA Lucia 4	SA Lucia 5	SA Lucia 9	Swa Map 14	Swa Map 15	Moz Nyam	Zamb Shandu	Zamb Kato	Zamb Kano	Zim Mazo
Cam Ayos 1 WO2	90	90	92	92	91	92	92	92	92	84
Cam Ekona W12	90	90	92	92	91	92	92	92	92	84
Cam Penda-Boko W16	90	90	92	92	91	92	92	92	92	83
Ivory Coast cassava	85	85	86	87	86	87	87	87	86	86
Ivory Coast okra	82	80	82	82	82	82	82	82	82	97
Cam Banga-Bak 2 okra	81	81	82	83	82	83	82	83	82	98
SA99(B) South Africa	82	82	83	84	83	84	84	84	83	93
IS (B) Israel	82	82	83	84	83	84	84	84	84	94
Morocco2	82	82	83	83	83	83	83	84	83	96
Tanzania93 Cabbage	96	97	99	99	98	99	97	99	97	83
Tanzania93 Pepper	96	96	98	98	97	98	97	99	97	82
T. vaporariorum	73	73	75	75	75	75	75	75	75	74

Whitefly isolates from St. Lucia region in South Africa (SA) formed a tight cluster of closely related genotypes that also include whitefly sequences from Swaziland, Zambia and Mozambique (Figure 3.1). The South African whiteflies share a nt sequence percentage identity of > 95% (data not shown), suggesting that they represent a single topotype. Furthermore, the SA whitefly isolates share a mtCOI sequence identity of > 95% with mtCOI sequences of whitefly isolates from Mozambique, Swaziland and Zambia, thereby suggesting that *B. tabaci* from southern African countries shared a common topotype lineage.

Comparisons of mtCOI sequences of whitefly from cassava and tomato (SA99B, Figure 3.1) in South Africa (Pietersen *et al.* 2000), revealed a nt sequence identity of only 81-82% (Table 3.2). The B biotype from tomato is an invasive population introduced only recently to the country from ornamental plants. Also interesting is that cassava *B. tabaci* from the southern African group i.e. SA, Mozambique, Swaziland and Zambia are closely related (96% nt sequence identity) to pepper and cabbage whitefly from Tanzania (Table 3.2). Whitefly from the 6 regions within Zambia also separate as a tight cluster of highly related individuals (>98% nt sequence identity in the COI regions), their closest relatives (>93% nt sequence identity) being *B. tabaci* from SA, Swaziland, Mozambique and Tanzania.

The whitefly mtCOI sequences from Cameroon, while closely related (90-92% nt sequence identity) to the southern African cassava *B. tabaci*, form a separate tight group of highly related genotypes (>91%). Comparisons of mtCOI sequences of *B. tabaci* from two different host plants, okra and cassava, in

Cameroon, demonstrated 83 % nt sequence identity (Table 3.2). Comparison of cassava whitefly mtCOI sequences from Cameroon and Ivory Coast in west Africa, show a nt sequence identity of only 89% which is lower than the identity exhibited by the southern African whitefly COI sequences (>95%). It was also evident from the phylogenetic tree (Figure 3.1) that the whitefly isolates from the Penda-Boko and Bambui regions of Cameroon, while closely related (95%), showed a comparatively low mtCOI sequence identity (89%) with whitefly isolates from other regions of Cameroon.

3.5 Discussion

In this study, the mitochondrial COI sequences were useful for distinguishing between geographical genotypes (topotypes) of whitefly on cassava in Africa. COI markers clearly separated Old World, southern African and Cameroon whitefly into three groups (Figure 3.1), and clearly demonstrated a notable divergence in sequences of *B. tabaci* that infest cassava in these regions. Molecular analyses using mtCOI sequences to distinguish the global variants of *B. tabaci*, and biochemical, behavioural and transmission studies, have suggested that *B. tabaci* is a cryptic whitefly that probably constitutes a species complex (Brown *et al.*, 1995a; Frohlich *et al.* 1999; Brown *et al.*, 2001). Our results support this hypothesis, and provide the first molecular evidence for the existence of geographically-limited genotype groups, or 'topotypes', within the *B. tabaci* colonising cassava in southern Africa.

It was not surprising that *B. tabaci* isolates collected from cassava from St. Lucia, South Africa, were more closely related to genotypes from cassava in Cameroon (>90%) compared with a B biotype whitefly isolated from tomato (SA99B) (82%) also in South Africa, since the B biotype is not indigenous to South Africa and is likely to have been introduced onto tomato. A host feeding preference for cassava is also supported in southern Africa, since host feeding preferences have been shown in Ivory Coast, where a cassava-restricted whitefly "biotype" demonstrated a limited host range comprising wild eggplant and cassava (Burban 1992), and since it was also observed that whitefly collected from tomato in SA would not feed on cassava (Dr. Gerhard Pietersen, Plant Protection Institute, SA, pers. comm.).

What is also interesting to note is that whitefly from SA were highly related (96%) in their mtCOI sequences to whitefly from cabbage and pepper in Tanzania (collections kindly provided, courtesy J. Legg, IITA). What may also be significant from this study is the observation that whitefly from cassava in Zimbabwe while exhibiting 80-86% nt sequence identity with whitefly from Cameroon and southern African countries, was closely related to mtCOI sequences of whitefly from eggplant from Ivory Coast and Cameroon (97% and 98%, respectively) and tomato from SA (93%).

Thus it is conceivable that whitefly may be able to move from cassava to alternate hosts, such as okra, tomato, pepper and cabbage, and undergo host-adaptation. While it is not impossible that the reverse may occur i.e. whitefly may disperse to other plant hosts, evidence to date has shown that cassava-

adapted whitefly are monophagous or near-monophagous, and many experiments conducted in our laboratory in SA in the late 1980's showed that whitefly isolated from tomato, tobacco and several other plant hosts, would not feed on cassava (data not shown). However, careful studies to determine host range and host preferences need to be performed in order to further elucidate questions regarding cassava-whitefly adaptations.

B. tabaci from nine of the eleven regions in Cameroon (Table 3.1) grouped as a tight cluster of closely related COI sequences (>90% nt sequence identity) while those from Bambui and PendaBoko, formed a separate branch of somewhat more divergent genotypes (89%). Because all of the geographical regions in Cameroon from where samples were collected are in the south-western parts of the country, except for Bertoua, the reason for this discrepancy is unknown. However, it would be informative to compare these whitefly with those from other regions within Cameroon in subsequent studies.

In addition to reports of the major species of cassava begomoviruses, namely ACMV, SACMV and EACMV in Africa (Berrie *et al.* 1998, 2001; Hong *et al.* 1993; Stanley and Gay 1983; Zhou *et al.* 1997), recent studies have further established high sequence diversity among cassava begomoviruses in southern and eastern Africa (Zhou *et al.* 1998; Berry and Rey 2001).

The demonstration of cassava begomovirus diversity, and the evidence in this study for the existence of genetic diversity amongst the *B. tabaci* species isolated from a host plant, cassava, from several geographical regions, supports the hypothesis that virus-vector co-adaptation may have occurred in the cassava-

begomovirus pathosystem. The fact that cassava viruses have few alternate hosts, the continual breeding of new cassava cultivars, and geographical movement of infected cassava stakes across Africa are probably factors driving the evolution of 'new' whitefly topotypes within the highly virulent *B. tabaci* species complex. Interbreeding between these highly related topotypes is highly likely since successful offspring production has been demonstrated between different biotypes (Brown *et al.*, 2001; Byrne *et al.*, 1995; Costa *et al.*, 1993; De Barro and Hart 2000). Studies on the virus transmission of EACMV-Tanzania, ACMV-Uganda, UgV-Uganda and ICMV-India by different geographically-isolated whitefly colonies, demonstrating different transmission efficiencies based on whitefly population and virus strain/species, further supports a co-adaptation hypothesis (McGrath and Harrison 1995, Maruthi *et al.* 2000).

This study presents convincing molecular evidence for three distinct groups of *B. tabaci* associated with cassava in southern Africa and Cameroon. In general, groups are delineated by geographic affiliation though evidence also suggests that distinct genotypes are not as tightly delimited by sub-geographies within southern Africa, as might have been expected. As has been noted for begomoviruses (Rybicki 1994; Padidam *et al.* 1995), their whitefly vector 'types' also appears to be delimited by geographical boundaries that correspond to similar extant origins. However, a direct comparison of begomovirus species (or strains) and vector 'topotypes' (and biotypes) is needed to further evaluate the

hypothesis that *B. tabaci* has a compelling influence on the transmission and evolution of cassava-infecting begomoviruses in Africa.

3.6 Acknowledgements

CHAPTER FOUR

DIFFERENTIATION OF CASSAVA-INFECTING BEGOMOVIRUSES USING HETERODUPLEX MOBILITY ASSAYS

This chapter is a modification of the published paper:

4.1 ABSTRACT

Cassava mosaic disease is an enormous problem affecting the viability and productivity of cassava in all growing regions in Africa. The disease is now known to be caused by a number of cassava-infecting begomoviruses. New viruses and viral strains continue to be discovered and due to the lack of cloning and sequencing facilities in many African laboratories, a simple, rapid and sensitive technique is needed for screening of cassava plantations. Here we report on the development of a heteroduplex mobility assay (HMA) which could be used for testing of viral-infected cassava. The assay involves amplifying the highly conserved core region of the coat protein gene of field isolates followed by denaturing and annealing with a number of reference strains. The HMA profiles in this study were able to differentiate four different viral species and eleven different virus strains, and show a good correlation with sequencing results and phylogenetic comparisons with other sequenced cassava viruses. This technique is sensitive and rapid and has the added advantage of being able to detect mixtures of viruses in field-grown cassava.

4.2 INTRODUCTION

The *Geminiviridae* is a family of plant viruses with ssDNA genomes that are divided into three genera based on genome organisation, host range and insect vector (Harrison *et al.* 1977, Van Regenmortel *et al.* 1997). The majority of the viruses belonging to the genus *Begomovirus* are bipartite (have two genomic components, DNA-A and -B) although a small minority have been found that are monopartite (only having DNA-A) (Navot *et al.* 1991). Within this genus there exist a number of viruses that have been found to affect cassava (*Manihot esculenta* Crantz).

Cassava is the third most abundant carbohydrate food source in the World and is the most important food crop grown in Africa (Fauquet *et al.* 1990). To date, eighteen different viruses have been found that affect cassava, and of these the most significant are those that cause Cassava Mosaic Disease (CMD) (Thresh *et al.* 1994). Cassava Mosaic Disease has recently been shown to be caused by a number of cassava-infecting begomoviruses. The first and most well studied is *African cassava mosaic virus* (ACMV) which was demonstrated as the causal agent in 1978 (Bock *et al.* 1978).

Early serological (Harrison *et al.* 1986) and nucleotide sequence (Hong *et al.* 1993) comparisons showed three distinct virus groups, those from East Africa referred to as *East African cassava mosaic virus* (EACMV), those from West Africa as ACMV and the more distantly related viruses from India and Sri-Lanka named *Indian cassava mosaic virus* (ICMV).

However this notion of virus groups only occurring in defined geographical areas has been challenged due to the discovery of distinct new viruses, such as *South African cassava mosaic virus* (SACMV) in South Africa (Berrie *et al.* 1998) and a recombinant virus (the Uganda variant, UgV) causing the current epidemic affecting cassava in Uganda (Zhou *et al.* 1997). Furthermore, heterogeneity within viral species has been identified and in addition mixtures of viruses have been found co-existing simultaneously in single cassava plants (Fondong *et al.* 2000).

Geminiviruses have recently been found to evolve much more rapidly than expected. Changes occur by mutations, recombination of dsDNA intermediates and pseudo-recombination (re-assortment of genome components) (Harrison 1985, Padidam *et al.* 1999, Fondong *et al.* 2000). This has led to an increase in virus diversity in field cassava and has necessitated the need for development of an effective, rapid and reliable technique for differentiating between theses numerous, constantly emerging viral isolates. Currently available immunological and molecular techniques, although useful, are expensive, time-consuming, and require specialised and expensive equipment. One possible technique to overcome these problems may be Heteroduplex Mobility Assays (HMAs).

The Heteroduplex Mobility Assay is a sequence difference detection technique that has been used for detecting polymorphisms in a number of genes (Cai *et al.* 1991) as well as for characterising a number of human RNA viruses, such as measles virus (Kreis *et al.* 1997), poliovirus (Chezzi *et al.* 1996) and most

widely for Human Immunodeficiency Virus (HIV) (Delwart *et al.* 1993). HMAs were found to be rapid, relatively inexpensive, easy to use and had a high degree of specificity (Delwart *et al.* 1995). Heteroduplex mobility analysis involves the differentiation of two viral samples based on their degree of sequence complementarity. Heteroduplexes (two different viral strains/species) are distinguished from homoduplexes (two identical viruses) by their reduced mobility during electrophoresis (Innis *et al.* 1995). A semi-quantitative relationship exists between the degree of retardation and the degree of sequence divergence.

Other mutation detection techniques, such as single stranded conformation polymorphism (SSCP) (Rosner *et al.* 1999) and S1 nuclease treatment (Rosner *et al.* 1998) have recently been used for characterising plant RNA viruses such as *Prunus necrotic ringspot virus* (PNRSV) and *Potato Virus Y* (PVY) respectively. The HMA technique has been used to analyse variability in *Zucchini yellow mosaic virus* (Lin *et al.* 2000) but to our knowledge, no work has been done on using HMAs for characterising plant DNA viruses.

The aim of this work was to develop and optimise this technique for differentiating between strains of cassava-infecting begomoviruses. The region of the genome used for analysis was the core region of the coat protein (CCP) gene on DNA-A. The CCP region was chosen because of it's ideal size of approximately 550 bp, it's highly conserved nature and evidence that the CCP region sequence is as informative in terms of phylogeny as the sequence of the entire genome (Wyatt *et al.* 1996).

4.3 MATERIALS AND METHODS

4.3.1 Virus samples

Cassava samples suspected of being infected with cassava-infecting begomoviruses and displaying the characteristic stunting and mosaic leaf symptoms, were collected from a number of cassava-growing regions in South Africa (St. Lucia, Tonga and Hoxane) and subjected to heteroduplex analyses.

4.3.2 Total Nucleic Acid Extraction

Total nucleic acid (TNA) was extracted from the leaf samples according to the method of Doyle and Doyle (1987) and resuspended in 50 ul 1x TE buffer containing 20 ug RNase A.

4.3.3 PCR

The 550 bp CCP region of DNA-A was amplified by PCR using a pair of degenerate primers, AV514 and AC1048 (Wyatt *et al.* 1996). These primers bind to highly conserved regions within the core region of the coat protein and are specific for most begomoviruses. Amplification was carried out in a final reaction volume of 50 ul containing approximately 500 ng TNA, 30 pmol of each primer, 1.5 mM $MgCl_2$, 0.5% Tween 20, 0.001% gelatin, 200 uM each dNTP and 2.5 U *Taq* DNA Polymerase (Roche Molecular Biochemicals). Amplification was carried out in a Hybaid Express Thermal Cycler using a modified 'hot start' and cycling conditions of: 95°C for 1 min., 55°C for 2 min.

and 72°C for 1 min. for 30 cycles followed by a final extension of 72°C for 5 min. The amplification products were examined by electrophoresis on 1% agarose gels which were then stained in ethidium bromide and viewed on a UV transilluminator.

4.3.4 Heteroduplex Mobility Assays

Eleven isolates collected from the various geographical areas were used for analysis by comparison to three reference viruses viz. SACMV (SACMV-CP clone, kindly supplied by L. Berrie, University of the Witwatersrand, South Africa), ACMV-Nigeria (pNIG008 DNA-A clone, courtesy of J. Stanley, John Innes Centre, UK) and EACMV-Malawi (courtesy of D. Robinson, Scottish Crop Research Institute, Scotland) which represent three of the main groups of cassava-infecting begomovirus species currently known to infect cassava in Africa.

