INVESTIGATIONS INTO THE USE OF MAIZE STREAK VIRUS AS A GENE VECTOR

BY

KENNETH EDWARD PALMER

A thesis submitted in fulfillment of the requirements
for the degree of

Doctor of Philosophy

Department of Microbiology
University of Cape Town

MARCH 1997

INVESTIGATIONS INTO THE USE OF MAIZE STREAK VIRUS AS A GENE VECTOR

TABLE OF CONTENTS

ABSTRACT

This thesis describes investigations into the potential use of the Subgroup I geminivirus, maize streak virus (MSV), as a gene vector. These involved testing MSV-based replicons in transgenic cell lines, in transient expression assays in maize cells and in an infectious gene expression system in plants.

MSV vectors which contained three different versions of a *bar* (bialaphos resistance) expression cassette in place of the viral movement and coat protein genes were used to generate transformed maize cell lines. A high proportion of these contained MSV-based episomes at high copy number. However, embryogenic maize tissue of the HiII line was not regenerable when an MSV-based replicon was present, possibly due to toxicity of the viral replication associated protein. In non-regenerable Black Mexican sweetcorn cell lines some of the MSV-*bar* episomes, which ranged in size from 3.15 kb to 4.78 kb, replicated for periods of over two years, and appeared structurally stable. The cellular levels of the *bar* gene product, phosphinothricin acetyl transferase (PAT), were significantly enhanced in lines where the gene was amplified by linkage to an MSV replicon in comparison with lines where the same gene was not amplified. Northern blot analysis also showed that higher levels of *bar* mRNA were produced in lines where the gene was amplified. However, the 3- to 5-fold enhancement in gene expression was less than was anticipated, based on results from similar Subgroup III geminivirus-based transgene amplification systems.

Several mutants of the MSV genome were generated to investigate the extent to which genome amplification contributes to the expression of the viral coat protein gene. The introduction of an *Nco*I restriction site at the start of the coat protein gene facilitated fusion of the *gus* marker gene with the coat protein upstream transcription and translation regulatory sequences. In one viral construct the plus strand origin of replication was inactivated by insertion of a short oligonucleotide; in another, the viral *rep* gene was inactivated by a frameshift mutation. These constructs were used to show that the MSV coat protein promoter has low, but measurable constitutive activity in the absence of genome amplification, but that viral replication enhances coat protein expression about 45-fold. I found no evidence for Rep-mediated transactivation of the coat protein promoter.

Transformed cell lines were generated which separately contained an MSV genome carrying a CaMV 35S promoter-*bar* expression unit as a selectable marker and another containing the *gus* gene under the control of the coat protein 5′ expression regulatory elements. Bialaphos resistant, GUS-positive lines were selected by histochemical assay: most contained high copy numbers of both the *bar* and GUS replicons. There was evidence for *trans*-replication of replication-deficient GUS constructs and also for recombination between the two different episomes. I found that GUS expression levels in transformed cell lines with high copy number MSV-GUS episomes were enhanced up to 90-fold.

To test whether infectious gene vectors derived from MSV were feasible, sweetcorn seedlings were agroinfected with various versions of a construct containing the CaMV 35S promoter and *bar* gene in place of the MSV virion sense ORFs: this replicon was exactly the same size as the MSV genome. Movement and coat protein genes were provided in *trans* by a wild type genome or a replication deficient MSV mutant, cloned on the same *Agrobacterium* binary vector. The wild type virus could complement the ssDNA formation-negative phenotype of the *bar* gene replacement mutant; however neither the wild type nor the replication-deficient mutant could complement the movement of the recombinant virus *in planta*. Rather, I noted efficient generation of replication- and movement-competent virus by homologous recombination between complementing mutants. MSV *bar*-carrying replicons were only present in the first two to four leaves of agroinfected plants, whether or not virus which could complement the deleted movement functions was present. To investigate whether the absence of movement was due to deletion of an encapsidation signal in the recombinants, I co-infected maize with both Digitaria streak virus (DSV) and MSV, which have different insect vectors: encapsidation of the viral genome is a prerequisite for viral transmission. The MSV vector, *Cicadulina mbila*, was able to transmit DSV from doubly infected plants, implying that the DSV genome was *trans*-encapsidated by the MSV coat protein. Therefore, either there is no specific encapsidation signal, or this is conserved between DSV and MSV.

The major conclusions of this research were that MSV has good potential for use as a gene vector for enhancing both transient and stable expression of foreign genes in cereals, but that its use as an infectious gene expression system is limited to the first few leaves in agroinfected plants.

ACKNOWLEDGMENTS

I gratefully acknowledge the following people and organisations for their contributions to my life and work as a Ph.D. student:

First, and foremost, my supervisor and mentor, Professor Ed Rybicki, for welcoming me and my ego into his lab, and giving us space to develop. Thanks Ed, for providing a fun and academically stimulating environment, for your enormous contribution to my intellectual growth, for giving me the freedom to pursue my own interests in this project, and of course, most latterly for thesis proofreading.

Professor Jennifer Thomson, head of the Microbiology Department, for facilitating much of the research presented here in terms of providing me with the maize transformation facilities and technical assistance with the tissue culture load, for her considerable professional mentoring, and for proofreading this thesis.

Professor Karl Kunert of AECI Ltd, for his support and constructive criticism of the research presented here. I am also grateful to AECI for giving me a generous Postgraduate Research Scholarship; to UCT for three years' awards of a Postgraduate Research Associateship; and the FRD for their support in the last 18 months.

Chantal Daniels and Bernice Hobbs, for cheerful technical assistance with media preparation and routine tissue culturing. Sandy Lennox taught me plant tissue culture, and I am grateful to her for all the advice, help and discussion so freely given. I thank Di James for her help with DNA sequencing and many colleagues in the Microbiology Department for advice on varied technical matters. Thanks to Professor George Lindsey of the Biochemistry Department for help with PAT assays. I acknowledge Professor Barbara von Wechmar for assistance with leafhopper transmissions and for providing plant growth room facilities and Mohamed Jaffer for electron microscopy.

I must particularly acknowledge the generosity of Dr Brad Roth and the company he works for, Pioneer Hi-Bred, Inc. (Des Moines, Iowa), for several plasmid constructs, tissue culture cell lines and invaluable advice on maize transformation and transient expression assays. I also thank Dr Margaret Boulton for valuable discussions, for giving me permission to cite a lot of her unpublished research in this thesis and for sharing her insight into Subgroup I geminiviruses. Thanks also to geminivirologists who gave me reprints of publications not easily available in South Africa, and to those people who provided plasmids, who are specifically acknowledged in the text.

The camaraderie and friendship I experienced from all of the members, past and present, of the Rybicki Lab enriched my whole Ph.D. experience. I acknowledge Janet Willment especially for valuable scientific discussions and comments on Chapter 1 and Darren Martin for his comments on Chapter 1 and advice on various computer matters.

Grateful thanks to Anne van Zyl, Margaret Lee and Robyn Hey for caring, and for looking after me so well over the last few years. Thanks especially to my brother Paul for his generosity and help whenever I had technical hassles and to my mother for copy-editing. Lastly, I thank my parents and brothers for their support, without which I would never have achieved as much over the last twenty six and a half years.

TABLE OF ABBREVIATIONS

Virus Names and Abbreviations

AbMV	Abutilon mosaic virus
ACMV	African cassava mosaic virus
BBTV	banana bunchy top virus
BGMV	bean golden mosaic virus
BDMV	bean dwarf mosaic virus
BCTV	beet curly top virus
BeYDV	bean yellow dwarf virus
CaMV	cauliflower mosaic virus
CoYMV	Commelina yellow mottle virus
CSMV	Chloris striate mosaic virus
DSV	Digitaria streak virus
FMV	figwort mosaic virus
HrCTV	horseradish curly top virus
HPV	human papillomavirus
MiSV	Miscanthus streak virus
MSV	maize streak virus
MSV-K	maize streak virus, Kenyan isolate
MSV-Kom	maize streak virus, Komatipoort, South Africa, isolate
MSV-N	maize streak virus, Nigerian isolate
MSV-Nm	maize streak virus, Nigerian mild isolate (Boulton *et al.*, 1991a; b)
MSV-Ns	maize streak virus, Nigerian severe isolate (Boulton *et al.*, 1991a; b)
MSV-R	maize streak virus, Reunion isolate
MSV-SA	maize streak virus, South African isolate (Lazarowitz, 1988)
MSV-Set	maize streak virus, isolate from *Setaria* sp.
PanSV	Panicum streak virus
PVX	potato virus X
PVY	potato virus Y
PYMV	potato yellow mosaic virus
RTBV	rice tungro bacilliform virus
SqLCV	squash leaf curl virus
SSV	sugarcane streak virus
SV40	simian virus 40
TEV	tobacco etch virus
TGMV	tomato golden mosaic virus
ToMoV	tomato mottle virus
TLCV	tomato leaf curl virus
TMV	tobacco mosaic virus
TPCTV	tomato pseudo-curly top virus
TYDV	tobacco yellow dwarf virus
TYLCV-Is	tomato yellow leaf curl virus, Israeli isolate
TYLCV-Sr	tomato yellow leaf curl virus, Sardinian isolate
WDV	wheat dwarf virus
WDV-CJI	wheat dwarf virus, Czechoslovakian isolate, John Innes Institute
WDV-F	wheat dwarf virus, French isolate
WDV-S	wheat dwarf virus, Swedish isolate

General Abbreviations

2,4-D	2,4-dichlorophenoxy acetic acid
A, C, G, T, U	nucleotides: adenine, cytosine, guanosine, thymidine, uridine
ABA	abscisic acid
Ap.	ampicillin
bar	bialaphos resistance gene
bla	beta lactamase gene, encoding ampillicin resistance
BMS	Black Mexican sweetcorn
bp	base pair
° C	degrees Celsius
ccc	covalently closed circular
CP	coat protein
cv.	cultivar
CAT	chloramphenicol acetyl transferase
CR	common region
d.a.p.	days after pollination
DEPC	diethylpyrocarbonate
DI	defective interfering
dNTPs	deoxynucleotide triphosphates
ds	double stranded
EDTA	ethylenediaminetetraacetic acid (disodium salt)
g	gram
Gm	gentamycin
gus or *gusA*	*Escherichia coli* β-glucuronidase (*uidA, gus* or *gusA*) gene
GUS	β-glucuronidase protein, gene product of the *gus* gene.
IR	intergenic region
k	prefix kilo-
kb	kilobases
kbp	kilobase pair
kDa	kilo Dalton
Km	kanamycin
l	litre
LA	Luria agar
LB	Luria broth
LIR	long intergenic region
LU	light units
m	prefix milli-
min	minute
mm	millimetre
M	molar
ml	millilitre
MP	movement protein
Mr	molar ratio
MS	Murashige and Skoog tissue culture media (Murashige and Skoog, 1962)
MW	molecular weight
μ	Greek letter Mu: prefix micro-
μl	microlitre
μg	microgram
n	prefix nano-
N/A	not applicable

