

**THE USE OF MAIZE STREAK VIRUS (MSV) REPLICATION-
ASSOCIATED PROTEIN MUTANTS IN THE DEVELOPMENT OF
MSV-RESISTANT PLANTS**

Dionne Natalie Shepherd

Thesis presented for the degree of Doctor of Philosophy in the Department of
Molecular and Cell Biology,
University of Cape Town
February, 2003

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Certification by Supervisors

In terms of paragraph GP9 of the regulations for the degree of Doctor of Philosophy at the University of Cape Town, I certify that I approve of the inclusion in this thesis of material already published, or submitted for publication by candidate Dionne Natalie Shepherd.

Signature Removed

Professor J. A. Thomson

Department of Molecular and Cell Biology

University of Cape Town

Signature Removed

Professor E. P. Rybicki

Department of Molecular and Cell Biology

University of Cape Town

February, 2003

Abstract

Maize streak virus (MSV) is the type member of the *Mastrevirus* genus of the *Geminiviridae*. As the causal agent of maize streak disease (MSD), MSV is the most significant pathogen of maize in Africa, resulting in crop yield losses of up to 100%. Transmitted by leafhoppers (*Cicadulina* spp.), MSV is indigenous to Africa and neighbouring Indian Ocean Islands. Despite maize being a crucial staple food crop in Africa, the average maize yield per hectare in Africa is the lowest in the world, leading to food shortages and famine. A major contributing factor to these low yields is MSD.

To genetically engineer MSV-resistant maize using the pathogen-derived resistance (PDR) strategy, the viral replication-associated (Rep) protein gene was targeted, whose multifunctional products Rep and RepA are the only viral proteins essential for replication. *Rep* constructs had previously been made containing deleterious mutations in several conserved amino acid motifs. In this study, these mutants and the wild type *Rep* gene were truncated to remove key motifs involved in viral replication. A quantitative PCR assay was developed to determine the effects of the mutant and truncated *Reps* on viral replication in black Mexican sweetcorn (BMS) suspension cells. The MSV-sensitive grass *Digitaria sanguinalis* was then transformed with *Rep* constructs that inhibited MSV replication in BMS, and transgenic lines were tested for virus resistance. Several plants of a *D. sanguinalis* line transgenic for a mutated full-length *Rep* gene showed excellent resistance (immunity) to MSV, but the transgene had negative effects on aspects of plant growth and development. Transformation with a mutated/truncated *Rep* gene resulted in healthy fertile transgenic *D. sanguinalis* plants, many of which showed good MSV resistance. Fertile maize (Hi-II) T₁ transgenic plants expressing the truncated/mutated *Rep* gene have been obtained, the offspring of which will be tested for resistance to MSV. Considering the success in achieving MSV-resistant *D. sanguinalis*, there is good reason to believe that the transgenic maize will too be resistant to MSV.

The transient expression studies in BMS provided some interesting insights into the mechanics of MSV replication and its interaction with host factors. A *Rep* construct with a mutation (*Rb*⁻) abolishing the protein's interaction with the host retinoblastoma-related (RBR) protein, previously thought to be required for viral replication, surprisingly supported high-level viral replication in BMS. A virus carrying the *Rb*⁻ mutation was infectious in maize; however, one of the nucleotides of the three-nucleotide mutation reverted at an extremely high frequency. A study was carried out to

determine the time taken for the nucleotide reversion to occur, and the point at which the revertant population superseded the original mutant population. These data will possibly enable the mutation rate of MSV to be calculated, which is valuable information when attempting to design resistance strategies that cannot be overcome by mutation of the viral genome. Further studies were aimed towards determining the selective advantage and the absolute requirement of the single nucleotide reversion for viral infectivity in maize, as well as its effect on viral replication in BMS.

In an attempt to further define the role of RepA in the virus' lifecycle, separate intronless Rep and RepA constructs were made, both wild type and containing the Rb⁻ mutation. The effects of the presence and absence of RepA on viral replication and infectivity in maize were determined. It was concluded that the Rb⁻ mutation had no effect on the role of Rep or RepA in the initiation of viral replication in BMS suspension cells. Overexpression of RepA inhibited the replication functions of Rep in BMS, and the absence of RepA had the effect of slightly reducing replication levels. These data indicate that a precise balance of Rep and RepA is required for optimal replication of the viral genome, and altering the wild type ratios of the two proteins has a negative effect on viral replication. RepA was required for efficient infectivity of MSV in maize, although results indicated that a mutant virus unable to express RepA may be capable of establishing a very weak infection.

Acknowledgements

It is difficult to know where to begin: there are so many people who have contributed to this project.

A good starting point would be to thank my previous lecturer and supervisor at Rhodes University, **Prof. Don Hendry**. This wonderful man was responsible for kindling my interest in virology, so that when it came to choosing a PhD project at UCT, working on *Maize streak virus* was an obvious choice.

From then on it was a pleasure to work with my co-supervisors, **Prof. Jennifer Thomson** and **Prof. Ed Rybicki**, who provided a wonderful mixture of leaving me free to pursue my own day-to-day goals in my own weird hours, but with constant guidance, support and enthusiasm. Although I was reluctant at the time, I must thank Jennifer for making me down tools and start writing this thesis. If it wasn't for her I would still be in the lab trying to "tie up loose ends"! Thanks also to Jennifer and Ed for their generous financial support throughout this project, and for their invaluable constructive criticism of this thesis.

The work in this project could never have been achieved without the help of **Marion Bezuidenhout**, who was more than simply my tissue culture technician, she was indispensable in my life! Only a week before the printing of this thesis Marion was still helping me by extracting DNA from maize leaves. Although she was also the technician of another lab, she never hesitated in helping me, from subculturing the enormous amount of tissue culture involved in this project, to looking after the transgenic plants (her "babies") and helping with hundreds of DNA extractions. Finally, she is a very warm friend. Thank you Marion.

Darren Martin played an enormous role in this project. He was a hands-on supervisor, someone who was always ready to give advice, and responsible for many of the ideas that led to the work in this thesis. I am also extremely grateful to Darren for proofreading the entire Literature Review – an odious task – and for his constructive criticism of many of the results presented in this thesis.

There are many people who contributed towards the work presented in this thesis. Those to be thanked are **David McGivern** for the yeast-two-hybrid analysis; **Rikus Kloppers** for looking after and self-pollinating the maize; **Ziyaad Valley-Omar** for the doing the RNA extractions, **Eric van der Walt** for the isolation of the viral genomes from leafhopper-inoculated maize and doing the sequence alignments; **Di James** for doing all the sequencing reactions in this project, and **Darren Martin** for performing most of the agroinoculations and symptom analysis.

Drs **Margaret Boulton** and **David McGivern** are to be thanked for their help and friendship when I worked in their lab in Norwich, UK, and for their personal communications of various results presented in the Literature Review.

For their friendship, entertainment and assistance over the years I thank my co-workers **Janet Willment**, **Leigh Berrie**, **Arvind Varsani**, **Eric van der Walt** and **Darren Martin**. Janet, I also thank you for taking me under your wing when I first joined the lab and always having time to help and encourage me. I hope that I can play a similar role in a young, unsure, unconfident student's life.

Thank you to the departmental assistants for helping with the day-to-day running of the lab and plant growth rooms, especially **Sheila Koeries**, **Jackie McCann**, **Donald Solomons** and **Melvin Keating**.

Finally, a big thank you to my family: **Mom**, **Geoff**, **Dad** and **Suzie** for all being the most caring and supportive parents I could hope for (you must I have thought I would remain a student forever!); **Carla** and **Dave** for apparently being interested and impressed by my work, and **Ryan** and **Georgia** for just being cute!

The biggest thank you of all to **Richard**, who has been a pillar of support throughout the writing of this thesis and indeed the whole project. He has worn many hats – financier, driver, friend and lover. What would I do without you!

Contents

Chapter 1: Literature Review	1
Chapter 2: Inhibition of MSV Replication by Transient and Transgenic Expression of MSV Rep Mutants	62
Chapter 3: The Effect of Mutations in the MSV Rep RBR Protein-Interaction Motif on Viral Infectivity in Maize	107
Chapter 4: A Comparison of the Replication Efficiencies of MSV Rep Variants.	140
Chapter 5: General Conclusion	158
Appendix A: Summary of Plasmid and Construct Properties.....	162
Appendix B: Key for Recognition of Severity of Maize Streak Disease (MSD) Symptoms on Maize Leaves	168
References	170

Chapter 1

Literature review

Contents

1.1 INTRODUCTION	2
1.2 THE MOLECULAR BIOLOGY OF GEMINIVIRUSES	3
1.2.1 The <i>Geminiviridae</i>	3
1.2.1.1 The evolutionary origin of geminiviruses	8
1.2.2 The <i>Maize Streak Virus</i> Genome	9
1.2.2.1 The long intergenic region	9
1.2.2.2 The short intergenic region	20
1.2.2.3 The complementary sense genes (<i>Rep</i> and <i>RepA</i>)	21
1.2.3 Replication	39
1.2.4 The MSV Lifecycle	42
1.3 THE EPIDEMIOLOGY AND CONTROL OF MAIZE STREAK DISEASE	47
1.3.1 The Epidemiology of Maize Streak Disease	47
1.3.2 Strategies for the Control of Maize Streak Disease	51
1.3.2.1 Classical cross-breeding for MSV resistance	52
1.3.2.2 Genetic engineering of plants for geminivirus resistance	54
1.4 PROJECT AIMS	60

1.1 INTRODUCTION

"The disorder of the mealie plant, locally described as 'Mealie Blight', 'Mealie Yellows' or 'Striped Leaf Disease', belongs to a group of plant troubles arising from obscure causes..." was how maize streak disease (MSD) was first described by Fuller (1901) in Natal, South Africa. Fuller mistakenly attributed the disease to a soil disorder, but in retrospect it is quite clear that the "mealie variegation" he described and drew in minute detail can be attributed to *Maize streak virus* (MSV). In the hundred years since this first report, scientists have come a long way in identifying and analysing the causal agent of MSD, to the point where we can now design effective strategies to control or even eliminate the disease in maize.

The first milestone in MSD research was reached in 1924, when H.H Storey determined that a virus obligately transmitted by leafhopper species of the genus *Cicadulina* was the causal agent of MSD (Storey, 1924). Storey named the virus *Maize streak virus*. Storey (1931) was also the first to describe the genetic basis of transmission of MSV by *Cicadulina mbila*, and that resistance to MSD in maize could be inherited (Storey and Howland, 1967).

MSV particles were first purified and visualised by Bock *et al.* (1974): they were found to have a novel twinned quasi-icosahedral (geminiate) shape, from which the name 'geminivirus' was born. This was followed by the unexpected discovery in 1977 that geminivirus particles contain circular single-stranded DNA (ssDNA), a genome type never before observed in plant viruses (Goodman *et al.*, 1977a, b; Harrison *et al.*, 1977). These novel characteristics led to the proposal of a new virus group - the geminiviruses - consisting of MSV and other viruses with geminate particle morphology and ssDNA genomes. This was officially accepted by the International Committee for Virus Taxonomy in 1978. The group was subsequently given the status of a taxonomic family in 1995, and by 2000 the family *Geminiviridae* had three genera (Briddon and Markham, 1995; Rybicki *et al.*, 2000), with MSV as the type member of the genus *Mastrevirus*.

Following the visualisation of circular ssDNA-containing geminivirus particles in the 1970s, the next major advance in geminivirus research came in the early 1980s with the cloning and sequence analysis of the first geminivirus genomes. The characterization and sequencing of the bipartite genome of *African cassava mosaic virus* (ACMV) in 1983 (Stanley and Gay, 1983) was followed by monopartite MSV in 1984 (Mullineaux *et al.*, 1984). This important development

led to a new age of geminivirus research, with the start of intensive investigation of geminivirus molecular biology.

1.2 THE MOLECULAR BIOLOGY OF GEMINIVIRUSES

The majority of this section will focus on the wealth of information on geminivirus molecular biology obtained in the 20 years since the first geminivirus genome was sequenced, with particular emphasis on MSV. However, in cases where there is little or no information on certain aspects of MSV biology, analogies will be drawn from more intensively studied geminiviruses. The taxonomy and general properties of the different geminivirus genera are summarised below as an introduction to the family.

1.2.1 The *Geminiviridae*

The taxonomic family *Geminiviridae* is classified into four genera (*Mastrevirus*, *Begomovirus*, *Curtovirus* and *Topocuvirus*; see Table 1 and Figure 1) on the basis of their host range, genome organisation and vector specificity. In addition to these biological characteristics, there is also substantial phylogenetic support for the existence of these genera (Rybicki, 1994; Padidam *et al.*, 1995). Mastreviruses (type member *Maize streak virus*) have monopartite genomes and are transmitted by different leafhopper species (Homoptera: family *Cicadellidae*) and generally infect monocotyledonous plants. Begomoviruses (type member *Bean golden mosaic virus*; BGMV), which comprise the largest genus of the family, are transmitted by a single whitefly species (*Bemisia tabaci*), and all infect dicotyledonous plants. Most have bipartite genomes (called DNA A and DNA B), although there are some viruses in this genus that have monopartite genomes (Rybicki *et al.*, 2000). Curtoviruses (type member *Beet curly top virus*; BCTV) occupy an intermediate position between mastreviruses and begomoviruses in that they have monopartite genomes and are transmitted by leafhoppers, but infect only dicotyledonous hosts. The latest addition to the family, topocuviruses (type member *Tomato pseudo curly top virus*) are similar to curtoviruses in genomic organisation but are transmitted by treehoppers.

TABLE 1.1 Classification and general properties of the *Geminiviridae*

Genus	No. of approved members	Examples of members (Type species first)	Genome size (nt)/Arrangement	Host Range	Vector
<i>Mastrevirus</i>	12 ^a	<i>Maize streak virus</i> (MSV)	2684-2701/ monopartite	Poaceae (cereals, grasses)	Leafhopper
		<i>Wheat dwarf virus</i> (WDV)	2749/ monopartite	Poaceae	Leafhopper
		<i>Sugarcane streak virus</i> (SSV)	2706/ monopartite	Poaceae	Leafhopper
		<i>Bean yellow dwarf virus</i> (BeYDV)	2561/ monopartite	Leguminosae	Leafhopper
		<i>Tobacco yellow dwarf virus</i> (TYDV)	2580/ monopartite	Solanaceae	Leafhopper
<i>Curtovirus</i>	4 ^b	<i>Beet curly top virus</i> (BCTV)	2933/ monopartite	Dicot plants	Leafhopper
		<i>Horseradish curly top virus</i> (HrCTV)	3080/ monopartite	Horseradish, Shepherd's purse	Leafhopper
<i>Topocuvirus</i>	1 ^c	<i>Tomato pseudo-curly top virus</i> (TPCTV)	2861/ monopartite	Solanaceae	Treehopper
<i>Begomovirus</i>	103 ^b	<i>Bean golden mosaic virus</i> (BGMV)	A: 2644; B: 2609/ bipartite	Leguminosae	Whitefly
		<i>Tomato golden mosaic virus</i> (TGMV)	A: 2588; B: 2522/ bipartite	Solanaceae	Whitefly
		<i>African cassava mosaic virus</i> (ACMV)	A: 2779; B: 2724/ bipartite	Euphorbiaceae	Whitefly
		<i>Tomato leaf curl virus</i> (ToLCV)	2766 / monopartite	Solanaceae	Whitefly
		<i>Squash leaf curl virus</i> (SqLCV)	A: 2634; B: 2606/ bipartite	Cucurbitaceae	Whitefly
		<i>Tomato yellow leaf curl virus</i> (TYLCV)	2743-2790/ monopartite	Solanaceae	Whitefly

^aBoulton and Davies (2002)^bFauquet *et al.* (2003)^cBriddon (2002)

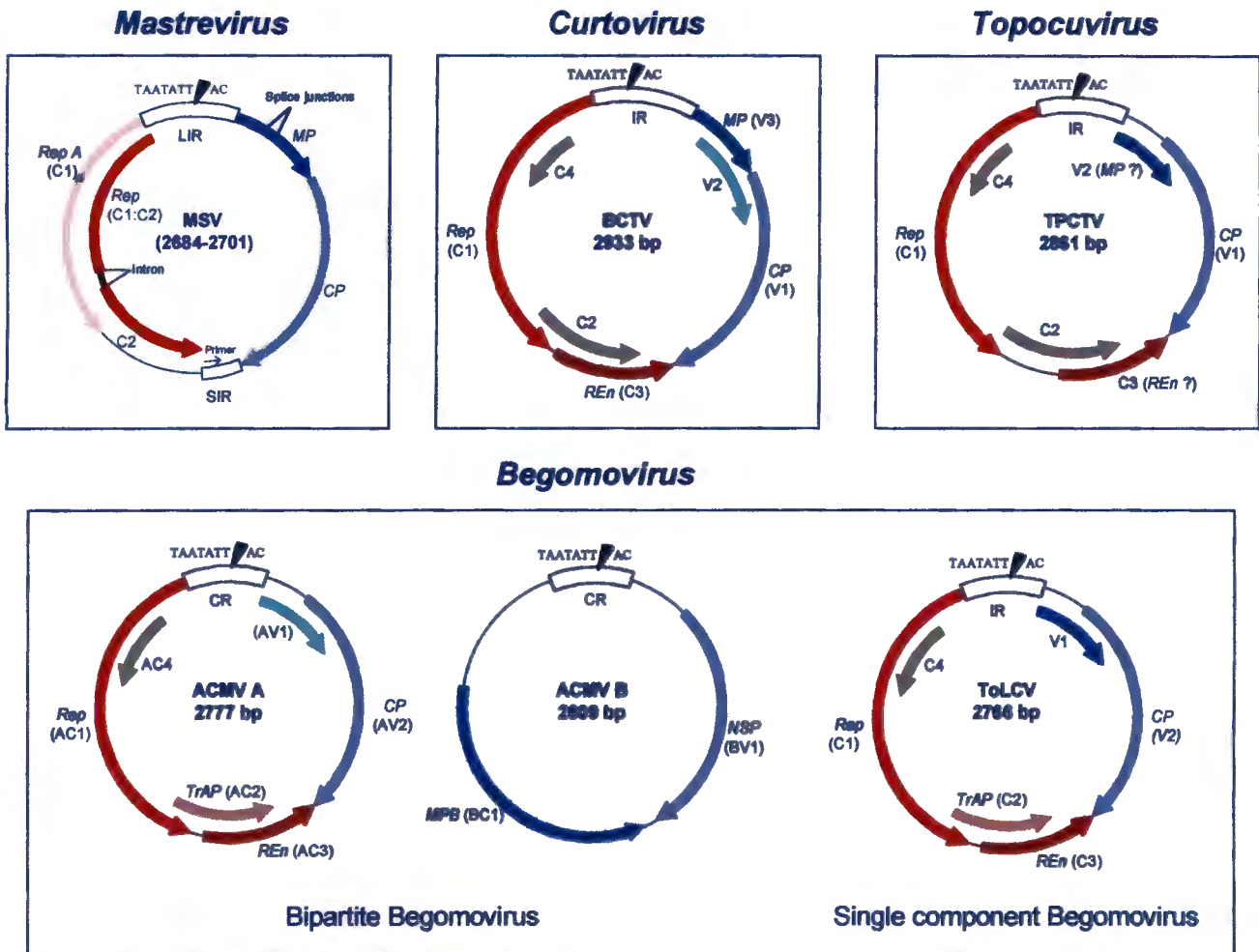


FIGURE 1.1 Genome organization of representative members of the four geminivirus genera. Curved arrows indicate open reading frames, diverging in the complementary (C) and virion (V) senses from an intergenic region (IR) in curtoviruses and topocoviruses, a long intergenic region (LIR) in mastreviruses, and a common region (CR) in bipartite begomoviruses. The CR is a ~200 bp sequence in the begomoviral IR that is conserved between the A and B components of a bipartite begomovirus. The position of the plus strand rolling circle replication (RCR) initiation site (TAATATTAC), situated within the loop of a stem-loop structure, is indicated within each genus' IR. To initiate RCR, a nick (▼) is introduced by the replication associated protein (Rep) at the penultimate A nucleotide of the invariant nonanucleotide sequence. In mastreviruses, bidirectional transcription initiates in the LIR and terminates mainly in the short intergenic region (SIR), which contains signals for polyadenylation. The SIR also functions as the C sense (negative strand) origin of replication. A small (~ 80 nt) ssDNA molecule annealed in the SIR to the plus strand is thought to act as the primer for negative strand replication. Genes have been named according to either their function, if known, or their genetic location. Genes in shades of red have a function in the early stage of the virus life cycle (DNA replication, regulation of transcription, and most likely, interfering with cellular process needed for the replicative cycle). Genes in shades of blue in general have movement and structural functions. Where the function of an ORF is not known, it is coloured in grey. Rep is found in all geminiviruses, although in mastreviruses a unique variant of Rep may also be expressed, called RepA. In this genus, Rep (C1:C2) is expressed from a spliced transcript of ORFs C1 and C2, whereas RepA (C1) is potentially expressed from the unspliced transcript. Other genes with known functions are CP (coat protein gene found in all geminiviruses), MP (movement protein gene found in mastreviruses and curtoviruses), TrAP (transcription activator protein gene in begomoviruses); REn (replication enhancer gene found in begomoviruses and curtoviruses); and MPB and NSP (movement protein gene and nuclear shuttle protein gene, respectively, found on the B component of bipartite begomoviruses). The C4 ORF appears to have different functions in different genera. In curtoviruses it may be involved in tumour induction (Latham *et al.*, 1997), while in monopartite begomoviruses such as TYLCV it is apparently involved in movement (Jupin *et al.*, 1994). Two contrasting studies in the bipartite begomovirus, TGMV, have indicated either a role for C4 in transcriptional regulation of Rep (Groning *et al.*, 1994) or no function at all (Pooma and Petty, 1996). The function of C4 in topocoviruses is unknown. The AV1 ORF in bipartite begomoviruses, sometimes called the pre coat protein gene, or *PreCP*, and the V2 ORF in curtoviruses, may both be involved in regulation of ssDNA accumulation. The V2 ORF in topocoviruses may have a function in movement, but this has not been proven. The topocovirus C3 gene has substantial sequence homology to REn, but its function is unknown. Similarly, although the C2 gene of curtoviruses has homology to TrAP, its function in curtoviruses is unknown. The role of C2 in topocoviruses has not been determined.

It is generally accepted that geminiviruses replicate in the nucleus using a rolling circle replication (RCR) mechanism (Saunders *et al.*, 1991; Stenger *et al.*, 1991). Convincing evidence for this model is the production of supercoiled, open circular, and linear double-stranded (ds) DNA species, a hallmark of the RCR strategy employed by circular ssDNA bacteriophages (Gruss and Ehrlich, 1989; Kornberg and Baker, 1992) and a class of eubacterial plasmids (Baas and Jansz, 1988). As additional evidence for the RCR model, the N-terminal portions of all geminiviral replication associated proteins (Reps) contain three motifs conserved among the replication initiator proteins from other known RC replicons (Koonin and Ilyina, 1992). More recently, electron microscopic visualisation and two-dimensional gel analysis of *Abutilon mosaic virus* (AbMV) confirmed that geminiviruses replicate via rolling circles (Jeske *et al.*, 2001); however, there is a strong possibility that a recombination-related process (recombination-dependent replication; RDR) is an additional replication strategy employed by geminiviruses (Jeske *et al.*, 2001).

As can be seen in Figure 1.1, all geminivirus genomes contain an intergenic region (IR) from which viral genes diverge in both the virion (V) and complementary (C) senses. The IR contains RNA-polymerase II-type promoter sequences that drive the transcription of genes in both genome senses (bidirectional transcription). In general, genes encoded in the virion sense ("V" genes) have functions in virus movement and encapsidation ("late" functions), whereas genes encoded in the complementary sense ("C" genes) are involved in virus replication and transactivation of the virion-sense promoter ("early" functions). The IR of all geminiviruses contains a stem-loop structure within which an invariant nonanucleotide sequence (TAATATT↓AC) contains the initiation site (↓) of RCR (Heyraud-Nitschke *et al.*, 1995; Laufs *et al.*, 1995a; Stanley, 1995).

The only genes shared by all geminiviruses are the Rep and coat protein (CP) genes (Rybicki, 1994; Rybicki *et al.*, 2000), the remainder of each genus' genes differing both in number and arrangement. Mastreviruses have a unique genome composition (reviewed in detail in the next section), being the only geminiviruses to possess a small, (or short) intergenic region (SIR), a Rep transcript with an intron which, depending on whether or not it is spliced, is capable of expressing a RepA protein (unspliced) or a full-length Rep protein (spliced); and virion sense transcripts that are also spliced, encoding the movement protein (MP) and the CP.

The only geminiviruses with two genome components are begomoviruses, such as the Old World (OW) begomovirus, ACMV, and the New World (NW) begomovirus, BGMV. The C sense

strand of the A component of all bipartite begomoviruses encodes Rep, a transcription activator protein (TrAP), and a replication enhancer protein (REn), while the CP is expressed from the A component's V sense strand. In addition, the OW begomovirus A component contains a C4 gene (AC4) on the C sense strand, whose function is unknown, and a "pre-coat protein" gene (PreCP) on the V sense strand. The latter is apparently involved in ssDNA accumulation (Wartig *et al.*, 1997) and/ or movement (Padidam *et al.*, 1996). The B component encodes two proteins involved in viral movement: on the C sense BC1, or otherwise known as the movement protein (MPB), and on the V sense BV1, or nuclear shuttle protein (NSP). The NSP, although unique to bipartite begomoviruses, shows some sequence homology to geminivirus CPs (Rybicki, 1994). Monopartite begomoviruses such as ToLCV have an arrangement of C sense genes identical to that on the A component of bipartite begomoviruses, while the virion strand contains two overlapping coding regions (V1 and V2), both of which are required for infectivity but not for replication (Rigden *et al.*, 1993; Padidam *et al.*, 1996). Satellite DNAs are frequently associated with monopartite begomovirus infections (Dry *et al.*, 1997; Monsoor *et al.*, 1999; Saunders and Stanley, 1999). Whereas the satellite DNAs appear to play no essential role in the disease caused by their associated begomovirus, a recently detected DNA (named DNA β) associated with the otherwise weakly virulent monopartite begomovirus, *Ageratum yellow vein virus* (AYVV) massively enhances the virus' virulence (Saunders *et al.*, 2000).

Curtoviruses, exemplified by BCTV, have an organization of C sense genes similar to that of begomoviruses, while the V sense organisation is similar to that of the monopartite geminiviruses. This has prompted speculation that curtoviruses arose from a recombination event between mastreviruses and begomoviruses (Rybicki, 1994), although more recent evidence suggests that the recombination event involved the 5' portion of *Rep* only, resulting in curtoviruses having a begomovirus-like Rep (Padidam *et al.*, 1999a; Martin *et al.*, unpublished). BCTV has three virion-sense ORFs: these are V1 (CP), V2 and V3 (MP). V1 and V3 are required for infection but not for replication, suggesting that they both possibly have a role in movement (Bridson *et al.*, 1989), while V2 may modulate the conversion of dsDNA to ssDNA (Hormuzdi and Bisaro, 1993).

The topocovirus TPCTV has a genomic organization very similar to that of curtoviruses, although the functions of some of its genes have yet to be elucidated. Although the C sense strand has four ORFs (C1, C2, C3 and C4), only the product of the C1 gene (Rep) is known. While the C3 gene encodes a protein with substantial sequence homology to the begomovirus and curtovirus REn, its

function in topocoviruses has not been determined. The topocovirus V sense strand encodes two potential proteins, V1 (CP) and V2. The latter most likely functions as a MP, although this has not been proven.

1.2.1.1 The evolutionary origin of geminiviruses

There is some similarity between the plant-infecting geminiviruses and nanoviruses, and the vertebrate-infecting circoviruses: all groups have circular ssDNA genomes that multiply by RCR, with each genome component having a stem-loop structure containing a very similar nonanucleotide sequence in the loop. However, nanovirus genomes consist of at least six circular ssDNAs, all about 1 kb in size (Aronson *et al.*, 2000). There is also detectable sequence homology between the Rep proteins of each group, suggesting a similar evolutionary origin. Because of the similarity between these Reps and the replication initiator proteins of prokaryotic ssDNA replicons, such as the bacteriophage ϕ X174, it has been suggested that geminiviruses have a prokaryotic origin (Koonin and Ilyina, 1993; Rigden *et al.*, 1996). Observations lending support to this hypothesis are the detection of replicative form (RF) AbMV DNA in chloroplasts, implying that this virus can replicate in plastids (Groning *et al.*, 1987, 1990), and the fact that ToLCV, TYLCV and ACMV can replicate efficiently in *Agrobacterium tumefaciens* and to low levels in *Escherichia coli* (Rigden *et al.*, 1996; Selth *et al.*, 2002).

Since geminivirus multiplication is heavily reliant on the host DNA replication machinery (Rep being the only virus-encoded protein indispensable for the process), this implies that geminivirus genomes have retained the capacity to be replicated by prokaryotic enzymes. However, this convincing argument has been complicated by the recent discovery of a category of eukaryotic DNA transposons (called *Helitrons*), which transpose by RCR in *Arabidopsis thaliana*, *Oryza sativa* (rice) and *Caenorhabditis elegans* genomes (Kapitonov and Jurka, 2001). *Helitrons* encode a replication initiator protein similar to the Rep from other RC replicons, as well as their own helicase and single-strand binding protein (SSB; in plants only). This structure is more like that of geminivirus Reps, which also have a putative helicase domain, while the Reps of prokaryotic RC replicons are usually assisted by host DNA helicases and SSBs. This finding suggests that geminiviruses might have evolved from plant RC transposons rather than from prokaryotic RC replicons (Kapitonov and Jurka, 2001).

Adding to the debate on the origin of geminiviruses, and of eukaryotic ssDNA viruses in general, is the discovery of a plasmid, obtained from a phytoplasma, that encodes a Rep protein whose N-

terminus has similarities to the Rep of prokaryotic RC replicons, while its C terminus is similar to the helicase domain of the Rep of eukaryotic viruses, especially circoviruses (Oshima *et al.*, 2001). The authors speculate that this phytoplasma plasmid may either be an ancestor of eukaryotic ssDNA viruses, or intriguingly may have arisen out of a recombination event between a prokaryotic plasmid and a eukaryotic virus.

Even more surprising information on the evolutionary history of viruses came from comparative analysis of the 3D NMR structure of the catalytic domain of TYLCV Rep with other viral proteins, which revealed a conserved architecture for a number of functionally diverse proteins (Campos-Olivas *et al.*, 2002). These include the RNA binding domain from U1A and other RNA-binding proteins, and the DNA-binding domains from SV40 T-ag and E1 and E2 from papillomaviruses. This structural conservation suggests there is an evolutionary relationship between primordial ssRNA-binding proteins, RCR initiator proteins (prokaryotic and eukaryotic), and mammalian tumour virus proteins such as SV40 T-ag. Accordingly, the development of ancient RCR elements to the more sophisticated DNA tumour viruses mirrors the evolution of their hosts (Campos-Olivas *et al.*, 2002). No doubt further investigations will reveal more interesting links in the evolutionary histories of viruses and plasmids.

1.2.2 The Maize Streak Virus Genome

The MSV genome is deceptively simple: at ~2.7 kb and with only four genes, it is one of the smallest virus genomes known. However, for this very reason, the regulation and interactions of MSV genes are extremely complex. To perform the tasks required by the virus to establish an infection, MSV genes and their products must have several distinct activities rather than each gene encoding a product with a single function. They must also be capable of interaction with each other, with viral DNA and with host factors, and have a role in regulating the virus life cycle. As will be seen in the following description of the MSV genome, MSV *Rep* is a perfect example of multifunctionality, playing a pivotal role in the virus life cycle.

1.2.2.1 The long intergenic region

The MSV LIR contains bidirectional C and V sense promoters and associated transcriptional regulatory elements, sites for the binding of Rep and plant nuclear factors for gene expression and replication, and the plus strand (V sense) origin of replication. The latter is found in the highly conserved nonanucleotide sequence TAATATTAC, situated in the loop of a stem-loop structure.

The plus-strand nonanucleotide sequence, present in all geminivirus IRs, is very similar to that found in the origins of nanoviruses (Bell *et al.*, 2002), circoviruses, and prokaryotic RC systems (Baas, 1987). The stem-loop structure in all geminivirus IRs is essential for replication; however, the sequence of the stem is not so important as long as the structure is maintained (Orozco and Hanley-Bowdoin, 1996; Willment, 1999). The sequence does however contribute to the efficiency of replication, as stem variants replicate less efficiently than wild type structures (Orozco and Hanley-Bowdoin, 1996; Schnippenkoetter *et al.*, 2001). The loop can also tolerate some mutations (Schneider *et al.*, 1992; Stanley, 1995); however the loop's nonanucleotide sequence is less flexible: insertions and deletions are not tolerated, although in MSV a mutation of TAATATTAC to TAATACTAC resulted in only a slightly less fit virus (Schneider *et al.*, 1992). Apart from the one notable exception of WDV, the stem-loop structure appears to be required for both initiation and termination of replication. Surprisingly, a WDV mutant with a deletion of the entire stem-loop sequence has been shown to initiate replication. There was, however, a defect in termination of replication, resulting in high molecular weight concatemeric forms of viral DNA (Kammann *et al.*, 1991). Heyraud *et al.* (1993b) subsequently found that the mutant virus was able to initiate replication from a second site (TACCC) resembling the nicking site in the stem-loop. However, this second initiation site is apparently unique to WDV, and it is likely that the stem-loop is required for replication initiation and termination by all other geminiviruses.

In order to understand the features of the mastrevirus LIR, it is helpful to first summarise the more-intensively studied IR of begomoviruses, also known in bipartite begomoviruses as the common region, or CR.

The begomovirus plus strand origin was first mapped by Lazarowitz *et al.* in 1992, and then further defined by Orozco *et al.* (1998), to an 89-bp sequence in the CR. This fragment, which is conserved between the A and B components of begomoviruses, includes the stem-loop sequence and an adjacent 60 nt located on the left side (5') of the CR (Fig. 1.2). Around the same time, Fontes *et al.* (1992 and 1994a) showed that TGMV Rep binds specifically to a 13-bp directly repeated motif (5'-GGTAGTAAGGTAG) that is essential for replication. This site, located 34 bp upstream of the hairpin and between the TATA box and transcription start site of the *Rep* promoter, also mediates transcriptional repression of the *Rep* gene (Eagle *et al.*, 1994), possibly through interference with assembly or activity of the transcription pre-initiation complex. In this way, TGMV Rep negatively regulates its own expression. Related virus-specific Rep-binding

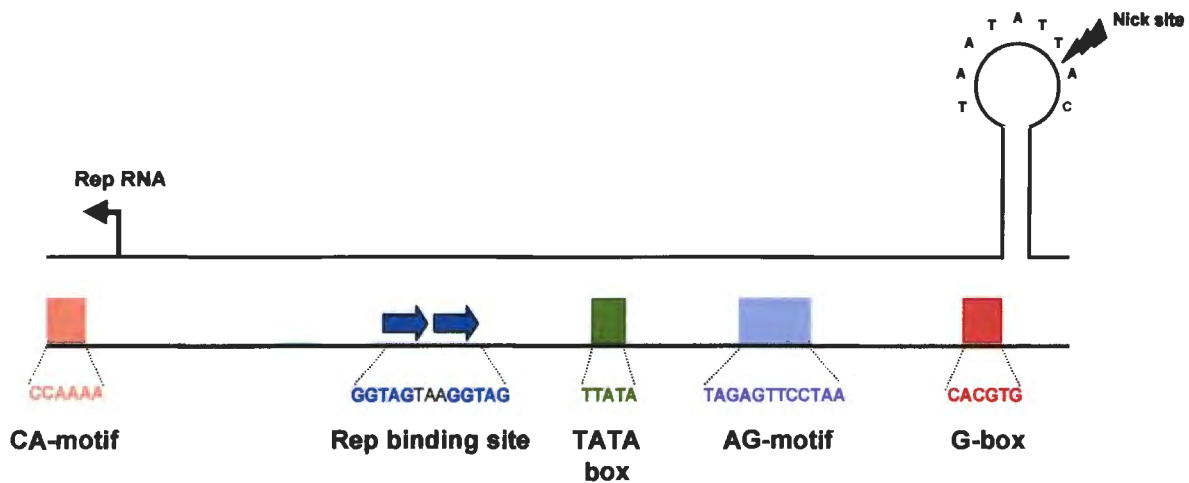


FIGURE 1.2 A schematic representation of the six known functional elements of the *Begomovirus* plus strand origin of replication. These are: (1) a conserved nonanucleotide motif in the loop of a stem-loop structure, which contains the replication initiation site; (2) a G box 5' of the stem which binds host transcription factors; (3) an AG motif between the G-box and (4) the Rep TATA box, which may bind host factors to facilitate initiation of replication; (5) a high affinity Rep-binding site between the *Rep* TATA box and transcription start site, which confers replication specificity; and (6) a CA motif upstream of the Rep binding site, which contributes to efficient replication possibly by binding host factors. The two repeats within the Rep-binding site have different functions, the 3' repeat being essential while the 5' repeat enhances Rep binding (Fontes *et al.*, 1994a).

sites are found in the IRs of all other begomoviruses (Arguello-Astorga *et al.*, 1994a; Fontes *et al.*, 1994b) and curtoviruses (Choi and Stenger, 1996), although the distance between the Rep-binding site and the stem-loop in different viruses is varies from 23 to 82 bp (Palmer and Rybicki, 1998). These Rep binding sites act as origin recognition elements to allow virus-specific replication.

Further analysis of the IR sequences from different begomoviruses (Arguello-Astorga *et al.*, 1994a and 1994b) revealed a series of repeated sequence elements of 8 - 12 nt, called iterons, which have sequence homology to the Rep-binding site identified by Fontes *et al.* (1994a, b). The nucleotide sequence of the iterons is virus specific, but the organization (number, orientation and spacing) is conserved within the CR of dicot-infecting begomoviruses. Based on these observations, Arguello-Astorga *et al.* (1994a) proposed that Rep binds specifically to these sites, and that they have a role in viral replication and/ or transcription.

The distance of the Rep-binding sites from the replication initiation site in the hairpin loop raises the question of how Rep catalyzes cleavage in the loop while bound to its distantly located binding sites. In their computer analysis of the begomovirus IR, Arguello-Astorga *et al.* (1994a)

identified a G-box-like site adjacent to the 5' end of the stem-loop (Fig. 1.2). G-boxes are known to be transcriptional *cis*-regulatory element in plants. Taking into account that the iterons are closely associated with the Rep TATA box, the authors proposed a model by which initiation of replication can occur: A host transcription factor bound at the G-box adjacent to the stem-loop interacts with the TATA-binding protein (TBP), resulting in sequence looping, bringing the Rep complex bound at the TATA-proximal iterons in contact with the cleavage site of the loop to initiate RCR. This process may be assisted by the replication enhancer protein, REn, which has been shown to interact nonspecifically with Rep (Settlage *et al.*, 1996) and possibly to increase the affinity of Rep for the origin (Fontes *et al.*, 1994a; Gladfelter *et al.*, 1997). One proposal is that REn bound to the stem-loop could recruit Rep to the nonanucleotide cleavage site (Hanley-Bowdoin *et al.*, 1996). However, there is no evidence for specific REn interaction with the stem-loop, and as yet there is only speculation on the mechanism by which the protein enhances replication of begomoviruses and curtoviruses.

In addition to the G-box, two other key elements have been identified in the begomoviral IR (Fig. 1.2). One element, the AG motif, is between the G-box and TATA box, and is essential for origin function (Orozco *et al.*, 1998). The second element, a CA motif, is located outside of the minimal origin immediately upstream of the Rep binding site, and may act as an efficiency element (Orozco *et al.*, 1998).

The six known functional elements in the begomovirus IR contributing to origin function are illustrated in Fig. 1.2. The fact that these elements are closely spaced, and that changes in spacing affects origin activity (Orozco *et al.*, 1998), suggests that they interact with each other during initiation of replication, most likely through the proteins that bind to them (Hanley-Bowdoin *et al.*, 1999).

The IR of mastreviruses differs in several ways from that of begomoviruses. Whereas the stem sequence of the IR stem-loop is highly conserved between different begomoviruses, there is low sequence homology between the LIR stem sequences of different mastreviruses (Rybicki, 1994; Padidam *et al.*, 1995). This suggested to Arguello-Astorga *et al.* (1994a) that the mastrevirus Rep binding sites could reside in the stem; and indeed, sequence analysis of eight mastrevirus LIRs showed that a GC-rich sequence in the stem was iterated elsewhere in the LIR (Fig 1.3 [WDV] and 1.4A [MSV]). The positions of these iterons (one between the TATA box of the Rep gene and its transcription start point, and two on either side of the stem) were conserved in all the

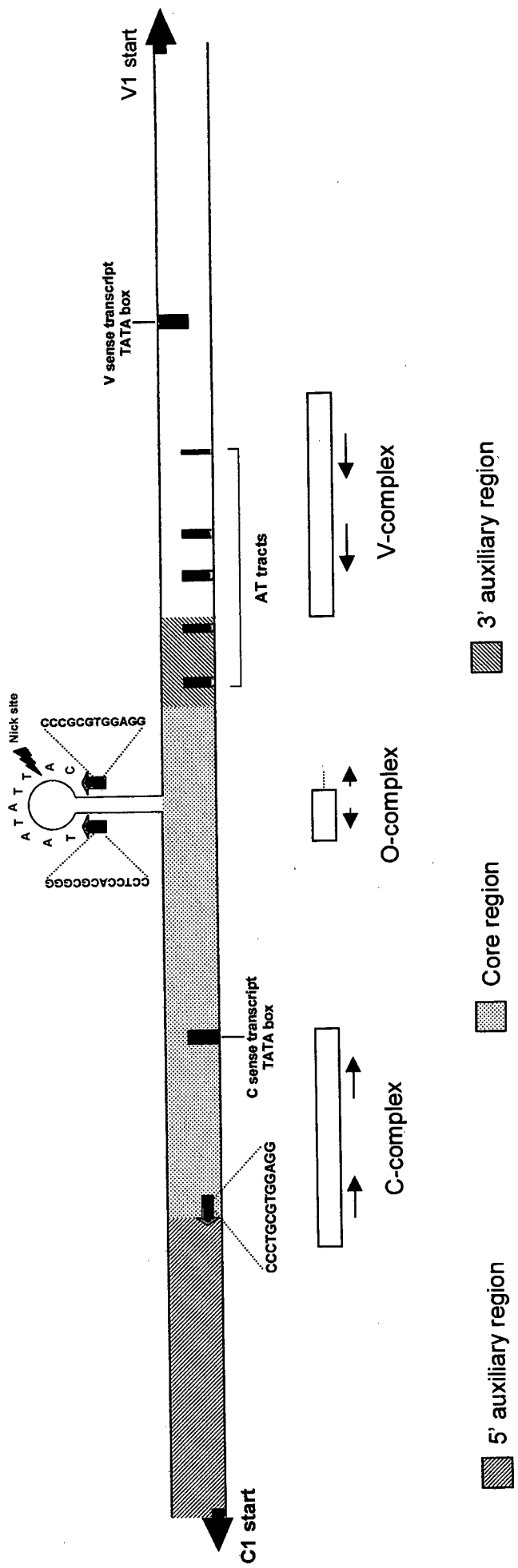


FIGURE 1.3 The main features of the *Wheat dwarf virus* (WDV) long intergenic region (LIR). These include (1) a stem-loop structure with an invariant nonanucleotide sequence where plus strand replication is initiated by an endonucleolytic nick introduced by Rep; (2) the TATA boxes of the two promoter elements controlling C and V sense transcription, 5' and 3' of the initiation site respectively; (3) the start of the C1 (5') and V1 (3') ORFs; (4) iterons represented by blue arrows, with the iterated sequences indicated in the V sense; (5) *cis* elements which include 5' and 3' auxiliary sequences (diagonal stripes) flanking a core sequence element (blue dots) essential for origin function. The core sequence includes the stem-loop structure and invariant nonanucleotide sequence. Also shown are the Rep binding sites defined by mutational, DNase I and footprinting studies. The C and V complexes have high affinity, and the O complex low affinity, footprints. The black (C and V complex) and grey (O complex) arrows indicate G+T-rich direct repeats that may be Rep recognition sequences. Because of their positions within the LIR, it is thought that WDV C and V complexes may be involved in regulation of C and V sense transcription respectively, while the role of Rep bound at the O complex, which is capable of introducing an endonucleolytic nick in the invariant nonanucleotide sequence, may be to initiate plus strand replication

mastreviruses the authors examined. By analogy with the iterated sequences observed in begomoviruses and curtoviruses, the authors proposed that the *Mastrevirus* iterons are also Rep recognition sequences. However, despite some results by Castellano *et al.* (1999; see later in this section) supporting this hypothesis, as yet a direct, specific interaction between iterons and Rep has not been experimentally demonstrated. There is also some evidence that, contrary to the prediction of Arguello-Astorga *et al.* (1994a), in MSV the stem-loop is not a specificity determinant, although a wild type stem sequence does enhance replication (Willment, 1999), a situation similar to that in begomoviruses.

The best-characterised mastrevirus LIR is that of WDV (Suarez-Lopez *et al.*, 1995; Suarez-Lopez and Gutierrez, 1997; Sanz-Burgos and Gutierrez, 1998; Castellano *et al.*, 1999, Missich *et al.*, 2000), the features of which are summarized below and illustrated in Fig. 1.3. Although the WDV LIR differs from that of MSV, some parallels may be drawn in an attempt to fully describe the MSV LIR.

A *cis* element unique to the mastrevirus LIR is a potential static DNA curvature first discovered in WDV (Suarez *et al.* 1995), produced by an 80-bp cluster of "A-T tracts" and located between the stem-loop and MP start codon (Fig 1.3). Since bent DNA sequences have been shown to be necessary for origin activity in a number of systems, the authors hypothesized that the bending locus could be a regulatory element of WDV replication. However, it was later shown that the DNA-bending locus has only a minimal impact on the replication of WDV (Suarez-Lopez and Gutierrez, 1997). It is more likely that bending of this region is involved in regulation of transcription of V sense genes (Castellano *et al.*, 1999).

The minimal origin of WDV replication (a ~200 bp core) was defined by Sanz-Burgos and Gutierrez (1998) as spanning a region ~170 and 28 bp upstream and downstream, respectively, from the initiation site in the stem-loop. This minimal *cis*-acting element, which includes the iterons identified by Arguello-Astorga *et al.* (1994b), is flanked by two auxiliary elements (5' aux and 3' aux) that enhance WDV replication. While the features of the 5' aux region are not known, the 3' aux region contains part of the AT-rich sequence conferring a static DNA curvature of the LIR (Fig. 1.3). Interestingly, visualization of WDV Rep/DNA complexes in the minimal *cis*-acting region by electron microscopy revealed a high affinity Rep-binding site located upstream of the stem-loop, between the C sense TATA box and transcription start site. This location is very similar to the high affinity Rep-binding site of TGMV identified by Fontes *et al.* (1994b), and

that of BCTV (Choi and Stenger, 1996), suggesting that the general configuration of plus strand origins of all geminivirus genera is similar (Hanley-Bowdoin *et al.*, 1999). Moreover, the binding of Rep in the proximity of the TATA box of the C sense promoter suggests that, like *Begomovirus* Reps, the Reps of mastreviruses may regulate their own expression (Sanz-Burgos and Gutierrez, 1998).

WDV Rep binding sites within the LIR were mapped further by Castellano *et al.* (1999). These authors used electron microscopic visualization and DNase I footprinting to identify three Rep-DNA complexes, which they named C, V and O (Fig. 1.3). The C complex lies between the TATA box for C sense transcription and the transcription start site (hence the name "C complex"), which is in agreement with the data of Sanz-Burgos and Gutierrez (1998), and similar to the location of TGMV Rep binding sites. The V complex, so called because it is located 52 bp upstream from the TATA box for V sense transcription, interestingly coincides with some of the A-T tracts conferring a static DNA curvature on this region identified by Suarez-Lopez *et al.*, (1995). Since WDV Rep (and/or RepA) has been implicated in up-regulation of V sense transcription (Hofer *et al.*, 1992; Collin *et al.*, 1996), and DNA curvature is known to play a role in transcriptional regulation, it is likely that the V complex is involved in positive transcriptional regulation of the V sense genes. This would partly explain the unique architecture of the mastrevirus LIR (e.g. bending of the LIR DNA and unique Rep binding sites), since in the other geminivirus genera Rep is not directly involved in V sense transcription. Conversely, the role of the C complex may be to down-regulate the expression of Rep; the location of the complex (encompassing both the C sense TATA box and transcription start site) is consistent with this theory. The C complex is also highly likely to have a role in DNA replication, as observed by Sanz-Burgos and Gutierrez (1998), considering its location on the 5' side of the minimal *cis*-acting core sequence required for WDV replication

Microscopic visualization by Castellano *et al.* (1999) of the C and V complexes revealed large spherical nucleoprotein structures, suggesting that the complexes consist of Rep oligomers. This is consistent with the observation that TGMV Rep oligomerizes *in vitro* to form octomeric complexes (Settlage *et al.*, 1996; Orozco *et al.*, 1997) and that MSV Rep monomers interact with one another in yeast (Horvath *et al.*, 1998).

WDV RepA also forms DNA-protein complexes (RepA C and V complexes) in a location similar to those of the Rep-DNA complexes, but with distinct DNaseI footprints (Missich *et al.*, 2000).

The Rep and RepA C complexes are approximately the same size, but the RepA V complex footprint is 10 nt smaller than that of the Rep V complex. In addition, Rep-RepA hetero-oligomers may be involved in complex formation, which is consistent with their dual and possibly co-operative roles in viral replication and transcriptional activation of the V sense promoter. The effect of Rep/RepA homo- and hetero-oligomerization on replication and transcription is discussed in the section on the C sense genes.

Whereas the C and V complexes are high-affinity complexes, Rep interacts with low affinity with WDV DNA in the region of the stem-loop, to form an O complex that is capable of carrying out the cleavage reaction necessary for the initiation of RCR. This experimental result supports the hypothesis of Arguello-Astorga *et al.* (1994a) that the stem-loop of mastreviruses contains specific Rep binding sites. However, although the footprints of the Rep and RepA C complexes also cover an iteron (proximal to the Rep TATA box, i.e. the Rep proximal iteron), the V complex does not encompass any sequence-predicted iterons. Interestingly, the footprints of both C and V complexes cover a G+T-rich repeated sequence, GTGTGAN₂₂₋₂₃GTG(G)TC that may be the actual Rep recognition sequence (Castellano *et al.*, 1999). A similar, although non-repeated, sequence occurs in the stem (GTGG(T)GG); the fact that it consists of half of the C- and V-complex consensus sequence may be the reason for the low affinity of the O complex (Castellano *et al.*, 1999).

The fact that analogous G+T-rich Rep recognition sequences have not been found in the MSV LIR makes it difficult to predict if similar MSV LIR-Rep complexes form. However, Willment (1999) mapped replication specificity determinants (RSD), which presumably include specific Rep binding sites, to a region spanning the Rep proximal iteron and the stem loop (Fig 1.4A). Interestingly, the major RSD corresponds to the region in WDV covered by the C complex, while a minor RSD is in the vicinity of the O complex, perhaps providing indirect evidence for the existence of these complexes in MSV. Another similarity between the WDV and MSV LIRs appears to be the minimal sequence required for replication, which Willment (1999) found to include sequences 5' of the stem-loop up to and including the Rep proximal iteron, and a 25 bp region 3' of the stem-loop, which is very similar to the WDV minimal *cis*-acting core sequence required for replication delineated by Sanz-Burgos and Gutierrez (1998). However, the essential MSV sequence 3' of the stem-loop does not include the AT tracts that in WDV confer DNA bending. The region 3' of the stem-loop essential for replication appears to be peculiar to mastreviruses, since in begomoviruses only the elements on the complementary side of the stem-

loop are required for replication.

An element found in the MSV LIR that is conserved throughout all the geminivirus genera, is a GC-box (directly repeated in MSV) at the 5' base of the stem (Fig. 1.4A). This binds maize nuclear factors and comprises part of the V sense gene promoter (Fenoll *et al.*, 1990). While the V sense promoter core spans the start of the CP gene through to the 3' side of the stem-loop, optimal expression of the CP requires a region 5' of the stem-loop (530 nt upstream of the CP start codon) called the upstream activator sequence (UAS; Fenoll *et al.*, 1988, 1990). Within the UAS, the region containing the GC-boxes, called the rightward promoter element (*rpe1*), is required for efficient replication as well as transcription of V sense genes, probably by recruiting nuclear factors to the region (Arguello-Astorga *et al.*, 1994b). Transcription of V sense genes is probably directed by two TATA boxes 26 and 214 nt upstream from the MP start codon (Wright *et al.*, 1997). Two 3' co-terminal bicistronic transcripts are produced (one large and one small) from which both CP and MP can be expressed, although CP is more efficiently expressed from the smaller, more abundant, transcript.

While begomoviruses encode a protein – TrAP - that transactivates the CP promoter (Sunter and Bisaro, 1991; 1992; 1997), in mastreviruses it is thought that Rep and RepA act together to activate the CP promoter (Zhan *et al.*, 1993; Collin *et al.*, 1996; Mazithulela *et al.*, 2000; McGivern, 2002). There is far more known about TrAP than the transactivational properties of mastrevirus Rep/RepA. TrAp is a zinc binding protein with an acidic transcriptional activation domain at its carboxy terminus (Hartitz *et al.*, 1999), which activates the CP promoter in mesophyll cells, but acts to de-repress the CP promoter in phloem tissue, and probably interacts with cellular proteins to recognize its target promoters (Sunter and Bisaro, 1997). There is also evidence for tissue specificity of the MSV V sense promoters (Mazithulela *et al.*, 2000; Gooding *et al.*, 1999) but it is unknown exactly what role Rep and/or RepA have to play in conferring the specificity. In MSV, further control of V sense transcription is conferred by the presence of an intron within the MP gene that prevents the production of MP from spliced transcripts, whereas the CP can be expressed from both spliced and unspliced transcripts. It is thought that the intron may enhance expression of the CP gene by an intron-mediated enhancement mechanism, such as that conferred by introns in cereal transgene expression cassettes (Wright *et al.*, 1997). Thus, the ratio of CP to MP is controlled in a number of ways: first, the size of the V sense transcripts determines which gene is more efficiently expressed (CP being expressed from the more abundant shorter transcript); second, splicing of an intron in the V sense transcripts enhances the

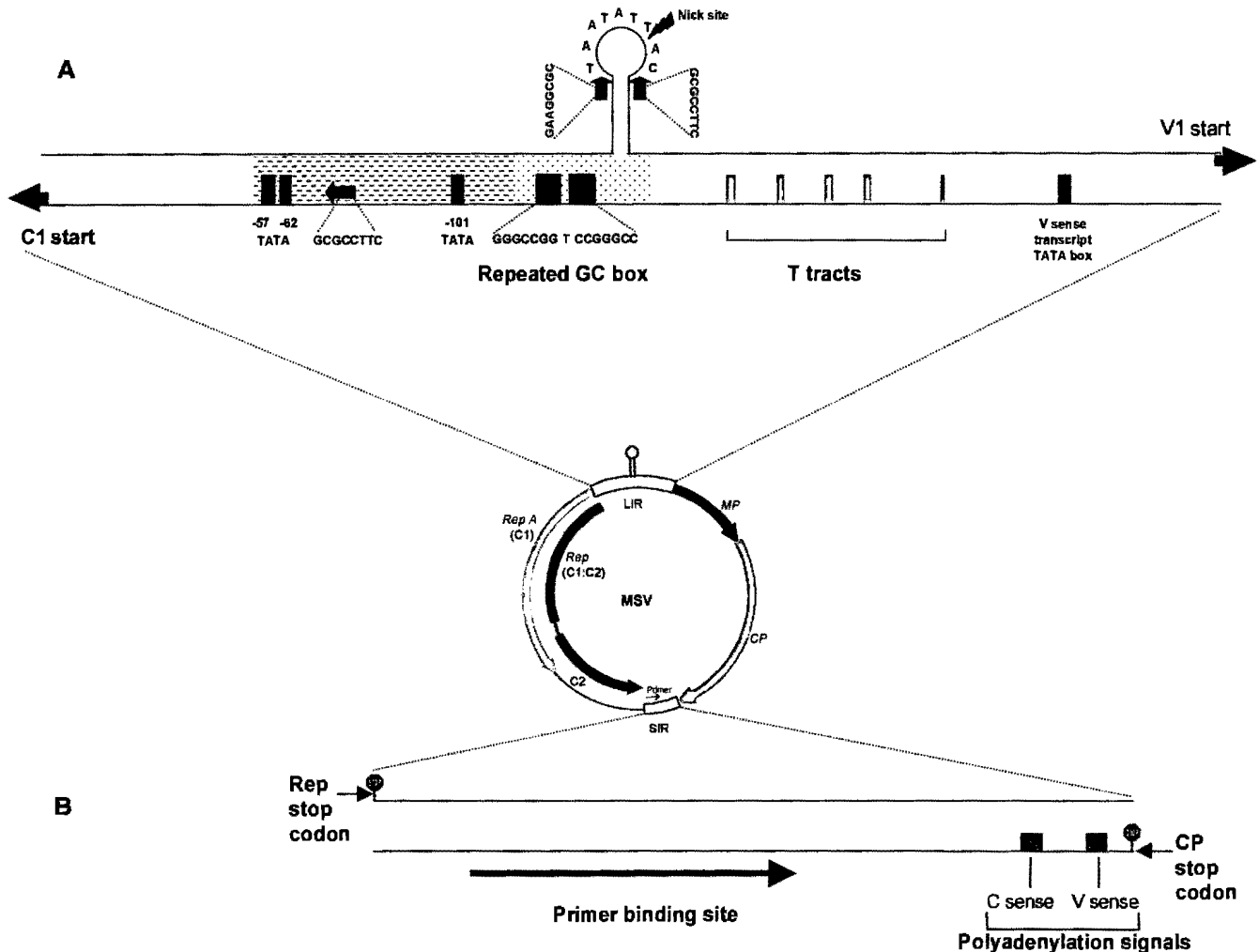


FIGURE 1.4 A schematic representation of the MSV long intergenic region (A) and short intergenic region (B), shown in context with the MSV genome. In (A) the main features of the MSV LIR are shown. These include a stem-loop structure with the loop's nonanucleotide sequence conserved amongst all geminiviruses and other rolling circle systems. The site at which Rep introduces an endonucleolytic nick to initiate plus strand replication is shown. Iterated sequences (iterons) are shown in the V sense, with blue arrows indicating their location in the LIR. Iterons are potentially specific Rep-recognition sequences via which Rep may bind to the LIR. 5' of the stem-loop is a repeated GC-box, which binds host transcription factors, and three TATA boxes from which C sense transcription can potentially be initiated. A series of T tracts 3' of the stem-loop may be involved in DNA bending of this region of the LIR. Also shown 3' is the TATA box for V sense gene transcription. Additional features that may be present in the MSV LIR, but are not shown in (A) include C, V and O complexes that were identified in the *Wheat dwarf virus* (WDV) LIR. These Rep-DNA complexes have not been identified in the MSV LIR. Because of differences in size between the MSV and WDV LIRs, it is difficult to infer analogous positions of Rep binding within the MSV LIR. However, they may correspond to a major replication specificity determinant (major RSD), highlighted in (A) with blue dashes (= C complex?) and a minor RSD, shaded with blue dots (= O complex?) identified by Willment (1999). The minor RSD also includes the stem-loop structure. In (B), the main features of the SIR include polyadenylation signals for V and C sense transcripts, and a primer binding site on the plus strand. A ~80 bp DNA primer-like molecule, encapsidated with the viral genome and annealed to this site, is thought to be involved in initiating negative strand replication. Both the MSV LIR and SIR are essential for viral replication.

production of CP; and third, Rep and RepA appear to transactivate the CP, but not MP, promoter (McGivern, 2002). These areas of control result in CP generally being present at much higher concentrations than MP in infected cells (Mullineaux *et al.*, 1988). The transactivation properties of Rep and RepA are further explored in the section on the C sense genes.

An important difference between the MSV and WDV LIRs is the presence in MSV of three TATA boxes from which C sense transcription can potentially be initiated (Fig 1.4A), located at positions -101, -62 and -57 (relative to the *Rep* ATG): in WDV it can only be initiated from two overlapping TATA boxes at positions -131 and -129. Transcripts initiated at different TATA boxes in MSV could represent one of a few ways in which the relative levels of Rep and RepA are controlled. For example, transcripts initiated from the -62 and -57 TATA boxes are predominantly of 1.2 kb (terminating in the C2), whereas predominantly 1.5 kb transcripts (terminating in the SIR) are initiated from the -101 TATA box (M.L. Boulton, pers. comm). Only RepA can be expressed from the shorter transcript, whereas both Rep and RepA are capable of being expressed from the 1.5 kb transcript. This represents another possible level of control of Rep/RepA expression, whereby splicing of an intron in the larger transcript is necessary to produce Rep, with RepA being expressed from the unspliced 1.5 kb transcript. Thus, modulation of splicing is an obvious way of regulating Rep and RepA expression. Whereas control at the level of splicing can occur in WDV, transcriptional control of Rep/RepA expression may only be possible in MSV. For example, if a Rep C complex does occur in MSV, it is conceivable that binding of Rep oligomers near the -101 TATA box could force C sense transcription from the -57 or -62 TATA boxes, simultaneously enhancing the expression of RepA and inhibiting the expression of Rep. In WDV, however, the C complex encompassing the overlapping -131 and -129 TATA boxes is likely to prevent expression of both Rep and RepA.

Finally, there are differences between the MSV and WDV LIRs in terms of size. The distance between the replication initiation site in the stem-loop and the Rep start codon is larger in WDV than in MSV, as is the distance between the respective Rep proximal iterons and TATA boxes, and the distance of the TATA boxes from the replication initiation sites in each virus. There is also the unique second replication initiation site in the WDV LIR that sets it apart from other mastreviruses. These differences are reflected in the properties of the viral proteins that interact with the WDV and MSV LIRs, as will be seen in the section on the C sense genes, and in the results presented in this thesis. Nevertheless, it is useful to draw analogies from knowledge of the

WDV LIR in order to better understand the interactions of viral and host proteins with the lesser-known MSV LIR. As will be seen throughout this thesis, this is important in determining the multiple functions of the MSV C sense genes.

1.2.2.2 The short intergenic region

The *Mastrevirus* negative (or complementary) strand origin of replication, located in the SIR (Fig 1.4B) and activated by a priming event, is involved at a very early stage of viral DNA replication. Both C sense gene expression and the conversion of plus strand DNA into dsDNA intermediates require activation of the negative strand origin of replication. Essential to this event is a ~ 80bp DNA primer that binds to the SIR plus strand (Fig 1.4B), and is encapsidated with the viral genome (Donson *et al.*, 1984; Hayes *et al.*, 1988; Kammann *et al.*, 1991). In MSV the primer has ribonucleotides linked to its 5' end, suggesting that the DNA is primed from a longer RNA primer molecule (Palmer and Rybicki, 1998). Once virus particles containing plus strand DNA have entered a cell and uncoated, the primer, conveniently already annealed to the plus strand, may initiate synthesis of RF dsDNA, completing the first stage of the viral replication process.

While both the SIR and LIR are required by mastreviruses for efficient genome amplification, in the other geminivirus genera the IR contains the *cis*-acting signals required for both negative and positive strand replication initiation, and therefore constructs containing only the IR support efficient begomoviral replication (Lazarowitz *et al.*, 1992). Although negative strand DNA replication was shown to be RNA-primed in ACMV (Saunders *et al.*, 1992), begomovirus and curtovirus genomes have no virion-associated DNA primer. The origin of the mastrevirus ~80 nt primer-like molecule is unknown, as are the *cis* elements regulating negative sense DNA replication.

Apart from the fact that in addition to its role in replication, the SIR also contains polyadenylation and termination signals of the V and C sense transcripts (Fig 1.4B), little else is known about the shorter of the two mastrevirus IRs. Further elucidation will require more thorough investigation, for example into the minimal sequences required to activate the negative strand origin both within the SIR of mastreviruses and the IR of curtoviruses and begomoviruses, as well as the SIR's potential role in control of mastrevirus gene expression.