For heteroduplex analysis under optimal conditions, equal amounts (500 ng, ±5 ul) of each of the type and reference strains PCR products were added together in thin-walled Hybaid PCR tubes with 1.1 ul of 10x heteroduplex annealing buffer (1M NaCl, 100 mM Tris-HCl pH 8.0, 20 mM EDTA). These heteroduplex mixtures were then heated at 94°C for 2 min. in a Hybaid thermal cycler and rapidly cooled on ice for 10 min. A 2 ul aliquot of 6x gel-loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 40% (w/v) sucrose) was added to each mixture and then loaded onto 5% polyacrylamide gels (prepared from a 30:0.8 acrylamide :bisacrylamide stock) with 0.5x TBE (0.045M Tris-borate,

0.001M EDTA) running buffer and electrophoresed at 250 V for 4-4.5 hrs, until the bromophenol blue dye front reached the bottom of the gel. The gels were then removed and stained for 10 min. in 0.5x TBE containing 0.5 ug/ml ethidium bromide and the homo- and heteroduplexes visualised on a UV transilluminator and photographed. For the control reactions, PCR products from each isolate, without the reference strains, were individually heated and cooled.

4.3.5 Optimisation of HMA reaction conditions

A number of different reaction conditions were investigated in order to fully optimise the HMAs for cassava begomoviruses. The parameters investigated were: the effect of denaturing and annealing temperatures, the effect of DNA concentration, the effect of salt concentration and the differences between non-denaturing polyacrylamide gels and Mutation Detection Enhancement (MDE™) gels (AT Biochem, Malvern, PA, USA). MDE™ gel matrix is a polyacrylamide-like matrix that has a high sensitivity to DNA conformational differences. It is claimed to be able to detect sequence differences from as low as 15% up to 80% and reduce the time required for optimising electrophoretic conditions, as compared to polyacrylamide gels (FMC Product Catalogue).

Five isolates were chosen for the optimisation reactions. All five had previously produced heteroduplex patterns using the HIV protocol (Delwart *et al.* 1993) and all five heteroduplex patterns were sufficiently different to each other. For testing the effect of heating and cooling on heteroduplex formation

the reaction mixtures were heated for different times (either 2 min or 5 min) using both a waterbath and a thermal cycler and cooled, either rapidly cooled on ice or slow cooled.

The amount of DNA used for the HMAs was important. Since the heteroduplex bands often stained only 25-30% of the intensity of the homoduplex bands, sufficient DNA needed to be loaded in order to be able to visualise the heteroduplexes. However loading too much DNA on the gel also had the effect of the hetero- and homoduplexes merging. Three different concentrations of PCR products were tested i.e. 500 ng, 600 ng and 800 ng, in order to find an ideal concentration which gave visible heteroduplexes without smearing.

The standard heteroduplex annealing buffer contained 1M NaCl. It appeared that NaCl was crucial for heteroduplex formation (Helba Bredell, National Institute for Virology, South Africa, pers. comm.) and that decreasing the salt concentration might be useful for comparison between more distantly related strains. Thus concentrations of 0.5 M and 1M NaCl were tested.

4.3.6 Cloning and Sequencing

The CCP PCR products of the virus isolates that were characterised by heteroduplex analyses were blunt-end cloned into a pBluescript cloning vector using standard methods (Sambrook *et al.* 1989) and sequenced by automated sequencing at the University of the Witwatersrand sequencing unit (ABI PRISM Big Dye Terminator Cycle Sequencing, Perkin Elmer). Sequences were

analysed in order to establish a correlation between sequence identity and heteroduplex mobility profiles. Sequence identity comparisons were made with other cassava-infecting begomoviruses (see list and accession numbers in Figure 4.7 legend). Phylogenetic trees were constructed using the DNAMAN version 4.0 (Lynnon Biosoft ©, Quebec, Canada) full optimal alignment and neighbour-joining method options with 5000 bootstrap replications.

4.3.7 Construction of Standard Curves

From the results of the HMAs using cloned CCP PCR products, standard curves were constructed for each of the three reference viruses. The standard curves were least-squares exponential curves of heteroduplex mobility vs. genetic distance. The heteroduplex mobilities were determined by calculating the ratios of the distances migrated by the heteroduplexes over that migrated by the homoduplexes. The genetic distances were determined by aligning the CCP sequences of the various isolates to the reference sequences using DNAMAN.

Since there was no acceptable means of weighting gaps introduced by the software package, when aligning the cassava virus sequences in this study all gaps that would have been introduced by the software were removed. Phylogenetic trees were then constructed using the genetic distance data from the standard curve and a distance-based phylogenetic method (FITCH, Fitch-Margoliash and Least-Squares Distance) from the PHYLIP (Phylogeny Inference Package, version 3.5c) software package.

4.4 RESULTS

4.4.1 Optimisation of HMA reaction conditions

CCP amplification products were amplified from several infected leaf samples and it was found that the PCR products were of sufficient yield and purity for use in HMAs (Fig. 4.1). It was however noted that variations occurred in the amount of PCR products depending on the amount of viral DNA in the field sample. Therefore in subsequent HMAs, the concentration of PCR products were adjusted to allow equal concentrations of reference and test samples.
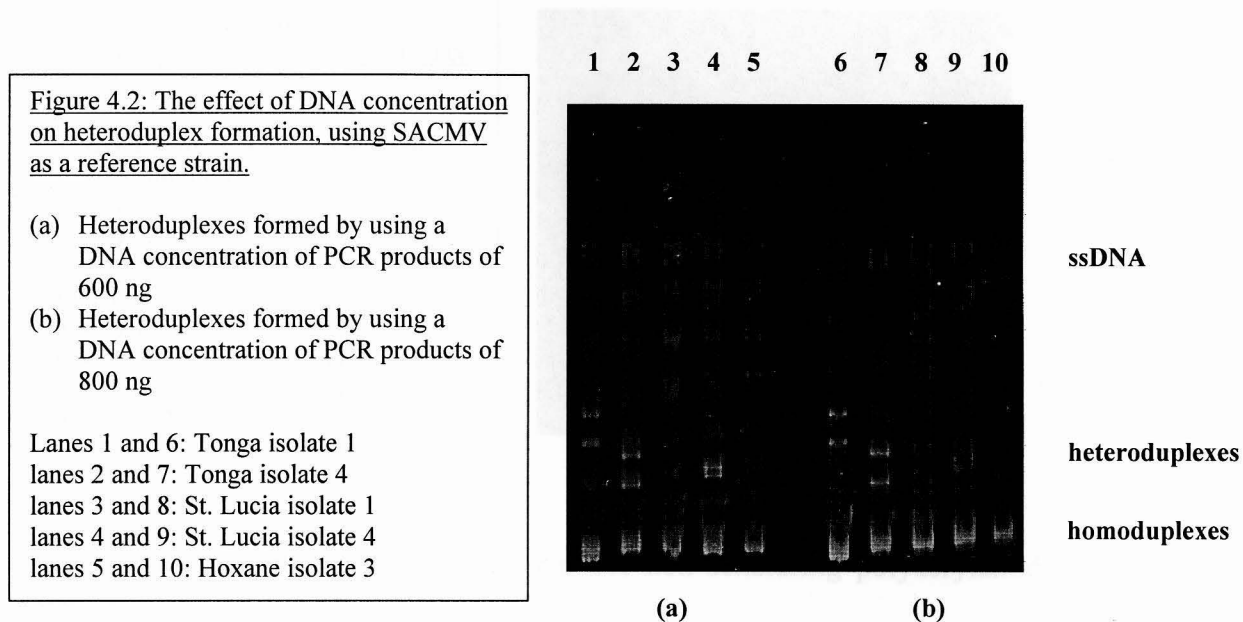

Figure 4.1: Amplification of the 580 bp CCP fragment from infected cassava plants by PCR.

Lane 1: Tonga isolate 1
lane 2: Tonga isolate 4
lane 3: St. Lucia isolate 1
lane 4: St. Lucia isolate 4
lane 5: Hoxane isolate 3
lane 6: *Pst* I-cut lambda DNA Mw markers.

580 bp CCP
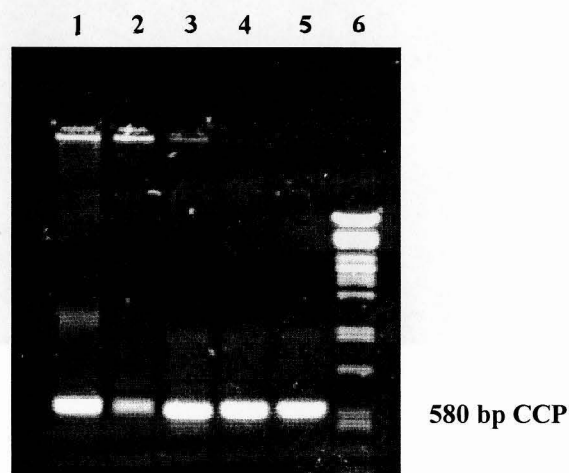
The initial protocol for developing an optimal HMA employed an already-established protocol used for subtyping HIV-1 isolates (Delwart *et al.* 1993, Helba Bredal, NIV, per. comm.), which was then altered and optimised for our own use. The initial use of this protocol with our viral isolates was promising, however a number of conditions known to affect HMAs needed to be further investigated in order to assess their relative importance and improve on the technique. In this study it was found that the length of time used for denaturing
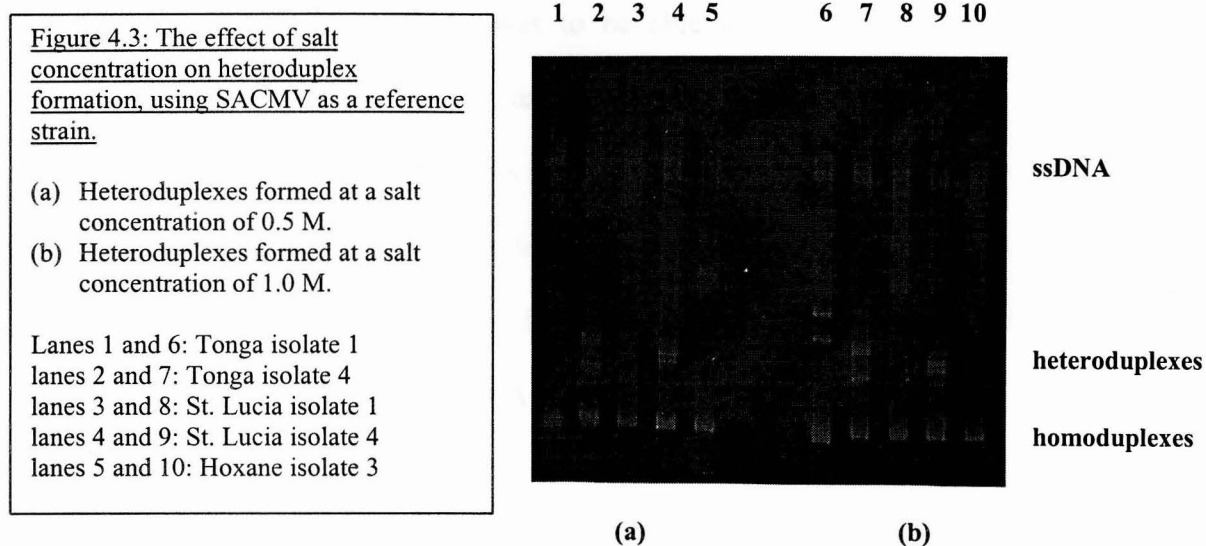
had little effect, provided the time was sufficient for adequate denaturing of dsDNA. However, the heteroduplex bands produced by using the thermal cycler, as opposed to the waterbath, were much more distinct and clearer. Slow cooling, which is more conducive to the formation of heteroduplexes between closely related samples, resulted in the occurrence of no heteroduplexes for any of our isolates tested, in contrast to rapid cooling on ice which yielded better results.

Figure 4.2: The effect of DNA concentration on heteroduplex formation, using SACMV as a reference strain.

(a) Heteroduplexes formed by using a DNA concentration of PCR products of 600 ng
(b) Heteroduplexes formed by using a DNA concentration of PCR products of 800 ng

Lanes 1 and 6: Tonga isolate 1
lanes 2 and 7: Tonga isolate 4
lanes 3 and 8: St. Lucia isolate 1
lanes 4 and 9: St. Lucia isolate 4
lanes 5 and 10: Hoxane isolate 3

In the case of DNA concentration it was ascertained that using 5ul of PCR product (approximately 500 ng) per sample was optimal. Increasing the amount of DNA from 500 ng to 600 ng or 800 ng made no significantly visible difference in most cases (Fig. 4.2, lanes 1-3, 5, 6-8, 10). Only in one instance (Fig. 4.2, lanes 4 and 9) was a visible difference noted.

Results of varying the salt concentration demonstrated that decreasing the salt concentration to 0.5 M still gave heteroduplex bands although they appeared more faint and less clear, particularly the Tonga isolates 1 and 4 (Fig. 4.3 lanes

1, 2, 6, 7). However the St. Lucia isolates (lanes 3, 4, 8, 9) and Hoxane isolate (lanes 5, 10) did not appear to vary significantly. Therefore either concentration of NaCl could be used in the annealing buffer but it is advisable that 0.5 M NaCl be used if some of the isolates, particularly the more distantly related ones, fail to produce heteroduplex patterns upon analysis.

Figure 4.3: The effect of salt concentration on heteroduplex formation, using SACMV as a reference strain.

(a) Heteroduplexes formed at a salt concentration of 0.5 M.
(b) Heteroduplexes formed at a salt concentration of 1.0 M.

Lanes 1 and 6: Tonga isolate 1
lanes 2 and 7: Tonga isolate 4
lanes 3 and 8: St. Lucia isolate 1
lanes 4 and 9: St. Lucia isolate 4
lanes 5 and 10: Hoxane isolate 3

Electrophoresing the samples on 25 cm 5% non-denaturing polyacrylamide gels gave satisfactory results with good separation between the homo- and heteroduplexes. The polyacrylamide gels were made from a 30:0.8 acrylamide:bisacrylamide stock solution, the decreased amount of bisacrylamide is thought to result in less cross-linking which improves heteroduplex mobility shifts. Increasing the percentage of the acrylamide to 8%, in an attempt to increase the resolution, did not significantly improve the heteroduplex patterns obtained and, in some cases, the heteroduplexes appeared fainter (results not shown). No significant improvement using the MDE™ gels compared to polyacrylamide gels was observed (data not shown). Using MDE™, the samples

took much longer to run (20 hrs for MDE™ compared to 4 hrs for polyacrylamide) and did not stain as well. Polyacrylamide gels gave better resolution in a shorter space of time and were cheaper than MDE™ gels.

4.4.2 Heteroduplex analysis of cassava samples

The primary aim of using HMAs was to be able to undertake large-scale screenings of infected field-collected cassava using TNA containing viral DNA as the starting material. The first HMAs in this study were performed using CCP amplification products from plasmid-derived CCP clones of a proven single species entity to check accuracy and reliability. These results were then compared to those obtained using TNA from infected leaf tissue as the source material. Cloned and sequenced CCP genes of eleven viral isolates, collected from various regions within South Africa, were used in the HMA by comparison to three distinct reference viruses: SACMV, ACMV-Nigeria and EACMV-Malawi. These same isolates were also amplified from corresponding TNA and compared to the same three reference viruses.

Heteroduplex patterns and ratios of the isolates from the two different sources were compared (results not shown). The results agreed well with one another, thus amplification from TNA as a starting source using the degenerate primers produces amplified products of a single species that mirrors that of the cloned products, except when the plant contained mixtures of viruses.

Heteroduplex patterns were obtained by comparing various isolates from South Africa to the SACMV reference virus. From the results (Fig. 4.4) a

number of distinct heteroduplex patterns can be observed. The four isolates from the St. Lucia area (lanes 7-10) produced heteroduplexes with similar mobilities to each other, signifying that they were very closely related, if not identical, viruses. The three isolates from Tonga (lanes 4-6) showed significantly different profiles to each other, suggesting that these three isolates, although collected from the same area, were different from each other and were distinguishable from SACMV. The third group of samples, collected from Hoxane (lanes 1-3), exhibited multiple bands in lanes 2 and 3 demonstrating a mixed infection of two or more viral species or strains. The presence of two viral strains was indeed confirmed by sequencing of the cloned CCP products (data not shown).