N.D.	not determined
nm	nanometres
nptII	neomcin phosphotransferase II gene
nt	nucleotide(s)
ORF	open reading frame
pat	phosphinothricin acetyl transferase gene
PAT	phosphinothricin acetyl transferase protein
PCR	ploymerase chain reaction
psi	pounds per square inch
RCR	rolling circle replication
REn	replication enhancer protein
rep	replication-associated protein gene
Rep	replication-associated protein
RF	replicative form
rif	rifampicin
RLU	relative light units
RNase	ribonuclease
RF	replicative form
SIR	short intergenic region
sp.(spp.)	species (plural)
ss	single stranded
SSC	salt-sodium citrate buffer
T-DNA	transferred DNA
TE	Tris-EDTA
Ti	tumour inducing
TrAP	transcription activator protein
Tris	tris(hydroxymethyl)aminomethane
U	unit
UCT	University of Cape Town
UV	ultraviolet
vir	*Agrobacterium* virulence gene
Vir	*Agrobacterium* virulence protein
v/v	volume per volume
w/v	weight per volume
X-Gluc	5-bromo-4-chloro-3-indolyl-β-glucuronide

CHAPTER 1

GENERAL INTRODUCTION AND LITERATURE REVIEW

TABLE OF CONTENTS

1.1 INTRODUCTION

Viruses are obligate parasites which have evolved molecular mechanisms allowing them to subvert cellular processes to their own ends: for the production and amplification of their genetic material, expression of their genes, and movement to, and subsequent infection of, new host cells. With the development of recombinant DNA technology in the last two decades, a major theme of research in molecular virology has been the understanding and exploitation of viral mechanisms for controlling the genetics and physiology of their host cells, for their ultimate use in biotechnology.

DNA viruses have generally proven to be the most readily amenable to genetic manipulation, largely because of the ease with which one can manipulate their genomes *in vitro*. True DNA viruses–with no RNA phase, and which replicate in the nucleus of infected cells–are particularly attractive candidates for development of gene vectors. These viruses need to exploit the normal cellular transcription, translation and DNA replication machinery, and as such generally use the same sorts of mechanisms for regulation of gene expression as their host. The small single-stranded DNA viruses of the family *Geminiviridae* belong to one of only two families of true plant DNA viruses, and have attracted wide interest in their potential for application in both transient and transgenic amplification of genes of interest in plants, as well as for furthering our understanding of the control of the plant cell cycle, DNA replication machinery and transcriptional control of plant genes.

In recent years, interest has grown in the concept of using transgenic plants as "bioreactors" for the production of recombinant proteins and secondary metabolites which could be useful in industry. It is becoming clear that plants have the capacity for expressing a variety of foreign proteins and performing most required post-translational modifications. There is therefore considerable potential for cost-effective production of many different proteins in transgenic plants, if they are produced at sufficiently high levels. This latter point may, in many cases, prove to be the limiting factor in the use of transgenic plants for production of valuable proteins: the cost of purification of many recombinant proteins from cells of transgenic plants makes switching to transgenic plant production from conventional microbial fermentation processes somewhat unattractive to industry. The challenge in plant biotechnology, therefore, lies in developing ways to enhance the

levels of production of foreign proteins in plants. Many researchers have therefore looked to plant viruses to gain a better understanding of how to achieve high level expression of foreign genes in plant cells.

There have been several reports of the use of recombinant plant RNA viruses to achieve expression of large amounts of foreign protein and coat protein fusion peptides in infected plants. Perhaps the most notable successes for the production of full length foreign proteins in plants and isolated plant cells have been with recombinant bromoviruses, hordeiviruses, tobamoviruses and potexviruses (see, for example, Chapman *et al.* 1992; Donson *et al.*, 1993; French *et al.*, 1986; Joshi *et al.*, 1990; Kumagai *et al.*, 1993 and Mori *et al.*, 1993). As an alternative system, both cowpea mosaic virus (CPMV) and tobacco mosaic virus (TMV) have been used successfully for the production of coat protein fusion peptides for use in epitope display systems, ultimately for use as safe recombinant vaccines (e.g. Lomonossoff and Johnson, 1992; Porta *et al.*, 1994; Turpen *et al.*, 1995 and Usha *et al.*, 1993). However, there are some drawbacks associated with the use of RNA viruses as gene vectors, not the least of which are biosafety concerns about the field use of infectious and potentially vector-transmissible recombinant viruses. Other, not insurmountable, problems include genetic instability of some recombinant RNA viruses and the (theoretically) high mutation rates associated with RNA-dependent RNA replication mechanisms; these factors may limit the number of times a recombinant virus may be passaged before the gene/s of interest are lost. Thus, although it is likely that plant RNA virus vectors will play a very important role in plant biotechnology in the future, recombinant products derived from RNA viral vectors will need to be subject to strict quality control. It is for these reasons that the major players in plant biotechnology have preferred to focus on the use of conventionally produced transgenic plants for the production of valuable proteins. However, genetic elements of plant RNA viruses have still found great utility in enhancing gene expression in transgenic plants: for example, RNA leader sequences of viruses like TMV, and various potyviruses, amongst others, can act as translational enhancers, and viral 3′ pseudoknot structures may enhance RNA transcript stability and translation (reviewed by Mushegian and Shepherd, 1995).

Geminiviruses are DNA viruses which replicate to very high copy number, via a double stranded DNA replicative intermediate, in the nuclei of infected cells, and which rely entirely on host DNA replication and gene expression machinery. A method for

amplification of gene/s of interest by linking them to a geminivirus-based "replicon" could potentially incorporate positive aspects of both transgenic plant and infectious RNA viral expression systems for the design of useful transgenic and transient plant gene expression systems. Conventional logic suggests that the potentially massive gene amplification afforded by linkage to a geminivirus replicon should drive transcription of linked genes towards the maximum level achievable in a plant cell, limited only by the availability of transcription factors. Although increasing the copy number of nuclear genes frequently results in transgene "silencing", geminiviruses naturally promote significant enhancement of their own genes as well as of marker genes incorporated into the viral genome. There is certainly a great deal that we can learn from geminiviruses about the enhancement of foreign gene expression, as well as about basic plant cell biology, and it is this rationale that I have used in the development of this thesis on my investigations into the use of the Subgroup I geminivirus, maize streak virus (MSV), as a gene vector for cereals.

1.2 LITERATURE REVIEW

Before geminiviruses may be put to practical use as gene vectors, we require a thorough understanding of their molecular biology. The *Geminiviridae*, their replication, movement and their use as gene vectors have been extensively reviewed in the last few years by several authors (Bisaro, 1996; Hanley-Bowdoin *et al.*, 1997; Laufs *et al.*, 1995a; Lazarowitz, 1992; Mullineaux *et al.*, 1992; Sanderfoot and Lazarowitz, 1996; Stanley, 1993; Timmermans *et al.*, 1994). These reviews have been presented primarily from the perspective of research into the molecular biology of viruses which infect dicotyledonous plants (Subgroup II and III geminiviruses), which have been the subject of far more intensive investigation than Subgroup I viruses, the members of which are primarily pathogens of cereals. I intend, therefore, to focus my review of the literature on the molecular biology of Subgroup I geminiviruses, drawing analogies where necessary from research on Subgroup II and III viruses. To place Subgroup I geminiviruses in taxonomic context and so that I can draw parallels between molecular biology of viruses from different genera, I will first discuss the *Geminiviridae* and the genetic relationships between geminiviruses. I will then review their molecular biology, with a focus on Subgroup I, before discussing the potential for use of geminiviruses as gene vectors.

1.2.1 The *Geminiviridae* and genetic relationships between geminiviruses

Geminiviruses, named for their unique geminate capsid morphology, have small single stranded circular DNA genomes which replicate in the nuclei of infected cells via a double stranded DNA intermediate. They are responsible for economically devastating diseases in a wide variety of crop species from cereals to legumes. Each geminate particle encapsidates a circular single stranded genomic component of between 2.5 and 3 kilobases. Viruses in the taxonomic family *Geminiviridae* are classified into three genera (Subgroups I-III), based on their host range, genome organisation and vector species (Briddon and Markham, 1995; Rybicki, 1994). Subgroup I geminiviruses, such as maize streak virus (MSV), wheat dwarf virus (WDV), Chloris striate mosaic virus (CSMV), Digitaria streak virus (DSV), Miscanthus streak virus (MiSV), Panicum streak virus (PanSV) and sugarcane streak virus (SSV) have monopartite genomes, are transmitted by different leafhopper species (Homoptera: family *Cicadellidae*) and generally infect monocotyledonous plants. Recently, two Subgroup I geminiviruses have been described which infect dicotyledonous hosts: an Australian virus, tobacco yellow dwarf virus (TYDV), has been completely sequenced (Morris *et al.,* 1992) and a South African bean-infecting virus called bean yellow dwarf virus (BeYDV) is in the process of being characterised at the molecular level at the John Innes Centre in Norwich, UK (Margaret Boulton and Gerhard Pietersen, personal communication).

Subgroup III geminiviruses–such as bean and tomato golden mosaic viruses (BGMV and TGMV), African cassava mosaic virus and tomato yellow leaf curl virus (ACMV and TYLCV)–are transmitted by a single whitefly (*Bemisia tabaci*) species complex and all infect dicotyledonous plants; most have bipartite genomes, although there are some viruses in this group which have monopartite genomes (Rybicki, 1994; Briddon and Markham, 1995). The two genomic components are named, by convention, DNA A and DNA B. The DNA A genome contains the coat protein gene as well as genes which are involved in replication of the viral genome and transactivation of the virion-sense genes, while DNA B encodes two movement protein genes. New World Subgroup III geminiviruses are all more closely related to each other than to the viruses found in the Old World, which are more diverse; all New World Subgroup III geminiviruses are bipartite, and all lack a V1 open reading frame (ORF). The presence of a pre-coat (V1) ORF, the gene product of which may function as an ancillary movement protein (Paddidam *et al.* 1996), is most

likely one of the factors which has allowed some Old World Subgroup III geminiviruses to dispense with their DNA B component, at least in certain hosts.