1.2.2.3. The complementary sense genes (*Rep* and *RepA*)

One of the earliest steps in the MSV life cycle after passage of the viral genome into the host cell nucleus is the host-directed, DNA-primed synthesis of a complementary (minus) strand, using the virion (plus) strand as the template. This event creates dsDNA, which serves as a template for transcription of viral genes and for RCR. These processes must be tightly controlled, since they cannot occur simultaneously: first the *Rep* gene must be expressed so that *Rep* can initiate replication of the genome to a high enough titer to initiate a systemic infection, and only then must the V sense genes be expressed to any significant level. Inappropriate expression of CP and MP would be likely to interfere with viral replication, for example by sequestering ssDNA or by moving the virus genome out of the nucleus. The role of the C sense genes in viral replication, transcription and in regulation of the viral life cycle is covered extensively in this section. First, the biochemical structure (and the functions ascribed to various motifs) of the MSV C sense gene products is discussed. Taken together, this knowledge is used to describe a model of viral replication, for which *Rep* is the only indispensable viral protein.

There are potentially two transcripts for the MSV C1 and C2 ORFs: these are one of 1.5 kb that results in translation of full-length *Rep* if spliced, or *RepA* if unspliced; and one of 1.2 kb that terminates in the C2 and is capable of expressing only *RepA*. It must be noted that expression of *RepA* has not been proven *in vivo*; however there is enough evidence of numerous important roles for *RepA* in the MSV life cycle that in this thesis it is assumed that *RepA* is an authentic protein.

The MSV C2 ORF, which encodes the carboxy terminus of *Rep*, is fused with the C1 ORF (encoding the *Rep* amino terminus) by the splicing of an intron in the C1:C2 transcript. Despite the presence of an ATG start codon in the MSV C2 ORF, it is not considered to encode an autonomous "RepB" protein, partly because there is no detectable transcript for such a protein, and because all other mastreviruses lack a C2 start codon.

As can be seen in Fig. 1.5, *Rep* and *RepA* share the same amino terminal 214 amino acids, but differ in their C termini. Although both *Rep* and *RepA* appear to be required to activate the promoters of the V sense genes (McGivern, 2002), *RepA* is unnecessary for viral replication (Schalk *et al.*, 1989; Collin *et al.*, 1996; results in this thesis). However, as will be seen in this section, *RepA* does appear to play an important role at various stages of the MSV replicative

cycle. These functions, separate and different from those of Rep, are reflected in each protein's unique C terminus.

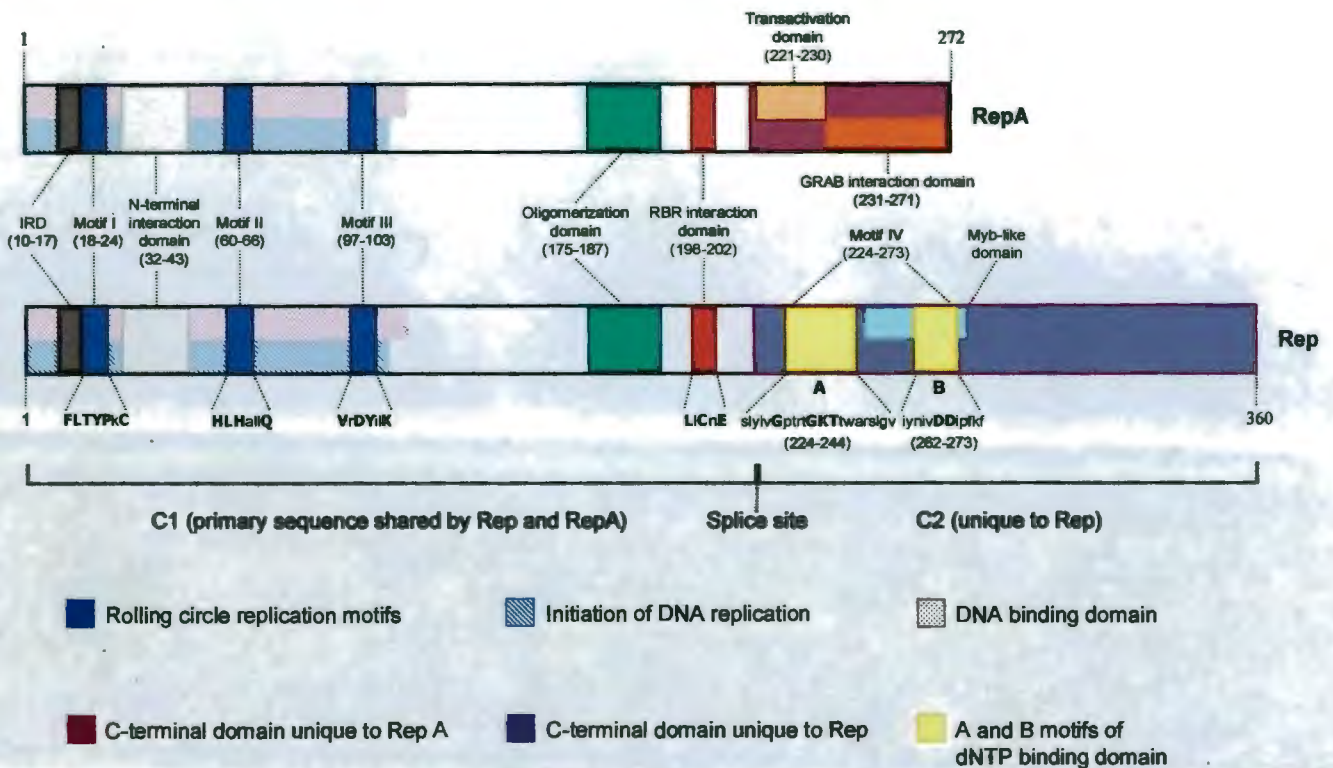


FIGURE 1.5 Predicted functional domains of MSV Rep and RepA proteins (based on Gutierrez, 2000 and Boulton, 2002). The sequences of motifs I, II, III, IV and the RBR protein-interaction domain are shown, with each motif's conserved residues indicated in bold uppercase. The unconserved amino acids (letters in lowercase) are represented by 'x' in the consensus sequences for each motif described in the text. Note that in MSV, the DNA-binding domain and GRAB-binding domain have not been identified, and are shown here in analogous positions to the Rep DNA-binding domain in begomoviruses and curtoviruses, and the GRAB-binding domain of WDV Rep. The N-terminal interaction domain of RepA, although unable to mediate self-association, can interact with the oligomerization domain in the RepA C terminus (Horvath *et al.*, 1998). A RepA fragment containing the N-terminal interaction domain but lacking the oligomerization domain can also interact with full-length Rep through the oligomerization domain of Rep. The biological role of the interaction between the N-terminal and oligomerization domains of both proteins is not clear (Horvath *et al.*, 1998). The details and function/s of each protein domain are discussed in the text

The N terminus of Rep and RepA contains three conserved amino acid sequences (Fig. 1.5) that are related to motifs found in initiator proteins of other RC systems. Although deletion and mutation of each of motifs I, II and III blocks the ability of Rep to cleave and replicate viral DNA, the precise roles of motifs I and II are not known. A tyrosine residue in motif III (VxDYxxK) participates in phosphodiester bond cleavage at the V sense origin of replication (the

loop in the MSV LIR stem-loop structure), and in the covalent linkage of Rep to the 5' terminus exposed by nicking (Laufs *et al.*, 1995b). The nicking-joining activity of geminivirus Reps is the main function of all replication initiator proteins from RC systems. The point of Rep-mediated nicking in the loop is between nucleotides 7 and 8 of the conserved nonanucleotide sequence: TAATATT^{8↓}AC (Heyraud-Nitschke *et al.*, 1995; Laufs *et al.*, 1995a; Stanley, 1995). Interestingly, although Rep binds to dsDNA to initiate replication, Laufs *et al.* (1995a) showed that the protein was unable to cleave a dsDNA origin *in vitro*. Although the authors state that there may have been problems in their experimental protocol, this result is corroborated by the finding of Orozco and Hanley-Bowdoin (1996) that formation of a stem-loop structure is required for viral replication. Presumably, transient melting to allow extrusion of a stem-loop structure would produce a ssDNA cleavage substrate in RF DNA.

Motif II (HLHxxxQ) may be involved in metal ion coordination through the histidine residues (Koonin and Ilyina, 1992); the finding that the covalent linkage of Rep to the 5' terminus of the nicked nonanucleotide motif *in vitro* requires Mg²⁺ or Mn²⁺ ions (Laufs, 1995a) supports this hypothesis.

Although no precise function has been ascribed to the conserved FLTYPx_nC signature of motif I, Arguello-Astorga and Ruiz-Medrano (2001) presented evidence that the region containing motif I may be involved in Rep recognition of iterons in the IR of geminiviruses. The most significant feature of this motif I-associated subdomain, which the authors called an iteron-related domain (IRD; see Fig.1.5), is that its primary structure differs among viruses harboring distinct iterons, while it is generally similar among viruses with identical iterons regardless of their differences in host range, insect vector, geographical origin or genome structure. The authors suggest that, together with motif I, the IRD may form the core of a DNA-binding domain whose secondary structure is apparently conserved in the replication proteins of nanoviruses, circoviruses, microviruses and archaeobacterial and eubacterial ssDNA plasmids. Indirect evidence for the IRD being a major component of the specific DNA recognition domain of geminivirus Rep, comes from the fact that the IRD is located within the Rep region where trans-acting replication specificity determinants have been mapped (see references within Arguello-Astorga and Ruiz-Medrano, 2001), and that deletion or mutation of the IRD-Motif I region of TGMV Rep eliminates its specific DNA-binding capability (Orozco *et al.*, 1997; 1998). More recently, the NMR-derived structure of the catalytic domain of TYLCV Rep has added structural evidence for this theory (Campos-Olivas *et al.*, 2002).

A fourth motif (EGX₄GKTX₃₂DD), conserved in the C2 of all geminivirus Reps (but absent from RepA; see Fig. 1.5) is a NTP-binding domain with typical A and B motifs that are found in proteins with kinase and helicase activities (Gorbalenya *et al.*, 1989). This motif exhibits ATPase activity that is required for replication, since mutations altering the lysine residue in the P loop of the TYLCV Rep NTP-binding domain impaired ATP hydrolysis *in vitro* and replication *in vivo* (Desbiez *et al.*, 1995). However, neither ATP binding nor hydrolysis is required for nicking and joining of ssDNA at the plus strand origin (Heyraud-Nitschke *et al.*, 1995), or for Rep-mediated transcriptional repression (Eagle *et al.*, 1994). Although a definite function has not been attributed to the NTP-binding domain, it has been speculated that it could enable Rep to act as a helicase that would unwind and displace positive strand DNA from the negative strand template in advance of the replication fork, or unwind and expose the origin to proteins of the replication apparatus (Bisaro, 1996). *In vitro* evidence for a topoisomerase function of Rep was provided by Pant *et al.* (2001). However, because the ATPase activity is DNA-independent (Desbiez *et al.*, 1995), it is still not certain whether its role *in vivo* is to contribute to the putative helicase activity of Rep or another, as yet undetermined ATPase-dependent activity of Rep.

As inferred when discussing motif I, the DNA recognition and binding domain resides in the N-terminus of Rep. While the specific location of this domain has not been mapped in MSV, in begomoviruses and curtoviruses the N-terminal 116 and 89 amino acids respectively are required for Rep to recognize specific DNA-binding sequences in its cognate origin. As mentioned in the section on the *Mastrevirus* LIR, MSV Rep and RepA also bind the origin DNA in a sequence-specific manner, and it is likely that the DNA-binding domain also resides in the N-terminus of both proteins (Fig. 1.5).

In addition to its catalytic and DNA binding activities, mastrevirus Rep is involved in several protein interactions, including homo-oligomerization, binding to RepA, and interaction with host proteins. These interactions and their established or putative roles in the mastrevirus life cycle are described below.

The various activities associated with geminivirus Reps appear to be determined in part by the aggregation state of Rep (Orozco *et al.*, 2000) and in the case of mastreviruses, Rep and/or RepA (Horvath *et al.*, 1998; Missich *et al.*, 2000). Studies have shown that TGMV Rep can form oligomers in solution (Settlage *et al.*, 1996; Orozco *et al.*, 1997), that MSV Rep monomers self-interact in yeast (Horvath *et al.*, 1998), and WDV Rep oligomers consisting of six to eight

monomers have been visualized bound to DNA (Sanz-Burgos and Gutierrez, 1998; Castellano *et al.*, 1999). Although Rep monomers can perform the cleavage and joining reactions *in vitro*, studies with TGMV have suggested that Rep-Rep interaction is a prerequisite for DNA binding (Orozco *et al.*, 1998). Furthermore, in a study of TGMV mutants defective for oligomerization, Orozco *et al.* (2000) discovered a clear correlation between the aggregation state of Rep and its function in replication and transcription. Mutations abolishing or impairing Rep-Rep interaction had the effect of inhibiting or severely impairing TGMV replication, at the same time enhancing Rep-mediated transcriptional repression of the C sense promoter. Interestingly, there are also examples of point mutations in papillomavirus E1 and E2 proteins that differentially affect replication and transcription (Cooper *et al.*, 1999). Replication initiation factors generally function as large protein complexes, whereas transcription factors frequently act as dimers or tetramers (Orozco *et al.*, 2000). Thus, it is conceivable that in geminiviruses different Rep complexes may be required for the two activities. Alternatively, mutations in the Rep oligomerization domain may result in a conformational change that favours repression complexes over replication complexes. For example, a conformational change might make a region that contacts the transcription apparatus more accessible, thus facilitating active repression (Eagle and Hanley-Bowdoin; 1997).

TGMV Rep also interacts with the viral replication accessory factor REn, which enhances viral DNA accumulation. REn can self-interact and oligomerize with Rep in a non virus-specific manner (Settlage *et al.*, 1996). Since the REn protein sequence does not appear to contain any homology to known enzymatic motifs, it is thought that the structure of the Rep/REn complex, rather than a catalytic activity of REn, may be important for replication (Hanley-Bowdoin *et al.*, 1999). Since experiments have indicated that REn increases the affinity of Rep for the origin (Fontes *et al.*, 1994a; Gladfelter *et al.*, 1997), possible functions of REn include (1) REn may direct Rep to its cleavage site in the origin, which in begomoviruses and curtoviruses is located distal from the Rep-binding site (analogous to the effect of papillomavirus E2 on E1 (Mohr *et al.*, 1990); (2) A REn/Rep complex may enhance DNA cleavage and ligation or putative helicase activities of Rep (Hanley-Bowdoin *et al.*, 2000), or (3) REn may play a similar role to transcription factors that enhance replication by stimulating the assembly of the initiation complex on the origin, in which case interaction with Rep would recruit REn to the origin, rather than vice versa. Although none of these possibilities has been proven *in vitro* or *in vivo*, it is very likely that interaction of REn with Rep plays an important role in the mechanism by which REn

enhances replication. Although mastreviruses do not encode a protein with homology to REn, it has been speculated that RepA, which interacts with *Mastrevirus* Rep, could play a similar role.

Yeast two-hybrid studies have identified the domains in MSV necessary for Rep and RepA homo- and hetero- oligomerization (Fig. 1.5; Horvath *et al.*, 1998). Although, as with begomoviruses, oligomerization of mastrevirus Rep is required to assemble a Rep-DNA complex at the origin, in WDV preformed Rep and RepA oligomers bind very poorly to DNA (Missich *et al.*, 2000). This situation is opposite to that reported in TGMV, where it was found that dimerization is required for Rep binding (Orozco and Hanley-Bowdoin, 1998). Interestingly, Missich *et al.* (2000) found that oligomerization of WDV Rep in solution is highly pH dependent within a small, perhaps physiological range. At pH 6.6–7.0, the predominant species is an oligomer (a >6 mer in the case of Rep and an octomer for RepA), while at pH 7.4–7.8 it is a monomer. The fact that preformed oligomers interact poorly with DNA, coupled with the pH-dependent oligomerization property of WDV Rep and RepA, strongly suggests that formation of large oligomers occurs in a stepwise manner. The first stage would be the interaction of a Rep monomer with DNA, mediated by the protein's DNA binding domain. The second stage, requiring the protein's oligomerization domain, would be the sequential addition of Rep monomers, which may lead to the stabilization of the oligomer assembled on the DNA (Missich *et al.*, 2000). Stepwise formation and pH-dependence of oligomeric structures in DNA has been proposed for other proteins, e.g. polyomavirus T-ag (Peng *et al.*, 1998) and SV40 T-ag (Runzler *et al.*, 1987). Furthermore, as has been proposed for geminivirus Repls, oligomerization and assembly of functional higher order complexes is crucial for the many different activities of SV40 T-ag forms (Missich *et al.*, 2000, and references within).

Rep and RepA have been implicated in activation of mastrevirus late gene expression (Hofer *et al.*, 1992; Zhan *et al.*, 1993, Collin *et al.*, 1996; McGivern, 2002), and it is likely that interaction of the two proteins plays an important role in this process. MSV RepA alone was shown to activate transcription of both the *HIS3* and *LacZ* reporter genes in yeast (Horvath *et al.*, 1998), and the activation domain was localized to the C terminus of RepA, which is not present in Rep (Fig. 1.5). However, no Rep/RepA-mediated activation of maize ubiquitin or *Cauliflower mosaic virus* (CaMV) 35S promoters was detected in maize cells (McGivern, 2002), suggesting that RepA does not directly activate plant host gene promoters. The same study found that RepA enhances CP gene expression, but only in the presence of Rep. Since RepA alone has transactivation ability in yeast, Rep may only be required for its replication ability, in order to

amplify the transcription template. However, Rep is likely to be required for additional functions, since in experiments using either Rep mutants unable to support replication, or a reporter construct unable to act as a replicon, expression of GUS in maize suspension cells was increased in the presence of Rep/RepA, but not RepA alone, independent of replication (McGivern, 2002). One possibility is that interaction of Rep with RepA may affect the DNA-binding property of RepA at the V complex. Another is that Rep bound at the V complex could recruit RepA to this site, similar to the proposed function of REn in recruiting Rep to the site of replication initiation. Further proposals include the requirement of a NLS, present in the C terminus of Rep but not in RepA, to target RepA to the nucleus (Boulton, 2002). This would explain why RepA alone can transactivate reporter gene promoters in yeast (Horvath *et al.*, 1998), but requires Rep to transactivate the CP promoter in plant cells (McGivern, 2002), since in yeast RepA is directed to the nucleus by the GAL4 nuclear targeting domain.

Rep, however, may simply be required for its own transcriptional activation activity. Two studies found that a region of the Rep C2, which is not present in RepA, can also activate reporter gene transcription in yeast in the absence of RepA (Hofer *et al.*, 1992; Horvath *et al.*, 1998), though Horvath *et al.* (1998) found that a deletion of the Rep C terminal 89 amino acids was required for transcription activation function. Interestingly, the transactivation region coincides with both the NTP-binding motif and a domain that shows homology to the DNA-binding domain of the avian myeloblastosis (*myb*) related class of plant transcription factors (Fig. 1.5; Hofer *et al.*, 1992). Although the transactivation of reporter genes by Rep was detected in yeast, this region may be functionally homologous in maize cells. In this case the activation region could allow Rep to activate either the CP promoter, or the promoters of cellular genes required during the replicative cycle.

The fact that full-length Rep is unable to activate reporter gene transcription implies that the activation domain is masked due to protein folding, but is exposed when Rep is truncated. Horvath *et al.* (1998) suggest that interaction between the C terminus of Rep and cellular proteins could expose the activation domain, thus ensuring that transcriptional activation does not occur until Rep has bound to the appropriate host proteins.

It is useful at this point to describe the protein responsible for transcriptional activation of the CP gene in begomoviruses, that is the highly conserved product of the AC2 (or C2) ORF, TrAP. The function of TrAP is not virus-specific, which suggests that either all begomovirus CP promoters

contain a common sequence element recognized by TrAP, or that TrAP interacts with cellular proteins common to all begomovirus plant hosts, or a combination of both. The transcriptional activation domain of TrAP, which resides in the acidic C terminus of the protein, also acts as a potent transactivator in mammalian and yeast cells (Hartitz *et al.*, 1999), which suggests that TrAP interacts with factors conserved in yeast, mammalian and plant cells. Optimal interaction of TrAP with viral ssDNA (which occurs in a sequence non-specific manner) requires the binding of zinc, which may facilitate the formation of TrAP-ssDNA complexes (Hartitz *et al.*, 1999). In addition, TrAP is phosphorylated; thus alternative phosphorylation may generate TrAP isoforms that stimulate transcription by different mechanisms. For example, phosphorylation may play a role in the mechanism by which TrAP activates *CP* expression in mesophyll cells, but derepresses the *CP* promoter in vascular tissue (Sunter and Bisaro, 1997). The most likely function of TrAP is to bind and recruit components of the transcription machinery, e.g. general transcription factors (GTFs) and TATA binding-protein-associated factors (TAFs) to the *CP* promoter. It has been shown that GTFs (including TATA binding protein) are contacted by acidic activation domains; thus the TrAP activation domain may stimulate transcription by interacting with similar factors (Hartitz *et al.*, 1999). It would be interesting to determine if the *Mastrevirus* RepA, in combination with Rep, plays a similar role to the begomoviral TrAP in recruiting components of the transcription machinery to the *CP* promoter.

A domain unique to RepA that mediates interaction with cellular proteins is a so-called GRAB (geminivirus RepA binding) protein-binding domain (Xie *et al.*, 1999). By using WDV RepA as a bait in the yeast two-hybrid system, Xie *et al.* (1999) isolated a family of proteins (GRAB), the N-terminus of which exhibits a significant amino acid homology to the NAC (non-apical-meristem, ATAF and CUC2 genes) domain present in a family of plant-specific proteins that are involved in a variety of processes, ranging from lateral root formation to development and senescence. Although this study was done in WDV, the residues required for interaction with GRAB proteins are located in the C terminal domain of RepA (Fig. 1.5), a region that has a significant degree of conservation in all mastrevirus RepA proteins (Xie *et al.*, 1999), except one, that is *Miscanthus streak virus* (MiSV) which lacks the C terminal region (Boulton, 2002). Thus, it is likely that the same interaction occurs in other mastreviruses. The fact that expression of GRAB proteins severely interferes with WDV DNA replication (Xie *et al.*, 1999) points to a role of GRAB proteins in cellular pathways that negatively affect viral replication. By binding to GRAB proteins, RepA may remove the inhibitory block on viral replication. However, a mutant MSV genome expressing a RepA protein that lacks the C terminal 45 amino acids (and therefore

the putative GRAB binding domain) is able to infect maize (Boulton, 2002). Thus, if this domain does exist in MSV, it is dispensable for viral infection of maize.

Further associations of WDV Rep with cellular proteins, identified in a similar manner to that described above, include an interaction with the wheat replication factor C complex (TmRFC-1) (Luque *et al.*, 2002). In eukaryotes RFC is crucial for the recruitment of DNA polymerase δ . During initiation of viral replication, Rep generates a 3'-OH terminus, to which cellular replication factors must be recruited for the assembly of an elongation complex. The data of Luque *et al.* (2002) suggest that WDV Rep, having introduced a nick at the initiation site in the LIR and produced a 3'-OH terminus, stimulates the recruitment of RFC by binding to the RFC large subunit, leading eventually to the recruitment of DNA polymerase δ . This model is further explored in the section on geminivirus replication.

The method employed by Xie *et al.* (1999), using RepA as a bait to bind to host proteins, is a potentially very useful way of identifying cellular proteins involved in processes related to plant growth and the cell cycle, since one of RepA's main functions is thought to be to influence the plant cell regulatory cycle to the benefit of viral replication (Xie *et al.*, 1995). An important domain in *Mastrevirus* Rep and RepA that is the focus of much ongoing research in mastreviruses is the retinoblastoma-related (RBR) protein interaction domain (Fig. 1.5). The retinoblastoma (Rb) tumour suppressor protein, the founding member of the family of so-called "pocket proteins", has long been known to play a pivotal role in the regulation of the human cell cycle. In particular, Rb negatively regulates the cellular G1/S transition of the proliferative cell cycle and is required for proper differentiation of certain cell types. For example in skeletal muscle Rb is required both for expression of late stage differentiation markers and for irreversible exit from the cell cycle (Adams, 2001). The role of Rb is mediated, at least in part, by binding to the E2F family of transcription factors via the Rb "A/B pocket" domain. E2F binding sites have been identified in the promoters of a number of genes involved in DNA replication and progression of the cell cycle. Thus, by binding to E2F transcription factors Rb exerts a block on cell proliferation (de Jager and Murray, 1999). Although *in vitro* Rb can inactivate E2Fs simply by binding and masking the E2F transactivation domain (Ross *et al.*, 1999), Rb also represses transcription by binding to histone deacetylase (HDAC), which removes acetyl groups from the tails of histone octomers (Dahiya *et al.*, 2000). This histone deacetylation activity appears to facilitate condensation of nucleosomes into chromatin, which in turn blocks access of transcription factors, leading to gene repression. Therefore, active repression by the Rb-E2F

complex at the promoters of cell cycle genes is thought to be mediated at least in part by recruitment of HDAC by Rb; an IXCXE site in the C terminus of HDAC seems to be important in mediating association with Rb (Magnaghi-Jaulin *et al.*, 1998).

Rb is regulated by the activity of cyclin/cyclin-dependent kinases (cdks), which are heterodimers of a catalytic cdk and a regulatory cyclin subunit. D-type cyclins, whose expression is strongly induced by growth regulatory signals, bind to and activate their specific cdk partners (particularly cdk4). These kinase complexes interact with the Rb A/B pocket via a conserved N-terminal Rb-binding motif consisting of the amino acids LxCxE in the D-type cyclins. As a result of this interaction, D-cyclin kinases direct the phosphorylation of Rb on multiple cdk phosphorylation sites. Phosphorylated Rb is inactive, since it can no longer bind to E2Fs, and E2F-regulated genes are thereby released from the transcriptional silencing induced by Rb, allowing cell cycle progression into S phase to occur (Weinberg, 1995). In turn, control of the kinase complexes is exerted by cdk inhibitory proteins, which inhibit the activity of the D-cyclin kinases in the quiescent or differentiated state (de Jager and Murray, 1999). Recent evidence suggests that this already complicated sequence of events is even more complex, in that Rb is a multifunctional protein which can inhibit transcription through various mechanisms, and that these functions are progressively and cooperatively inactivated by multiple cyclin/cdk complexes during G1 and S phase (Adams, 2001). However, for the purposes of this discussion, the simpler model will suffice.

The *RB* gene is frequently implicated in tumour formation; indeed it was originally identified because individuals who inherit a mutant copy of the gene are predisposed to develop childhood retinoblastoma (Adams, 2001). The *RB* gene is also associated with a number of other cancers, due to mutation of *RB* or abrogation of its function. As an example of the latter, DNA tumour viruses (such as SV40, adenovirus and oncogenic human papillomavirus (HPV) subtypes) produce oncoproteins that bind to the Rb A/B pocket, displacing E2Fs which then allows E2F-directed transcription of the genes involved in DNA replication and which are necessary for efficient viral replication (Weinberg, 1995; de Jager and Murray, 1999).

Until a few years ago, the powerful role of Rb in mammalian cell cycle control, cellular differentiation and development, together with its involvement in cellular transformation and cancer, were thought to be unique to the vertebrates (de Jager and Murray, 1999). That was until the surprising discovery of homologues of the human Rb protein in maize (Grafi *et al.*, 1996; Xie

et al., 1996; Ach *et al.*, 1997) and subsequently other plants, including tobacco, *Chenopodium rubrum* and *Arabidopsis* (de Jager and Murray, 1999). Analysis of the predicted amino acid sequence of the maize protein, ZmRb, revealed a striking conservation of the domain organization of the plant protein with human Rb and the Rb-related proteins, p107 and p130, particularly the A and B domains of the pocket region, which show 50-65% similarity (Gutierrez, 1998). As described above, this region is necessary and sufficient for Rb to bind to and negatively regulate E2Fs, to bind to and be inactivated by D cyclin-cdk complexes, and to bind to and be sequestered by viral oncoproteins. The conservation of ZmRb with mammalian Rb proteins over the A/B pocket raised the possibility that function may also be conserved at the molecular level (de Jager and Murray, 1999).

Even before the discovery of retinoblastoma-related protein (Rb) in plants, evidence was already accumulating for the existence of an homologous Rb pathway in the regulation of the plant cell cycle. First, a plant D-type cyclin (CycD) homologue was isolated from *Arabidopsis* (Soni *et al.*, 1995). Although the *Arabidopsis* CycDs share only a low level of sequence identity with mammalian D-type cyclins, of particular interest was the presence of the Rb-binding LxCxE motif at the N-terminus of the plant CycDs, which raised the possibility that their target could be a plant Rb homologue. The subsequent isolation of CycD genes from a wide range of plant species, showing that the LxCxE motif is an invariant feature (de Jager and Murray, 1999), supported this view.

The second line of evidence came from geminiviruses, when WDV Rep and RepA were found to contain an Rb-binding LxCxE motif, mutation of which reduced viral replication efficiency (Xie *et al.*, 1995). These findings suggested that Rep and/ or RepA may be required to bind a plant Rb homologue, thereby sequestering it in an inactive form to remove a negative block on viral replication, analogous to the strategy used by animal oncoviruses such as SV40, adenoviruses and HPV.

Although ZmRb was able to bind (albeit weakly) to both human E2F-1 and the *Drosophila* E2F homologue dE2F, and to negatively regulate E2F-dependent transcription (Huntley *et al.*, 1998), the identification of plant E2F-like proteins proved elusive. Finally, in 1999 Ramirez-Parra *et al.* isolated from wheat a cDNA that encoded an E2F homologue, providing further evidence that the components of the Rb pathway are present in plants. In addition, ZmRb was found to interact much more strongly with this plant E2F than with human E2F-1 (Inze *et al.*, 1999), and putative

E2F binding sites were identified in some plant promoters that show S-phase specific expression (Inze *et al.*, 1999).

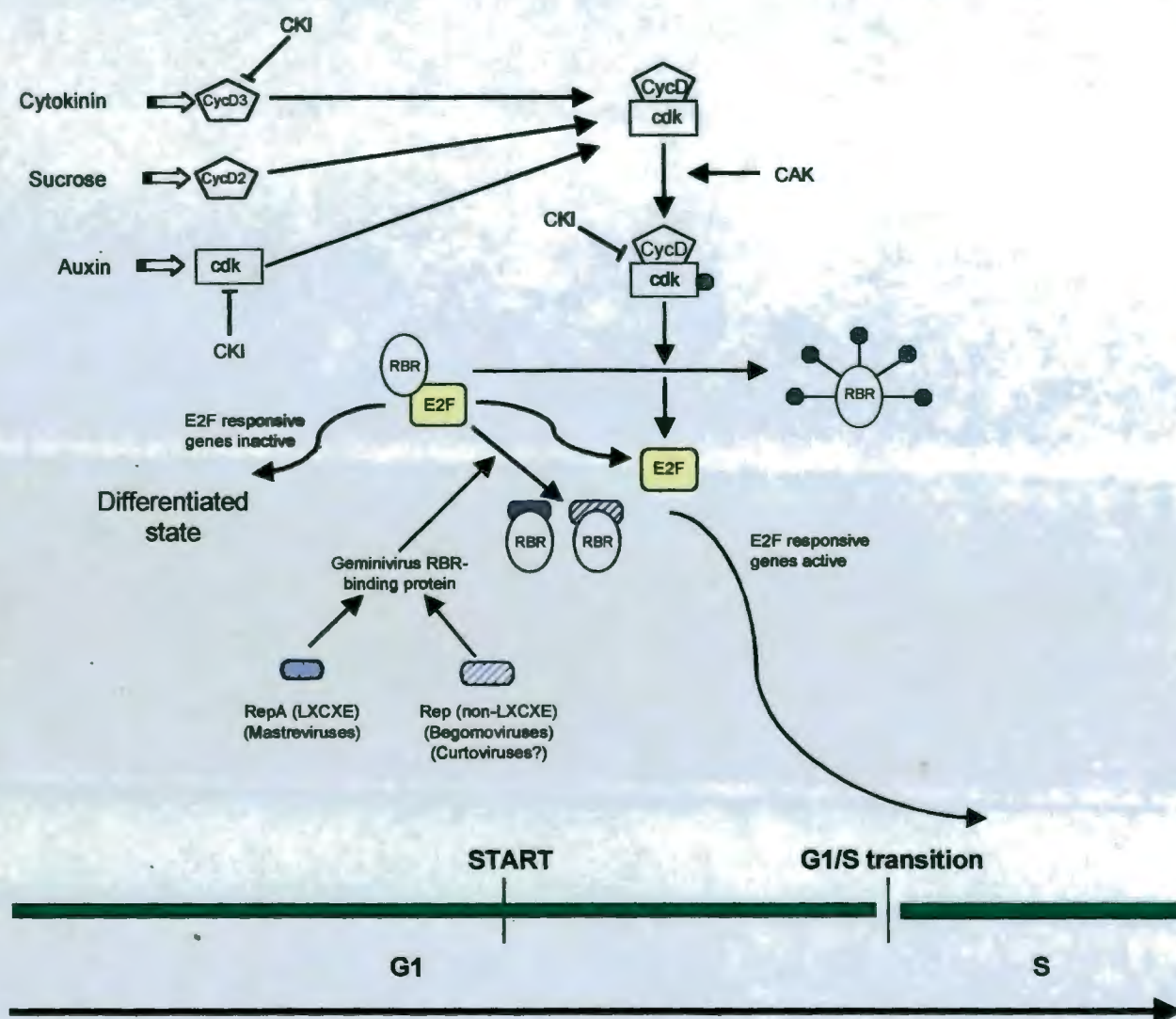


FIGURE 1.6 Model for the control of the G1-to-S phase transition in plant cells, and for the proposed interference of geminivirus Rep proteins with the retinoblastoma-related (RBR) pathway (adapted from the diagrams of Meijer and Murray, 2000; and Gutierrez, 2000). The key components of the pathway are shown. These include CycD cyclins, which are plant homologues of the mammalian D-type cyclins; cyclin-dependent kinases (cdk); cdk inhibitory proteins (CKI); E2F transcription factors that activate the expression of S-phase-specific genes; and retinoblastoma-related (RBR) protein, the plant homologue of mammalian retinoblastoma (Rb) protein. RBR is regulated by the activity of CycD/cdk complexes, whose expression is strongly induced by growth regulatory signals (hormones/nutrients). The CycD/cdk complex is activated by specific phosphorylation of the cdk component by cdk-activating kinase (CAK), and cdk in turn directs the phosphorylation of RBR. This results in the dissociation of E2F from RBR, allowing expression of S-phase genes and cell cycle progression. The normal RBR phosphorylation pathway may be bypassed in geminivirus-infected cells by the action of RepA (mastreviruses) and Rep (begomoviruses and potentially curtoviruses). By binding to RBR, these proteins may drive cells into S phase (or enable the activation of specific S-phase factors) by promoting release of E2F from RBR.

Further results pointing to a functional conservation of plant and mammalian Rb include the kinase activity displayed against ZmRb by human D cyclin-cdk, that requires an intact ZmRb A/B pocket (suggesting a specific interaction with the A/B pocket is required to exert their kinase function); the binding of SV40 LTA and HPV E7 proteins to ZmRb (de Jager and Murray, 1999); and the binding of *Arabidopsis* CycD cyclins to ZmRb via their LXCXE motif (Huntley *et al.*, 1998). Recently, direct evidence that ZmRb is a cell cycle regulator was provided by Gordon-Kamm *et al.* (2002), who demonstrated that expression of ZmRb inhibits cell division in tobacco cell cultures.

Taken together these results suggest that, although specific mechanistic details may differ (Mironov *et al.*, 1999; Ramirez-Parra *et al.*, 1999), the major components of the mammalian and plant Rb pathways are functionally equivalent. In the plant RBR pathway model (illustrated in Fig. 1.6), growth stimulatory mechanisms such as sucrose and cytokinins induce the expression of CycD cyclins. These form active kinase complexes targeting RBR for inactivation and dissociation from the promoter-bound E2Fs, allowing expression of S phase genes and culminating in DNA replication and cell cycle progression (de Jager and Murray, 1999). The RBR pathway may also be important in plant differentiation and development. For example, in developing maize leaves, which show a gradient of cell proliferation from actively dividing cells at the leaf base to differentiated cells nearer the tip, the ZmRb protein is abundant in the differentiating cells at the leaf tip and almost undetectable in the proliferating cells of the leaf base (Huntley *et al.*, 1998).

The excitement associated with the discovery of an Rb pathway conserved in mammals and plants is understandable, as it may help our understanding of plant and animal evolution. Given that the components of the Rb pathway are not present in yeast or fungi, which use unrelated proteins to control the same processes, a common pathway relating cell division and differentiation may have arisen only once in the evolution of all higher eukaryotes. This suggests that the invention of the Rb pathway may have been the factor that allowed multicellular organisms to develop complex body plans (de Jager and Murray, 1999).

The discovery of the LxCxE RBR protein interaction motif and other cellular protein-binding domains (such as the GRAB-binding domain) in WDV RepA has prompted some intense research, with the aim of not only elucidating the components and mechanisms of the viral life cycle, but also of providing insights into DNA replication, cell cycle and growth control in plants.

Now that a background has been given on the Rb protein function and the discovery of its homologue in plants, the relationship between the RBR protein, geminivirus Reps and associated effects on the plant cell cycle in geminivirus-infected plants can be described.

Geminivirus replication depends, apart from Rep, entirely on host cell proteins. These cellular replication proteins are generally absent or not functional in differentiated cells. Geminivirus replication is excluded from the meristem, which contains actively proliferating cells that naturally express the replication factors required for geminivirus replication. Thus, even before the relationship between the plant RBR protein and the WDV Rep LXCXE motif was known, it was suggested that geminiviruses might induce a cellular state permissive for viral replication. There were two main observations supporting this idea. The first was that dsDNA replication intermediates are significantly more abundant in S-phase nuclei than in nuclei from other phases of the cell cycle (Acotto *et al.*, 1993), suggesting that cellular DNA replication (the S phase of mitosis rather than actual cell division) is important for viral replication. The second observation was that proliferating cell nuclear antigen (PCNA), which is virtually undetectable in differentiated cells, accumulates to high levels in fully differentiated cells expressing TGMV Rep (Nagar *et al.*, 1995). PCNA, whose expression is regulated by E2F, has an essential role in DNA replication, functioning as a homotrimeric clamp at the origin, which facilitates the recruitment and processivity of DNA polymerase δ (Sever-Chroneos *et al.*, 2001).

However, in some geminiviruses a correlation was not found between viral and host DNA replication. By comparison of the distribution of replicative forms of MSV DNA with the expression of the S-phase-specific host gene, H2b, Lucy *et al.* (1996) showed that host DNA synthesis is not a prerequisite for MSV replication; this lack of correlation was most notable in the developmentally mature leaf laminal tissues where MSV replication could be detected. This suggests that although different geminiviruses may interfere with control of cell proliferation, more than one mechanism or strategy may have evolved. Although it appears that certain geminiviruses such as BCTV do have the capacity to initiate host cell division (causing tumorigenic growths in infected plants; Latham *et al.*, 1997), the data of Lucy *et al.* (1996) suggest that one or more factors associated with dividing cells, rather than cell division itself, may be required for MSV replication.

Given that geminiviruses may induce a cellular state that is permissive for viral DNA replication, either by inducing cells to enter the S phase or by activating some S phase function/s, the

discovery of RBR in plants and the LxCxE RBR-interaction motif in WDV RepA provided a possible mechanism by which this cell-cycle interference could occur (Fig. 1.6). The LxCxE motif is present in the RepA of most mastreviruses, including dicot-infecting members such as BeYDV (Liu *et al.*, 1999a). Although the mastrevirus Rep shares the LxCxE motif with RepA (Fig.1.5), Rep does not interact with RBR (Horvath *et al.*, 1998; Liu *et al.*, 1999a; results in this thesis), most likely due to steric hindrance induced by the C terminal domain of Rep (Gutierrez, 2000). This raises the possibility that RepA may be required in the viral life cycle for its ability to provide a cellular environment competent for viral replication through interaction with RBR and other cellular factors, such as the GRAB proteins mentioned earlier. There is convincing evidence supporting this view. First, both BeYDV (Liu *et al.*, 1998), and MSV (Boulton, 2002) mutants unable to express RepA are also unable infect plants. Second, expression of WDV RepA stimulates cell division in tobacco cell cultures (Gordon-Kamm *et al.*, 2002). The stimulatory effect is observed not only in mitotically active cells, but RepA also appears to overcome cell arrest in the G₀/G₁ phase of the cell cycle. Furthermore RepA (but interestingly not Rep) expression stimulates maize embryogenic callus growth and increases transformation efficiency, most likely due to stimulated cell division. Overexpression of ZmRb suppresses RepA-stimulated cell division, suggesting that the effects of RepA on the cell cycle and on transformation are due to the removal of RBR protein-mediated repression of the cell cycle. However, no mastrevirus-induced host cell proliferation has yet been observed, so it is unlikely that in a natural infection these viruses induce total progression from G₁ phase through S phase to mitosis.

RepA-RBR interaction may not only have an effect on viral replication. As mentioned previously, RepA has been implicated in the transactivation of the mastrevirus *CP* promoter. This activity could be facilitated by the binding of RepA to RBR protein, thus releasing E2F transcription factors that could then be involved in activating the *CP* promoter. Interestingly, the E1a protein from adenoviruses uses this strategy to promote activation of viral promoters (La Thangue, 1994).

Mutational analysis of the mastrevirus RepA LxCxE motif has shown the importance of the three conserved residues in mediating binding to RBR. In one study, a mutation of C to G had a significant effect on the ability of WDV RepA to bind to Rb, while changing the E residue to K abolished binding (Xie *et al.*, 1995). In a detailed mutational analysis of the motif in BeYDV RepA, Liu *et al.* (1999a) confirmed the importance of all three conserved residues: while mutants containing changes of L to I; C to S; and C to G retained the ability to bind to RBR, they did so

with reduced efficiency. Furthermore, in a result similar to that of Xie *et al.* (1995), mutation of the E residue (in this case E to Q) had the effect of drastically reducing binding efficiency, by 95%. The LXCXE motif also mediates binding of a nanovirus gene component (Clink, for "cell cycle link") to RBR (Aronson *et al.*, 2000), and the importance of the motif in the oncoproteins of animal tumour viruses has already been mentioned. Thus, it was surprising to find that TGMV Rep, which encodes no LXCXE motif, interacts with RBR (Ach *et al.*, 1997) through a different motif (Kong *et al.*, 2000). There are several lines of evidence that TGMV alters the cell-cycle controls of its host plant, *Nicotiana benthamiana*. Indications that TGMV may trigger re-entry of mature cells into the cell cycle include the incorporation of high levels of bromodeoxyuridine into both viral and host DNA in infected cells, suggestive of progression into S phase and DNA replication (Egelkrout *et al.*, 2001), and the fact that a large fraction of TGMV-infected cells contains condensed chromatin, which is characteristic of early mitotic prophase (Bass *et al.*, 2000). The importance of TGMV Rep-RBR interaction was established by Kong *et al.* (2000), who found that mutants with impaired binding to RBR accumulate less viral DNA and cause chlorosis that is confined to the veins, and that the mutations result in altered tissue specificity.

A clue to the mechanism by which binding of TGMV Rep to RBR could alter the cellular environment came from analysis of the *N. benthamiana* PCNA promoter. As mentioned previously, PCNA is an S phase gene that is regulated by E2F. In a study of transgenic plants carrying a mutation of an E2F consensus element in the *N. benthamiana* PCNA promoter, Egelkrout *et al.* (2001) demonstrated that E2F functions as a negative regulatory element to repress PCNA transcription in mature leaves. Thus, E2F can repress as well as activate PCNA promoter activity. The mechanism by which this occurs may be similar to the E2F-mediated control of the tobacco ribonucleotide reductase small subunit gene (RNR2) expression. The RNR2 promoter contains multiple E2F elements, two of which activate transcription and one that represses transcription out of S phase (Egelkrout *et al.*, 2001). It has previously been shown that PCNA expression in infected cells is tightly linked to the capacity of TGMV Rep to interact with RBR (Kong *et al.*, 2000). Taken together, these data suggest that TGMV Rep may disrupt the RBR/E2F complexes that repress PCNA promoter function and in so doing induce PCNA transcription in mature leaves (Egelkrout *et al.*, 2001).

Rep is not the only TGMV protein to interact with RBR protein. Settlage *et al.* (2001) established that the TGMV replication enhancer protein, REn, binds to RBR protein as well as Rep. However, unlike Rep, REn alone cannot induce expression of PCNA (Settlage *et al.*, 2001). The

protein domains that mediate the interactions between Rep, REn and RBR overlap, suggesting that REn serves a dual role in enhancing geminivirus replication and that these two functions are co-ordinated by shared-protein domains. Although Rep oligomerization may be required for REn binding (Settlage *et al.*, 2000) and has been demonstrated to be a prerequisite for RBR binding (Kong *et al.*, 2000), oligomerization of REn is not required for interaction with Rep or RBR. Settlage *et al.* (2001) suggest that Rep and REn may serve different roles in the host induction process. One possibility is that REn regulates Rep/RBR interactions through a partially shared protein interaction domain. Both Rep and RBR are predicted to bind to REn within its first 35 amino acids (Settlage *et al.*, 2001). Given that RBR is greater than 100kDa in size and that Rep probably binds REn as an oligomer, REn may not be able to interact simultaneously with both proteins. The authors propose that Rep/REn, but not Rep/RBR complexes are functional for initiation of geminivirus replication. According to this model, REn modulates the stoichiometry of different Rep complexes, and hence Rep activity, through its binding to both Rep and RBR protein. It would be interesting to determine if RepA, which also binds to both Rep and RBR, has a similar function in mastreviruses.

Apart from RBR protein, geminivirus Rep may also interact with other host proteins involved in cell division and development. For example, interaction of TGMV Rep with a kinase and a kinesin from *Arabidopsis*, both of which are potentially involved in the progression of the cell cycle through the G2 and M phases, possibly prevents transit through the G2 phase, stalling the cell cycle in an S-like phase that results in endoreduplication rather than cell division (Kong and Hanley-Bowdoin, 2002).

Further interactions with plant factors were determined in the nanovirus *Faba bean necrotic yellows virus* (FBNYV) genetic component, Clink. In addition to the LxCxE RBR protein interaction motif, Aronson *et al.* (2000) identified in Clink an F-box that binds to a plant SKP1 homologue, a constituent of the ubiquitin-protein turnover pathway. F-box proteins serve as substrate-specific adapter subunits to recruit various substrates to be ubiquitinated, in preparation for degradation by the 26S proteasome. Therefore, FBNYV Clink may target RBR protein for ubiquitin-mediated degradation, as was described for HPV-16 E7 protein (Boyer *et al.*, 1996). As suggested for TGMV and other geminiviruses, nanoviruses probably achieve viral replication without the completion of mitosis (for example by the process of endoreduplication), since neither group causes uncontrolled host cell proliferation. Thus, targeting the degradation of proteins (including the degradation of Clink itself) in a later phase of the cell cycle could restore a

semblance of normal cell cycle control in the host cells, once a critical amount of viral genome products has accumulated (Aronson *et al.*, 2000).

It would seem that interference with the RBR pathway is an important part of the geminivirus life cycle. However, there is substantial evidence to suggest that the RBR-binding activity of Rep is dispensable for both viral replication in cultured cells and infectivity in host plants. Although Xie *et al.* (1995) found a correlation between the ability of WDV RepA to bind to RBR and the replication efficiency of the virus in wheat suspension cells, no such relationship has been detected in other mastreviruses. Liu *et al.* (1999a) found that all BeYDV mutants, in which RBR binding was impaired by varying degrees up to 95%, were able to replicate in tobacco protoplasts and to systemically infect *N. benthamiana* and bean. In addition, the FBNYV genetic component that binds to RBR, Clink, is not absolutely required for viral replication in *N. benthamiana*, although the LxCxE sequence of Clink does enhance FBNYV replication (Aronson *et al.*, 2000). Further examples of the redundancy of the LxCxE motif in viral genes include those found in potyviruses (a group of plant RNA viruses). Although the potyviral RNA-dependent RNA polymerase (NIb) carries the LxCxE motif, it does not interact with any RBR proteins, and mutation of the highly conserved E residue to K has no effect on viral replication (Oruetxebarria *et al.*, 2002). Although this is not so surprising considering that potyvirus replication is largely independent of the cellular replication activities, the conservation of this motif in a range of plant and animal RNA viruses (a similar motif is also found in the replication proteins of bymoviruses, potexviruses and *Rubella virus*) is suggestive of an alternative function, such as the global folding of the NIb protein necessary for interactions with other proteins during virus replication (Oruetxebarria *et al.*, 2002). Although it may be for a different reason, the possibility exists that in geminiviruses and in particular the mastreviruses, either the LxCxE motif is required for an activity other than RBR binding, or at least one other functional domain may overlap the LxCxE motif, resulting in its conservation in mastrevirus Reps.

Clearly, there is a variety of possible strategies for geminiviruses to interfere with the regulation of the cell cycle and other host pathways, mediated by a number of interactions between viral and host proteins, many possibly as yet undiscovered. Adding to the confusion, the exact function and purpose of the RBR-interaction motifs in geminiviruses, particularly mastreviruses, is yet to be determined. Part of the work in this thesis attempts to provide further insight into the role of the LXCXE motif in MSV Rep.

1.2.3 Replication

Geminiviruses replicate via double stranded circular intermediates, which form minichromosomes within the nuclei of infected cells (Abouzid *et al.*, 1988). The generally accepted model is that geminiviruses replicate using a rolling circle mechanism (RCR). However, recent evidence suggests that geminiviruses might also multiply via recombination-dependent replication (RDR; Jeske *et al.*, 2001). The focus of this section is on the characteristics of RCR, which have been more extensively studied; however it is important to bear in mind that an alternative route of replication by recombination most likely exists.

In the RCR model, amplification of the viral genome occurs in three stages (Fig. 1.7). The first stage is the conversion of the genomic circular ssDNA [(c)ssDNA] into supercoiled covalently closed circular dsDNA [(ccc)dsDNA] intermediates, or replicative form I (RFI). This involves the host-directed, DNA-primed (or RNA-primed in the case of begomoviruses) synthesis of a complementary (minus) strand, as described in the section on the mastrevirus SIR. The second stage is the amplification of the RFI by RCR, which is discussed in detail below. The production and encapsidation of mature genomic (c)ssDNA into viral particles represents the third and final stage of the geminivirus replicative cycle.

The viral RFI serves as a template for both viral transcription (leading early on to the expression of Rep and RepA), and for further DNA replication steps. As mentioned previously, the replication initiation site is contained in the loop of a stem-loop structure within the mastrevirus LIR. DNA sequences around the initiation site comprise the replication origin, which exhibits a modular organization. Upon interaction with specific DNA sequences at the viral origin, Rep introduces a sequence-specific endonucleolytic nick in the V sense genomic strand of the RFI. In the case of WDV, Rep binds with low affinity at the stem of the stem-loop, constituting an O complex, and with high affinity ~140 nucleotides upstream from the initiation site in the vicinity of the C sense promoter, constituting a C complex. Formation of the O complex is sufficient to carry out sequence-specific cleavage at the loop (Castellano *et al.*, 1999). However, it is possible that Rep molecules bound at the C and O complex may interact to form a higher order complex at the origin (Gutierrez, 1999). The initiation reaction involves the nucleophilic attack by the OH group of the conserved tyrosine residue in motif III of Rep, to the phosphodiester bond between the last T and A residues of the invariant nonanucleotide plus strand origin sequence (TAATATT↓AC; Heyraud-Nitschke *et al.*, 1995; Stanley, 1995; Laufs *et al.*, 1995a).

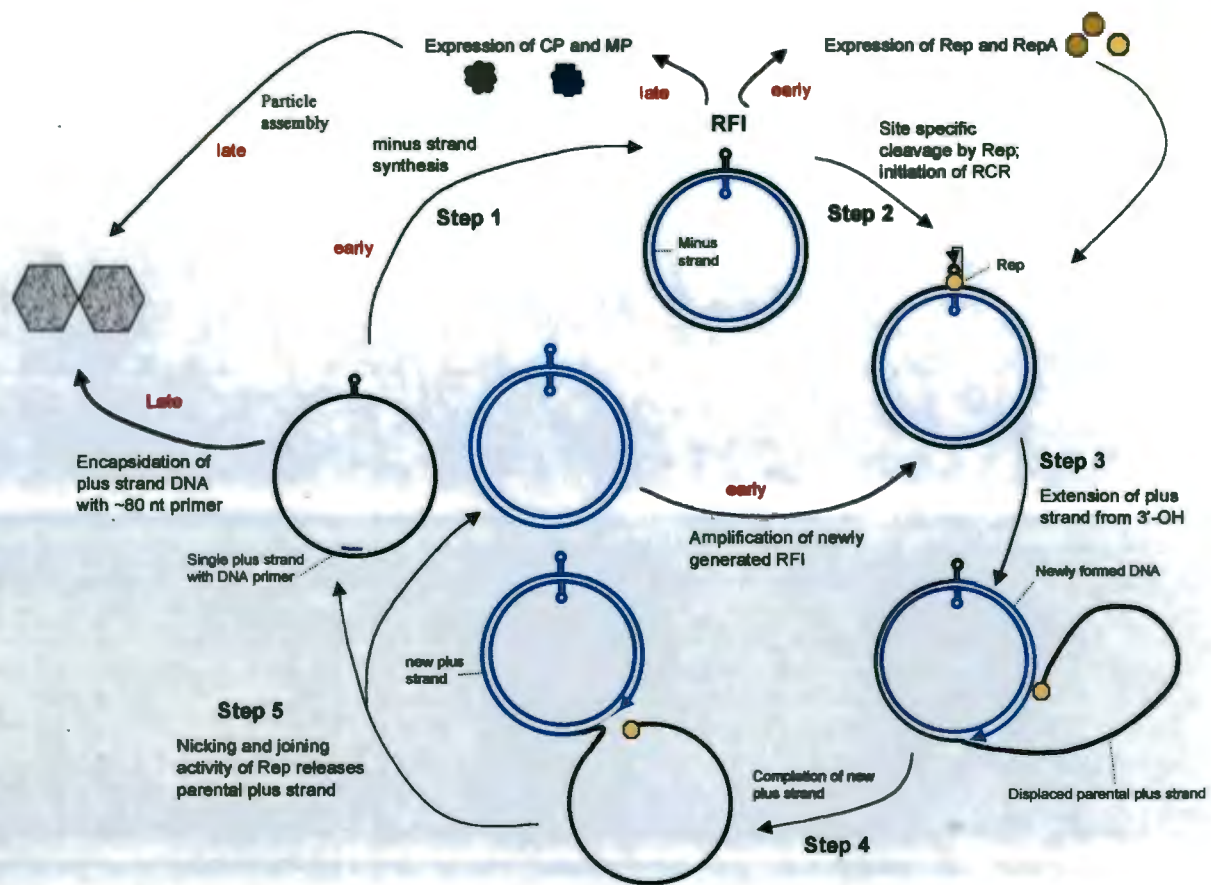


FIGURE 1.7 Summary of the MSV replicative cycle. Early in the cycle genomic circular ssDNA is converted into a dsDNA replicative form intermediate (RFI; step 1). RFI serves as a template for transcription of Rep and RepA (early in the cycle) and CP and MP (late in the cycle). Rep initiates rolling circle replication (RCR; step 2) by binding to the viral origin and introducing a nick in the loop of the plus strand DNA stem-loop structure. The initiation reaction results in the production of a free 3'-OH terminus which is used as a primer for synthesis of a new plus strand (step 3), while Rep remains covalently linked to the phosphorylated 5'-OH end. Once the new plus strand is synthesised by the host replication machinery (step 4), the parental plus strand is displaced from the negative strand template, and termination occurs whereby the nicking-joining activity of Rep simultaneously releases the parental plus strand and liberates a (c)ssDNA molecule (step 5). Early in the cycle the newly liberated plus strand is converted back into RFI to begin the cycle once again, while late in the cycle it is encapsidated and moved out of the cell to establish a systemic infection.

The Rep-mediated initiation reaction results in the production of a free 3'-OH terminus (which is used as a primer for synthesis of a new plus strand), while Rep remains covalently linked to the phosphorylated 5'-OH end (Laufs *et al.*, 1995b). In eukaryotes, the efficient binding of a processive DNA polymerase complex to a 3'-OH primer-terminus depends on the function of the

RFC clamp loader. RFC facilitates loading of the PCNA clamp, which eventually recruits DNA polymerase δ . It is known that in WDV (and most probably other geminiviruses), Rep interacts with the host RFC, thus stimulating its recruitment to the viral origin after initiation. Subsequently, PCNA could be incorporated into the pre-elongation complex, or alternatively, a preformed RFC/PCNA complex could directly be recruited onto the newly formed 3'-OH primer-terminus. Finally, a DNA polymerase would be recruited, concomitantly with ATP hydrolysis and release of RFC, leading to the assembly of an elongation complex that can extend the primer (Luque *et al.*, 2002).

Once the new strand is synthesized, the parental plus strand is displaced from the intact negative strand template, possibly mediated by the putative helicase activity of Rep. The mechanism of termination, whereby the replication cycle is resolved to release a (c)ssDNA molecule, is unknown. By analogy with other rolling circle replicons, geminivirus RCR may be a continuous process, as is the case for phage Φ X174, or a noncontinuous process, as for example in plasmid pC194. In the former case, completion of the new plus strand regenerates the origin of replication, which again is nicked by Rep, this time acting as a terminase to displace the parent plus strand. This Rep becomes covalently linked to the 5'-AC end of the new nonanucleotide, and the 5' phosphoryl group of the displaced strand is then transferred to the newly generated 3'-OH group to liberate a (c)ssDNA molecule. In this model, two active tyrosines are required for switching from initiation to termination. Since Rep contains a single conserved active tyrosine, this mechanism would require two Rep molecules, possibly within the same bound oligomer: in this way an active Rep molecule can always remain attached to the DNA via alternating tyrosines. Alternatively, the mechanism for resolution could be discontinuous, requiring only one tyrosine. As in the continuous model, after nicking Rep becomes linked to the 5' end of the cleaved DNA via a phosphotyrosine linkage. After one round of DNA synthesis, the release of the (c)ssDNA is mediated by a non-tyrosine residue in the same Rep molecule. The newly synthesised origin is then cleaved, and the 5' end that was linked to the Rep active site tyrosine is transferred to the newly created 3'-OH end. Thus, one tyrosine is sufficient to effect nicking and resolution of (c)ssDNA, but the next cycle of replication must be re-initiated by a different Rep molecule, rendering it non-continuous (Palmer and Rybicki, 1998).

Depending on the stage of the replication cycle, the newly released plus strand may either be incorporated into the replication pool to be converted into another RFI (early in the cycle), or will be accumulated as (c)ssDNA, destined for encapsidation or movement to surrounding cells,

thereby removing it from the replication pool (late in the cycle). This aspect is covered in the section on the viral life cycle.

1.2.4 The MSV Life Cycle

An MSV infection is presumably initiated when a leafhopper introduces at least one virus particle (encapsidated ssDNA) into a nucleated plant cell, probably a phloem companion cell. Virus particles delivered into the phloem of leaves are apparently transported within sieve tubes to regions distal from the leafhopper feeding site (Peterschmitt *et al.*, 1992); a sufficient number move out of an inoculated leaf within a few hours to initiate a productive infection (Storey, 1938). It is unknown whether an infection can be established upon entry of a virus particle into any nucleated cell, or whether the cell must be undergoing active cell division and have DNA replication enzymes available. Plant DNA replication and cell division are confined to apical meristems, developing leaves and the cambium of mature plants (Hanley-Bowdoin *et al.*, 1999). Accordingly some geminiviruses, such as AbMV, are restricted to the phloem (Abouzid *et al.*, 1988), possibly replicating in procambial cells. However, other geminiviruses are not confined to vascular tissue, and may modify differentiated cells to induce the synthesis of replication enzymes. For example, MSV DNA can be detected in vascular and mesophyll cells of mature leaves (Lucy *et al.*, 1996), although paradoxically MSV DNA is not detected in meristematic cells.

Once introduced into a permissive cell the first step would be for the virion to uncoat, followed by the movement of the released ssDNA to the nucleus, where viral replication takes place. Nuclear import of viral DNA must then occur in each subsequently infected cell. In MSV, movement of probably partially uncoated ssDNA into the nucleus is facilitated by the CP, which has DNA-binding activity (Liu *et al.*, 1997) and a nuclear localization signal (NLS; Liu *et al.*, 1999b).

Once in the nucleus, the ssDNA viral genome is converted to (ccc)dsDNA (RFI) as described in the previous section. It is assumed that no modification of host cell gene expression is necessary at this stage, since nick repair enzymes and cofactors presumably responsible for synthesizing the negative strand are constitutively expressed (Palmer and Rybicki, 1998).

Once transformed into RFI, the first priority would be to express Rep and RepA. Rep is essential to initiate RCR, and RepA may be required to induce the expression of host enzymes and co-factors required for the completion of RCR. As discussed previously, Rep and/or RepA may directly induce the promoters of certain host genes required for viral replication (since both proteins are transcriptional activators), or they may interfere with cell cycle regulatory systems to indirectly induce the host genes required for virus replication. It is interesting to note that in maize suspension cells the MSV C sense promoter is most active in the early S phase of the cell cycle (before the start of histone H4 transcription), while the CP promoter shows the highest activity in early G2 (Nikovics *et al.*, 2001). The difference in the expression timing from these promoters is consistent with the functions of the MSV gene products. For example, a relatively high level of expression from the C sense promoter by early S phase could result in RepA-mediated manipulation of the host cell cycle by binding to RBR, thereby potentially releasing factors regulating the transition from the G1 to the S phase of the cell cycle. Similarly, Rep is necessary to initiate RCR, which is likely to be accomplished during the S phase. In contrast, CP is required for encapsidation of MSV DNA and systemic infection (Liu *et al.*, 1997) and thus is not needed in the early stages of MSV replication (Nikovics *et al.*, 2001). Interestingly, the C sense promoter is re-activated in the late G2 phase, which may reflect a requirement for Rep or RepA to interfere with progression through G2 phase, thereby locking infected cells into S phase (Nikovics *et al.*, 2001; Nagar *et al.*, 1995).

Apart from their function in the replication of the viral genome, Rep and RepA may have an integral role in the regulation of the entire virus infection cycle. While RepA may be required early in the infection process to prepare the cellular environment for replication, it also potentially plays an important role in activating the CP promoter. Since the CP is only needed later in the infection cycle, this could be another reason for the re-activation of the C sense promoter in the late G2 phase.

Although Rep expression would be expected to rise with the increase in copy number of its gene due to viral replication, Rep transcripts are relatively rare in infected cells (Wright *et al.*, 1997). This implies that Rep expression is tightly controlled. It is known that begomoviral Rep represses its own promoter (Eagle *et al.*, 1994). Since MSV Rep binds near the C sense TATA box (the C complex), it has been suggested that Rep binding in this area interferes with initiation of Rep transcription, creating a negative feedback mechanism (Arguello-Astorga *et al.*, 1994a, 1994b). Thus, even though the Rep gene copy number increases exponentially during replication of the

genome, Rep expression is kept fairly constant. However, there is no direct evidence that mastreviral Rep does autoregulate its expression.

RepA has also been implicated in the down-regulation of viral replication (Collin *et al.*, 1996; Liu *et al.*, 1998). Since Rep and RepA have different functions in the mastrevirus life cycle, control of the infection process may be achieved through altering the relative proportions in which Rep and RepA are expressed. This may occur in several ways. Differential splicing is one obvious method of altering the ratios of Rep:RepA. The unspliced transcripts from which RepA is produced comprise approximately 80% of the total C sense transcripts in MSV-infected maize tissues (Wright *et al.*, 1997). Expression of Rep and RepA may also be influenced by the size of the C sense transcript produced: a long transcript (terminating in the SIR) can be translated to produce either Rep or RepA, while a short transcript (terminating in the C2) can only be translated to produce RepA. Control of Rep/RepA levels may also be achieved at the level of transcription, depending on which of three TATA boxes the transcripts are initiated from. For example, transcripts initiated from the two TATA boxes closest to the Rep/RepA start codon are predominantly of the shorter type that are capable of producing only RepA (M.I. Boulton, personal communication).