Figure 4.4: Heteroduplex Mobility Assay of cassava-infecting begomovirus isolates amplified from TNA tested against SACMV reference strain

Lane1: Hoxane isolate 1
lane 2: Hoxane isolate 2
lane 3: Hoxane isolate 5
lane 4: Tonga isolate 1
lane 5: Tonga isolate 3
lane 6: Tonga isolate 2
lane 7: St. Lucia isolate 1
lane 8: St. Lucia isolate 2
lane 9: St. Lucia isolate 3
lane 10: St. Lucia isolate 5
lane 11: St. Lucia isolate 6
lane 12: control (SACMV)
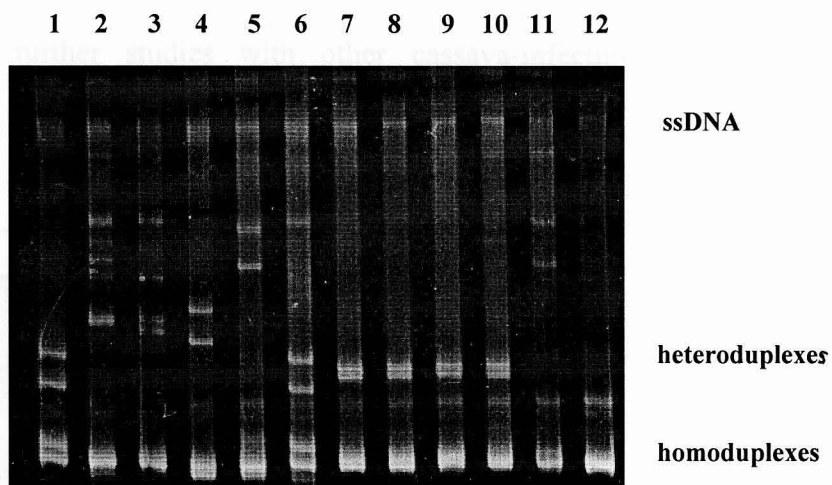
Comparison of the same eleven isolates against the two other reference viruses (EACMV-Malawi and ACMV-Nigeria) (results not shown) further supported the validity of the HMA technique. The isolates from St. Lucia showed similar heteroduplex profiles to each other and their patterns suggested that they were more closely-related to EACMV-Malawi than ACMV-Nigeria which was shown by their clustering with EACMV-Malawi in the phylogenetic tree (Fig. 4.7).

Two isolates (Hoxane 1 and Tonga 2) formed heteroduplexes when hybridized with EACMV-Malawi but failed to produce heteroduplexes with ACMV-Nigeria. The Hoxane 1 and Tonga 2 isolates differed in nucleotide sequence similarity from EACMV-Malawi by 6.73% and 5.27% respectively, illustrating that the sensitivity of this technique was ~5%. However, lower detection limits of 2.3% were achieved in further studies with other cassava-infecting begomovirus field isolates (results not shown).
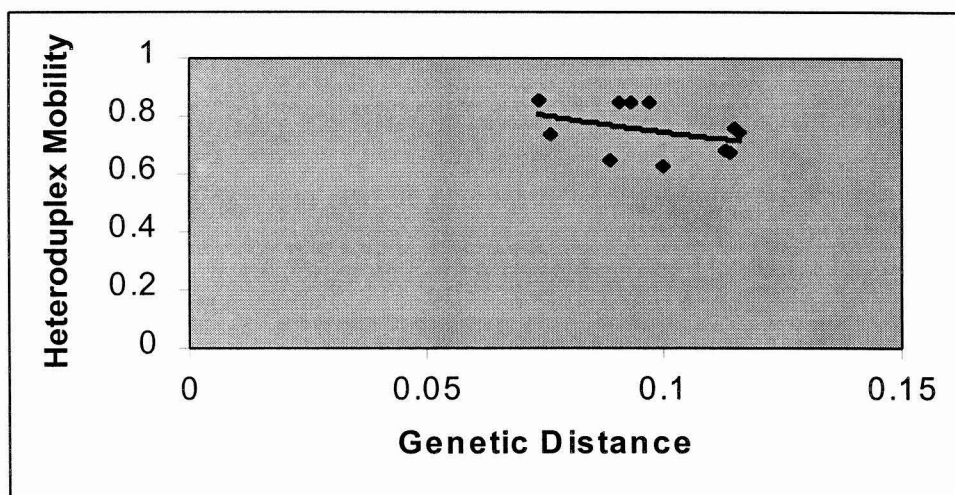

Figure 4.5: Standard Curve generated by plotting heteroduplex mobility vs. genetic distance from the analysis of cloned CCP products of field isolates tested against the SACMV reference strain.

After performing HMAs using cloned PCR products, standard curves were constructed as described in Materials and Methods. In Figure 4.5 we see an example of one such curve using SACMV as the reference. The curve was linear with the heteroduplex mobility values decreasing as the genetic distances increased. The distribution and segregation of the eleven isolates along the curve were clearly visible. The viral isolates exhibited genetic differences to SACMV in the range of 7.5%-12.5% which resulted in heteroduplex mobilities of between 0.6 and 0.9. When standard curves were constructed for these same isolates compared with EACMV and ACMV reference strains similar trends were observed (results not shown).

Examination of the distance-based phylogenetic tree (Fig. 4.6) showed that the three isolates from St. Lucia segregated very closely together and are considered almost identical (nt sequence similarity of 97-98%), and therefore we would expect them to give very similar heteroduplex patterns and mobilities. This was indeed clearly observed in Figure 4.4 (lanes 7-9). Hence the St. Lucia virus isolates can be accurately classified as closely related strains, different from SACMV. Similarly, this can also be demonstrated by comparing the clustering of the three isolates from Tonga with their corresponding HMA results (Fig. 4.4 lanes 4-6). The three Tonga isolates fell within separate branches (nt sequence similarity of 75-82%) and were therefore deemed to be distinct species. This was reiterated in the HMA results where we see that the three isolates all gave different heteroduplex patterns and mobilities, with the most closely related ones
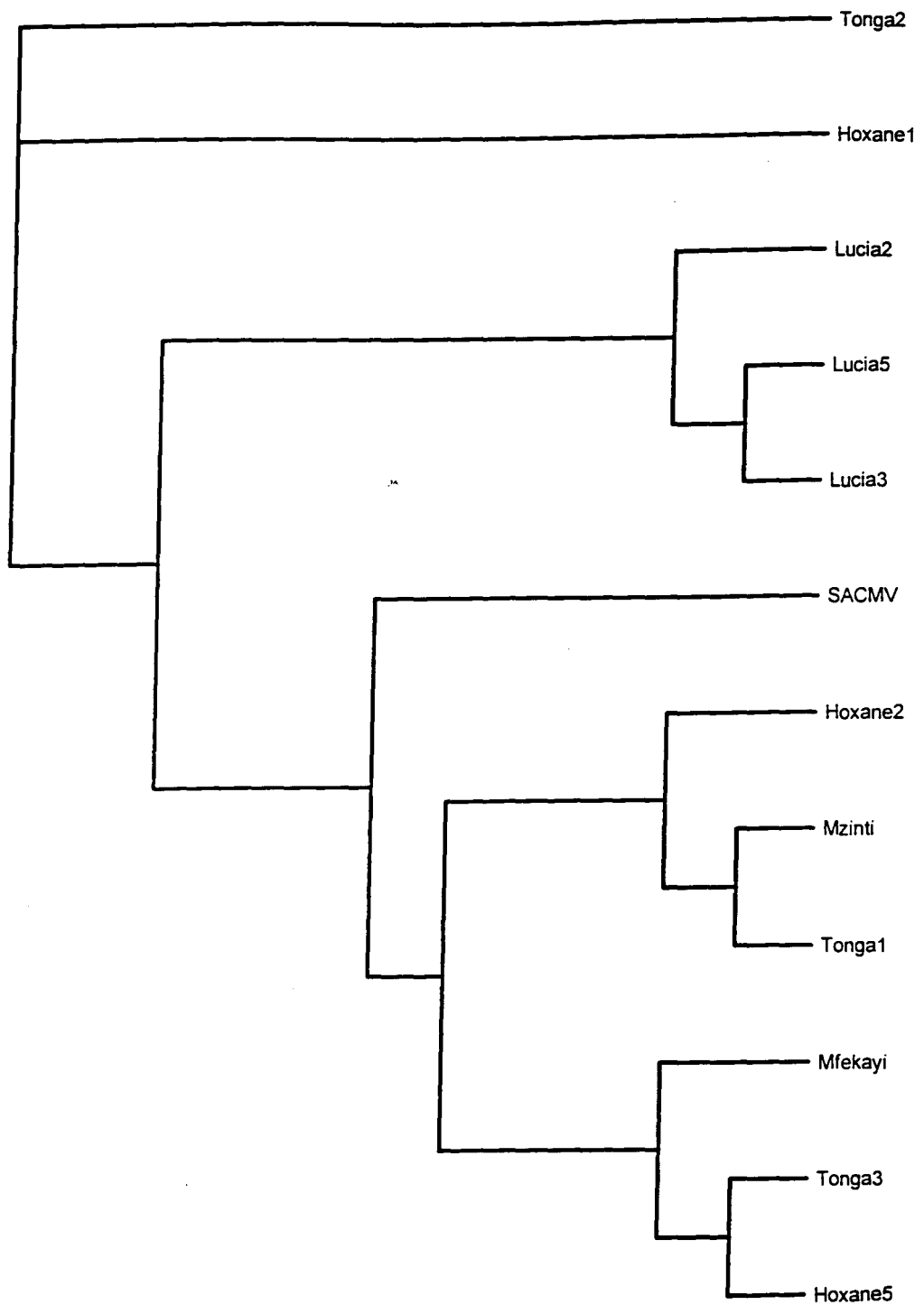
Figure 4.6: Phylogenetic Tree constructed using the distance-based phylogenetic method (FITCH) showing the relationships between cassava-infecting begomovirus isolates compared to SACMV reference strain

having a faster heteroduplex mobility compared to the more distantly related ones.

The degree of conservation in the coat protein gene between begomoviruses that infect cassava ranges from 67% to 100% at the nucleotide level (Brown 1996). Generally the percentage similarities between ACMV and EACMV species in the coat protein are from 80%-85%. Any virus isolate having more than 90% sequence identity to a previously characterised virus genome is proposed to be a strain of an already described virus species and an isolate with less than 90% similarity considered a new strain (Padidam *et al.* 1995). The identity of the eleven virus isolates from South Africa were confirmed in comparisons with other characterised cassava-infecting begomoviruses from Africa (Fig. 4.7). Phylogenetic analysis demonstated that three isolates were closely related to strains of ACMV, six clustered together with EACMV strains from Malawi and the remaining two segregated with other EACMV strains from Kenya, Tanzania and Cameroon. This confirmed the HMA results, where field isolates were differentiated from the three distinct reference viruses, ACMV, EACMV and SACMV.

4.5 DISCUSSION

Recent increases in the incidences and severity of geminiviral infections in certain parts of the World have highlighted the need for continual surveillance and monitoring of the epidemiology of these viruses. One technique available
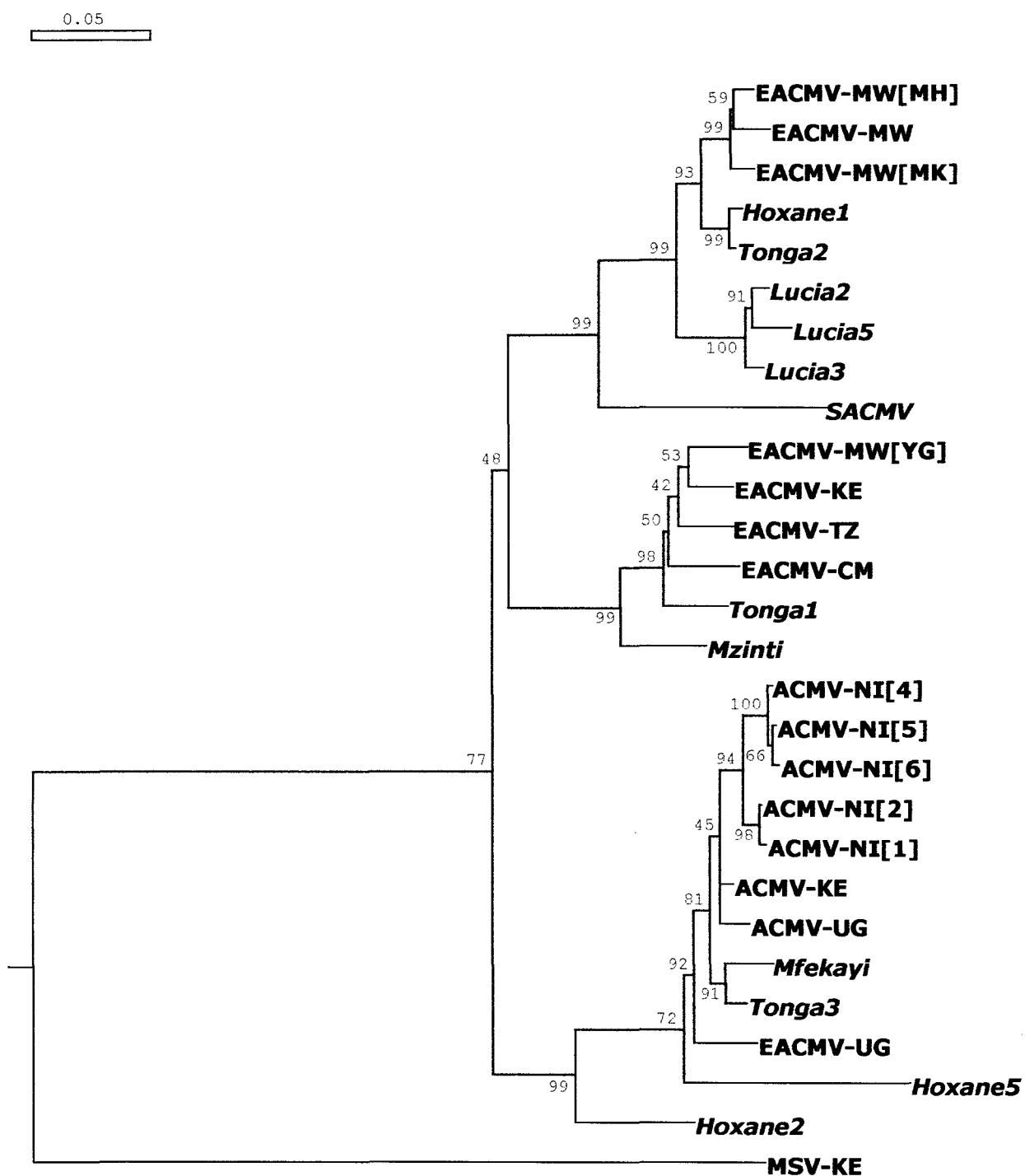
Figure 4.7: Phylogenetic tree demonstrating the nucleotide sequence similarity of the
CCP of South African cassava-infecting begomovirus isolates compared to other
sequenced African cassava mosaic viruses.
Cassava reference viruses: EACMV-Malawi (isolate MH) [AJ006459], EACMV-Malawi
(isolate MK) [AJ006460], SACMV [AF155807], EACMV-Malawi (isolate YG)
[AJ006461], EACMV-Kenya [AJ006458], EACMV-Tanzania [Z83256], EACMV-
Cameroon [AF112354], ACMV-Nigeria (clone pPCR4) [X68320], ACMV-Nigeria
(clone pPCR5) [X68319], ACMV-Nigeria (clone pPCR6) [X68318], ACMV-Nigeria
[X17095], ACMV-Kenya [J02057], ACMV-Uganda [Z83252]. The outgroup used was
Maize Streak Virus-Kenya [X01089].

for differentiating virus species or strains is the use of specific PCR primers designed to non-conserved regions of the genome. While the use of degenerate primers to the DNA-A component for PCR amplification has been widely used for the detection and characterisation of geminivirus species (Briddon *et al.* 1994, Deng *et al.* 1994), these techniques do not necessarily distinguish different strains or mixtures of viruses. Geminiviruses are useful plant viruses with which to test the HMA technique since they have DNA-containing genomes which make them simpler and easier to work with, and they can be considered analogous to HIV in that they also appear to mutate or recombine fairly frequently and numerous 'new' isolates continue to be discovered (Berrie *et al.* 1997, Zhou *et al.* 1997, Fondong *et al.* 2000).