Subgroup II geminiviruses–such as beet curly top, horseradish curly top and tomato pseudo curly top viruses (BCTV, HrCTV and TPCTV)–occupy an intermediate position between Subgroup I and Subgroup III in that they have monopartite genomes and are transmitted by leafhopper (BCTV and HrCTV) or treehopper (TPCTV) species, but only infect dicotyledonous hosts. BCTV has a remarkably broad host range for a plant virus: it is reported to infect over 100 plant families. However, recent data which show that different BCTV "strains" are unable to complement each other's replication functions (Choi and Stenger, 1995; Stenger, 1994) would seem to suggest that viruses known as strains of BCTV, in fact represent a number of distinct virus species. Rybicki (1994) and other authors have speculated, on the basis of phylogenetic analysis of viral sequences, that Subgroup II geminiviruses arose as a result of a recombination event between ancient Subgroup I and Subgroup III-like viruses.

The genomic organisation of geminiviruses from Subgroups I, II and III is illustrated in Figure 1.1. There is convincing evidence that geminiviruses replicate by a rolling circle mechanism (Saunders *et al.*, 1991; Stenger *et al.*, 1991). Specific viral proteins are involved in the initiation of rolling circle replication, *trans*-activation of the virion sense genes' promoters, production of ssDNA and viral movement functions. Examination and comparison of the different viral ORFs reveal several features which are conserved amongst all three genera of the *Geminiviridae*, and several which are subgroup-specific. In the interests of clarity, I have adopted a nomenclature system based on those proposed by Rochester *et al.* (1993), Rybicki (1994) and Bisaro (1996) for the discussion of functions of geminiviral gene products.

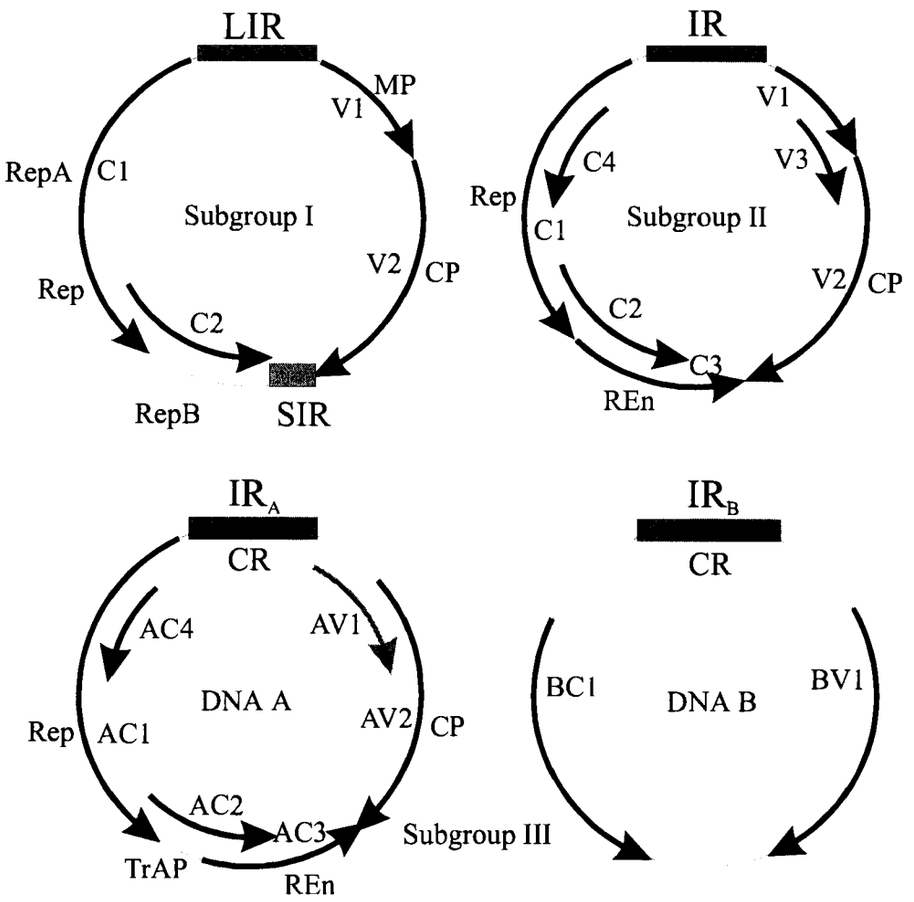


Figure 1.1: Genome organisation of geminiviruses.

The diagram depicts the double stranded replicative forms of consensus Subgroup I, II and III geminiviruses. Grey boxes indicate the intergenic regions (IR) which contain the origin of replication and transcription regulatory regions for bidirectional transcription. The part of the intergenic region which is identical in both Subgroup III genome components is called the common region (CR); as Subgroup I geminiviruses have two intergenic regions, the region which contains the origin of replication and viral gene promoter sequences is called the *long* intergenic region (LIR). The complementary strand origin of replication in Subgroup I geminiviruses is in the short intergenic region (SIR). Open reading frames (labels on inside of genome diagrams) are shown by arrows, which indicate the direction in which the ORF is transcribed. Where a gene's function is known, the name of the gene product is indicated (labels on the outside of diagrams). CP: coat protein; MP: movement protein; Rep: replication initiator protein; TrAP: virion-sense genes' transcriptional activator protein; REn: replication enhancer protein. In Subgroup II viruses, the C2 ORF does not seem to have transcriptional activator activity. An AV1 ORF is present in Subgroup III geminiviruses from the Old World. The (A)V2 ORF is indicated as the coat protein gene in all cases, whether or not an AV1 ORF is present.

All geminiviral genomes contain an intergenic region (IR) from which viral genes diverge in both the virion and complementary sense. The IR contains RNA polymerase II-type promoter sequences responsible for the transcription of genes in both genome senses. The

IR also contains the viral plus-strand origin of replication. All geminiviruses' IRs contain an inverted repeat sequence which is capable of forming a hairpin-loop structure which is required for replication (Orozco *et al.,* 1996). A conserved nonanucleotide sequence which is found in the loop of the hairpin-loop sequence of all geminiviruses has been shown to be the sequence where the replication initiator protein (Rep) induces a specific nick in the plus strand of the viral genome to initiate rolling circle replication (Heyraud-Nitschke *et al.,* 1995; Laufs *et al.,* 1995c; Stanley, 1995). The negative-sense origin of replication lies in the short intergenic region (SIR) of Subgroup I geminivirus genomes, where a short (approximately 80 nucleotides) DNA primer, with a few ribonucleotides at its 5′ end, binds and primes the synthesis of the complementary sense strand. This primer molecule is encapsidated within the virion (Andersen *et al.*, 1988; Donson *et al.,* 1984; 1987; Hayes *et al.,* 1988a; Morris *et al.,* 1992). A similar DNA primer has not been identified for Subgroup II or III viruses, but Saunders *et al.* (1992) presented evidence that second strand synthesis of ACMV is primed by an RNA primer which probably binds within the common region and which may be synthesised after uncoating of the viral DNA.

In general, genes encoded in the virion sense of monopartite geminiviruses and on the DNA A of bipartite Subgroup III viruses have functions in virus movement and encapsidation, while genes encoded in the complementary sense are involved in virus replication and *trans* activation of the virion sense promoter. One could therefore make a broad generalisation by saying that geminiviral "early" functions are mediated by the complementary sense genes, while the virion sense genes specify "late" functions. Viruses belonging to Subgroup III of the *Geminiviridae* are the only geminiviruses with a second DNA component (DNA B), which carries two ORFs, one in the virion sense (BV1) and one in the complementary sense (BC1). Both the *bv1* and *bc1* genes encode movement proteins.

Using the nomenclature convention of Rochester *et al.* (1993) and Rybicki (1994), the coat protein gene is encoded by the (A)V2 ORF. Sequence homology between the coat protein and the BV1 protein may imply that, in bipartite geminiviruses, BV1 has taken over at least some of the coat protein's functions in viral movement, but this does not necessarily obviate the coat protein's role (Rybicki, 1994).

In monopartite geminiviruses, other proteins involved in viral movement are encoded in an ORF which precedes the coat protein gene (V1); in the bipartite geminivirus TLCV-India, this ORF is required for efficient movement in *Nicotiana benthamiana*, even in the presence of DNA B (Paddidam *et al.*, 1996). However, in ACMV, viruses which carry mutations in the AV1 ORF have apparently wild-type infectivity (Etessami *et al.* 1989), at least in *N. benthamiana*. The V1 movement protein of Subgroup I geminiviruses as well as BC1 localises to the cell wall of infected cells, and is probably involved in modification of plasmodesmata to allow cell-to-cell movement of viral nucleoprotein complexes (Dickinson *et al.*, 1996; Nouiery *et al.*, 1994; Pascal *et al.*, 1993). The Subgroup II viruses BCTV and HrCTV have a further "precoat" ORF (V3), which appears, in the case of BCTV, to be involved in the regulation of ssDNA levels (Stanley *et al.*, 1992; Hormuzdi and Bisaro, 1993). This is a role which is played primarily by the coat protein in other geminiviruses. It is interesting to note that the other Subgroup II virus species, TPCTV, does not possess a V3 (ssDNA regulator) ORF (Briddon *et al.*, 1996).

All geminivirus genomes encode a replication initiator (Rep) protein. Rep has no DNA polymerase activity and only functions in initiating DNA replication, which otherwise is mediated by the DNA replication machinery of the host cell. The *rep* gene is encoded by the AC1 ORF in Subgroup III, the C1 ORF in Subgroup II and by both the C1 and C2 ORFs in Subgroup I geminiviruses. Rep is the only viral protein which is absolutely required for viral DNA replication in all genera of the *Geminiviridae*; the only *cis* DNA requirements for DNA replication are the intergenic regions. Geminivirus Rep proteins bear some distant relationship to replication initiator proteins of some single stranded DNA plasmids and of the ssDNA phage ϕX174 (Ilyina and Koonin, 1992). These observations have led to speculation that geminiviruses evolved from prokaryotic ssDNA replicons. Rigden *et al.* (1996) showed that the Rep protein of the Subgroup III geminivirus tomato leaf curl virus (TLCV-Australia) can initiate rolling circle replication of the viral genome in *Agrobacterium tumefaciens*. This lends support to the hypothesis of a prokaryotic origin of geminiviruses, and has already led to some interesting speculation on the involvement of *Agrobacterium*, or a related bacterium, in the original transfer of the progenitor geminivirus to plants.