Regulation of the infection cycle may also be achieved through control of the relative levels of CP and MP expression. Enhancement of the expression of CP relative to MP is achieved in three ways: (1) activation of the CP promoter by RepA; (2) expression of CP from both spliced and unspliced versions of a long and short transcript (with particularly high expression from the short, most abundant transcript), while MP can only be expressed with similar efficiencies from the two unspliced transcripts (Wright *et al.*, 1997); and (3) splicing of the MP transcript, which not only has the effect of decreasing the amount of transcript available for MP expression, but splicing of the MP intron may also greatly enhance CP expression (Wright *et al.*, 1997).

The cell-cycle, phase specific expression of MSV Rep and RepA may provide a clue to the mechanism by which these proteins could control the viral life cycle. Although the experiments of Nikovics *et al.* (2001) could not identify differential expression of Rep and RepA, it is likely that regulation of both proteins' expression at different stages of the infection process (and correspondingly at different phases of the host cell cycle) plays a vital role in the co-ordination of the viral life cycle. RepA appears to be required at all stages of the infection process, from preparing the cellular environment early on for replication, to down-regulating expression of Rep

(thus influencing replication levels) later in the cycle, to activating the expression of CP during the later stages of infection. Mechanisms regulating RepA expression or activity may operate at the transcriptional, post-transcriptional (intron splicing) or post translational stage. The latter may include post-translational processing, as well as biological aspects such as the aggregation state of RepA. Thus, it is clear that MSV has evolved sophisticated mechanisms to ensure tight regulation of both V and C sense expression, possibly co-ordinated by RepA.

Late in the replication cycle, ssDNA is removed from the replication pool and accumulated. It is likely that CP plays a role in this, possibly binding to plus strand DNA released during RCR, arresting the synthesis of new RFI DNAs (Donson *et al.*, 1984). An MSV mutant unable to produce the 13 C-terminal amino acids of the CP fails to accumulate ssDNA (Boulton *et al.*, 1989). However, it is unknown whether this is due to lack of ssDNA sequestration within capsids or whether the mutants lack a specific genetic switch that shifts the infection process from replication to ssDNA accumulation (Palmer and Rybicki, 1998). Interestingly, an MSV mutant containing a point mutation within the CP is able to accumulate ssDNA even though it is unable to form geminate particles (Liu *et al.*, 2001). Although this mutant is competent for all other known functions of the CP (including binding ss and ds DNA, and interaction with the MP), it is incapable of systemic infection. This suggests that (a) encapsidation is not necessary for the accumulation of ssDNA, and (b) encapsidation is necessary for long distance movement of MSV in maize.

After ssDNA has accumulated (but probably before encapsidation is complete), it may need to be moved into neighbouring cells. In MSV this appears to involve both CP and MP (Boulton *et al.*, 1989). Although cell-to-cell movement of both ssDNA and dsDNA is believed to occur in some geminivirus species, ssDNA appears to be the most common form in which geminiviruses move their genomes. MSV CP can bind and transport to the nucleus both ssDNA and dsDNA (Liu *et al.*, 1997), suggesting that it may be the functional equivalent of the bipartite begomovirus nuclear shuttle protein (NSP). The NSP, which also localizes to the nucleus and binds both ss and dsDNA, is believed to mediate the movement of ssDNA into and out of nuclei (Ward and Lazarowitz, 1999). However, the ability of the MSV CP to act as a nuclear shuttle protein is unproven. The reason that the CP binds both ss and dsDNA is unclear. One suggestion is that transport of dsDNA from the nucleus of one cell to that of an adjacent cell might provide a convenient transcription template for immediate early gene expression (Boulton, 2002).

The product of the MSV V1 ORF (MP gene) appears to function as a classical MP: it promotes cell-to-cell movement of viral DNA, it localizes to the cell walls and plasmodesmata in infected maize leaves (Dickinson *et al.*, 1996), and it appears to be capable of modifying plasmodesmatal exclusion limits (Kotlizky *et al.*, 2000). MP also appears to interact with a CP:DNA complex to prevent nuclear import of the viral DNA. The experiments of Kotlizky *et al.* (2000) indicate that MP is able to redirect a proportion of CP:DNA complexes from the nucleus to the cell periphery. A possible model for the role of the CP and MP in the late stage of viral infection is that CP, transports viral DNA to the nucleus for replication and transcription, and then accumulates in the nucleus to eventually encapsidate ssDNA, thus sequestering it from the replication pool. When MP interacts with CP:DNA complexes that form prior to encapsidation, the complexes are targeted to the cell periphery. The CP:DNA complex can then be transported through the plasmodesmata and, following release of the MP, be directed to the nucleus of an adjacent cell to begin another round of replication (Boulton, 2002). The reversal in the directionality of transport (exiting rather than entering the nucleus) may be achieved in a number of ways. For example, interaction of MP with CP may mask the CP NLS, or there may be an as yet undiscovered nuclear export signal on the CP. Alternatively, the balance between nuclear import and export may be regulated by post-translational modification of the CP (Boulton, 2002).

To establish a systemic infection a plant virus moves in two phases: cell-to-cell movement via plasmodesmatal connections, and long distance movement as part of the flow of photoassimilates in the plant vascular system, usually the phloem. Although it is likely that intracellular transport and cell-to-cell movement of MSV DNA is in the form of a nucleoprotein complex, it is not clear whether systemic movement relies on normal cell-to-cell movement to deliver genomic DNA into the phloem, or whether viral DNA is specifically packaged for long distance transport. The data of Liu *et al.* (2001) certainly suggest that encapsidation is required for systemic infection.

The mechanisms involved in the encapsidation of geminiviruses are not well understood. There is apparently no encapsidation signal within geminivirus genomes, so it is possible that any circular ssDNA of approximately the right size can be encapsidated (Mansoor *et al.*, 1999; Saunders and Stanley, 1999). MSV plus strands, each with an attached ~80 nt primer molecule (Donson *et al.*, 1984), were shown by Pinner *et al.* (1993) to be packaged into particles that congregate to form large paracrystalline nuclear inclusions. Recently, the structure of MSV particles was resolved by cryo-electron microscopy and 3-D image reconstruction. This showed the MSV particles to consist of two incomplete icosahedra containing 110 copies of the CP arranged with 52-point

symmetry (Zhang *et al.*, 2001). Modelling of the CP revealed that it consists of an eight-stranded antiparallel β -barrel motif, and that the N terminus consists of an α helix containing the putative MSV DNA binding domain. The authors suggest that this region is important for maintaining the geminate particle architecture through interactions with the viral genome (Zhang *et al.*, 2001; Boulton, 2002).

The coat protein is the sole genetic determinant of vector specificity (Briddon *et al.*, 1990). Thus, virus transmission occurs when leafhoppers feed on symptomatic tissues and pick up encapsidated ss MSV DNA, possibly undergoing long distance movement in the phloem. The particles attach at and are transported across the leafhopper hindgut wall, enter the hemocoel and are transported to the salivary glands, thus completing the acquisition process. The leafhopper feeds from the mesophyll and phloem of a healthy plant and delivers the virions to a permissive cell to begin the infection cycle again.

1.3. THE EPIDEMIOLOGY AND CONTROL OF MAIZE STREAK DISEASE

Just as the development of molecular biology techniques in the early 1980s, starting with the cloning and sequencing of geminivirus genomes, led to an intensive era of research on geminivirus molecular biology, the 1990s saw the rapid development of a new age of biotechnology with the potential to revolutionize an important area of research: this was the control of geminivirus diseases. Thanks to the development of genetic engineering, whereby a gene with a desired trait can be transferred into a particular genetic background, the control or even elimination of geminivirus-induced diseases in economically important crops is now viable. However, this cannot be achieved using biotechnology alone. It needs an integrated approach involving, among others, scientists, plant breeders, seed companies, and farmers. For the approach to succeed, a detailed knowledge of the epidemiology of the disease as well as the molecular biology of the viral genome is needed. This section focuses on the epidemiology of maize streak disease (MSD), as well as current and future efforts to achieve its control in maize.

1.3.1 The Epidemiology of Maize Streak Disease

Although geminiviruses can be linked to plant diseases reported over a hundred years ago, it is really only since the early 1990s that these viruses have emerged as a group of serious pathogens that are devastating crops worldwide (Moffat, 1999). Economically important plants as diverse as

tomato, cotton, cassava, wheat, maize, sugarcane, bean, tobacco, beet and horseradish are infected by geminiviruses, sometimes resulting in the destruction of the entire crop. Serious outbreaks of geminivirus-induced disease have occurred in Africa, India, Pakistan, southern Europe, South and Central America, the Caribbean and the USA (Moffat, 1999).

Maize (*Zea mays* L.) was first introduced to Africa in Ghana by Portuguese traders in the 16th century (Gorter, 1953), and has become Africa's most important staple food crop, increasingly replacing traditional food crops such as sorghum and millet. Whereas the worldwide average maize yield is ~4 tons hectare⁻¹ (with highly industrialized nations averaging ~8 tons hectare⁻¹), the average maize yield in Africa is the lowest in the world at ~1.7 tons hectare⁻¹ (Wambugu and Wafula, 1999). Maize pathogens coupled with outmoded agricultural practices are the main reasons for poor yields in Africa. Of the many pathogens infecting maize, MSV is considered the most important and widespread (Thottappilly *et al.*, 1993). Indigenous to sub-Saharan Africa and the neighbouring Indian Ocean islands of Madagascar, Mauritius and La Réunion (Bosque-Pérez, 2000), MSV can result in maize yield losses of up to 100% (Wambugu and Wafula, 1999).

The symptoms of MSD first appear on the lowest exposed portion of the maize leaf as roughly circular spots (Bock *et al.*, 1974), which develop into chlorotic streaks as the leaf expands. Symptoms on mature leaves range from narrow veinal streaks a fraction of a millimeter in width to complete leaf chlorosis. The chlorosis is caused by the failure of chloroplasts to develop in the tissue surrounding the vascular bundles, which results in reduced photosynthesis and increased respiration, leading to a reduction in leaf length and plant height. Symptom severity depends on the MSV strain, the host genotype, and the age of the plant at the time of infection. Highly sensitive varieties or plants infected at an early stage become severely stunted, producing undersized, deformed cobs or giving no yield at all. In the most severe cases, chlorosis of the entire leaf is followed by progressive necrosis and plant death (Bosque-Pérez, 2000).

Many factors determine the severity and frequency of an outbreak of MSD. In particular, the disease is dependent on the complex relationship between the host plant/s (maize, sugarcane and over 80 grass species), a large, unknown number of MSV strains (Konate and Traore, 1992), and the leafhopper vector (*Cicadulina* spp.), nine species of which are able to transmit MSV (Bosque-Pérez, 2000). Environmental factors that have an influence on the leafhopper population also play an important role in MSD epidemiology. For example, MSD outbreaks are often associated with drought conditions followed by irregular rains at the beginning of growing seasons (Efron *et al.*,

1989), as in the savanna regions of West Africa in 1983 and 1984 (Rossel and Thottappilly, 1985), or in Kenya in 1988-89 (Njuguna *et al.*, 1990). The relative abundance of different *Cicadulina* species with differing abilities to transmit the virus in different parts of Africa, is influenced by altitude, temperature, and rainfall (Dabrowski *et al.*, 1987). In addition, late rainfall favours the development of leafhopper nymphs during the winter (Stanley *et al.*, 1998). The interplay of all these factors makes MSD rather erratic, being devastating some years and insignificant in others (Efron *et al.*, 1989).

Although MSV is transmitted by at least nine *Cicadulina* species, *C. mbila* (Naudè) is the species most often implicated in MSD outbreaks (Dabrowski, 1987). This is because *C. mbila* is the most widely distributed species, covering the entire continent of Africa, as well as parts of Asia, Australia, the Indian and Pacific Islands and northern parts of South America (Rose, 1978). Additionally, a larger proportion of *C. mbila* populations have the ability to transmit MSV, compared with other *Cicadulina* species (Storey, 1928, 1933; Markham *et al.*, 1984). This is partly due to the proportion of *C. mbila* females, which are better transmitters, being 2-3 times higher than in other species (Wambugu and Wafula, 1999).

Leafhoppers do not breed on maize; their favoured hosts are annual grasses. Grasses are most probably MSV's natural reservoir sources, from which they "emerged" into maize when the crop was introduced into Africa (Rybicki and Pieterse, 1999). *C. mbila* can feed on more than 138 grass species, ~70% of which are potential MSV hosts (Konate and Traore, 1992). Thus, grass species have a great influence on MSD epidemiology. For example, the species composition and age distribution of grasses in an area may strongly influence the amount of MSV inoculum available for transfer in that area.

When feeding on an infected plant, a leafhopper can acquire the virus by feeding directly on chlorotic lesions and on the phloem where the virus circulates. The length of time required for a leafhopper to acquire MSV while feeding is between 5 and 20 seconds, while the minimum transmission time is between 5 and 10 minutes (Storey, 1925, 1938). The leafhopper becomes viruliferous within 30 hours following acquisition of the virus, although at 30°C the latent period can be as short as 6 to 12 hours (Storey, 1938). While there is no evidence of transovarial transmission of MSV in *Cicadulina* (Storey, 1928), leafhopper nymphs are able to acquire MSV soon after emerging from the egg, and retain the ability to acquire and transmit the virus throughout their lifetime. However, virus titres within a leafhopper decrease over the lifetime of

the insect, indicating that MSV replication does not occur within leafhoppers (Reynaud and Peterschmitt, 1992).

The mechanisms by which MSV particles are transported into the leafhopper's haemocoel and then into its salivary glands are unknown, but they are likely to play a role in determining the ability of the leafhopper to transmit the virus, which is an inherited, dominant sex-linked characteristic (Storey, 1932). Thus, a leafhopper may acquire the virus from the phloem of infected plants, but if the virus cannot, for example, attach to the leafhopper hindgut wall or be transported across the wall to the haemocoel, it cannot be transmitted to another plant. In addition, studies of feeding activities of *C. mbila* on different hosts by Mesfin *et al.* (1995) have revealed vector preferences for certain hosts, which may also play an important role in virus transmission from one host to another. For example, inoculation of MSV into healthy maize occurs when insects salivate into the phloem tissue (Kimmins and Bosque-Pérez, 1996); the time taken to reach the phloem and transmit MSV may take as long as one to three hours from initial access (Bosque-Pérez, 2000). Since leafhoppers making brief probes may not reach the phloem, transmission of MSV thus requires long-duration probing. In this way, the feeding behaviour of *Cicadulina* on a maize genotype has an influence on the resistance of the variety to MSV. On hosts from which the leafhopper does not prefer to feed, the ability to transmit MSV is reduced due to shorter probing times (Bosque-Pérez, 2000).

The flight behaviour of leafhopper populations is another factor that determines the incidence and severity of MSD. Leafhoppers move either within a crop, or migrate from maturing crops or from perennial alternative hosts into younger crop or grass hosts (Bosque-Pérez, 2000). Distinct long and short distance flight morphs have been detected amongst *Cicadulina* populations in Zimbabwe (Rose, 1972), which determine the distance that MSV spreads from the source of inoculum. The long flight morphs, believed to be the migratory form, may play an important part in the long distance spread of virulent MSV variants (Rose, 1978). In turn, migratory movement is probably influenced by environmental conditions such as rainfall and temperature, as well as wind, since *Cicadulina* disperse downwind (Rose, 1972). In Zimbabwe and Nigeria, population densities of *Cicadulina* are low at the onset of the rainy season and then rise gradually as host plants become abundant and succulent. Populations decline sharply in the dry season, probably due to both adverse environmental conditions and an absence of hosts on which to feed (Bosque-Pérez, 2000). Other factors influencing leafhopper flight behaviour are the season, time of day,

gender (females fly further than males), presence of mature ova in the females, age of the leafhopper, and condition of the plants on which the leafhoppers feed (Rose, 1972).

In summary, outbreaks of MSD appear to occur only when favourable weather conditions allow leafhopper survival and population build-up, and where MSV infects both grass and maize hosts. Certain agricultural practices increase the chances of epidemics occurring. During the past 15 years the area of maize cultivation in Africa has greatly increased, and the crop is now often grown as a monoculture (Bosque- Pérez, 2000). In addition, emphasis in the past has been on breeding to improve yields, resulting in high yielding varieties that are also highly susceptible to MSV (Wambugu and Wafula, 1999). Indeed, the introduction of new susceptible genotypes and the increased area under maize are believed to be two of the main factors leading to increased MSV and *Cicadulina* occurrences in Africa (Bosque- Pérez, 2000). Growing crops year-round under irrigation has also contributed to the increased incidence of MSD. For example, in Zimbabwe cereal crops such as wheat are cultivated in the dry season, serving as a host for leafhoppers that later move to early-planted maize (Rose, 1973). Similarly, in many parts of Africa (especially West Africa) maize is grown all year round. Both instances provide year-round food and suitable oviposition sites for leafhoppers. The continuous presence of these crops and the grasses associated with them has serious implications for MSD epidemiology (Bosque-Pérez, 2000). The following section discusses the combined strategies required for the effective, durable control of MSD. These include changing cultural practices, and creating resistant cultivars, either by classical cross-breeding or through genetic engineering, or a combination of both.

1.3.2 Strategies for the Control of Maize Streak Disease

Historically, attempts at control of MSD have focused on evasive measures and the breeding of maize for naturally occurring resistance. Evasive measures include control of the leafhopper population using insecticides, and various cultural practices. Insecticides have been largely unsuccessful due to the need for frequent spraying, which is not only ecologically undesirable, but is expensive and can lead to insecticide resistance within leafhopper populations (Stanley *et al.*, 1998). Cultural practices suggested for control include barriers of bare ground between early- and late-planted maize fields to reduce leafhopper movement and subsequent spread of MSV (Bosque- Pérez, 2000), avoiding maize plantings downwind from older cereal crops, and the use of crop rotations that will minimize invasion by viruliferous leafhoppers (Rose, 1978). Of the

traditional control measures, resistance breeding is perceived as the most practical solution for the control of MSV.

1.3.2.1 Classical cross-breeding for MSV resistance

Resistance in maize was noted as early as 1931 in South Africa, in the variety "Peruvian Yellow" (Fielding, 1933), and several other varieties have since been found to have varying degrees of resistance (or tolerance). Resistance usually manifests itself as reduced symptom severity combined with low virus titres, leading to low virus incidence in the field. Resistant varieties are therefore much poorer sources of inoculum during secondary disease spread (Rodier *et al.*, 1995). In some cases resistant varieties yield well even when infected (Bosque-Pérez, 2000).

Several national programmes in Africa (including South Africa, Zimbabwe, Nigeria, Kenya, and La Réunion) are breeding for resistance to MSV, combining the resistance with other desirable characteristics (Rybicki and Pieterse, 1999). For example, Nigeria's International Institute of Tropical Agriculture (IITA), in cooperation with Zimbabwe's International Maize and Wheat Improvement Center (CIMMYT) and National Programs in Africa, have incorporated MSV resistance into high-yielding varieties and varieties traditionally grown in various African countries (Efron *et al.*, 1989). In addition, the Pannar seed company of South Africa has developed and released MSV-resistant hybrids in several African countries (Bosque-Pérez, 2000).

Despite these extensive efforts to breed resistant maize, there has been only limited success in the field (Stanley *et al.*, 1998). Although in certain countries such as Nigeria there are examples of MSV-resistant varieties being largely unaffected by MSV over many years (Bosque-Pérez, 2000), there are also a number of reports of severe infection of so-called MSV-tolerant maize (Stanley *et al.*, 1998). This occurs particularly when varieties are grown under environmental conditions different from those in which the plants were selected. In addition, some maize varieties known to be resistant elsewhere are susceptible to certain viral strains/isolates, as reported in La Réunion (Rodier *et al.*, cited by Bosque-Pérez, 2000). In Kenya, despite having one of Africa's leading national agricultural research institutions, the Kenya Agricultural Research Institute (KARI), in which plant pathologists, entomologists and plant breeders have worked on MSV for two decades, there has recently been an escalation in MSD incidence (Wambugu and Wafula, 1999). For example, whereas in 1978 a survey revealed low and sporadic incidence of MSD in maize farms in 34 districts in Kenya, in 1988 infection rates as high as 70-80% were reported in the Central Highlands, and in 1998 a survey of ten farms in southwest Kenya showed a MSD

incidence of 80-100% (Wambugu and Wafula, 1999). The increase of MSD incidence over the years could be due to several factors. Possibilities include the emergence of new, more virulent MSV strains, an increase in MSV-susceptible maize varieties grown by farmers, and the breakdown of resistance in the "resistant" varieties. New MSV strains in a particular area could arise from mutation or recombination of the viral genome, or migration via the leafhopper, and have the potential to break the resistance developed against less virulent strains. Natural inherited virus resistance in plants is thought to be due to one or more viral genes (e.g. the genes encoding the coat protein, replicase, or movement protein) encoding an avirulence factor that elicits resistance controlled by a cognate dominant host gene. Some resistance-breaking virus variants have merely a single nucleotide replacement in their avirulence gene (Harrison, 2002). The probability of a resistance-breaking variant appearing depends on the type of resistance and the number of resistance genes to be overcome. The fact that the maize genes involved in MSV resistance are poorly understood makes it difficult to create a variety that is resistant to a broad range of viral strains, especially when the virus frequently mutates and recombines (D. Martin, pers. comm.). Thus, it remains to be seen whether the MSV resistance of commercially available maize varieties is durable and effective over a long period.

Even if long-term resistance can be achieved by classical cross-breeding techniques, there are drawbacks to conventional breeding programmes, which often prove to be difficult and time-consuming. One of the major constraints facing breeders is the need to maintain crop quality and yield while introducing resistance traits (Frischmuth and Stanley, 1993). Genetic resistance is not usually readily available in the desired cultivar, and as mentioned previously, is not necessarily effective against all the strains of a given virus. Resistance genes derived from other plant lines or species are difficult to transfer because they are usually associated with undesirable characteristics, or they can be polygenic in nature (Hemenway *et al.*, 1989).

Modern gene transfer techniques are faster and more precise than classical plant breeding (Moffat, 1999). In recent years the transformation of many plant species has become routine, and genetic engineering is becoming an important part of plant breeding programmes. There are two main advantages of genetic engineering. These are: (1) the ability to transfer single genes directly without linkage to undesired genes, and (2) the ability to construct novel genes that are unlikely to exist in nature (Gadani *et al.*, 1990). This makes it easier to design a resistance strategy that is more likely to succeed against a broad range of virus strains, and that is less likely to be broken than natural resistance. Using genetic engineering, a number of promising strategies (discussed in

the following section) have been developed to introduce geminivirus resistance into economically important crops. Most work has focused on begomovirus diseases due to the recalcitrance of cereals to transformation, but the principles can be readily applied to mastrevirus diseases such as MSD.

1.3.2.2 Genetic engineering of plants for geminivirus resistance

The first approach to genetically engineering plants for virus resistance was to mimic the natural phenomenon of "cross protection", first observed over 70 years ago by McKinney (1929), who showed that plants already infected with a virus are normally protected against infection by a related strain of the virus. Thus, plants infected with a mild strain could be protected against infection by severe isolates or strains of that virus. There were a number of reasons for the limited use of this form of protection, including the impracticality of inoculating plants with a live virus on a large scale, and the possibility that the mild strain might mutate to a more virulent form. In 1985, Sanford and Johnson used the principle of cross protection to develop the concept of parasite- or pathogen-derived resistance (PDR). A simple but elegant concept, PDR is the process whereby resistance to a pathogen may be engineered by transforming a susceptible plant with a gene derived from the pathogen itself. In general, PDR is thought to operate through the expression of the viral gene product at either an inappropriate time, in inappropriate amounts, or in an inappropriate form during the infection cycle, thereby perturbing the ability of the pathogen to sustain an infection (Lomonossoff, 1995). The first demonstration of virus-derived resistance in transgenic plants made use of the coat protein (CP) gene of *Tobacco mosaic virus* (TMV; Powell-Abel *et al.*, 1986). CP-mediated protection, which has subsequently been used successfully against a number of viruses, is thought to operate through the inhibition of virion disassembly in the initially infected cells (Baulcombe, 1996).

Other viral genes shown to be capable of conferring PDR are the replicase gene, first demonstrated with TMV and subsequently found to be effective with numerous other viruses, and the movement protein gene from tobamoviruses, bromoviruses and potexviruses (Lomonossoff, 1995). The PDR concept has been applied mainly to viruses with positive-strand RNA genomes, and most examples of CP-mediated resistance are based on the transgenic expression of wild-type genes. Although there is one report of wild type CP-mediated resistance against a geminivirus (TYLCV; Kunik *et al.*, 1994), this is not considered to be the best strategy to engineer resistance against geminiviruses. However, other effective strategies are based on variations of the PDR concept, including the transformation of susceptible plants with viral genes containing dominant

negative mutations (Herskowitz, 1987), defective interfering DNA derived from the viral genome, and antisense viral RNA. There are also examples of PDR occurring as a result of gene silencing (Baulcombe, 1996). The following are examples of the successful application of engineered resistance strategies against geminiviruses.

(1) Defective Interfering DNA

Subgenomic DNA molecules that are related to and dependent on the parent virus for their proliferation have been found associated with geminivirus infection. Because some subgenomic DNAs have the ability to delay and attenuate infection symptoms, they are referred to as defective interfering (DI) DNA (Frischmuth *et al.*, 1997). Transgenic *N. benthamiana* plants containing tandemly repeated, integrated copies of ACMV (Stanley *et al.*, 1990; Frischmuth and Stanley, 1991) or BCTV (Frischmuth and Stanley, 1994) DI-DNAs exhibit ameliorated symptoms when challenge-inoculated with the homologous parent virus. The mechanism of resistance is thought to occur through the mobilization of extrachromosomal copies of DI DNA following virus infection, through Rep-mediated release. Subsequent episomal amplification of the DI DNA occurs at the expense of the genomic viral DNA, resulting in reduced virus amplification and symptom amelioration (Stenger, 1994). This strategy is limited, since it relies on the ability of the infecting virus to replicate the DI DNA. Thus, plants show resistance only to closely related strains of virus from which the DI DNA is derived (Stanley *et al.*, 1998).

(2) Antisense RNA

RNA molecules capable of annealing to a given mRNA are a means of natural and artificial gene regulation by silencing the expression of the corresponding gene (Bendahmane and Gronenborn, 1997). This phenomenon has been successfully exploited to target and selectively suppress the expression of specific genes, in both therapeutics and in the prevention of viral diseases. It is likely that antisense RNAs anneal with sense RNAs to form a double strand complex, which is rapidly degraded or which inhibits the translation of the RNA. The nuclear-replicating geminiviruses are potentially promising targets for antisense RNA-mediated suppression, since the antisense sequences do not have to be directed to the cytoplasm as they do with cytoplasmic RNA viruses (Frischmuth and Stanley, 1993).

Antisense RNAs are usually targeted at the C sense genes that not only are indispensable for viral replication and required early in the infection cycle, but are also expressed from transcripts of relatively low abundance. These factors make the Rep gene an ideal target for obtaining

geminivirus resistance by suppressing its expression. This approach has met with some success in TYLCV- and TGMV-susceptible plants. In transgenic *N. tabacum* plants expressing the TGMV Rep antisense RNA, symptoms were less severe compared with non-transformed control plants when challenged with TGMV (Day *et al.*, 1991). In addition, most antisense lines contained fewer symptomatic plants than control plants and transgenic lines expressing Rep sense RNA. Interestingly, Bejarano and Lichtenstein (1994) showed that TGMV antisense RNA was also effective against BCTV, but not ACMV. The authors concluded that a minimal contiguous complementarity between the antisense RNA and the mRNA target is required for efficient suppression. *N. benthamiana* plants that are less susceptible to TYLCV infection have also been produced by expressing TYLCV Rep antisense RNA (Bendahmane and Gronenborn, 1997). Some of the resistant lines were symptomless, and the replication of TYLCV almost completely suppressed. These reports suggest that antisense RNA is a promising resistance strategy, but the effectiveness of this approach against other geminiviruses depends on the level of homology between the target sequence and the antisense RNA (Stanley *et al.*, 1997).

(3) Expression of Viral Proteins

The expression of geminivirus proteins or their derivatives in transgenic plants has conferred resistance against a number of geminiviruses. An ideal candidate for interference with virus infection is Rep, which is multifunctional, required early in the infection cycle and expressed at low levels from the Rep promoter. However, a novel resistance strategy based on the suppression of virus movement was developed for bipartite geminiviruses by von Arnim and Stanley (1992a, b). They found that the TGMV MP, which cannot complement movement of ACMV, specifically inhibits ACMV systemic spread. It is likely that the ACMV MP functions as a multimeric component of a movement complex that is disrupted by the inclusion of one or a limited number of copies of the inhibitory protein (Frischmuth and Stanley, 1993). A variation of this strategy was employed by Duan *et al.* (1997), who transformed *N. tabacum* with a mutated MP gene from *Tomato mottle virus* (TMoV). Transgenic plants expressing the defective MP showed resistance to both TMoV and *Cabbage leaf curl virus* (CabLCV). The degree of resistance correlated with the level of expression, suggesting that the defective protein functions as a dominant negative mutant of a movement function. In a similar experiment, Hou *et al.* (2000) found that tomato plants transformed with a mutated *Bean dwarf mosaic virus* (BDMV) MP showed a delay in ToMV infection compared with non-transformed plants. The fact that the resistance in the cases of both Duan *et al.* (1997) and Hou *et al.* (2000) extended to heterologous geminiviruses, suggests that this approach may result in broader spectrum resistance than strategies that target

viral DNA replication (Frischmuth and Stanley, 1993). However, expression of wild type MP, which is a pathogenicity determinant, can have deleterious effects of various aspects of plant development (Covey and Al-Kaff, 2000; Hou *et al.*, 2000), necessitating the use of a defective MP transgene in order to regenerate phenotypically normal plants.

One of the first steps in virus multiplication is the replication of the viral genome. Therefore, blocking this phase should be one of the most efficient ways to protect plants from geminivirus infection. Noris *et al.* (1996) and Brunetti *et al.* (1997) showed that expression of a truncated TYLCV Rep gene (T-Rep), in transgenic *N. benthamiana* and tomato respectively, interfered with TYLCV infection. However, plants expressing TYLCV T-Rep were not protected against ToLCV, suggesting that the resistance mechanism is specific. Brunetti *et al.* (2001) determined that T-Rep acts as a *trans*-dominant negative mutant, inhibiting both viral transcription and replication. They proposed a model of TYLCV resistance conferred by T-Rep. Initially, T-Rep inhibits but does not abolish Rep transcription by recognizing and binding to the cognate transcriptionally active RF DNA. The limited amount of newly synthesized Rep cannot properly synthesise the viral plus strand since it is out-competed by T-Rep for utilization of the required sequence; as a result, virus replication is inhibited.

Hong and Stanley (1996) used a similar approach in conferring resistance to ACMV in *N. benthamiana*. They found that transient expression of ACMV Rep or a truncated N-terminal portion of the protein caused a significant reduction in the level of viral DNA in *N. tabacum* protoplasts. Subsequent transformation of *N. benthamiana* with the full-length Rep gene resulted in plants that were less susceptible to infection by ACMV. None of the transformed lines showed resistance to TGMV or BCTV, again demonstrating the specific nature of the resistance mechanism. The authors suggest that the transgenic Rep protein, which is highly expressed by the CaMV 35S promoter, could affect viral DNA replication by disturbing the equilibrium between monomeric and multimeric forms of the viral Rep complexes. Resistance to ACMV has also been achieved in *N. benthamiana* by the high-level expression of the Rep gene containing a mutation of the NTP-binding domain (Sangaré *et al.*, 1999).

Transient expression of a truncated ToLCV Rep protein was found to inhibit homologous viral DNA accumulation in tobacco protoplasts and in *N. benthamiana* plants (Chatterji *et al.*, 2001). This protein, which contains the N-terminal 160 amino acids of Rep and therefore the sites for DNA cleavage, DNA binding, and protein oligomerization, was shown to interfere with DNA

binding and oligomerization activities during virus infection. Surprisingly, the truncated protein also reduced accumulation of *Pepper huasteco yellow vein virus* (PHYVV) and *Potato yellow mosaic virus* (PYMV), although to a lesser extent than ToLCV. The truncated ToLCV Rep formed oligomers with the Rep proteins of the heterologous geminiviruses, which suggests that it may function as a dominant negative mutant, interfering with one or more of the multiple functions of the wild type Rep oligomers.

As well as published reports, there have been several patent applications relating to begomovirus resistance in transgenic plants conferred by Rep genes mutated in one or more of the conserved motifs I-IV (Hanson *et al.*, 1998; Gronenborn, 2000, Stout *et al.*, 2001; Hanley-Bowdoin *et al.*, 2002a) as well as a mutant form of the begomoviral accessory replication protein, REn (Hanley-Bowdoin *et al.*, 2002b). In general, the mutant Rep (or REn) in all these cases may interfere with the replication activity of the wild type protein expressed by infecting begomoviruses, in this way behaving as dominant negative mutants.

(4) Virus-induced Cell Death

Using a mechanism involving a viral-activated antiviral protein, Hong *et al.* (1996) conferred resistance to ACMV in *N. benthamiana* by the expression of a ribosome inactivating protein (RIP) from the ACMV V sense promoter. The RIP used was dianthin, a potent plant cytotoxin, which rapidly kills infected cells, thereby containing the infection. This resistance mechanism depends on the transactivation property of TrAP, which induces the expression of the ACMV CP. Thus, expression of dianthin under the control of the CP promoter should be activated specifically once TrAP is expressed in virus-infected cells, avoiding constitutive expression of the RIP and ensuring transgene expression is localized to virus-infected cells. Hong *et al.* (1996) found that plants containing the dianthin gene were less susceptible to virus infection and accumulated only low levels of viral DNA. A drawback to this method is that a low level of constitutive expression from the CP promoter can occur in the absence of TrAP. Another is that the resistance was confined to ACMV isolates, implying that TrAP activity is virus-specific (Stanley *et al.*, 1998)

(5) Gene Silencing

The examples of PDR discussed above are those in which expression of the transgene leads to the resistance phenotype. However, in some types of PDR gene silencing is involved in and is responsible for the resistance mechanism (Baulcombe, 1996). Gene silencing is a cellular

mechanism that targets specific nucleic acid sequences for down-regulation or degradation (Covey and Al-Kaff, 2000). First observed in plants containing transgenes, gene silencing can operate both at the transcriptional and post-transcriptional levels. Transgene-associated post-transcriptional gene silencing (PTGS) targets the transgene RNA and RNA from homologous endogenous genes for degradation in the cytoplasm (Covey and Al-Kaff, 2000). Experiments with transgenic plants expressing viral sequences led to the discovery of virus-associated PTGS, whereby, upon infection by the virus, the virus-derived transgene leads to the suppression of expression of both the transgene and the homologous viral gene. If the virus gene is an essential component of its lifecycle, for example the Rep gene in the case of geminiviruses, PTGS of the gene leads to virus inactivation. Although most cases of genetically engineered geminivirus resistance have been associated with high levels of transgene expression and transgene product, there are cases of transgene-mediated resistance to geminiviruses that may have been mediated at the RNA level. For example, resistance to TMoV was found in transgenic plants expressing the TMoV CP gene with a 5' deletion of 30 bp (Sinisterra *et al.*, 1999, cited by Covey and Al-Kaff, 2000), but the resistance did not appear to be the result of the classical CP-mediated protection. In those plants where a resistance phenotype was observed, transgene RNA, but no transgene protein, was detected. The authors concluded that this resistance may have been mediated at the RNA level, although specific characteristics typical of PTGS, such as co-reduced levels of transgene and viral RNA, were not determined. Although there is not much evidence for gene silencing elicited by transgenic expression of individual geminiviral genes or fragments, geminivirus replicons or vectors do have the ability to induce PTGS. For example, when Kjemtrup *et al.* (1998) used TGMV as a vector to carry a foreign sequence in place of the CP gene, it triggered gene silencing of both the vector sequence and the homologous endogene when introduced into *N. benthamiana*. Gene silencing elicited from an integrated geminiviral vector releasing the mastrevirus TYDV (an autonomously replicating multicopy plant episome, or MPE) has also been demonstrated (Atkinson *et al.*, 1998).

Gene silencing appears to be one of several host mechanisms operating against viruses. Since it is clear that geminiviruses can elicit PTGS targeted against viral RNAs, it could be possible to trigger PTGS targeted against an essential geminiviral RNA, leading to viral inactivation. Thus, virus-derived sequences could be used as transgenes to provide resistance based on PTGS. However, there are some drawbacks to this resistance mechanism. Because gene silencing is homology-dependent, there is likely to be only a limited amount of sequence non-identity between the transgene and the challenge virus that can be tolerated by the silencing mechanism. It

is also becoming clear that some viruses, including geminiviruses, contain anti-silencing genetic determinants, which lead to the suppression of gene silencing (Kasschau and Carrington, 1998; Voinnet *et al.*, 1999). One such determinant is the ACMV AC2 gene, encoding TrAP. The anti-silencing determinants associated with AC2 were assayed in transgenic plants containing a green fluorescent protein (GFP) transgene that had been systemically silenced by PTGS. Subsequent infection by a potato virus X (PVX) vector expressing the ACMV AC2 gene led to suppression of silencing of the GFP gene (Voinnet *et al.*, 1999).

It is clear from the above example that there is a more balanced interaction between plants and viruses than previously thought, involving host defense and viral counter-defensive strategies. The best, most durable viral resistance is therefore likely to have the following characteristics: (1) protection against a wide range of viral strains; (2) multiple mutations (preferably more than two) in the viral genome required to overcome host resistance, and (3) a resistance mechanism that confines the virus to the inoculated cell (Harrison, 2002). Examples of the latter include blocking replication or cell-to-cell movement of the virus, or inhibiting the expression of a protein required early in the infection cycle by an RNA-mediated resistance mechanism such as antisense RNA or PTGS. Development of this sort of resistance, together with sound crop management as discussed at the beginning of this section, may be the only long-term means of successfully combating MSD.

1.4 PROJECT AIMS

The main aim of this project was the development of a genetic engineering approach, to ultimately result in the development of MSV-resistant maize. The strategy chosen was that of pathogen-derived resistance (PDR), using the multifunctional viral Rep gene as a target for mutagenesis and truncation. Rep constructs had previously been made containing deleterious mutations in several conserved motifs and had been shown in transient expression assays to be effective in inhibiting MSV replication. In this study, the first aim was to truncate these mutants as well as the wild type Rep gene, in order to remove key motifs involved in viral replication and in the interaction with host factors that could potentially interfere with plant development. A three-step approach was designed to test the effects of these constructs on viral replication. First, truncated and mutated Rep constructs would be co-bombarded with a widespread representative MSV strain (MSV-Kom) into black Mexican sweetcorn (BMS) suspension cells, and the effects of the constructs on viral replication determined by developing a quantitative PCR assay. Second,

Rep constructs that inhibited viral replication in BMS would be used to transform an MSV-sensitive grass, *Digitaria sanguinalis*. Because of the recalcitrance of maize to tissue culture, *D. sanguinalis*, which is easily transformable and regenerable, is useful as a model system to test the effect on viral replication of the transgenic expression of a range of mutated or truncated Rep proteins. To do this, transgenic plants would be challenged with MSV using viruliferous leafhoppers, the vectors of the virus. The third step of the approach, assuming MSV-resistant *D. sanguinalis* was produced, would be to transform the maize cultivar High Type II (Hi-II) with Rep constructs that resulted in the best, most durable resistance in *D. sanguinalis*.

In the process of carrying out the above transient and transgenic assays, some interesting results were obtained with some of the mutated and truncated Rep constructs, which had unexpected effects on viral replication when co-bombarded with MSV-Kom into BMS. As a sideline to the main aim, a particular focus was a Rep gene containing a mutation in the retinoblastoma-related (RBR) protein interaction domain, which contrary to published reports at the time, surprisingly supported high-level viral replication in BMS, while *in planta* one of the nucleotides of the three-nucleotide mutation reverted at an extremely high frequency. The remaining aims of the project were to carry out a study of the dynamics of the mutant and revertant viral populations in maize, and to determine the selective advantage of the single nucleotide reversion.

Chapter 2

Inhibition of *Maize streak virus* (MSV) Replication by Transient and Transgenic Expression of MSV Replication-Associated Protein Mutants

Contents

ABSTRACT	63
2.1 INTRODUCTION	63
2.2 MATERIALS AND METHODS.....	66
2.2.1 Clone Construction.....	66
2.2.2 Yeast Two-Hybrid Analysis of the RBR Protein Interaction Properties of Mutant, Truncated and Wild Type <i>Rep</i> Constructs ...	71
2.2.3 Transient Replication Assays	71
2.2.4 <i>Digitaria sanguinalis</i> Tissue Culture.....	73
2.2.5 Transgene Expression Analysis.....	75
2.2.6 Test for Resistant <i>D. sanguinalis</i> by Challenge with MSV	76
2.2.7 Maize Tissue Culture	77
2.2.8 Analysis of the MSV Strain Composition Transmitted by Viruliferous Leafhoppers.....	78
2.3 RESULTS AND DISCUSSION.....	79
2.3.1 Interactions between MSV Rep and RBR Proteins.....	79
2.3.2 The Effects of Transiently Expressed Rep Proteins on MSV Replication in Black Mexican Sweetcorn (BMS).....	80
2.3.3 Transformation of <i>D. sanguinalis</i>	84
2.3.4 Transformation of Maize HiII.....	88
2.3.5 Challenge of <i>D. sanguinalis</i> Transgenic Plants with MSV	91
2.3.6 Expression of Transgenes in <i>D. sanguinalis</i> Lines used in MSV- Challenge Experiments	100
2.3.7 Analysis of the MSV Strain Composition Transmitted by Viruliferous Leafhoppers.....	102
2.4 CONCLUSION	104

ABSTRACT

Maize streak virus (MSV), the type member of the *Mastrevirus* genus of the *Geminiviridae*, is the worst viral disease-causing agent affecting maize in Africa, resulting in crop yield losses of up to 100%. To genetically engineer MSV-resistant maize using the pathogen-derived resistance (PDR) strategy, the viral replication-associated (Rep) protein gene was targeted, whose multifunctional products are the only viral proteins essential for replication. *Rep* constructs had previously been made containing deleterious mutations in several conserved motifs. In this study, these mutant *Reps* and the wild type *Rep* gene were truncated to remove key motifs involved in viral replication. A quantitative PCR assay was developed to determine the effects of the mutant and truncated *Reps* on viral replication in black Mexican sweetcorn (BMS) suspension cells. The MSV-sensitive grass *Digitaria sanguinalis* was then transformed with *Rep* constructs that inhibited MSV replication in BMS, and transgenic lines were tested for virus resistance. Several plants of a *D. sanguinalis* line transgenic for a mutated full-length *Rep* gene showed excellent resistance (immunity) to MSV, but the transgene had negative effects on aspects of plant growth and development. Transformation with a mutated/truncated *Rep* gene, however, resulted in healthy fertile transgenic *D. sanguinalis* plants, many of which showed good MSV resistance in challenge experiments. Resistance phenotypes included a delay in symptom development, a recovery from early symptoms, and an absence of virus symptoms at all stages. Expression of the transgenes in lines that were challenged with MSV was determined by RT-PCR and by histochemical staining for GUS (the transforming plasmid contained the *uidA* gene). The maize cultivar Hi-II was transformed with the mutated/truncated *Rep* gene, and three transgenic lines produced fertile T₁ offspring. Considering the success in achieving MSV-resistant *D. sanguinalis*, there is good reason to believe that the transgenic maize will too be resistant to MSV.

2.1 INTRODUCTION

Maize is Africa's most important staple food crop and is increasingly replacing traditional food crops such as sorghum and millet. In most of sub-Saharan Africa poor maize yields are usually linked to food shortages and famine. Despite being a crucial staple food crop, the average maize yield per hectare in Africa is the lowest in the world. A major contributing factor to these low yields is the causal agent of maize streak disease (MSD), *Maize streak virus* (MSV). Transmitted by leafhoppers (*Cicadulina* sp.), MSV is indigenous to sub-Saharan Africa and neighbouring Indian Ocean Islands.

Methods of reducing yield losses caused by MSV include the use of insecticides to control leafhopper populations, and the use of MSV-resistant maize genotypes. Insecticides are expensive and are beyond the means of most farmers in Africa. Frequent spraying of insecticides is undesirable as it may lead to insecticide resistance, and is ecologically unfavourable. Conventional breeding programmes are protracted, and there has been limited success in combining economically important traits such as yield with resistance characteristics. Another important drawback to conventional breeding is that the resistance can be circumvented by virus variation.

A more efficient, cost-effective and sustainable solution could be the development of MSV-resistant maize by genetic engineering. Transformation of plants with viral genes can give rise to lines of plants that are resistant to the virus from which the sequence was derived. Although this phenomenon, termed "pathogen derived resistance" (PDR), has been successfully applied for resistance to viruses of the *Begomovirus* genus of the *Geminiviridae*, there are no published reports of any transgenic plants resistant to mastreviruses using this or any other approach. With begomoviruses, coat protein-mediated protection against *Tomato yellow leaf curl virus* (TYLCV) has been achieved in tomato (Kunik *et al.*, 1994); interference with *African cassava mosaic virus* (ACMV) replication by defective-interfering viral genomes in transgenic plants has been reported (Stanley *et al.*, 1990; Frischmuth and Stanley, 1991), and defective movement protein has resulted in resistance to *Tomato mottle virus* (ToMoV, Duan *et al.*, 1997) and *Bean dwarf mosaic virus* (BDMV; Hou *et al.*, 2000). Other approaches for interfering with begomovirus replication have included expression of the antisense RNA of *Tomato golden mosaic virus* (TGMV, Day *et al.*, 1991; Bejarano and Lichtenstein, 1994) and TYLCV (Bendahmane and Gronenborn, 1997), and truncations of the viral replication-associated protein (Rep). There are two reports of resistance being achieved (although limited) with the latter approach, against TYLCV (Noris *et al.*, 1996) and ACMV (Hong and Stanley, 1996). In addition, a mutant viral Rep protein was used by Sangare *et al.* (1999) to develop resistance against ACMV in tobacco.

To genetically engineer MSV-resistant plants, the strategy employed was that of virus-derived resistance, by means of dominant negative mutant interference (Herskowitz, 1987) with virus replication. The viral gene chosen for mutagenesis and truncation was *Rep*, whose multifunctional products, Rep and RepA, are the only viral proteins essential for replication.

MSV replicates in the nucleus by a rolling circle replication (RCR) mechanism (Saunders *et al.*, 1991; Stenger *et al.*, 1991). This is initiated by the binding of Rep to the origin of replication, where the protein introduces a sequence-specific nick in the loop of a stem-loop structure. Host replication enzymes then complete the RCR process.

Rep is a product of transcription of two open reading frames (ORFs), *C1* and *C2*. The *C1/C2* transcript has an intron, and depending on whether or not it is spliced, expresses either Rep (from the spliced transcript) or RepA (from the unspliced transcript). Rep and RepA have several distinct domains with diverse biochemical activities (see Fig. 2.1), among them DNA-binding, nicking-joining, transactivation (all activities shared by Rep and RepA), interaction with the host retinoblastoma related (RBR) protein (RepA only), and NTP-binding/ATPase activities (Rep only).

In a previous study from this laboratory, the following conserved motifs of MSV Rep were chosen as targets for PCR-mediated mutagenesis: motif III, which is essential for nicking and closing the DNA during rolling circle replication; the LxCxE motif of the RBR protein-interaction domain, which may enable the virus to create an optimal cellular environment for virus replication; and a motif in the NTP binding domain, which may be necessary for Rep helicase activities. Single, double (two motifs mutated) or triple (all three motifs mutated) Rep mutants were generated by PCR mutagenesis of the Rep gene (T. Mangwende, 2001). In the present study, mutated and wild type *Rep* constructs were truncated to yield N-terminal fragments missing the entire *C2* ORF and the C terminus of the RepA gene.

To determine the effects of the mutant and truncated Reps on viral replication, a three-system approach was designed. First, a quantitative PCR assay was developed to accurately compare the effect of the mutant and truncated Reps on viral replication in black Mexican sweetcorn (BMS) suspension cells. Second, an MSV-sensitive grass, *Digitaria sanguinalis*, was transformed with *Rep* constructs that had inhibited MSV replication in BMS, and transgenic lines were tested for virus resistance. Third, the maize cultivar Hi-II was transformed with the constructs that resulted in MSV-resistant *D. sanguinalis*.

2.2 MATERIALS AND METHODS

2.2.1 Clone Construction[§]

All *Rep*-based clones were derived from an infectious MSV plasmid, pKom602, which is a partial tandem dimer of the MSV-Kom genome cloned in pUC19 (Schnippenkoetter *et al.*, 2001). The C1/C2 ORFs of pKom602 were amplified by PCR using the forward primer C1: 5' TTAGGATCCCTCAGCCTCAACCCTCC, which introduced a *Bam*HI restriction enzyme site (underlined) 26 bp upstream of the C1 start codon, and the reverse primer C2: 5' ACGCAAACAATACAGGGGGGTAGTC, which binds in the SIR. The PCR product was cloned into the *Bam*HI/ *Hind*III site of pBluescript SK+ (pSK; Stratagene, La Jolla, CA) and all subsequent mutations and truncations were performed on this wild type (wt) *Rep* construct (pSK*Rep*). Unless otherwise stated, all restriction enzymes were obtained from Boehringer Mannheim (Mannheim, Germany). Sequencing of clones was carried out by D. James at the University of Cape Town (UCT), using an ALF Express automated sequencer (Pharmacia Corporation, Peapack, NJ). Sequence analysis was carried out using DNAMAN (version 4.0; Lynnon BioSoft, Quebec). Standard cloning techniques were used as in Sambrook *et al.* (1989).

2.2.1.1 PCR site-directed mutagenesis

As part of a detailed analysis of the effects of changing amino acids within various *Rep* motifs, pSK*Rep* was used as template DNA for mutagenic PCR (T. Mangwende, 2001). In the present study, *Rep* constructs containing novel mutations in motif III and the RBR protein interaction domain of MSV *Rep* were chosen for transformation of plants because they completely abolished MSV replication in black Mexican sweetcorn (BMS). Mutagenic primer sequences and introduced enzyme sites used to detect the mutations are shown in Table 2.1. All mutated clones were confirmed to be correct by sequence analysis.

TABLE 2.1 Mutations introduced into pSK*Rep*.

Mutagenic Primer Sequence (altered nucleotides are shown in lowercase)	Amino Acid Changes	Introduced enzyme site	Mutated clone name
5'-GATTTACTTTGTctTaAGTCAATCAAC-3'	^a LLCN ²⁰¹ E ²⁰² to LLCL ²⁰¹ K ²⁰²	<i>Bfr</i> I	pSK <i>Rep</i> ^{III+Rb+NTP+}
5'-AGAGTGAGGGccTAtTTCTCAAGGAAC-3'	^b VRD ⁹⁹ YI ¹⁰¹ LKE to VRA ⁹⁹ YF ¹⁰¹ LKE	<i>Hae</i> III	pSK <i>Rep</i> ^{III+Rb+NTP+}

^aMutation introduced into the *Rep* RBR-interaction domain. Amino acid numbering is relative to the *Rep* start codon.

^bMutation introduced into the *Rep* motif III. Amino acid numbering is relative to the *Rep* start codon.

[§] Properties of all plasmids and constructs used in Chapters 2 to 4 are summarized in Appendix A.

2.2.1.2 Truncated Rep constructs

To generate C-terminal truncated Rep genes, pSKRep (wt) and pSKRep^{III-Rb+NTP+} (III⁻ mutant) were digested with *Bam*H1/*Hind*III, creating 537-bp N-terminal fragments that were subsequently cloned into the *Bam*H1/*Hind*III site of pSK. The resulting plasmids were called pSKRep^{ΔRbΔC2} and pSKRep^{III-ΔRbΔC2} respectively (see Fig.2.1 for a diagrammatic illustration). In addition, pSKRep and the Rb⁻ mutant, pSKRep^{III+Rb-NTP+}, were subjected to a partial digest with *Bam*H1/*Hind*III, and the resulting 658 bp N-terminal fragments cloned into pSK to create the plasmids designated pSKRep^{ΔC2} and pSKRep^{Rb-ΔC2} respectively (Fig. 2.1). All constructs were confirmed to be correct by sequencing.

2.2.1.3 Construction of plant vectors for *trans*-replication analysis of mutant and truncated Rep proteins

A 1.3-Kb *Bam*HI-*Bgl*II fragment (containing the full-length Rep gene; see Fig 2.1) from each of the plasmids pSKRep^{III-Rb+NTP+} (III⁻ mutant), pSKRep^{III+Rb-NTP+} (Rb⁻ mutant), pSKRep^{III-Rb-NTP+} (III⁻ Rb⁻ double mutant) (T. Mangwende, 2001) and pSKRep (wt Rep gene; this thesis) was cloned into the *Bam*HI site of a 5' dephosphorylated plasmid, pAHC17 (Christensen and Quail, 1996). The resulting plasmids were designated pRep^{Mut}, where Mut = the III⁻, Rb⁻, and III⁻Rb⁻ mutations (T. Mangwende, 2001), or pRep (wt Rep). To create an antisense Rep gene, pSKRep was cut with *Bam*HI and *Bgl*II to release the full-length Rep gene, and cloned into the *Bam*HI site of pAHC17 in the antisense orientation (designated pRep^{III+Rb+NTP+(AS)}; T. Mangwende, 2001). To clone the truncated Rep genes into pAHC17 (this thesis), a *Bam*HI site was inserted at the C-terminus of the truncated genes. The truncated genes were then cut with *Bam*HI, and inserted into the same site of pAHC17. Vectors with the Rep genes in the sense orientation were selected, and designated pRep^Δ or pRep^{MutΔ}, where Δ= a deletion, and Mut = Rb⁻ or III⁻ mutations. To assay the effect of truncated antisense Rep genes on viral replication, Rep^{ΔC2} and Rep^{ΔRbΔC2} cloned in the *Bam*HI site of pAHC17 in the antisense orientation were selected and designated pRep^{ΔC2(AS)} and pRep^{ΔRbΔC2(AS)} respectively.

In most cases, transformation of *D. sanguinalis* was carried out by co-bombardment of pRep^{Mut} or pRep^{III+Rb+NTP+(AS)} or pRep^{MutΔ} (all in pAHC17) with pAHC25, which contains the bar (bialophos resistance) gene and the GUS (*uidA*) gene under the control of separate maize ubiquitin promoters (Ubi-Bar/ Ubi-Gus; Christensen and Quail, 1996). However, with the co-bombardment strategy, both plasmids have to integrate into the genome of one cell to regenerate bialophos resistant/ Rep transgenic plants. To dispense with co-bombardment, the GUS gene in

pAHC25 was replaced with Rep^{III-Rb-NTP+}. The Rep gene was isolated from pRep^{III-Rb-NTP+} as a *Pst*I (whose sites flank the Rep gene in pAHC17) fragment, and the GUS gene was excised out of pAHC25 with *Sma*I and *Sac*I. The Rep^{III-Rb-NTP+} fragment was then blunt-cloned into the *Sma*I/*Sac*I site of pAHC25. To do this, the *Sma*I/*Sac*I-digested vector and the *Pst*I-Rep fragment were first treated with Klenow DNA polymerase, followed by dephosphorylation of the vector with SAP (shrimp alkaline phosphatase). The Rep^{III-Rb-NTP+} fragment was then ligated into the dephosphorylated/ blunt ended vector. Rep^{III-Rb-NTP+} in the sense orientation was selected, and designated pAHCRep^{III-Rb-NTP+} (Ubi-Bar/ Ubi-Rep).

2.2.1.4 Construction of intronless Rep and RepA

pSKRep was used as a template to create an intronless Rep gene, by inverse PCR. The forward primer Rep Δ I-F (5' -TCATCAGATGAAAGATCAAGAAAGC - 3') amplified the Rep gene from the 3' end of the intron through the C2, while the reverse primer Rep Δ I-R (5'-CTGGAAGATGTTAGGCTGGAGCC - 3') amplified the gene from the 5' end of the intron through the C1. In this way the whole template plasmid was amplified, minus the intron. The PCR product was self-ligated to create the intronless plasmid pSKRep Δ I, which was then confirmed to be correct by sequencing. The same procedure was followed to create an intronless RepRb⁻ gene, this time using pSKRep^{Rb-} as the template, creating the plasmid pSKRep^{Rb-} Δ I.

The RepA gene contains the entire intron at its C-terminus, therefore the first step in making a RepA construct was to prevent the possibility of splicing occurring. This was done using a Rep plasmid (pMB1657, provided by Dr. M. Boulton, John Innes Centre, Norwich, U.K.) with a 3' splice site mutation of A⁷³³G⁷³⁴ to T⁷³³C⁷³⁴ (Wright *et al.*, 1997). A fragment containing the splice site mutation was excised from pMB1657 with *Xho*I and *Bgl*III, and swapped with the same wt fragment from pKom602, resulting in pKomMB1657. Although pMB1657 contains the MSV-Ns strain, the *Xho*I/*Bgl*III fragment is 100% homologous to the same fragment from MSV-Kom. Therefore, the only mutation introduced into pKom602 by the fragment swap was the desired Rep 3' splice site mutation. The RepA gene was amplified from pKomMB1657 using the C1 forward primer, which introduced a *Bam*H1 site at the 5' end, and a reverse primer (*RepABgl*, provided by D. McGivern, John Innes Centre, Norwich, U.K.), which introduced a *Bgl*III site immediately after the RepA stop codon. The *RepABgl* primer has the following sequence: 5' -TTATAGATCTCTAGGCTTCTGG - 3'. Once amplified, the RepA gene was cloned into the *Bam*H1 site of pSK to form pSKRepA, and subsequently confirmed to be correct by sequencing.

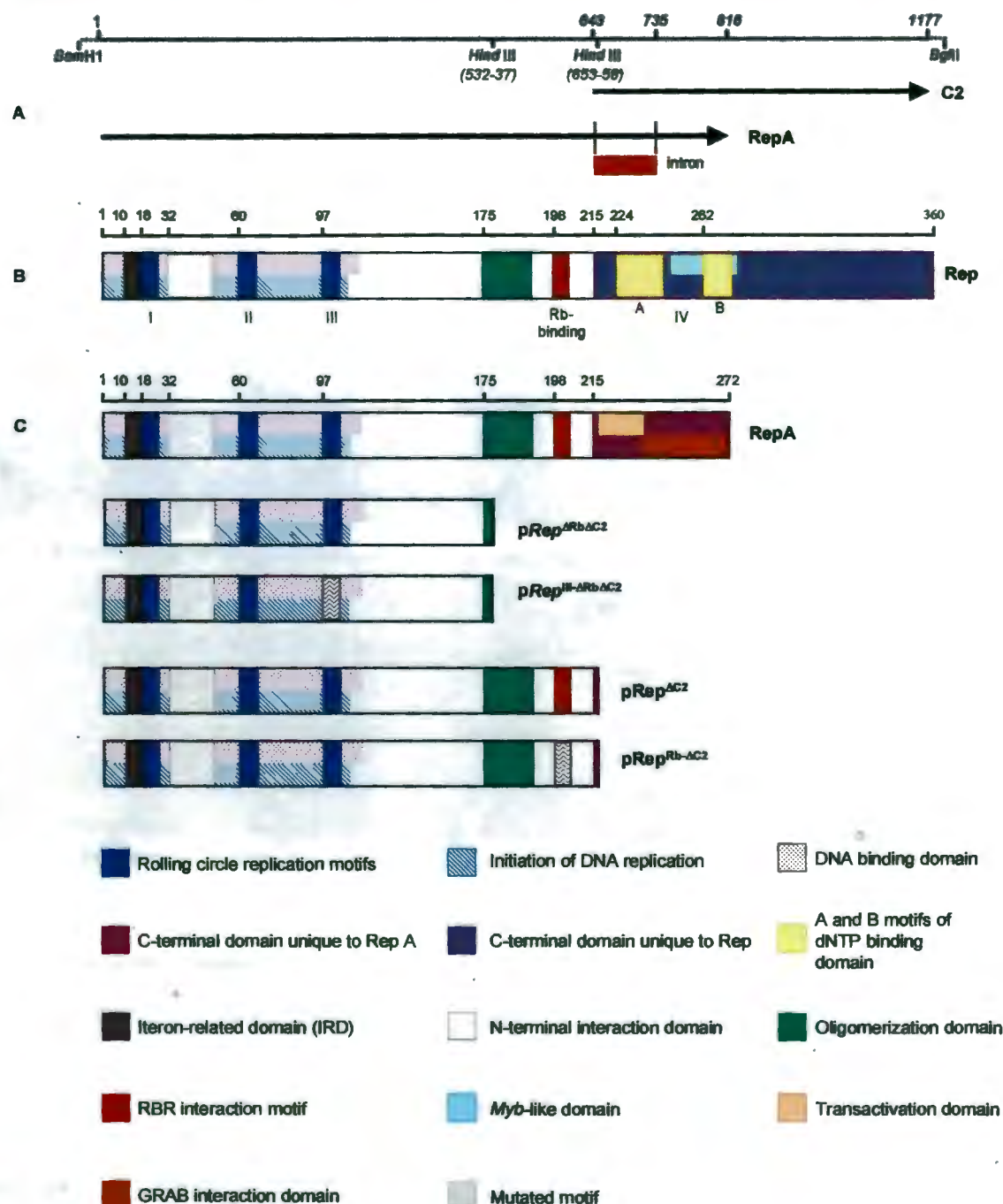


FIGURE 2.1 MSV Rep genes and gene products expressed in transient replication assays. (A) The MSV-Kom RepA (C1) and C2 ORFs are shown, including selected important nucleotide co-ordinates. Numbering of nucleotides is relative to the Rep ATG (A = 1). The positions of restriction sites used in the cloning of various Rep genes are also shown (see text for details). The *Bam*H1 site was introduced upstream of the Rep ATG, and the *Bgl*III site is in the MSV SIR. (B) The known sequence motifs and functional domains of the mastrevirus Rep protein (expressed from the spliced C1/C2 ORFs). Amino acid numbering is relative to the Rep start codon. (C) Truncations of RepA (shown relative to the full-length RepA protein) used in the replication assays. The motifs present in each Rep variant are shown, with mutated motifs represented by grey zigzags.

The same procedure was followed to obtain the RepA version of the Rb⁻ mutant, using pKomRb⁻

(described in chapter 3) as a starting template in place of pKom602. The resulting RepA plasmids, both containing the 3' splice site mutation, were called pSKRepA (wt) and pSKRepA^{Rb-} (Rb⁻ mutant)

2.2.1.5 Clone construction for yeast two-hybrid analysis

The Rep genes of pSKRep^{Δ1}, pSKRep^{Rb-Δ1}, pSKRepA, pSKRepA^{Rb-}, pSKRep^{ΔC2} and pSKRep^{Rb-ΔC2} were cloned in frame with the GAL4 activation domain into pGAD424 (CLONTECH, CA) and in frame with the GAL4 binding domain into pBD-GAL4-CAM (Stratagene, La Jolla, CA) to create Rep-GAL4 fusion products. To create the appropriate enzyme sites in order to clone the Rep genes in frame with the GAL4 binding domain, primers were used (Table 2.2A) which added on the *SalI* restriction site (forward primer) and the *PstI* restriction site (reverse primer) to the 5' and 3' ends respectively of each Rep gene. The same procedure was followed to create the appropriate enzyme sites in order to clone the Rep genes in frame with the GAL4 activation domain (Table 2.2B), this time adding on *SalI* to the 5' end and *BglII* to the 3' end of each Rep gene. Rep genes in pGAD424 were designated pADRep^Δ, pADRep^{MutΔ} or pADRep^{Mut}, while those cloned in pBD-GAL4-CAM were designated pBDRep^Δ, pBDRep^{MutΔ} or pBDRep^{Mut} (where Δ = a deletion, and Mut = Rb⁻ mutation). All constructs were confirmed to be correct by sequencing.