The methodology of HMAs for the purpose of virus identity differentiation requires several parameters to be optimised for individual systems. One of the important parameters is the choice of primer and size of amplified PCR products. From these results, it was demonstrated that the CCP primers and size of PCR products (550 bp) were ideal for HMA analyses. The CCP primers target highly conserved regions of the coat protein gene and thus amplify many different types of geminiviruses as well as producing amplified products within the suggested optimal size range (200-600 bp) (AT Biochem protocol).

The size of the PCR products used for HMAs affect the sensitivity of the method, with sensitivity generally decreasing as the fragments increase in size. Other researchers have used PCR products of 480 bp (poliovirus), 590 bp (measles virus), 760 bp (*Zucchini yellow mosaic virus*) and 1.2 kb (HIV-1) for

their respective HMAs. The effects of fragment size depends on the type and context of sequence differences between the unknown and reference samples and the position of mismatches (Wang *et al.* 1991, Delwart *et al.* 1993, Chezzi *et al.* 1996, Kreis *et al.* 1997, Lin *et al.* 2000).

The optimisation of reaction conditions is important since the conditions for heteroduplex analyses need to be reproducible and consistent. The consistency is necessary so that results obtained from different gels can be correlated, which is essential if phylogenetic inferences are to be made from unknown isolates in relation to reference viruses. In this study it was found that the duration and temperature of annealing appeared to be the most important factors. Annealing is affected by the reaction environment (such as salt concentration and pH) and temperature. Rapid cooling is favourable if there is a significant degree of variation between the type and reference strains whereas slow cooling favours heteroduplex formation between more closely related strains.

For the majority of the cassava virus isolates examined in this study, rapid cooling on ice was preferable. Other factors such as DNA- and salt concentration, although important, were less crucial than that of denaturing and annealing temperatures. Reaction conditions were optimised so that this technique was sensitive enough to detect sequence differences between the eleven isolates of 5.2 % which is at the more sensitive end of the range of 5-25% which has been suggested as the range for good discrimination of viral isolates (Delwart *et al.* 1995). However later studies were able to distinguish differences as low as 2.3% (results not shown).

Other methods such as, SSCP, RNase cleavage, RFLP and denaturing gradient gel electrophoresis (DGGE) have also been used for detecting sequence differences/mutations (Innis *et al.* 1995). RFLPs and RNase cleavage techniques are useful if single base pair changes are being investigated, whereas HMA and SSCP techniques are used for larger DNA size fragment analyses. The sensitivities of HMAs and SSCP are comparable (in the range of 10-25%), however the major limitation to SSCP is that smaller size (<200 bp) fragments are needed for high sensitivity (Maynard *et al.* 1998).

The heteroduplex patterns and profiles obtained by HMAs showed a good correlation with sequencing and phylogenetic results. Thus isolates collected from field samples can be rapidly and accurately assigned to a viral species/strain in relation to known reference viruses, such as SACMV, ACMV or EACMV. Construction of standard curves showing the relationships between genetic distance and heteroduplex mobility for all the possible pair-wise combinations of sequenced virus strains and species, enables quicker identifica-tion of field isolates since unknown test samples can be easily referenced with regard to established identities depicted on the graphs. This has advantages of enabling large-scale screening of field samples and reducing the need and cost of having to clone and sequence all the isolates.

Recent reports (Fondong *et al.* 2000) have shown that cassava exhibiting severe symptoms often have mixtures of ACMV and EACMV and that these mixtures of viruses not only act synergistically to result in more severe mosaic disease but complicate viral diagnosis. Heteroduplex mobility assays could play

a role in diagnosing plants with multiple infections since mixtures of viruses within plants can be demonstrated by the appearance of a number of multiple heteroduplex bands, as seen in this study (Fig. 4.4). Thus using the right combinations of primers and HMAs it is possible to detect mixtures of begomoviruses infecting cassava plants which would then increase the 'recovery' of isolates and improve the accuracy of diagnosis.

In conclusion, we have been able to show that heteroduplex mobility assays are rapid and sensitive in differentiating between cassava-infecting begomoviruses and is useful for preliminary screening of large numbers of field samples. Heteroduplex mobility assays may be useful in detecting sequence diversity in a range of DNA and RNA plant viruses. The only drawback is that the use of a single primer pair may not detect recombination of large DNA fragments in other regions of the genome. It is recommended that as wide a range as possible of reference samples be used enabling quicker, more detailed identification of unknown isolates.

4.6 ACKNOWLEDGEMENTS

well as Leigh Berrie for the SACMV-CP clone. This work would not have been

possible without funding from the British Council, for which we are grateful.

CHAPTER FIVE

PHAGE ANTIBODIES TO SOUTH AFRICAN

CASSAVA MOSAIC VIRUS

5.1 ABSTRACT

Whitefly transmitted geminiviruses are important plant pathogens worldwide. Cassava mosaic disease, which is caused by a group of whitefly-transmitted, cassava-infecting begomoviruses, is a serious problem affecting cassava production in Africa. The epidemiology and spread of this disease has been well studied in several African countries particularly with the use of serological methods. However current serological methods and antibodies are limited in their usefulness and specificity and new antibodies need to be developed to detect all the possible viral species. The objective of this study was to use the technique of differential phage display to develop phage antibodies to South African cassava mosaic virus. From the results, the viability of using phage antibodies to detect geminiviruses proved promising as a number of phage clones were isolated and characterised. These clones, when used in combination, were able to differentiate between several cassava-infecting begomoviruses. However a number of improvements on this technique would need to be implemented before it became an acceptable method for producing antibodies to identify and distinguish between the various cassava-infecting begomovirus species and strains.

5.2 INTRODUCTION

Monoclonal antibodies (MAbs) are antibodies of a single type produced by populations of genetically identical plasma cells (Prescott, Harley and Klein 1993). In 1975, Kohler and Milstein described the production and characterisation of the first somatic cell hybrids capable of indefinite production of antibodies of predetermined specificity. This technology (hybridoma technology) involved fusing B-lymphocyte cells from the spleens of suitably stimulated animals with immortal myeloma cells to produce hybridomas which were capable of producing indefinite amounts of antibodies of predefined specificity.

Monoclonal antibodies have been widely used in plant pathology and antibodies have been produced to more than 60 plant viruses in 20 different virus groups (Torrance 1995). These antibodies have been used for the detection and diagnosis of plant pathogens (viruses, bacteria, fungi and mycoplasma-like organisms) as well as the study of host-pathogen interactions and molecular ecology of viruses and vectors (Torrance 1995).

Whitefly-transmitted geminiviruses (WTGs) are ssDNA viruses that are transmitted by a whitefly (*Bemisia tabaci* Genn) vector and infect a wide range of tropical and sub-tropical plants worldwide (Harrison 1985). In 1986, a panel of 10 monoclonal antibodies was produced to an *African cassava mosaic virus* isolate (ACMV-JI) which could distinguish whitefly-transmitted geminiviruses from leafhopper-transmitted geminiviruses as well as distinguish between ACMV-JI (type strain) and ACMV-C (coastal strain) (Thomas *et al.* 1986). The coat protein gene of WTGs is highly conserved

and all WTGs irrespective of host range or geography share serological relatedness which complicates viral identification and diagnosis.

Currently 17 MAbs have been prepared to ACMV (SCR 11-33) and another 10 MAbs to *Indian cassava mosaic virus* (ICMV) (SCR 52-68) (Aiton and Harrison 1989). Fortunately a number of these antibodies can discriminate between different virus species: SCR 20 is a general MAb that detects a wide range of WTGs, SCR 14 detects both ACMV and *East African cassava mosaic virus* (EACMV) and distinguishes them from other geminiviruses, SCR 23 and 33 when used together can differentiate between ACMV and EACMV and SCR 60 detects ICMV.

This panel of MAbs has been widely used for identifying and characterising WTGs from many different plant hosts and from many different parts of the world such as Burkina Faso (Konate *et al.* 1995), India (Harrison *et al.* 1991, Swanson *et al.* 1992a), the Americas (Swanson *et al.* 1992b) and Europe (Macintosh *et al.* 1992). Serological relationships between geminiviruses were previously examined using immunodiffusion (Sequeira and Harrison 1982) and immunosorbent electron-microscopy (Roberts *et al.* 1984) but it was found that using ELISAs with polyclonal antibodies and unconjugated MAbs gave more sensitive and reliable results (Harrison *et al.* 1991).

Since the advent of hybridoma technology, MAbs to many different antigens and for many different purposes have been created although one important aim i.e. the production of therapeutic human Abs has not been realised. However advances in recombinant DNA technology and increased knowledge about antibody gene sequences has resulted in the production of

human antibodies using phage display as an alternative to animal immunisation (Sheets *et al.* 1998). This technology overcomes many of the difficulties associated with immunising animals, is suitable for antigens with poor immunogenicity and high toxicity and minimises production costs and time (Torrance 1998).

Differential phage display bypasses the immune system and involves the *in vitro* production and tailoring of antibodies. This is achieved by fusing the coding sequences of antibody variable region genes to the amino-terminus of bacteriophage minor coat proteins such that the resulting fusion protein (Ab) is expressed on the outer coat of bacteriophages (Hoogenboom 1997). Phage display libraries have been constructed which have a large repertoire of phages ($>10^8$) displaying different single chain Fv fragment antibodies on their surfaces. These libraries have been used to produce antibodies to a variety of both human (self and non-self) and foreign antigens (Nissim *et al.* 1994, Sheets *et al.* 1998, Wilson and Finlay 1998). Phages displaying antibodies that recognise specific antigenic epitopes are selected by binding to immobilised antigen and recovered and enriched by successive rounds of selection.

Differential phage display libraries have been used to create antibodies to a number of plant viruses such as: *Beet necrotic yellow vein virus* (Uhde *et al.* 2000), *Black currant reversion associated virus* (Susi *et al.* 1998), *Cucumber mosaic cucumovirus* (Ziegler *et al.* 1995), *Cucumber mosaic virus* (Gough *et al.* 1999), *Potato leafroll virus* (Harper *et al.* 1999) and *Tomato spotted wilt virus* (Griep *et al.* 2000). Peptide libraries have even been used for producing

phages that bind to the SCR 20 MAb as a substitute for geminiviruses as a positive control for use in ELISAs (Ziegler *et al.* 1998).

Currently available monoclonal antibodies can only distinguish between 2 cassava-infecting begomovirus species viz. ACMV and EACMV. However two additional viruses, *South African cassava mosaic virus* (SACMV) and Ugandan recombinant virus (UgV) also infect cassava and these viruses are serologically indistinguishable from EACMV and ACMV respectively. Therefore the aim of this work was to use a differential phage display library to produce MAbs to SACMV for use in routine virus screening and diagnosis.

5.3 MATERIALS AND METHODS

5.3.1 Virus purification

Nicotiana benthamiana plants, which were agro-inoculated with full-length SACMV DNA-A and -B dimers cloned in pBIN19 plant transformation vector (kindly supplied by L. Berrie, University of Cape Town) were grown in growth chambers at 25°C with 16 hr:8 hr light:dark photoperiod. After approximately 4 weeks the plants were harvested and virus extractions performed according to the method of Sequeira and Harrison (1982). Viral preparations were tested for viruses by electron microscopy (using formvar, carbon-coated grids and 2% uranyl acetate and PTA pH 3.5 and 7.1 negative staining) and ELISA (using SCR 17 and -20 MAbs, supplied by D. Robinson, SCRI, Dundee).

5.3.2 Phage Display

The Human Synthetic VH and VL scFv Library (Griffin 1 Library, Medical Research Council, Cambridge, England) was used for isolating phages displaying antibodies to SACMV. The Griffin 1 library is a scFv library containing approximately 1×10^{10} clones made from the cloning of heavy and light chain variable regions of synthetic V-gene segments. The library was grown in 500 ml 2xTY media containing 100 mg/ml ampicillin and 1% glucose and infected with helper phage (ratio of 1 bacteria:20 helper phage) (M13-K07 helper phage, Pharmacia). Phages from the library were then isolated, concentrated and resuspended in 5 ml PBS.

Three rounds of selection were performed. For the 1st round of selection a 1/10 dilution of SACMV antigen in 50 mM sodium hydrogen carbonate pH 9.6 was used to coat the immunotube (NUNC Maxisorp Immunotube), for subsequent selections the amount of antigen was reduced to 1/100 (2nd round) and 1/500 (3rd round). Selection involved overnight coating of the immunotubes with antigen, 2hrs of blocking with 2% Marvel-PBS, addition of 1ml of phage to the immunotube, washing the tube with PBS and 0.1% PBS-Tween to remove unbound phage, eluting bound phage with 1 ml 100 mM triethylamine and infection of exponentially growing *E. coli* TG1 which were then plated on 2xTY media and grown overnight.

5.3.3 Screening of potential phage clones

Populations of phages produced at the end of the 2nd and 3rd rounds of selection were then screened for their affinity to SACMV. Individual clones were isolated and grown and monoclonal phage particles rescued. Seventy

two clones from the 2nd round and 98 clones from the 3rd round were initially screened by ELISA for their affinity to SACMV from cassava (positive control) and healthy *N. benthamiana* (negative control).

ELISAs were performed by coating microtitre plates (NUNC Immunosorp 96-well plates) with duplicates of antigen diluted 1/20 in equal volumes of extraction buffer (0.05 M Tris-HCl, 0.005 M EDTA pH 8.0, 2% PVP, 0.05% Tween-20) and carbonate coating buffer pH 9.6. The next day the plates were washed with PBS and 50 ul of phage added to each well, afterwhich anti-M13 Mab (1:2000) was added, followed by the addition of alkaline phosphatase-conjugated anti-mouse MAb (1:5000) and detection with 1 mg/ml p-nitrophenyl phosphate and 10% diethanolamine. The plates were then read on a Titertek Multiscan Plate Reader at a wavelength of 405 nm.

Those clones which exhibited high absorbance values (> cut-off value, cut-off value=mean of controls+2 S.D.) for SACMV and low absorbance values (< cut-off value) for healthy *N. benthamiana* were chosen for further screening. Twenty four clones were selected and their specificity for SACMV was examined by testing them against 4 other begomoviruses that infect cassava: an isolate from Swaziland (shown by sequence analysis to be most closely related to SACMV), an isolate from Mzinti (belonging to EACMV1 group), an isolate from St. Lucia (EACMV2) and an isolate from Mfekayi (ACMV).

These same clones were then also used to infect *E. coli* HB2151 which resulted in the production of soluble antibody fragments and were again tested by ELISA to the 5 different viruses. For the final screening procedure the 8 clones that gave the most promising results were used. These clones

were tested against other related cassava begomoviruses: ACMV-Kenya, EACMV-Kenya, ACMV-Nigeria, ACMV+UgV mix and UgV (kindly provided by S. Winter, DSMZ, Plant Virus Division, Braumschweig, Germany).

5.3.4 Optimisation of ELISA conditions

Several experimental conditions affecting ELISAs using phage-bound antibodies and ELISAs using soluble antibodies were optimised (Table 5.1).