The C1 and C2 ORFs of Subgroup I geminiviruses together constitute the *rep* gene; both ORFs are transcribed as a single RNA species (Accotto *et al.*, 1989; Dekker *et al.*, 1991;

Morris-Krsinich *et al.*, 1985; Mullineaux *et al.*, 1990). It seems that two different mature mRNA species are produced: one results from splicing of the C1 and C2 ORFs after the removal of a small intron which overlaps the end of the C1 ORF and the beginning of the C2 ORF, and the other is an unspliced RNA species. The spliced RNA is probably translated to produce the full length (41-kDa) Rep protein, while the unspliced RNA could be translated to produce the N-terminus of the Rep protein (RepA) only (Accotto *et al.*, 1989; Schalk *et al.*, 1989). The protein product of the (A)C3 ORF, found in Subgroup II and Subgroup III geminiviruses, significantly enhances the replication of the viral genome. Bisaro (1996) has therefore proposed that this protein be called the Replication Enhancer protein (REn). There is no REn protein homologue found in Subgroup I geminiviruses.

An interesting facet of control of geminivirus virion-sense gene expression is the transactivation phenomenon first reported in TGMV by Sunter and Bisaro (1991; 1992), which is mediated by the protein product of the AC2 ORF. Bisaro (1996) has proposed that this protein be called TrAP. It seems that the Rep protein of Subgroup I geminiviruses may play a similar role in transactivation of the coat protein promoter (Collin *et al.*, 1996; Hofer *et al.* 1992; Zhan *et al.*, 1993). However, the C2 ORF of BCTV does not seem to have TrAP activity (Hormuzdi and Bisaro, 1995) and as yet has no specifically ascribed functions.

The C4 ORF, which significantly overlaps the *rep* gene in dicot-infecting geminiviruses, presents some perplexing questions about its role in viral pathogenesis. Although an AC4 ORF is found in some bipartite geminiviruses, it seems to play no role in the virus life cycle, and may represent a vestigial or pseudogene (Etessami *et al.*, 1991; Pooma and Petty, 1996). In monopartite Subgroup III viruses and in Subgroup II geminiviruses, however, the C4 ORF plays an important role in symptom development (Jupin *et al.*, 1994; Stanley and Latham, 1992; Stanley *et al.*, 1992). Jupin *et al.* (1994) speculated that the C4 ORF plays a role in monopartite TYLCV movement; Gröning *et al.* (1994), on the other hand, found that it contributed to suppression of *rep* expression in TGMV, a bipartite virus. However, Pooma and Petty (1996) raise doubts as to whether the TGMV AC4 gene is functional and expressed at all.

In comparison with Subgroup I geminiviruses, Subgroup II and III viruses seem to have a far more complex set of genes and regulatory mechanisms. This may only reflect the fact

that these viruses have been more comprehensively investigated; nevertheless, it is clear that all members of the *Geminiviridae* have common "life" strategies, but in some cases Subgroup II and III viruses have gained extra genes to help them accomplish similar functions, perhaps in a more efficient manner. Subgroup I viruses have a simpler genomic organisation than Subgroup II and III viruses, with only three functional genes (compared with up to eight genes in some Old World Subgroup III viruses): *rep*, a movement protein gene and a coat protein gene. Subgroup I viruses therefore probably represent the most ancient genus in the *Geminiviridae* in evolutionary terms; this is borne out by the observation that there is more genetic diversity between Subgroup I viruses than between viruses in the other two genera, and that there is more diversity in the number of insect vector species which are involved in transmission of Subgroup I viruses, compared with a single vector species for Subgroup III (*Bemisia tabaci*) (Rybicki, 1994). In the next section, I review what is known about the molecular biology of Subgroup I geminiviruses. Since research on this genus has lagged behind that on Subgroup III viruses, where necessary I have drawn analogies from the latter to speculate on certain aspects of the molecular biology of Subgroup I geminiviruses.

1.2.2 The molecular biology of Subgroup I geminiviruses

Very early events in the life cycle

The first stage in the Subgroup I geminivirus life cycle involves its injection into a cell, presumably a phloem sieve element, by its leafhopper vector. It then presumably uncoats, and the virus ssDNA genome is somehow transported into the nucleus. Whether this is an active process, mediated by the coat protein or a host factor, or whether it is a passive process, is not known. The Subgroup III geminivirus BV1 protein, which is evolutionarily related to the coat protein, acts as a nuclear shuttle protein: it transports viral ssDNA in and out of the cell nucleus (Pascal *et al.*, 1994; Sanderfoot *et al.*, 1996). If Subgroup I virus coat protein fulfills at least some of the roles of BV1 in viral movement, then it is reasonable to suppose that transport of ssDNA into the nucleus is mediated by coat protein in Subgroup I geminiviruses.

Once in the nucleus, the ssDNA viral genome must be converted to the dsDNA replicative intermediate form (RF-DNA) , as this is the form which acts as a template for transcription

of viral genes and for replication. The presence of a virion-associated DNA primer molecule (Andersen *et al.*, 1988; Donson *et al.*, 1984; 1987; Hayes *et al.*, 1988a; Morris *et al.*, 1992) probably facilitates the host-mediated process of making the virus DNA double stranded. Like the small circular dsDNA papovaviruses, geminivirus RF-DNA is supercoiled, associated with histones and thus packaged into "minichromosomes" (Abouzid *et al.*, 1988; Pilartz and Jeske, 1992). Although there have been no reports of Subgroup I geminivirus RF-DNA in chromatin-like structures, it is reasonable to assume that the replicative intermediate form of all geminiviruses is packaged into minichromosomes, as this would be consistent with geminiviruses' absolute dependence on host DNA replication and transcription machinery.

Transcription of viral genes and RNA processing

The first priority for a Subgroup I geminivirus once it has entered the host cell nucleus and been transformed into its dsDNA replicative intermediate form, is to express the single gene involved in the early phase of its life cycle (the *rep* gene), so that it may generate a high enough titre of its genome to initiate a systemic infection. It is probably not to the virus's advantage at this stage to express the virion sense genes to any significant level, as these genes do not seem to be involved in replication. Moreover, their expression may well interfere with viral replication, if their protein products are involved in sequestration of ssDNA and in moving the virus genome out of the nucleus, whether as single stranded or dsDNA.

There are transcript maps available for three Subgroup I geminiviruses: MSV (Morris-Krsinich *et al.*, 1985), DSV (Accotto *et al.*, 1989; Mullineaux *et al.*, 1990) and WDV (Dekker *et al.*, 1991). I should point out that none of these studies took into account that different transcripts might be present in different tissues, representing genes expressed at different stages of the virus life cycle. Rather, RNAs were isolated from expanded leaf tissues, several days after inoculation. Consequently, the transcript maps probably represent an approximation of the RNA species which are present at various stages of the virus infection of the plant; in fact, the RNA preparations used were probably enriched for "late" transcripts.

Consensus promoter sequences for the C1 and V1 ORFs have been identified in the long intergenic regions of all Subgroup I geminiviruses sequenced to date (Timmermans *et al.*, 1994). In Figure 1.2, I have presented a consensus transcript map, based on the maps of MSV, DSV and WDV. According to these, several different complementary sense transcripts may be produced in all three viruses. Due to low levels of RNA present, however, some of these could represent S1 nuclease mapping artifacts (Mullineaux *et al.*, 1990). In WDV two different complementary sense transcripts were 5′ co-terminal; one was spliced, and the other unspliced (Dekker *et al.*, 1991). These authors did note some heterogeneity in the 3′ termini of the WDV complementary sense transcripts, with some minor unspliced transcripts which appeared to be truncated, and thus capable of expressing only RepA. Mullineaux *et al.* (1990) presented similar results for DSV complementary sense transcripts.

The published transcript map of MSV (Morris-Krsinich *et al.*, 1985) showed only one complementary sense RNA of 1.2 kb; the 5′ end of this transcript was roughly mapped to near nt 2360 on the MSV-N genome, with its 3′ end at 1160. Since this transcript maps downstream of the ATG of the C1 ORF, the authors proposed that the C1 protein product would be expressed from an ATG 99 bp downstream of the C1 start codon. Subsequent transcript mapping of MSV-N has shown that the transcript map for the complementary sense published in 1985 was incorrect. There are three potential TATA boxes which could constitute part of the MSV *rep* gene promoter, one at position -101 relative to the *rep* ATG, and two which almost overlap at -57 and -62, respectively. Results from the MSV laboratory at the John Innes Centre show two transcripts for the complementary sense gene of MSV; one of approximately 1.2 kb and the other of 1.5 kb. As found for WDV and DSV, the shorter transcript terminates in the C2 ORF and could only result in translation of RepA. The 1.5-kb transcript spans the entire Rep gene, and initiates mainly from the -101 TATA box. On the other hand, the 1.2-kb transcript probably initiates mainly from the -57 or -62 TATA box. So, there are not one, but two transcription start points for the *rep* gene (M.I. Boulton, personal communication; Wright, 1995).

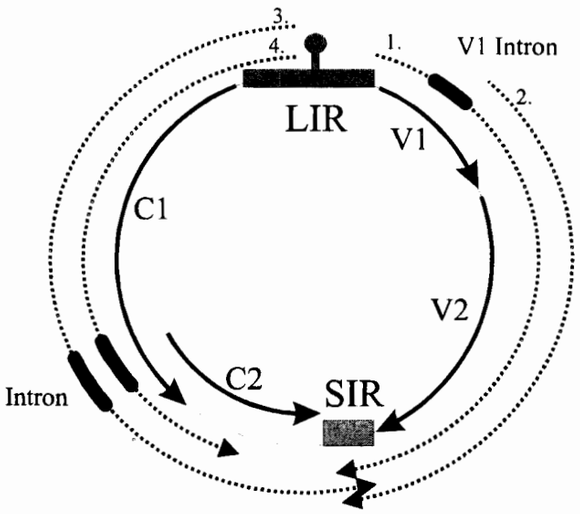


Figure 1.2: Consensus transcript map of Subgroup I geminiviruses

The four predominant RNA species are indicated. Other minor transcripts have also been mapped, but may represent S1 nuclease mapping artifacts. Transcript 1 is present in WDV, DSV and MSV. Transcript 2 is present in DSV and MSV, not WDV. This may represent the spliced version of transcript 1. Transcript 3 is the *rep* gene transcript, which may be spliced or unspliced. The spliced version could be translated to produce Rep, while the unspliced version can only translate RepA. There may be a minor transcript (4) which terminates within the C2 ORF and encodes RepA or truncated Rep only. The existence of transcript 4 is not certain for all Subgroup I geminiviruses. In MSV, transcripts 3 and 4 can be initiated from 3 different TATA boxes in the LIR; at -101, -57 and -62.