TABLE 2.2 Primers used in the cloning of various Rep genes into (A) pBD-GAL4-CAM and (B) pGAD424

Forward Primer for all Rep genes:	
(A) pBD-GAL4-CAM Rep Template pSKRep ^{Δ1} , pSKRep ^{Rb-Δ1} pSKRep ^{ΔC2} , pSKRep ^{Rb-ΔC2} pSKRepA, pSKRepA ^{Rb-}	BD <i>SalI</i> C1(F) 5'-CCGGGTCGACTCATGGCCTCCTCCTCATCC-3'
	Reverse Primer
	BD <i>PstI</i> C2 5'-CTTGGCTGCAGTTTACACTTCCTCCGTAGGAGG-3'
	BD <i>PstI</i> ΔC2 5'-CTTGGCTGCAGTAAGCTTCGGGACTAACCTGG-3'
BD <i>PstI</i> RepA 5'-CTTGGCTGCAGTCTAGGCTTCTGGCCCAAG-3'	
(B) pGAD424 Rep Template pSKRep ^{Δ1} , pSKRep ^{Rb-Δ1} pSKRep ^{ΔC2} , pSKRep ^{Rb-ΔC2} pSKRepA, pSKRepA ^{Rb-}	Reverse Primer
	AD <i>BglII</i> C2 5'-TTCATAGATCTCTTACACTTCCTCCGTAGGAGG-3'
	AD <i>BglII</i> ΔC2 5'-TTCATAGATCTCAAGCTTCGGGACTAACCTGG-3'
	AD <i>BglII</i> RepA 5'-TTCATAGATCTCCTAGGCTTCTGGCCCAAG-3'

The cDNA of the maize RBR protein (pZmRb1; Xie *et al.*, 1996) cloned into pGBT9 (pGBT9ZmRb1) was provided by Dr G. Horvath (described in Horvath *et al.*, 1998).

2.2.2 Yeast Two-Hybrid Analysis of the RBR Protein Interaction Properties of Mutant, Truncated and Wild Type *Rep* Constructs

While the cloning of the *Rep* constructs described above for yeast two-hybrid analysis was done as part of this thesis, the actual yeast transformation and analysis of the RBR-interaction properties of the *Rep* constructs was performed by D. McGivern (John Innes Centre, Norwich, U.K.) as described below.

Plasmids containing *GAL4* binding domain fusions (*trp1* transformation marker) and *GAL4* activation domain fusions (*leu2* transformation marker) were introduced separately into *Saccharomyces cerevisiae* strains CG1945 and Y187 respectively as described by Gietz and Woods (1994). The transformation mixture was plated onto yeast drop-out selection media lacking the appropriate amino acid to select for transformants. Yeast strains CG1945 (MAT α ; transformed with pZmRb1) and Y187 (MAT α ; transformed with *Rep* gene derivatives) were mated according to a protocol modified from the CLONTECH Yeast Protocols Handbook (CLONTECH Laboratories Inc., Palo Alto, Ca, USA). The mated yeast was grown on selective drop-out medium lacking tryptophan (Trp) and leucine (Leu), and drop-out medium lacking Trp, Leu and histidine (His) and containing 5 mM 3-amino-1,2,4-triazole (3AT). Only strains containing interacting fusion proteins can grow on the latter medium.

2.2.3 Transient Replication Assays

2.2.3.1 Maintenance of black Mexican sweetcorn (BMS) suspension cultures

BMS suspension culture cells were maintained in the dark at 26°C with constant rotary shaking (120 r.p.m.) in BMS liquid medium (pH 5.75), which is MS medium (Murashige and Skoog, 1962) supplemented with 3% sucrose, 0.01% myoinositol and 2 mg/L of 2,4-dichlorophenoxyacetic acid (2,4-D). Cells were subcultured at a 1:6 dilution every two weeks.

2.2.3.2 Transfection of BMS by particle bombardment

BMS cells were subcultured at a 1:3 dilution three days prior to bombardment. Four hours before bombardment, 1.0 mL packed volume of cells was filtered onto Whatman # 4 filter paper, which

was placed on BMS high osmoticum solid medium (BMS medium containing 0.2 M mannitol, 10 $\mu\text{g}/\text{mL}$ silver nitrate and 0.8% agar). An aliquot of 1 μg of each plasmid was precipitated onto 1 μm gold particles (50 μL of 60 mg/mL gold suspended in 50% glycerol) according to the protocol of Dunder *et al.* (1995). Cells were bombarded using a Biorad/DuPont PDS1000-He system at a pressure of 650 psi under a vacuum of 27 inHg. The gap distance was 6 mm, the macrocarrier travel distance 5 mm, and the target distance 6 cm. Cells on each plate were bombarded twice, each shot delivering approximately 167 ng of each plasmid. In each replication experiment, nine plates of BMS were co-bombarded with pKom602 and one of the Rep constructs, pRep^{Wt}, pRep^{Mut}, pRep ^{Δ} or pRep^{Mut Δ} provided in *trans*, and nine plates of BMS were co-bombarded with pKom602 and pAHC17. Since pAHC17 is the vector in which the Rep genes were cloned, it provided a non-Rep co-bombardment control. From here on, co-bombardment of pKom602 and pAHC17, with no Rep gene provided in *trans*, is referred to as bombardment of "pKom602 alone". After bombardment, cells were incubated in the dark at 26°C for 24 hours, after which they were transferred to BMS solid medium and incubated in the same conditions for four days.

2.2.3.3 Quantitative PCR

Total DNA was extracted from BMS cells four days after bombardment, using the method of Dellaporta *et al.* (1983) except that, having resuspended the chromosomal DNA in 50 mM Tris-HCl, 10 mM EDTA (pH 8), 600 $\mu\text{g}/\text{mL}$ RNaseA was added and the mixture incubated for 1 hour at 37°C. After a second precipitation with isopropanol, DNA was finally resuspended in water and diluted to 50 $\text{ng}/\mu\text{L}$. To ensure all samples were the same concentration, equal amounts of total DNA were electrophoresed through a 0.8% agarose gel, and band intensity measured using the virtual densitometer computer program, GelTrak (Dennis Maeder, University of Cape Town). The relative amounts of viral replication were determined using a quantitative PCR-based assay (Fig. 2.2). Primers were designed to amplify the region corresponding to nucleotides 1595-209 in the MSV-Kom genome (relative to the unique *Bam*H1 site at the beginning of the V1 gene). Primer sequences were: MSV-DEG1 5'-TTGGVCCGMVGATGTASAG -3' and MSV-DEG2 5'-CCAAKDTTCAGCTCCTCCG -3' (Willment *et al.*, 2001). These primers, which are able to amplify viral DNA once it has replicated, cannot amplify linear MSV-Kom DNA from the pKom602 input plasmid (see Fig 2.2). To confirm this, total DNA extracted immediately after bombardment (as a control for input plasmid DNA amplification) was subjected to the same PCR. Each PCR reaction was "spiked" with pKep177 (obtained from K.E. Palmer, UCT) of known concentration as an internal control. The amount of pKep177 added to each reaction was

in the range of 20 to 100 pg, and the amount of total DNA was 100ng, depending on the level of viral replication approximated previously by non-quantitative PCR. Optimum ratios of spike to viral DNA were required to avoid the one out-competing the other. The relative concentration of replicated DNA was calculated by determining the ratio of the replicative form (RF) virus band intensity to the pKep177 competitor band intensity, using GelTrak. pKep177 has a *Pst*I site inserted 72 bp downstream from the start codon, which distinguishes it from the RF viral band. The principle of the quantitative PCR assay is explained in Fig. 2.2.

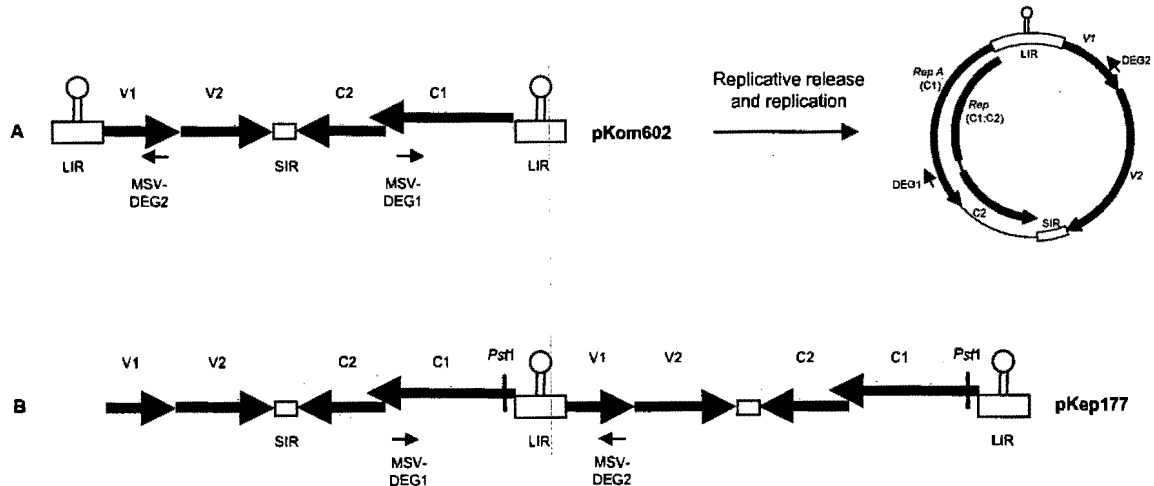


Figure 2.2 Plasmids and primers used for quantitative PCR assay. (A) The primers MSV-DEG1 and MSV-DEG2 do not amplify linear MSV-Kom DNA from pKom602. In the nuclei of BMS cells the viral DNA is replicatively released from the vector, and as circular RF DNA it can serve as a template for the amplification of a 1314 bp fragment. (B) Included in the PCR reaction is an internal control, pKep177, of known concentration. This competitor is a tandem dimer of the MSV-Kom genome, with a *Pst*I site inserted at the start of the *C1* ORF. Being a dimer, pKep177 is amplified by the MSV-DEG1 and MSV-DEG2 primers and competes with viral RF DNA for primers and other PCR components. Digestion of amplified pKep177 with *Pst*I yields two bands of 604 and 710 bp, allowing the competitor to be distinguished from amplified MSV-Kom DNA. The relative concentration of replicated viral DNA is calculated by determining the ratio of MSV-Kom DNA band intensity to that of pKep177, whose concentration is known.

2.2.4 *Digitaria sanguinalis* Tissue Culture

2.2.4.1 Callus induction

Young unemerged inflorescences of *D. sanguinalis* were surface-sterilised by soaking in 70% ethanol for 1 minute followed by 0.35% sodium hypochlorite for 20 minutes, and were then washed four times with sterile distilled water. The inflorescences were cut into 5 mm segments and transferred to callus induction medium, which is MS medium (adjusted to pH 5.8) supplemented with 3% sucrose and 2.5 mg/L 2,4-D, and solidified with 0.8% agar. Plates

containing the inflorescences were maintained in the dark at 26°C, as were the calli once they had initiated.

2.2.4.2 Transformation of *D. sanguinalis*

Embryogenic *D. sanguinalis* calli were transformed by particle bombardment using the Biorad/DuPont PDS1000-He system. Sixteen hours prior to bombardment, embryogenic calli were transferred to high osmoticum medium (MS medium containing 2.5 mg/L 2,4-D, 100 mg/L myo-inositol, 0.2 M mannitol and 10 mg/L silver nitrate). For each bombardment, two µg of plasmid DNA were precipitated onto gold particles as described for transfection of BMS. Each mutant or truncated *Rep* construct was co-bombarded with pAHC25 (Ubi-Bar/UbiGus) at a 1:1 weight ratio. *Rep* plasmids chosen along with pAHC25 to co-transform *D. sanguinalis* were pRep^{III-Rb+NTP+}, pRep^{III-Rb-NTP+}, pRep^{Rb-ΔC2} and pRep^{III+Rb+NTP+(AS)}. However, in bombardments with pAHCRep^{III-Rb-NTP+} (Ubi-Bar/Ubi-Rep), there was no co-bombardment with pAHC25. The settings on the biolistics device were as follows: the gap distance was 6 mm, the macrocarrier travel distance 5 mm, and the target distance 6 cm. Each target plate was bombarded twice at a pressure of 900 psi, each shot delivering approximately 333 ng per plasmid. Twenty-four hours after bombardment the calli were transferred from high osmoticum to MS maintenance medium. Non-bombarded calli, and calli bombarded with pAHC25 alone were used as controls in all experiments.

2.2.4.3 Selection and regeneration of transgenic calli

Seven days after bombardment, calli were transferred to regeneration medium (MS medium containing 0.1 mg/L naphthaleneacetic acid and 10 mg/L benzylaminopurine [NAA and BAP, Sigma]) with selection (3 mg/L bialophos). Once on regeneration/ selection medium, calli were kept in the dark for five days at 26°C, followed by 16 hours diffuse light for nine days, after which they were exposed to 16 hours full light per day. Shooting callus was then transferred to rooting medium (MS medium without hormones or plant growth regulators). Once roots had grown, plantlets were hardened off in a 33:33:33 mix of sand, compost and palm peat and finally transferred to potting soil.

2.2.4.4 Seed germination

Seed collected from transgenic plants were incubated at 37°C overnight. They were then surface sterilised by shaking in 70% ethanol for five minutes, followed by 3.5% sodium hypochlorite, 0.02% Triton X-100 for ten min, and then washed five times with sterile distilled water. The

sterilised seeds were suspended in 5 mL of 0.1% agar and pipetted onto a plate containing PNS medium (Plant Nutrient agar medium with Sucrose (pH 5.5) solidified with 0.75 % nutrient agar. The seeds on PNS plates were kept in the light at 26°C until they germinated. The plantlets were then hardened off as described for *D. sanguinalis* regeneration.

2.2.5 Transgene Expression Analysis

2.2.5.1 Reverse transcription-PCR (RT-PCR)

Total RNA was extracted from stock transgenic callus material (previously initiated from T₀ plants of each transgenic line), using the reagent TRIzol® (GIBCOBRL). ~200 mg of callus material was homogenized in 1 mL Trizol, followed by incubation at room temperature (RT) for 5 min. Insoluble material was removed from the homogenate by centrifugation at 12,000 g for 10 min at 4°C. 200 µL of chloroform was added to the supernatant and shaken vigorously for 3 min at RT. The samples were centrifuged at 12,000 g for 15 min at 4°C. RNA was precipitated by the addition of 0.5 mL isopropanol to the supernatant, which was incubated at RT for 10 min, followed by centrifugation at 12,000 g at 4°C. The RNA pellet was washed with 1 mL 75% ethanol, and resuspended in 50µl DEPC-treated water (that is 0.5 mL diethyl pyrocarbonate [Sigma] added to 1 L distilled water, left to stand overnight, and autoclaved).

As a positive control for the RT-PCR reaction, the Rep gene (from pGEMTRep, cloned in frame with the T7 promoter in the pGEM®-T Easy Vector (Promega, WI) was transcribed using the Ribomax™ RNA production system (Promega, WI). pGEMTRep was linearized with *Pst*I, precipitated from the digest reaction mixture using 4M LiCl, and the resuspended DNA was used in the transcription reaction, as described in the kit manual (also available at www.promega.com). After transcription, 1 µL of DNase (Promega, WI) was added to the transcription reaction and incubated for 30 min at 37°C to remove the original DNA template, after which the DNase was inactivated at 90°C for 5 min.

The RNA from each transgenic sample was also DNase-treated as described above before undergoing RT-PCR, to remove any contaminating transgene DNA that could be amplified by the DNA polymerase. A control reaction was carried out using the same RT-PCR conditions except that the reverse transcriptase enzyme was omitted to confirm that amplification was not due to contaminating genomic DNA. Primers were designed that could amplify all Rep transgenes (truncated and full-length) from the different transgenic samples. These were RepΔC2F: 5' ATGGCCTCCTCCTCATCCAAC 3' and RepΔC2R: 5' AAGCTTCGGGACTAACCT 3'. The

bar transcript was amplified from each transgenic sample by RT-PCR using the primers: BAR1 (forward) 5' CGTCAACCACTACATCGAG 3' and BAR2 (reverse) 5' GAAACCCACGTCATGCCAG 3' obtained from T. Dube, UCT.

RT-PCR was carried out using the Access RT-PCR System (Promega, WI). The RT-PCR cycles were as follows: 48°C for 30 min (first strand synthesis); 94°C for 2 min (inactivation of reverse transcriptase and denaturation of primers/cDNA); [94°C for 30 sec (denaturation), 60°C for 30 sec (annealing), 68°C for 1 min (extension)] x40 cycles; 68°C for 2 min (final extension).

2.2.5.2 GUS assays

Expression of the GUS gene (*uidA*, encoding β -glucuronidase) in both transient and transgenic assays was analysed by histochemical and protein (spectrofluorometric) assays. In the histochemical assay, transient GUS activity in *D. sanguinalis* calli that had been bombarded with pAHC25 was visualized by addition of the GUS substrate X-gluc (Sigma) to the calli three days after bombardment (described in Jefferson *et al.*, 1987). GUS expression in transgenic calli and in leaves from transgenic plants was analysed in the same way. In the spectrofluorometric assay, protein was extracted from calli three days after bombardment, and GUS activity was determined according to Jefferson (1987). Protein concentrations were determined by the method of Bradford (1976) and GUS activity was corrected for protein concentration.

2.2.6 Test for Resistant *D. sanguinalis* by Challenge with MSV

Transgenic plants were challenged with MSV using viruliferous leafhoppers (*C. mbila*), obtained from Dr. Mike Barrow (Pannar Ltd, Greytown, South Africa). The leafhopper population at Pannar Ltd had been fed on symptomatic maize leaves collected from various locations in South Africa and Zimbabwe in order to develop a mixed (albeit unknown) MSV population, and were subsequently maintained at UCT on *Zea mays* cv Jubilee. In the challenge experiments, the viruliferous leafhoppers were placed in a small vial containing a ~1 mm slit, through which a single leaf from a transgenic plant was inserted, ensuring that the leafhoppers in a particular vial could only feed on one leaf of one plant. Four leafhoppers were placed in each vial, and three vials were placed at different positions on each plant. Thus, there were three points of entry for the virus, and in total 12 leafhoppers feeding on each plant. One challenge experiment, using plants transgenic for *pRep*^{III-Rb-NTP+}, was carried out at Pannar Ltd in Greytown, South Africa, under the supervision of Dr Rikus Kloppers. In that case, five plants from each of two lines and a non-transgenic control plant were challenged by feeding leafhoppers on three different leaves of

each plant, as explained above. The rest of the challenge experiments, carried out at UCT, Cape Town, were slightly different. Leafhoppers in vials were fed first on a non-transgenic or pAHC25-transformed plant for two days, then transferred to a transgenic plant and allowed to feed for two days. Since the same leafhoppers feed on both the transgenic and control plants, in the case of a transgenic plant not becoming infected while the control does get infected, this method ensures that the lack of infection is not due to non-viruliferous leafhoppers, but rather to the plant being resistant to viral infection.

2.2.7 Maize (Hi-II) Tissue Culture

Embryogenic high type II (Hi-II) calli, obtained from Dr. W. Gordon-Kamm (Pioneer Hi-Bred, International, Inc., Johnston) were maintained on N6 medium (Chu, 1978), pH 5.8, supplemented with 3% sucrose, 0.3% proline, 0.01% casamino acids and 2 mg/L 2,4-D, solidified with 2.4g/L Gelrite.

2.2.7.1 Transformation of Hi-II

Actively growing embryogenic calli were selected as target tissue for gene delivery by particle bombardment, using the Biorad/ DuPont PDS1000-He system. The calli were placed onto target plates containing high osmoticum medium (N6 maintenance medium with 36.4 g/L mannitol, 36.4 g/L sorbitol and 10 mg/L silver nitrate) four hours prior to bombardment. Two µg of plasmid DNA were precipitated onto gold particles as described for transfection of BMS. The *Rep* plasmid chosen to transform Hi-II, $pRep^{Rb-ΔC2}$, was co-bombarded with pAHC25 (Christenson and Quail, 1996) at a pressure of 1100 psi. The biolistic device settings were as follows: 8 mm between the rupture disc and macrocarrier, 10 mm between the macrocarrier and the stopping screen, and 7 mm between the stopping screen and the target. After the first shot, a recovery time of four hours was allowed before the calli were shot a second time. Approximately 333 ng of each plasmid were delivered per shot. Twenty-four hours after bombardment the calli were transferred from high osmoticum to maintenance N6 medium. Non-bombarded calli, and calli bombarded with pAHC25 alone were used as controls in all experiments.

2.2.7.2 Selection and regeneration of transgenic calli

Four days after bombardment, calli were transferred to a gentle selection medium (N6 maintenance medium with 1 mg/L bialophos) on which they were maintained for two weeks. All subsequent selection was more stringent at 3 mg/L bialophos. Selection continued for 6-8 weeks

in the dark, after which bialophos-resistant calli were transferred to shoot initiation medium (MS medium with 5 mg/L BAP, 0.25 mg/L 2,4-D, 3% sucrose and 3 mg/L bialophos). After five days in the dark followed by nine days under diffuse light, the calli were kept in the light for 16 hours a day. Once shoots emerged they were transferred to shoot elongation medium (shoot initiation medium without BAP or 2,4-D) and finally root elongation medium (MS supplemented with 1.0% sucrose and 3 mg/L bialophos). Stringent selection was maintained during the whole regeneration process. Once roots had grown, plantlets were hardened off in a 33:33:33 mix of sand, compost and palm peat and finally transferred to potting soil. Adult transgenic plants were sent to Pannar Ltd, Greytown, SA, for controlled self-pollination and setting of seed.

2.2.8 Analysis of the MSV Strain Composition Transmitted by Viruliferous Leafhoppers.

0.5 g to 2 g of infected maize leaf material was frozen in liquid nitrogen and finely ground. 6 mL of Extraction Buffer (0.1 M Tris-HCl; 0.1 M NaCl; 0.1 M EDTA; 1% SDS w/v; pH7) was added, the mixture shaken and incubated at 65 °C for 5 min. After centrifuging at 10 000 G for 5 min, 5.5 mL of the supernatant was added to 5.5 g CsCl and shaken until the CsCl dissolved, followed by centrifugation at 10 000G for 5 min at RT. 5 mL of the supernatant, plus 10 µg of pSK and 400 µl of 10 mg/mL ethidium bromide were added to a 5 mL Quick-Seal tube and centrifuged in a Beckman Vti65 rotor at 266 805 G overnight at 20°C. After the separation of linear ds DNA (plant genomic DNA, upper band) from ds ccDNA (viral and plasmid DNA, lower band) the lower band, visualized using ultra violet (UV) light at 310 nm, was collected from the tube using a sterile wide-bore needle on a sterile 2 mL syringe. The ethidium bromide was extracted from the collected fraction by the addition of an equal volume of salt-saturated phenol (300 mL of 5M NaCl in TE buffer [10mM Tris, 1mM EDTA] added to 600 mL of isopropanol), followed by thorough mixing and centrifugation for 1 min in a microcentrifuge. The upper isopropanol phase containing the ethidium bromide was discarded. This procedure was repeated twice, before adding two volumes of sterile distilled water to the DNA-containing phase. The DNA was pelleted by the addition of 1 volume isopropanol, incubation on ice for 10 min and centrifugation for 15 min. The pellet was washed in 70% ethanol and resuspended in sterile distilled water.

50 µL of DNA (of unknown concentration) was digested with *Bam*HI, followed by precipitation of the DNA with 4 M LiCl. The DNA was then cloned into the *Bam*HI site of pSK and transformed into *E. coli* DH5α. Colonies were picked and replica-plated onto two plates, one containing solidified bacterial growth medium with ampicillin selection and the other containing

a Hybond-N+ membrane (Amersham Pharmacia, UK). 100 colonies were screened by colony hybridization using DIG-labelled (Boehringer Mannheim) whole genome probes. An equal mixture of four probes consisting of the genomes of MSV-Kom, MSV-Set, MSV-Mat and MSV-VW was used. The colony hybridization was carried out according to the Hybond-N+ manufacturer's instructions (Amersham Pharmacia). Chemiluminescent detection of colonies containing DNA that hybridized to the genome probes was carried out according to the DIG System User's Guide (Boehringer Mannheim). Positive colonies picked from the replica-plated bacteria on growth medium were subjected to PCR using primers that annealed to either side of the pSK polycloning site. Primer sequences were: M13 uni Rev 5' AGCGGATAACAATTTTCACACAGG 3' and M13 uni Fwd 5' CCCAGTCACGACGTTGTAAAACG 3'. PCR-amplified fragments were analysed by restriction length polymorphism (RFLP) to distinguish different virus strains from one another, as described in Willment *et al.* (2001). Plasmid DNA was extracted from colonies containing different MSV genomes identified by RFLP of the PCR products, and the whole genomes were sequenced, first using the M13 Fwd and Rev primers described above, and then internal primers in the forward and reverse directions. Using DNAMAN, the sequences were aligned with a variety of African streak virus genomic sequences, and the relationship between the genomes was determined using a rooted neighbour joining tree.

2.3 RESULTS AND DISCUSSION

2.3.1 Interactions between MSV Rep and RBR proteins

Since some of the *Rep* constructs used in the replication assays included mutations in the RBR protein interaction domain, it was important to first determine whether the RBR protein-binding ability of these mutants was abolished. In order to separate Rep from RepA, intronless *Rep* (in effect behaving as a spliced *Rep* and therefore capable of expressing Rep only) and *RepA* genes were made, both wt and with the Rb⁻ mutation. The ability of truncated, mutated and wt Rep and RepA gene products to bind to the maize RBR protein (ZmRb1) was determined using a yeast two-hybrid assay (Table 2.3). Interaction of a Rep product fused in-frame to the Gal4 activation domain with ZmRb1 fused in-frame to the GAL4 binding domain allows growth of yeast on drop-out media lacking histidine, by inducing expression from the *HIS3* promoter. Of the Rep constructs, only yeast co-transformed with pADRepA and pGBT9ZmRb1 grew on media lacking histidine (Table 2.3). The Rb⁻ mutants (each with a mutation of LxCxE to LxCxK) pADRep^{Rb-Δ1}, pADRep^{Rb-ΔC2}, and pADRepA^{Rb-} did not interact with pGBT9ZmRb1, which is consistent with

results of Xie *et al.* (1995), and Liu *et al.* (1999a). However, yeast co-transformed with pADRep^{ΔC2} and pGBT9ZmRb1 was also unable to grow in the absence of histidine, indicating that Rep^{ΔC2} cannot interact with ZmRb1 despite having an unmutated RBR-interaction motif. This is contrary to the data of Horvath *et al.* (1998), who found that a similar-sized Rep from MSV-Ns does interact with ZmRb1. The discrepancy may reflect the differences in the assay conditions. For example, different yeast strains were used, as well as different amounts of 3-AT. The yeast strain used in this assay was more sensitive to 3-AT than the one used by Horvath *et al.* (1998), possibly resulting in this assay being too stringent (D. McGivern, pers. comm.). The different strains of MSV used could also play a role in the disparate results, although this is unlikely since MSV-Ns and MSV-Kom are closely related. Nonetheless, the main objective of the yeast two-hybrid assay was achieved, that is confirmation that the LxCxE to LxCxK mutation did indeed abolish the RBR-binding ability of the Rb⁻ mutant Reps.

TABLE 2.3 Interactions between MSV Rep protein variants and the maize retinoblastoma protein (ZmRb1).

Binding domain fusion	Activation domain fusion	Growth of Yeast (Trp ⁻ Leu ⁻ , His ⁻ , + 5mM 3-AT)	Growth of Yeast (Trp ⁻ Leu ⁻ , + 5mM 3-AT)
pGBT9ZmRb1	pADRepA	++	++
pGBT9ZmRb1	pRepΔC2	-	++
pGBT9ZmRb1	pADRepRb ⁻ ΔC2	-	++
pGBT9ZmRb1	pADRepARb ⁻	-	++
pGBT9ZmRb1	pADRepRb ⁻ ΔI	-	++
pGBT9ZmRb1	pGAD424	-	++
pLamC	pGAD424	-	++
p53	pSV40	+++	++
p53	pADRepA	-	++

Yeast cells co-transformed with pGBT9ZmRb1 and pGAD424-Rep fusion products were grown on drop-out medium supplemented with 5mM 3-AT in the presence or absence of histidine. Only strains containing interacting fusion proteins can grow in the absence of histidine. As negative controls, pGBT9 alone (GAL4 binding domain not fused to the ZmRb1 protein) was co-transformed with each pGAD424-Rep plasmid, as was pLamC, which is unable to interact with either ZmRb1 or MSV Rep. All negative controls grew in the presence, but not the absence, of histidine (data not shown). As a positive control, yeast was transformed with plasmids expressing p53 (BD fusion) and SV40 T-Ag (AD fusion). This assay was repeated with the same results

2.3.2 The Effects of Transiently Expressed Rep Proteins on MSV Replication in black Mexican sweetcorn (BMS)

Figure 2.3 uses the data of a truncated Rep construct (having been co-bombarded with pKom602 into BMS) as an example of how quantitative PCR (QPCR) was used to determine the level of

viral replication achieved in the presence and absence of various Rep constructs. In this example the Rep construct (Rep^{ΔRbΔC2}) enhanced replication of MSV-Kom five-fold compared with MSV-Kom alone (discussed later). This procedure was followed for each of the Rep constructs presented in Figs 2.4 and 2.5, and each bombardment was repeated at least once to ensure consistency of results. Because there is a large amount of variation between bombardments (the efficiency of one bombardment can differ greatly from another), the effect of each Rep construct on viral replication is presented relative to the replication of MSV-Kom alone (100% replication). For a diagrammatic representation of the Rep constructs used in the replication assays, refer to Fig. 2.1.

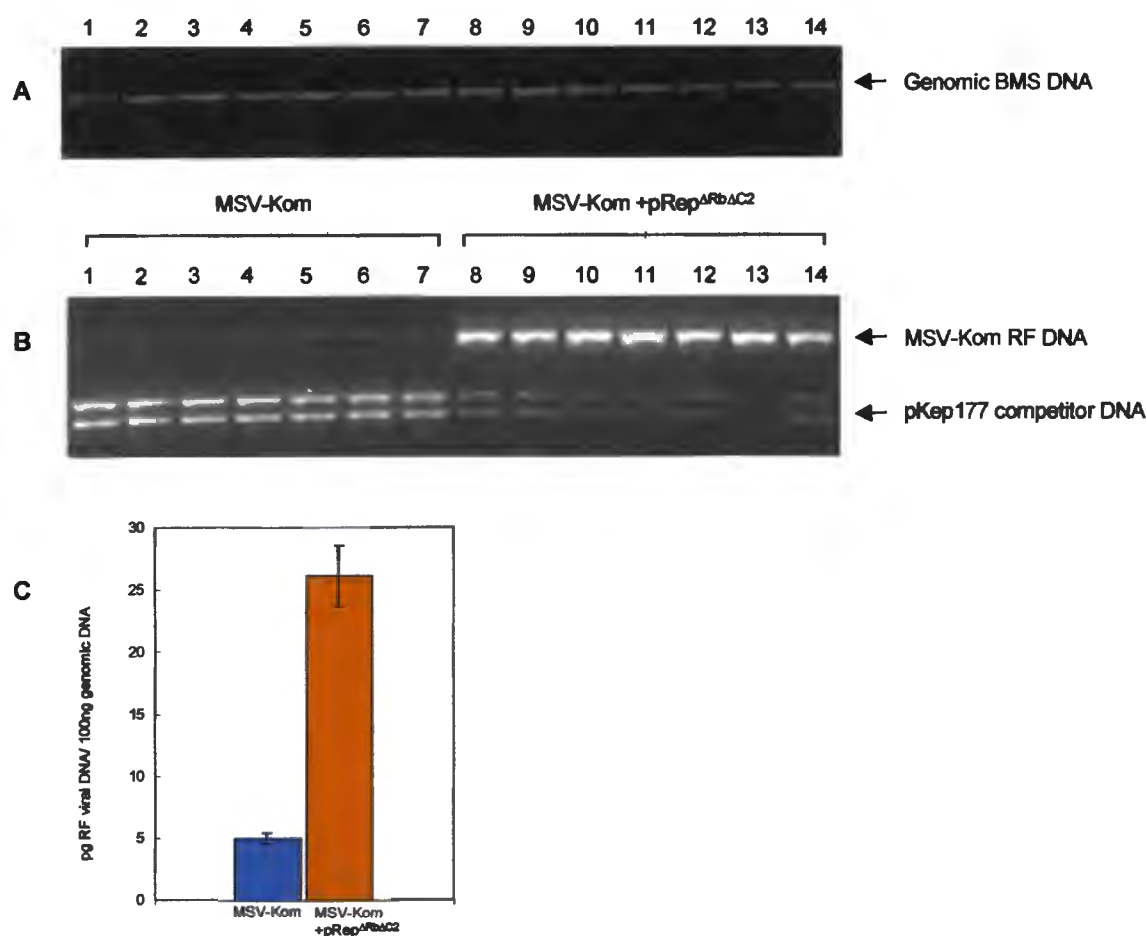


FIGURE 2.3 Determination of MSV replication levels in BMS by a quantitative PCR assay (QPCR). (A) BMS genomic DNA (extracted three days after bombardment) is quantified on a gel to ensure uniform amounts of each sample are used in the QPCR. (B) Equal amounts of genomic DNA are subjected to QPCR, each reaction being spiked with an internal control (pKep177) of known concentration. After DNA amplification, equal amounts of the QPCR reaction are digested with *Pst*I to distinguish the spike from the viral bands, and run on a gel. (C) A densitometry programme (GelTrak) is used to determine the concentration of viral DNA in each sample, by calculating the ratio of the band intensity of viral DNA to that of the spike DNA. This is expressed as picograms (pg) of RF viral DNA present in 100 nanograms (ng) of genomic DNA.

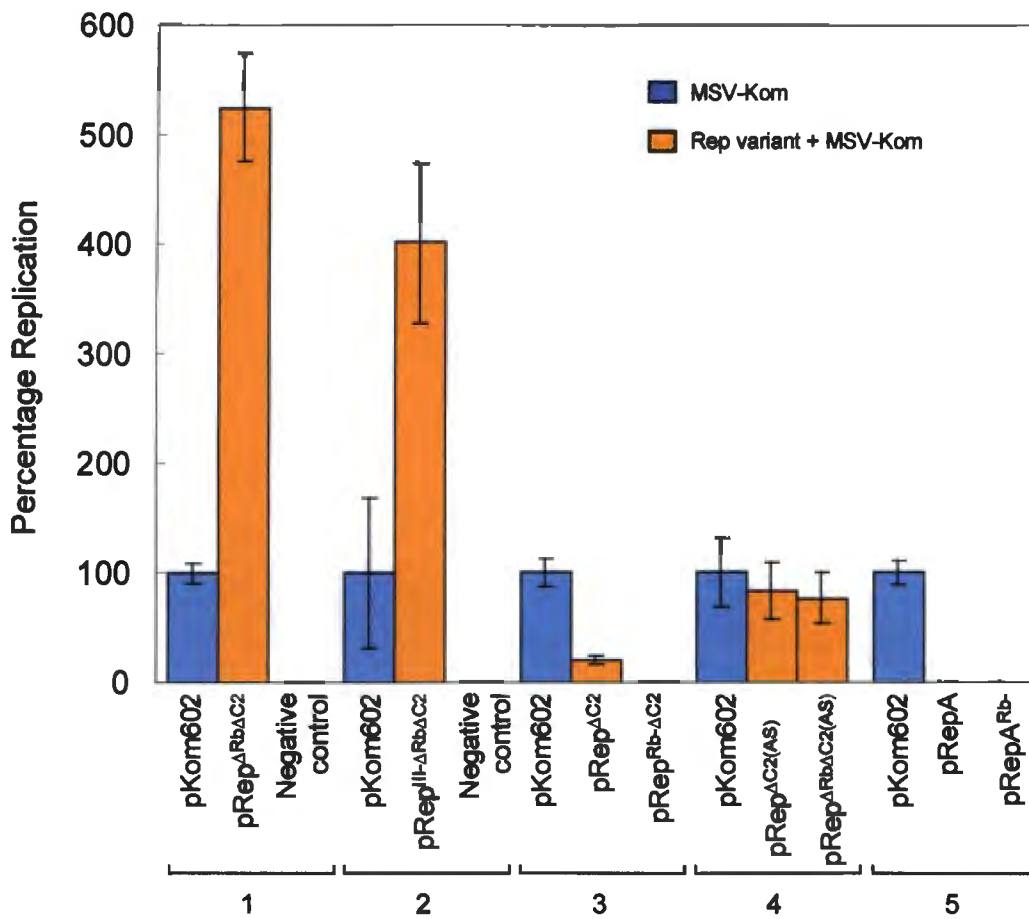


FIGURE 2.4 The effects of RepA mutants and truncations on MSV-Kom replication in black Mexican sweetcorn (BMS). In each bombardment experiment (which was repeated at least once), up to nine BMS samples were bombarded with pKom602 alone and with pKom602 + pRep^A (assays 1 and 3), pRep^{mutΔ} (assay 2), pRep^{Δ(AS)} (assay 4) and pRepA^{wt}/pRepA^{mut} (assay 5). From individual replicate data calculated using QPCR, an average amount in pg of replicated virus in the presence and absence of each RepA derivative was calculated. The average value for MSV-Kom alone was then taken as 100% replication, and the values for MSV-Kom + each RepA derivative expressed relative to 100%. In assays 1 and 2, the negative control was genomic BMS DNA extracted immediately after bombardment with pKom602, showing that the QPCR does not amplify input plasmid pKom602. Error bars represent 95% confidence intervals.

The effects of the Rep constructs on viral replication varied enormously, from ~30-fold enhancement of replication (pRep and pRep^{III+Rb-NTP+}) to 100% inhibition (pRep^{III-Rb+NTP+}, pRep^{III-Rb-NTP+}, pRep^{Rb-ΔC2}, pRepA and pRepA^{Rb-}). Still others enhanced (pRep^{ARbΔC2} and pRep^{III-ARbΔC2}) or inhibited (pRep^{ΔC2}) replication to a lesser extent, while the truncated antisense Reps (pRep^{ΔC2(AS)} and pRep^{ARbΔC2(AS)}) had no significant effect. The Rb⁻ mutation, while abolishing RBR protein interaction, had no effect on the *trans*-replicating ability of pRep^{III+Rb-NTP+}, which enhanced the

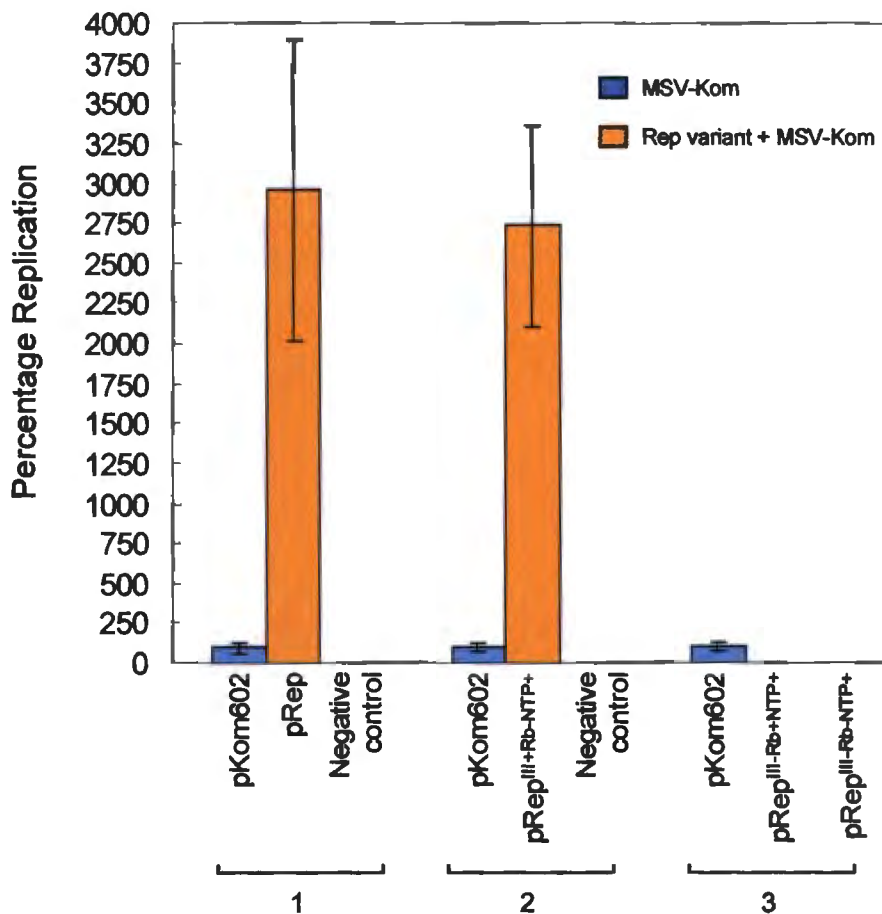


FIGURE 2.5 The effects of full-length Rep (wild-type and mutants) on MSV-Kom replication in black Mexican sweetcorn (BMS). For assays 1 and 2, the same procedure as described in Fig 2.4 was followed. Error bars represent 95% confidence intervals. Assay 3 was carried out by T. Mangwende (2000), using a Southern blot and subsequent densitometry analysis in place of QPCR.

virus' replication to the same levels as did wt Rep. A striking correlation can be seen between the size of the truncated Reps and their effects on MSV-Kom replication. The two Reps of 179 aa had the effect of enhancing replication when supplied in *trans*, while the opposite was true of the slightly larger Reps of 219 aa, which inhibited replication (see Fig. 2.1 for a diagram of these truncated Reps). The motif III' mutation in $pRep^{III-\Delta Rb\Delta C2}$ had little effect on the replication-enhancing abilities of the protein, there being only a slightly significant difference in the level of viral replication achieved in the presence of $pRep^{III-\Delta Rb\Delta C2}$ and the non-mutated $pRep^{\Delta Rb\Delta C2}$. The Rb' mutation in $pRep^{Rb-\Delta C2}$ appeared to have a more significant effect on viral replication than the non-mutated $pRep^{\Delta C2}$, the former inhibiting replication completely, the latter by 80%. However, this is unlikely to be due to the abolition of RBR protein interaction, since the yeast-two hybrid

assay demonstrated that RBR protein binds to neither $pRep^{\Delta C2}$ nor $pRep^{Rb-\Delta C2}$ (or at the very least the RBR protein- $pRep^{\Delta C2}$ interaction was so weak that it was beneath the detection levels of the assay). It is possible that the LxCxE to LxCxK mutation had another effect on the function/s of Rep apart from abolishing RBR interaction. This is the subject of Chapters 3 and 4, and will not be discussed further here.

It is likely that the Rep constructs with inhibitory effects on viral replication are behaving in a dominant negative manner. While $pRep$ and $pRep^{III+Rb-NTP+}$ replicated MSV-Kom to very high levels, the addition of the III⁻ mutation to both Reps led to the complete inhibition of MSV-Kom replication. Since the mutant Reps should be expressed at high levels from the ubiquitin promoter (transient expression levels, however, were not determined), they probably interfere with the functions of MSV-Kom's wt Rep (which is expressed at low levels) in various possible ways. These include the competitive occupation of viral DNA binding sites by the mutant Reps, which, once bound to the LIR, cannot initiate RCR due to the III⁻ mutation. Although this was not directly tested on the III⁻ mutants, the fact that $pKom602$ containing the III⁻ mutation cannot replicate in BMS or establish an infection in maize (Mangwende, 2001) suggests that the mutation has abolished the ability of Rep to initiate replication. Another possibility for replication interference is the binding of the mutant Reps to wt Rep, disrupting the functions of oligomerization complexes. This could be the case with the 219 aa truncated Reps, which contain the oligomerization domain (see Fig. 2.1) and inhibit replication. Conversely, the 179 aa truncated Reps are missing the oligomerization domain (see Fig 2.1) and enhance replication. The fact that the antisense versions of both the longer and the shorter truncated Rep genes had no effect on viral replication, suggests that the inhibition and enhancement effects of the sense versions are as a result of protein expression. While oligomerization of the 219 aa Reps with MSV-Kom Rep could account for the replication inhibitory effects of the truncated proteins, the enhancement of viral replication by the 179 aa Reps is more difficult to explain. This is further explored in Chapter 5, since it is not relevant to this discussion. The inhibition of viral replication by RepA and RepA^{Rb-} is also expanded on in Chapter 4, which attempts to further define the role of RepA in the virus' life cycle.

2.3.3 Transformation of *D. sanguinalis*

Before transforming *D. sanguinalis* with viral replication-interfering Rep constructs, it was important to establish the optimum pressure of the bombardment. The GUS gene (in pDPG208

under the control of the CaMV 35S promoter) was bombarded into *D. sanguinalis* calli at a pressure of 900psi and 1100psi, and three days later GUS expression was determined using a spectrofluorometric assay. In this way, a pressure of 900psi was determined to be optimum (data not shown).

From the results of the transient replication assays, $pRep^{III-Rb+NTP+}$, $pRep^{III-Rb-NTP+}$, $pRep^{III+Rb+NTP+(AS)}$ and $pRep^{Rb-\Delta C2}$ were chosen to transform *D. sanguinalis* for their ability to inhibit MSV replication (see Table 2.4 for a summary of the properties of the transgenic lines obtained). In the first two bombardments, $pRep^{III-Rb+NTP+}$ and $pRep^{III-Rb-NTP+}$ were each co-bombarded with pPHP7503, which contains the bar gene under the control of the CaMV 35S promoter. However, after assaying GUS expression from pAHC25 (ubiquitin promoter) compared with that from pDPG208 (CaMV 35S promoter), it was concluded that genes under the control of the ubiquitin promoter rather than the CaMV 35S promoter were more efficiently expressed in *D. sanguinalis* (data not shown). For the remainder of the *D. sanguinalis* transformations, pAHC25 (Ubi-Bar/ Ubi-Gus) was co-bombarded with $pRep^{III-Rb-NTP+}$, $pRep^{Rb-\Delta C2}$ or $pRep^{III+Rb+NTP+(AS)}$, with the exception of pAHCRep^{III-Rb-NTP+} (Ubi-Bar/ Ubi-Rep) which was bombarded alone.

Having bombarded *D. sanguinalis* with $pRep^{III-Rb+NTP+}$, it was speculated that the non-mutated RBR-interaction motif in the transgene could interfere with plant regeneration due to interaction with the plant RBR protein, which is important in the control of the cell cycle as well as plant differentiation and development (Huntley *et al.*, 1998). For this reason, after selecting the bombarded calli for six weeks on bialophos-containing media (MS-Bi), each bialophos-resistant callus piece (which had arisen from one transformation event) was divided into two, and one half plated onto regeneration media, while the other half was replica-plated and maintained on MS-Bi.

Some of the calli on regeneration media formed leaves and shoots, but none grew into plants, and all eventually died. After 5 months of maintenance on MS-Bi, the replica-plated calli were tested for the presence of the transgene. Out of 27 callus pieces (each an independent line), 22 were positive for $Rep^{III-Rb+NTP+}$ and all were positive for the bar gene. Although five callus pieces contained only the bar gene which should not interfere with regeneration, the number of full-grown plants regenerated from transgenic calli is usually far from 100% efficient even when the transgene is not toxic to growth (as was found with calli transformed with pAHC25). Out of the 22 callus pieces transgenic for $Rep^{III-Rb+NTP+}$, however, one would expect some regeneration of

TABLE 2.4 Properties of transgenic *D. sanguinalis*

¹ Plasmids bombarded	² No. of transgenic lines regenerated/ no. of calli bombarded		Comments
	Rep+Bar	Bar	
pRep ^{III-Rb+NTP+} + pPHP7503	0/ 450 (0%)	0/ 450 (0%)	Regeneration was inhibited in callus pieces shown by PCR to be transgenic for pRep ^{III-Rb+NTP+}
pRep ^{III-Rb+NTP+} + pPHP7503	4 /450 (0.9%)	0/450 (0%)	All lines transgenic for pRep ^{III-Rb+NTP+} were stunted and infertile.
pRep ^{III-Rb+NTP+} + pAHC25	1/ 450 (0.2%)	5/ 450 (1.1%)	Out of 6 regenerated lines, only one was transgenic for Rep+bar. This line was infertile and stunted in comparison with the bar transgenic lines
pAHCRep ^{III-Rb+NTP+}	12/ 900 (1.3%)	N/A	Most (92%) transgenic lines were stunted and infertile
pRep ^{III+Rb+NTP+(AS)} + pAHC25	0/ 450 (0%)	3/450 (0.7%)	Only plants transgenic for the bar gene were obtained.
pRep ^{Rb-ΔC2} + pAHC25	20/ 900 (2.2%)	2/ 900 (0.2%)	Good transformation and regeneration efficiency; a large percentage (70%) of lines transgenic for pRep ^{Rb-ΔC2} were phenotypically normal and fertile
pAHC25	N/A	40/ 900 (4.4%)	Very good transformation and regeneration efficiency; a large percentage of lines were phenotypically normal and fertile

¹pPHP7503 contains the bar gene under the control of the CaMV 35S promoter; pAHC25 contains the bar gene under the control of the maize ubiquitin promoter

²In cases where the Rep construct was co-bombarded with a separate bar construct, not all transgenic plants contained both bar and Rep genes. Thus, the regenerated plants from each co-bombardment have been separated into those transgenic for Rep+bar, and those transgenic for bar alone. Because bar was the selectable marker, there were no cases of a plant being transgenic for Rep alone. All plants transgenic for pAHCRep^{III-Rb+NTP+} contained both Rep and bar genes since they were present on the same bombarded plasmid. The presence of each transgene was determined by PCR using Rep- or bar- specific primers.

plants. Thus, it is possible that the Rep^{III-Rb+NTP+} gene inhibited regeneration, most likely through interaction with RBR via the LxCxE motif.

Lending support to this theory is the fact that transformants of *D. sanguinalis* containing the double mutant Rep^{III-Rb+NTP+}, which cannot interact with RBR, did regenerate. However, the transformation efficiency of all bombardments with this construct was low, and as with the III mutant Rep, the double mutant Rep variant appeared to have a negative effect on aspects of plant growth. Most lines transgenic for pRep^{III-Rb+NTP+} were severely stunted in terms of leaf size and

height (see Fig. 2.6), and all were infertile. At the same time, the identical source calli transformed with pAHC25 alone produced phenotypically normal and fertile transgenic plants, and the transformation efficiency was high. This indicates that the stunting and infertility of plants transgenic for $pRep^{III-Rb-NTP+}$ were caused by the transgene, and were not due to any negative properties of the callus from which the plants were regenerated. Although the LxCxE motif in the RBR-interaction domain was rendered non-functional in $pRep^{III-Rb-NTP+}$, a second motif that may affect the normal functioning of plant cells is found within the NTP-binding domain of the C2. This motif, which contains amino acids that are conserved in the N-terminal DNA-binding domain of the avian myeloblastosis (*myb*) oncogene-homologous regulatory genes, has transcriptional activation activity (Horvath *et al.*, 1998) and it is possible that it was the cause of the stunting and infertility phenotypes. In addition, the C-terminus of the mastrevirus RepA has been shown to interact with plant GRAB (for geminivirus RepA binding) proteins, which are involved in a variety of processes, ranging from lateral root formation to development and senescence (Xie *et al.*, 1999). Interaction with GRAB proteins could therefore also interfere with plant development.

With this in mind, two truncated Rep genes were made of different sizes, both missing the entire C2 ORF and the C terminus of the RepA gene (see Fig. 2.1). As can be seen in Fig. 2.4, $pRep^{Rb-\Delta C2}$ completely inhibited viral replication in BMS. Since this construct cannot interact with RBR and is missing the putative GRAB interaction domain as well as the entire C2 ORF, it was considered the best option for transformation of *D. sanguinalis*. Indeed, a large percentage of plants transgenic for $pRep^{Rb-\Delta C2}$ regenerated into phenotypically normal, fertile adults. It was also notable that while many plants regenerated from calli bombarded with $pRep^{III-Rb-NTP+}$ + pAHC25 and $pRep^{III+Rb+NTP+(AS)}$ + pAHC25 contained only the bar gene (indicating that bar transgenics had a selective advantage over Rep+Bar transgenics), 20 out of 22 transgenic lines regenerated from calli bombarded with $pRep^{Rb-\Delta C2}$ + pAHC25 contained both Rep and bar genes.

Figure 2.6 shows three representative *D. sanguinalis* lines transformed with $pRep^{Rb-\Delta C2}$ (A), $pRep^{III-Rb-NTP+}$ (B) and pAHC25 (C), illustrating the effects of the different transgenes on plant growth and development. Note that (B) is very stunted in comparison with A and C. All Lines of B were infertile.



Figure 2.6 Comparison of transgenic *D. sanguinalis* lines of $pRep^{Rb-\Delta C2}$ (A), $pRep^{III-Rb-NTP+}$ (B) and pAHC25 (C). The plants photographed were typical of most of the plants transgenic for each *Rep* or *bar* construct.

2.3.4 Transformation of Maize Hi-II

Judging by the effects of various *Rep* transgenes on *D. sanguinalis* development and growth, $pRep^{Rb-\Delta C2}$ was chosen to transform maize Hi-II, with pAHC25 carrying the selectable marker. 17 lines transgenic for both $Rep^{Rb-\Delta C2}$ and pAHC25 were regenerated. These were sent to Pannar Ltd for controlled self-pollination and setting of seed. Table 2.5 summarises the properties of the 17 lines transgenic for $pRep^{Rb-\Delta C2}$, called MTA 1-17. Only five of these lines were fertile, and of those only three produced T₁ offspring. These were analysed by PCR for the presence of the $Rep^{Rb-\Delta C2}$ and *bar* transgenes (see Table 2.6 and Fig. 2.7). Fertile T₁ plants were self-pollinated, and the number of kernels that each produced is presented in Table 2.6.

As can be seen in Table 2.5, many T₀ plants were stunted or infertile, and five of the 17 T₀ lines died while being maintained at Pannar Ltd. Although these negative effects could have been caused by the truncated *Rep* gene, many Hi-II plants transgenic for pAHC25 alone showed similar characteristics of stunting and infertility (data not shown). Thus, Hi-II may be more sensitive than *D. sanguinalis* to the presence of any transgene, be it *bar* or *Rep*-based. For example, integration of the transgene into an inappropriate position in the Hi-II genome could disrupt the normal functioning of important developmental genes. Also, the time spent in tissue culture, as well as the bombardment itself, could negatively affect Hi-II more than the hardier *D. sanguinalis*.

TABLE 2.5 The properties of T₀ maize lines (MTA 1-17) transgenic for *Rep*^{Rb-ΔC2} and *bar*

T ₀ Line of MTA	Condition of regenerated plant	Fertility			
		No. of kernels produced by T ₀ lines	No. of kernels planted	No. of T ₁ offspring	No. of fertile T ₁ offspring
1	Good	13	10	1	1
2	Good	0	N/A	N/A	N/A
3	Good	0	N/A	N/A	N/A
4	Good	0	N/A	N/A	N/A
5	Good	1	1	0	N/A
6	Variegated	1	1	0	N/A
7	Good	142	10	10	9
8	Good	41	10	9	4
9	Died	N/A	N/A	N/A	N/A
10	Stunted	0	N/A	N/A	N/A
11	Stunted	0	N/A	N/A	N/A
12	Died	N/A	N/A	N/A	N/A
13	Died	N/A	N/A	N/A	N/A
14	Stunted	0	N/A	N/A	N/A
15	Died	N/A	N/A	N/A	N/A
16	Stunted	0	N/A	N/A	N/A
17	Died	N/A	N/A	N/A	N/A

TABLE 2.6 Transgene segregation and fertility of T₁ maize lines

T ₁ Line of MTA	¹ Presence of <i>Rep</i>	¹ Presence of <i>Bar</i>	No. of kernels produced by T ₁ lines
MTA1 a	×	×	0
MTA7 a	✓	✓	39
MTA7 b	×	✓	16
MTA7 c	✓	✓	42
MTA7 d	×	✓	10
MTA7 e	✓	✓	108
MTA7 f	✓	×	13
MTA7 g	✓	✓	87
MTA7 h	✓	✓	1
MTA7 I	×	×	1
MTA8 a	? (See Fig. 2.7)	×	104
MTA8 b	? (See Fig. 2.7)	×	52
MTA8 c	✓	×	0
MTA8 d	✓	✓	37
MTA8 e	✓	✓	0
MTA8 f	×	×	0
MTA8 g	✓	✓	61
MTA8 h	✓	✓	0
MTA8 I	✓	✓	0

¹× = absence of transgene

✓ = presence of transgene

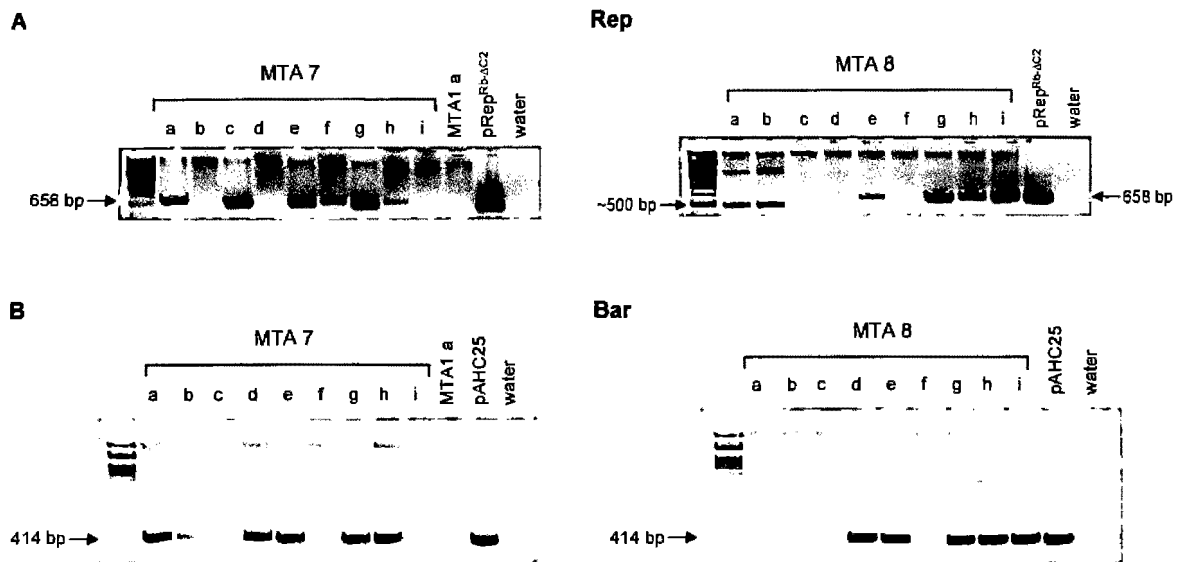


Figure 2.7 Inheritance of *Rep*^{Rb-ΔC2} and *bar* transgenes in T₁ maize. T₁ maize lines were analysed for the presence of *Rep*^{Rb-ΔC2} and *bar* transgenes by PCR using RepΔC2-specific primers (658 bp product, A) and Bar-specific primers (414 bp product, B). MTA1 a, MTA7 a-i, and MTA8 a-i are the offspring of the T₀ MTA1, 7 and 8 lines. The first lane of each gel contains λ DNA digested with *Pst*I as a size marker. The plasmids pRep^{Rb-ΔC2} (A) and pAHC25 (B) are positive controls. Note the ~500 bp band for MTA8 a and b, which may be a non-specific product, or may be an amplification product of a transgene with an internal deletion.

Of the nine T₁ MTA7 lines, six contained the *Rep*^{Rb-ΔC2} gene, and seven contained the *bar* gene. Five out of the nine lines were positive for both *Rep*^{Rb-ΔC2} and *bar* genes. Similarly, of the nine T₁ MTA 8 lines, six contained the *Rep*^{Rb-ΔC2} gene, and five contained the *bar* gene. As with MTA7 lines, five of the nine MTA8 T₁ lines were positive for both *Rep*^{Rb-ΔC2} and *bar* genes. These ratios indicate Mendelian inheritance. However, it must be noted that a ~500 bp product was amplified from MTA 8 a and b lines (which were not taken as being *Rep*^{Rb-ΔC2}-positive) using RepΔC2-specific primers, which could indicate the presence of a transgene with an internal deletion. This can be confirmed by sequencing the PCR product.

The fact that some lines inherited the *Rep*^{Rb-ΔC2} but not the *bar* gene, and vice versa, indicates that the co-bombarded plasmids did not integrate at the same loci. This could be useful when developing transgenic maize for commercial purposes, since lines transgenic for *Rep* but not *bar* could be selected, thus removing the undesirable bialophos resistance gene marker from the genome.

Two T_0 lines, MTA7 and MTA8, produced fertile T_1 offspring. The seed from nine MTA7 T_1 lines and four MTA8 T_1 lines will be tested for resistance to MSV by *Agrobacterium*-mediated delivery of a variety of MSV strains (agroinoculation; Grimsley *et al.*, 1987) into three-day old T_2 seedlings. The principle of agroinoculation is explained in Chapter 3. Before challenging maize with MSV, however, it was considered important to first test selected transgenic *D. sanguinalis* lines for resistance to MSV.

2.3.5 Challenge of *D. sanguinalis* Transgenic Plants with MSV

Due to the large number of *D. sanguinalis* lines transgenic for $pRep^{III-Rb-NTP^+}$ (a total of 17) and $pRep^{Rb-\Delta C2}$ (a total of 20), a few selected lines (chosen for their robustness compared with other lines) were challenged with MSV. These were A6A1, A6B14, A6B15, A6B16, A6C8 and A6D10 (A6 lines, all transgenic for $pRep^{III-Rb-NTP^+}$) and TA1, TB1, and TB2 (T lines, all transgenic for $pRep^{Rb-\Delta C2}$). Because all A6 lines were infertile, plants were regenerated from callus initiated from the parent transgenic lines, as an alternative to T_1 offspring. Callus was also initiated from the T lines even though they produced viable seed, for a number of reasons: (1) it was useful to keep a stock of each line as callus; (2) it was simple to regenerate many plants of each line from the stock callus, all of which would be genetically identical to the parental line; (3) the challenge of A6 and T lines would be more comparable if the plants came from the same source (i.e. callus instead of seed) and (4) the seed from the T lines could be kept for future work if the line looked promisingly resistant to MSV.

The first two challenges (challenge experiments 1 and 2) with MSV, carried out on lines A6A1, A6B14, A6B15 and A6B16 by exposure to viruliferous leafhoppers carrying MSV, were subsequently considered to be flawed (explained below). However, because the parental A6B14, A6B15 and A6B16 lines were eventually lost (due to fungal contamination of the stock callus) and therefore they could no longer take part in any further trials, the results from these two experiments are summarised in Table 2.8.

In the first challenge experiment, the parental T_0 plants (regenerated from bombarded callus) A6A1, A6B14, A6B15 and A6B16 were placed in a gauze-covered wooden cage, along with a *D. sanguinalis* plant transgenic for pAHC25, and a non-transgenic maize plant (cv Jubilee) as controls. ~ 20 viruliferous leafhoppers were introduced into the cage, and every second day were

TABLE 2.8 Results from two experiments challenging lines transgenic for pRep^{III-Rb-NTP+}

Experiment	¹ Plant Lines	² Symptoms	³ Presence of viral DNA
1	A6A1	-	-
1	A6B14	-	-
1	A6B15	-	-
1	A6B16	-	-
1	<i>D. sanguinalis</i> control	+++	+++
1	Maize control	+++	+++
2	A6A1 I	+++	+++
2	A6A1 II	+++	+++
2	A6A1 III	+++	+++
2	A6A1 IV	+++	+++
2	A6A1 V	-	+
2	A6A1 VI	+++	+++
2	A6A1 VII	-	+
2	A6A1 VIII	+++	+++
2	A6A1 IX	+++	+++
2	A6A1 X	+++	+++
2	<i>D. sanguinalis</i> control I	+++	+++
2	<i>D. sanguinalis</i> control II	+++	+++

¹ Lines of A6A1, A6B14, A6B15 and A6B16 were transgenic for pRep^{III-Rb-NTP+}. In experiment 1, parental lines were used. In experiment 2, plants regenerated from callus that had been initiated from line A6A1 were used. In both experiments, the *D. sanguinalis* control was transgenic for pAHC25.

² Symptoms were scored four weeks after the start of the trial, and rated using the following scale: - = no streak; + = mild streak and recovery; ++ = moderate stippled streak to continuous streak; +++ = severe continuous streak.

³ The presence of viral DNA was analysed by PCR four weeks after the start of the trial. + = very faint band (barely detectable); ++ = faint, but easily detectable band; +++ = very bright band.

shaken off the plants and allowed to resettle to ensure that most plants would be fed on. After four weeks, DNA was extracted from each of the plants' leaves and analysed for the presence of MSV by PCR. The same procedure was followed for the second challenge experiment, this time using 10 plants regenerated from stock callus initiated from the A6A1 parental line, and two control plants transgenic for pAHC25. As can be seen in Table 2.8, some of the *Rep* transgenic plants did not become infected with MSV while all the controls did. In particular, all four parental lines appeared resistant, although only two out of ten plants regenerated from the A6A1 callus were unsusceptible to MSV. However, it could not be discounted that the uninfected plants were not fed on by the leafhoppers, or at least for long enough for viral transmission to occur. The time taken for the leafhopper's proboscis to reach the phloem and transmit the virus can take as long as three hours from initial access (Bosque-Pérez, 2000). Because the *Rep* transgenic plants' leaves were very small and perhaps not very appetising to the leafhoppers in comparison to the control plants (see Fig 2.6), the leafhoppers may have preferentially fed on the control plants, skewing the results. Therefore, in subsequent challenge experiments, the viruliferous leafhoppers were placed in a small vial containing a ~1 mm slit, through which a single leaf from a transgenic plant

was inserted, ensuring that the leafhoppers in a particular vial could only feed on one leaf of one plant. Three vials were placed at different positions on each plant (Fig. 2.8A). In total the same number of leafhoppers were fed on each plant, as far as possible ensuring uniformity of the MSV inoculum received by the plant.