5.4 RESULTS

Agro-inoculation of approximately 100 *N. benthamiana* plants resulted in a large percentage (>60%) of the plants becoming infected and displaying characteristic symptoms of leaf curling and mosaic patterning on the leaves. Using *N. benthamiana* for propagating cassava begomoviruses is advantageous since it has been found that virus concentration in cassava reaches only 1-7% of that in *N. benthamiana* (Sequeira and Harrison 1982) and extracting viruses from cassava has been proved to be extremely difficult. Electron microscopy of the viral preparations using uranyl acetate revealed geminivirus-like particles of the correct size, although the particles did not appear entirely intact, probably as a result of the omission of the sucrose gradient step in the virus purification protocol which might have resulted in the viral particles being broken up due to the high speed centrifugation. Viral particle confirmation was achieved serologically by ELISA using SCR 17 and SCR 20 MAbs (results not shown).

Table 5.1 Parameters investigated in attempts to optimise ELISA conditions

ELISAs using phage-bound antibodies	ELISAs using soluble antibodies
Antigen: 1/10, 1/20, 1/50	Antigen: 1/10, 1/20, 1/50
anti-M13 1° Mab: 1:500, 1:1000, 1:2000	anti-myc 9E10 MAb: 1:100, 1:500, 1:750
	Alk. Phosph. 2°MAb: 1:1000, 1:2000, 1:5000

Buffer 1: 0.05 M Tris-HCl, 0.005 M EDTA pH 8.0, 2% PVP, 0.05% Tween-20   and
Buffer 2: 0.05 M Tris-HCl, 0.06 M sodium sulphite pH 8.25 + 0.1 % (v/v) 2-mercaptoethanol.

Carbonating coating buffer   and
no carbonate coating buffer

Polyclonal antibody   and
no polyclonal antibody

no. of washes (2 , 4 and 6)

Temperature (25°C and 37°C)

Once the preparations were found to be viral they were used for producing antibodies by differential phage display.

After performing the 3 rounds of selection and enrichment, individual phage clones were screened for their affinity to SACMV. Of the 170 individual clones tested, 24 clones that gave the best results i.e. displayed high absorbance values when tested against SACMV from cassava and low absorbance values when testing healthy *N. benthamiana*, were used for further screening. These clones were then tested for their specificity towards SACMV and 4 other cassava begomoviruses from southern Africa. The results (Figures 5.1 and 5.2) showed that the 24 clones, although all isolated from a library primed specifically towards SACMV, all had different specificities and affinities towards SACMV.
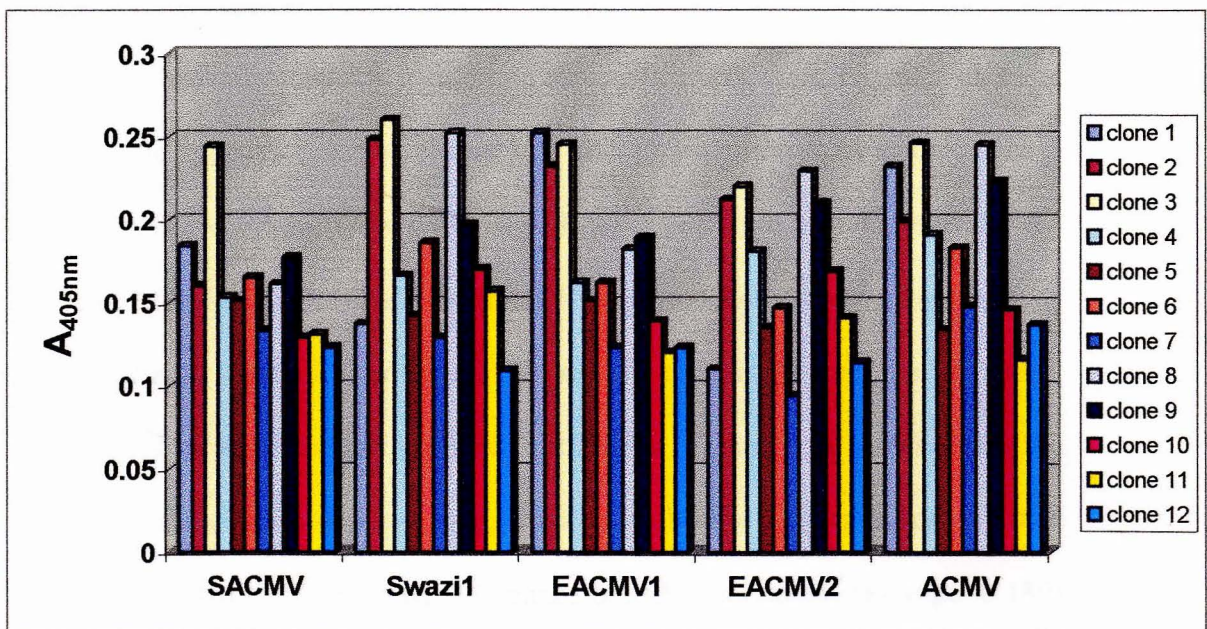

Figure 5.1 Chart showing the A_{405} values of phage clones 1-12 when tested against SACMV and 4 other cassava begomoviruses from southern Africa
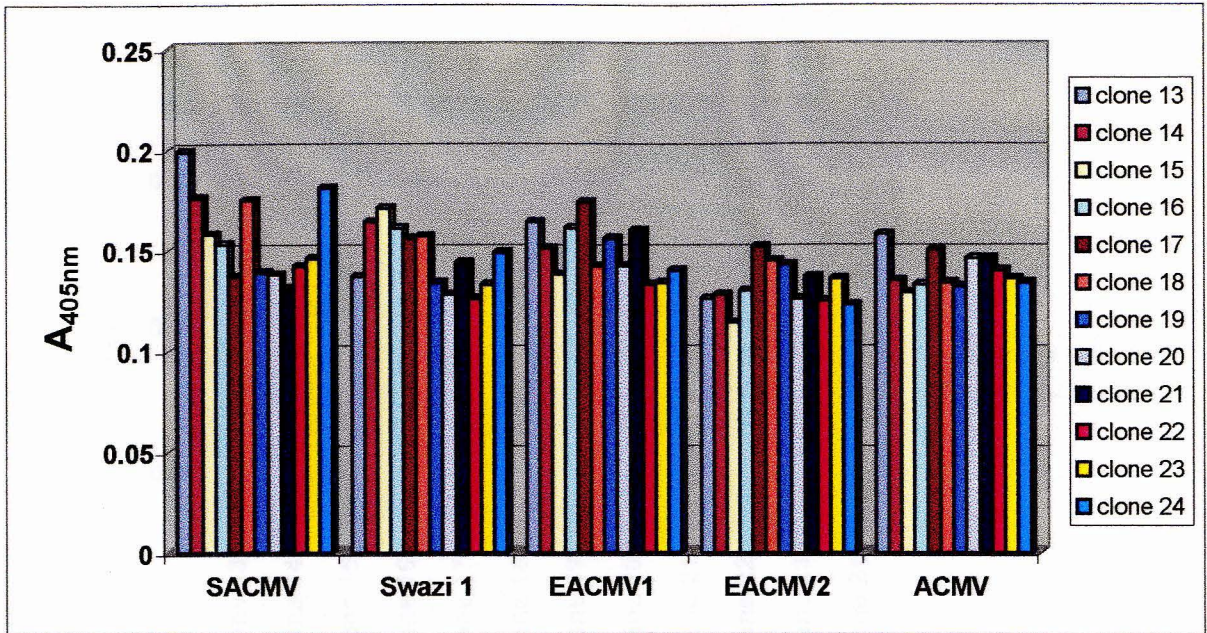
Figure 5.2 Chart showing the A_{405} values of phage clones 13-24 when tested against SACMV and 4 other cassava begomoviruses from southern Africa

The different antibodies recognised different epitopes on the viruses, some recognised epitopes general to all the viruses (such as clones 2, 3 and 8 which had affinities for 3 or more viruses), others were targeted to more specific epitopes (such as clones 1, 9 and 17 which recognised only 1 or 2 viruses) and others (clones 13, 18 and 24) appeared to be more specific for SACMV. Soluble antibody fragments were produced by infecting the non-suppressor strain of *E. coli* HB2151 with the phagemids from the 24 clones. Results of ELISAs using soluble antibodies (Figure 5.3) showed that the absorbance values obtained using soluble antibodies were in all cases lower (average=0.120) compared to phage-bound antibodies (average=0.189) although the variances were not that significant (p=0.020 using Fishers Exact Test).
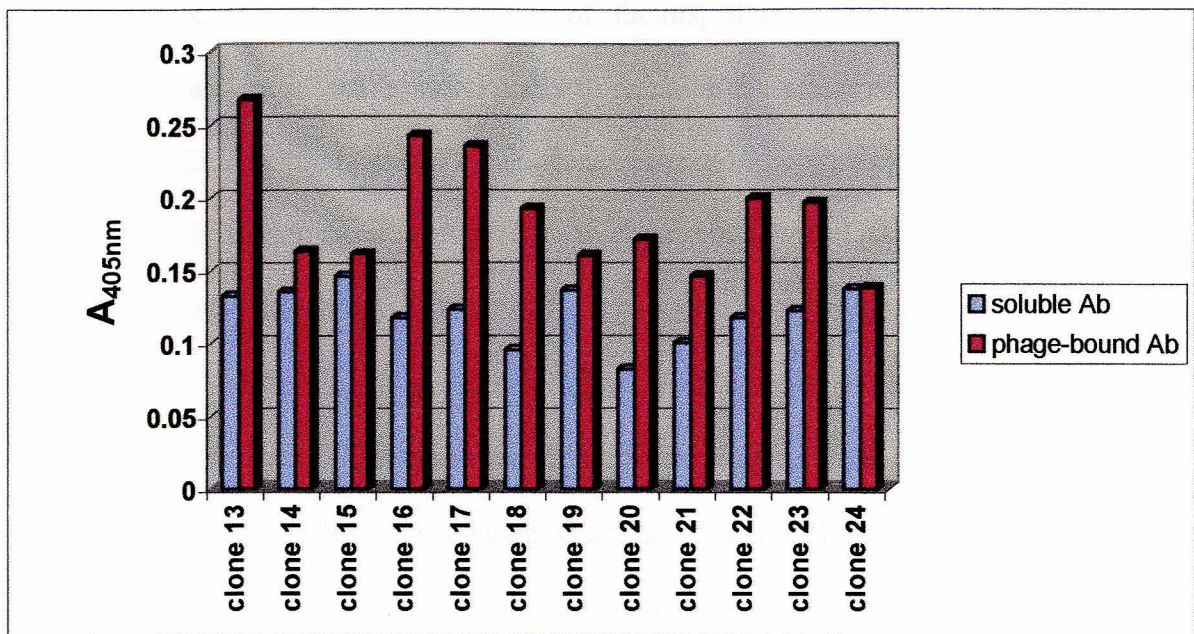
Figure 5.3: Chart showing the difference in A_{405} values of selected soluble and phage bound antibody clones when tested against SACMV

To verify that soluble antibodies were being produced, SDS-PAGE followed by Coommassie Blue and Silver staining was performed. These results (data not shown) suggested that soluble antibodies were indeed being induced and expressed but in lesser amounts compared to phage-bound antibodies. Further improvement of the results was attempted by optimising the soluble-antibody ELISA conditions. It was found that varying the concentrations of SACMV antigen (1/10, 1/20 and 1/50) and anti-myc 9E10 antibody (1:100, 1:500 and 1:750) had little effect on the results. Increasing the amount of alkaline-phosphatase labelled 2nd antibody from 1:5000 to 1:1000 gave false positive results and 1:5000 was adopted for further tests. Since the results obtained using the soluble antibodies were not significantly better than that using phage-bound Abs it was decided to continue the rest of the screenings using phage-bound antibodies.

The relatively weak reactions between the antibodies and SACMV antigen could be attributed to a number of factors, thus it was necessary to investigate these factors individually to examine the effect of each of these on the ELISA conditions. The choice of extraction buffer is very important since it has been shown that tobacco leaf curl virus from *Lonicera japonica* could only be detected when reducing agents were added to the leaf extraction media (Macintosh *et al.* 1992) and clarification of cassava extracts by the addition of chloroform decreased the inhibitory effects of cassava sap and increased the absorbance values (Fargette *et al.* 1987).

Comparison of the values obtained using the 2 different extraction buffers (see Table 5.1) (results not shown) indicated that using buffer 1, as opposed to buffer 2, gave better results, suggesting that the low readings obtained were not as a result of viral particles being degraded or excessive effects of inhibitory compounds.

The addition of carbonate coating buffer to the extraction buffer also seemed to have little effect on antigen adsorption as absorbance values were not noticeably affected. One factor which did appear to play a role was the use of polyclonal antibodies for precoating the microtitre plates. It was found that using polyclonal antibodies to coat the plates before the addition of antigen resulted in the A_{405} values obtained for SACMV with phage clone 1 increasing from 0.342 (without polyclonal antibody) to 0.443 (with polyclonal antibody).

It was also hypothesised that the use of 6 washes between each step was too excessive as it could result in the breaking of the already weak binding between phage Abs and antigen. Decreasing the number of washes to 4 and 2

washes did not seem to substantially affect the results thus 6 washes was adopted to maintain the stringency and prevent false positive results. Increasing the incubation temperature from the recommended 25°C to 37°C gave increased variability in the results between duplicate wells. Varying the amounts of antigen (1/10, 1/20 and 1/50) and anti-M13 antibody (1/500, 1/1000 and 1/2000) made little difference, and better results were obtained using antigen diluted 1:20 and anti-M13 at 1:1000.

The final screening procedure was performed using the optimised ELISA conditions and the 8 clones (clones 1, 3, 5, 8, 9, 13, 14, and 24) that gave the best results against the 5 cassava begomovirus isolates (Figure 5.1 and 5.2).
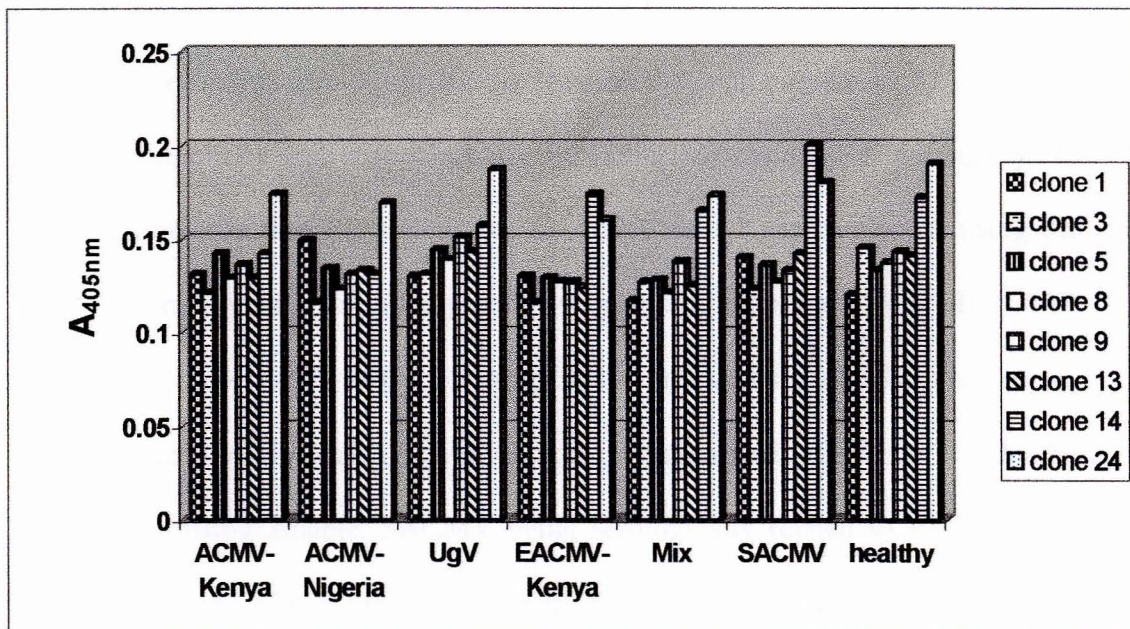
Figure 5.4: Chart displaying the specificities of the 8 'best' clones against SACMV and other African cassava begomoviruses from eastern and western Africa

These results (Figure 5.4) showed that clones 1, 3, 5, 8, 9, 13 and 24 exhibited no apparent increased specificity towards SACMV compared to the other 5 viruses and hence were not as discriminatory as hoped for. Clone 14 showed the most promising results with increased specificity towards

SACMV (A_{405}= 0.200) as opposed to EACMV-Kenya (A_{405}= 0.174), ACMV+UgV mix (A_{405}= 0.165) and healthy cassava (A_{405}= 0.171).