It is unlikely that any autonomous protein encoded by the C2 ORF is expressed, as there has been no transcript mapped to this region, and MSV is the only Subgroup I geminivirus which has an ATG start codon for this ORF. The C2 ORF is, therefore, probably only expressed as the carboxyl terminus of Rep, from the spliced *rep* transcript. Schalk *et al.* (1989) and Accotto *et al.* (1989) independently demonstrated the presence of the intron in the *rep* gene of WDV and DSV, respectively. WDV required splicing for replication, and a mutant virus with the native *rep* gene replaced with a cDNA copy replicated efficiently in *Triticum monococcum* protoplasts (Schalk *et al.*, 1989). This implies that the full length Rep protein, expressed from a spliced mRNA, is both essential and sufficient for replication, in protoplasts at least. These results were corroborated by Collin *et al.* (1996). There have been no reports of whether these mutant "intronless" viruses are infectious in plants.

The presence of an intron in the *rep* gene is a common feature of all Subgroup I geminiviruses sequenced thus far, including TYDV (Morris *et al.*, 1992). All of the investigations into the processing of the *rep* intron of cereal-infecting geminiviruses (Schalk *et al.*, 1989; Accotto *et al.*, 1989; Dekker *et al.*, 1991 and Mullineaux *et al.*, 1990) have shown that this intron is rather inefficiently processed, resulting in an unusually high proportion of unspliced to spliced mature mRNAs in infected tissues. These results could be interpreted as indicating that the *rep* gene transcript(s) are inefficiently spliced to facilitate production of both Rep and RepA. It is interesting to note, however, that Morris *et al.* (1992) found that the TYDV *rep* intron was efficiently spliced in infected tobacco and bean leaves.

The absence of good consensus TATA box elements near the start of the coat protein open reading frame of many Subgroup I geminiviruses, including DSV, MSV and WDV, presents somewhat of a mystery. Of the three Subgroup I viruses with transcript maps, both "African streak"-group viruses (MSV and DSV) seem to express two 3′-co-terminal transcripts: a major, approximately 0.9-kb, transcript from which coat protein could be translated and a longer, less abundant, ~1.05-kb transcript from which the V1 (movement protein) could be expressed (Accotto *et al.*, 1989; Morris-Krsinich *et al.*, 1985). In contrast, only a long, bicistronic, virion sense transcript is present in the WDV transcript map (Dekker *et al.*, 1991).

Recently, Heckel (1996) reported the presence of another intron in the MSV genome, within the V1 ORF (M. I. Boulton, personal communication). These authors also found putative intron sequences within the V1 ORF of other Subgroup I geminiviruses: this intron interestingly spans the region of the V1 ORF which encodes the "transmembrane" domain of the Subgroup I geminivirus movement protein (Boulton *et al.*, 1993). It is significant that the 3′ splice site of the V1 intron coincides exactly with the 5′ end of the shorter virion-sense transcript identified for both MSV and DSV (personal observation; Accotto *et al.*, 1989; Morris-Krsinich *et al.*, 1985). On this basis, I would predict that the shorter transcripts do not exist, but rather represent an artifact of the S1 nuclease mapping procedure: the authors in both cases may have identified an S1 nuclease-sensitive site which represents the point where a gapped duplex started between the protecting DNA fragment and the spliced mRNA. The possibility therefore exists that as for WDV, MSV and DSV express both the movement protein and coat protein genes from the same

bicistronic RNA transcript. Whether there is specific control of the splicing efficiency is an interesting question which will hopefully be addressed in the future. An additional point to be raised here is that Boulton and colleagues have not detected transcript splicing in WDV (or TYDV), and no smaller V2 "transcript" was detected in the transcript mapping of this virus, although the intron and splice donor and acceptor sites are still present in the WDV sequence (Heckel, 1996; M.I. Boulton, personal communication). These observations would lend support to the proposal that the MSV and DSV 0.9-kb transcripts represent S1 nuclease mapping artifacts as a result of splicing of the longer RNA species. This group also could not find V1 intron processing in TYDV, but there is no transcript map available for this virus species.

The discovery of the V1 intron now raises the issue of how the coat protein ORF is expressed from a biscistronic RNA. Unlike the situation in MSV and DSV, the V1 and V2 ORFs of WDV overlap somewhat. Dekker *et al.* (1991) found that the pattern of codon usage in the region where the two ORFs overlap could indicate that frameshifting occurs in this region to allow the translation of the coat protein ORF. There are two adjacent rare codons in the 3′ end of the movement protein ORF, immediately 5′ to the start codon of the coat protein ORF, as shown in Figure 1.3. The authors (Dekker *et al.*, 1991) postulate that the low availability of tRNAs for the AGA (arginine) and UUA (leucine) codons in the V1 ORF could induce a +2 frameshift to allow the translation of V2. This hypothesis requires that the two ORFs then be cleaved by an endopeptidase (Dekker *et al.*, 1991). An alternative could be that the coat protein ORF is translated by a "relay race" type of ribosomal frameshifting, as proposed for cauliflower mosaic virus (CaMV; Dixon and Hohn, 1984).

Processing of the virion sense RNA would remove part of the V1 ORF, so that the movement protein could not be expressed from the spliced transcript. However, the movement protein gene start codon would still be present in the spliced RNA. The transcript maps of both MSV and DSV show that the "longer" virion sense transcript starts very close to the ATG of the movement protein gene: between 1 and 4 nt. This is an exceptionally short RNA leader for a eukaryotic gene. It is probably significant that there are no ATG triplets between the movement protein initiation codon and the coat protein ATG in the sequences of MSV, DSV or any other "African streak" geminivirus (personal observation). Therefore, if ribosomes frequently miss the first ATG in the transcript

because it is too close to the 5′ end of the mRNA, the coat protein *ATG would be the first* to be detected, if the "ribosome scanning" model for initiation of translation holds. This may explain how the coat protein is expressed from a bicistronic RNA, and the observation that coat protein is generally expressed at a much higher level than movement protein in infected cells (Mullineaux *et al.*, 1988). There may also be a virion sense transcript which initiates further upstream than the -1 transcript of MSV, which could be the main movement protein messenger (M.I. Boulton, personal communication).

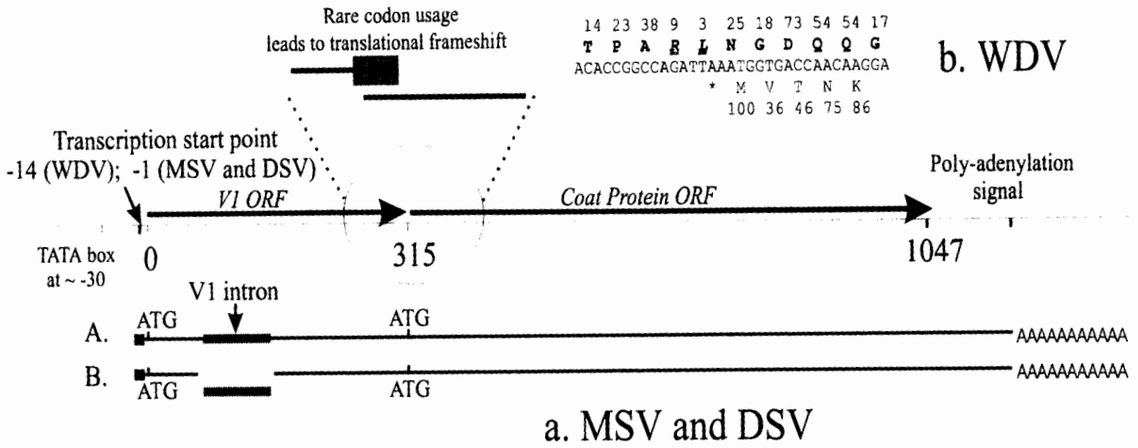


Figure 1.3: Schematic representation of the strategy for expression of the virion sense genes used by Subgroup I geminiviruses MSV, DSV and WDV.

A. and B. represent unspliced and spliced transcripts, respectively. The putative frameshifting strategy employed by WDV (b.) is depicted above the genome diagram; the region of overlap between the movement protein and coat protein genes in WDV is indicated, with rare codons at the C-terminus of the movement protein italicised and underlined. Numbers above amino acids indicate frequency of usage of the codon in genes from monocotyledonous plants. The splicing strategy probably used by MSV and DSV is shown in (a.) below the representation of the Subgroup I geminivirus virion sense genes. Genome co-ordinates shown are for MSV-N.

The discovery of an interesting reduced-fitness variant of the Nigerian MSV isolate has given Margaret Boulton and colleagues some useful insight into MSV gene expression. The genome of this mild isolate, termed MSV-Nm, had three nucleotide substitutions compared with the severe isolate (MSV-Ns) (Boulton *et al.*, 1991a; b). Two of the three mutations seemed to be responsible for the reduced symptom severity and narrow host range phenotype observed in MSV-Nm. The first of these mutations was an A→G transition at nt 2473 in the LIR; this mutation altered the -101 TATA box associated with the *rep* gene, and seemed to result in increased levels of the 1.2-kb complementary sense

transcript relative to the 1.5-kb transcript (M.I. Boulton, personal communication). The other mutation, a transition mutation from C→T at nt 40, affected the efficiency of splicing of the V1/V2 transcript. In this mutant, the intron was *more* efficiently processed than in the wild type. This was a very surprising observation, as nt 40 lies outside the intron sequence, and this is the first report of a single nucleotide substitution in an exon affecting the efficiency of splicing of an RNA transcript (M.I. Boulton, personal communication). More efficient splicing of the virion sense transcript would result in *lower* levels of the intact movement protein being produced; this is consistent with the narrow streak phenotype associated with the nt 40 mutation in MSV-Nm. The inclusion of introns in chimaeric genes constructed for expression in cereals, particularly in the 5′ untranslated leader sequence, has been shown to greatly enhance the expression of the gene (e.g. Callis *et al.*, 1987; Luehrsen and Walbot, 1991). In this light, it would be of great interest to determine if there is any enhancement of coat protein expression associated with the presence of the V1 intron discovered by Boulton and colleagues.

Replication of Subgroup I geminiviruses

It is clear that expression of geminivirus genes is inextricably linked to viral replication, so it is appropriate that I now introduce what is known about the replication of Subgroup I geminiviruses, in order to discuss the control of viral gene expression in finer detail. Based on computer analyses of common region or long intergenic region sequences, together with published observations on Rep-intergenic region interactions, Argüello-Astorga *et al.* (1994a; b) proposed a widely accepted model for the mechanism of geminivirus replication, integrated with control of gene expression. This model was based almost entirely on literature on replication of Subgroup III geminiviruses, so the authors' extrapolation of the model to include Subgroup I viruses was not based on much experimental evidence. Since the proposal of the model by Argüello-Astorga *et al.* (1994a; b), a considerable body of data on biochemical aspects of replication of Subgroup III geminiviruses has been published, and has resulted in a refined model for replication, proposed by Hanley-Bowdoin *et al.* (1997).

Rep Protein Structure and Function

It is now well established that geminiviruses replicate by a rolling circle replication (RCR) mechanism, analagous to that used by eubacterial ssDNA plasmids and ssDNA phages.