A preliminary challenge employing this method was carried out at Pannar Ltd by Dr. R. Kloppers, using five plants of line A6A1 and three plants of line A6C8, as well as a field-grown non-transgenic *D. sanguinalis* control (challenge no. 3). Whereas the non-transgenic control plant developed symptoms 10 days after the start of the challenge, no symptoms were observed on any of the transgenic plants throughout the two-month trial (Fig. 2.8B and C). Viral DNA levels in the challenged transgenic plants were shown by PCR to be either non-existent or significantly lower than those in the non-transgenic control plant (Fig. 2.8D). The A6C8 line appeared to be particularly resistant. Not only did the plants develop no symptoms, but viral replication was also completely inhibited in two and significantly reduced in one out of the three challenged A6C8 plants. Viral replication in three out of five A6A1 plants was greatly reduced compared with that in the non-transgenic control plant, and two out of the five A6A1 plants contained no detectable viral DNA.

In a fourth challenge, six plants of line TB1 and one plant of line TB2, both transgenic for $pRep^{Rb-AC2}$, and seven control plants transgenic for pAHC25 were challenged with MSV (Table 2.9). Leafhoppers in vials were fed first on the control plants for two days and then transferred to the plants of line TB1 and TB2 and allowed to feed for two days. The plants were monitored for symptom development for five weeks. The same procedure was followed for challenges five (four control plants and four plants of line TB1) and six (one control plant and one plant of line A6C8). Symptoms and viral DNA amplified by PCR from total DNA extracted from challenged plants' leaves were scored as in Table 2.8.

In the experiments shown in Table 2.9, each *bar* transgenic control plant can be directly compared with its *Rep*-transgenic counterpart (e.g. control 1 with TB1 I; control 2 with TB1 II, etc), since the same leafhoppers, transferred from the control to the corresponding *Rep* transgenic, fed on both plants. Leafhoppers remain viruliferous for their lifetime, and therefore should transmit the virus throughout the four days they feed on the plants (two days on the *bar* controls followed by two days on the *Rep* transgenics). It can be seen that in experiment 4, *bar* controls 1,

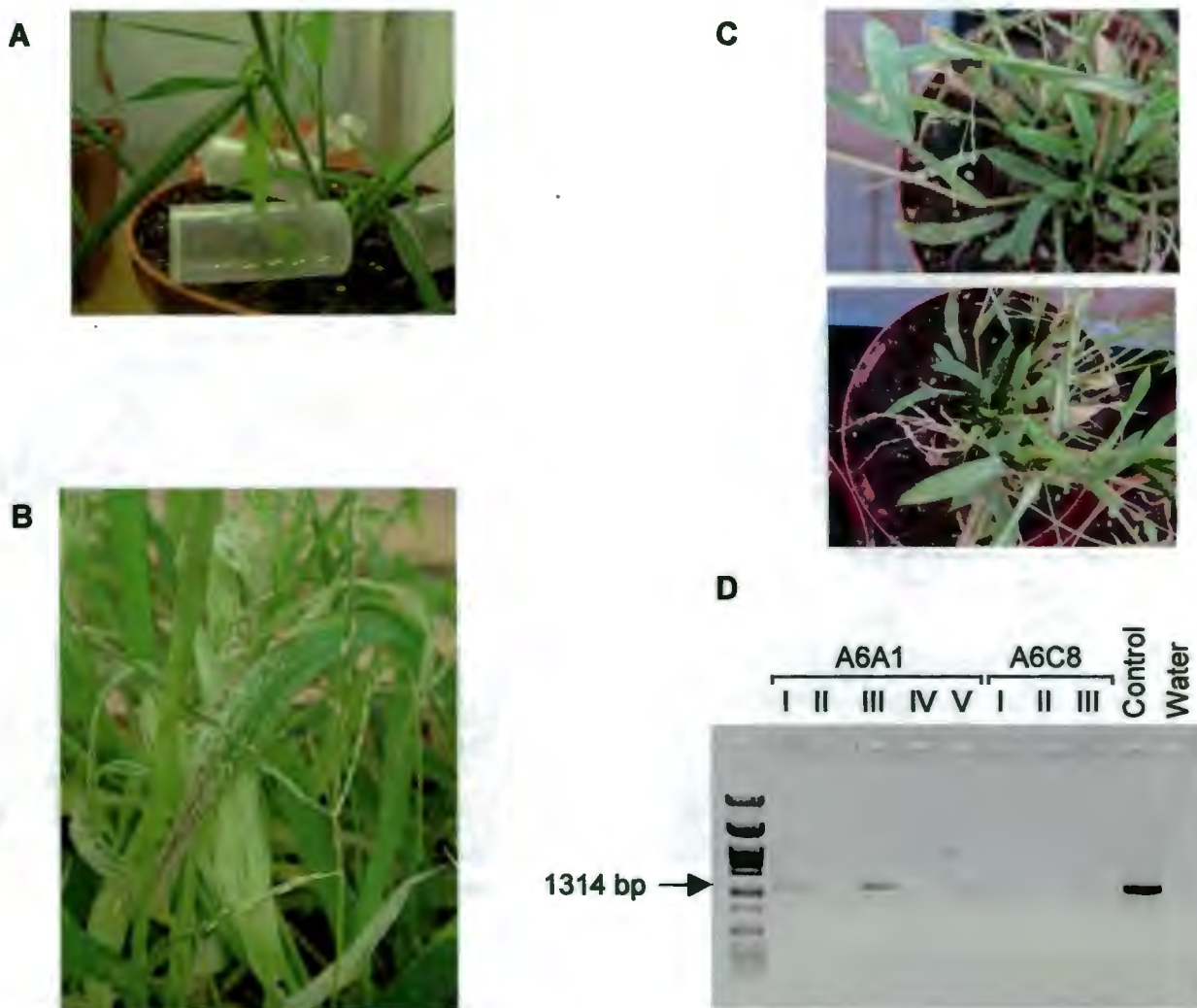


FIGURE 2.8 Challenge with MSV of lines A6A1 and A6C8, both transgenic for $pRep^{III-Rb-NTP+}$. Four leafhoppers were placed in each vial, and three vials placed at different positions on each plant (A). The *D. sanguinalis* control became severely infected (B) while the *Rep* transgenics were asymptomatic (C, showing two lines of A6A1). PCR analysis of viral replication in leaves of the transgenic and non-transgenic *D. sanguinalis* (D) showed greatly reduced levels of viral replication in the transgenics' leaves compared with the control. The degenerate primers MSVDEG1 and MSVDEG2 (described in section 2.2.3.3) were used to amplify the putatively mixed MSV population transmitted by the leafhoppers. These primers were designed to amplify DNA fragments from all African streak viruses sequenced up until 1996 (Willment *et al.*, 2001). Note in (C) that the leaves of A6A1 plants are stunted in comparison with the control in (B). Although faint speckles can be seen on some of the transgenics' leaves, these looked more like leafhopper-feeding damage rather than symptoms. This was confirmed by the PCR results.

3, and 6 developed symptoms 13 days after the start of the challenge, while the corresponding TB1 I, III, and VI plants never developed an infection.

TABLE 2.9 Results of challenge experiments 4, 5 and 6

Challenge experiment number	Bar transgenic plants (controls)	Time to develop symptoms (Days)	¹ Symptom severity after 35 days	² Presence of viral DNA	³ Rep transgenic plants	Time to develop symptoms (Days)	¹ Symptom severity after 35 days	² Presence of viral DNA
4	1	13	+	-	TB1 I	-	-	-
	2	12	+++	++	TB1 II	14	+	+
	3	13	++	+	TB1 III	-	-	-
	4	12	+	+	TB1 IV	14	+	-
	5	12	+	+	TB1 V	14	+	-
	6	13	+++	+++	TB1 VI	-	-	-
	7	13	+++	+++	TB2 I	14	+++	+++
5	1	11	+++	ND	TB1 VII	15	++	ND
	2	-	-	ND	TB1 VIII	-	-	ND
	3	11	+++	ND	TB1 IX	11	++	ND
	4	11	+++	ND	TB1 X	11	+++	ND
6	1	11	+++	+++	A6C8 VI	-	-	-

¹Symptoms were scored five weeks after the start of the trial, and rated using the following scale: - = no streak; + = mild streak and recovery; ++ = moderate stippled streak to continuous streak; +++ = severe continuous streak.

²The presence of viral DNA was analysed by PCR five weeks after the start of the trial. + = very faint band (barely detectable); ++ = faint, but easily detectable band; +++ = very bright band. ND = Not determined

³TB1 and TB2 plants are transgenic for pRep^{Rb-ΔC2}; A6C8 is transgenic for pRep^{III-Rb-NTP+}

Others, such as TB1 II, did become infected after 14 days, but showed a recovery phenotype (+) and correspondingly a low level of viral DNA was amplified from this plant's leaves. Conversely control 2, which became infected after 12 days, had a severe infection (+++), and high levels of viral DNA in its leaves. The remainder of the TB1 plants in challenge 4 did not differ greatly in their response to viral infection from their control counterparts. The line TB2 was highly susceptible to virus infection, as was its corresponding control.

The TB1 plants in challenge experiment no. 5 showed less resistance to viral infection than those in challenge 4, all but one (TB1 VIII) becoming infected. The fact that TB1 VIII was not symptomatic by day 35 was probably due to weakly- or non-virulent leafhoppers, since the corresponding control plant also remained symptomless. This illustrates the importance of having the same leafhoppers feed on both control and Rep transgenic plants. The infection of TB1 VII was slightly delayed compared with its control, although only by four days, and symptoms were slightly less severe than those of the control. TB1 VII also developed symptoms in fewer stalks than the control. The same effects could be seen in TB1 IX, whose symptoms were slightly attenuated and limited to three stalks, compared with ten in the control, which developed a very severe infection.

In challenge no. 6, A6C8 developed no symptoms (as in challenge 3) and no viral DNA was detected in the plant's leaves, whereas the control developed a severe infection and high levels of viral DNA were amplified from its leaves by PCR.

In a seventh and final trial, five plants initiated from a line transgenic for $pRep^{Rb-\Delta C2}$ (TA1) and five non-transgenic *D. sanguinalis* plants were challenged with MSV. The analysis of the symptoms was different from the previous trials, in that symptom severity was scored using a key of symptoms (see Appendix B) ranging from 1% (very mild) to 95% (very severe) chlorosis.

All control plants developed symptoms between 15 and 17 days after the start of the challenge, whereas the transgenic TA1 plants had different responses to the challenge (Table 2.9). All TA1 plants showed some kind of resistance, from immunity (TA1 V) to delayed symptom development (TA1 II) to attenuated symptoms compared with the corresponding control (TA1 I and III). The severity of the symptoms of TA1 IV equaled that of the control. However, whereas the control developed symptoms in all leaves that emerged after MSV transmission, only the leaves of one stalk of TA1 IV developed symptoms. A similar pattern could be seen in all the other TA1 plants, indicating that the spread of the virus from initially infected leaves was inhibited. The restriction of symptoms in plant TA1 I compared with the corresponding control plant is shown in Fig. 2.9A, B and C.

To analyse the viral infection at a molecular level, the DNA of three leaves taken from each plant 47 days after the start of the trial was subjected to PCR (Fig. 2.9D). From the TA1 plants showing symptoms, a young (newly-emerged) symptomatic leaf (A), a young asymptomatic leaf immediately adjacent to the symptomatic leaf (B), and a young asymptomatic leaf on a different stalk distant from the symptomatic leaf (C) was analysed by PCR to determine if there was any virus spread from the symptomatic leaf. From the non-transgenic control plants three young leaves were taken from three different areas of the plant. Figure 2.9D shows that viral DNA was present only in the symptomatic leaves of the transgenic plants TA1 Ia, IIIa, IVa and all control plants. In the TA1 plants, even leaves immediately adjacent to the symptomatic leaves contained no viral DNA. TA1 V at this stage in the experiment (47 days) did not have any new symptomatic leaves - the one stalk containing an infected leaf had died and no new infection emerged, which correlates to the lack of viral DNA in any of the leaves analysed. Visible symptoms on TA1 III only emerged on one leaf on day 47, and subsequently viral DNA was amplified from this leaf (IIIA) by PCR.

TABLE 2.9 Results of challenge experiment no. 7 using five plants of line TA1, transgenic for $pRep^{Rb-\Delta C2}$, and five non-transgenic control plants

Challenged Plant	Time to develop symptoms (days)	¹ Symptom severity after 35 days	Comments
Control 1	17	93.7±2.5	All new leaves subsequent to the challenge were symptomatic
Control 2	17	76.6±8	All new leaves subsequent to the challenge were symptomatic
Control 3	15	85±4.6	All new leaves subsequent to the challenge were symptomatic
Control 4	17	41.3±6.1	All new leaves subsequent to the challenge were symptomatic
Control 5	17	86.3±6.1	All new leaves subsequent to the challenge were symptomatic
TA1 I	19	59.1±5.8	Symptoms restricted to three stalks. The rest of the plant remained symptomless
TA1 II	47	-	Very mild symptoms seen only after 47 days near the end of the trial period; restricted to 1 stalk. The rest of the plant remained symptomless
TA1 III	22	36.4±11.1	Symptoms restricted to 1 stalk. The rest of the plant remained symptomless
TA1 IV	18	42.4±15.3	Symptoms restricted to 1 stalk. The rest of the plant remained symptomless
TA1 V	-	-	No symptoms developed throughout the trial

¹average % chlorosis per leaf; 95% CI

In conclusion, the challenges of various transgenic lines with MSV produced varied although encouraging results, summarised below.

Of 16 challenged A6A1 plants in three separate challenge experiments, eight showed MSV resistance, manifesting itself in these plants as no symptom development and vastly lowered viral replication compared with controls.

The parental lines A6B14, A6B15 and A6B16 did not become infected in challenge no.1, but it could not be discounted that the “resistance” may have been due to non-transmission of MSV by the leafhoppers due to preferential feeding on the more succulent-looking control plants.

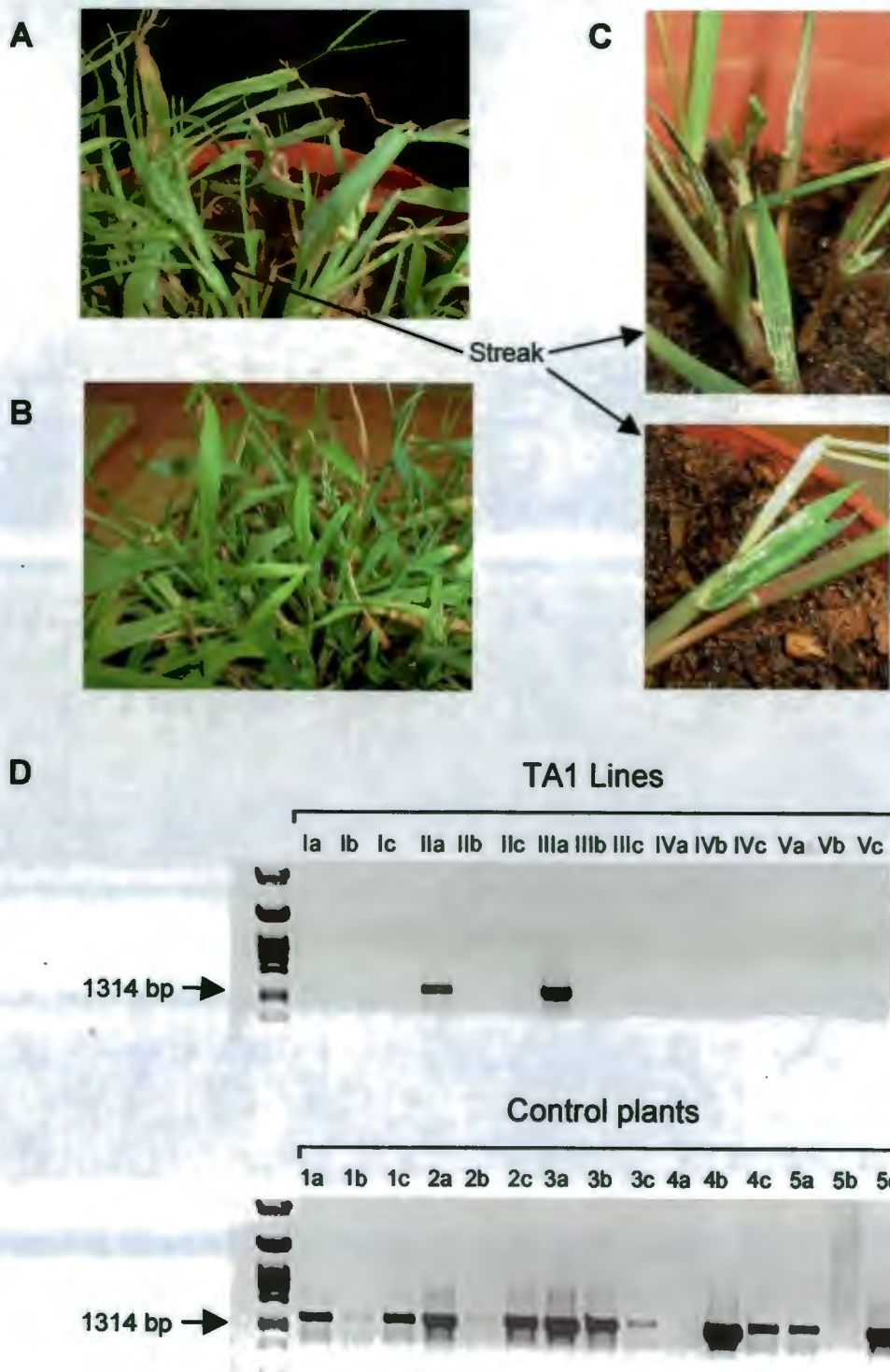


FIGURE 2.9 The challenge of line TA1, transgenic for $pRep^{RB-AC2}$, with MSV. Plant TA1 I is shown in (A) and (B). In (A), symptoms on a leaf of one stalk are shown. In (B), it is clear that leaves on other stalks of the plant are symptomless. In contrast, the leaves of all stalks of the corresponding control plant (C) are severely infected, and the plant is close to dying. The photographs were taken 35 days after the start of the challenge. Viral replication in the plants' leaves was analysed by PCR (D) 47 days after the start of the challenge. For TA1 plants, a = symptomatic leaf (except in the case of non-symptomatic plants), b = adjacent non-symptomatic leaf; c = distant non-symptomatic leaf. For control plants, a, b and c are leaves taken from 3 distant areas of the plant.

Three lines transgenic for $pRep^{Rb-\Delta C2}$, TB1, TB2 and TA1 had different responses to the challenge with MSV. While TB2 was highly susceptible, some plants of TB1 showed differing degrees of resistance to MSV, as seen in Table 2.9. Overall, however, TB1 does not look like a promising line in terms of reliable, consistent MSV resistance.

The $pRep^{Rb-\Delta C2}$ transgene in all TA1 plants appeared to confer phenotypes of delayed, attenuated symptoms, as well as potentially restricting movement or spread of the virus. The TA1 plants that developed symptoms only did so in three or fewer stalks, which were most likely the points of entry for the virus; i.e. the stalks on which the viruliferous leafhoppers were fed. From the point of entry it is possible that the virus was not able to replicate to titres high enough to enable spread to surrounding stalks.

The most encouraging line in terms of consistent MSV resistance was A6C8. Four out of four plants in two separate challenge experiments never developed symptoms, while the controls were severely infected. In addition, only one of the four challenged plants contained any viral DNA detectable by PCR. Since the possibility that the titre of MSV in viruliferous insects may decrease when fed first on control plants (non-transformed or pAHC25-transformed) was considered, a larger-scale challenge experiment involving more A6C8 plants was attempted, in which the leafhoppers were to be fed first on the transgenic plants. Unfortunately, after four years in tissue culture, the callus appears to have lost its ability to regenerate into plants. Indeed, the reason only one A6C8 plant was used in challenge no. 6 is because it was the only one to regenerate from a plateful of callus. Thus, the question of whether the efficiency of transfer of MSV to the transgenic plants is affected by being passaged first through non-transgenic plants remains to be answered. However, bearing in mind that A6C8 plants developed no symptoms at all, it is perhaps unlikely that virus titres in leafhoppers would decrease so dramatically over a period of two days as to render the leafhoppers non-viruliferous.

Another challenge, using plants of line A6D10 (transgenic for $pAHCRep^{III-Rb-NTP+}$) was aborted because the regenerated plants were too stunted to enable the attachment of leafhopper-containing vials at three separate positions on each plant, and at an early stage in their development the plants stopped growing any new leaves. This could have been due to effects of the $Rep^{III-Rb-NTP+}$ transgene, or to too much time spent in tissue culture as callus (or a combination of both). There are many more transgenic lines being maintained as callus, and even if they cannot regenerate plants they may still be useful in identifying transgenes that confer MSV resistance. For example,

different transgenic callus lines can be bombarded with MSV, and the levels of replication supported by the transgenic callus compared with control callus can be determined using quantitative PCR, as in the transient assays in BMS.

The differences in the degree of resistance of the challenged transgenic plants could be due to differing expression levels of the transgenes. This was determined using reverse transcription PCR, described in section 2.3.6.

2.3.6 Expression of Transgenes in *D. sanguinalis* Lines used in MSV-Challenge Experiments

The expression of GUS (in cases where the transforming plasmid contained the *uidA* gene), Bar and Rep was determined in the transgenic lines that were challenged with MSV, by histochemical staining and RT-PCR of RNA extracted from the transgenic callus lines. Figure 2.10 shows GUS stains of lines A6B14, A6B15, A6B16, A6C8, TA1, TB1 and TB2. A6A1 and A6D10 were not included since neither line was transformed with the GUS gene. Since there were no Rep antibodies available for analysis of Rep expression by western blotting, expression of GUS gives an idea of how well the transgenes are being expressed by each line, and RT-PCR of the *Rep* transcript provides indirect evidence that the protein is being expressed. As can be seen in Fig. 2.10, GUS expression was not detected in leaf material as clearly as it was in callus. In (A), the leaves of A6B15 and TA1 were not positive for GUS, while blue spots could be seen in calli of the same lines. The leaf of A6C8, although positive for GUS, also did not stain as well as the A6C8 callus. In a previous study Chen (1996) found that while GUS activity in transgenic plants' leaves was not detectable by histochemical staining, a more sensitive fluorescence assay (where the protein was extracted from the leaves) detected GUS activity in the majority of the same plants. Thus, the lines TB1 and TB2 may be expressing GUS, albeit at lower levels than A6C8, even though the leaves gave negative results. Despite the fact that genetically identical cells should comprise the calli of an individual line, expression of GUS was sporadic in all lines but A6C8. This could provide a clue as to why plants regenerated from callus had different responses to MSV infection, from sensitivity to immunity. If Rep expression in callus cells follows the same pattern as that of GUS, it is conceivable that some plants may regenerate from cells expressing Rep, while others may arise from cells not expressing Rep. The majority of A6C8 calli expressed GUS, which is consistent with the fact that 100% of plants regenerated from A6C8 calli were resistant to MSV infection, i.e. it is probable that all the plants were expressing *Rep*^{III-Rb-NTP+}. Unfortunately, leaf samples taken before the plants were challenged with MSV

were used for DNA extractions to confirm the presence of the transgene, and not for transgene expression analysis.

While A6C8 and A6D10 contain relatively high levels of Rep transcript (Fig. 2.11), cDNA bands in TA1, TB1 and TB2 samples are very faint. The Rep transcript levels of all lines correlate with the GUS expression levels shown by histochemical staining of the calli. Since the RNA was extracted from calli, it follows that calli in which most cells are expressing the transgene, such as A6C8, will collectively have higher levels of transcript, while in other lines a low percentage of callus cells may be expressing the transgene, diluting the transcript out. Only calli expressing Bar can survive on media containing bialophos (on which the callus lines were constantly maintained) explaining the higher levels of bar transcript in all lines.

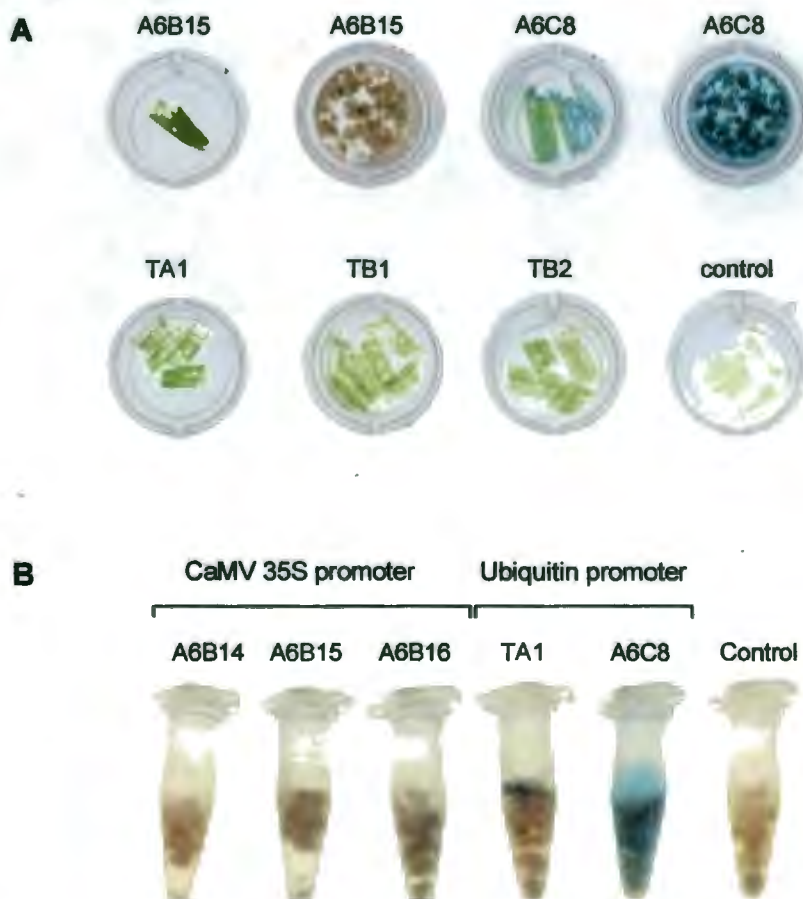


FIGURE 2.10 GUS stains of callus and leaf material of *Rep* transgenic lines. A6B14, A6B15, A6B16 and A6C8 are lines transgenic for *Rep*^{III-Rb-NTP+}. TA1, TB1 and TB2 are lines transgenic for p*Rep*^{Rb-ΔC2}. The control is leaf or callus material from a non-transgenic plant. In (A) leaf material was treated with the X-Gluc substrate and then destained in order to see blue spots more easily. Most of the leaf material did not stain blue, but blue-stained calli of two lines, A6B15 and A6C8, are shown to compare the sensitivity of the histochemical staining technique in leaves and calli. In (B) stained calli are shown. A6B14, A6B15 and A6B16 were transformed with the GUS gene under the control of the CaMV 35S promoter, while in A6C8 and TA1 GUS is expressed from the maize ubiquitin promoter.

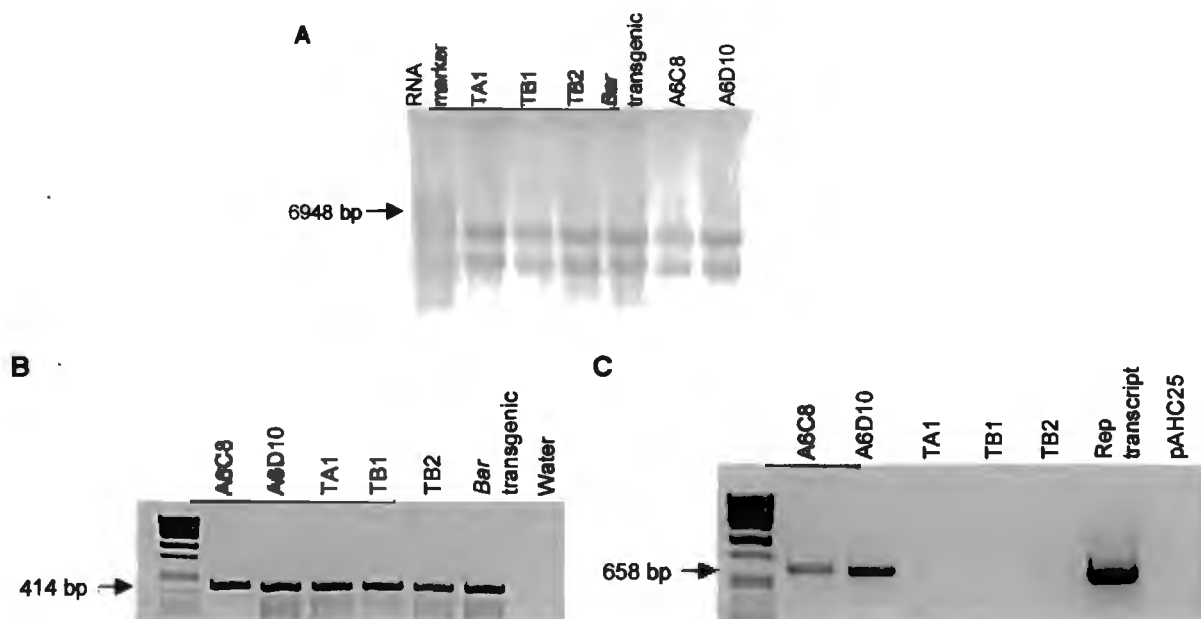


FIGURE 2.11 Analysis of transgene expression by RT-PCR. RNA was extracted from callus of lines A6C8, A6D10, TA1, TB1, TB2 and a line transgenic for Bar (A), and RT-PCR was carried out using primers specific for the Bar gene (414 bp product, B) and the Rep gene (658 bp product, C). RNA from (A) was diluted 1 in 10, and 1 μ l of each diluted sample was used in the RT-PCR reaction, therefore RNA expression levels in B and C can be compared with RNA concentration in (A). Before RT-PCR, the RNA samples were treated with DNase, and a control reaction was carried out using the same RT-PCR conditions except that the reverse transcriptase enzyme was omitted to confirm that amplification was not due to contaminating genomic DNA. Since there was no amplification product from any of the DNase-treated samples in this control reaction (data not shown), the bands in (B) and (C) are clearly products from reverse transcription of the transcript and subsequent PCR of the cDNA.

2.3.7 Analysis of the MSV Strain Composition Transmitted by Viruliferous Leafhoppers.

The viruliferous leafhoppers used to challenge *D. sanguinalis* transgenic plants were obtained from the seed company Pannar Ltd, where they had been fed on symptomatic maize leaves collected from various locations in South Africa and Zimbabwe. The leafhoppers were subsequently maintained at UCT on *Zea mays* cv Jubilee. Leafhoppers from the same population have been used by Pannar Ltd to challenge the company's MSV-resistant hybrids, which have shown resistance in Nigeria, Kenya, Zambia, Malawi, Zimbabwe, Mozambique and South Africa (M. Barrow, pers. comm). In order to determine which viral strains were inoculated by the leafhoppers into the challenged *D. sanguinalis* plants, total DNA was extracted from maize leaves containing leafhopper-transmitted MSV, from which 21 full-length genomic clones were isolated. PCR-amplified fragments from these viruses were analysed by restriction length polymorphism (RFLP) to distinguish different virus strains from one another, as described in

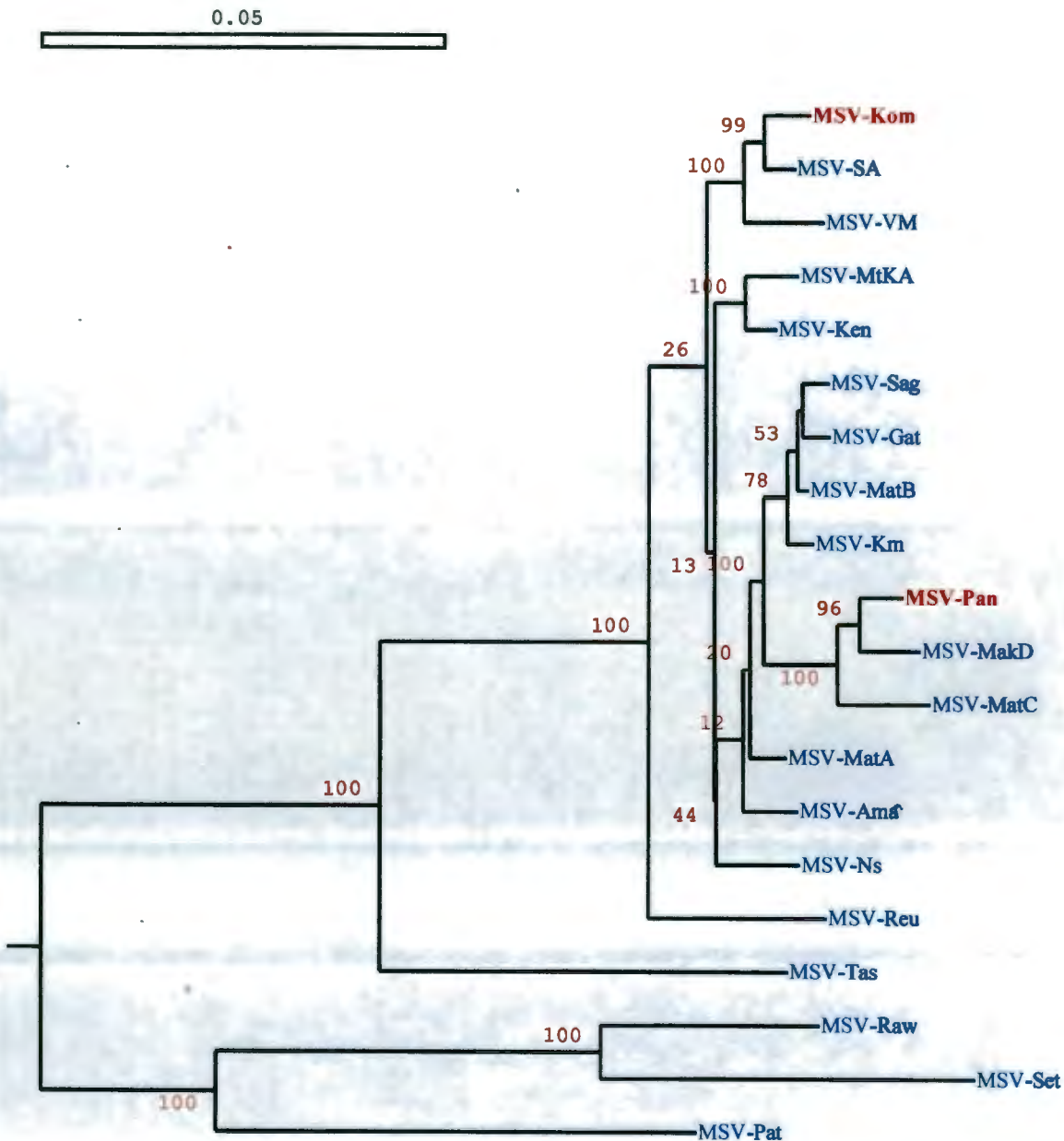


FIGURE 2.12 A rooted neighbour joining tree showing the grouping with other African streak viruses of an MSV strain (MSV-Pan) isolated from a leafhopper-transmitted virus population in maize. The majority of strains isolated from the maize sample consisted of this strain, which groups with MSV-MaKD and MSV-MatC, and is classified as an A₃ subtype (Martin *et al.*, 2001; see text for details). Numbers associated with nodes represent the percentage of 1000 bootstrap iterations supporting the nodes.

Willment *et al.* (2001). From the RFLP it was clear that the MSV population present in the maize sample consisted entirely of one type of MSV. Two of these clones were sequenced, confirming that both full-length clones were identical. The sequence of the full-length viral genome was aligned with a wide variety of other African streak virus genomic sequences, and the relationship

between the MSV genomes was determined using a rooted neighbour joining tree (Fig. 2.12). The MSV sequences in Fig. 2.12 have been divided into five major strain groupings (Martin *et al.*, 2001) designated A-E. In this MSV diversity study, the overwhelming majority of MSVs obtained from maize were strain A isolates. The MSV-A isolates were further classified into six subtypes, A₁-A₆. The isolate isolated from maize in this study (called MSV-Pan) was closely related to MSV-MaKD and MSV-MatC, and was classified as an A₅ subtype, one of the two most severe subtypes analysed in Martin *et al.* (2001). The A₅ subgroup is also the most divergent MSV A genotype grouping yet found in Africa. It is interesting to note that while isolates in the A₅ subgroup produce severe symptoms in the moderately resistant maize genotype PAN6099 (Martin *et al.*, 2001), in this study several transgenic *D. sanguinalis* plants challenged with MSV-Pan developed no symptoms at all.

Although the RFLP and sequencing analysis of the 21 virus clones isolated in this study indicate that the majority of the virus population introduced into *D. sanguinalis* plants by leafhoppers were of one type, there may have been other, less abundant genotypes present that were not isolated. It is likely that a mixed MSV population originally existed, but the isolates best adapted to maize overtook the population while being maintained in *Zea mays*. The fact that the MSV-A isolates are the most widely distributed maize-infecting MSVs, and that MSV-Pan did not group with MSV-Kom (the strain from which the transgenes were derived), indicates that the protection afforded by the mutated and truncated Rep transgenes will be effective against other MSV strains prevalent in Africa, including the most virulent A₅ subtypes.

2.4 CONCLUSION

The three-system approach proved to be highly effective in reaching the goal of producing MSV-resistant plants. The transient expression studies not only served as a very useful guide for the direction to take in developing MSV-resistant transgenic plants, they also provided some interesting insights into the mechanics of MSV replication. For example, a Rep construct with a mutation abolishing the protein's interaction with the host RBR protein, previously thought to be required for viral replication (Xie *et al.*, 1995), surprisingly supported the *trans*-replication of the viral genome to levels 30 fold higher than when pKom602 was bombarded alone into BMS. Even more surprising, a highly truncated 179-aa Rep protein with a deleterious mutation in motif III (which is essential for the initiation of RCR) enhanced viral replication 4-fold when provided in *trans* with pKom602. It must be borne in mind that in all the replication assays presented in

this study, MSV-Kom would express the wt Rep gene at much lower levels than the Rep genes under the control of the ubiquitin promoter provided in *trans*. Therefore, the effects of the truncated Rep constructs on viral replication (enhancement and inhibition depending on the size of the truncation) are likely to be as a result of interaction or interference with the wt Rep protein, or the result of effects, positive or negative, on the wt protein's function.

The transgenic *D. sanguinalis* lines also provided some insights into the interaction of Rep with host factors. For instance, none of the calli bombarded with $pRep^{III-Rb+NTP+}$ regenerated, even though they were shown by PCR to be transgenic for the Rep and bar genes. The inhibition of regeneration was probably due to the transgene having an intact RBR-interaction domain, which may negatively affect regeneration through interference with the plant cell cycle. Significantly, calli transformed with $pRep^{III-Rb-NTP+}$ did regenerate, although at a lower efficiency than calli containing the truncated transgene. In addition, $pRep^{III-Rb-NTP+}$ appeared to have a negative effect on aspects of plant growth and development. Various motifs that may affect the normal functioning of plant cells include the putative GRAB-binding domain at the C-terminus of RepA and the *myb*-binding domain in the C2. In addition, numerous other interactions between geminivirus Rep and host proteins involved in cell division and development have been reported (Kong and Hanley-Bowdoin, 2002). The fact that plants transformed with the truncated Rep gene $pRep^{Rb-\Delta C2}$, in which amino acid sequences such as the GRAB- and *myb*-binding domains were deleted, supports the theory that domains residing in the RepA C terminus and the Rep C2 ORF have a negative impact on plant cell biology.

Results from experiments challenging the *D. sanguinalis* transgenic lines with MSV are encouraging, with all plants from the TA1 line (transgenic for $pRep^{Rb-\Delta C2}$) and the A6C8 line (transgenic for $pRep^{III-Rb-NTP+}$) showing phenotypes of resistance, including a delay in symptom development, a recovery from early symptoms, and an absence of virus symptoms at all stages.

GUS expression in all callus lines except A6C8 indicates that it is unlikely that all cells of a callus line express the transgene. This could explain why genetically identical plants regenerated from cells of a callus line, such as TB1 and A6A1, had different responses to MSV infection, some plants showing resistance and others susceptibility. The low levels of transcript amplified by RT-PCR in TA1, TB1 and TB2 could also be due to a low percentage of cells (comprising the callus from which RNA was extracted) expressing the transgene. There are alternative possibilities for the low expression levels of the truncated Rep genes compared with those of the full-length A6C8 and A6D10 Rep genes. For example, the truncated transcripts may be more

unstable than the full-length counterparts, or they may be perceived as aberrant and subsequently silenced, while the full-length transcripts are perceived as "normal". However, the resistance mechanisms in both the transient and transgenic assays point to dominant negative interference of MSV wt Rep by the transgene product. For example, while the sense version of $pRep^{Rb-\Delta C2}$ inhibited viral replication in BMS, the antisense version of the same gene had no effect on viral replication. In the transgenic assays, the high levels of *Rep* transcript and GUS expression in A6C8 was reflected in this line showing immunity in 100% of challenged plants. In future, these questions can be answered by determining *Rep* transcript levels in individual plants, rather than the callus line, and correlating that to the resistance or susceptibility phenotype of each plant. Detection of the Rep protein will also be very useful, to which end work is in progress to produce Rep antibodies.

Fertile Hi-II maize transgenic for $pRep^{Rb-\Delta C2}$ has been regenerated. Future work will include challenging T₂ Hi-II seedlings with different MSV strains by agroinoculation, a useful way of determining the broadness of the resistance. The resistance of *D. sanguinalis* plants to the leafhopper-transmitted virus MSV-Pan, closely related to the severest MSV strains isolated thus far, indicates that any resistance achieved in maize should be durable and effective against even the most severe of MSV isolates.

Chapter 3

The Effect of Mutations in the MSV Rep RBR Protein-Interaction Motif on Viral Infectivity in Maize

Contents

ABSTRACT	108
3.1 INTRODUCTION	109
3.2 MATERIALS AND METHODS	111
3.2.1 Clone Construction	111
3.2.2 Analysis of Infectivity in Maize	113
3.2.3 Analysis of the Interaction of RepARb^{C(601)A} with RBR Protein using the Yeast Two-Hybrid Assay	116
3.3 RESULTS AND DISCUSSION	117
3.3.1 Infectivity of Wild Type and Mutated MSV-Kom in Maize	118
3.3.2 Analysis of the C(601)A Reversion	123
3.3.3 Investigations into the Possible Selective Advantages of the C(601)A Reversion	128
3.4 CONCLUSION	136

ABSTRACT

Previous reports indicated that geminivirus infectivity requires the interaction between the viral complementary (C) sense gene products (replication associated protein [Rep] or RepA) and the host retinoblastoma-related (RBR) protein. RBR protein controls cell-cycle progression by sequestering transcription factors required for entry into S phase. It was therefore suggested that one of the functions of the mastrevirus RepA protein is to provide a cellular environment competent for viral replication through interaction with RBR protein. In the previous chapter (Chapter 2), I showed that a mutation introduced into the RBR protein interaction motif (LxCxE to LxCxK) of the MSV C sense gene products abolished the interaction of RBR with RepA, but had no effect the ability of the C sense gene products to initiate replication of the MSV genome in black Mexican sweetcorn (BMS) suspension cells. In this study, the effect of the LxCxK mutation on MSV infectivity in maize was determined by agroinoculation of the mutant (Rb⁻) and wild type genomes into maize, followed by objective and precise symptom evaluation by quantification of infection rates, stunting, and chlorotic leaf areas. While the Rb⁻ mutant virus established a systemic infection at the same rate as the wild type virus, symptoms such as percentages of leaf areas covered by chlorotic lesions and stunting were attenuated by approximately 50%. The Rb⁻ mutant virus was transmissible to maize by leafhoppers, the natural vector of MSV, but symptom development took twice as long as in wild type infections, and symptoms such as leaf chlorosis and stunting were attenuated. Viral DNA was cloned from infected maize samples, and the Rep gene sequenced. In 100% of clones sequenced, a single nucleotide reversion of C(601)A had occurred in the three-nucleotide mutated sequence, while the two other mutated nucleotides were unchanged. A study was carried out to determine the time taken for the nucleotide reversion to occur, and the point at which the A(601) revertant virus population superseded the original C(601) mutant population. Further studies were aimed towards determining the selective advantage and the absolute requirement of the C(601)A reversion for viral infectivity in maize. These included the agroinoculation and symptom evaluation in maize of four different viruses containing either a G, T, A or C at position 601 in the Rep gene, and sequence analysis of the Rep genes cloned from infected maize three months after agroinoculation, to determine which nucleotides at position 601 were maintained and which reverted to A. The results indicate that while there is an extremely strong selective advantage of A at position 601, T(601) or G(601) is tolerated, while the major selective factor was a complete intolerance of C(601). The apparent important influence of the nucleotide at position 601 in the Rep gene is discussed in the context of the putative functions of RepA in the viral life cycle.

3.1 INTRODUCTION

While not absolutely required for replication, there is substantial evidence that RepA is a multifunctional protein with unique features that are required at different stages during the replicative cycles of mastreviruses (Gutierrez, 1999). Possible functions unique to RepA include downregulation of MSV replication (Collin *et al.*, 1996), and its interaction with cellular retinoblastoma-related (RBR) protein (Xie *et al.*, 1995; Horvath *et al.*, 1998) and a group of host GRAB (geminivirus RepA binding) proteins (Xie *et al.*, 1999), both of which may influence host pathways for the benefit of viral processes (Gutierrez, 2000). Functions ascribed to both Rep and RepA include sequence-specific DNA binding (Castellano *et al.*, 1998; Missich *et al.*, 2000); oligomerization (Horvath *et al.*, 1998); and transactivation of V-sense gene promoters (Hofer *et al.*, 1992; Zhan *et al.*, 1993; Collin *et al.*, 1996; McGivern, 2002).

Together, Rep and RepA have many features in common with viral oncoproteins, such as SV40 T-antigen (Ludlow, 1993), adenovirus E1A (Moran, 1994) and human papillomavirus type 16 E7 (Vousden, 1993). Like these oncoproteins, Rep and RepA have an essential role in both viral replication and transcription, and for these processes rely on host proteins such as DNA polymerase α . The most striking similarity between the animal virus oncoproteins and the Rep/RepA of most mastreviruses is the presence of a motif, LxCxE, via which the oncoproteins interact with the mammalian retinoblastoma (Rb) protein, and the mastrevirus Rep/RepA interact with the plant RBR protein (Xie *et al.*, 1995). In addition, mammalian (Dowdi *et al.*, 1993; Ewen *et al.*, 1993) and plant (Dahl *et al.*, 1995; Soni *et al.*, 1995) D-type cyclins contain the LxCxE motif, through which they also interact with Rb and RBR respectively. The same motif also mediates binding of a nanovirus gene component (Clink, for "cell cycle link") to plant RBR (Aronson *et al.*, 2000). Although mastrevirus Rep shares the LxCxE motif with RepA, only RepA interacts with RBR (Horvath *et al.*, 1998; Liu *et al.*, 1999).

The mammalian Rb protein controls the cell cycle transit of G1 to S phase by inactivating the E2F family of transcription factors that transactivate the promoters of many S-phase genes. In this way Rb prevents inappropriate cell proliferation. Inactivation of Rb by phosphorylation results in the release of the E2F transcription factors and the onset of cell division. By producing proteins that bind to Rb, thereby releasing E2F transcription factors, animal oncoviruses create a molecular environment in the nucleus suitable for viral replication and transcription, bypassing the need for Rb phosphorylation.

Geminivirus replication, which has been detected in terminally differentiated cells (Lucy *et al.*, 1996), is dependent on host cell proteins. These cellular replication proteins are generally absent or not functional in differentiated cells. Paradoxically, geminivirus replication is excluded from the meristem, which contains actively proliferating cells that naturally express the replication factors required for geminivirus infection. Since the major components of the mammalian Rb pathway are functionally equivalent in plants (see de Jager and Murray, 1999, for a review on the plant RBR pathway), the discovery of the LxCxE RBR-interaction motif in WDV RepA provided a possible mechanism by which geminiviruses, like mammalian oncoviruses, may induce a cellular state that is permissive for viral DNA replication. This would be by inducing terminally differentiated cells to enter the S phase, or by activating S phase-specific genes.

There are several lines of evidence geminiviruses alter the cell-cycle controls of their hosts. First, Accotto *et al.* (1993) found that dsDNA replication intermediates are significantly more abundant in S-phase nuclei than in nuclei from other phases of the cell cycle. Second, proliferating cell nuclear antigen (PCNA), which is virtually undetectable in differentiated cells, accumulates to high levels in fully differentiated cells expressing *Tomato golden mosaic virus* (TGMV) Rep (Nagar *et al.*, 1995). PCNA, whose expression is regulated by E2F, has an essential role in DNA replication, functioning as a homotrimeric clamp at the origin, which facilitates the recruitment and processivity of DNA polymerase δ (Sever-Chroneos *et al.*, 2001). Third, the importance of TGMV Rep-RBR interaction was established by Kong *et al.* (2000), who found that mutants with impaired binding to RBR accumulate less viral DNA and cause chlorosis that is confined to the veins, and that the mutations result in altered tissue specificity. Kong *et al.* (2000) also showed that PCNA expression in infected cells is tightly linked to the capacity of TGMV Rep to interact with RBR. Fourth, Gordon-Kamm *et al.* (2002) found that expression of WDV RepA stimulated cell division in tobacco cell cultures and overcame cell arrest in the G₀/G₁ phase of the cell cycle. These findings seem to provide compelling evidence for a geminiviral influence on host gene expression affecting the cell cycle.

Mutational analysis of the mastrevirus RepA LxCxE motif has shown the importance of the three conserved residues in mediating binding to RBR. In one study, a mutation of E to K abolished the ability of WDV RepA to bind to Rb (Xie *et al.*, 1995), and in a detailed mutational analysis of the motif in *Bean yellow dwarf virus* (BeYDV) RepA, Liu *et al.* (1999a) confirmed the importance of all three conserved residues. In a result similar to that of Xie *et al.* (1995), a mutation of the E

residue (in this case E to Q) in BeYDV Rep had a particularly drastic effect, reducing binding efficiency by 95%.

Although Xie *et al.* (1995) found a correlation between the ability of WDV RepA to bind to RBR and the replication efficiency of the virus in wheat suspension cells, no such relationship has been detected in other mastreviruses. Liu *et al.* (1999a) found that all BeYDV mutants, in which RBR binding was impaired by varying degrees up to 95%, were able to replicate in tobacco protoplasts and to systemically infect *N. benthamiana* and bean. In addition, by comparison of the distribution of replicative forms of MSV DNA with the expression of the S-phase-specific host gene, H2b, Lucy *et al.* (1996) showed that host DNA synthesis is not a prerequisite for MSV replication; this lack of correlation was most notable in the developmentally mature leaf laminal tissues where MSV replication could be detected. These data suggest that, although different geminiviruses may interfere with control of cell proliferation, more than one mechanism or strategy may have evolved, and that one or more factors associated with dividing cells, rather than cell division itself, may be required for MSV replication.

In the study presented in Chapter 2, an MSV Rep construct with a mutation (LxCxK) which abolished the protein's interaction with the host RBR protein, supported the *trans*-replication of the viral genome in black Mexican sweetcorn (BMS) suspension cells to levels 30-fold higher than the replication of MSV alone. This suggested that in MSV, like BeYDV, RepA-RBR interaction was not required for viral replication. In this chapter, the effect of the LxCxK mutation, as well as further mutations that were introduced into the RBR-binding motif, on viral infectivity in maize is investigated.

3.2 MATERIALS AND METHODS

3.2.1 Clone Construction

3.2.1.1 PCR Mutagenesis

In Chapter 2 an MSV-Kom *Rep* construct (pSKRep^{III+Rb-NTP+}) containing a mutation in the LxCxE motif of the RBR-interaction domain was described. In this study, pSKRep^{III+Rb-NTP+} (for simplicity referred to in this chapter as pSKRepRb⁻) was used as a template for further mutagenesis of the RBR-interaction domain (Table 3.1). The single nucleotide (nt) changes, introduced by PCR using mutagenic primers, were from C(601) (in pSKRepRb⁻) to G(601) and

T(601). The mutagenic primer sequences were: RepDeg(TG)Forward 5'-TGATTACTTTGT(TG)TTAAGTCAATCAACGATTG-3' and RepDeg(CA)Reverse 5'-TTGACTTAA(CA)ACAAAGTAAATCAGGTGAGG-3'. These overlapping degenerate primers were designed to introduce either a T or a G (or an A or C in the complementary strand) at position 601, shown in brackets in the primer sequences. The design was such that the primers amplified the whole pSKRepRb⁻ template. The PCR product was then digested with *BfrI*, whose restriction site (CTTAAG) in pSKRepRb⁻ was destroyed by the mutation of C(601) to G or T (see Table 3.1). Thus, digest of the PCR product with *BfrI* linearised the template DNA, ensuring that only re-ligated PCR-amplified plasmids could be transformed into the *Escherichia coli* strain DH5 α . Clones containing either the G(601) mutation or the T(601) mutation were identified by sequence analysis, and designated pSKRepRb^{-C(601)G} and pSKRepRb^{-C(601)T} respectively. Sequencing and sequence analysis was carried out as described in Chapter 2. Standard cloning techniques were used as in Sambrook *et al.* (1989).

TABLE 3.1 Mutations introduced into the LLCLK Motif of pSKRep^{Rb-}

¹ Rep Construct	² Amino Acid Sequence	³ Nucleotide Sequence
pSKRep (wild type)	¹⁹⁸ LLCNE	⁵⁹² TTACTTTGTAATGAG
pSKRepRb ⁻ (template)	¹⁹⁸ LLCLK	⁵⁹² TTACTTTGTCTTAAG
pSKRepRb ^{-C(601)G}	¹⁹⁸ LLCVK	⁵⁹² TTACTTTGTgTTAAG
pSKRepRb ^{-C(601)T}	¹⁹⁸ LLCFK	⁵⁹² TTACTTTGTtTTAAG

¹The mutations introduced into the template pSKRepRb⁻ are shown in relation to the wild type (wt) Rep amino acid and nucleotide sequence.

²Amino acid numbering is relative to the Rep start codon. Amino acid changes relative to the wt Rep sequence are shown in bold type, while amino acid changes relative to the pSKRepRb⁻ sequence are shown in bold and underlined in pSKRepRb^{-C(601)G} and pSKRepRb^{-C(601)T}.

³Nucleotide numbering is relative to the Rep ATG (A = 1). Nucleotide changes, relative to the wt Rep, in pSKRepRb⁻ are shown in bold capital letters. Nucleotide changes relative to pSKRepRb⁻ are shown in bold lower case letters.

3.2.1.2 Production of agroinfectious MSV constructs

Although natural transmission of MSV is mediated exclusively by leafhoppers, Grimsley *et al.* (1987) developed an alternative means of transmitting MSV using *Agrobacterium tumefaciens*-mediated transfer of tandemly cloned complete or partial dimers of the MSV genome in binary cloning vectors (referred to as agroinfectious constructs). Using this technique, called agroinfection or agroinoculation, maize plants can be efficiently and reproducibly infected with precise concentrations of single MSV isolates, enabling an accurate evaluation and comparison of the pathogenicity of different MSV isolates (Martin *et al.*, 1999).

All agroinfectious MSV constructs were based on pKom602, which contains one full MSV-Kom genome bounded by two full long intergenic regions (1.1-mer; Schnippenkoetter *et al.*, 2001), permitting efficient replicative release upon agroinfection.

The LLCN²⁰¹E²⁰² to LLCL²⁰¹K²⁰² (Rb⁻) mutation was introduced into pKom602 by exchanging the wt *NsiI/XhoI* fragment of pKom602 with the mutated *NsiI/XhoI* fragment of pSKRepRb⁻, to create pKomRb⁻. Similarly, the plasmids pKomRb^{-C(601)G} and pKom^{Rb-C(601)T} were created by introducing the *NsiI/XhoI* fragment from pSKRepRb^{-C(601)G} and pSKRepRb^{-C(601)T} respectively into pKomRb⁻. Initial identification of correct clones was carried out by digesting the plasmids with *BfrI*, whose restriction site was introduced with the Rb⁻ mutation into pKomRb⁻. Conversely, introduction of the C(601)G and C(601)T mutation into pKomRb⁻ was identified by the destruction of the *BfrI* site. These clones were then confirmed to be correct by sequencing.

A virus genome isolated from maize that had been agroinfected with pKomRb⁻, containing a naturally-occurring single nucleotide reversion of C(601)A, was converted into an agroinfectious 1.1-mer in the same way, by replacing the *NsiI/XhoI* fragment of pKomRb⁻ with the same fragment from the revertant virus, resulting in the agroinfectious construct pKomRb^{-C(601)A}.

The 1.1-mer genomes of pKom602 and each of the mutant viruses described were cloned into the *EcoRI* and *XbaI* sites of the binary vector pBI121 (CLONTECH, CA) to obtain pBIKom602, pBIKomRb⁻, pBIKomRb^{-C(601)G}, pBIKomRb^{-C(601)T} and pBIKomRb^{-C(601)A}. *A. tumefaciens* C58C1 (Konec and Schell, 1986) was transformed with these constructs using the method of An *et al.* (1988).

3.2.2 Analysis of Infectivity in Maize

3.2.2.1 Agroinoculation

pBIKom602, pBIKomRb⁻, pBIKomRb^{-C(601)A}, pBIKomRb^{-C(601)T} and pBIKomRb^{-C(601)G} were introduced into maize plants by agroinoculation, essentially according to Grimsley *et al.* (1987), with the modifications of Martin *et al.* (1999). For each agroinfection experiment, transformed *A. tumefaciens* was injected into the apical meristem of a group of 14 three-day old maize seedlings (cv. Jubilee), and a group of 14 control plants was injected with sterile distilled water. The agroinfection of each construct was repeated in three separate experiments. Inoculated seedlings were grown in a plant growth room maintained at between 21 and 22°C, at 80% relative

humidity, with 16 hours of light per day. The majority of agroinoculations were done by D. Martin, UCT.

3.2.2.2 Analysis of symptoms

Disease severity was measured in terms of percentages of leaf areas covered by chlorotic lesions in infected plants as described in Martin *et al.*, (1998). Percentage leaf areas covered by chlorotic lesions in each infected plant were estimated for leaves two through six using a microcomputer-based image analysis technique (Martin *et al.*, 1998). The percentage chlorotic areas of leaves two and three were assessed 15 days after agroinoculation; the percentage chlorotic areas of leaves four, five and six were assessed 22, 29 and 35 days after agroinoculation respectively. For each MSV construct, the mean percentage chlorotic area of the 2nd to 6th leaf was used as a representative measure of chlorosis. Analysis of symptoms was carried out by D. Martin, UCT.

3.2.2.3 Leafhopper transmission

Symptomatic maize plants that had been agroinoculated with pKom602 and pKomRb⁻ were placed in separate gauze-covered cages containing non-viruliferous leafhoppers (*C. mbila*), which were allowed to acquire the virus over a three-day period. A latency period of seven days was allowed after which the viruliferous leafhoppers were fed on maize plantlets (cv. Jubilee) at the three-leaf stage for a period of one to two weeks. Symptoms were assessed visually.

3.2.2.4 Viral DNA analysis

Total genomic DNA was extracted from infected maize plants, and the Rep gene was PCR-amplified from each sample, cloned into the pGEM®-T Easy vector (Promega, WI), and sequenced. This was to ensure that the mutations in the Rep gene had been retained, or to check for any reversions or mutations that may have arisen in the viral population during the maize infection. To ensure PCR fidelity, each sample was subjected to PCR at least twice, and with two different sets of primers. The primers used were the C1 and C2 primers (described in Chapter 2), which amplify the entire Rep gene, and TYR-F and NTP-R primers, which amplify the region between motif III in the C1 and the NTP binding domain in the C2. Primer sequences were: TYR-F 5'-GCCAAGTCAGTTAACAGAGTGAGGG-3' and NTP-R 5'AACCCCTAGGCTTCTGGCCC-3'. All PCR-amplified Rep fragments were sequenced in the forward and reverse directions to ensure that any perceived mutations were not due to sequencing errors. Nucleotide changes were only taken as being genuine mutations of the viral genome if the

same changes occurred in Rep fragments amplified from more than one PCR reaction and if they occurred in both directions sequenced.

3.2.2.5 Analysis of the C(601)A reversion in pKomRb⁻

Fourteen maize seedlings were agroinoculated with pBIKomRb⁻ as described in 3.2.2.1. As symptoms emerged and thereafter every 10 days, leaf samples were taken from each plant. Genomic DNA extracted from these samples was quantified on a 0.8% gel as described in Chapter 2 (2.2.3), and diluted to 50 ng/μl. Primers were designed with the last 3' nt of the primer sequence being an A or a C to amplify the Rep gene containing specifically the A(601) or C(601) nucleotide respectively. Primer sequences were: Rb⁻C (F) 5' CACCCTCCTCACCTGATTACTTTGTC 3' and Rb⁻revA (F) 5' CCACCCTCCTCACCTGATTACTTTGTA 3'. The reverse primer was the C2 primer (described in Chapter 2, 2.2.1), which amplifies the Rep gene from the SIR through the C2.

Quantification of viral DNA and mutational analysis was performed on a LightCycler instrument (Roche, Mannheim, Germany). Real time PCR was performed using primers designed to amplify a 300-bp fragment spanning the mutated area of the Rep gene (nt 446 to 746 relative to the Rep ATG). Primer sequences were: MSVmut-F 5'- CAAAGAAGAGTACCTCTCCAT -3', and MSVmut-R 5'- TTCATCTGATGACTGGATACAG -3'. Fluorescent probes were designed to hybridize to the target sequence in a head-to-tail arrangement on the same strand of amplified products. The donor probes were labelled at the 5' end with fluorescence, while the acceptor probe was labelled at the 5' end with LCRed640 and modified at the 3' end by phosphorylation to block extension. The gap between the 3' end of the donor probe and the 5' end of the acceptor probe was one base. One of the donor probes (F-Sensor [C]) was designed to be 100% complementary to the sequence of the C(601) mutation, and another donor probe (F-sensor [A]) to be 100% complementary to the sequence of the A(601) reversion. The sequence of the F-sensor [C] probe was 5'- TCACCTGATTACTTTGTCTTAAGTCAATC -3' and the F-sensor [A] probe was 5' - TCACCTGATTACTTTGTATTAAGTCAATC -3', both covering nt 583 to 612. The sequence of the acceptor probe (LC-Anchor) was 5'- LCRed640-ACGATTGGCTCCAGCCTAACATCTTC-p and covered nt 614 to 637. All primers and probes were synthesized and purified by reverse-phase HPLC by TIB MOLBIOL (Berlin, Germany).

PCR amplification using the LightCycler was carried out in 20 μl of reaction mixture consisting of a master mixture containing FastStart Taq DNA polymerase, dNTP mixture and reaction

buffer (LightCycler - FastStart DNA Master Hybridization Probes; Roche), 5 mM MgCl₂, 1 μM of primers, 0.4 μM of each probe and 2 μl (100ng) of template genomic DNA in a LightCycler capillary. For amplification of the 300-bp *Rep* fragment, a pre-incubation step at 95°C for 10 minutes was carried out to activate the FastStart Taq DNA polymerase, followed by 40 rounds of amplification at 95°C (0sec) for denaturation, 58°C (10 sec) for annealing and 72°C (10 sec) for extension, with a temperature slope of 20°C/sec performed in the LightCycler. Real-time PCR monitoring was achieved by measuring the fluorescent signal at the end of the annealing phase for each cycle. External standards were prepared by 10-fold serial dilutions of *Rep* plasmid DNA (corresponding to a range of 0.1 to 1000 pg/ μl).