5.5 DISCUSSION

Several methods, other than serology, exist for identifying and characterising cassava-infecting begomoviruses, including PCR using degenerate primers that anneal to highly conserved regions of the genome (Rojas *et al.* 1993, Deng *et al.* 1994, Wyatt and Brown 1996), nucleic acid hybridisation (Roberts *et al.* 1984) and cloning and sequencing of full-length genomes (Hong *et al.* 1993, Zhou *et al.* 1998).

Nevertheless serological testing remains the most popular in Africa due to it's simplicity and cost-effectiveness. Hence the majority of the surveys of cassava-infecting begomoviruses in Africa have been conducted using the SCR monoclonal antibody panel (Swanson and Harrison 1994, Offei *et al.* 1999, Ogbe *et al.* 1999). However these antibodies can only discriminate between ACMV and EACMV and other viruses such as SACMV and UgV are indistinguishable, SACMV being serologically related to EACMV and UgV to ACMV. Since no other monoclonal antibodies are available for testing it would be beneficial if additional antibodies could be made that could distinguish SACMV and UgV.

The easiest, quickest and cheapest way to do this would be to use differential phage display technology to produce scFv antibodies which could then be used in field testings and routine screenings. However producing MAbs to geminiviruses has proven difficult due to the poor

immunogenicity of these viruses, low levels of virus in cassava, the tendency of viral particles to aggregate and/or attach to plant material which significantly reduces the number of viral particles isolated (Sequeira and Harrison 1982) and interference by components of cassava leaf extracts particularly in the older leaves (Sequeira and Harrison 1982, Fargette *et al.* 1987).

From the results of this study we were able to isolate 4 clones (clones 13, 14, 18 and 24) which showed some limited degree of affinity and specificity towards SACMV compared to healthy controls and other related cassava begomoviruses. Three clones (13, 18 and 24) could distinguish SACMV from other southern African virus isolates whereas clone 14 differentiated between SACMV and other east African virus isolates.

Factors which may have contributed to the limited ability to produce MAbs with high specificity and affinity to SACMV could be: plant material used for testing and screening not being sufficiently fresh (S. Winter, pers. comm.); affinities (i.e. binding) of MAbs being too low hence their loss after washing steps; MAbs not being specific enough to discriminate between different cassava begomoviruses due to coat protein conservation and small size and monovalency of scFv antibodies compared to 'normal' antibodies resulting in weaker signalling.

Ways to improve on this technique in future could be: clarifying cassava sap to remove the effects of inhibitors (Fargette *et al.* 1987); increasing the number of rounds of selections from 3 to 4 or 5; increasing the time of elution of phages during the selection process thereby selecting for the very high affinity binders which might otherwise have been lost (de Bruin *et al.*

1999); multimerizing the soluble scFv fragment antibodies to enhance avidity and slow the rate of dissociation (Nissim *et al.* 1994); improve the affinity of the antibodies by shuffling of the heavy and light chain genes (Marks *et al.* 1992) or by manipulating antigen binding sites of antibodies by site-directed mutagenesis or error-prone PCR (Hoogenboom 1997); re-engineering scFv proteins by introducing peptide linkers or incorporating into fusion proteins (Whitlow *et al.* 1991). The probability of actually isolating a clone is multifactorial and depends on the size of the library, it's diversity and the threshold binding affinity (Hoogenboom *et al.* 1992) as well as the number and efficiency of rounds of selection (Marks *et al.* 1991).

In conclusion, using 4 clones in combination, it might be possible to tentatively identify and characterise cassava virus isolates. Although phage display showed potential in identifying cassava-infecting begomoviruses, the technique still needs improvement and refinement before it can be adopted as a reliable screening method.

5.6 ACKNOWLEDGEMENTS

CHAPTER SIX

SUMMARY

Three major objectives of this study were:

1. To study the genetic diversity of cassava-infecting begomoviruses in selected countries within southern Africa.

2. Examine the genetic diversity of *Bemisia tabaci* Gennadius populations in southern Africa.

3. Develop techniques for screening and differentiating cassava begomoviruses from field samples.

6.1 Genetic Diversity of Cassava-Infecting Begomoviruses

Investigations into the causative agent/s responsible for cassava mosaic disease in Africa have revealed three distinct virus species, viz. *African cassava mosaic virus* (ACMV) (Hong *et al.* 1993), *East African cassava mosaic virus* (EACMV) (Zhou *et al.* 1998) and *South African cassava mosaic virus* (SACMV) (Berrie *et al.* 1998). Furthermore, several different strains have been reported, for example, the Ugandan variant (UgV) (Zhou *et al.* 1997) which is a recombinant virus of ACMV and EACMV.

The results of our study through examination of the genetic diversity of cassava-infecting begomoviruses from 6 southern African countries (Angola, Mozambique, South Africa, Swaziland, Zambia and Zimbabwe), revealed that in countries such as Zimbabwe where previous serological evidence suggested only EACMV, there is now evidence of ACMV as well; in Angola, Mozambique and South Africa where previously only ACMV were reported, there now exists EACMV.

In addition, EACMV viruses from southern Africa could be further separated into 2 distinct groups (EACMV1 and EACMV2). The EACMV2 viruses are believed to be 'newer' species that arose from recombination with the closely related *Tomato yellow leaf curl viruses* (Zhou *et al.* 1998). Our results also demonstrate, for the first time, the appearance of SACMV in Swaziland. SACMV, which is a recently discovered recombinant virus, had thus far only been found to occur once in South Africa (Berrie *et al.* 1998) but phylogenetic analysis has now identified a virus isolate (CBV-SZ[Ma1]) highly related (>97%) to SACMV. The exact extent and nature of possible recombination within this virus isolate from Swaziland warrants further investigation and may shed light on the origin/s of SACMV.

Examination of the nucleotide identities of cassava begomoviruses from the different geographical regions showed that significant movement of infected cassava material must have occurred. This is exemplified by isolates from one country (e.g. CBV-ZA[T1], South Africa) showing high degrees of relatedness to isolates from both neighbouring (CBV-MZ[Bez2], Mozambique) and distant (CBV-AO, Angola) countries. Thus the differentiation and separation of cassava-infecting begomoviruses from southern Africa is complex and not geographically limited, most probably as a result of the extensive trafficking of cassava stakes between and within neighbouring countries.

The results of our study also show that the recombinant virus, UgV, is present in southern Africa and occurs in a mixed infection in a substantial number of infected plants. These results are rather surprising as UgV was thought to be mainly confined to Uganda (Zhou *et al.* 1997) and parts of

Guinae (Winter 1998) and Sudan (Harrison *et al.* 1997). It was disturbing to find that UgV-like viruses were present further south, and this has huge implications for disease diagnosis and control.

This suggests that 'hot-spots' of recombination are most probably occurring in all areas where cassava begomoviruses are found. Recombination is now recognised to occur between geminiviruses and contribute to geminivirus evolution and diversity (Rybicki *et al.* 1994, Deng *et al.* 1997, Zhou *et al.* 1997).

The opportunities exist within southern Africa for epidemics similar to that in Uganda to occur, particularly if the continued movement of cassava stakes occurs, if the spread of the more destructive, polyphagous B-biotype continues and eventually reaches southern Africa, and if extremely susceptible cassava varieties becomes widely adopted and planted. Continued vigilance of the incidences and severity of cassava mosaic outbreaks, increased and broader epidemiological surveys as well as the continued development of increasingly resistant cassava cultivars (either through plant breeding or genetic engineering) could prevent similar catastrophes occurring in southern Africa.

6.2 Genetic Diversity of *Bemisia tabaci* Gennadius

Cassava mosaic disease is spread by the planting of infected stem cuttings as well as by transmission by a whitefly vector (*Bemisia tabaci* Gennadius) (Homoptera: Aleyrodidae). There are more than 1100 species of *Bemisia* (whitefly) yet only one species, *B. tabaci*, transmits begomoviruses that infect cassava.

One of the aims of our work was to examine the genetic diversity of populations of *Bemisia tabaci* that exist on cassava in southern Africa. Phylogenetic analyses of the mitochondrial cytochrome oxidase I (mtCOI) gene molecular marker revealed that whitefly populations affecting cassava in southern Africa formed two distinct groups: those from Mozambique, South Africa, Swaziland and Zambia formed a closely-related (95-99%), distinct clade, whereas those collected from Zimbabwe formed a separate group and were more closely related to whitefly infecting okra from Cameroon and the Ivory Coast. Whitefly from cassava in Cameroon segregated as a separate cluster of closely related isolates.

These results suggest that geographical genotypes (topotypes) of *B. tabaci* exist on cassava in Africa. This geographical separation of cassava-colonising *B. tabaci* correlates directly with that observed when comparing cassava begomovirus coat protein sequences (Hong and Harrison 1995).

This is a significant finding but not unexpected since experimental evidence has demonstrated both direct and indirect evidence that the coat protein region plays a major role in influencing viral-vector transmission (Briddon *et al.* 1990, McGrath and Harrison 1995). Future research is essential to investigate transmission efficiencies of different virus species in conjunction with different *B. tabaci* species and correlate the effect of virus genetic diversity with that of whitefly population diversity.

6.3 Development of New Identification Techniques

There exists numerous methods and techniques for the diagnosis, identification and characterisation of geminiviruses: serological analysis

(Harrison *et al.* 1991, Swanson *et al.* 1992); immunosorbent electron microscopy (Sequeira and Harrison 1982); nucleic-acid hybridisation using cloned probes to DNA-A and –B (Roberts *et al.* 1984) and PCR (Rojas *et al.* 1993, Deng *et al.* 1994, Wyatt and Brown 1996, Brown *et al.* 2001).

An objective of this project was to develop more sensitive and specific methods for differentiating between field isolates of cassava-infecting begomoviruses. One such method was the Heteroduplex Mobility Assay (HMA) which was successful in identifying and differentiating cassava begomoviruses without the need to clone and sequence viral gene products. The second method which showed promising results, was the development of monoclonal antibodies using differential phage display technology.

In this study we were able to develop the HMA technique for distinguishing cassava-infecting begomoviruses. The HMA profiles were able to differentiate four different viral species and eleven different virus strains and show a good correlation with sequencing results and phylogenetic comparisons with other sequenced cassava viruses. This technique is sensitive (detecting sequence differences as little as 2.3%) and rapid and has the added advantage of being able to detect mixtures of viruses in field-grown cassava.

It is envisaged that this technique could become widely adopted as the method-of-choice for large-scale screening of field material and complement the results obtained using other more traditional methods such as ELISAs with the SCR (Scottish Crop Research) monoclonal antibody panel or newly developed phage antibodies.

The highly conserved nature of the coat protein of WTGs allows not only the serological identification of strains of viruses but also enables antisera produced to one geminivirus to detect other members of a group.

The current SCR panel of monoclonal antibodies can distinguish cassava-infecting begomoviruses from begomoviruses from other hosts as well as differentiate between ACMV and EACMV. However using these antibodies it is not possible to discriminate between UgV and SACMV as UgV and SACMV are serologically related to ACMV and EACMV respectively (Thomas *et al.* 1986, Zhou *et al.* 1997, Berrie *et al.* 1998).

Using differential phage display as a possible solution, monoclonal antibodies to SACMV were produced relatively quickly (3-4 weeks) and with relative ease. From our experiments we were able to isolate 4 phage clones which showed a degree of specificity towards SACMV. Using these 4 clones in combination it might be possible to tentatively differentiate SACMV in field cassava. However a number of parameters need to be further investigated and improved upon before this technique can realise it's full potential and become widely-used for the production of monoclonal antibodies on demand.

In conclusion, the diversity exhibited by both the cassava-infecting begomoviruses and their whitefly vectors, will have significant impacts in both the control and spread of disease, affecting the success of developing cassava plants resistant to infection. It is hoped that these newly-developed diagnostic methods will be adopted and assist in studying the epidemiology

of this disease in an effort to bring it under control and ensure cassava's

future as a secure, stable food source for Africa.

CHAPTER SEVEN

REFERENCES

Adipala, E., Byabakama, B.A., Ogenga-Latigo, M.W. and Otim-Nape, G.W. (1998) Effect of planting date and varietal resistance on the development of cassava mosaic virus disease in Uganda. *African Plant Protection* 4: 71-79.

Aiton M.M. and Harrison B.D. (1989) Monoclonal antibodies to Indian cassava mosaic geminivirus (ICMV). *Report of the Scottish Crop Research Institute for 1988* pp.175.

Alleman J., Coertze A., Viljoen J. (1996) Cassava-a crop of many uses. *Farmers Weekly* October 4, pp. 16-18.

Arguello-Astorga, G., Herrera-Estrella, L. and Rivera-Bustamante, R. (1994a) Experimental and theoretical definition of geminivirus origin of replication. *Plant Molecular Biology* 26: 553-556.

Arguello-Astorga, G.R., Guevara-Gonzalez, R.G., Herrera-Estrella, L.R. and Rivera Bustamante, R.F. (1994b) Geminivirus Replication Origins Have a Group-Specific Organization of Iterative Elements: A Model for Replication. *Virology* 203: 90-100.

AT Biochem., Short Protocol: Heteroduplex Analysis using MDE Gel. Appendiz V, pp. 42-45.

Bedford, I.D., Briddon, R.W., Markham, P.G., Brown, J.K., Rosell, R.C. (1992) *Bemisia tabaci*- Biotype Characterisation and the Threat of the Whitefly species to agriculture, *Brighton Crop Protection Conference-Pests and Diseases*, pp. 1235-1240.

Bedford, I.D., Briddon, R.W., Markham, P.G., Brown, J.K., Rosell, R.C. (1992) A new species of *Bemisia* or biotype of *Bemisia tabaci* (Genn.) as a future pest of European agriculture. Proc. Plant Health and the European Single Market. *BCPC Monograph* 54: 381-386.

Bedford, I.D., Briddon, R.W., Brown, J.K., Rosell, R.C., Markham, P.G. (1994) Geminivirus transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. *Annals of Applied Biology* 125: 311-325.

Berrie, L.C., Palmer, K.E., Rybicki, E.P. and Rey, M.E.C. (1998) Molecular characterisation of a distinct South African cassava infecting geminivirus. *Archives of Virology* 143: 2253-2260.

Berry, S. and Rey, M.E.C. (2001) Differentiation of cassava-infecting begomoviruses using Heteroduplex Mobility Assays. *Journal of Virological Methods* (in press).

Bird, J. and Maramorosch, K. (1978) Viruses and virus diseases associated with whiteflies. *Advances in Virus Research* 22: 55-110.

Bock, K.R., Guthrie, E.J., Meridith, G. (1978) Distribution, host range and purification of cassava latent virus, a geminivirus. *Annals of Applied Biology* 90: 361-367.

Bock, K.R. and Woods, R.D. (1983) Etiology of African Cassava Mosaic Disease. *Plant Disease* 67: 994-995.

Briddon, R.W., Pinner, M.S., Stanley, J. and Markham, P.G. (1990) Geminivirus Coat Protein Gene Replacement Alters Insect Specificity. *Virology* 177: 85-94.

Briddon, R.W. and Markham, P.G. (1994) Universal Primers for the PCR Amplification of Dicot-Infecting Geminiviruses. *Molecular Biotechnology* 1: 202-205.

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. (1994) The status of *Bemisia tabaci* Genn. as a pest and vector in world agroecosystems. *FAO Plant Protection Bulletin* 42: 3-32.

Brown, J.K., Frohlich, D.R., Rosell, R.C. (1995a) The sweetpotato or silverleaf whiteflies: Biotypes of *Bemisia tabaci* or a species complex? *Annual Review of Entomology* 40: 511-534.