The first step in the geminivirus replication cycle–conversion of the ssDNA genome into the dsDNA replicative intermediate–is fairly poorly characterised, although it is likely that the primer-like molecule associated with Subgroup I geminivirus virions participates in this process. The second stage of geminivirus replication–production of single stranded circular DNA from the RF-DNA–has been extensively studied in recent years (reviewed by Bisaro, 1996; Hanley-Bowdoin *et al.*, 1996 and Laufs *et al.*, 1995a). The geminiviral Rep protein shares the property of being a sequence specific DNA binding protein with site specific cleavage and joining activity, with the RCR initiator proteins of prokaryotic ssDNA replicons (del Solar *et al.*, 1993; Ilyina and Koonin, 1992; Koonin and Ilyina, 1992; 1993). The RCR mechanisms employed by many different ssDNA replicons bear striking similarities. In all cases, the replication initiator protein binds RF-DNA at specific sequences, and nicks the plus strand at a specific point. This process is strictly dependent on the replication initiator protein, but all other enzymes involved in plus strand synthesis are supplied by the host.

Both Laufs *et al.* (1995c) and Stanley (1995) proved that the point of Rep-mediated nicking in the geminivirus loop sequence is between nucleotides 7 and 8 of the conserved nonanucleotide sequence, i.e. at TAATATT$\downarrow^{8}$AC. Using a biochemical approach, Laufs *et al.* (1995c) showed that TYLCV Rep expressed in *E. coli* could both cleave and join oligonucleotides which contained the canonical TAATATTAC motif. Interestingly, TYLCV could also cleave oligonucleotides with WDV origin of replication sequences, and vice versa, albeit at reduced frequency (Heyraud-Nitschke *et al.*, 1995; Laufs *et al.*, 1995c). It is perhaps significant that Rep was unable to cleave a dsDNA origin of replication *in vitro*, as this points to the possibility that extrusion of a stem loop structure which allows presentation of a ssDNA cleavage substrate in RF-DNA, is required. This would support the findings of Orozco and Hanley-Bowdoin (1996) that formation of a stem-loop structure is required for replication of Subgroup III geminiviruses. However, Laufs *et al.* (1995c) could not rule out possible problems in their experimental protocol being responsible for their observations on the lack of Rep-mediated cleavage of its dsDNA substrate.

After nicking the origin of replication, the Rep protein becomes covalently linked to the 5′ end of the nicked strand, via a phosphotyrosine linkage. In the geminivirus case, Rep would be linked to the last A nucleotide of the TAATATTAC nonanucleotide; it remains linked there throughout DNA polymerisation. The 3′ -OH of the last T in the

nonanucleotide is then available to prime the synthesis of a new plus strand, using the negative strand as a template, and displacing the old plus strand as it is synthesised. On completion of one full replication cycle, the origin of replication is reconstituted. There are now two options for how the replication cycle is resolved to release a single stranded circular DNA molecule.

The first model is based on the mechanism used by phage ϕX174; this requires two tyrosines in close proximity to each other to achieve resolution and release of a circular ssDNA molecule. As Rep has only one tyrosine in its active site, this implies that Rep functions as at least a dimer, which would support the finding of Settlage *et al.* (1996), that TGMV Rep forms oligomers. After one round of synthesis, the newly synthesised nonanucleotide sequence is cleaved again by a second Rep molecule associated with the one which had initially nicked the origin of replication; this second Rep becomes linked to the 5′-AC end of the new nonanucleotide. The 5′ phosphoryl group of the displaced strand is then transferred from the first Rep molecule to the newly generated 3′-OH group to liberate a circular ssDNA molecule (Laufs *et al.*, 1995a; b; c). Thus, an active Rep molecule always remains attached to the DNA via alternating tyrosines.

The second model for resolution is similar to one elucidated for rolling circle plasmid pC194 (Noirot-Gros *et al.*, 1994). This DNA replication system is, unlike the ϕX174 one, non-continuous. One active tyrosine is sufficient to effect nicking and resolution of circular ssDNA. As in the model described first, Rep becomes linked to the 5′ end of the cleaved DNA via a phophotyrosine linkage. However, after one round of DNA synthesis, the release of the circular single strand is mediated by the gamma carboxylate group of another amino acid. The newly synthesised origin is then cleaved in a nuclease-like reaction, and the 5′ end which was previously linked to the Rep active site tyrosine is transferred to the newly created 3′-OH end (Laufs *et al.*, 1995b; Noirot-Gros *et al.*, 1994). Thus, the termination step of the pC194 replication cycle renders it non-continuous, as Rep will need to re-initiate replication at an origin of replication. Both replication systems initiate rolling circle replication in the same way, but their mechanisms of termination differ. It is not yet clear which model of RCR termination occurs in geminiviruses.

The domain of Rep responsible for cleavage of the nonanucleotide motif appears to be in the N-terminus of Rep of both Subgroup I (WDV) and Subgroup III (TYLCV) geminiviruses (Heyraud-Nitschke *et al.*, 1995). A fusion protein with RepA of WDV functions in both origin cleavage and joining, as does a 24-kDa C-terminal truncated version of TYLCV Rep. In addition, both Jupin *et al.* (1995) and Choi and Stenger (1995; 1996) have shown that the N-terminal domain of Rep of Subgroup II and III geminiviruses is also involved in specific origin recognition. Jupin *et al.* (1995) mapped this region to the N-terminal 116 amino acids of TYLCV, while Choi and Stenger (1995; 1996) showed that 89 amino acids at the N-terminus of BCTV Rep allows Rep to recognise its specific DNA binding sequences in the origin of replication. It will be interesting to determine whether the N-terminus of the Subgroup I geminivirus Rep protein is also involved in sequence-specific DNA binding.

Geminivirus Rep proteins share four highly conserved amino acid motifs: motifs I to III are conserved in RCR replication initiator proteins (Ilyina and Koonin, 1992; Koonin and Ilyina, 1992) while motif IV is a putative nucleoside triphosphate (dNTP) binding site (P-loop) commonly found in kinases and DNA helicases (Gorbalenya and Koonin, 1989). The function of motif I, with a completely conserved "FLTY" signature, is as yet unknown. Motif II may be involved in metal ion coordination (Koonin and Ilyina, 1992); this possibility is supported by the finding of Laufs *et al.* (1995c) that Rep-mediated covalent linkage of Rep-cleaved oligonucleotides *in vitro* required Mg^{2+} or Mn^{2+} ions. Motif III contains a conserved tyrosine residue which participates in phosphodiester bond cleavage and in the covalent linkage of Rep to the 5′ terminus of the nicked nonanucleotide motif (TAATATT $\downarrow^{8}$AC) (Laufs *et al.*, 1995b). Desbiez *et al.* (1995) showed that TYLCV Rep has ATPase activity associated with dNTP binding motif IV; this hints at the possibility that Rep has helicase activity. Mutational analysis shows that the putative NTP-binding domain in motif IV is required for geminivirus replication (Desbiez *et al.*, 1995; Hanson *et al.*, 1995), as is the nicking motif (III) (Hoogstraten *et al.*, 1996).

The recently discovered ability of geminiviruses, like DNA tumour viruses, to modify the host cell cycle has raised a great deal of interest for the potential that these viruses now hold for elucidating mechanisms of plant DNA replication and cell cycle control. Nagar *et al.* (1995) found that TGMV could cause the accumulation of proliferating cell nuclear antigen (PCNA) in terminally differentiated cells, where PCNA is not normally detected.

Studies with plants transgenic for TGMV Rep confirmed that this effect was Rep-mediated. PCNA is a DNA replication processivity factor which is associated with the S-phase of the cell cycle; this implies that Rep induces DNA replication machinery in cells which would not otherwise be undergoing DNA synthesis. These observations agree with those of Accotto *et al.* (1993), who found that RF-DNA forms of Subgroup I geminivirus DSV were much more abundant in S-phase nuclei than in G-phase nuclei. However, Lucy *et al.* (1996) were able to detect *rep* gene transcripts, probably indicative of replication, in cells in which transcripts of the S-phase specific gene histone H2b were definitely absent. This finding suggests that host DNA replication is not absolutely necessary for MSV replication.

Xie *et al.* (1995) showed that WDV Rep was able to bind human retinoblastoma protein (Rb), a member of a protein family which controls cell cycle progression by sequestering transcription factors necessary for entry of the cell cycle into S-phase. This is a function assumed by DNA tumour virus proteins like SV40 large T antigen (T-Ag), adenovirus E1A and human papillomavirus type 16 (HPV-16) E7 protein to modify the cellular environment to one that allows viral DNA replication. It is logical that geminiviruses, which depend on host DNA replication enzymes to effect viral DNA replication, should use a similar strategy to be able to replicate in plant cells which would not normally be undergoing DNA replication. The viral oncoproteins E1A, T-Ag and E7 all interact with Rb through the Rb binding motif LeuXCysXGlu (LXCXE), where X represents any amino acid. Xie *et al.* (1995) identified Rb binding motifs in Rep proteins of all Subgroup I geminiviruses except SSV. According to results presented by Xie *et al.* (1995) and Collin *et al.* (1996), the Rb binding motif is essential for WDV replication. In this light, it is worthwhile noting that the clone of SSV which lacks an intact LXCXE motif is not infectious (Hughes *et al.*, 1993). The position of the Rb binding domain in MSV as a representative Subgroup I geminivirus, is shown in Figure 1.4.; it is present in both Rep and RepA (Collin *et al.*, 1996; Xie *et al.*, 1995).

Although there is no LXCXE motif in Rep of Subgroup II or III geminiviruses, unpublished data from Linda Hanley-Bowdoin's laboratory show that TGMV Rep nonetheless interacts with a plant Rb homologue (Ach *et al.*, in preparation, cited in Hanley-Bowdoin *et al.*, 1997). Further exciting developments in the Rep-Rb interaction saga are independent reports from two groups (Grafi *et al.*, 1996 and Xie *et al.*, 1996) of

cloning of an Rb homologue from maize, and identification of this protein as a *bona fide* member of the retinoblastoma protein family based on its ability to interact with WDV Rep. As yet, there are no data available which correlate expression of Subgroup I virus Rep with the accumulation of proteins like PCNA in infected cells, but it is likely that Subgroup I virus Rep will have a similar effect as TGMV Rep on this, and possibly other proteins implicated in progression of the cell cycle from G_1 into S phase. It is not clear whether the TGMV Rep-mediated induction of PCNA is due to its interaction with Rb or a related protein, or due to a direct effect of TGMV Rep as a transcription factor.