Each run consisted of five external standards, a negative control (plant genomic DNA uninfected by MSV), a water control (water in place of template DNA) and plant samples with unknown concentrations of MSV. Quantitation of MSV DNA in each sample was performed automatically by reference to the standard curve constructed each time according to the LightCycler software. In order to distinguish between MSV-KomRb⁻ and MSV-KomRb^{-C(601)A} in each sample, a melting curve analysis, which genotypes single nucleotide polymorphisms, was performed after the completion of the amplification cycles. The peak areas of the melting curves were related to the relative amount of each virus' DNA present in the plant samples by using standard plasmid samples containing differing ratios of pKomRb⁻ (C):pKomRb^{-C(601)A} (A) (100A:0C, 80A:20C; 60A:40C; 50A:50C; 40A:60C; 20A:80C; 0A:100C). The peak areas of these standards were plotted as a standard curve, and used to determine the ratios of the two viruses present in the plant samples at different time points.

3.2.3 Analysis of the Interaction of RepArb^{-C(601)A} with RBR protein using the Yeast-Two Hybrid Assay

Since it is now well established that RBR interacts with RepA but not Rep (Horvath *et al.*, 1998; Liu *et al.*, 1999; C. Gutierrez, pers. comm.; results from Chapter 2 of this thesis), the ability of RepA (rather than Rep) containing the C(601)A reversion to bind to RBR was investigated. In order to clone the revertant Rep gene, the sequence encompassing the C(601)A reversion (obtained from a previously cloned and sequenced PCR fragment amplified from an infected maize sample) was introduced into pSKRepRb⁻ by exchanging the *NsiI/XhoI* fragment of pSKRepRb⁻ with the *NsiI/XhoI* fragment from the revertant Rep gene, to create pSKRepRb^{-C(601)A}. The RepA version of pSKRepRb^{-C(601)A} was constructed first by introducing a 3' splice site

mutation into pSKRepRb^{-C(601)A} to prevent the possibility of intron splicing (since RepA contains the entire intron sequence), and then amplifying the RepA ORF from the full-length gene by PCR, and cloning the PCR product into pSK, to create pSKRepARb^{-C(601)A}. The details of the creation of RepA constructs can be found in Chapter 2 (2.2.1.4). The strategy for the cloning of RepA genes into the yeast-two hybrid vectors is described in detail in Chapter 2. Briefly, the RepARb^{-C(601)A} gene was cloned in frame with the GAL4 activation domain into pGAD424 (CLONTECH, CA) to create a RepARb^{-C(601)A}-GAL4 fusion product, called pADRepARb^{-C(601)A}. The cDNA of the maize RBR protein (pZmRb1; Xie *et al.*, 1996) cloned into pGBT9 (pGBT9ZmRb1) was provided by Dr. G. Horvath (described in Horvath *et al.*, 1998). pADRepA (described in Chapter 2) was used as a positive control for interaction with RBR. The yeast transformation and analysis of the RBR-interaction properties of the RepA constructs was performed by D. McGivern (John Innes Centre, Norwich, UK), as described in Chapter 2.

3.3 RESULTS AND DISCUSSION

In a study of the biological role of the WDV Rep LxCxE motif, Xie *et al.* (1995) reported that the replication of WDV mutants in wheat suspension cells was eliminated when the motif was altered to LxCxK. Conversely, in the replication assay presented in Chapter 2, an MSV Rep LxCxK mutant (RepRb⁻) was able to *trans*-replicate MSV-Kom in black Mexican sweetcorn (BMS) suspension cells to levels 30-fold higher than when MSV-Kom was bombarded alone. This was comparable to the *trans*-replication of MSV-Kom by wt Rep. Both Rep and RepRb⁻ were highly expressed from the maize ubiquitin promoter, indicating that the enhanced levels of viral replication in their presence was due to the greater number of Rep molecules available to initiate RCR of the MSV-Kom genome, compared with the far lower expression of Rep from its own promoter when MSV-Kom was bombarded alone. Although in the co-bombardments of pKom602+pRepRb⁻, wt Rep, (or more specifically RepA) expressed by pKom602 was available to bind RBR, in assays involving the *trans*-replication of a mutant MSV unable to express its own Rep, RepRb⁻ *trans*-replicated the mutant genome to the same levels as those achieved by wt Rep (results presented in Chapter 4). These combined data indicate that the Rb⁻ mutation in RepRb⁻ has no effect on its ability to initiate RCR of the MSV genome. Since this mutation was demonstrated in yeast two-hybrid assays to abolish the interaction of RepARb⁻ with RBR (Chapter 2), it appears that unlike what happens in WDV, RBR-RepA interaction is not necessary for MSV replication. However, since the replication assays in both Xie *et al.* (1995) and this study employed the use of actively dividing cell cultures that were already competent for DNA replication, the proposed function of the RepA-RBR interaction (that is to drive cells into S

phase) is likely to be redundant. A more relevant assay to determine the effect of the Rb⁻ mutation was designed, in which the mutation was introduced into an agroinfectious wt virus, pKom602, to produce pKomRb⁻, and the infectivity of wt and mutant viruses in maize was compared.

3.3.1 Infectivity of Wild Type and Mutated MSV-Kom in Maize

The infectivity of pKom602 and pKomRb⁻ was established by agroinoculation of the MSV-sensitive maize cv. Jubilee, and symptom severity was measured in terms of percentages of leaf areas covered by chlorotic lesions in infected plants, determined using a microcomputer-based image analysis technique (Fig. 3.1). As can be seen in Figs. 3.1 and 3.2, MSV-KomRb⁻ was able to systemically infect maize, although symptom severity in leaves 2 to 6 was attenuated by approximately 50% when compared with the infection of MSV-Kom. In addition, maize plants infected with MSV-KomRb⁻ were less stunted than those infected with the wt virus (see Fig. 3.2). The timing of symptom development and the percentages of agroinoculated plants that became infected, however, were the same in both mutant and wt infections.

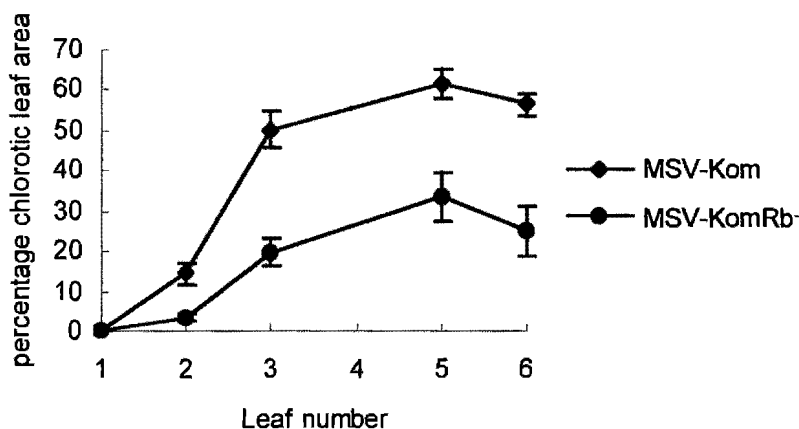


FIGURE 3.1 Percentage chlorotic areas recorded on leaves 2 through 6 of maize cv. Jubilee agroinoculated with agroinfectious MSV-Kom (blue symbols and lines) and MSV-KomRb⁻ (red symbols and lines) constructs. 14 maize plants were agroinoculated with each virus in three separate experiments. Error bars represent 95% confidence intervals of the mean.

In previous studies several observations suggested that geminivirus infectivity in differentiated cells required the interaction between Rep or RepA and RBR proteins. It has been established that

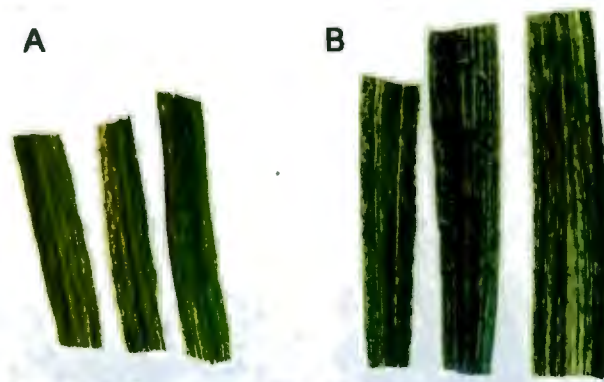


FIGURE 3.2 symptoms occurring on the 3rd, 4th and 5th leaves of a representative maize plant (cv. Jubilee) agroinoculated with (A) MSV-Kom and (B) MSV-KomRb⁻. Leaves were cut transversely into four segments of equal length, and photographs were taken of the second segment from the base of each leaf blade. It can be seen that leaves from (A) are more stunted and have a greater chlorotic area than leaves from (B).

TGMV Rep-RBR protein interaction and associated PCNA accumulation is important for TGMV infectivity in *N. benthamiana* (Kong *et al.*, 2000). In mastreviruses, suggestions that RepA is required in the viral lifecycle for its ability to provide a cellular environment competent for viral replication through interaction with RBR protein, came from the fact that MSV and BeYDV mutants unable to express RepA are unable to infect plants (Liu *et al.*, 1998; Boulton, 2002); that RepA, but not Rep, binds RBR (Horvath *et al.*, 1998; Liu *et al.*, 1999), and that RepA can stimulate cell division in tobacco cell cultures (Gordon-Kamm *et al.*, 2002). However, Figs. 3.1 and 3.2 show that an MSV mutant that cannot interact with RBR protein apparently retains its ability to establish a systemic infection in maize. Similarly, Liu *et al.* (1999) found that BeYDV tolerates similar mutations in the RBR-interaction motif of its Rep. However, unlike the infectivity of MSV-KomRb⁻ in maize, the BeYDV mutants were able to produce a wt infection in *N. benthamiana* and bean. This may reflect the sensitivity and accuracy of the symptom analysis technique used in this study, rather than relying on subjective analysis of symptoms by eye. Alternatively, the difference in behaviour between BeYDV and MSV mutants may be attributable to adaptations of these viruses to dicotyledonous and monocotyledonous hosts respectively.

It is likely that the attenuated symptoms caused by MSV-KomRb⁻ were as a result of the LxCxK mutation, since the three altered nucleotides comprising the mutation were the only differences between the mutant and wt virus. The pattern of chlorotic streak-like lesions on maize leaves can be directly correlated with the pattern of virus accumulation within the leaves (Lucy *et al.*, 1996).

Since a general characteristic of symptom development in MSV-KomRb⁻ infected plants was the development of slightly narrower streaks, it is possible that the mutation altered tissue specificity of the virus, for example restricting its replication to vascular tissue, where the virus could use pre-existing host replication factors present in some phloem-associated cells. However, there is also a direct correlation between streak width and virus titre (M. Boulton, pers comm.). Although the amount of viral MSV-KomRb⁻ DNA present in maize samples or within streak lesions was not accurately quantitated or compared with that of MSV-Kom DNA, the narrower streaks and the lower degree of chlorosis within lesions caused by the mutant virus may have been a result of lower virus titres due to lowered replication efficiency in all tissues.

With *Agrobacterium*-mediated delivery of viral strains, the virus is injected into the maize seedling's apical meristem where cellular replication factors may enable the replication of MSV-KomRb⁻ to high enough titres for the virus to then invade the photosynthetic cells of the leaf lamina, making RBR-RepA interaction dispensable. In this case the presence of the mutant virus along the entire length of the leaf (as evidenced by the extension of chlorotic streaks along the entire leaf lamina) could have been a consequence of viral movement without replication in the developed leaf. The ability of the mutant virus to be transmitted to maize by leafhoppers, which deposit virus particles into mature leaf cells, was determined by feeding non-viruliferous leafhoppers on leaf material that had been agroinfected with MSV-KomRb⁻, and following acquisition of the virus and a seven-day latency period, feeding the newly-viruliferous leafhoppers on maize plantlets at the three-leaf stage. The same procedure was followed for MSV-Kom to be used as a positive control and to compare infectivities. This was done in three separate experiments (Table 3.2).

Although MSV-KomRb⁻ was transmissible to maize by leafhoppers, in experiments 1 and 2 symptom development took twice as long compared with the wt infection. Symptoms were also attenuated in terms of plant height, leaf length and degree of chlorosis of the leaf. However, since leafhopper transmission is not as accurate or reproducible as agroinoculation; for example it is impossible to tell how much MSV inoculum leafhoppers delivered to each plant, these observations are purely subjective. The most interesting observation from the leafhopper transmissions was the fact that in experiments 1 and 2, samples taken from maize 19–24 days after agroinoculation with MSV-KomRb⁻ contained a mutant viral population, as evidenced by the PCR-amplified Rep products being cut by *BfrI*, while the Rep gene amplified from viral populations in leaf samples taken from the leafhopper-inoculated maize 30-39 days after

leafhopper transmission no longer contained the *BfrI* site. Interestingly, the Rep gene amplified from a sample of leaf 5 taken from one plant 24 days after agroinoculation with MSV-KomRb⁻ contained the *BfrI* site, while in a sample of the same leaf taken 10 weeks later, the Rep gene was no longer cut by *BfrI*.

TABLE 3.2 Transmission of wild type and mutant MSV-Kom to maize by leafhoppers

¹ EXPERIMENT 1					
Agroinfected virus	Viral DNA extracted (days after agroinoculation)	Cut with <i>BfrI</i>	Time to develop symptoms after leafhopper transmission	DNA extracted (days after leafhopper transmission)	Cut with <i>BfrI</i>
MSV-Kom	19 days (leaf 4)	✗	10 days	39 days (leaf 9)	✗
MSV-KomRb ⁻	19 days (leaf 4)	✓	20 days	39 days (leaf 9)	✗
² EXPERIMENT 2					
MSV-Kom	24 days (leaf 5)	✗	5 days	30 days (leaf 8)	✗
MSV-KomRb ⁻	24 days (leaf 5)	✓	10 days	30 days (leaf 8)	✗
³ EXPERIMENT 3					
Leafhopper-transmitted virus	DNA extracted (days after leafhopper transmission)	Cut with <i>BfrI</i>	Time to develop symptoms after 2 nd leafhopper transmission	DNA extracted (days after 2 nd leafhopper transmission)	Cut with <i>BfrI</i>
MSV-Kom	30 days (leaf 8)	✗	7 days	23 days (leaf 7)	✗
MSV-KomRb ⁻	30 days (leaf 8)	✗	9 days	23 days (leaf 7)	✗

¹A leaf sample was taken from a symptomatic maize plant 19 days after agroinoculation, and leafhoppers were fed on the same plants to acquire the virus. These leafhoppers were then fed on maize plantlets, from which leaf samples were taken 39 days after leafhopper transmission of the virus. Total genomic DNA was extracted from all the leaf samples, from which the Rep gene was amplified by PCR, followed by digestion with *BfrI*. Rep fragments containing the Rb⁻ mutation should be cut by *BfrI*, while wt Rep fragments should remain uncut.

²Having taken a sample from leaf 5 of a symptomatic maize plant 24 days after agroinoculation with MSV-KomRb⁻, leafhoppers were fed on the same leaf to acquire the virus, and then fed on maize plantlets at the three-leaf stage. The same procedure was followed for an MSV-Kom-infected plant. Samples were taken from leaf 8 of leafhopper-inoculated maize plants 30 days after transmission of the virus. DNA was extracted from all the leaf samples, the Rep gene PCR-amplified, and digested with *BfrI* to determine the presence or absence of the Rb⁻ mutation.

³Leafhoppers were fed on leafhopper-inoculated maize from experiment 2 to acquire the virus, and in a second round of leafhopper transmissions, were fed on maize plantlets at the three-leaf stage. 23 days after transmission of the virus, DNA samples were taken from leaf 7 of each plant, the Rep gene PCR-amplified, and digested with *BfrI*.

The Rep genes from the samples that had lost the *BfrI* site, as well as from samples taken from leaves 5 and 12 of maize 13 weeks after agroinoculation with MSV-KomRb⁻, were cloned and sequenced. All Rep genes sequenced from MSV-KomRb⁻ infected samples contained a single nucleotide reversion of C(601) to A(601), while the wt Rep genes from MSV-Kom-infected

plants were unchanged (Table 3.3). Although it is unlikely that MSV can autonomously replicate in *A. tumefaciens*, pKomRb⁻ was extracted from the *A. tumefaciens* culture that was used for the agroinoculations, the RepRb⁻ gene was PCR-amplified from the plasmid, and cut with *Bfr*I. The fact that the RepRb⁻ amplification products were completely digested by *Bfr*I (data not shown) confirmed that the reversion did not occur while being maintained in the bacterium, but rather occurred after inoculation into the maize seedlings.

TABLE 3.3 Single nucleotide reversion occurring in the Rep gene of agroinfected and leafhopper-transmitted MSV-KomRb⁻ populations.

Rep gene	¹ Nucleotide Sequence	² Amino Acid Sequence
<i>Rep</i> (wild type)	⁵⁹² TTACTTTGT A ATGAG	¹⁹⁸ LLCNE
<i>RepRb</i> ⁻ (mutant)	⁵⁹² TTACTTTGT C TAA G	¹⁹⁸ LLCLK
<i>RepRb</i> ^{-C(601)A} (revertant)	⁵⁹² TTACTTTGT A TAA G	¹⁹⁸ LLCIK

¹The single nucleotide reversion (highlighted in red) in *RepRb*^{-C(601)A} is shown in comparison with the original mutated sequence in *RepRb*⁻ and the wt sequence in *Rep*. Note that the A(602)T and G(604)A mutations (shown in bold) are unaltered in *RepRb*^{-C(601)A}. Note that the *Bfr*I site (underlined) in *RepRb*⁻ is lost in *RepRb*^{-C(601)A}.

²The RBR-binding motif amino acid sequence in wt Rep, *RepRb*⁻ and *RepRb*^{-C(601)A} is shown, with the amino acid change resulting from the C(601)A nt reversion highlighted in red.

While the C(601)A nt reversion explains the loss of the *Bfr*I site in the leafhopper-transmitted virus samples, it may also provide a clue why the time taken for symptoms to develop in the second round of leafhopper transmissions was closer to the wt infection (Table 3.2). In the first leafhopper transmission, leafhoppers presumably carried a viral population consisting mainly of MSV-KomRb⁻, since samples taken from the same agroinfected leaf from which the leafhoppers acquired the virus contained PCR-amplified Rep genes that were cut by *Bfr*I. Although there was a small amount of uncut DNA in these samples, it is difficult to conclude whether this was due to incomplete digestion or whether some of the population contained the C(601)A reversion. In the second round of leafhopper transmissions, the viral population transmitted by the leafhoppers consisted mainly (or possibly entirely) of the C(601)A revertant virus. It is possible, although entirely speculative, that in the first transmission only viruses containing the C(601)A reversion could be transmitted by the leafhoppers. Since the majority of the population in the first transmission consisted of MSV-KomRb⁻, this would explain why symptom development took twice as long as the wt infection, whereas when the C(601)A revertant population was transmitted in the second round of leafhopper transmissions, symptom development was closer to the wt infection.

3.3.2 Analysis of the C(601)A Reversion

In order to study the nt reversion more thoroughly, 14 maize seedlings were agroinoculated with MSV-KomRb⁻. As symptoms emerged on each leaf, and thereafter every 10 days for four months, samples were taken and genomic DNA extracted. Primers were designed to specifically amplify Rep genes containing either the A(601) or the C(601) nucleotide. Samples amplified by the C(601)-specific primer but not by the A(601)-specific primer clearly consisted of only the original mutant population, while the A(601)-specific primer was used to determine when the nt reversion first appeared in each sample (Table 3.4). Samples were only analysed up to 35 days after agroinoculation, since the objective was simply to determine when the nt reversion first appeared.

Of the nine symptomatic agroinfected plants, all but one contained the A(601) revertant virus by day 35 after agroinoculation. In some plants, such as no. 3, 4, 7 and 12, the revertant virus was detected as early as 10 days after agroinoculation, at the same time as the detection of the agroinoculated C(601) mutant virus.

Samples from plants 1, 4, and 5 were chosen for a more in-depth analysis of the kinetics of the C(601)A reversion. These plants were chosen for their differences in the time taken for the revertant virus to appear in the mutant population. The composition of the viral population in each sample was determined by real-time, sequence-specific analysis and quantification of PCR-

TABLE 3.4 Detection of C(601) mutant and A(601) revertant viral populations in plants agroinoculated with pKomRb⁻

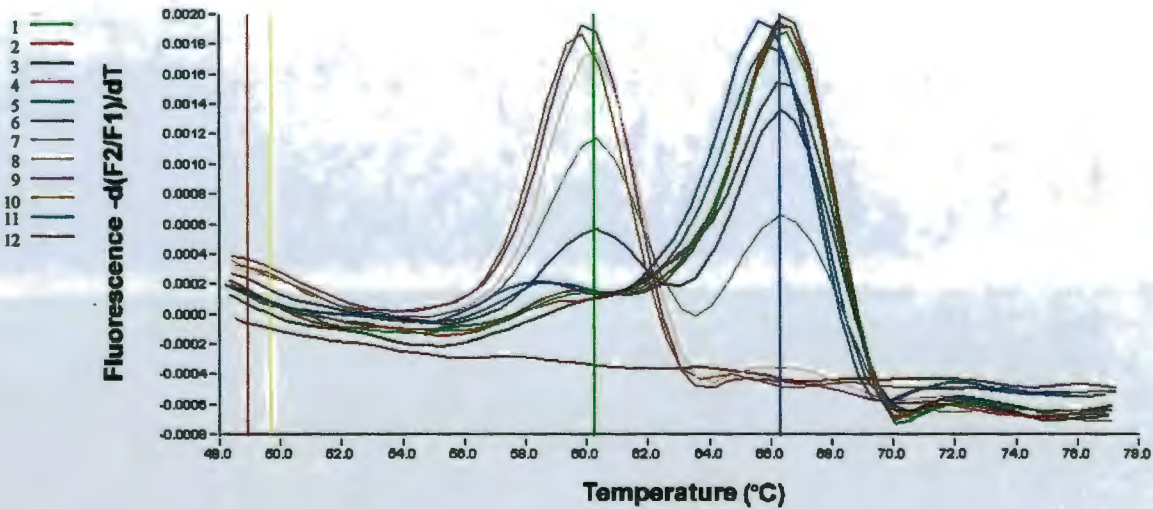
Plant	First appearance of symptoms (days after agroinoculation)	First detection of C(601) virus (days after agroinoculation)	First detection of A(601) virus (days after agroinoculation)
1	32 days, leaf 5	32 days, leaf 4	-
2	-	-	-
3	10 days, leaf 3	10 days, leaf 3	10 days, leaf 4
4	10 days, leaf 3	10 days, leaf 2	10 days, leaf 2
5	20 days, leaf 4	10 days, leaf 2	30 days, leaf 6
6	-	-	-
7	20 days, leaf 5	10 days, leaf 3	10 days, leaf 4
8	-	-	-
9	10 days, leaf 3	10 days, leaf 2	12 days, leaf 4
10	-	-	-
11	12 days, leaf 3	10 days, leaf 2	32 days, leaf 4
12	12 days, leaf 2	10 days, leaf 2	10 days, leaf 3
13	-	-	-
14	10 days, leaf 4	10 days, leaf 4	33 days, leaf 4

amplified viral DNA in a LightCycler, using hybridisation probes specific for either the A(601) nt or the C(601) nt. This method can, during the amplification reaction, identify single-base mutations. Initially each sample underwent two PCR reactions: one with the C(601)-specific probe and one with the A(601)-specific probe, and using a standard concentration curve the concentration of amplification products containing the A(601) nt (as detected by the fluorescent signal from the A-specific probe) and the C(601) nt (as detected by the C-specific probe) in each sample was determined. However, a more accurate method whereby both C(601) and A(601) amplification products in each sample could be detected in a single PCR reaction was used to produce the data shown in Figs. 3.3, 3.4 and 3.5. In each PCR reaction, only the C(601) specific probe was used. A melting curve analysis, which genotypes single nucleotide polymorphisms, was performed after the completion of the amplification cycles, separating fragments containing C(601) from those containing A(601). The latter PCR-amplified fragments melted at a lower temperature due to the polymorphism (shown in Fig. 3.3A for plant 4). The peak area from each curve was calculated using LightCycler software, and related to the relative amount of each virus' DNA present in the plant samples by using standard plasmid samples containing differing ratios of pKomRb⁻ [C(601)] : pKomRb^{-C(601)A} [A(601)], from 100C(601): 0A(601) to 0C(601): 100A(601) (data not shown). The peak areas of these standards were plotted as a standard curve, and used to determine the ratios of the two viruses present in each leaf sample at different time points.

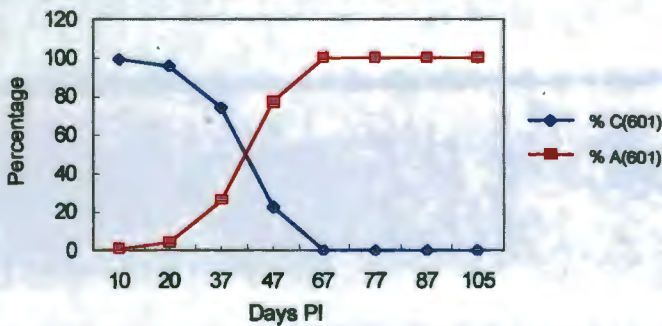
The extremely high frequency at which the C(601)A reversion occurred indicates that there is an exceptionally strong selective advantage of an A over a C at position 601 of the Rep gene. The virus population in all plants initially infected with MSV-KomRb⁻, whether by agroinoculation or leafhopper inoculation, was eventually completely superseded by the A(601) revertant virus. This was determined by non-digestion of PCR products with *Bfu*I (Table 3.2) and from the data in Figs. 3.3 and 3.4. Even in plant 1 (Fig. 3.5), the C(601) viral population, which had not started to revert by day 32 PI (confirming the data in Table 3.4) was being overtaken by the A(601) population by day 73 PI. The A(601) reversion started to appear in plant 4 just 10 days PI (although it only represented 1% of the population at that time, which is too little to be seen in Fig. 3.3B), and in plant 5 20 days PI (representing 2% of the population at that time). Although this was earlier than detected in plant 5 by the A(601)-specific primer using conventional PCR (Table 3.4), amplification using the LightCycler is much more sensitive and efficient than conventional PCR. When constructing a concentration standard curve by PCR using the LightCycler, plasmid DNA as little as 0.1 pg was amplified and easily detectable as a band of the

correct size on a gel. Accordingly, Fig. 3.4A shows that the A(601) virus population comprised 15% of the total population by day 30, which is when it was first detected using conventional PCR.

A



B



C

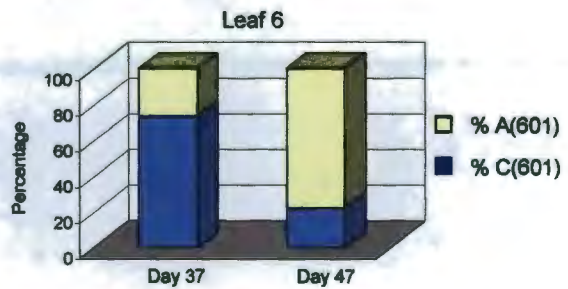


FIGURE 3.3 Change in composition of virus population in plant 4 and in individual leaves over time. (A) Melting curves of seven leaf samples from plant 4. Also included are positive and negative controls, and two leaf samples from plant 5. Using the C(601)-specific probe, PCR-amplified Rep fragments containing C(601) melt at 66.3°C (blue cursor), while fragments containing A(601) melt at 60.2°C (green cursor), clearly separating the two populations. Melting curves with a T_m below 60°C represent primer dimers. 1 and 2= samples from plant 5 taken 20 and 32 days post inoculation (PI) respectively. 3-9 = Plant 4 samples: 3= leaf 2, day 10; 4 = leaf 3, day 10; 5= leaf 4, day 10; 6 = leaf 5, day 20; 7 = leaf 6, day 37; 8 = leaf 7, day 67; 9 = leaf 11, day 105. 10 = 100A:0C (positive control for A(601) samples); 11 = 100C:0A (positive control for C(601) samples); 12 = uninfected maize genomic DNA (negative control). (B) Change in composition of virus population in plant 4 over time. At increasing time points from 10 to 105 days PI the percentage of the PCR-amplified Rep fragments containing C(601) or A(601) in leaf samples is shown. (C) Change in composition of virus population in leaf 6 of plant 4 over a period of 10 days. A sample was taken when symptoms were first observed in leaf 6 (37 days PI) and then 10 days later (47 days PI).

The melting curve experiment shown in Fig. 3.3A is just one of a few such experiments, to illustrate how the graphical data in Figs 3.3, 3.4 and 3.5 were obtained.

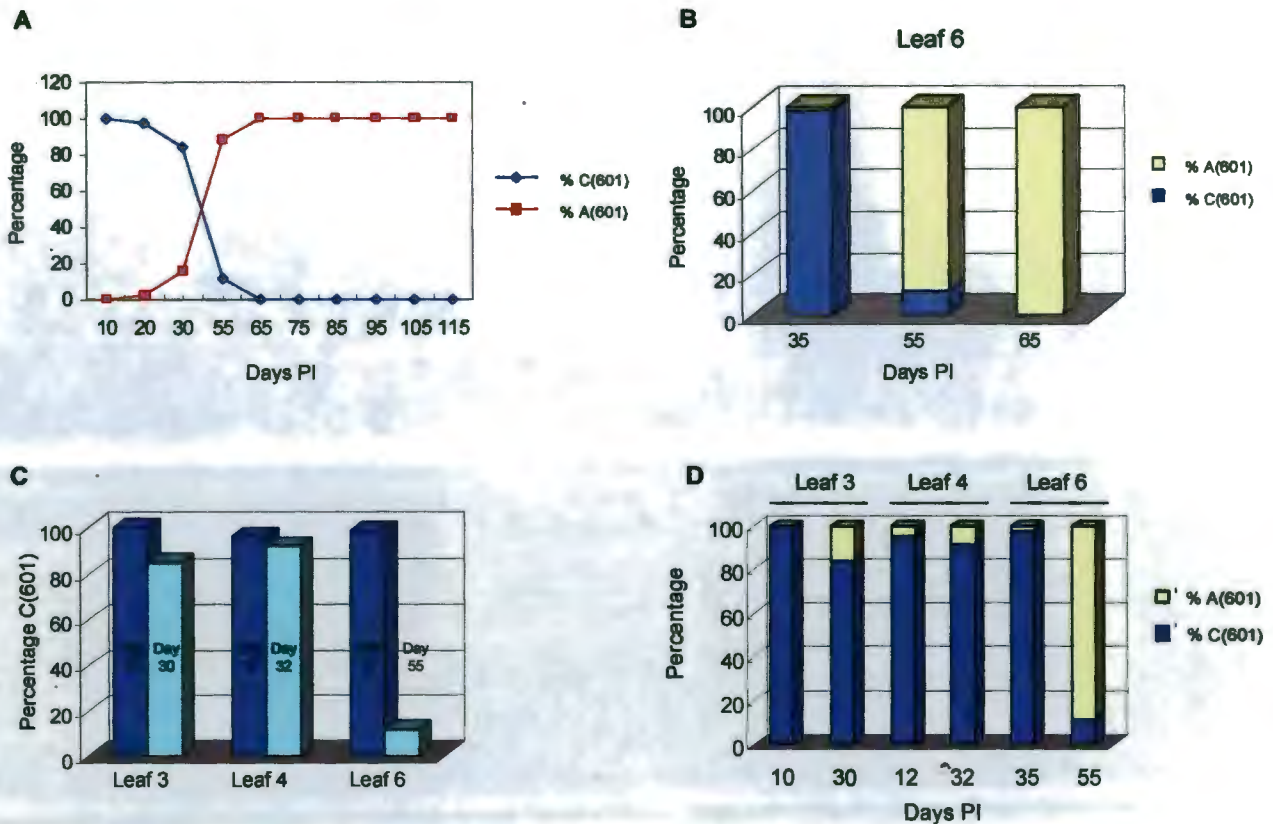


FIGURE 3.4 Change in composition of virus population in plant 5 (A) and in individual leaves of plant 5 (B, C and D) over time. (A) At increasing time points from 10 to 115 days PI the percentage of PCR-amplified Rep fragments containing C(601) or A(601) in leaf samples is shown. (B) Change in virus population in leaf 6 of plant 5 over a 30-day period. A sample was taken when symptoms were first observed on leaf 6 (35 days PI) and then 20 and 30 days later. (C) Change in percentage of virus population consisting of MSV-KomRb^r [C(601)] in individual leaves over a 20 day period. Samples were taken when symptoms were first observed on leaf 3, 4 and 6 (10, 12 and 35 days PI respectively), and then 20 days later (day 30, 32 and 55 for leaf 3, 4 and 6 respectively). (D) Change in virus population composition in individual leaves over a 20-day period.

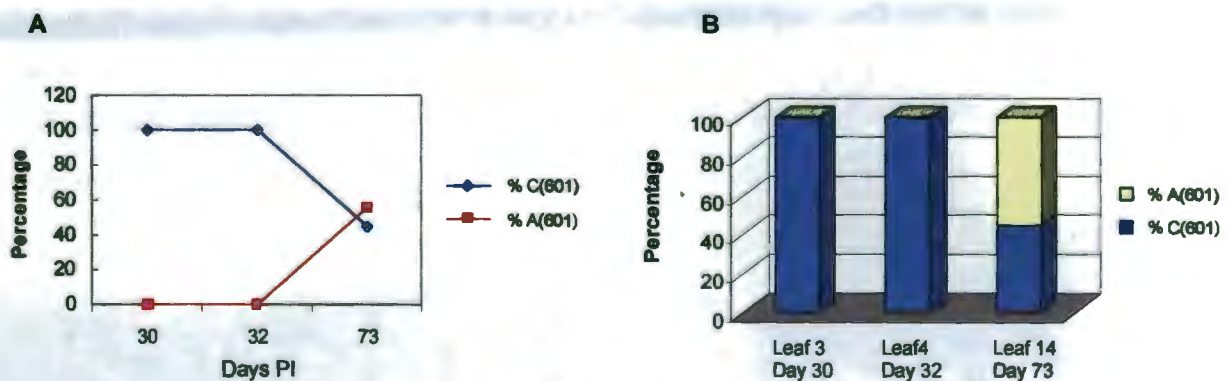


FIGURE 3.5 Change in composition of virus population in plant 1 (A) and in individual leaves of plant 1(B) over a time period of 30 to 73 days PI.

The rapid appearance of the A(601) virus population indicates that the mutation of C to A had occurred and been maintained at position 601 in the Rep gene very soon after agroinoculation, certainly before each leaf emerged from the whorl. Not only does this reflect the extremely strong selective advantage of the A(601) nt, but it also indicates that the mutation rate of MSV, whose genome was previously thought to be relatively stable, is extremely high. Whereas the mutation rate of viruses with RNA genomes is known to be high, it is more unusual for a DNA virus to mutate at the kind of rate inferred by Figs. 3.3 and 3.4. Although the actual mutation rate of MSV-Kom was not determined, it could be estimated by a similar method to the one presented, using a greater number of plant samples taken at closer time intervals than ten days. For example, it should be possible to pinpoint in each plant the exact time the nt reversion occurs, the exact point at which the total virus population consists of 50% A(601) and 50% C(601), and the point at which the A(601) population replaces the C(601) population. Even using the limited number of samples in this study, it can be seen that in plant 4 the A(601) virus population comprised 50% of the total population at a time point between day 37 and 47 days PI (Fig. 3.3B), and in plant 5 between day 30 and 55 (Fig. 3.4A). The A(601) virus comprised 100% of the total population by day 67 PI in plant 4 and by day 65 PI in plant 5. If samples had been taken at closer time points, the likely outcome would have been that the A(601) virus population would have reached 50% and 100% at the same time in both plants.

The A(601) virus population was originally expected to progressively overtake the C(601) population leaf by leaf as the plant aged, at any given time point younger leaves containing a higher proportion of A(601) than older leaves. It was therefore interesting to note that the ratio of the A(601) to C(601) virus population changed in individual leaves over time. For example, when leaf 6 of plant 4 first emerged from the whorl 37 days PI, 26% of the virus population in the leaf consisted of the A(601) virus, whereas 10 days later the A(601) virus comprised nearly 78% of the total population in the same leaf (Fig. 3.3C). Similarly, when leaf 6 of plant 5 first emerged from the whorl 35 days PI, the A(601) virus only comprised 1.6 % of the population; 20 days later it made up 88% of the population, and 30 days later the population consisted entirely of the A(601) virus (Fig. 3.4B). At similar time points in plant 5 (Fig 3.4D), the A(601) virus comprised 15.5% of the total viral population in leaf 3 (day 30 PI); leaf 4 of the same plant contained a viral population consisting of 8% A(601) (day 32 PI), while the total viral population in leaf 6 only consisted of 1.6% A(601) (day 35 PI). This indicates that, although the A(601) reversion probably originated in the leaf primordia (with the A(601) virus appearing with the emergence of each leaf from the whorl), the reversion continued occurring and being maintained at different rates in individual leaves, even after they had fully expanded.

3.3.3 Investigations into the Possible Selective Advantages of the C(601)A Reversion

The first option investigated was whether the C(601)A reversion in the RBR-binding domain increased the binding affinity of the mutant Rep to RBR. In order to clone the revertant Rep gene, the sequence encompassing the C(601)A reversion (obtained from a previously cloned and sequenced PCR fragment amplified from an infected maize sample) was introduced into pSKRepRb⁻ to create pSKRepRb^{-C(601)A}. Since RBR binds to RepA rather than Rep, a RepA gene containing the C(601)A reversion (RepARb^{-C(601)A}) was made as described in Chapter 2 for wt RepA and RepARb⁻. The ability of the RepARb^{-C(601)A} gene product to bind to the maize RBR protein (ZmRb1) was determined using a yeast two-hybrid assay (Table 3.5).

While wt RepA interacted with RBR, RepARb^{-C(601)A} did not. Therefore it appears that the loss of RBR interaction ability in RepRb⁻ was not restored by the C(601)A reversion. It must be noted that the same yeast strain and the same amount of 3-AT was used as in the yeast two-hybrid assay in Chapter 2. Since a truncated Rep protein containing a wt RBR interaction domain (RepΔC2) did not interact with RBR, it was concluded in that study that the assay may have been too stringent for the yeast strain used, and that weak interactions may not have been detected. The

TABLE 3.5 Interactions between MSV RepA and RepARb^{-C(601)A} and the maize retinoblastoma protein (ZmRb1).

Binding domain fusion	Activation domain fusion	Growth of Yeast (Trp ⁻ Leu ⁻ , His ⁻ , + 5mM 3-AT)	Growth of Yeast (Trp ⁻ Leu ⁻ , + 5mM 3-AT)
pGBT9	pADRepA	-	++
pGBT9ZmRb1	pADRepA	++	++
pLamC	pADRepA	-	++
pGBT9	pADRepARb ^{-C(601)A}	-	++
pGBT9ZmRb1	pADRepARb ^{-C(601)A}	-	++
pLamC	pADRepARb ^{-C(601)A}	-	++
pGBT9	pGAD424	-	++
pGBT9ZmRb1	pGAD424	-	++
pLamC	pGAD424	-	++
p53	pSV40	+++	++
p53	pADRepA	-	++

Yeast cells co-transformed with pGBT9ZmRb1 and pGAD424-RepA fusion products were grown on drop-out medium supplemented with 5mM 3-AT in the presence or absence of histidine. Only strains containing interacting fusion proteins can grow in the absence of histidine. As negative controls, pGBT9 alone (GAL4 binding domain not fused to the ZmRb1 protein) was co-transformed with each pGAD424-RepA plasmid, as was pLamC, which is unable to interact with either ZmRb1 or MSV RepA. As a positive control, yeast was transformed with plasmids expressing p53 (BD fusion) and SV40 T-Ag (AD fusion). This assay was repeated with the same results

same may apply to this assay; for example the C(601)A reversion may have allowed the interaction of RepRb^{-C(601)A} and RBR protein with a very low, undetectable, affinity. Decreasing the amount of 3-AT or using a strain that is less sensitive to 3-AT may allow the detection of low affinity interactions. However, considering that the C(601)A reversion did not restore the wt amino acid in the RBR interaction motif, and that the amino acid change effected by the nt reversion (underlined in LLCLK to LLCIK) is not one of the conserved amino acids that comprise the RBR protein interaction motif (LxCxE), it is not so surprising that it did not have an effect on RBR protein-Rep interaction.

It is highly unlikely that the amino acid change caused by the C(601)A reversion (leucine to isoleucine) provided a selective advantage, since the change was a very conservative one. Examples of where the viral nucleotide rather than amino acid sequence could be important for viral infectivity include codon usage in maize, spliceosome recognition sites for intron splicing, and *Rep* DNA or RNA secondary structure.

Codon usage is unlikely to be the reason for the nt reversion. The codon for leucine (CUU) introduced by the Rb⁻ mutation does not have rare usage in maize (Fennoy *et al.*, 1993), and there is no more bias for usage of the codon for isoleucine (AUU) introduced by the C(601)A reversion.

The nucleotide sequence of *Rep* may have an influence on the efficiency with which the *Rep* transcript is spliced. One of the essential steps in splicing of introns is the precise recognition of 5' and 3' splice sites by various snRNPs and various protein factors. RNA binding factors called SR proteins are involved in 5' and 3' splice site recognition as well as being crucial for the assembly of the spliceosome (Lopato *et al.*, 1999). This early function involves sequence-specific binding by the SR protein to the pre-mRNA. SR proteins bind to RNA sequences known as splicing enhancers with distinct specificities. Splicing enhancers, which usually contain a purine-rich element, most frequently lie within exons and facilitate splicing of the upstream intron (Manley and Tacke, 1996). Since SR proteins interact with RNA in a sequence specific manner, altering the enhancer sequence could affect the affinity with which the SR protein binds the RNA.

Exonic sequences upstream of the intron could also have an effect on splicing efficiency by influencing RNA secondary structure. While there is no evidence that the sequence upstream of the MSV *Rep* intron encompassing the C(601) mutation (the mutation is 41 bp upstream of the 5'

splice site) contains an enhancer element or is recognised by the spliceosome, there is evidence of a single nucleotide mutation ~50 bp upstream of the *MP* 5' splice junction that increased splicing efficiency of the mutant's *MP* transcript (M. Boulton, pers. comm.). If the C(601) mutation in Rep also increased the splicing efficiency of the transcript, it could explain why the mutation had no effect on viral replication in BMS but required a reversion of C(601)A when infecting maize. Increased splicing efficiency would result in a reduction in the expression of RepA, which should not have an effect on viral replication in BMS since the two main functions of RepA appear to be to interact with RBR protein and to transactivate the *CP* promoter, neither of which should be required for the virus to replicate in BMS suspension cells. To infect maize, however, interaction with RBR is possibly required (or at least advantageous) early in the infection cycle, and later in the cycle the CP is required along with the MP for movement of the viral genome. In particular, CP transports viral DNA to the nucleus for replication and transcription, and interacts with MP to target the DNA to the cell periphery to be transported through the plasmodesmata into another cell (Liu *et al.*, 1996; Kotlizky *et al.*, 2000). Thus, the C(601)A reversion could be required for infectivity of MSV-KomRb⁻ in maize to lower the splicing efficiency of the *Rep* transcript, allowing increased RepA expression and transactivation of the *CP* promoter. The actual effect of the Rb⁻ mutation and C(601)A reversion on splicing efficiency of the *Rep* transcript was not determined. However, in Chapter 4 transient replication assays in BMS using intronless Rep genes (which in effect produce spliced Rep transcripts) containing the Rb⁻ mutation and the C(601)A reversion are described.

The theoretical influence of the Rb⁻ mutation and the C(601)A reversion on the secondary structure of the Rep transcript was investigated using the RNA folding programme, RNADRAW. As can be seen in Figs. 3.6 and 3.7, the theoretical Rep transcript secondary structure was radically changed in the area of the three-nt Rb⁻ mutation, including the potential intron structure. A single nt change of C(601)A, however, potentially changed the mutant structure to one closely resembling the wt structure (Fig.3.6C). In particular, the branch and other structures to the left of the mutated area were restored in RepRb^{-C(601)A} to the wt structure (see the black-boxed area in Fig. 3.6A and C). The potential intron structure, however, remained the same in the A(601) and C(601) transcript structures. As can be seen in Fig. 3.7, the nucleotides comprising the RBR-interaction motif potentially form a hairpin structure, which is destabilized by the three-nt Rb⁻ mutation and restored, although not exactly as in the wt, by the C(601)A reversion. The structures to the left of the mutation are completely restored to wt by the single nt reversion. Using the RNADRAW computer program to change the C(601) nt to G(601) or T(601), however, did not

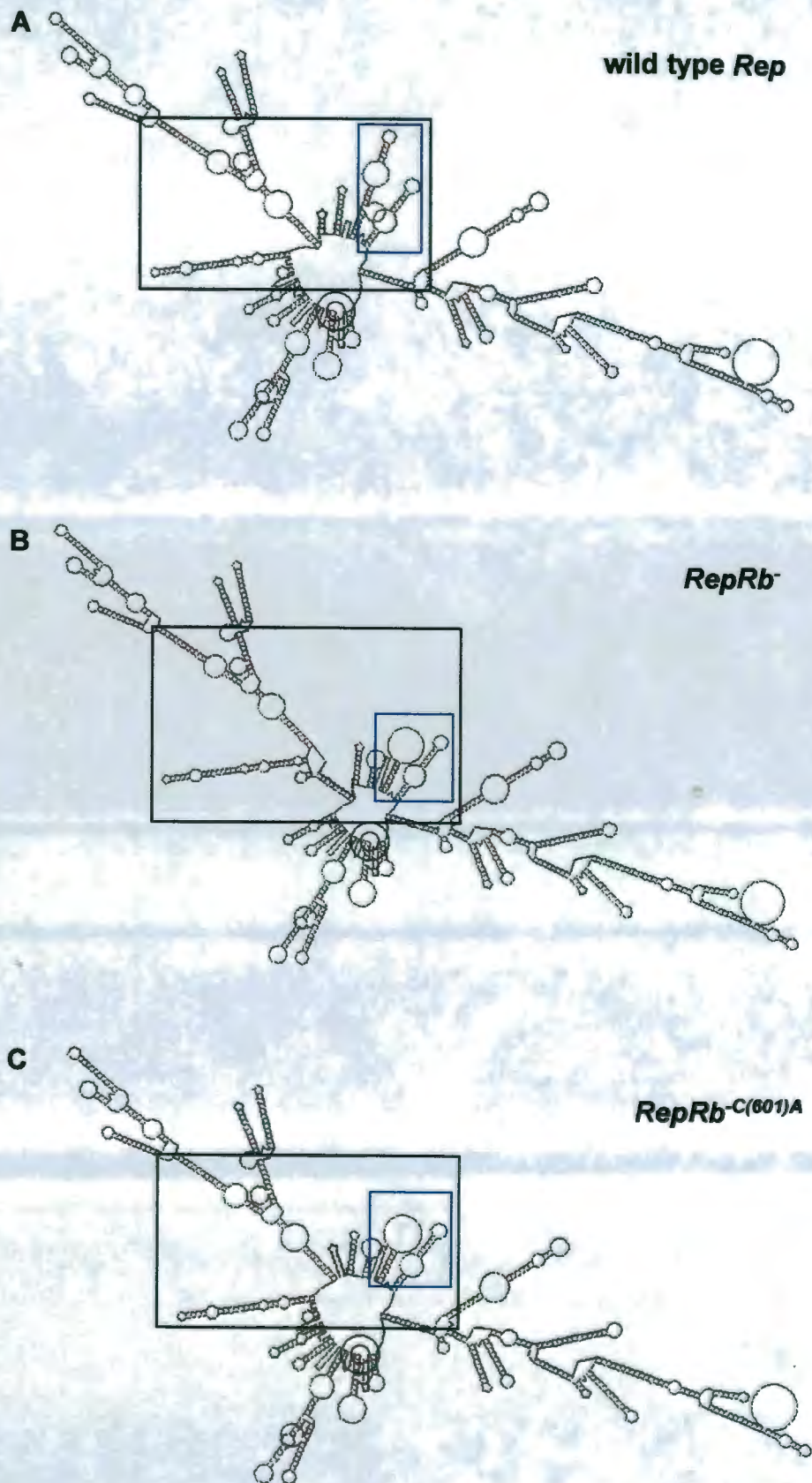


FIGURE 3.6 The predicted structure of RNA transcripts of (A) wt *Rep*; (B) *RepRb⁻* and (C) *RepRb^{-C(601)A}*. The black-boxed area, containing the Rb⁻ mutation and C(601)A reversion, is magnified in Fig. 3.7. The blue-boxed area is the putative structure of the intron, which is spliced in the mRNA.

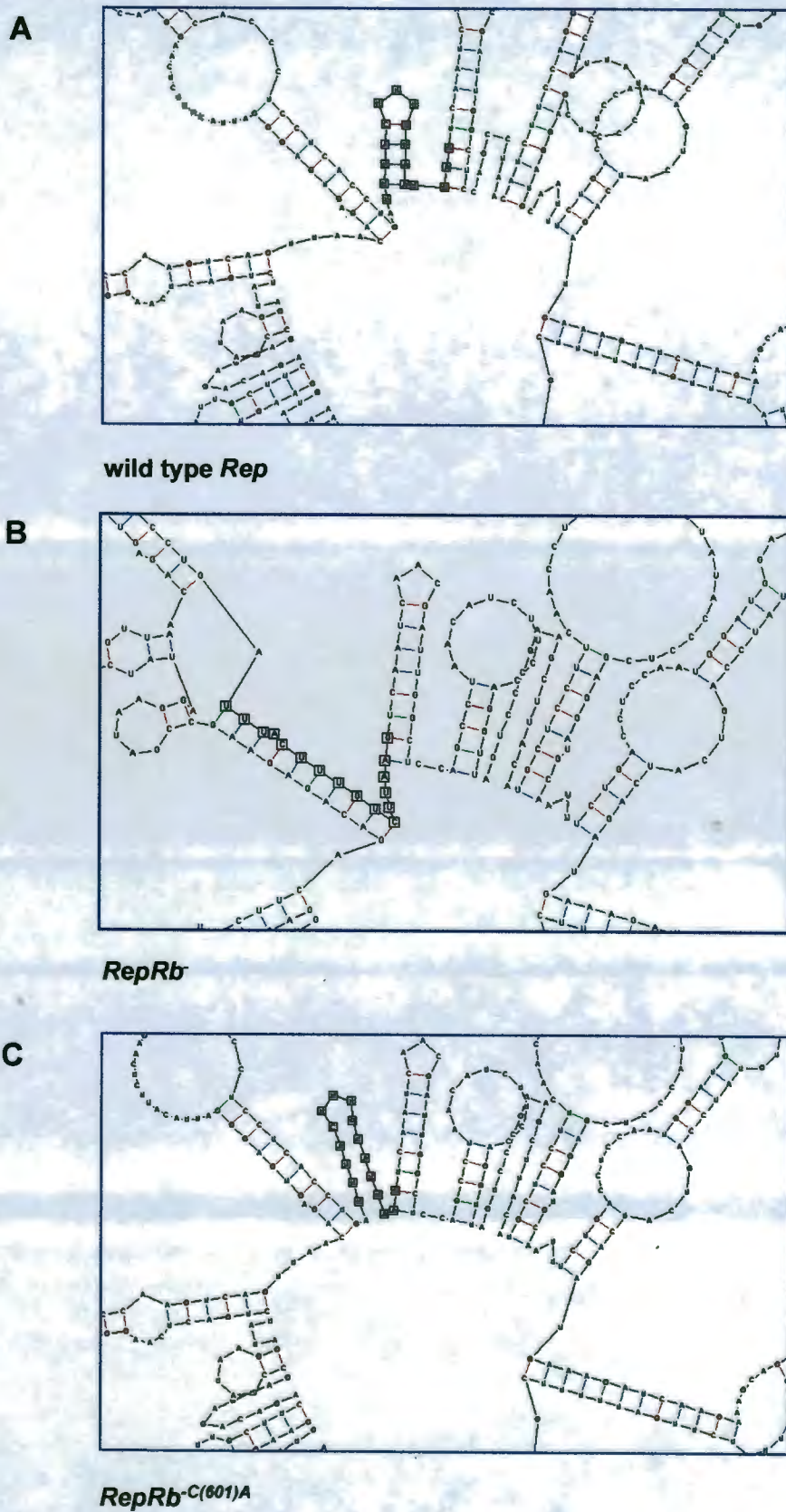


FIGURE 3.7 Predicted structure of the mutated area of the RNA transcript of (A) *Rep*, (B) *RepRb⁻* and (C) *RepRb^{-C(601)A}*. The nucleotides comprising the RBR-interaction motif are boxed.

restore the wt-like secondary structure (data not shown). This seemed to provide evidence, at least in theory, that the A(601) reversion was required to stabilize the RNA secondary structure, or perhaps, using splicing efficiency as an example, to restore exon structural recognition sites for the splicing apparatus.

To determine experimentally the absolute requirement in *Rep* of an A at position 601, *Rep* mutants were made with either a G or a T at position 601, and the *Rep* gene from the C(601)A revertant virus was cloned. The mutations were then introduced into pKomRb⁻, resulting in four different viruses with the same A(602)T and G(604)A mutations as in pKomRb⁻ (see Table 3.3), but with either a C, G, T or A at position 601. These were made into agroinfectious constructs, and 14 maize seedlings were agroinoculated with each construct. This was repeated in three separate experiments. The percentage chlorotic area of leaves two to six of each plant (Fig. 3.8) was analysed as in 3.3.1. Three months after agroinfection, viral DNA was cloned from three maize plants infected with each mutant virus, and the *Rep* gene from all the clones sequenced to confirm whether or not the C, G, T or A nucleotides had been maintained at position 601.

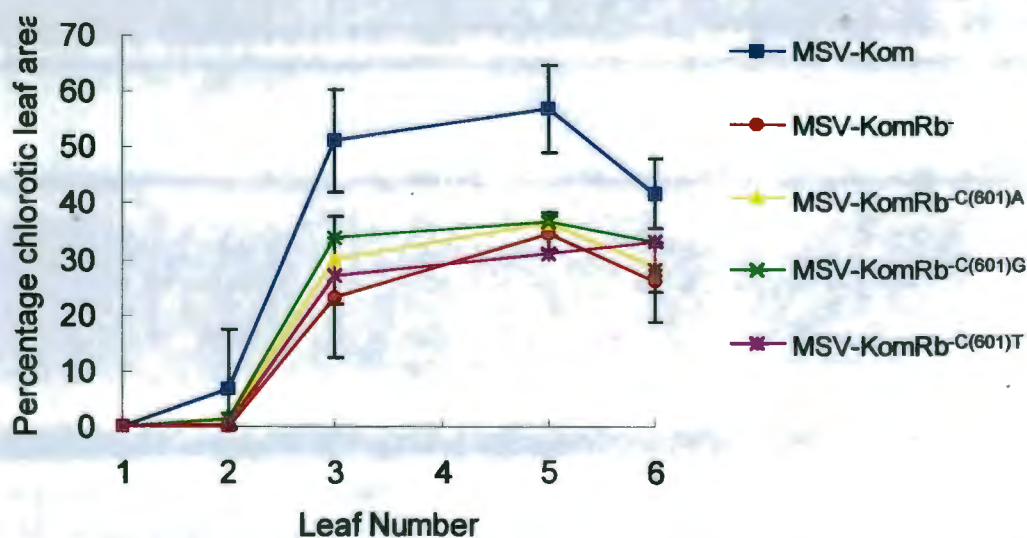


FIGURE 3.8 Infectivity of MSV-Kom mutants in maize cv. Jubilee in comparison with wild type MSV-Kom. Percentage chlorotic areas were recorded on leaves 2 through 6 of 14 maize plants agroinoculated with each virus in three separate experiments. Error bars represent 95% confidence intervals of the mean.

The infectivity of all the mutant viruses was surprisingly no different: whether they contained a C, T, G or A at position 601, the percentage leaf chlorosis caused by each virus was attenuated to the same degree when compared with the chlorosis caused by the wt virus (Fig. 3.8). Despite MSV-KomRb^{-C(601)A} having a "head start", being agroinoculated at the same time as MSV-

KomRb⁻, having an A(601) it did not improve its infectivity. As can be seen in Table 3.5, all Rep genes amplified from maize plants agroinoculated with MSV-KomRb⁻ contained the C(601)A reversion 3 months PI. However, this was not the case with viruses containing G(601) and T(601). All Rep genes amplified and cloned from one out of three plants agroinoculated with pKomRb^{-C(601)T} contained a T(601)A reversion, but it did not occur in the other two plants analysed. A G(601)A reversion was detected in one plant agroinoculated with pKomRb^{-C(601)G}, but in only one clone out of nine sequenced. The other two plant samples analysed contained no reversions. Clones isolated from these plants did, however, contain numerous mutations in the Rep gene, and one clone isolated from a plant agroinoculated with pKomRb^{-C(601)T} contained a reversion of T(602)A, which never occurred in plants agroinoculated with pKomRb⁻. Most of the naturally occurring mutations resulted in amino acid changes, although without determining the effect of the changes on protein structure it is difficult to conclude why these mutations occurred and were maintained. One interesting mutation occurred in more than one clone in both pKomRb^{-C(601)T} and pKomRb⁻. A nt mutation of T(662)G in the intron changed a methionine (M) at position 221 to an arginine (R). It could be speculated whether the methionine is a start codon for a "RepB" protein produced from unspliced transcripts, but there are two other methionines closer to the C2 ORF that could serve the same purpose, if it exists at all. In RepA, M(221) is the first amino acid of the transactivation domain delineated by Horvath *et al.* (1998). While it is pointless to speculate on the effect of the M(221)R mutation on the transactivation properties of RepA without experimental validation, the fact that the same mutation occurred in two different viruses in separate agroinoculations indicates that it may have some selective advantage. The remainder of the mutations did not occur in any recognisable domains. However, of interest was a silent mutation of T(516)C, coding for phenylalanine [F(172)] in plant 2 agroinoculated with KomRb^{-C(601)G}. The original codon for F, UUU, was changed to UUC by the mutation, which has much greater usage in maize than UUU (Fennoy *et al.*, 1993). The relative synonymous codon usage (RSCU), which mathematically describes the disproportionate use of synonymous codons in maize, is 0.5 for UUU and 1.5 for UUC. In maize, codons ending in G or C are most frequently used, reflecting the high GC content of the maize genome (Fennoy *et al.*, 1993). It must be born in mind that Rep genes containing a C, G, T or A at position 601 cannot be directly compared in terms of the single nucleotide change, since each mutation resulted in a change in amino acid residue at position 201. The change in amino acid from aliphatic L in pKomRb⁻ (LLCLK) to aromatic F in KomRb^{-C(601)T} (LLCFK) is particularly radical. The change is more conservative in KomRb^{-C(601)G}, from aliphatic L to aliphatic valine (V) (LLCVK). The wt amino acid residue at

TABLE 3.5 Mutations and reversions in the Rep gene of agroinoculated mutant viruses

Virus inoculated/ Plant no.	¹ Mutation		Location	² No. of clones/No. sequenced
KomRb^{-C(601)T}				
	Nucleotide	amino acid		
Plant 1	T(601) (original mutation)	F(201)	Rb-binding domain	0/6
	T(601)A (reversion)	F(201)I	Rb-binding domain	6/6
Plant 2	T(601) (original mutation)	F(201)	Rb-binding domain	8/8
	T(602)A (reversion)	F(201)Y	Rb-binding domain	2/8
	A(452)C	E(151)A	RepA	6/8
Plant 3	T(601) (original mutation)	F(201)	Rb-binding domain	8/8
	T(662)G	M(221)R	Intron (RepA-activation domain)	2/8
	C(447)T	S(149)S	RepA	2/8
KomRb^{-C(601)G}				
Plant 1	G(601) (original mutation)	V(201)	Rb-binding domain	8/9
	G(601)A (reversion); A(605)G	V(201)I; K(202)R	Rb-binding domain	1/9
Plant 2	G(601) (original mutation)	V(201)	Rb-binding domain	2/2
	T(516)C	F(172)F	RepA	2/2
Plant 3	G(601) (original mutation)	V(201)	Rb-binding domain	2/2
KomRb⁻				
Plant 1	C(601) (original mutation)	L(201)	Rb-binding domain	0/4
	C(601)A (reversion)	L(201)I	Rb-binding domain	4/4
	T(662)G	M(221)R	RepA-activation domain (intron)	1/4
Plant 2	C(601) (original mutation)	L(201)	Rb-binding domain	0/2
	C(601)A (reversion)	L(201)I	Rb-binding domain	2/2
KomRb^{-C(601)A}				
Plant 1	A(601) (original reversion)	I(201)	Rb-binding domain	4/4
MSV-Kom				
Plant 1	A(601) (wt sequence)	N(201)	Rb-binding domain	1/1
Plant 2	A(601) (wt sequence)	N(201)	Rb-binding domain	1/1

¹For each plant the number of virus clones containing the original mutation (or A nucleotide in the case of the agroinoculated revertant and wt viruses) at position 601 is shown first. The number of virus clones containing mutations or reversions in the Rep genes are shown next. Nucleotide changes were only taken as being genuine mutations of the viral genome if the same changes occurred in Rep fragments amplified from more than one PCR reaction, if they occurred in both forward and reverse directions sequenced, and if they occurred in more than one clone.

²The number of clones containing the original mutation (or wt sequence in the case of MSV-Kom) and naturally occurring mutations in the Rep gene are shown relative to the total number of virus clones sequenced from each plant.

position 201 in MSV-Kom is the acidic glutamic acid (N) (LLCNE). It is therefore surprising that these amino acid mutations, especially in MSV-KomRb^{-C(601)T}, were tolerated at all.

Despite, in most cases, there being no apparent reason for the Rep gene mutations presented in Table 3.5, three general patterns can be seen. The first is that, although an A at position 601 of *Rep* is not absolutely required for infectivity in maize, a C is not tolerated, since it was the only nucleotide that was never maintained in that position. The second is that viruses with an A at position 601 in Rep are more stable than those with alternative nucleotides in the same position. Neither MSV-KomRb^{-C(601)A} nor MSV-Kom, both containing an A(601), contained any mutations in the entire Rep gene. Similarly, all clones isolated from plant 1 agroinoculated with pKomRb^{-C(601)T}, contained the T(601)A reversion and no other mutations. Only one clone out of 4 in one plant agroinoculated with MSV-KomRb⁻ contained a mutation other than the C(601)A reversion. In that case the mutation may have been advantageous since the same mutation, residing in the RepA activation domain, also occurred in a different virus. Viruses containing no reversions to A(601), however, contained numerous mutations for no apparent reason. These may be compensatory mutations or may simply be mutations that were tolerated. It is possible that changing the A at position 601 in Rep to another nucleotide decreases the fidelity of the host DNA polymerase, although it would be hard to explain how this would occur. The third pattern observed was that all the mutations that occurred, even if they apparently were not related to the original Rb⁻ mutation, were in the RepA ORF. Not a single mutation was detected in the C2 ORF of the Rep gene. Thus, if the mutations did serve any purpose, they potentially affected functions of RepA, as most probably did the Rb⁻ mutations and the C(601)A reversion. Future work will include modelling of the wt and mutated proteins to determine if the mutations have any important effects on the tertiary structure of the Rep and RepA proteins. The 3D NMR structure of the conserved catalytic domain of the TYLCV Rep protein has been solved (Campos-Olivas *et al.*, 2002), and may be used as a basis to predict the structure of the MSV wt and mutated Rep proteins.

3.4 CONCLUSION

While the MSV Rep LxCxE to LxCxK mutation had no effect on viral replication in BMS suspension cells, it reduced the infectivity of MSV-KomRb⁻ in maize, symptom severity being attenuated by approximately 50% compared with the wt infection. This may have been as a result of the abolition of RBR protein-binding ability in the mutant Rep gene, which may have had an

effect on tissue specificity of the virus (for example, limiting the number of tissue types in which the virus could replicate) or may have lowered viral replication efficiency in differentiated cells. It was interesting to note that MSV-KomRb⁻ was transmissible to maize by leafhoppers, which deposit the virus into developmentally mature leaf cells, where an interaction with RBR would be expected to be required for viral replication. Thus, even in natural transmission of MSV, interaction with RBR is not absolutely required, although it may be beneficial, for the virus to establish a systemic infection.