Brown, J.K., Coats, S., Bedford, I.D., Markham, P.G., Bird, J., Frohlich, D.R. (1995b) Characterisation and distribution of esterase electromorphs in the whitefly, *Bemisia tabaci* (Genn.) (Homoptera: Aleyrodidae). *Biochemical Genetics* 33: 205-214.

Brown, J.K. (1996) The Biology and Molecular Epidemiology of the Geminiviridae Subgroup III. In: Plant-Microbe Interaction. ed. Stacey, G. and Keen, N. pp. 125-195.

Brown, J. K., J. Bird, D. R. Frohlich, R. C. Rosell, I. D. Bedford, and P. G. Markham. (1996). The relevance of variability within the *Bemisia tabaci* species complex to epidemics caused by Subgroup III geminiviruses. Pages 77-92 in:*Bemisia '95: Taxonomy, Biology, Damage, Control and Management* (D. Gerling and R. T. Mayer, eds). Intercept Publications, Wimborne, UK.

Brown, J.K., Torres-Jerez, I., Caballero, R. and Brentt, D. (1998) First report of *Wolbachia* spp. in the whitefly, *Bemisia tabaci*, and unidirectional production of female offspring due to apparent cytoplasmic incompatibility. Third

International Symposium on Molecular Insect Science, Snowbird, Utah, June 5-10, 1998.

Brown, J.K. (2001) Proceedings of the European Whitefly Symposium, Ragusa, Sicily, Italy, 27 February-3 March 2001, pp. 92.

Brown, J.K., Idris, A.M., Torres-Jerez, I., Banks, G.K., Wyatt, S.D. (2001) The core region of the coat protein gene is highly useful for establishing the provisional identification and classification of begomoviruses. *Archives of Virology* (in press).

Burban, C., Fishpool, L.D.C., Fauquet, C., Fargette, D., 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, F.J., Cahill, M., Denholm, I. and Devonshire, A.L. (1995) Biochemical identification of interbreeding between B-type and non B-type strains of the tobacco whitefly *Bemisia tabaci*. *Biochemical Genetics* 33: 13-23.

Cai, S.P., Eng, B., Kan, Y.W., Chui, D.H.K. (1991) A rapid and simple electrophoretic method for the detection of mutations involving small insertion and deletion: Applications to β-thalassemia. *Human Genetics* 87: 728-720.

Campbell, B.C., Steffen-Campbell, J.D. and Gill, R.J. (1996) Origin and radiation of whiteflies: an initial molecular phylogenetic assessment. Pages 29-51 In: *Bemisia*: 1995 Taxonomy, Biology, Damage, Control, and Management. D. Gerling and R.T. Mayer (Eds.), Intercept Ltd., Andover, Hants, UK. pp. 702.

Chezzi, C. and Schoub, B.D. (1996) Differentiation between vaccine-related and wild-type polioviruses using a heteroduplex mobility assay. *Journal of Virological Methods* 62: 93-102.

Christiaensen, L., Tollens, E. and Ezedinma, C. (1995) Development Patterns under Population Pressure: Agricultural Development and the Cassava-Livestock Interaction in Smallholder Farming Sytems in Sub-Saharan Africa. *Agricultural Systems* 48: 51-72.

Coats, S.A., Brown, J.K., Hendrix, D.L. (1994) Biochemical characterization of biotype-specific esterases in the whitefly *Bemisa tabaci* Genn. (Homoptera: Aleyrodidae). *Insect Biochemistry and Molecular Biology* 24: 723-728.

Cock, M.J.W. (ed.) (1986) *Bemisia tabaci*-a literature survey on the cotton whitefly with an annotated bibliography *FAO/CAB,* pp.121

Colvin, J., Otim-Nape, G.W., Holt, J., Omongo, C., Seal, S., Stevenson, P., Gibson, G., Cooter, R.J., Thresh, J.M. (1999) Factors driving the current epidemic of severe cassava mosaic disease in East Africa. In: Proc. VIIth International Plant Virus Epidemiology Symposium; Plant Virus Epidemiology: Current Status and Future Prospects April 11-16, 1999, Aguadulce (Almeria), Spain, pp 76. Plant Virus Epidemiology Committee of the International Society of Plant Pathology, pp 193.

Costa, A.S. and Russell, L.M. (1975) Failure of *Bemisia tabaci* to breed on cassava plants in Brazil (Homoptera: Aleyrodidae). *Cien. e Cult.* 27: 388-390.

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. *Entomological experimental applications* 61: 211-219.

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

Czosnek, H., Ber, R., Antigus, Y., Cohen, S., Navot, N., Zamir, D. (1988) Isolation of tomato yellow leaf curl virus, a geminivirus. *Phytopathology* 78: 508-512.

Dahniya, M.T. (1994) An Overview of Cassava in Africa. *African Crop Science Journal* 2: 337-343.

Daphne, P. (1980) Cassava, a South African venture. *Optima* 1: 61-68.

De Barro, P.J. and Hart, P.J. (2000) Mating interactions between two biotypes of the whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Australia. *Bulletin of Entomological Research* 90: 103-112.

De Bruin, R., Spelt, K., Mol, J., Koes, R. and Quattrocchio, F. (1999) Selection of high-affinity phage antibodies from phage display libraries. *Nature Biotechnology* 17: 397-399.

Delwart, E.L., Shaper, E.G., Louwagie, J., McCutchan, F.E., Grez, M., Rubsamen-Waigmann, H., Mullins, J.I. (1993) Genetic Relationships Determined by a DNA Heteroduplex Mobility Assay: Analysis of HIV-1 env Genes. *Science* 262: 1257-1261.

Delwart, E.L., Herring, B., Rodrigo, A.G., Mullins, J.I. (1995) Genetic Subtyping of Human Immunodeficiency Virus Using A Heteroduplex Mobility Assay. Manual Supplement in: PCR Methods and Applications, Cold Spring Harbor Laboratory, NY, pp. 2-15.

Deng, D., McGrath, P.F., Robinson, D.J., Harrison, B.D. (1994) Detection and differentiation of whitefly-transmitted geminiviruses in plants and vector insects by the polymerase chain reaction with degenerate primers. *Annals of Applied Biology* 125: 327-336.

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

Doyle, J.J. and Doyle, J.L. (1987) A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11-15.

Enete, A.A., Nweke, F.I. and Okorji, E.C. (1995) Trends in Food Crop Yields under Demographic Pressure in Sub-Saharan Africa: the Case of Cassava in Southeast Nigeria. *Outlook on Agriculture* 24: 249-254.

Etessami, P., Saunders, K. Watts, J. and Stanley, J. (1991) Mutational analysis of complementary-sense genes of African cassava mosaic virus DNA A. *Journal of General Virology* 72: 1005-1012.

Fargette, D., Thouvenel, J.C. and Fauquet, C. (1987) Virus content of leaves of cassava infected by African cassava mosaic virus. *Annals of Applied Biology* 110: 65-73.

Fargette, D., Jeger, M., Fauquet, C. and Fishpool, L.D.C. (1993) Analysis of Temporal Disease Progress of African Cassava Mosaic Virus. *Phytopathology* 84: 91-98.

Fargette, D. and Thresh, J.M. (1994) The Ecology of African Cassava Mosaic Geminivirus. Pages 269-282. In: Ecology of Plant Pathogens. Blakeman, J.A. and Williamson, B. (eds.), CABI International, Wallingford.

Fargette, D., Fauquet, C. and Thresh, J.M. (1994) Analysis and Modelling of the Temporal Spread of African Cassava Mosaic Virus and Implications for Disease Control. *African Crop Science Journal* 2: 449-458.

Fargette, D. and Vie, K. (1995) Simulation of the Effects of Host Resistance, Reversion, and Cutting Selection on Incidence of African Cassava Mosaic Virus and Yield Losses in Cassava. *Phytopathology* 85: 370-375.

Fargette, D., Colon, L.T., Bouveau, R. and Fauquet, C. (1996) Components of Resistance of Cassava to African Cassava Mosaic Virus. *European Journal of Plant Pathology* 102: 645-654.

Fauquet, C. and Fargette, D. (1990). African Cassava Mosaic Virus: Etiology, Epidemiology, and Control. *Plant Disease* 74, 404-411.

Fauquet, C.M., Maxell, D.P., Gronenborn, B., Stanley, J. (2000) Revised proposal for naming geminiviruses. *Archives of Virology* 145: 1743-1761.

Felsenstein, J. (1989) PHYLIP- Phylogeny Inference Package. *Cladistics* 5: 164-166.

Fitch, W.M. and Margoliash, E. (1967). Construction of Phylogenetic Trees. *Science* 155: 279-284.

Fondong, V.N., Pita, J.S., Rey, C., Beachy, R.N., Fauquet, C.M. (1998) First Report of the Presence of East African Cassava Mosaic Virus in Cameroon. *Plant Disease* 82: 1172.

Fondong, V.N., Pita, J.S., Rey, M.E.C., de Kochko, A., Beachy, R.N., Fauquet, C.M. (2000) Evidence of synergism between African cassava mosaic virus and a new double-recombinant geminivirus infecting cassava in Cameroon. *Journal of General Virology* 81: 287-297.

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

FMC Product Catalogue., (1998) Mutation Detection. 85-91

Frischmuth, T., Roberts, S., von Arnim, A. and Stanley, J. (1993) Specificity of Bipartite Geminivirus Movement Proteins. *Virology* 196: 666-673.

Frohlich, D. R., Brown, J.K., Bedford, I.D. and Markham, P.G. (1996) Mitochondrial 16S ribosomal subunit as a molecular marker in *Bemisia*, and implications for population variability. Pages 143-145 in: *Bemisia '95: Taxonomy, Biology, Damage, Control and Management* (D. Gerling and R. T. Mayer, eds). Intercept Publications, Wimborne, UK.

Frohlich, D.R., Torres-Jerez, I., Bedford, I.D., Markham, P.G. and Brown, J.K. (1999) A phylogeographic analysis of the *Bemisia tabaci* species complex based on mitochondrial DNA markers. *Molecular Ecology* 8:1593-1602.

Gawel, N.J. and Bartlett, A.C. (1993) Characterisation of differences between whiteflies using RAPD-PCR. *Insect Molecular Biology* 2: 33-38.

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.

Gibson, R.W. and Otim-Nape, G.W. (1997) Factors determining Recovery and Reversion in Mosaic-affected African Cassava Mosaic Virus Resistant Cassava. *Annals of Applied Biology* 131: 259-271.

Gill, R.J. (1992) A review of the sweetpotato whitefly in southern California. *Pan-Pacific Entomologist* 68:144-152.

Goodman, R.M. (1977) Single-Stranded DNA Genome in a Whitefly-Transmitted Plant Virus. *Virology* 83: 171-179.

Gough, K.C., Cockburn, W. and Whitelam, G.C. (1999) Selection of phage-display peptides that bind to cucumber mosaic virus coat protein. *Journal of Virological Methods* 79: 169-180.

Griep, R.A., Prins, M., van Twisk, C., Keller, J.H.G., Kerschbaumer, R.J., Kormelink, R., Goldblach, R.W. and Schots, A. (2000) Application of Phage Display in Selecting *Tomato spotted wilt virus*-Specific Single-Chain Antibodies (scFvs) for Sensitive Diagnosis in ELISA. *Phytopathology* 90:183-190.

Haley, A., Zhan, X., Richardson, K., Head, K. and Morris, B. (1992) Regulation of the Activities of African cassava mosaic virus Promoters by the AC1, AC2, and AC3 Gene Products. *Virology* 188: 905-909.

Haley, A., Richardson, K., Zhan, X. and Morris, B. (1995) Mutagenesis of the BC1 and BV1 genes of African cassava mosaic virus identifies conserved amino acids that are essential for spread. *Journal of General Virology* 76: 1291-1298.

Hanley-Bowdoin, L., Settlage, S.B., Orozco, B.M., Nagar, S., Robertson, D. (1999) Geminiviruses: Models for Plant DNA Regulation, Transcription, and Cell Cycle Regulation. *Critical Reviews in Plant Sciences* 18: 71-106.

Harper, K., Toth, R.L., Mayo, M.A. and Torrance, L. (1999) Properties of a panel of single chain variable fragments against *Potato leafroll virus* obtained from two phage display libraries. *Journal of Virological Methods* 81: 159-168.

Harrison, B.D., Barker, H., Bock, K.R., Guthrie, E.J., Meredith, G., Atkinson, M. (1977) Plant viruses with circular single-stranded DNA. *Nature* 270: 760-762.

Harrison, BD., (1985) Advances in Geminivirus Research. *Annual Review of Phytopathology* 23: 55-82.

Harrison, B.D., Lennon, A.M., Massalski, P.R., Robsinson, D.J., Thomas J.E. (1986) Geographical Variation in African Cassava Mosaic Virus. Proceedings of 10th Plant Virus Epidemiology Workshop, pp. 9-10.

Harrison, B.D., Lennon, A.M., Massalski, P.R., Robinson, D.J., Thomas, J.E. (1986) Geographical variation in geminivirus isolates associated with cassava mosaic disease. *Annual Report of the Scottish Crop Research Institute* pp. 179-180.

Harrison, B.D., Muniyappa, V., Swanson, M.M., Roberts, I.M. and Robsinson, D.J. (1991) Recognition and differentiation of seven whitefly-transmitted geminiviruses from India, and their relationships to African cassava mosaic and Thailand mung bean yellow mosaic viruses. *Annals of Applied Biology* 118: 299-308.

Harrison, B.D., Zhou, X., Otim-Nape, G.W., Liu, Y., Robinson, D.J. (1997) Role of a novel type of double infection in the geminivirus-induced epidemic of severe cassava mosaic in Uganda. *Annals of Applied Biology* 131: 437-448.

Harrison, B.D. and Robinson, D.J. (1999) Natural Genomic and Antigenic Variation in Whitefly-Transmitted Geminiviruses (Begomoviruses). *Annual Review of Phytopathology* 37: 369-398.

Herren, H.R. (1994) Cassava Pest and Disease Management: An Overview. *African Crop Science Journal* 2: 345-353.

Hofer, P., Bedford, I.D., Markham, P.G., Jeske, H. and Frischmuth, T. (1997) Coat Protein Gene Replacement Results in Whitefly Transmission of an Insect Nontransmissable Geminivirus Isolate. *Virology* 236: 288-295.

Hong, Y.G., Robinson, D.J., 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.

Hong, Y.G. and Harrison, B.D. (1995) Nucleotide sequences from tomato leaf curl viruses from different countries: evidence for three geographically separate branches in evolution of the coat protein of whitefly-transmitted geminiviruses. *Journal of General Virology* 76: 2043-2049.

Hong, Y. and Stanley, J. (1995) Regulation of African cassava mosaic virus complementary-sense gene expression by N-terminal sequences of the replication-associated protein AC1. *Journal of General Virology* 76: 2415-2422.

Hoogenboom, H.R., Marks, J.D., Griffiths, A.D. and Winter, G. (1992) Building Antibodies from their Genes. *Immunological Reviews* 130: 41-67.

Hoogenboom, H.R. (1997) Designing and optimizing library selection strategies for generating high-affinity antibodies. *Tibtech* 15: 62-70.

Howarth, A.J. and Vandemark, G.J. (1989) Phylogeny of Geminiviruses. *Journal of General Virology* 70: 2717-2727.

Innis, M.A., Gelfand, D.H., Snisky, J.J. (1995) Heteroduplex Mobility Assays for Phylogenetic Analysis. In: PCR Strategies, Academic Press, pp. 287-297.

Jericho, C. and Thompson, G.J. (1999) Characterization of viruses infecting cassava in South Africa. 37th Congress South African Society of Plant Pathology, 17-20 January, University of Natal, Pietermaritzburg.

Klinkenberg, F.A., Ellwood, S. and Stanley, J. (1989) Fate of African Cassava Mosaic Virus Coat Protein Deletion Mutants after Agroinoculation. *Journal of General Virology* 70: 1837-1844.

Klinkenberg, F.A. and Stanley, J. (1990) Encapsidation and spread of African cassava mosaic virus DNA A in the absence of DNA B when agroinoculated to *Nicotiana benthamiana. Journal of General Virology* 71: 1409-1413.