There are presently no reports of any transcription activation activity associated with the Rep proteins of Subgroup II or III geminiviruses. However, the Rep proteins of Subgroup I do seem to have an additional role as a transcription factor. WDV and CSMV Rep and/or RepA may induce expression of the coat protein promoter (Hofer *et al.*, 1992; Zhan *et al.*, 1993; Collin *et al.*, 1996). Rep proteins of all geminiviruses show some homology to the DNA binding domain of the *myb*-related class of plant transcription factors (Hofer *et al.*, 1992); this domain is encoded in the C2 ORF in Subgroup I viruses. Hofer *et al.* (1992) showed that this region may be required for Rep transactivation of the coat protein promoter of WDV, but it is not known whether Rep exerts transcription factor activity on the promoters of any cellular genes. I have summarised the main structural features of Subgroup I geminivirus Rep proteins in the diagram of MSV Rep in Figure 1.4.

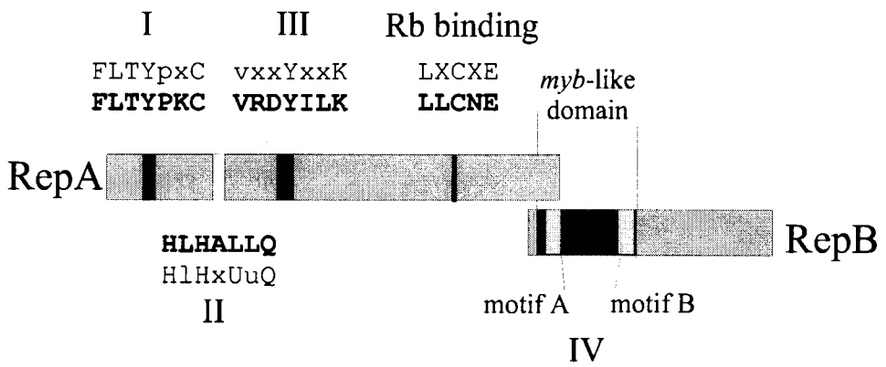


Figure 1.4: Conserved motifs in Rep of MSV-Kom
Consensus amino acid motifs are indicated for motifs I (FLTY), II (putative metal ion co-ordination), III (nicking motif), and the Rb binding domain. I have also indicated the actual amino acid sequence in the MSV-Kom Rep protein for these motifs. The position of motif IV, the dNTP binding domain (P-loop), in RepB is indicated, as is the domain of RepB which shares homology with plant *myb*-like transcription factors.

The structure of the large intergenic region of Subgroup I geminiviruses and its interaction with Rep.

The LIR of Subgroup I geminiviruses must contain promoter elements for expression of both virion sense and complementary sense genes, binding sites for plant nuclear factors involved in gene expression and DNA replication, specific sequences which constitute Rep binding sites, and the stem-loop sequence which functions as the plus strand origin of replication. The Subgroup I geminivirus Rep binding site/s are equivocally determined, but are presumed to be part of certain iterative elements or "iterons" identified by Argüello-Astorga *et al.* (1994a; b). These authors proposed a plausible and widely accepted model for geminivirus replication, integrated with control of viral gene expression; the model has since been expanded upon to incorporate a function for REn (AC3) protein in a review of Subgroup III geminivirus replication by Hanley-Bowdoin *et al.* (1997). The factors involved in interaction between Rep and the LIR of Subgroup I geminiviruses are poorly characterised, so I will explain the currently accepted model for replication of Subgroup III geminiviruses, and then point out the parallels and differences between the Subgroup III and Subgroup I situations.

Lazarowitz *et al.* (1992) mapped the minimal origin of replication of the Subgroup III geminivirus, squash leaf curl virus (SqLCV), to a ~90-bp fragment which includes the conserved stem-loop sequence and about 60 bp of the common region proximal to the *rep* gene. No sequence on the right hand side of the stem loop was required for replication. The Subgroup III geminivirus Rep binding site was localised to a 52-base pair sequence on the left hand side of the TGMV common region by Fontes *et al.* (1992). This group later identified the sequence of the Rep binding site to be a direct repeat of a specific 5-bp sequence, separated by a spacer of two or three nucleotides (Figure 1.5; Fontes *et al.*, 1994a; b).

At the time it was surprising to discover that the Rep-binding site was distinct from the stem-loop sequence; in TGMV these two sites are separated by 34 bp, but this distance is highly variable between different viruses (23 to 82 bp). The high affinity Rep binding site is, in all Subgroup III geminiviruses, situated between the *rep* gene TATA box and the stem-loop sequence (Figure 1.5; Argüello-Astorga *et al.*, 1994a; b). The presence of this Rep binding site is necessary for the negative autoregulation effect which Rep exerts on the

expression of its own gene; this probably occurs through some sort of interference with assembly or activity of the transcription preinitiation complex (Eagle *et al.*, 1994).

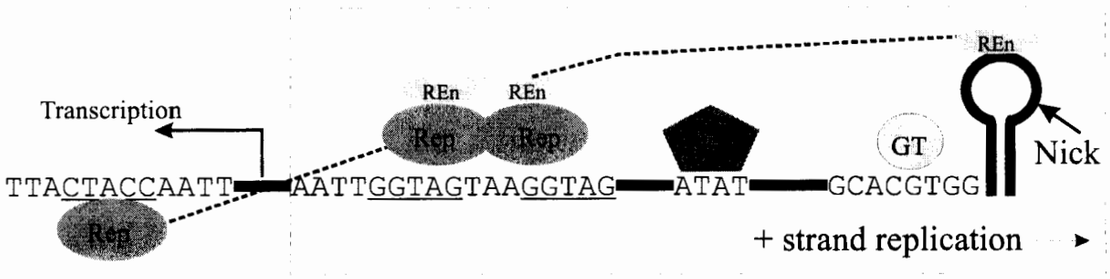


Figure 1.5: Model for initiation of plus strand replication of Subgroup III geminiviruses.

Redrawn from Hanley-Bowdoin *et al.* (1996). The top strand of the origin is shown with the minimal origin outlined by the dotted box. The sequences designate protein binding sites. The initiation sites and directions of replication and transcription are marked. Proteins predicted to interact with the origin are shown (Rep, REn, TATA binding protein or TBP, G-box transcription factor or GT.) Rep and REn protein interactions are indicated, including potential long range interactions (dotted lines) that might change origin structure.

In a computer analysis of the intergenic regions of different Subgroup III geminiviruses, Argüello-Astorga *et al.* (1994a; b) found that the virus-specific Rep binding sites identified by Fontes *et al.* (1994a; b) formed part of sequences which were iterated a number of times within the viral common regions of all Subgroup III geminiviruses. They called these sequences "iterons". The arrangement (but not the sequence) of the iterons is highly conserved within the common regions of dicot-infecting geminiviruses, including BCTV, with slightly different iteron arrangements found between Old World and New World viruses (Argüello-Astorga *et al.*, 1994a). On the basis of their sequence homology with the high affinity Rep binding site near the stem-loop sequence, it is presumed that all of the other iteron sequences also interact with Rep, perhaps to mediate transcriptional control of the *rep* gene. It is likely, but unproven, that Rep binding causes extrusion of the stem-loop sequence in dsDNA, as occurs on binding of replication initiator proteins of ssDNA plasmids like pT181 (Noirot *et al.*, 1990). Laufs *et al.* (1995c) presented evidence that Rep only cleaved single stranded, not dsDNA oligonucleotides with origin of replication sequences, which also points to a requirement for extrusion of stem-loop sequences, if not for nicking, at least for termination of RCR (Heyraud *et al.*, 1993a). In addition, strong support for the necessity for a stem-loop structure was provided by Orozco and Hanley-Bowdoin (1996), who found, by replacing bases in the stem, that the ability to form the

stem-loop structure was essential for replication and the sequence of the stem only contributes to the efficiency of replication.

Argüello-Astorga *et al.* (1994a; b) speculated that there may be involvement of plant transcription factors in facilitating initiation of replication. Interaction between a host transcription factor bound at a G-box homologous site near the stem-loop sequence and a TATA box binding protein (TBP) might induce sequence looping, which brings the Rep complex bound at the high affinity binding site in contact with the loop sequence to initiate RCR (see Figure 1.5. and Argüello-Astorga *et al.* (1994b). There must also be some involvement of the viral replication enhancer (REn) protein in this process.

The mechanism of action of the protein encoded by the Subgroup II and III geminivirus (A)C3 ORF (REn), has not been fully elucidated. While viruses which carry mutations in the AC3 ORF are still infectious, symptoms are greatly attenuated (Etessami *et al.*, 1991). In transient replication assays, REn enhances replication approximately 50-fold (Sunter *et al.*, 1990). It is known that this protein's function is not virus-specific, as complementation of REn function by heterologous viruses (from both Subgroups II and III) is possible in protoplasts (Sunter *et al.*, 1994). Their findings that REn forms homo-oligomers and also interacts non-specifically with Rep from different viruses, has led Hanley-Bowdoin and her co-workers to propose a mechanism by which this protein might act in initiation of viral plus strand replication (Hanley-Bowdoin *et al.*, 1997). The sequence of the stem-loop region of Subgroup II and III geminiviruses is highly conserved, unlike Subgroup I viruses. If REn binds specifically to the stem-loop sequence, then its ability to interact with Rep would provide an elegant mechanism for recruiting Rep to the nonanucleotide cleavage-joining substrate in the loop sequence. This hypothesis would answer questions about REn's apparent lack of virus-specificity, and at the same time provide an explanation for the fact that the Rep binding sites do not coincide with the origin of replication. As yet, however, there is no proof for specific REn interaction with the stem-loop, let alone evidence for its recognition of the stem-loop being on the basis of sequence or structure. Nevertheless, the hypothesis is an attractive one, and is not exclusive of the proposal of transcription factor and TBP involvement.

While there are some parallels that one could draw between the structure of (long) intergenic regions of Subgroup I, II and III geminiviruses, there are also some striking

differences between LIRs of Subgroup I viruses and those of viruses from the other two genera. In Figure 1.6, I have shown the sequence of the MSV-Kom LIR, from the initiation codon of the *rep* gene to the initiation codon of the V1 ORF, and have identified the main elements which may be of importance in origin function. MSV-Kom is an isolate of MSV, sequenced in this laboratory (M. D. James, F. L. Hughes and E. P. Rybicki, unpublished) and is most closely related to MSV-SA (Lazarowitz, 1988).