Taking into account the high-frequency occurrence of a reversion of C(601)A in the mutated Rep gene, which did not have an effect on RBR-protein binding efficiency, the three-nt mutation in the LxCxE motif may have had effects separate from RBR protein interaction. Considering the small size of the MSV genome, it would not be surprising for sequence motifs to have more than one function or role in the viral life cycle. A definite conclusion as to what function the three-nucleotide mutation, in particular that of A(601)C, interfered with could not be reached, but there were several promising theories that could be further explored. The fact that the Rb⁻ mutation had no effect on viral replication in BMS indicates that it may not have affected the functions of Rep, but rather RepA. Since the C(601)A reversion in RepARb^{-C(601)A} did not restore the protein's RBR-binding ability, a plausible function of RepA that would affect viral infectivity if interfered with is that of transactivation of the CP promoter. While the Rb⁻ mutation may not have directly interfered with the RepA transactivation domain itself (see Fig. 2.1 for the positions of the RBR-interaction motif and the transactivation domain in RepA), it may have resulted in a reduction in the level of RepA expression. An example of where a single nucleotide change could have an effect at both the RNA and protein level is that of splicing of the Rep transcript, as discussed in the section 3.3.3. Although the effects of the Rb⁻ mutation and the C(601)A reversion on splicing efficiency of the Rep transcript was not determined experimentally, for the purposes of this discussion it is taken as the most plausible reason for the C(601)A reversion, since it is the one theory that reconciles all the data presented in this chapter.

First, a C(601)A reversion leading to increased levels of RepA (expression of RepA having been reduced by the C(601) mutation) could result in the infectivity graphs presented in Figs. 3.1. and 3.8. For example, the nt reversion could have allowed the virus to systemically infect maize by increased expression of RepA and subsequent transactivation of CP, but the virus did not reach wt levels possibly because it cannot interact with RBR, which may have restricted virus replication to certain tissues, or lowered its replication efficiency.

Second, the C(601)A reversion theoretically had a positive effect on the RNA secondary structure, potentially restoring the structure from one that was destabilized by the C(601) mutation to one very similar to the wt structure. While the nucleotide sequence upstream of the 5' splice site may affect splicing efficiency, so could the RNA secondary structure.

Third, if a C at position 601 increases splicing efficiency, it would explain why any nucleotide other than a C at that position is tolerated. Although there is probably a selective advantage of the wt A over a G or a T at position 601, a G or a T may also have the effect of reducing splicing efficiency simply because they have replaced the C.

Fourth, the fact that the infectivities of Rb⁻ mutant viruses containing either a A, C, G or T at position 601 were the same, may reflect the speed at which the reversion in MSV-KomRb⁻ occurred, coupled with that fact that the effects of the reversion may only be required later in the infection cycle, for transactivation of the CP promoter.

While a model for the C(601)A reversion is described below, it is completely speculative and needs to be validated experimentally. However, it is notable that a naturally occurring mild MSV-Ns isolate contained a single nucleotide mutation in the MP gene which led to increased splicing of the MP transcript, resulting in decreased expression of MP and a subsequently attenuated infection compared with the naturally severe infection caused by the wt MSV-Ns (M. Boulton, pers. comm.)

In the alternate splicing model, the C(601) mutation in MSV-KomRb⁻ does not affect replication of the virus early in the infection cycle, especially if the virus is inoculated into cells that are competent for DNA replication, as is the case with agroinoculation. This allows the virus to reach wt titres early in the infection. Later in the infection cycle, the revertant virus starts appearing, the selective advantage being that it can express RepA to higher levels than the Rb⁻ mutant. By the time the infection cycle has reached the stage of requiring movement of the genome, enough reversions have occurred for expression of RepA to transactivate the CP promoter. Once CP has been expressed, it would presumably allow movement of the MSV-KomRb⁻ genome as well as the revertant virus' genome, which is why MSV-KomRb⁻ can be detected up to leaf 6 in infected maize plants. Beyond leaf 6 and around 60 days PI, however, the selective advantage of the revertant virus is so great that it completely takes over the population, presumably through the

ability to systematically invade new cells and start new infections cycles, while the Rb⁻ mutant cannot complete its own infection cycle without the aid of the revertant virus. As mentioned previously, the revertant virus does not reach titres as high as the wt virus because it cannot interact with RBR protein, at least to an extent detectable by yeast two hybrid assays. There are likely to be other domains in RepA and Rep that interfere with the cell cycle and other cellular pathways, such as the GRAB-binding domain in RepA and the *myb*-like domain in Rep, making a wt interaction with RBR protein, although advantageous, not essential. However, the possibility that there was a low level of RBR protein binding activity that as not detected in the yeast two hybrid assays cannot be excluded.

Apart from potential insights into the functions of RepA, the high frequency of the C(601)A reversion has great potential for determining the mutation rate of MSV under natural infection conditions, and investigating aspects of MSV evolution such as factors that influence the rates at which variants with increased fitness supersede their parents. For example, the influence of different host genotypes or species on the rate of the reversion could be determined. This could be valuable information when attempting to design resistance strategies that cannot be overcome by mutation of the viral genome.

Chapter 4

A Comparison of the Replication Efficiencies of MSV-Kom Rep Variants

Contents

ABSTRACT	140
4.1 INTRODUCTION.....	142
4.2 MATERIALS AND METHODS.....	143
4.2.1 Clone Construction.....	143
4.2.2 Transient Replication Assays	144
4.4.3 Agroinoculation	145
4.3 RESULTS AND DISCUSSION.....	146
4.3.1 The Effect of Rep Variants on MSV-Kom and MSV-KomRb⁻ Replication in BMS	146
4.3.2 The Effect of Mutations in the Rep RBR Protein Interaction Domain on Viral Replication in BMS	148
4.3.3 The Ability of Rep Variants to Trans-Replicate a Rep-Deficient MSV Genome.....	149
4.3.4 The Replication Efficiency of Rep Variants in the Presence and Absence of RepA.....	150
4.3.5 The Infectivity of MSV-Kom and MSV-KomRb⁻ in the Absence of RepA	153
4.4 CONCLUSION.....	155

ABSTRACT

In Chapter 3, the comparative infectivities of MSV-Kom, MSV-KomRb⁻ and MSV-KomRb^{-C(601)A} were determined. While MSV-Kom was more pathogenic than the latter two viruses, there was no difference between the pathogenicity of MSV-KomRb⁻ and MSV-KomRb^{-C(601)A} in maize. However, a reversion of C(601)A in the Rep gene of MSV-KomRb⁻ consistently occurred at a rapid rate. In order to determine the effect, if any, of the Rep gene C(601)A reversion on viral replication alone, a series of transient replication assays was undertaken involving *trans*-replication of MSV-Kom and MSV-KomRb⁻, and *trans*-replication of a Rep-deficient MSV-Kom genome (pKep177). In addition, in order to study more thoroughly the effect of the Rb⁻ mutation and the C(601)A reversion on viral replication in the absence of wt Rep, the replication of MSV-Kom, MSV-KomRb⁻, and MSV-KomRb^{-C(601)A} was assayed in BMS without any Rep constructs provided in *trans*. The overall trend of these experiment indicated that the C(601)A reversion somehow increased the replication efficiency of Rep, even above wt levels. As seen in previous replication assays, there was no significant difference between the replication efficiency of wt Rep and RepRb⁻. Further studies included assaying the *trans*-replication of pKep177 by wt Rep and RepRb⁻ in the absence and presence of RepA and RepARb⁻, in an attempt to determine whether the Rb⁻ mutation affected replication functions of Rep or RepA, and what effect the presence and absence of RepA has on the replication efficiency of Rep. A general trend could be seen in which intronless Rep and RepRb⁻ constructs replicated pKep177 to lower levels than did the equivalent full-length Reps, and the addition of equal amounts of RepA or RepARb⁻ reduced the replication efficiency even further. The aggregation state of Rep and RepA possibly plays an important role in each protein's functions, and it is likely that upsetting this balance, whether by the addition or removal of RepA from the system, resulted in lowered replication efficiency when compared with wt expression of Rep and RepA. The effect of abolishing RepA expression on viral infectivity in maize was determined by the removal of the Rep intron from MSV-Kom and MSV-KomRb⁻, which were then agroinoculated into maize seedlings. The inability to express RepA had the effect of greatly lowering, but apparently not abolishing, the infectivity of these viruses in maize.

4.1 INTRODUCTION

One of the earliest steps in the MSV life cycle after passage of the viral genome into the host cell nucleus is the host-directed, DNA-primed synthesis of a complementary (minus) strand, using the virion (plus) strand as the template. This event creates dsDNA (RFI), which serves as a template for transcription of viral genes and for rolling circle replication (RCR), processes that require tight control. Once transformed into RFI, the first priority would be to express Rep, which initiates RCR, and RepA, which may have a role in inducing the expression of host enzymes and co-factors required for the completion of RCR. Although RepA is not essential for viral replication (Schalk *et al.*, 1989; Collin *et al.*, 1996), the protein does appear to play an important role at various stages of the MSV replicative cycle.

In addition to its catalytic and DNA binding activities, the mastrevirus Rep is involved in several protein interactions, including homo-oligomerization, binding to RepA, and interaction with host proteins. The various activities associated with Rep appear to be determined in part by the aggregation state of Rep and/or RepA (Horvath *et al.*, 1998; Missich *et al.*, 2000). Studies have shown that MSV Rep monomers self-interact in yeast (Horvath *et al.*, 1998), and WDV Rep oligomers, consisting of six to eight monomers, have been visualized bound to DNA (Sanz-Burgos and Gutierrez, 1998; Castellano *et al.*, 1999). In TGMV, there is a clear correlation between the aggregation state of Rep and its function in replication and transcription (Orozco *et al.*, 2000). It is also likely that interaction between mastrevirus Rep and RepA plays an important role in processes such as the transcriptional activation of the *CP* promoter.

Apart from their function in the replication of the viral genome, Rep and RepA may have an integral role in the regulation of the entire virus infection cycle. RepA appears to be required at all stages of the infection process, from preparing the cellular environment early on for replication by binding to RBR and other cellular factors (Gutierrez, 2002), to down-regulating expression of Rep (thus influencing replication levels) later in the cycle (Collin *et al.*, 1996), to activating the expression of CP during the later stages of infection (McGivern, 2002). Mechanisms regulating RepA expression or activity may operate at the transcriptional, post-transcriptional (intron splicing) or post-translational stage. The latter may include post-translational processing, as well as biological aspects such as the aggregation state of RepA.

Thus, it is clear that MSV has evolved sophisticated mechanisms to ensure tight regulation of both V and C sense expression, possibly co-ordinated by RepA.

In this chapter, the effects on viral replication of the LLCNE to LLCLK mutation in pRepRb⁻ are studied in more detail, as well as the effect on replication of the C(601)A reversion in pRepRb⁻C^{(601)A}. In addition, the functions of Rep and RepA in viral replication are further analysed by the creation of separate intronless Rep and RepA constructs, both wt and containing the Rb⁻ mutation. The ability of MSV-Kom to establish an infection in maize while unable to express RepA is also determined. In this way the effect of abolishing RepA expression on viral replication in BMS and on viral infectivity in maize can be compared in an attempt to further understand the separate roles of RepA in viral replication and in establishing a systemic infection.

4.2 MATERIALS AND METHODS

4.2.1 Clone Construction

Construction of intronless Rep and RepRb⁻ genes, and RepA and RepARb⁻ genes was described in Chapter 2 (2.2.1.4).

The cloning of the C(601)A revertant Rep gene (pSKRepRb⁻C^{(601)A}) was described in Chapter 3 (3.2.3).

The introduction of the Rb⁻ mutation (LLCNE to LLCLK) into wt MSV-Kom to produce MSV-KomRb⁻, and the introduction of the C(601)A reversion into MSV-KomRb⁻ to produce MSV-KomRb⁻C^{(601)A} was described in Chapter 3 (3.2.1.2).

Intronless Rep genes were introduced into pKom602 and pKomRb⁻ by replacing each template's *NsiI/BglII* fragment, which spans the intron area, with the *NsiI/BglII* fragment from pSKRep Δ I and pSKRepRb⁻ Δ I respectively. To confirm the introduction of the intronless Rep gene into each viral genome, constructs were digested with *XhoI*, whose restriction site is in the *Rep* intron, and *BglII*, whose restriction site is in the SIR. Neither enzyme cuts elsewhere in pKom602 or pKomRb⁻. Constructs containing intronless Rep genes linearise, while a 480 bp fragment is liberated in constructs containing the *Rep* intron. The resulting clones, pKom Δ I and KomRb⁻ Δ I, were confirmed to be correct by sequencing the Rep genes using the TYR primer (forward

sequence) and NTP primer (reverse sequence), which anneal to sites either side of the intron area. The TYR and NTP primers were described in Chapter 3 (3.2.2.4).

RepRb^{-C(601)A}, RepΔI, RepRb^{-ΔI}, RepA, and RepARb⁻ were liberated from pSK and introduced into the expression vector pAHC17 by cloning a *Bam*HI/*Bgl*II fragment from each gene into the *Bam*HI site of the 5' dephosphorylated pAHC17, as described in Chapter 2 (2.2.1.3). The Rep and RepA gene variants all contain a *Bam*HI site 26 bp upstream of the *C1* start codon, introduced by the *C1* forward primer (described in Chapter 2, section 2.2.1). A *Bgl*II site was introduced into the RepA genes immediately after the RepA stop codon (described in Chapter 2.2.1.4). Thus, in each case, digestion with *Bam*HI and *Bgl*II excises the entire Rep, intronless Rep or RepA genes from pSK, in order to be cloned into the *Bam*HI site of pAHC17, downstream of the ubiquitin promoter. Rep gene variants in the sense orientation were selected, and designated pRepRb^{-C(601)A}, pRepΔI, pRepRb^{-ΔI}, pRepA and pRepARb⁻. The cloning of Rep and RepRb⁻ genes into pAHC17 has been described (2.2.1.3).

pKomΔI and pKomRb^{-ΔI} were made into agroinfectious constructs as described for MSV-Kom variants in Chapter 3 (3.2.1.2), by cloning each genome into the *Eco*RI and *Xba*I sites of the binary vector pBI121 (CLONTECH, CA) to obtain pBIKomΔI and pBIKomRb^{-ΔI}. *A. tumefaciens* C58C1 was transformed with these constructs using the method of An *et al.* (1988).

4.2.2 Transient Replication Assays

The abilities of pRep, pRepRb⁻, pRepRb^{-C(601)A}, pRepΔI and pRepRb^{-ΔI} to *trans*-replicate a replication-deficient MSV construct (pKep177) were determined by co-bombardment of each construct with pKep177 into BMS. A premature stop codon was introduced into the Rep gene of pKep177, which is a tandem dimer of the MSV-Kom genome, by the insertion of a *Pst*I site 72 bp downstream from the *C1* start codon (K.E. Palmer, UCT; see Fig. 2.2 for a diagrammatic representation of pKep177). Since the wt Rep gene is not expressed by pKep177, the ability of Rep constructs to replicate the viral genome in the absence of wt Rep was determined using the transient replication assay described in Chapter 2 (2.2.2). In each bombardment experiment, pKep177 was co-bombarded with equal amounts of pRep, pRepRb⁻, pRepRb^{-C(601)A}, pRepΔI or pRepRb^{-ΔI}; as well as equal amounts of pRepΔI + pRepA, pRepΔI + pRepARb⁻, pRepRb^{-ΔI} + RepA and pRepRb^{-ΔI} + pRepARb⁻.

In the quantitative PCR assay used in Chapter 2, the internal control, or spike, of known concentration was pKep177, which could be distinguished from the RF viral band of unknown concentration by digest with *Pst*I (see Fig. 2.2). Considering that in this Chapter the RF viral DNA amplified from BMS genomic DNA was from pKep177, quantitative PCR could not be used to determine the relative concentrations of replicated virus. Instead, each PCR run included a reaction sample containing 200 pg of plasmid pKep177. Band intensity (determined using GelTrak) of the RF viral bands of unknown concentration was converted to pg of RF viral DNA/100ng of BMS genomic DNA by determining the ratio of the band intensity of the 200 pg plasmid control with that of each viral sample.

The effect of pRep, pRepRb⁻, and pRepRb^{-C(601)A} on replication of pKom602 was determined by co-bombardment of equal amounts of each plasmid with pKom602 into BMS, allowing four days for viral replication, followed by quantitative PCR on BMS genomic DNA exactly as described in Chapter 2. The same procedure was followed to analyse the effect of pRep and pRepRb⁻ on replication of pKomRb⁻ in BMS.

The ability of pKom602, pKomRb⁻ and pKomRb^{-C(601)A} to replicate in BMS was analysed by bombarding equal amounts of each plasmid separately into BMS and determining the amount of replicated viral DNA present in each bombarded sample by quantitative PCR on BMS total DNA as described in Chapter 2. In addition, total DNA was extracted from BMS bombarded with pKomRb⁻ 10 days after bombardment, and the Rep gene was amplified from the replicated viral DNA in two separate PCR reactions using two different sets of primers (C1 and C2, and TYR and NTP primers). PCR-amplified Rep fragments were cloned into the pGEM®-T Easy vector (Promega, WI), and sequenced in the forward and reverse directions to determine if any mutations or reversions had occurred in the mutated RBR protein interaction domain during viral replication in BMS.

4.2.3 Agroinoculation

pBIKom602, pBIKomΔI, and pBIKomRb⁻ΔI were introduced into maize plants by agroinoculation, as described in Chapter 3 (3.2.2.1). A group of 14 three-day old maize seedlings was agroinoculated with each construct, and a group of 14 control plants was injected with sterile distilled water. Agroinfections were performed by D. Martin, UCT.

Four weeks after agroinoculation, leaf samples were taken from all agroinoculated plants, total DNA extracted and subjected to PCR with the C1 and C2 primers. These served to simultaneously detect the presence of viral DNA in the leaf samples and to determine whether any detected viral DNA contained the intronless Rep gene (by difference in size between the wt and intronless Rep genes).

4.3 RESULTS AND DISCUSSION

In order to determine whether the Rep gene C(601)A reversion described in Chapter 3 has an effect on viral replication alone, without having to take into account functions required for viral infectivity in maize such as movement of the genome, a series of transient replication assays was undertaken involving *trans*-replication of MSV-Kom and MSV-KomRb⁻ (Fig.4.1), and *trans*-replication of a Rep-deficient MSV-Kom genome (pKep177; Fig 4.4). In addition, in order to study more thoroughly the effect of the Rb⁻ mutation on viral replication in the absence of wt Rep, the replication of MSV-Kom, MSV-KomRb⁻, and MSV-KomRb^{-C(601)A} was assayed in BMS (Fig. 4.3) without any highly expressed Rep constructs provided in *trans*.

4.3.1 The Effect of Rep Variants on MSV-Kom and MSV-KomRb⁻ Replication in BMS

In Fig. 4.1 it can be seen that while Rep and RepRb⁻ *trans*-replicate MSV-Kom to similar levels (as found in Chapter 2), RepRb^{-C(601)A} is nearly twice as efficient at replicating the viral genome. This result was unexpected, since although RepRb^{-C(601)A} has an A nucleotide at position 601 which was shown in Chapter 3 to provide a selective advantage over a C(601), RepRb^{-C(601)A}, like RepRb⁻, contains mutated nucleotides at positions 602 and 604, and in maize the infectivity of MSV-KomRb^{-C(601)A} was reduced compared with the wt virus.

Although it appears in Fig 4.2 that RepRb⁻ was more efficient than wt Rep at replicating MSV-KomRb⁻, the increased replication levels were not highly significant. Surprisingly, neither Rep construct enhanced the replication of MSV-KomRb⁻ to a great extent above the replication of MSV-KomRb⁻ alone, especially when compared with the ~30-fold enhancement of MSV-Kom replication by Rep and RepRb⁻ seen in Chapter 2. While the data in Chapter 2 were the result of many bombardments, the data in Fig. 4.2 came from just one bombardment, which may need repeating for greater accuracy.

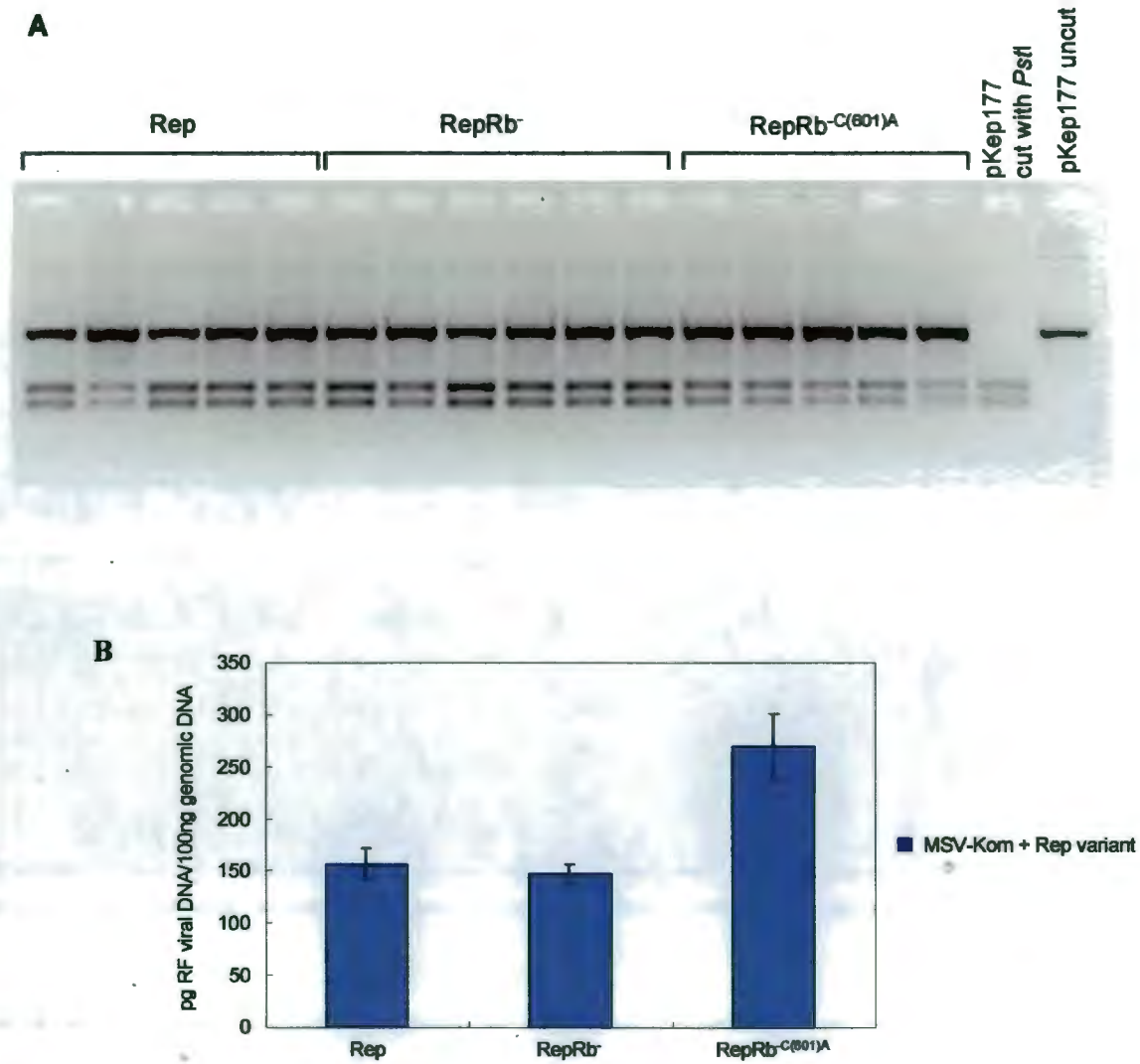


FIGURE 4.1 Replication of MSV-Kom in the presence of Rep variants. In each of two separate bombardment experiments, six plates of BMS were co-bombarded with each Rep construct and MSV-Kom. (A) Quantitative PCR (QPCR) gel from one bombardment experiment. Each PCR reaction was spiked with an internal control (pKep177) of known concentration. After DNA amplification, equal amounts of the QPCR reaction were digested with *PstI* to distinguish the spike from the RF viral DNA bands, and run on a gel. (B) A densitometry programme was used to determine the concentration of viral DNA in each sample, by calculating the ratio of the band intensity of viral DNA to that of the spike DNA. From individual replicate data an average amount in pg of replicated MSV-Kom in the presence of each Rep construct was calculated. Error bars represent 95% confidence intervals. $P = <0.05$.

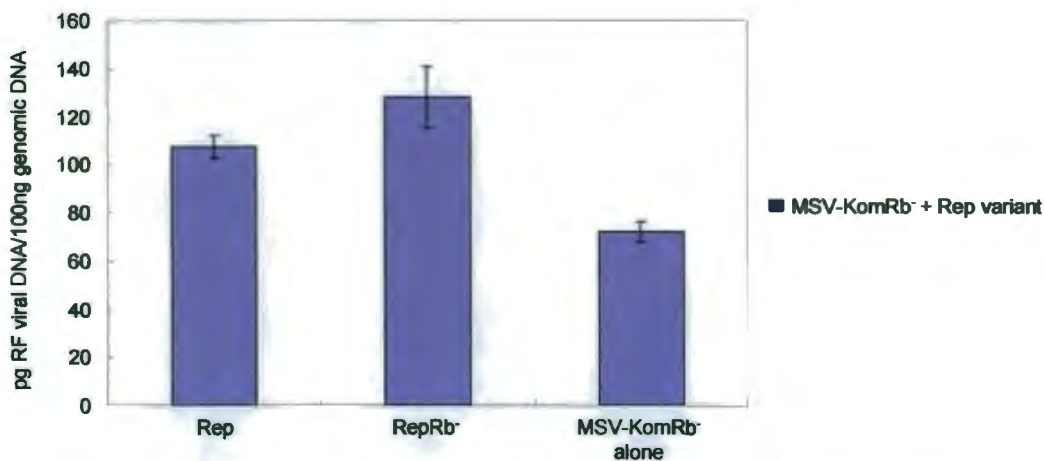


Fig. 4.2 Replication of MSV-KomRb⁻ in the presence of Rep and RepRb⁻. The same QPCR procedure was followed as in Fig. 4.1. Error bars represent 95% confidence intervals. $P = <0.05$

4.3.2 The Effect of Mutations in the Rep RBR Protein Interaction Domain on Viral Replication in BMS

In Fig. 4.3, MSV-KomRb^{-C(601)A} appeared to replicate to the highest levels; however the difference in replication levels of all three viruses was not significant. Hence, the Rb⁻ mutation and C(601)A reversion had no significant effect on viral replication in BMS. This is in contrast to the data in Fig. 4.1. However, in that case RepRb^{-C(601)A} was co-bombarded with MSV-Kom; thus, an advantageous interaction of RepRb^{-C(601)A} or RepARb^{-C(601)A} with wt Rep and/or RepA expressed by MSV-Kom may have resulted in the enhanced replication seen in Fig. 4.1.

MSV-KomRb⁻ was bombarded into BMS and allowed to replicate for 10 days, before genomic DNA was extracted from the samples. The Rep gene was amplified from the replicated viral DNA and sequenced to determine if the C(601)A reversion occurred in the mutated RBR-interaction domain during viral replication in BMS. Unlike in maize agroinfected with MSV-KomRb⁻, no C(601)A reversion was detected in the bombarded viral sample. Although ten days is not a long time span in which to detect a reversion, DNA extracted from BMS samples later than ten days after bombardment no longer contained any detectable viral DNA, probably due to the replication of BMS cells that did not receive the virus during the bombardment diluting out the virus-containing cells. Although in Chapter 3 the C(601)A reversion could be detected in many plants within ten days of agroinoculation, it is likely that the difference between the two systems reflects the ability of the revertant virus to be amplified in plants by movement to

adjacent cells. Since this does not happen in cultured cells individual events such as the C(601)A reversion are less likely to be detected.

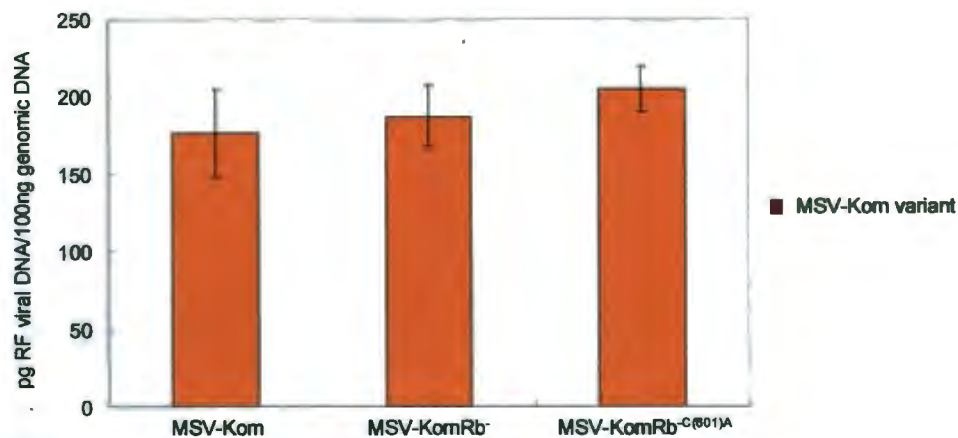


FIGURE 4.3 Replication of MSV-Kom, MSV-KomRb⁻ and MSV-KomRb^{-C(601)A}. In each of two separate bombardment experiments, six plates of BMS were bombarded separately with equal amounts of each virus. Four days were allowed for viral replication, before extraction of genomic DNA from the BMS samples and QPCR. Error bars represent 95% confidence intervals. $P = <0.05$.

4.3.3 The Ability of Rep Variants to *Trans*-Replicate a Rep-Deficient MSV Genome

Apart from Fig. 4.3, the replication assays in this chapter and Chapter 2 involved co-bombardment of highly expressed Rep variants with wt MSV-Kom. When considering all the results from these assays, it has to be taken into account that expression of wt Rep from MSV-Kom could influence the perceived effects of the Rep variants on replication. For example, although RepARb⁻ expressed from pRepRb⁻ could not interact with RBR protein, that function could still be performed by the wt RepA expressed by MSV-Kom. In order to determine the effect of the Rep gene Rb⁻ mutation and C(601)A reversion on viral replication in the absence of wt Rep, each Rep construct was co-bombarded with pKep177, which cannot express Rep or RepA. The *trans*-replication of pKep177 by wt, Rb⁻ and C(601)A revertant Reps is shown in Fig. 4.4. Much like the results in Figs. 4.1 and 4.3, the replication efficiency of RepRb^{-C(601)A} was higher than that of wt and Rb⁻ Reps. However, the increased replication level of pKep177 achieved by RepRb^{-C(601)A} was only just significant compared with RepRb⁻, and not significant compared with wt Rep.

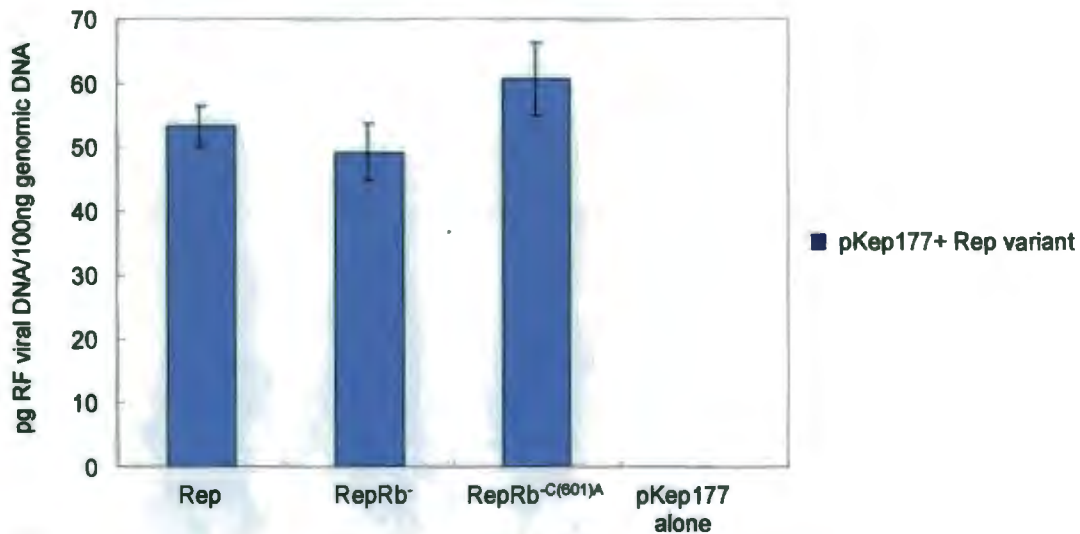


FIGURE 4.4 The *trans*-replication of a Rep-deficient MSV-Kom by Rep variants. Six plates of BMS were co-bombarded with each Rep construct and pKep177. Since pKep177 could not be used as an internal control for QPCR, each PCR run included a reaction sample containing 200 pg of plasmid pKep177. Band intensity of each RF viral band of unknown concentration was converted to pg of RF viral DNA/100ng of BMS genomic DNA by determining the ratio of the band intensity of the 200 pg plasmid control with that of each viral sample. Error bars represent 95% confidence intervals. $P = <0.05$.

4.3.4 The Replication Efficiency of Rep Variants in the Presence and Absence of RepA

In each of the replication assays presented in this chapter and Chapter 2, bombardment of each Rep construct would result in expression of Rep and RepA. In order to separate the functions of the Rep variants, intronless *Rep* and *RepA* constructs were made, as described in Chapter 2. Bombardment of intronless Rep genes with pKep177 should mimic the replication efficiency resulting from a *Rep* transcript that is 100% spliced. Indirectly, this could be used to infer the effect of increased or decreased splicing efficiency on viral replication in a natural infection.

In both Fig. 4.5 and 4.6 a general trend can be seen in which the intronless Reps replicated pKep177 to lower levels than did the wt Reps, and the addition of equal amounts of RepA reduced the replication efficiency even further. However, the lowered replication of pKep177 by RepRb Δ I (when compared with RepRb) was not as significant as the decrease in replication efficiency of Rep Δ I compared with Rep.

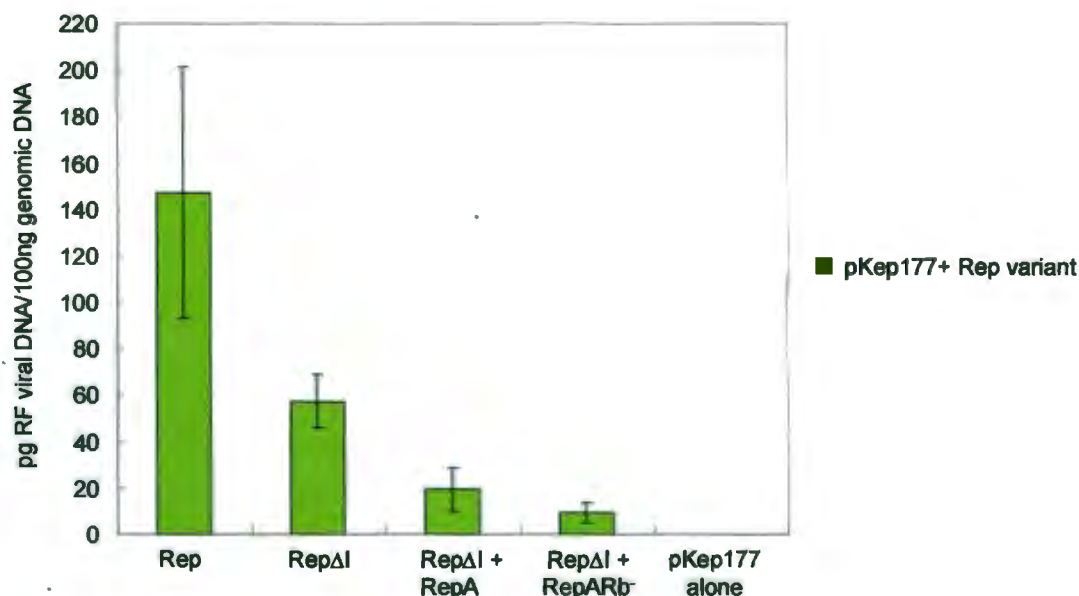


FIGURE 4.5 The replication efficiency of Rep in the presence and absence of RepA variants. In each of two separate bombardment experiments, six plates of BMS were co-bombarded with equal amounts of each Rep construct and pKep177. Error bars represent 95% confidence intervals. $P = <0.05$.

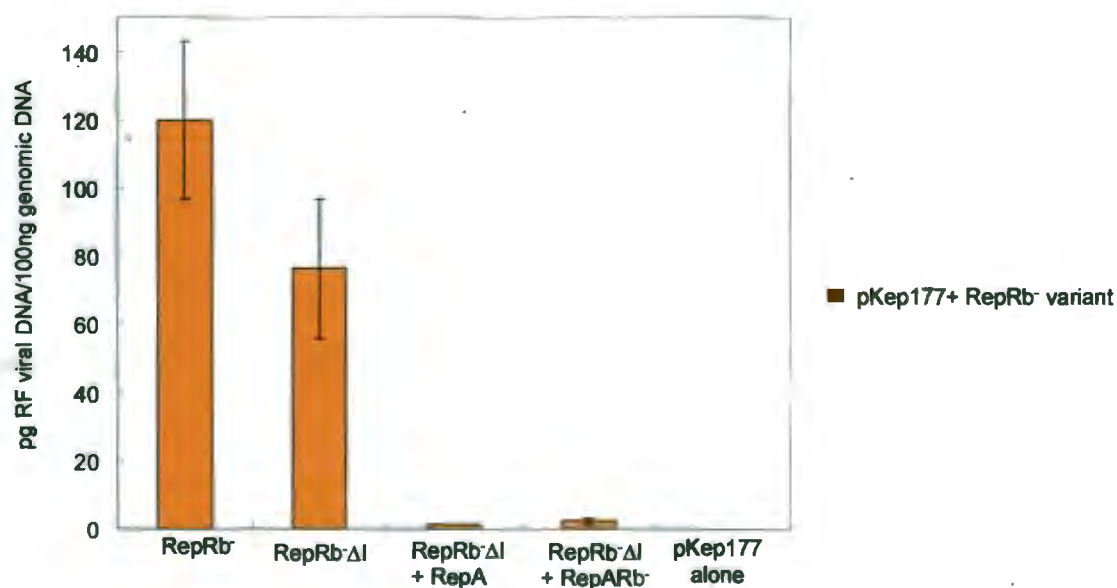


FIGURE 4.6 The replication efficiency of RepRb in the presence and absence of RepA variants. In each of two separate bombardment experiments, six plates of BMS were co-bombarded with equal amounts each Rep construct and pKep177. Error bars represent 95% confidence intervals. $P = <0.05$.

The poorer replication efficiency of the intronless Rep is contrary to the data of Collin *et al.* (1996) and Liu *et al.* (1998), who reported that Rep mutants unable to express RepA replicated more efficiently than the wt genomes. However, in both these reports the replication ability of replicons was determined rather than the *trans*-replication of a Rep-deficient genome by highly expressed intronless Reps as in Figs. 4.5 and 4.6. The difference in results could also reflect the different viruses under investigation: these were WDV in the case of Collin *et al.* (1996) and the dicot-infecting mastrevirus BeYDV in the case of Liu *et al.* (1998). Contrasting results between these viruses and MSV have already been observed several times in this thesis: for example, the dependency of WDV Rep interaction with RBR protein for efficient replication in wheat suspension cells, with no such dependence observed in MSV, and the wt infectivity of BeYDV mutants unable to bind RBR protein, compared with lowered pathogenicity of MSV mutants unable to bind RBR protein. Each case could represent differences in experimental conditions, or different responses from the three diverse mastreviruses.

In this study, the replication efficiency in BMS of MSV-Kom genomes containing intronless Reps (MSV-Kom Δ I and MSV-KomRb $^{-}\Delta$ I) was not determined. However, the infectivity of these viruses in maize was tested by agroinoculation (4.3.5).

In Fig. 4.5 there was no significant difference in the inhibition of the wt intronless Rep's replication efficiency by RepA or RepARb $^{-}$. Similarly, in Fig. 4.6 there was no significant difference in the inhibition of the replication efficiency of the intronless RepRb $^{-}$ by wt RepA or RepARb $^{-}$. One observable difference was that RepA and RepARb $^{-}$ lowered the replication efficiency of RepRb $^{-}\Delta$ I to a greater extent than they did Rep Δ I. This requires confirmation by repeating the bombardments.

In Chapter 2, RepA and RepARb $^{-}$ expressed at high levels by the maize ubiquitin promoter completely inhibited the replication of MSV-Kom. It is notable that in Fig. 4.5 and 4.6, RepA and RepARb $^{-}$ did not inhibit replication to the same extent when co-bombarded with equal amounts of the intronless Reps. This indicates that RepA and RepARb $^{-}$ behave in a dominant negative manner, in the former case out-competing the wt Rep which was expressed at low levels from the Rep promoter. When expressed at the same level as the intronless Reps, however, the negative effects of RepA and RepARb $^{-}$ were not so dominant. A plausible way in which RepA could compete with Rep would be competition for DNA binding sites in the LIR. WDV Rep and RepA both form DNA-protein complexes of approximately the same size in similar locations in the LIR

(Castellano *et al.*, 1999; Missich *et al.*, 2000); thus it is conceivable that RepA, when expressed at higher levels than in a wt infection, could swamp these binding sites, out-competing the binding of Rep.

It is not known at what ratio Rep:RepA is expressed in a wt situation, but these results indicate that the exact ratio is important for efficient viral replication. Expression of 100% Rep and 0% RepA slightly lowered replication efficiency, while expression of RepA at the same (Fig. 4.5 and 4.6) or a higher (Fig. 2.4) ratio to Rep led to greatly reduced viral replication or complete inhibition of viral replication respectively. Thus, it would appear that optimal replication, at least in the case of MSV-Kom, requires the presence of RepA at a lower ratio than Rep. In a wt infection these ratios could be controlled at the level of splicing, as well as potentially at the level of transcription. For example, in MSV two *Rep* transcripts are produced, one terminating in the C2 that can only express RepA, and one terminating in the SIR that is capable of expressing both Rep and RepA (M. Boulton, pers. comm.). By bombarding a range of ratios of Rep:RepA with pKep177, it may be possible to estimate the optimum ratio for viral replication. Since Rep and RepA form hetero-oligomers, possibly consisting of ~ eight monomers (Castellano *et al.*, 1999; Missich *et al.*, 2000), the optimum number of RepA molecules in a hetero-oligomer could possibly be extrapolated. However, in reality it would not be so simple, since the MSV replication cycle is likely to require the involvement of Rep-Rep and RepA-RepA homo- and Rep-RepA hetero- oligomers. The aggregation state of Rep and RepA possibly plays an important role in each protein's functions, and it is likely that upsetting this balance resulted in the lowered replication efficiency seen in Figs. 4.5 and 4.6 compared with wt expression of Rep and RepA.

4.3.5 The Infectivity of MSV-Kom and MSV-KomRb⁻ in the Absence of RepA

Agroinfectious pKom Δ I and pKomRb⁻ Δ I constructs were agroinoculated into 14 three-day old maize seedlings, monitored daily for symptom development, and 4 weeks after agroinoculation tested for the presence of viral DNA by PCR of agroinoculated maize leaf samples. Clear symptoms were not observed on any of the plants, while plants agroinoculated with MSV-Kom at the same time established severe infections in which 100% of the plants developed symptoms). Fig. 4.7 shows that bands of varying sizes were PCR-amplified from some of the test maize samples, while a sample from a control plant inoculated with water contained no PCR amplification products at all. The C1 and C2 primers (which amplify the entire Rep gene) were

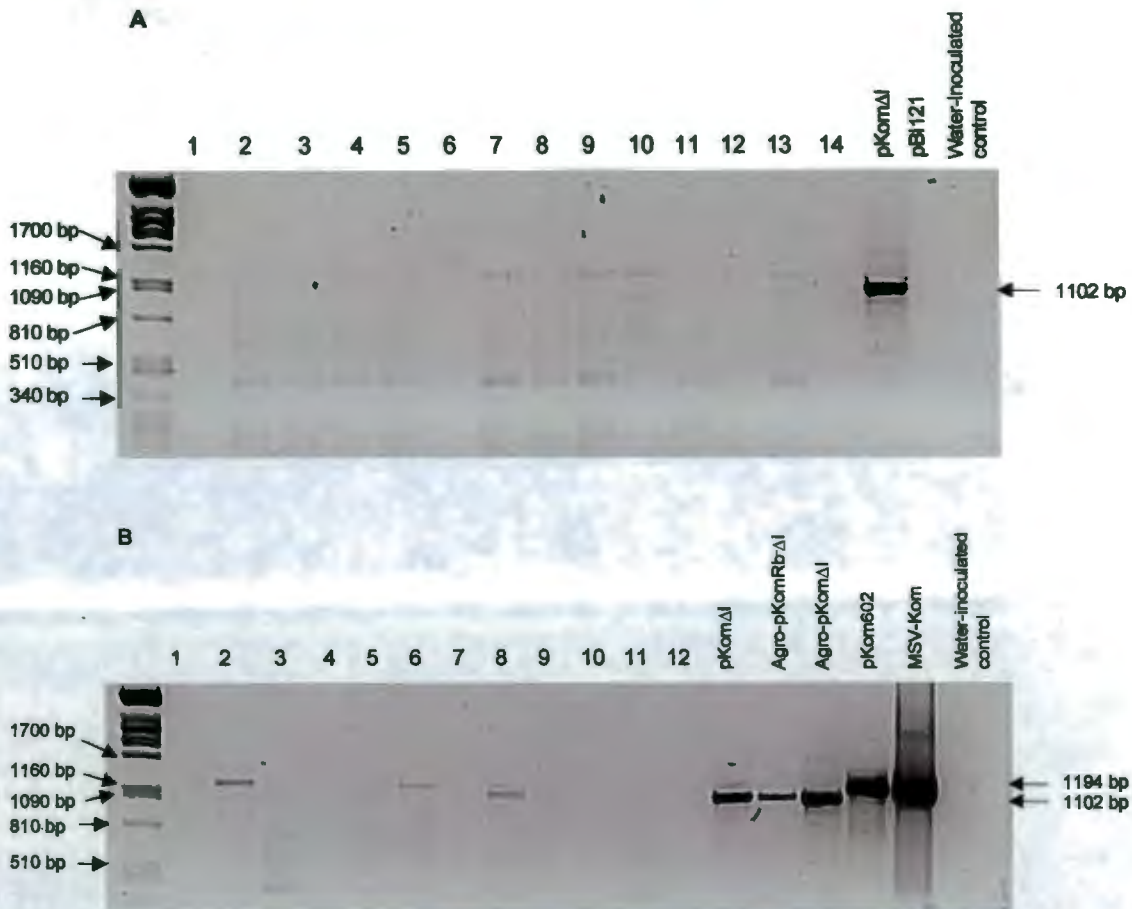


FIGURE 4.7 PCR amplification of viral fragments from agroinoculated maize using C1 and C2 primers, which amplify the entire Rep gene. (A) 1-14 = Plants agroinoculated with MSV-KomΔI. pKomΔI = plasmid used to transform *A. tumefaciens*, included as a size control for the intronless Rep gene (1102 bp). pBI121 = vector that the viral genomes were cloned into, included as a negative control. The water inoculated control was one of 14 plants injected with water in place of the virus constructs and grown under the same conditions. (B) 1-12 = Plants agroinoculated with MSV-KomRbΔI. pKomΔI = plasmid used to transform *A. tumefaciens*, included as a size control for the intronless Rep gene (1102 bp). Agro-pKomRbΔI and Agro-pKomΔI = plasmids extracted from the *A. tumefaciens* culture that was used to agroinoculate maize seedlings. pKom602 = size control for the full-length Rep gene (1194 bp). MSV-Kom = sample extracted from maize agroinoculated with wt MSV-Kom. The size marker in lane 1 of each gel is λ digested with *Pst*I.

used with the aim of differentiating between the intronless Rep (1102 bp) and full-length Rep (1194 bp) genes. While a low volume of the MSV-Kom positive control in Fig. 4.7 (B) was loaded in order to distinguish a size difference between MSV-Kom Rep and MSV-KomΔI Rep, the difference is not clear due to the high band intensity of MSV-Kom Rep. However, the Rep fragment amplified from the plasmid control (pKom602) is clearly slightly larger than the intronless Rep fragments amplified from pKomΔI controls. Using the plasmid amplification

products as size markers, it can be seen in Fig. 4.7 that some PCR products from samples agroinoculated with pKom Δ I (A) and pKomRb Δ I (B) are approximately the size of the full-length Rep gene. One sample (B8) contains a band of the correct size for the intronless Rep gene, as well as a faint band the size of the full-length Rep gene. It is possible that these are amplification products of the inoculated *A. tumefaciens* carrying the pKom Δ I and pKomRb Δ I plasmids, which may have persisted in the plant. If, however, they are the products of a very low-level infection by MSV-Kom Δ I and MSV-KomRb Δ I, two interesting observations may be made. First, previous reports have suggested that BeYDV (Liu *et al.*, 1998) and MSV (Boulton, 2002) mutants unable to express RepA cannot infect plants. Second, while the Rep genes amplified from the pKom Δ I and pKomRb Δ I plasmids extracted from the *A. tumefaciens* cultures that were inoculated into maize are the correct size for the intronless Rep gene, many of the bands amplified from the maize samples are approximately the same size as the full-length Rep gene. This may indicate that in order to establish an infection, albeit an inefficient one, the viruses incorporated "stuffer" fragments into the Rep gene, either by recombination or taken from the maize genome, to restore the gene to wt size. This could be confirmed by sequencing the PCR products. Confirmation that the PCR-amplified viral bands are not from the inoculated *A. tumefaciens* but rather from replicative release and subsequent infection of MSV-Kom Δ I and MSV-KomRb Δ I could come from subjecting the maize samples to PCR using the degenerate primers (DEG1 and DEG2) described in Chapter 2, which can only amplify the MSV-Kom genome once it has been replicatively released from the plasmid and replicated. Unfortunately, due to time constraints these experiments could not be done as part of this thesis.

4.4 CONCLUSION

In previous replication assays and those in this chapter, the three-nt Rb- mutation in RepRb- did not affect its replication ability, whether trans-replicating wt MSV-Kom or a Rep-deficient MSV-Kom genome, or replicating its own genome while under the control of the Rep promoter in MSV-KomRb-. However, a reversion of one of the three mutated nucleotides in the RBR protein interaction domain, of C(601)A apparently, and surprisingly, increased the replication efficiency above wt levels. However, the increase was only slight and in some cases not significant. Confirmation of this result requires repeated experiments of the same nature.

The results from the *trans*-replication of pKep177 by Rep Δ I and by combinations of Rep Δ I+RepA or RepA Rb- indicate that the ratio of Rep:RepA is important for optimal replication.

The fact that *Rep* transcripts are relatively rare in infected cells (Wright *et al.*, 1997) implies that *Rep* expression is tightly controlled. It appears that even though the *Rep* gene copy number increases exponentially during replication of the genome, *Rep* expression is kept fairly constant. Thus, if one of the roles of *RepA* is to downregulate expression of *Rep* as part of the control mechanism, as inferred by Collin *et al.* (1996), it is easy to see why altering the levels of *RepA* upsets the seemingly delicate balance. For example, preventing the expression of *RepA* may remove the control on *Rep* expression, while over-expressing *RepA* may down-regulate *Rep* expression too much, or may inhibit the functions of *Rep* in a dominant negative manner.

The lower replication efficiency resulting from the inhibition of *RepA* expression could also reflect a possible role of *RepA* in the transactivation of host gene promoters required for viral replication. Although this has not been experimentally proven, Horvath *et al.* (1998) found that both *Rep* and *RepA* can transactivate reporter genes in yeast. However, a deletion of the *Rep* C-terminal 89 amino acids was required for transcription activation function, implying that the activation domain is masked by protein folding. It is possible that interaction of *Rep* with *RepA* could result in a conformational change, exposing the activation domain in *Rep*. This would also explain the requirement of both *Rep* and *RepA* to activate the *CP* promoter despite each protein having their own functional activation domain. Even if *Rep* and *RepA* do not directly activate the expression of host genes, it is possible that *RepA* plays a similar role to the begomoviral TrAP, which recruits components of the host transcriptional machinery to the promoter requiring activation (Hartitz *et al.*, 1999). Abolishing the putative function of *RepA* in activating host S-phase gene promoters may not have a drastic effect in BMS cells, in which S-phase specific genes are likely to be expressed by the host's own machinery, but in differentiated cells the direct activation of S-phase gene promoters (without *RepA* having to interact with RBR protein) could provide yet another reason for the requirement of *RepA* in establishing an efficient infection.

The abolition of *RepA* expression certainly had a drastic effect on viral infectivity in maize. While no definite symptoms were observed on plants agroinoculated with pKom Δ I or pKomRb Δ I, in some plants faint speckles or tiny spots could be seen, which in some cases correlated with the presence of *Rep*-specific bands when samples were amplified by PCR. If it is confirmed that these amplified fragments did indeed come from a very weak infection of MSV-Kom Δ I or pKomRb Δ I, they could provide further insight into the role of *RepA* in a viral infection. It would be especially interesting to determine if the amplified fragments the size of full-length *Rep* are intronless *Rep* genes that have been converted to wt size by the addition of stuffer DNA. If this is

the case, it would provide some insight into the replication requirements of Rep, as well as possibly being indicative of an exact size requirement in order for the genome to be packaged

Chapter 5

General Conclusion

Although maize streak disease (MSD) was first described over a hundred years ago, it is only in the past 15 years that several factors have combined to make MSV, the causal agent of MSD, the most significant pathogen of maize in Africa. These include the increased area of maize cultivation and the introduction of high-yielding, MSV-sensitive genotypes, both of which contribute to increased *Cicadulina* and MSV occurrences, as well as the possible emergence of more severe MSV strains by recombination and mutation of the genome. Although there are many factors that contribute to the complex epidemiology of MSD, it is generally believed that the generation and widespread use of high yielding MSV-resistant maize genotypes is a viable solution to the disease. However, one of the major constraints faced by breeders is that MSV resistance genes are often tightly linked with genes for resistance to yield. In addition, the nature of natural virus resistance in plants, where one or more viral genes encode an avirulence factor that elicits resistance controlled by a cognate dominant host gene, is such that it could easily be circumvented by virus variation. As seen in this thesis, MSV potentially mutates at a rapid rate, which could translate into the rapid evolution of more virulent, resistance-breaking strains.

In theory, the introduction of MSV resistance into maize by genetic engineering provides two main advantages over classical resistance breeding: these are the ability to transfer genes directly without linkage to undesired genes, and the ability to construct novel genes that are unlikely to exist in nature. Resistance strategies using genetic engineering should be easier to design and harder to break than natural resistance, and more likely to succeed against a broad range of virus strains. Some of the above criteria were met in this project. A resistance strategy using dominant negative mutant and truncated Rep constructs was designed and tested in transient and transgenic assays. The MSV resistance achieved in a model host, *D. sanguinalis*, was effective against a viral strain that is closely related to the most severe group of MSV genotypes ever isolated, which is also the most widely distributed and one of the most divergent groups in Africa.

However, far from the theoretical advantage of introducing the resistance phenotype without any undesired traits, the transformation of plants with the Rep gene, even when highly mutated, negatively affected aspects of plant growth and development, and most transgenic plants were

sterile. This was overcome by the transformation of *D. sanguinalis* with a mutated/truncated Rep gene, illustrating the ease with which genetic engineering strategies can be changed and optimised. Most plants transgenic for the mutated/truncated gene were phenotypically normal and fertile, and showed good resistance to MSV. This is the first report of the successful development of MSV-resistant plants by genetic engineering. However, the major test will be whether the strategy works in maize. Fertile Hi-II T₁ maize plants transgenic for the mutated/truncated Rep gene have been produced, the offspring of which will be tested for resistance to MSV. If these too are resistant, which is likely considering the success in producing MSV-resistant *D. sanguinalis*, a significant breakthrough in the fight against MSD will have been achieved.

Apart from being a useful guide for the direction to take in developing MSV-resistant plant, the transient replication assays provided some valuable insights into the factors influencing MSV replication. The effect of a three-nucleotide mutation in the Rep RBR protein interaction domain was thoroughly investigated, and it was concluded that the mutation had no effect on Rep or RepA in their role of initiating viral replication in black Mexican sweetcorn (BMS) suspension cells. However, the pathogenicity of a virus carrying the Rb⁻ mutation (MSV-KomRb⁻) was reduced by approximately 50% when compared with the wild type, indicating that the mutation had a negative effect on viral infectivity, possibly because of the inability of the mutant virus' Rep to interact with the host RBR protein.

The high frequency occurrence of a reversion of C(601)A in the mutated RBR protein interaction motif, which did not restore the protein's ability to interact with RBR protein, indicated that the three nucleotide mutation had effects separate from RBR protein interaction. While the reversion consistently occurred when MSV-KomRb⁻ was inoculated into maize, whether by agroinoculation or leafhopper transmission, it did not occur when the mutant virus was bombarded into BMS. This indicates that an A at position 601 in the Rep gene is required for infectivity in maize for an aspect separate from simple replication efficiency, such as movement of the viral genome. It is possible that the original A(601)C mutation increased splicing efficiency of the Rep transcript, and that a reversion of C(601)A was required to decrease splicing efficiency to optimum levels for expression of RepA, one of whose important roles in viral infectivity is thought to be to combine with Rep to transactivate the CP promoter. In turn, CP along with MP is required for movement of the viral genome in order for it to establish a systemic infection.

Replication assays determining the effect of removal of the *Rep* intron in a small way corroborated this theory. If the C(601) mutation in *RepRb*⁻ did indeed increase the splicing efficiency of the *Rep* transcript resulting in a reduction in the expression of RepA, one would expect that there would be less of a difference in the replication ability of the full-length *RepRb*⁻ and the *RepRb*⁻ΔI constructs. As seen in Chapter 4, this was indeed the case. While the replication ability of the wt Rep was significantly reduced by the removal of the intron, the lower replication levels achieved by *RepRb*⁻ΔI were not significant when compared with *RepRb*⁻. If the Rb⁻ mutation did increase splicing efficiency of the *Rep* transcript, it is unlikely that the efficiency would be increased to 100%, explaining the slight difference between *RepRb*⁻ and *RepRb*⁻ΔI in terms of their effect on viral replication.

Because Rep and RepA have different activities in mastreviruses it has been suggested that mechanisms altering the relative proportions in which they are expressed may control the progression of the infection cycle from replication initiation through to movement and encapsidation. The results in Chapter 4 indicate that a precise balance of expression of Rep and RepA is required, both for optimal viral replication and infectivity. It is likely that RepA contributes to the down-regulation of Rep expression, since the expression of RepA at a similar ratio to Rep reduced the replication efficiency of the latter approximately 3-fold. When expressed at a higher ratio to Rep, as was the case when RepA under the control of the ubiquitin promoter was co-bombarded with MSV-Kom, RepA completely inhibited the replication functions of Rep. However, if expressed at the correct ratio, it is likely that the regulation of Rep expression by RepA leads to optimal viral replication.

The effect of abolishing RepA expression was more profound on viral infectivity than on replication. Even if Rep-specific amplification products from PCR of maize samples agroinoculated with MSV-KomΔI and MSV-KomRb⁻ΔI, which cannot express RepA, are confirmed to be the products of a low-level viral infection, it can still be concluded that to establish an efficient infection, RepA expression is an essential component of the viral life cycle.

Finally, potential insight into the role of Rep and RepA in viral replication was gained from the transient replication studies in Chapter 2. A highly truncated 179-aa Rep protein and a Rep protein of the same size containing a mutation in motif III enhanced the replication of MSV-Kom approximately 4-fold. Conversely, two 219-aa Rep constructs, one containing a mutated RBR protein interaction motif, inhibited replication of MSV-Kom. A major difference between the

219-aa and 179-aa proteins is the presence of the entire oligomerization domain in the former, while it is deleted in the latter. Previous studies have found that expression of the Rep oligomerization domain interferes with viral replication, by interfering with the DNA binding and oligomerization activities of Rep (Chatterji *et al.*, 2001). Thus, formation of hetero-oligomers containing wt Rep and the 219-aa truncated Reps could sequester the wt Rep oligomers that otherwise would participate in initiation of viral replication.

The enhancement of viral replication by the 179-aa Reps could be due to the presence of the putative DNA-binding domain coupled with the absence of the oligomerization domain. It has been shown that the formation of Rep oligomers occurs in a stepwise manner, the first step being the binding of a Rep monomer with DNA, mediated by the protein's DNA binding domain, and the second step, requiring the protein's oligomerization domain, being the sequential addition of Rep monomers (Missich *et al.*, 2000). It has been suggested that the formation of a large oligomeric C complex near the Rep TATA box in the LIR could inhibit Rep transcription by interfering with the assembly or activity of the transcription pre-initiation complex (Castellano *et al.*, 1999). Thus, it is conceivable that the highly expressed 179-aa Rep protein could recognize the sequence-specific DNA binding site and interact with viral DNA in the vicinity of the C complex, preventing the binding of further Rep monomers at the same site. Unable to interact with other Rep monomers, the highly truncated Rep protein, due to its small size, does not interfere with the assembly of the transcription pre-initiation complex, removing the block on Rep transcription.

Clearly, in an attempt to explain some puzzling but interesting results involving Rep and RepA, there are several theories in this thesis that need to be tested experimentally. Whatever the outcome, future experimental results are sure to provide some unexpected insights into the multiple functions of Rep and RepA in viral replication and its lifecycle in general.

Appendix A

Summary of Plasmid and Construct Properties

Contents

TABLE A.1 Properties of Constructs used in Chapters 2 to 4	163
---	-----

TABLE A.1 Properties of Constructs used in Chapters 2 to 4.

Construct	Chapter Introduced	Properties	Reference
<i>Backbone constructs</i>			
pKom602	2	One full MSV-Kom genome bounded by two full long intergenic regions (1.1 mer); cloned in pUC19	Schrippenkoetter <i>et al.</i> (2001)
pSKRep	2	Entire wt Rep gene cloned in pSK	This thesis
pGEMTRep	2	Entire wt Rep gene cloned in front of the T7 promoter in pGEM®-T Easy Vector (Promega, WI)	This thesis
<i>Mutated/truncated Rep genes</i>			
pSKRep ^{III+Rb-NTP+/} pSKRepRb ⁻	2	Rep gene containing an Rb ⁻ mutation (LLCNE to LLCLK) cloned in pSK.	T. Mangwende (2001)
pSKRep ^{III-Rb+NTP+}	2	Rep gene containing motif III (III ⁻) mutation (VRDYILKE to VRAYFLKE) cloned in pSK	T. Mangwende (2001)
pSKRep ^{III-Rb-NTP+}	2	Rep gene containing III ⁻ and Rb ⁻ mutations cloned in pSK	T. Mangwende (2001)
pSKRep ^{ΔRbΔC2}	2	537-bp N-terminal Rep fragment with a deletion of the RBR protein binding domain and the entire C2; cloned in pSK	This thesis
pSKRep ^{III-ΔRbΔC2}	2	537-bp N-terminal Rep fragment containing the III ⁻ mutation, with a deletion of the RBR protein binding domain and the entire C2; cloned in pSK	This thesis
pSKRep ^{ΔC2}	2	658-bp N-terminal Rep fragment with a deletion of the RepA C ⁻ terminus and the entire C2; cloned in pSK	This thesis
pSKRep ^{Rb-ΔC2}	2	658-bp N-terminal Rep fragment containing the Rb ⁻ mutation, with a deletion of the RepA C ⁻ terminus and the entire C2; cloned in pSK	This thesis
pSKRep ^{ΔI}	2	Intronless Rep gene cloned in pSK	This thesis
pSKRep ^{Rb-ΔI}	2	Intronless Rep gene containing the Rb ⁻ mutation; cloned in pSK	This thesis
pKomMB1657	2	pKom602 containing a 3' splice site mutation of A ⁷³³ G ⁷³⁴ to T ⁷³³ C ⁷³⁴ in the RepA ORF	This thesis
pSKRepA	2	RepA gene containing the 3' splice site mutation; cloned in pSK	This thesis
pSKRepA ^{Rb-}	2	RepA gene containing the 3' splice site mutation and the Rb ⁻ mutation; cloned in pSK	This thesis

TABLE A.1 cont.