Kohler, G. and Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256:495.

Konate, G., Barro, N., Fargette, D., Swanson, M.M. and Harrison, B.D. (1995) Occurrence of whitefly-transmitted geminiviruses in crops in Burkina Faso, and their serological detection and diferentiation. *Annals of Applied Biology* 126: 121-129.

Kreis, S. and Whistler, T. (1997) Rapid identification of measles virus strains by the heteroduplex mobility assay. *Virus Research* 47: 197-203.

Lazarowitz, S.G. (1992) Geminiviruses: Genome structure and gene function. *Critical Reviews in Plant Science* 11: 327-349.

Legg, J.P. (1994) *Bemisia tabaci*: The whitefly vector of cassava mosaic geminiviruses in Africa: an Ecological Perspective. *African Crop Science Journal* 2: 437-448.

Lin, S.S., Hou, R.F., Yeh, S.D. (2000) Heteroduplex Mobility and Sequence Analyses for Assessment of Variability of *Zucchini yellow mosaic virus*. *Phytopathology* 90: 228-235.

Liu, S., Bedford, I.D., Briddon, R.W. and Markham, P.G. (1997) Efficient whitefly transmission of African cassava mosaic geminivirus requires sequences from both genomic components. *Journal of General Virology* 78: 1791-1794.

Macintosh, S., Robinson, D.J. and Harrison, B.D. (1992) Detection of three whitefly-transmitted geminiviruses occurring in Europe by tests with heterologous monoclonal antibodies. *Annals of Applied Biology* 121: 297-303.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D. and Winter, G. (1991) By-passing Immunization: Human antibodies from V-gene Libraries Displayed on Phage. *Molecular Biology* 222: 581-597.

Marks, J.D., Griffiths, A.D., Malmqvist, M., Clackson, T.P., Bye, J.M. and Winter, G. (1992) By-passing Immunization: Building High Affinity Human Antibodies by Chain Shuffling. *Biotechnology* 10: 779-783.

Maruthi M. N., Colvin, J. and Seal, S. (2000) Virus-vector co-adaptation in the cassava mosaic pathosystem. *EWSN (European Whitefly Studies Network) Newsletter*, November 2000, Issue No. 07.

Maynard, J.H. and Upadhyaya, M. (1998) High-Throughput Screening for the Detection of Unknown Mutations: Improved Productivity using Heteroduplex Analysis. *BioTechniques* 25: 648-651.

Morris, B., Richardson, K., Eddy, P., Zhan, X., Haley, A. and Gardner, R. (1991) Mutagenesis of the AC3 open reading frame of African cassava mosaic virus DNA A reduces DNA B replication and ameliorates disease symptoms. *Journal of General Virology* 72: 1205-1213.

Muniyappa, V. (1980) Chapter 3: Whiteflies, In: Harris K.F. and Maramorosch K. ed. *Vectors of Plant Pathogens,* Academic Press, pp. 39-73.

Navot, N., Pichersky, E., Zeidan, M., Zamir, D., Czonek, H. (1991) Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single genomic component. *Virology* 185: 151-161.

Nestel, B. (1993) A new outlook for an ancient crop. *Biotechnology* 1: 53-39.

Nissim, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D. and Winter, G. (1994) Antibody fragments from a 'single pot' phage display library as immunochemical reagents. *Embo Journal* 13: 692-698.

Nweke, F.I. and Spencer, D.S.C. (1995) Future Prospects for Cassava Root Yield in Sub Saharan Africa. *Outlook on Agriculture* 24: 35-42.

Offei, S.K. and Owuna-Kwakye, M. (1999) First Report of East African Cassava Mosaic Begomovirus in Ghana. *Plant Disease* 83: 877.

Ogbe, F.O., Songa, W. and Kamau, J.W. (1996) Survey of the incidence of African cassava mosaic and East African cassava mosaic viruses in Kenya and Uganda using a monoclonal antibody based diagnostic test. *Roots* 3: 10-14.

Ogbe, F.O., Atiri, G.I., Robinson, D., Winter, S., Dixon, A.G.O., Quin, F.M., Thotappilly, G. (1999) First Report of East African Cassava Mosaic Begomovirus in Nigeria. *Plant Disease* 83: 398.

Olsen, K.M. and Schaal, B.A. (1999) Evidence on the origin of cassava: Phylogeography of *Manihot esculenta. Proceedings of the Natural Academy of Sciences, USA* 96: 5586-5591.

Orozco, B.M., Gladfelter, H.J., Settlage, S.B., Eagle, P.A., Gentry, R.N. and Hanley Bowdoin, L. (1998) Multiple *Cis* Elements Contribute to Geminivirus Origin Function. *Virology* 242: 346-356.

Otim-Nape, G.W., Bua, A. and Baguma, Y. (1994) Accelerating the Transfer of Improved Production Technologies: Controlling African Cassava Mosaic Virus Disease Epidemics in Uganda. *African Crop Science Journal* 2: 479-495.

Otim-Nape, G.W., Thresh, J.M., Bua, A., Baguma, Y. and Shaw, M.W. (1998a) Temporal spread of casava mosaic virus disease in a range of cassava cultivars in different agroecological regions of Uganda. *Annals of Applied Biology* 133: 415-430.

Otim-Nape, G.W. and Thresh, J.M. (1998b) Cassava mosaic virus disease. *Roots* 4: 16-17.

Padidam, M., Beachy, R.N., Fauquet, C.M. (1995) Classification and identification of geminiviruses using sequence comparisons. *Journal of General Virology* 76: 249-263.

Padidam, M., Beachy, R.N. and Fauquet, C.M. (1996) The Role of AV2 ("Precoat") and Coat Protein in Viral Replication and Movement in Tomato Leaf Curl Geminivirus. *Virology* 224: 390-404.

Padidam, M., Maxwell, D.P., Fauquet, C.M. (1997) A proposal for naming geminiviruses. *Archives of Virology* 142: 2253-2262.

Padidam, M., Sawyer, S., Fauquet, C.M. (1999) Possible Emergence of New Geminiviruses by Frequent Recombination. *Virology* 265: 218-225.

Pascal, E., Sanderfoot, A.A., Ward, B.M., Medville, R., Turgeon, R. and Lazarowitz, S.G. (1994) The Geminivirus BR1 Movement Protein Binds Single-Stranded DNA and Localizes to the Cell Nucleus. *The Plant Cell* 6: 995-1006.

Paximadis, M., Dusterwald, G., Duyver, L., Rey, M.E.C. (1997) Preliminary investigation of the causal agent(s) of a disease causing leaf curl of tobacco in South Africa. *Plant Pathology* 46: 282-290.

Perring, T.M., Cooper, A.D., Rodriguez, R.J., Farrar, C.A. and Bellows, T.S.J. (1993) Identification of a whitefly species by genomic and behavioural studies. *Science,* 259: 74-77.

Pietersen, G., Idris, A.M., Krüger, K. and Brown, J.K. (2000) *Tomato curly stunt virus*, a new begomovirus within the TYLCV-IS cluster causing a severe disease of tomato in South Africa. *Plant Disease* 84: 810.

Prescott L.M., Harley, J.P. & Klein D.A. (1993) Microbiology: 2nd Edition. WCB Publishers, pp. G16.

Roberts, I.M., Robinson, D.J. and Harrison, B.D. (1984) Serological Relationships and Genome Homologies among Geminiviruses. *Journal of General Virology* 65: 1723-1730.

Robinson, D.J., Harrison, B.D., Sequeira, J.C., Duncan, G.H. (1984) Detection of strains of African cassava mosaic virus by nucleic acid hybridisation and some effects of temperature on their multiplication. *Annals of Applied Biology* 105: 483-493.

Rochester, D.E., Beachy, R.N., Fauquet, C.M., (1993) Geminivirus nomenclature: the need to set taxonomic standards. *Archives of Virology* 132: 221-224.

Rojas, M.R., Gilbertson, R.L., Russell, D.R., Maxwell, D.P. (1993) Use of Degenerate Primers in the Polymerase Chain Reaction to Detect Whitefly-Transmitted Geminiviruses. *Plant Disease* 77: 340-347.

Ronda, M., Adan, A., Beitia, F., Cifuentes, D. and Cenis, J.L. (2000) Interbreeding between biotypes of *Bemisia tabaci*. *EWSN (European Whitefly Studies Network) Newsletter*, November 2000, Issue No. 07.

Rosell, R.C., Bedford, I.D., Frohlich, D.R., Gill, R.J., Markham, P.G., Brown, J.K. (1997) Analyses of morphological variation in distinct populations of *Bemisa tabaci. Annals of the Entomological Society of America*, 90: 575-589.

Rosner, A., Spiegel, S., Maslenin, L., Levy, D. (1998) The use of S1 nuclease treatment of hybrid PCR products for the differentiation between PVY isolates. *Annals of Applied Biology* 132: 107-114.

Rosner, A., Maslenin, L., Spiegel, A. (1999) Double-stranded conformation polymorphism of heterologous RNA transcripts and its use for virus strain differentiation. *Plant Pathology* 48: 235-239.

Rybicki, E.P., (1994) A phylogenetic and evolutionary justification for three genera of *Geminiviridae. Archives of Virology* 139: 49-77.

Rybicki, E.P. (1998) A proposal for naming geminiviruses: a reply by the Geminivirdae Study Group Chair. *Archives of Virology* 143: 421-424.

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (2nd Edition). Cold Spring Harbor Laboratory Press.

Sanderfoot, A.A. and Lazarowitz, S.G. (1995) Co-operation in Viral Movement: The Geminivirus BL1 and Movement Protein Interacts with BR1 and Redirects It from the Nucleus to the Cell Periphery. *The Plant Cell* 7: 1185-1194.

Sanz, A.I., Fraile, A., Garcia-Arenal, F., Zhou, X., Robinson, D.J., Khalid, S., Butt, T.and Harrison, B.D. (2000) Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. *Journal of General Virology* 81: 1839-1849.

Saunders, K., Lucy, A. and Stanley, J. (1991) DNA forms of the geminivirus African cassava mosaic virus consistent with a rolling circle mechanism of replication. *Nucleic Acids Research* 19: 2325-2330.

Schaffer, R.L., Miller, C.G. and Petty, I.T.D. (1995) Virus and Host-Specific Adaptations in the BL1 and BR1 Genes of Bipartite Geminiviruses. *Virology* 214: 330-338.

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

Sheets, M.D., Amdersdorfer, P., Finnern, R., Sargent, P., Lindqvist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J.C. and Marks, J.D. (1998) Efficient construction of a large nonimmune phage antibody library: The production of high-affinity human single-chain antibodies to human antigens. *Proceedings of the National Academy of Science* 95: 6157-6162.

Sparks M. (1996) Cassava, the root of all good. The Star, Monday, 15th April, pp.6

Stanley, J. (1983) Infectivity of the cloned geminivirus genome requires sequences from both DNAs. *Nature* 305: 643-645.

Stanley, J., Frischmuth, T. and Ellwood, S. (1990) Defective viral DNA ameliorates symptoms of geminivirus infection in transgenic plants. *Proceedings of the Natural Academy of Sciences, USA* 87: 6291-6295.

Stanley, J. (1995) Analysis of African Cassava Mosaic Virus Recombinants Suggest Strand Nicking Occurs within the Conserved Nonanucleotide Motif during the Initiation of Rolling Circle DNA Replication. *Virology* 206: 707-712.

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

Sunter, G., Stenger, D.C. and Bisaro, D.M. (1994) Heterologous Complementation by Geminivirus AL2 and AL3 Genes. *Virology* 203: 203-210.

Susi, P., Ziegler, A. and Torrance, L. (1998) Selection of Single-Chain Variable Fragment Antibodies to Black Currant Reversion Associated Virus from a Synthetic Phage Display Library. *Phytopathology* 88: 230-233.

Swanson, M.M., Varma, A., Muniyappa, V. and Harrison, B.D. (1992a) Comparative epitope profiles of the particle proteins of whitefly-transmitted geminiviruses from nine crop legumes in India. *Annals of Applied Biology* 120: 425-433.

Swanson, M.M., Brown, J.K., Poulos, B.T. and Harrison, B.D. (1992b) Genome affinities and epitope profiles of whitefly-transmitted geminiviruses from the Americas. *Annals of Applied Biology* 121: 285-296.

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

Swofford, D. L. (1991) PAUP User's Manual. Illinois Natural History Survey, Champaign, IL.

Thomas, J.E., Massalski, P.R., Harrison, B.D. (1986) Production of Monoclonal Antibodies to African Cassava Mosaic Virus and Differences in Their Reactivities with Other Whitefly-transmitted Geminiviruses. *Journal of General Virology* 67: 2739-2748.

Thresh, M., Fargette, D., Otim-Nape, G.W. (1994a) The viruses and virus diseases of cassava in Africa. *African Crop Science Journal* 2: 459-478.

Thresh, J.M., Otim-Nape, G.W. and Jennings, D.L. (1994b) Exploiting resistance to African cassava mosaic virus. *Aspects of Applied Biology* 39: 51-60.

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

Torrance, L. (1995) Use of monoclonal antibodies in plant pathology. *European Journal of Plant Pathology* 101: 351-363.

Torrance, L. (1998) Developments in serological methods to detect and identify plant viruses. *Plant Cell, Tissue and Organ Culture* 52: 27-32.

Trench, T.N., Martin, M.M., Hemmes, E.A. (1985) An assessment of cassava African mosaic disease in South Africa and Swaziland. *South African Journal of Plants and Soil* 2: 169-171.

Uhde, K., Kerschbaumer, R.J., Koenig, R., Hirschl, S., Lemaire, O., Boonham, N., Roake, W. and Himmler, G. (2000) Improved detection of *Beet necrotic yellow vein virus* in a DAS-ELISA by means of antibody single chain fragments (scFv) which were selected by protease-stable epitopes from phage display libraries. *Archives of Virology* 145: 179-185.

Van Regenmmortel, M.H.V., Bishop, D.H.L., Fauquet, C.M., Mayo, M.A., Maniloff, J., Calisher, C.H. (1997) Guidelines to the demarcation of virus species. *Archives of Virology* 142: 1505-1518.

von Arnim, A., Frischmuth, T. and Stanley, J. (1993) Detection and Possible Functions of African Cassava Mosaic Virus DNA B Gene Products. *Virology* 192: 264-272.

Wang, Y.H. and Griffith, J. (1991) Effects of Bulge Composition and Flanking Sequence on the Kinking of DNA by Bulged Bases. *Biochemistry* 30: 1358-1363.

Whitlow, M. and Filpula, D. (1991) Single-Chain Fv Proteins and Their Fusion Proteins. *METHODS: A Companion to Methods in Enzymology* 2: 97-105.

Wilson, D.R. and Finlay, B.B. (1998) Phage display: applications, innovations, and issues in phage and host biology. *Canadian Journal of Microbiology* 44: 313-329.

Winter, G. and Milstein, C. (1991) Man-made antibodies. *Nature* 349: 293-299.

Winter, S. (1998) Analysis of Cassava Samples from Guinae to study Presence of Viruses contributing to Cassava Mosaic Disease (CMD). *WASEED Bulletin* 2: 7-8.

Wyatt, S.D. and Brown, J.K. (1996) Detection of Subgroup III Geminivirus Isolates in Leaf Extracts by Degenerate Primers and Polymerase Chain Reaction. *Phytopathology* 86: 1288-1293.

Zhou, X., Liu, Y., Calvert, L., Munoz, C., Otim-Nape, G.W., Robinson, D.J., Harrison, BD. (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.

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

Ziegler, A., Torrance, L., Macintosh, S.M., Cowan, G.H. and Mayo, M.M. (1995) Cucumber Mosaic Cucumovirus Antibodies from a Synthetic Phage Display Library. *Virology* 214: 235-238.

Ziegler, A., Mayo, M.A. and Torrance, L. (1998) Synthetic Antigen from a Peptide Library Can Be an Effective Positive Control in Immunoassays for the Detection and Identification of Two Geminiviruses. *Phytopathology* 88: 1302-1305.