Construct	Chapter Introduced	Properties	Reference
<i>Mutated/truncated Rep genes</i>			
pSKRepRb ^{-C(601)G}	3	Rb ⁻ mutant Rep gene with C(601) changed to G(601), but with the same T(602) and A(604) mutations as in pSKRep ^{Rb-} ; cloned in pSK	This thesis
pSKRepRb ^{-C(601)T}	3	Rb ⁻ mutant Rep gene with C(601) changed to T(601), but with the same T(602) and A(604) mutations as in pSKRep ^{Rb-} ; cloned in pSK	This thesis
pSKRepRb ^{-C(601)A}	3	Rb ⁻ mutant Rep gene with a naturally occurring reversion of C(601) to A(601), but with the same T(602) and A(604) mutations as in pSKRep ^{Rb-} ; cloned in pSK	This thesis
pSKRepARb ^{-C(601)A}	3	Rb ⁻ mutant RepA gene containing the 3' splice site mutation and a reversion of C(601) to A(601), but with the same T(602) and A(604) mutations as in pSKRepA ^{Rb-} ; cloned in pSK	This thesis
pKep177	2	Dimer of the MSV-Kom genome cloned in pUC19, with a premature stop codon introduced into the Rep gene by the insertion of a <i>Pst</i> I site 72 bp downstream from the <i>C1</i> start codon, rendering the construct replication deficient.	Referred to as pMSV- <i>Pst</i> I in Palmer and Rybicki (2001)
<i>Plant expression vectors</i>			
pAHC17	2	Expression vector containing the maize ubiquitin promoter and the NOS terminator	Christensen and Quail (1996)
pRep	2	Full-length Rep gene cloned in pAHC17	This thesis
pRep ^{III+Rb-NTP+/pRepRb⁻}	2	Rep gene with the Rb ⁻ mutation cloned in pAHC17	T. Mangwende (2001)
pRep ^{III-Rb+NTP+}	2	Rep gene with the III ⁻ mutation cloned in pAHC17	T. Mangwende (2001)
pRep ^{III-Rb-NTP+}	2	Rep gene with III ⁻ and Rb ⁻ mutations cloned in pAHC17	T. Mangwende (2001)
pRep ^{III+Rb+NTP+(AS)}	2	Rep gene cloned into pAHC17 in the antisense orientation	T. Mangwende (2001)
pRep ^{ΔRbΔC2}	2	537-bp N-terminal Rep fragment with a deletion of the RBR protein binding domain and the entire C2; cloned in pAHC17	This thesis
pRep ^{III-ΔRbΔC2}	2	537-bp N-terminal Rep fragment containing the III ⁻ mutation, with a deletion of the RBR protein binding domain and the entire C2; cloned in pAHC17	This thesis

TABLE A.1 cont.

Construct	Chapter Introduced	Properties	Reference
<i>Plant expression vectors</i>			
pRep ^{ΔC2}	2	658-bp N-terminal Rep fragment with a deletion of the RepA C terminus and the entire C2; cloned in pAHC17	This thesis
pRep ^{Rb-ΔC2}	2	658-bp N-terminal Rep fragment containing the Rb ⁻ mutation, with the RepA C terminus and the entire C2 deleted; cloned in pAHC17	This thesis
pRep ^{ΔRbΔC2(AS)}	2	Rep ^{ΔRbΔC2} gene cloned into pAHC17 in the antisense orientation	This thesis
pRep ^{ΔC2(AS)}	2	Rep ^{ΔC2} gene cloned into pAHC17 in the antisense orientation	This thesis
pAHC25	2	Vector containing the <i>bar</i> (bialophos resistance) and <i>uidA</i> (GUS) genes under the control of separate maize ubiquitin promoters	Christensen and Quail (1996)
pAHCRep ^{III-Rb-NTP+}	2	Rep ^{III-Rb-NTP+} gene cloned in place of the <i>uidA</i> gene in pAHC25	This thesis
pRepA	2	RepA gene containing a 3' splice site mutation, cloned into pAHC17	This thesis
pRepA ^{Rb-}	2	RepA gene containing the 3' splice site mutation and the Rb ⁻ mutation, cloned into pAHC17	This thesis
pRep ^{Rb-C(601)A}	4	Rb ⁻ mutant Rep gene containing a reversion of C(601) to A(601), but with the same T(602) and A(604) mutations as in pRepRb ⁻ ; cloned in pAHC17	This thesis
pRep ^{ΔI}	4	Intronless Rep gene cloned in pAHC17	This thesis
pRep ^{Rb-ΔI}	4	Intronless Rep gene containing the Rb ⁻ mutation; cloned in pAHC17	This thesis
<i>Yeast two- hybrid expression vectors</i>			
pGAD424	2	Yeast two-hybrid expression vector containing the GAL4 activation domain	CLONTECH, CA
pBD-GAL4-CAM	2	Yeast two-hybrid expression vector containing the GAL4 binding domain	Stratagene, La Jolla, CA
pADRepA	2	RepA gene containing the 3' splice site mutation, cloned in frame with the GAL4 activation domain in pGAD424	This thesis
pADRepA ^{Rb-}	2	RepA gene containing the 3' splice site mutation and the Rb ⁻ mutation, cloned in frame with the GAL4 activation domain in pGAD424	This thesis

TABLE A.1 cont.

Construct	Chapter Introduced	Properties	Reference
<i>Yeast two- hybrid expression vectors</i>			
pADRepARb ^{-C(601)A}	3	Rb ⁻ mutant RepA gene containing the 3' splice site mutation and a reversion of C(601) to A(601), but with the same T(602) and A(604) mutations as in pRepA ^{Rb-} ; cloned in frame with the GAL4 activation domain in pGAD424	This thesis
pADRep ^{ΔI}	2	Intronless Rep gene cloned in frame with the GAL4 activation domain in pGAD424	This thesis
pADRep ^{Rb-ΔI}	2	Intronless Rep gene containing the Rb ⁻ mutation; cloned in frame with the GAL4 activation domain in pGAD424	This thesis
pADRep ^{ΔC2}	2	658-bp N-terminal Rep fragment with a deletion of the RepA C terminus and the entire C2; cloned in frame with the GAL4 activation domain in pGAD424	This thesis
pADRep ^{Rb-ΔC2}	2	658-bp N-terminal Rep fragment containing the Rb ⁻ mutation, with a deletion of the RepA C terminus and the entire C2; cloned in frame with the GAL4 activation domain in pGAD424	This thesis
pBDRepA	2	RepA gene containing the 3' splice site mutation, cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis
pBDRepA ^{Rb-}	2	RepA gene containing the 3' splice site mutation and the Rb ⁻ mutation, cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis
pBDRepARb ^{-C(601)A}	3	Rb ⁻ mutant RepA gene containing the 3' splice site mutation and a reversion of C(601) to A(601), but with the same T(602) and A(604) mutations as in pRepA ^{Rb-} ; cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis
pBDRep ^{ΔI}	2	Intronless Rep gene cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis
pBDRep ^{Rb-ΔI}	2	Intronless Rep gene containing the Rb ⁻ mutation; cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis
pBDRep ^{ΔC2}	2	658-bp N-terminal Rep fragment with a deletion of the RepA C terminus and the entire C2; cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis
pBDRep ^{Rb-ΔC2}	2	658-bp N-terminal Rep fragment containing the Rb ⁻ mutation, with a deletion of the RepA C terminus and the entire C2; cloned in frame with the GAL4 binding domain in pBD-GAL4-CAM	This thesis

TABLE A.1 cont.

Construct	Chapter Introduced	Properties	Reference
<i>pKom602 mutants</i>			
pKomRb ⁻	3	pKom602 containing the Rb ⁻ mutant Rep gene	This thesis
pKomRb ^{-C(601)G}	3	pKom602 containing the Rb ⁻ mutant Rep gene with the C(601) changed to G(601), but with the same T(602) and A(604) mutations as in pKomRb ⁻	This thesis
pKomRb ^{-C(601)T}	3	pKom602 containing the Rb ⁻ mutant Rep gene with the C(601) changed to T(601), but with the same T(602) and A(604) mutations as in pKomRb ⁻	This thesis
pKomRb ^{-C(601)A}	3	pKom602 containing the Rb ⁻ mutant Rep gene with the C(601) reverted to the wt A(601), but with the same T(602) and A(604) mutations as in pKomRb ⁻	This thesis
pKomΔI	4	pKom602 containing the intronless Rep gene	This thesis
pKomRb ⁻ ΔI	4	pKom602 containing the intronless Rep gene and the Rb ⁻ mutation	This thesis
<i>Agroinfectious constructs</i>			
pBIKom602	3	MSV-Kom 1.1 mer from pKom602 cloned into the binary vector pBI121 (CLONTECH, CA)	This thesis
pBIKomRb ⁻	3	MSV-KomRb ⁻ 1.1 mer from pKomRb ⁻ cloned into pBI121	This thesis
pBIKomRb ^{-C(601)G}	3	MSV-KomRb ^{-C(601)G} 1.1 mer from pKomRb ^{-C(601)G} cloned into pBI121	This thesis
pBIKomRb ^{-C(601)T}	3	MSV-KomRb ^{-C(601)T} 1.1 mer from pKomRb ^{-C(601)T} cloned into pBI121	This thesis
pBIKomRb ^{-C(601)A}	3	MSV-KomRb ^{-C(601)A} 1.1 mer from pKomRb ^{-C(601)A} cloned into pBI121	This thesis
pBIKomΔI	4	MSV-KomΔI 1.1 mer from pKomΔI cloned into pBI121	This thesis
pBIKomRb ⁻ ΔI	4	MSV-KomRb ⁻ ΔI 1.1 mer from pKomRb ⁻ ΔI cloned into pBI121	This thesis

Appendix B

Key for Recognition of Severity of Maize Streak Disease (MSD) Symptoms on Maize Leaves

Contents

FIGURE B.1 Key used for recognition and quantification of maize streak disease (MSD) symptoms	169
--	-----

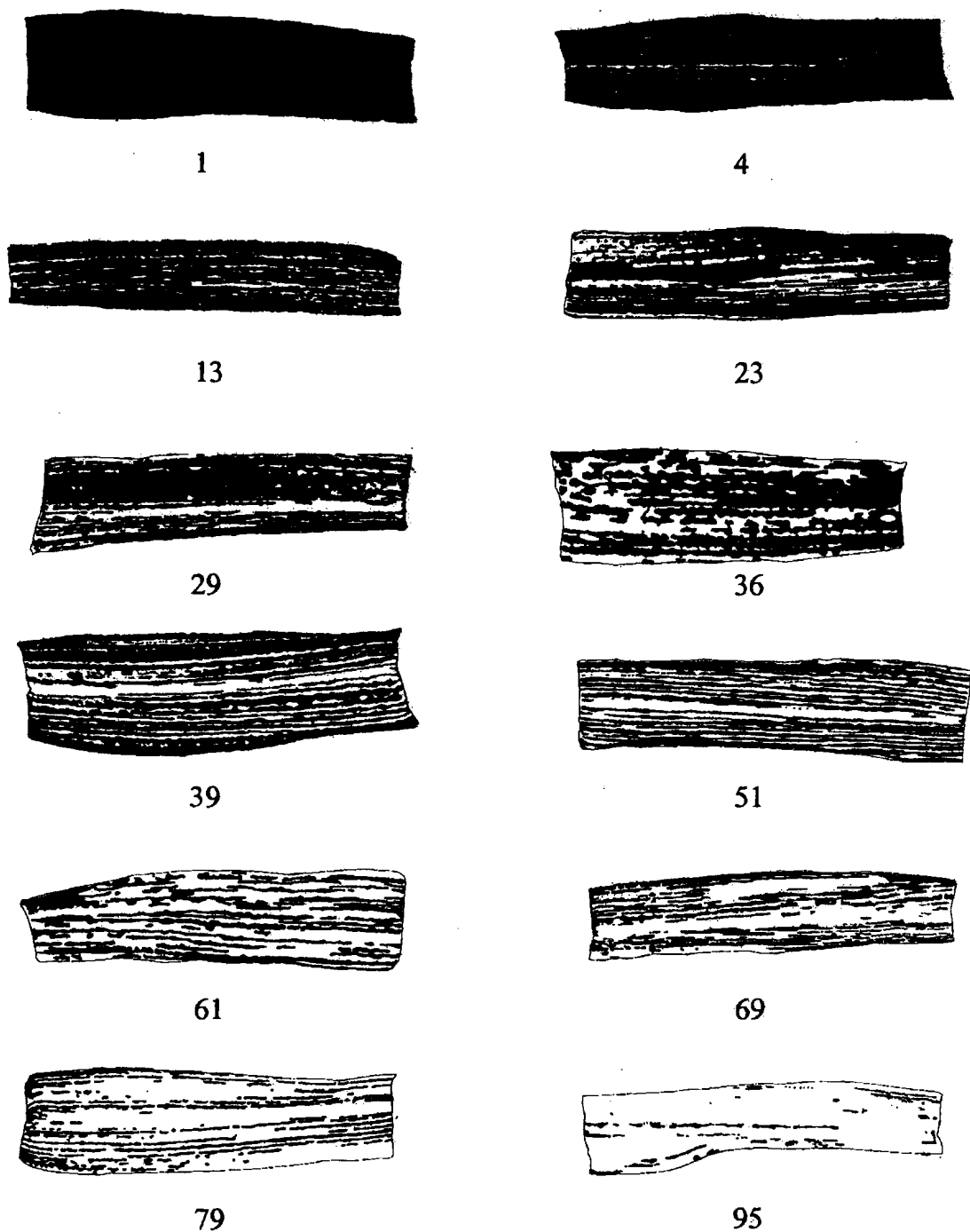


FIGURE B.1 Key used for recognition and quantification of maize streak disease (MSD) symptoms. The Figure shows visual representations of quantified chlorotic areas of maize leaves, ranging from 1% to 95%. Numbers below each leaf represent the percentage of leaf area covered by chlorotic lesions. From D. Martin, unpublished.

References

- Abouzid, A.M., Frischmuth, T., and Jeske, H. (1988). A putative replicative form of the *Abutilon mosaic virus* (gemini group) in a chromatin-like structure. *Mol. Gen. Genet.* **212**, 252-258
- Ach, R.A., Durfee, T., Miller, A.B., Taranto, P.; Hanley-Bowdoin, L., Zambryski, P.C.; and Grissem, W. (1997). RRB1 and RRB2 encode maize retinoblastoma-related proteins that interact with a plant D-type cyclin and geminivirus replication protein. *Mol. Cell. Biol.* **17**, 5077-5086
- Acotto, G.P., Mullineaux, P.M., Brown, P.M., and Marie, D. (1993). Digitaria streak geminivirus replicative forms are abundant in S-phase nuclei of infected cells. *Virology* **195**, 257-259
- Adams, P.D. (2001). Regulation of the retinoblastoma tumor suppressor protein by cyclin/cdks. *Biochimica et Biophysica Acta* **1471** M123-M133
- An, G., Ebert, P.R., Mitra, A., Ha, S.B. (1988). Binary vectors. Pages 87-88. *In: Plant Molecular Biology Manual* (Gelvin S.B. and Schilperoort R.A., eds.). Kluwer Dordrecht, The Netherlands.
- Arguello-Astorga, G.R., Guevara-Gonzalez, R.G., Herrera-Estrella, L.R., and Rivera-Bustamante, R.F. (1994a). Geminivirus replication origins have a group-specific organization of iterative elements: A model for replication. *Virology* **203**, 90-100
- Arguello-Astorga, G., Herrera-Estrella, L., and Rivera-Bustamante, R. (1994b). Experimental and theoretical definition of geminivirus origin of replication. *Plant Mol. Biol.* **26**, 553-556
- Aronson, M.N., Meyer, A.D., Gyorgyey, J., Katul, L., Vetten, H.J., Gronenborn, B., and Timchenko, T. (2000). Clink, a nanovirus-encoded protein, binds both pRB and SKP1. *J. Virol.* **74**, 2967-2972
- Atkinson, R.G., Bielski, L.R.F., Gleave, A.P., Janssen, B-J., and Morris, B.A.M. (1998). Post-transcriptional silencing of chalcone synthase in petunia using a geminivirus-based episomal vector. *Plant J.* **15**, 593-604

- Baas, P.D. (1987). Mutational analysis of the bacteriophage f174 replication origin. *J. Mol. Biol.* **198**, 51-61.
- Baas, P.D., and Jansz, H.S. (1988). Single strand DNA phage origins. *Curr. Top. Microbiol. Immunol.* **136**, 31-70
- Bass, H.W., Nagar, S., Hanley-Bowdoin, L., and Robertson, D. (2000). Chromosome condensation induced by geminivirus infection of mature plants. *J Cell Sci.* **113**, 1149-1160
- Baulcombe, D.C. (1996). Mechanisms of Pathogen-Derived Resistance to Viruses in Transgenic Plants. *Plant Cell.* **8**: 1833-1844
- Bell, K.E., Dale, J.L., Ha, C.V., Vu, M.T., and Revill, P.A. (2002). Characterisation of Rep-encoding components associated with banana bunchy top nanovirus in Vietnam. *Arch Virol.* **147**, 695-707.
- Bendahmane, M. and Gronenborn, B. (1997). Engineering resistance against *Tomato yellow leaf curl virus* (TYLCV) using antisense RNA. *Plant Mol Biol.* **33**, 351-357
- Bisaro, D.M. (1996). Geminivirus DNA replication. In *DNA Replication in Eukaryotic Cells*, pp 833-854. DePamphilis, M.L., Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Bock, K. R., Guthrie, E. J., and Woods, R. D. (1974). Purification of maize streak virus and its relationship to viruses associated with streak diseases of sugarcane and *Panicum maximum*. *Annals of Applied Biology* **77**, 289-296.
- Bosque-Pérez, N.A. (2000). Eight decades of maize streak virus research. *Virus Research* **71**, 107-121
- Boulton, M.I. (2002) Functions and interactions of mastrevirus gene products. *Physiological and Molecular Plant Pathology.* **60**, 243-255.
- Boulton, M.I. and Davies, J.W. (2002) Mastrevirus. In "The Springer Index of Viruses". Christian A. Tidona and Gholamreza Darai, Eds., pp355-361. Springer-Verlag Berlin Heidelberg.

- Boulton, M.I., Steinkellner, H., Donson, J., Markham, P.G., King, D.I., and Davies, J.W. (1989). Mutational analysis of the virion sense genes of maize streak virus. *J. Gen. Virol.* **70**, 2309-2323
- Boyer, S.N., Wazer, D.E., and Band, V. (1996). E7 protein of human papillomavirus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res.* **56**, 4620-4624.
- Bradford, MM. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye-binding. *Anal Biochem* **72**, 248-254
- Briddon, R.W. (2002). Begomovirus. In "The Springer Index of Viruses". Christian A. Tidona and Gholamreza Darai, Eds., pp355-361. Springer-Verlag Berlin Heidelberg.
- Briddon, R. W. and Markham, P. G. (1995). *Geminiviridae*. In "Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses." F. A. Murphy, C. M. Fauquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers, Eds., pp. 158-165. Springer-Verlag, Wien.
- Briddon, R. W., Watts, J., Markham, P. G., and Stanley, J. (1989). The coat protein of beet curly top virus is essential for infectivity. *Virology* **172**, 628-633.
- 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.
- Brunetti, A., Tavazza, M., Noris, E., Tavazza, R., Caciagli, P., Ancora, G., Crespi, S., and Accotto, G.P. (1997). High expression of truncated viral Rep protein confers resistance to *Tomato yellow leaf curl virus* in transgenic tomato plants. *MPMI*. **10**, 571-579
- Brunetti, A., Tavazza, R., Noris, E., Lucioli, A., Accotto, G.P., and Tavazza, M. (2001). Transgenically expressed T-Rep of *Tomato yellow leaf curl Sardinia virus* acts as a trans-dominant negative mutant, inhibiting viral transcription and replication. *J. Virol.* **75**, 10573-10581

- Castellano, M. M., Sanz-Burgos, A. P., and Gutierrez, C. (1999). Initiation of DNA replication in a eukaryotic rolling-circle replicon: identification of multiple DNA-protein complexes at the geminivirus origin. *J Mol Biol.* **290**, 639-652.
- Chatterji, A., Beachy, R.N., and Fauquet, C.M. (2001). Expression of the oligomerization domain of the replication-associated protein (Rep) of *Tomato yellow leaf curl New Delhi virus* interferes with DNA accumulation of heterologous geminiviruses. *J. Biol. Chem.* **276**, 25631-25638
- Chen, W. (1996). "Taxi", A New Vehicle for the Transfer of Genes into Cereals. PhD thesis. UCT: Cape Town, South Africa.
- Choi, I. R. and Stenger, D. C. (1996). The strain-specific cis-acting element of beet curly top geminivirus DNA replication maps to the directly repeated motif of the ori. *Virology* **226**, 122-126.
- Christenson, A.H. and Quail, P.H. (1996). Ubiquitin promoter-based vectors for high level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research* **5**:, 213-218
- Collin, S., Fernandez-Lobato, M., Gooding, P. S., Mullineaux, P. M., and Fenoll, C. (1996). The two nonstructural proteins from wheat dwarf virus involved in viral gene expression and replication are retinoblastoma- binding proteins. *Virology* **219**, 324-329.
- Cooper, C.S., Upmeyer, S.N., and Winokur, P.L (1998) Identification of single amino acids in the human papillomavirus 11 E2 protein critical for the transactivation or replication functions. *Virology* **241**, 312-322
- Covey, S.N, and Al-Kaff, N.S. (2000). Plant DNA viruses and gene silencing. *Plant Mol. Biol.* **43**, 307-322
- Dabrowski, Z.T. (1987). *Cicadulina ghaurii* (Hem., Euscelidae): distribution, biology and maize streak virus (MSV) transmission. *J. Appl. Entomol.* **103**, 489-496
- Dahiya, A., Gavin, M.R., Luo, R.X., and Dean., D. (2000). Role of the LXCXE Binding Site in Rb function. *Mol. Cell. Biol.* **20**, 6799-6805.

- Dahl, M., Meskiene, I., Bögre, L., Ha, D.T., Hubmann, R., Hirt, H., and Herberle-Bors, E. (1995). The D-type alfalfa cyclin gene *cycMs4* complements G1 cyclin-deficient yeast and is induced in the G1 phase of the cell cycle. *Plant Cell*. **7**, 1847-1857.
- Day, A.G., Bejarano, E.R., Buck, K.W., Burrell, M., and Lichtenstein, C.P. (1991). Expression of an antisense viral gene in transgenic tobacco confers resistance to the DNA virus tomato golden mosaic virus. *Proc. Natl. Acad. Sci. USA*. **88**, 6721-6725
- de Jager, S.M. and Murray, J.A.H. (1999). Retinoblastoma protein in plants. *Plant Mol Biol* **41**, 295-299
- Dekker, E.L., Woolson, C.J., Xue, Y., Cox, B., and Müllineaux, P.M. (1991). Transcript mapping reveals different expression strategies for the bicistronic RNAs of the geminivirus wheat dwarf virus. *Nucleic Acids Res.* **19**, 4075-4081.
- Desbiez, C., David, C., Mettouchi, A., Laufs, J., Gronenborn, B. (1995) Rep protein of tomato yellow leaf curl geminivirus has an ATPase activity required for viral DNA replication. *Proc. Natl. Acad. Sci. USA*. **92**, 5640-5644.
- Dickinson, V.J., Halder, J., and Woolston, C.J. (1996). The product of the maize streak virus ORF V1 is associated with secondary plasmodesmata and is first detected with the onset of viral lesions. *Virology*. **220**, 51-59.
- Donson, J., Morris-Krsinich, B. A. M., Mullineaux, P. M., Boulton, M. I., and Davies, J. W. (1984). A putative primer for second-strand DNA synthesis of maize streak virus is virion-associated. *EMBO Journal* **3**, 3069-3073.
- Dowdy, S.F., Hinds, P.W, Louie, K., Reed, S.I., Arnold, A., and Weinberg, R.A. (1993). Physical interaction of the retinoblastoma protein with human D cyclins. *Cell*. **73**, 499-511
- Dry, I. B., Krake, L. R., Rigden, J. E., and Rezaian, M. A. (1997). A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. *Proc Natl Acad.Sci U.S.A.* **94**, 7088-7093.

- Duan, Y.P., Powell, C.A., Webb, S.E., Purcifull, D.E., and Hiebert, E. (1997). Geminivirus resistance in transgenic tobacco expressing mutated BC1 protein. *MPMI*. **10**, 617-623
- Dunder, E., Dawson, J., Suttie, J., and Pace, G. (1995). Maize transformation by microprojectile bombardment of immature embryos, p 126-138. *In* Gene Transfer to Plants. I. Potrykus and G. Spengenberg (eds.), Springer-Verlag, Berlin and Heidelberg.
- Eagle, P. A., Orozco, B. M., and Hanley-Bowdoin, L. (1994). A DNA sequence required for geminivirus replication also mediates transcriptional regulation. *Plant Cell* **6**, 1157-1170.
- Eagle, P.A., and Hanley-Bowdoin, L. (1997). Cis-elements that contribute to geminivirus transcriptional regulation and the efficiency of DNA replication. *J. Virol.* **71**, 6947-6955
- Efron, Y., Kim, S.K., Fajemisin, J.M., Mareck, J.H., Tang, C.Y., Dabrowski, Z.T., Rossel, H.W., Thottappilly, G., and Buddenhagen, I.W. (1989). Breeding for resistance to maize streak virus: A multidisciplinary team approach. *Plant Breed.* **103**, 1-36.
- Egelkrout, E.M., Robertson, D., and Hanley-Bowdoin, L. (2001). Proliferating Cell Nuclear Antigen Transcription Is Repressed through an E2F Consensus Element and Activated by Geminivirus Infection in Mature Leaves. *Plant Cell.* **13**, 1437-1452.
- Ewen, M.E., Sluss, H.K., Sherr, C.J., Matsushime, H., Kato, J., and Livingston, D.M. (1993). Functional interactions of the retinoblastoma protein with mammalian D-type cyclins. *Cell.* **73**, 487-497.
- Fauquet, C.M., Bisaro, D.M., Briddon, R.W., Brown, J., Harrison, B.D., Rybicki, E.P., Stenger, D.C., and Stanley, J. (2003). Revision of taxonomic criteria for species demarcation in the *Geminiviridae* family, and a new updated list of begomovirus species. *Arch Virol* **148**(2), 405-421
- Fennoy, S.L., Surti, G., and Bailey-Serres, J. (1993). Analysis of synonymous codon usage in maize. *In*: Maize Genetic Cooperation Newsletter 67, pp 99-100.

- Fenoll, C., Black, D. M., and Howell, S. H. (1988). The intergenic region of maize streak virus contains promoter elements involved in rightward transcription of the viral genome. *EMBO J.* **7**, 1589-1596.
- Fenoll, C., Schwarz, J. J., Black, D. M., Schneider, M., and Howell, S. H. (1990). The intergenic region of maize streak virus contains a GC-rich element that activates rightward transcription and binds maize nuclear factors. *Plant Mol Biol* **15**, 865-877.
- Fielding, W.L. (1933). Field experimental work on rotation crops. Empire Cotton Growing Association Progress Report 1931-1932. pp10-14
- Fontes, E.P.B., Eagle, P.A., Snipe, P.A., Luckow, V.A., and Hanley-Bowdoin, L. 1994a. Interaction between a geminivirus replication protein and origin DNA is essential for viral replication. *J. Biol. Chem.* **269**, 8459-8465.
- Fontes, E.P.B., Gladfelter, H.J., Schaffer, R.L., Petty, I.T.D., and Hanley-Bowdoin, L. 1994b. Geminivirus replication origins have a modular organization. *Plant Cell* **4**, 597-608.
- Fontes, E.P.B., Luckow, V.A., and Hanley-Bowdoin, L. (1992). A geminivirus replication protein is a sequence specific DNA binding protein. *Plant Cell.* **4**, 597-608
- Frischmuth, T., and Stanley, J. (1991). African cassava mosaic virus DI DNA interferes with the replication of both genomic components. *Virology* **183**, 539-544
- Frischmuth, T., and Stanley, J. (1994). Beet curly top virus symptom amelioration in *Nicotiana benthamiana* transformed with a naturally occurring viral subgenomic DNA. *Virology* **200**, 826-830.
- Frischmuth, T., Engel, M., and Jeske, H. (1997). Beet curly top virus DI DNA-mediated resistance is linked to its size. *Molecular Breeding.* **3**, 213-217
- Fuller, C. (1901). Mealie variegation. In "1st Report of the Government Entomologist, Natal, 1899-1900".

- Gadani, F., Mansky, L.M., Medici, R., Miller, W.A., and Hill, J.H. (1990). Genetic engineering of plants for virus resistance. *Arch. Virol.* **115**, 1-21
- Gietz, R.D, and Woods, R.A. (1994). High efficiency transformation in yeast. *In: Molecular Genetics of Yeast: Practical approaches*, ed J.A. Johnson, Oxford University Press pp. 121-134
- Gladfelter, H. J., Eagle, P. A., Fontes, E. P., Batts, L., and Hanley-Bowdoin, L. (1997). Two domains of the AL1 protein mediate geminivirus origin recognition. *Virology* **239**, 186-197.
- Gooding, P.S., Batty, N.P., Goldsbrough, A.P., and Mullineaux, P.M. (1999). *Nucl. Acids Res.* **27**, 1709-1718.
- Goodman, R.M. (1977a). Infectious DNA from a whitefly transmitted virus of *Phaseolus vulgaris*. *Nature.* **266**, 54.
- Goodman, R.M. (1977b). Single stranded DNA genome in a whitefly-transmitted plant virus. *Virology.* **83**, 171.
- Gorbalenya, A.E. and Koonin, E.V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucl. Acids Res.* **17**, 8413-8440.
- Gordon-Kamm, W., Dilkes, B.P., Lowe, K., Hoerster, G., Sun, X., Ross, M., Church, L., Bunde, C., Farrell, J., Hill, P., Maddock, S., Snyder, J., Sykes, L., Li, Z., Woo, Y., Bidney, D., and Larkins, B.A. (2002) Stimulation of the cell cycle and maize transformation by disruption of the plant retinoblastoma pathway. *PNAS.* **99**, 11975-11980.
- Gorter, G.J.M.A. (1953). Studies on the spread and control of streak disease of maize. *Sci. Bull. Dept. Agric. & For. Union of S. Afr.* **134**: 20
- Grafi, G., Burnett, R.J., Helentjaris, T., Larkins, B.A., DeCaprio, J.A., Sellers, W.R., Kaelin, W.G Jr. (1996). A maize cDNA encoding a member of the retinoblastoma protein family: involvement in endoreduplication. *Proc. Natl. Acad. Sci. USA.* **93**, 8962-8967.
- Grimsley, N., Hohn, T., Davies, J.W., and Hohn, B. (1987). *Agrobacterium*-mediated delivery of infectious maize streak virus in maize plants. *Nature* **325**, 177-179.

Gronenborn, B. (2000). Phytopathogenic geminivirus resistant transgenic plants and seeds and methods for obtaining same by introduction of mutated C1 gene. United States Patent 6133505. Application no. 809103.

Groning, B.R., Abouzid, A., Jeske, H. 1987. Single-stranded DNA from abutilon mosaic virus is present in the plastids of infected *Abutilon sellovianum*. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8996-9000.

Groning, B.R., Hayes, R.J., and Buck, K.W. (1994). Simultaneous regulation of tomato golden mosaic virus coat protein and AL1 gene expression: expression of the AL4 gene may contribute to suppression of the AL1 gene. *J. Gen. Virol.* **75**, 721-726.

Groning, R.R., Frischmuth, T., and Jeske, H. (1990). Replicative form DNA of Abutilon mosaic virus is present in plastids. *Mol. Gen. Genet.* **220**:485-488.

Gruss, A., and Ehrlich, S.D. (1989). The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* **53**, 231-241.

Gutierrez, C. (1998). The retinoblastoma pathway in plant cell cycle and development. *Current opinion in plant biology.* **1**, 492-497.

Gutierrez, C. (2000). DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J.* **19**, 792-799.

Gutierrez, C. (2002) Strategies for geminivirus DNA replication and cell cycle interference. *Physiological and Molecular Plant Pathology.* **60**, 219-230.

Hanley-Bowdoin, L., Settlage, S. B., Orozco, B. M., Nagar, S., and Robertson, B. M. (1999). Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. *Critical Reviews in Plant Sciences* **18**, 71-106.

Hanley-Bowdoin, L. Orozco, B., Kong, L-J., and Gruissem, W.(2002a). Geminivirus resistant transgenic plants. United States Patent Application 20020138867

- Hanley-Bowdoin, L. and Settlage, S. (2002b). Geminivirus resistant transgenic plants. United States Patent Application 20020073442
- Hanson, S.F., and Maxwell, D.P (2000). Trans-dominant inhibition of geminiviral DNA replication by geminivirus rep gene. United States Patent 6118048
- Harris, S.F., and Botchan, M.R. (1999) Crystal structure of the human papillomavirus type 18 E2 activation domain. *Science* **284**, 1673-1677.
- Harrison, B. D., Barker, H., Bock, K. R., Guthrie, E. J., Meredith, G., and Atkinson, M. (1977). Plant viruses with circular single-stranded DNA. *Nature* **270**, 760-762.
- Harrison, B.D. (2002). Virus variation in relation to resistance-breaking in plants. *Euphytica* **124**, 181-192.
- Hartz, M.D., Sunter, G., and Bisaro, D.M. (1999). The *Tomato golden mosaic virus* transactivator (TrAP) is a single-stranded DNA and zinc-binding phosphoprotein with an acidic activation domain. *Virology* **263**, 1-14.
- Hayes, R. J., MacDonald, H., Coutts, R. H., and Buck, K. W. (1988). Priming of complementary DNA synthesis in vitro by small DNA molecules tightly bound to virion DNA of wheat dwarf virus. *J. Gen. Virol.* **69**, 1345-1350.
- Hemenway, C., Tumer, N.E., Powell, P.A. and Beachy, R.N. (1989). Genetic engineering of plants for viral disease resistance. *In Cell Culture and somatic cell genetics of plants* vol 6. pp 405-420.
- Herskowitz, I. (1987). Functional inactivation of genes by dominant negative mutations. *Nature* **329**, 219-222.
- Heyraud, F., Matzeit, V., Kammann, M., Schaefer, S., Schell, J., and Gronenborn, B. (1993b). Identification of the initiation sequence for viral-strand DNA synthesis of wheat dwarf virus. *EMBO J.* **12**, 4445-4452.

- Heyraud, F., Matzeit, V., Schaefer, S., Schell, J., and Gronenborn, B. (1993a). The conserved nonanucleotide motif of the geminivirus stem-loop sequence promotes replicational release of virus molecules from redundant copies. *Biochemie* **75**, 605-615.
- Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J., and Gronenbron, B (1995). Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucleic.Acids.Res* **23**, 910-916.
- Hofer, J.M.I., Dekker, E.L., Reynolds, H.V., Woolston, C.J., Cox, B.S., and Mullineaux, P.M. (1992). Coordinate regulation of replication and virion sense gene expression in wheat dwarf virus. *Plant Cell* **4**, 213-223.
- Hong, Y. and Stanley, J. (1996). Virus resistance in *Nicotiana benthamiana* conferred by *African cassava mosaic virus* replication-associated protein (AC1) transgene. *MPMI* **9**, 219-225.
- Hong, Y., Saunders, K., Hartley, M.R., and Stanley, J. (1996). Resistance to geminivirus infection by virus-induced expression of dianthin in transgenic plants. *Virology*. **220**, 119-127.
- Hormuzdi, S. G. and Bisaro, D. M. (1993). Genetic analysis of beet curly top virus: evidence for three virion sense genes involved in movement and regulation of single- and double- stranded DNA levels. *Virology* **193**, 900-909.
- Horvath, G. V., Pettko-Szandtner, A., Nikovics, K., Bilgin, M., Boulton, M., Davies, J. W., Gutierrez, C., and Dudits, D. (1998). Prediction of functional regions of the maize streak virus replication-associated proteins by protein-protein interaction analysis. *Plant Mol. Biol.* **38**, 699-712.
- Huntley, R., Healy, S., Freeman, D., Lavender, P., de Jager, S., Greenwood, J., Makkerh, J., Walker, E., Jackman, M., Xie, Q., Bannister, A.J., Kouzarides, T., Gutierrez, C., Doonan, J.H., and Murray, J.A.H. (1998). The maize retinoblastoma protein homologue ZmRb-1 is regulated during maize leaf development and displays conserved interactions with G1/S regulators and plant cyclin D (CycD) proteins. *Plant Mol. Biol.* **37**, 155-169.

- Inze, D., Gutierrez, C., and Chua, N-H. (1999). Trends in plant cell cycle research. *Plant Cell*. **11**, 991-994.
- Jefferson, R.A. (1987). Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Mol Biol Rep*. **5**, 387-405
- Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J*. **6**, 3901-3907
- Jeske, H., Lutgemeier, M., and Prei β , W. (2001). DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *EMBO J*. **20**, 6158-6167.
- Kammann, M., Schalk, H. J., Matzeit, V., Schaefer, S., Schell, J., and Gronenborn, B. (1991). DNA replication of wheat dwarf virus, a geminivirus, requires two cis- acting signals. *Virology*. **184**, 786-790.
- Kapitonov, V. V. and Jurka, J. (2001). Rolling-circle transposons in eukaryotes. *Proc.Natl.Acad.Sci.U.S.A.* **98**, 8714-8719.
- Kasschau, K.D., and Carrington, J.C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell*. **95**: 461-470
- Kimmins, F.M., and Bosque-Pérez, N.A. (1996). Electrical penetration graphs from *Cicadulina* spp. and the inoculation of a persistent virus into maize. *Entomol. Exp. Applic.* **80**, 46-49.
- Kjemtrup, S., Sampson, K.S., Peele, C.G., Nguyen, L.V., Conkling, M.A., Thompson, W.F., and Robertson, D. (1998). Gene silencing from plant DNA carried by a geminivirus. *Plant J*. **14**, 91-100
- Konate, G. and Traore, O. (1994). Variability of the maize streak virus (MSV) in the Sudano-Sahelian region. *Phytoprotection*. **75**, 91-99

- Koncz, C. and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* **204**, 383-396.
- Kong, L.J., Orozco, B.M., Roe, J.L., Nagar, S., Ou, S., Feiler, H.S., Durfee, T., Miller, A.B., Griussem, W., Robertson, D., and Hanley-Bowdoin, L. (2000). A geminivirus replication protein interacts with the retinoblastoma protein through a novel domain to determine symptoms and tissue specificity of infection in plants. *EMBO J.* **19**, 3485-3495.
- Kong, L.-J., and Hanley-Bowdoin, L. (2002). A geminivirus replication protein interacts with a protein kinase and a motor protein that display different expression patterns during plant development and infection. *Plant Cell.* **14**, 1817-1832.
- Koonin, E. V. and Ilyina, T. V. (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. *J. Gen. Virol.* **73**, 2763-2766.
- Koonin, E. V. and Ilyina, T. V. (1993). Computer-assisted dissection of rolling circle DNA replication. *Biosystems.* **30**, 241-268.
- Kornberg, A., and Baker, T.A. 1992. DNA Replication. 2ed. W.H.Freeman and Company, New York.
- Kotlizky, G., Boulton, M.I., Pitaksutheepong, C., Davies, J.W., and Epel, B.L. (2000). Intracellular and intercellular movement of maize streak geminivirus V1 and V2 proteins transiently expressed as green fluorescent protein fusions. *Virology.* **274**, 32-38.
- Kunik, T., Salomon, R., Zamir, D., Navot, N., and Zeiden, M. (1994). Transgenic tomato plants expressing the tomato yellow leaf curl capsid protein are resistant to the virus. *Bio/Technology* **12**, 500-504.
- La Thangue, N.B. (1994) DRTF1/E2F: an expanding family of heterodimeric transcription factors implicated in cell-cycle control. *Trends Biochem. Sci* **19**, 108-114
- Latham, J.R., Saunders, K., Pinner, M.S., and Stanley, J. (1997). Induction of plant cell division by beet curly top virus gene C4. *Plant J.* **11**, 1273-1283.

- Laufs, J., Jupin, I., Schumacher, S., Heyraud-Nitschke, F., and Gronenborn, B. (1995a). Geminivirus replication: genetic and biochemical characterisation of Rep protein function, a review. *Biochemie*. **77**, 765-773.
- Laufs, J., Schumacher, S., Geisler, N., Jupin, I., and Gronenborn, B. (1995b). Identification of the nicking tyrosine of geminivirus Rep protein. *FEBS letters* **377**, 258-262.
- Laufs, J., Traut, W., Heyraud, F., Matseit, V., Rogers, S.G., Schell, J., and Gronenborn, B. (1995c). In vitro cleavage and joining at the viral origin of replication by the replication initiator protein of tomato yellow leaf curl virus. *Proc. Natl. Acad. Sci. USA*. **92**, 3879-3883.
- Lazarowitz, S. G., Wu, L. C., Rogers, S. G., and Elmer, J. S. (1992). Sequence-specific interaction with the viral AL1 protein identifies a geminivirus DNA replication origin. *Plant Cell* **4**, 799-809.
- Liu, H., Boulton, M.I., and Davies, J.W. (1997). Maize streak virus coat protein binds single and double stranded DNA in vitro. *J. Gen. Virol.* **78**, 1265-1270
- Liu, H., Boulton, M.I., Thomas, C.L., Prior, D.A.M., Oparka, K.J., and Davies, J.W. (1999b). Maize streak virus coat protein is karyophyllic and facilitates nuclear transport of viral DNA. *Mol. Plant. Microb. Interact.* **12**, 894-900
- Liu, H., Lucy, A.P., Davies, J.W., and Boulton, M.I. (2001). A single amino acid change in the coat protein of maize streak virus abolishes systemic infection, but not interaction with viral DNA or movement protein. *Mol. Plant. Path.* **2**, 223-228
- Liu, L., Davies, J.W., and Stanley, J. 1998. Mutational analysis of *Bean yellow dwarf virus*, a geminivirus of the genus *Mastrevirus* that is adapted to dicotyledonous plants. *J. Gen. Virol.* **79**, 2265-2274.
- Liu, L., Saunders, K., Thomas, C.L., Davies, J.W., and Stanley, J. 1999a. *Bean yellow dwarf virus* RepA, but not Rep, binds to maize retinoblastoma protein, and the virus tolerates mutations in the consensus binding motif. *Virology*. **256**, 270-279

- Lomonossoff, G.P. (1995). Pathogen-derived resistance to plant viruses. *Annu. Rev. Phytopath.* **33**: 323-343
- Lopato, S., Gattoni, R., Fabini, G., Stevenin, J., and Barta, A. (1999). A novel family of plant splicing factors with a Zn knuckle motif: examination of RNA binding and splicing activities. *Plant Mol Biol.* **39**, 761-773
- Lucy, A.P., Boulton, M.I., Davies, J.W., and Maule, A.J. (1996). Tissue Specificity of *Zea mays* Infection by *Maize Streak Virus*. *Molecular Plant-Microbe Interactions.* **9**, 22-31
- Ludlow, J.W. (1993). Interactions between SV40 large-tumor antigen and the growth suppressor proteins pRB and p53. *FASEB J.* **7**, 866-871.
- Luque, A., Sanz-Burgos, A.P., Ramirez-Parra, E., Castellano, M., and Gutierrez, C. (2002). Interaction of geminivirus Rep protein with replication factor C and its potential role during geminivirus DNA replication. *Virology.* **302**, 83-94
- Magnaghi-Jaulin, J.L., Groisman, R., Naguibneva, I., Robin, P., Lorain, S., Le, V.J., Troalen, F., Trouche, D., and Harel, B.A. (1998). Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391**, 601-605.
- Mangwende, T. (2000). The Effect of Mutant Replication Proteins and Antisense RNA to the Gene on *Maize Streak Virus* Replication. PhD Thesis. UCT, Cape Town, South Africa.
- Manley, J.L. and Tacke, R. (1996). SR proteins and splicing control. *Genes and Development* **10**, 1569-1579
- Markham, P.G., Pinner, M.S., and Boulton, M.I. (1984). *Dalbulus maidis* and *Cicadulina* species as vectors of diseases in maize. *Maize Virus Disease Newsletter* **1**, 33-34.
- Martin, D.P., and Rybicki, E.P. (1998). Microcomputer-based quantification of *Maize streak virus* symptoms in *Zea mays*. *Phytopathology* **88**, 422-427.

- Martin, D.P., Willment, J.A., and Rybicki, E.P. (1999). Evaluation of *Maize Streak Virus* Pathogenicity in Differentially Resistant *Zea mays* Genotypes. *Phytopathology*. **89**, 695-700.
- Martin, D.P., Willment, J.A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., James, D., and Rybicki, E.P. (2001). Sequence Diversity and Virulence in *Zea mays* of *Maize Streak Virus* Isolates. *Virology* **288**, 247-255
- Mazithulela, G, Sudhakar, D., Heckel, T., Mehlo, L., Christou, P., Davies, J.W., and Boulton, M.I. (2000) The *Maize streak virus* coat protein transcription unit exhibits tissue-specific expression in transgenic rice. *Plant Science* **155**, 21-29.
- McGivern, D. (2002). Functional Analysis of the *Maize streak virus* Replication-associated Proteins. PhD thesis. UEA: Norwich, UK.
- McKinney, H.H. (1929). Mosaic diseases in the Canary Islands, West Africa and Gibraltar. *J Agricult Res* **39**, 557-578
- Mesfin, T., Bosque-Perez, N. A., Buddenhagen, I. W., Thottapilly, G., and Olojede, S. O. (1992). Studies of maize streak virus isolates from grass and cereal hosts in Nigeria. *Plant Disease*. **76**, 789-795.
- Mironov, V., De Vaylder, L., Van Montagu, M., and Inze, D. (1999). Cyclin-dependent kinases and cell division in plant: the nexus. *Plant Cell* **11**: 509-521.
- Missich, R., Ramirez-Parra, E., and Gutierrez, C. (2000). Relationship of oligomerization to DNA binding of *Wheat dwarf virus* RepA and Rep proteins. *Virology*. **273**, 178-188.
- Moffat, A.S. (1999). Geminiviruses emerge as serious crop threat. *Science* **286**, 1835
- Mohr, I.J., Clark, R., Sin, S., Androphy, E.J., MacPherson, P., and Botchan, M.R. 1990 Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with E2 transactivator. *Science* **250**, 1694-1699.

- Monsoor, S., Khan, S.H., Bashir, A., Saeed, M., Zafar, Y., Malik, K.A., Briddon, R., Stanley, J., and Markham, P.G. 1999. Identification of a novel circular single stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology*. **259**, 190-199.
- Moran, E. (1994). Mammalian cell growth controls reflected through protein interactions with the adenovirus E1A gene products. *Semin. Virol.* **5**, 327-340.
- Mullineaux, P. M., Boulton, M. I., Bowyer, P., van der Vlugt, R., Marks, M., Donson, J., and Davies, J. W. (1988). Detection of a non-structural protein of Mr 11 000 encoded by the virion DNA of maize streak virus. *Plant Mol. Biol.* **11**, 57-66.
- Mullineaux, P. M., Donson, J., Morris-Krsinich, B. A., Boulton, M. I., and Davies, J. W. (1984). The nucleotide sequence of maize streak virus DNA. *EMBO J.* **3**, 3063-3068.
- Mullineaux, P., Gooding, P., Batty, N., and Goldsmith, A. (1999). Two mechanisms of control of virion-sense gene expression in wheat dwarf virus.
- Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bioassay with tobacco tissue culture. *Plant Physiology* **15**, 473-497.
- Nagar, S., Pederson, T.J., Carrick, K.M., and Hanley-Bowdoin, L. (1995). A geminivirus induces expression of a host DNA synthesis protein in terminally differentiated plant cells. *Plant Cell.* **7**, 705-719.
- Nikovics, K., Simidjieva, J., Peres, A., Ayaydin, F., Pasternak, T., Davies, J., Boulton, M.I., Dudits, D. and Horvath, G.V. (2001). Cell-cycle, phase-specific activation of *Maize streak virus* promoters. *MPMI.* **14**, 609-617
- Njuguna, J.G.M., Theuri, J.M., and Ininda, J.M. (1990). The status of maize streak disease in Kenya. *In: Maize Improvement, Production and Protection in Eastern and Southern Africa.* Gebrekidan B (ed), pp 218 - 223. AMREF Printing Dept., Nairobi, Kenya.

- Noris, E., Accotto, G.P., Tavazza, R., Brunetti, A., Crespi, S., and Tavazza, M. (1996). Resistance to *Tomato yellow leaf curl* geminivirus in *Nicotiana benthamiana* plants transformed with a truncated viral C1 gene. *Virology*. **224**, 130-138.
- Orozco, B. M. and Hanley-Bowdoin, L. (1996). A DNA structure is required for geminivirus replication origin function. *J. Virol.* **70**, 148-158.
- Orozco, B. M., Gladfelter, H. J., Settlege, 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.
- Orozco, B.M., Kong, L.J., Batts, L.A., Elledge, S., Hanley-Bowdoin, L. (2000). The multifunctional character of a geminivirus replication protein is reflected by its complex oligomerization properties. *J. Biol. Chem.* **275**, 6114-6122.
- Orozco, B. M., Miller, A. B., and Settlege, S. B. (1997). Functional domains of a geminivirus replication protein. *J. Biol. Chem.* **272**, 9840-9846.
- Orozco, B.M., and Hanley-Bowdoin, L. (1998). Conserved sequence and structural motifs contribute to the DNA binding and cleavage activities of a geminivirus replication protein. *J. Biol. Chem.* **273**, 24448-244456.
- Oruetxebarria, I., Kvarnheden, A., and Valkonen, J.P.T. (2002). Analysis of putative interactions between potyviral replication proteins and plant retinoblastoma proteins. *Virus genes* **24**:1, 65-75.
- Oshima, K., Kakizawa, S., Nishigawa, H., Kuboyama, T., Miyata, S., Ugaki, M., and Namba, S. 2001. A plasmid of phytoplasma encodes a unique replication protein having both plasmid- and virus-like domains: clue to viral ancestry or result of virus/plasmid recombination? *Virology* **285**, 270-277.
- Padidam, M., Beachy, R. N., and Fauquet, C. M. (1995). Classification and identification of geminiviruses using sequence comparisons. *J. Gen. Virol.* **76**, 249-263.
- Padidam, M., Sawyer, S., and Fauquet, C.M. (1999a). Possible emergence of new geminiviruses by frequent recombination. *Virology* **265**, 218-225.

- Palmer, K. E. and Rybicki, E. P. (1998). The molecular biology of mastreviruses. *Adv. Virus Res.* **50**, 183-234.
- Palmer, K.E. and Rybicki, E.P. (2001). Investigation of the potential of *Maize streak virus* to act as an infectious gene vector in maize plants. *Arch Virol.* **146**, 1089-1104.
- Pant, V., Gupta, D., Choudhury, N.R., Malathi, V.G., Varma, A., and Mukherjee, S.K. (2001). Molecular characterisation of the Rep protein of the blackgram isolate of Indian mungbean yellow mosaic virus. *J. Gen Virol.* **82**, 2559-2567.
- Peng, Y.C. and Acheson, N.H. (1998). Polyomavirus large T antigen binds cooperatively to its multiple binding sites in the viral origin of DNA replication. *J. Virol.* **72**, 7330-7340.
- Peterschmitt, M., Quiot, J.B., and Reynaud, B. (1992). Detection of maize streak virus antigens over time in different parts of maize plants of a sensitive and so-called tolerant cultivar by ELISA. *Annals of Applied Biology* **121**, 641-653.
- Pinner, M.S., Medina, V., and Plaskitt, K.A. (1993). Viral inclusions in monocotyledons infected by maize streak and related geminiviruses. *Plant Pathology* **42**: 75-87
- Powell-Abel, P., Nelson, R.S., De, B., Hoffman, N., Rogers, S.G., Fraley, R.T., and Beachy, R.N. (1986). Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science* **232**, 738-743
- Ramirez-Parra, E., Xie, Q., Boniotti, M.B., and Gutierrez, C. (1999). The cloning of plant E2F, a retinoblastoma-binding protein, reveals unique and conserved features with animal G1/S regulators. *Nucl.Ac.Res.* **27**, 3527-3533.
- Reynaud, B. and Peterschmitt, M. (1992). A study of the mode of transmission of maize streak virus by *Cicadulina mbila* using an enzyme-linked immunosorbent assay. *Ann. Appl. Biol.* **121**, 85-94.
- Rigden, J. E., Dry, I. B., Krake, L. R., and Rezaian, M. A. (1996). Plant virus DNA replication processes in *Agrobacterium*: insight into the origins of geminiviruses? *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10280-10284.

- Rigden, J. E., Dry, I. B., Mullineaux, P. M., and Rezaian, M. A. (1993). Mutagenesis of the virion-sense open reading frames of tomato leaf curl geminivirus. *Virology* **193**, 1001-1005.
- Rodier, A., Assio, J., Marchand, J-L., and Hervo, Y (1995). Breeding maize lines for complete and partial resistance to maize streak virus (MSV). *Euphytica*. **81**: 57-70.
- Rose, D.J.W. (1972). Times and sizes of dispersal flight by *Cicadulina* spp. vectors of maize streak disease. *J. Anim. Ecol.* **41**, 494-506
- Rose, D.J.W. (1973). Field studies in Rhodesia on *Cicadulina* spp. (Homoptera: *Cicadellidae*), vectors of maize streak disease. *Bull. Entomol. Res.* **62**, 477-495.
- Rose, D.J.W. (1978). Epidemiology of maize streak disease. *Ann. Rev. Ent.* **23**: 250-282.
- Ross, J.F., Liu, X., and Dynlacht, B.D. (1999). Mechanism of transcriptional repression of E2F by the retinoblastoma tumor suppressor protein. *Mol. Cell* **3**: 195-205.
- Rossel, H.W. and Thottappilly, G. (1985). Virus diseases of important food crops in tropical Africa. IITA, Ibadan, Nigeria, 61 pp.
- Runzler, R., Thompson, S., and Fanning, E. (1987) Oligomerization and origin DNA-binding activity of simian virus 40 large T antigen. *J. Virol.* **61**, 2076-2083.
- Rybicki, E. P. (1994). A phylogenetic and evolutionary justification for three genera of *Geminiviridae*. *Arch.Virol.* **139**, 49-77.
- Rybicki, E. P., Briddon, R. W., Brown, J. E., Fauquet, C. M., Maxwell, D. P., Harrison, B. D., Markham, P. G., and Stanley, J. (2000). *Geminiviridae*. In "Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses." (C. M. Fauquet, D. H. L. Bishop, E. B. Carsten, M. K. Estes, S. M. Lemon, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner, Eds.), pp. 285-297. Academic Press, San Diego.

- Rybicki, E.P. and Pieterse, G. (1999). Plant virus disease problems in the developing world. In *Advances in virus research* **53**, 127-175. Maramorosch, K., Murphy, F.A., and Shatkin, A.J. (Eds). San Diego, Academic Press
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning, a laboratory manual*, 2nd edn. Cold Spring Harbour Laboratory Press, New York.
- Sanford, J.C. and Johnson, S.A. (1985). The concept of parasite-derived resistance - Deriving resistance genes from the parasite's own genome. *J. Theor. Biol.* **113**, 395-405.
- Sangare, A., Deng, C.M. Fauquet, and R.N. Beachy (1999). Resistance to *African cassava mosaic virus* conferred by a mutant of the putative NTP-binding domain of the Rep gene (AC1) in *Nicotiana benthamiana*. *Molecular-Breeding*. **5**(2), 95-102
- Sanz-Burgos, A. P. and Gutierrez, C. (1998). Organization of the cis-acting element required for *Wheat dwarf geminivirus* DNA replication and visualization of a Rep protein-DNA complex. *Virology*. **243**, 119-129.
- Saunders, K. and Stanley, J. (1999). A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides*: evidence for interfamilial recombination between plant DNA viruses. *Virology* **264**, 142-152.
- Saunders, K., Bedford, I. D., Briddon, R. W., Markham, P. G., Wong, S. M., and Stanley, J. (2000). A unique virus complex causes *Ageratum yellow vein disease*. *Proc. Natl. Acad. Sci. U S A*. **97**, 6890-6895.
- Saunders, K., Lucy, A., and Stanley, J. (1992). RNA-primed complementary-sense DNA synthesis of the geminivirus *African cassava mosaic virus*. *Nucleic Acids Res.* **20**, 6311-6315.
- 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. *Nucl. Acids Res.* **19**, 2325-2330.
- Schneider, M., Jarchow, E., and Hohn, B. (1992). Mutational analysis of the 'conserved region' of maize streak virus suggests its involvement in replication. *Plant Mol Biol* **19**, 601-610.

- Schnippenkoetter, W.H., Martin, D.P., Willment, J.A., and Rybicki, E.P. (2001). Forced recombination between distinct strains of *Maize streak virus*. *J. Gen. Virol.* **82**, 3081-090.
- Selth, L.A., Randles, J.W., and Rezaian, M.A. (2002). *Agrobacterium tumefaciens* supports DNA replication of diverse geminivirus types. *FEBS Letters* **516**, 179-182.
- Settlage, S. B., Miller, A. B., and Hanley-Bowdoin, L. (1996). Interactions between geminivirus replication proteins. *J. Virology* **70**, 6790-6795.
- Settlage, S.B., Miller, A.B., Gruissem, W., and Hanley-Bowdoin, L. (2001). Dual interaction of a geminivirus replication accessory factor with a viral replication protein and a plant cell cycle regulator. *Virology* **279**, 570-576.
- Sever-Chroneos, Z., Angus, S.P., Fribourg, A.F., Wan, H., Todorov, I., Knudsen, K., and Knudsen, E.S. (2001). Retinoblastoma Tumor Suppressor Protein Signals through Inhibition of Cyclin-Dependent Kinase 2 Activity to Disrupt PCNA Function in S Phase. *Mol. Cell. Biol.* **21**, 4032-4045.
- Sinisterra, X. H., Polston, J. E., Abouzid, A. M., and Hiebert, E. (1999). Tobacco plants transformed with a modified coat protein of tomato mottle begomovirus show resistance to virus infection. *Phytopathology* **89**, 701-706.
- Soni, R., Carmichael, J.P., Shah, Z.H., and Murrey, J.A.H. (1995). A family of cyclin D homologues from plants differentially controlled by growth regulators and containing the conserved retinoblastoma protein interaction motif. *Plant Cell*. **7**, 85-103.
- Stanley, J. (1995). Analysis of *African cassava mosaic virus* recombinants suggests strand nicking occurs within the conserved nonanucleotide motif during the initiation of rolling circle DNA replication. *Virology*. **206**, 707-712.
- Stanley, J. and Gay, M. R. (1983). Nucleotide sequence of cassava latent virus DNA. *Nature* **301**, 260-262.

- Stanley, J., Boulton, M.I., and Davies, J.W. (1998) *Geminiviridae*. In *Encyclopaedia of Life Sciences* pp 1-12. Stanley, J (Ed). Macmillan Reference Ltd.
- Stanley, J., Frischmuth, T., and Ellwood, S. (1990). Defective viral DNA ameliorates symptoms of geminivirus infection in transgenic plants. *Proc. Natl. Acad. Sci. USA*. **87**, 6291-6295
- Stenger, D.C. (1994). Strain-specific mobilization and amplification of a transgenic defective-interfering DNA of the geminivirus *Beet Curly Top Virus*. *Virology*. **203**, 397-402
- Stenger, D.C., Revington, G.N., Stevenson, M.C., and Bisaro, D.M. (1991). Replicational release of geminivirus genomes from tandemly repeated copies: evidence for rolling circle replication of a plant viral DNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8029-8033.
- Storey, H. H. (1924). The transmission of a new plant virus disease by insects. *Nature* **144**, 245.
- Storey, H.H. (1925). The transmission of streak disease of maize by the leafhopper *Balclutha mbila* Naudé. *Ann. Appl. Biol.* **12**, 422-439.
- Storey, H.H. (1928). Transmission of maize streak disease. *Ann Appl. Biol.* **15**, 1-25
- Storey, H.H. (1931). The inheritance by a leafhopper of the ability to transmit a plant virus. *Nature*. **127**, 928.
- Storey, H.H. (1933). Investigations of the mechanism of the transmission of plant viruses by insect vectors. *I. Proc. R. Soc. Lond. (Ser. B)* **113**, 463-485
- Storey, H.H. (1938). Investigation of the mechanism of the transmission of plant viruses by insect vectors. II. The part played by puncture in transmission. *Proc. Royal Soc.* **125**, 455-477
- Storey, H.H. and Howland, A.K. (1967). Inheritance of resistance in maize to the virus of streak disease in East Africa. *Ann. Appl. Biol.* **59**, 429-436.
- Stout, J.T., Luu, H.T., Hanson, S.F., Maxwell, D.P., Ahlquist, P.G., and Gilbertson, R.L. (2001). Transgenic plants expressing mutant geminivirus AC1 or C1 genes. United States Patent Application 20010011379.

- Suarez-Lopez, P. and Gutierrez, C. (1997). DNA replication of wheat dwarf geminivirus vectors: effects of origin structure and size. *Virology* **227**, 389-399.
- Suarez-Lopez, P., Martinnez-Salas, E., Hernandez, P., and Gutierrez, C. (1995). Bent DNA in the large intergenic region of wheat dwarf geminivirus. *Virology* **208**, 303-311.
- Sunter, G. and Bisaro, D. M. (1991). Transactivation in a geminivirus: AL2 gene product is needed for coat protein expression. *Virology* **180**, 416-419.
- Sunter, G. and Bisaro, D. M. (1992). Transactivation of geminivirus AR1 and BR1 gene expression by the viral AL2 gene product occurs at the level of transcription. *Plant Cell* **4**, 1321-1331.
- Sunter, G. and Bisaro, D. M. (1997). Regulation of a geminivirus coat protein promoter by AL2 protein (TrAP): evidence for activation and derepression mechanisms. *Virology* **232**, 269-280.
- Thottappilly, G., Bosque-Perez, N.A., and Rossel, H.W. (1993). Viruses and virus diseases of maize in tropical Africa. *Plant Pathol.* **42**, 494-509.
- Voinnet, O., Pinto, Y.M., and Baulcombe, D.C. (1999). Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *PNAS*. **96**, 14147-14152.
- von Arnim, A. and Stanley, J. (1992a). Determinants of tomato golden mosaic virus symptom development located on DNA B. *Virology* **186**, 286-293.
- von Arnim, A. and Stanley, J. (1992b). Inhibition of African cassava mosaic virus systemic infection by a movement protein from the related geminivirus tomato golden mosaic virus. *Virology*. **187**, 555-64.
- Vousden, K. (1993). Interactions of human papillomavirus transforming proteins with the products of the tumor suppressor genes. *FASEB J.* **7**, 872-879.

- Wambugu, F. and Wafula, J. (Eds) (1999) Advances in Maize Streak Virus Disease Research in Eastern and Southern Africa, Workshop Report, 15-17 September 1999, KARI and ISAAA AfriCenter, Nairobi, Kenya. ISAAA Briefs No.16. ISAAA: Ithica, NY. 43p.
- Ward, B.M., and Lazarowitz, S.G. (1999). Nuclear export in plants. Use of geminivirus movement proteins for a cell-based export assay. *Plant Cell* **11**, 1267-1276
- Weinberg, R.A. (1995) The retinoblastoma protein and cell cycle control. *Cell* **81**:323-330.
- Willment, J. A. (1999) Investigation of the molecular determinants of *Maize streak virus* replication. PhD Thesis. University of Cape Town, South Africa.
- Willment, J.A., Martin, D.P., and Rybicki, E.P. (2001). Analysis of the diversity of African streak mastreviruses using PCR-generated RFLPs and partial sequence data. *J. Virol. Methods* **93**, 75-87.
- Wright, E. A., Heckel, T., Groenendijk, J., Davies, J. W., and Boulton, M. I. (1997). Splicing features in maize streak virus virion- and complementary-sense gene expression. *Plant J.* **12**, 1285-1297.
- Xie, Q., Sanz-Burgos, A.P., Hannon, G.J., and Gutierrez, C. (1996). Plant cells contain a novel member of the retinoblastoma family of growth regulatory proteins. *EMBO J.* **15**, 4900-4908.
- Xie, Q., Suarez-Lopez, P. and Gutierrez, C. (1995). Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant virus: requirement for efficient viral DNA replication. *EMBO J.* **14**, 4073-4082.
- Zhan, X., Richardson, K. A., Haley, A., and Morris, B. A. (1993). The activity of the coat protein promoter of chloris striate mosaic virus is enhanced by its own and C1-C2 gene products. *Virology* **193**, 498-502.
- Zhang, W., Olson, N.H., Baker, T.S., Faulkner, L., Agbandje-McKenna, M., Boulton, M.I., Davies, J.W., and McKenna, R. (2001). Structure of the *Maize streak virus* geminate particle. *Virology*. **279**, 471-477