As expected, the conserved geminivirus loop sequence is essential for replication; no insertions or deletions are tolerated, but a single transition mutation to TAATA**C**TAC has only a minor effect on virus fitness in MSV (Schneider *et al.*, 1992). As yet, no Rep-binding sites in the LIRs of any Subgroup I geminiviruses have been identified. Functional Rep protein from WDV has been expressed in *E.coli* (Heyraud-Nitschke *et al.*, 1995), so it should be fairly straightforward to identify the regions where Rep interacts with the WDV LIR, perhaps by similar methods to those employed to determine specific Rep binding sites in TGMV (Fontes *et al.*, 1992; 1994a; b). However, Argüello-Astorga *et al.* (1994a) identified iteron sequences in the LIRs of eight Subgroup I geminiviruses which they hypothesised must represent Rep binding sites, by analogy with Subgroup III geminiviruses. The positions of the MSV-Kom iterons in the viral LIR are shown in Figure 1.6.

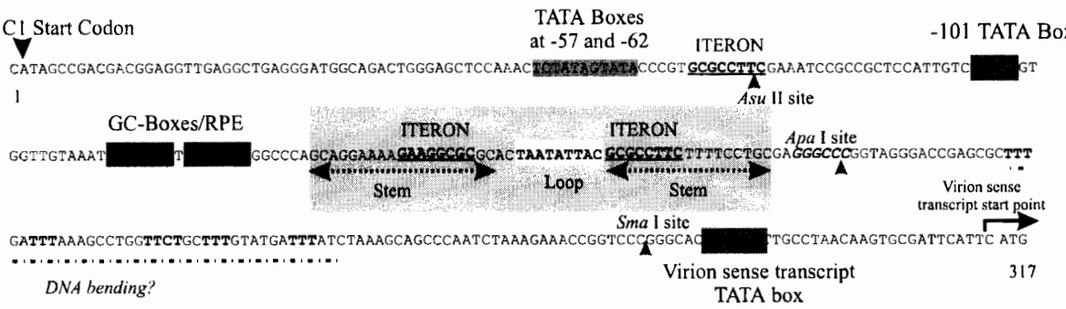


Figure 1.6: Functional organisation of the Long Intergenic Region of MSV

The conserved stem loop structure is shown boxed. I identified the iteron sequences in this virus, according to Argüello-Astorga *et al.* (1994a; b), and have indicated them underlined, and in boldface text. TATA boxes which may be part of virion and complementary sense promoters are indicated. The GC-box element shown by Fenoll *et al.* (1990) to be part of the *rightward promoter element* (RPE) is also labelled. A region of DNA which was shown to be involved in DNA bending in WDV is shown in MSV-Kom; however, the sequence does not seem to fit the requirements for DNA bending as well as WDV does (Suárez-López *et al.*, 1995). In addition I have shown the positions of two restriction sites (*Asu* II and *Apa* I) which Schneider *et al.* (1992) used insertional mutagenesis to determine the role of these sequences in viral replication.

In their analyses of genetic relationships between geminiviruses, both Rybicki (1994) and Padidam *et al.* (1995) found very low levels of sequence homology between LIRs of different Subgroup I geminivirus species. In contrast with Subgroup II and III geminiviruses, these differences even extend to the stem sequences of the stem-loop origin of replication sequence. This observation led Argüello-Astorga (1994a) to propose that the replication-specificity determinant for Subgroup I geminiviruses must lie in the stem structure. Their proposal was strengthened by showing that certain sequences in the stem of the LIR stem-loop structure were iterated at least three times in the LIR: once on either side of the stem sequence, and once between the TATA box of the *rep* gene and its transcription start point. The positions of these iterons are well conserved in the LIRs of all of the Subgroup I viruses the authors examined (see Figure 1.6). Given that the ability of Rep to bind to its specific recognition sequence in the common region of Subgroup II and III viruses is a determinant of replication specificity (Fontes *et al.*, 1992; 1994a; Lazarowitz *et al.*, 1992), it seems reasonable to propose that similar iterons found in Subgroup I geminiviruses also constitute specific Rep binding sites. This hypothesis is further strengthened by the fact that the position of the *rep* gene proximal iteron would allow Rep binding at this site to mediate transcriptional repression of its own gene, in the same way that Subgroup III Rep does (Eagle *et al.*, 1994; Hong and Stanley, 1995; Sunter *et al.*, 1993). Hanley-Bowdoin *et al.* (1997) proposed that it was logical that Subgroup I Rep should bind the stem-loop sequence specifically, as this obviates the need for a REn-like protein to direct Rep complexes to the origin of replication in this genus.

There is some indirect experimental evidence that the iteron sequences identified by Argüello-Astorga do comprise at least part of the Rep recognition sequences. Kamman *et al.* (1991b) made a surprising discovery that a mutant of WDV with a deletion of 127 bp of the LIR, including the stem-loop sequence, could still initiate replication. However, only high molecular weight concatemeric forms of the virus DNA were produced, which suggested that although initiation of replication was occurring, there was some defect in termination of RCR which resulted in the formation of concatamers of the virus genome. Upon further analysis, Heyraud *et al.* (1993a) found that this mutant virus was initiating replication at the sequence TACCC, which is found adjacent to the WDV *rep* gene proximal iteron and fortuitously resembles the nicking site of the geminivirus nonanucleotide. In an elegant experiment, Heyraud *et al.* (1993a) cloned the deleted WDV LIR in tandem with the wild-type virus genome. When this clone was introduced into

protoplasts, replicative release of a mutant virus resulted, mediated by Rep binding at the proximal iteron and nicking the TACCC sequence and nicking by another Rep complex at the TAATATTAC sequence in the wild type stem-loop. This resulted in the formation of a mutant virus with a sequence duplication in its LIR. This mutant virus had reconstituted a stem-loop structure which consisted entirely of a stem made up of the first and second iterons in the virus (instead of the second and third) in between the TAATATTAC sequence in the loop. The mutant virus replicated efficiently in protoplasts. These experiments provide good circumstantial evidence for Rep binding the iteron sequences and incidentally showed that, in Subgroup I geminiviruses, a stem-loop structure is not necessarily required for initiation of RCR, but the structure is required for termination and resolution.

In their identification of iterons as putative Rep binding sites in the LIRs of Subgroup I geminiviruses, Argüello-Astorga *et al.* (1994a) made three unproven suppositions: (1) like that of Subgroup II and III viruses, Rep from Subgroup I is highly specific and only binds to its own origin of replication; (2) the Rep binding sites of Subgroup I geminiviruses are found in the stem of the stem-loop sequence; and (3) again by analogy with the Subgroup III situation, the Subgroup I Rep must suppress its own expression. These assumptions would seem to be reasonable, but there are some indications, unfortunately mostly anecdotal, that they are not entirely correct.

There is some evidence that interaction between Subgroup I geminivirus Rep and the LIR may not be as specific as that seen in Subgroup III geminiviruses. The ability to form viable pseudorecombinants, that is for DNA A of one geminivirus to transreplicate the DNA B of another geminivirus, could be used to determine whether two different Subgroup III geminiviruses are separate species, or isolates of the same virus. Only viruses with the same, or very similar, iteron structure can form viable pseudorecombinants, such as is found for different isolates of ACMV (Stanley *et al.*, 1985) and two closely related species tomato mottle virus (ToMoV) and bean dwarf mosaic virus (BDMV) (Gilbertson *et al.*, 1993). One could use the same criterion for determining whether two Subgroup I geminiviruses are different species: strains of the same virus should be able to *trans* replicate each other, but different virus species should not. Generally, Subgroup III geminiviruses with more than about 90% homology at the nucleotide level are considered to be strains of the same virus species (Padidam *et al.*, 1995). For example, TYLCV-Sr

and TYLCV-I, which are 75% homologous at the nucleotide level, are not able to *trans* replicate each other (Jupin *et al.*, 1995), and should perhaps be considered different virus species rather than strains of TYLCV. Similarly, BCTV "strains" CFH and Logan, which share 82.5% overall nt sequence homology, are not able to complement each other's replication functions; the replication specificity determinant in this case maps to the viral IR iteron sequences (Choi and Stenger, 1995; 1996). On the other hand, I am aware of two cases where two Subgroup I geminiviruses with sequence divergence of greater than 10% between them, are still able to recognise each other's origin of replication.

The first example is that of the Swedish isolate of WDV (WDV-S) and a "barley adapted" strain of WDV called WDV-ER. The sequence of WDV-ER is unfortunately not available in any nucleotide sequence databases, so the only information I have been able to garner about this virus is from mention in two papers: Bendahmane *et al.* (1995) and Heyraud *et al.* (1993b). These two strains of WDV are 82.5% homologous at the nucleotide level and their LIRs exhibit only 75% sequence homology (Heyraud *et al.*, 1993b). Alignment of the sequences of the stem-loop regions of these two WDV strains shows that they share 82% sequence identity in this region, and that there are some minor differences in the iteron sequences of the two viruses. Yet WDV-S and WDV-ER are able to recognise the heterologous LIR, as shown by the fact that reciprocal clones with hybrid LIRs could replicate (Heyraud *et al.*, 1993b). Moreover, Bendahmane *et al.* (1995) cite unpublished data of Heyraud-Nitschke and Gronenborn which prove that WDV-S and WDV-ER complement each other for replication. Similarly, two distantly related strains of MSV, MSV-Kom and a mild isolate from *Seteria* sp. are also able to *trans*-replicate the other genome, despite significant sequence divergence in the LIR (W. H. Schnippenkoetter *et al.*, in preparation).

The evidence of Heyraud *et al.* (1993a) that the WDV iteron constitutes the Rep binding site is fairly convincing; it is also attractive to propose that Rep binds near its cleavage and joining site. A provisory point is that, if RepA is produced, and given evidence from Settlage *et al.* (1996) that Rep forms oligomers, there are potentially three different Subgroup I virus Rep complexes which could be formed: Rep homo-oligomers; Rep-RepA hetero-oligomers and RepA-RepA homo-oligomers. These different complexes could have different binding specificities.

It is not unreasonable to speculate that there is an element of structural recognition involved in binding of Rep to the stem-loop sequence to effect initiation and termination of replication. This could reconcile Rep and REn functions in the same protein: REn does not act virus-specifically, so it is plausible that this protein might be able to recognise the stem-loop structure, although the sequence of the stem-loop is very similar between different Subgroup III geminiviruses (Lazarowitz, 1992). It would also help to explain how Rep of the two different MSV and WDV variants can recognise the heterologous origin of replication. It is possible that Subgroup I Rep recognition of the viral origin of replication could be on the basis of multiple or additive factors, such as recognition of specific sequences (the iterons?), recognition of the stem-loop structure, and recognition of other proteins, possibly transcription factor/s bound at the GC-boxes found near the base of the stem-loop structure (Figure 1.6). Transcription factors are often involved in enhancing origin recognition by origin recognition proteins of vertebrate infecting DNA viruses like SV40, polyomavirus, papillomaviruses and adenovirus 2 (DePamphilis, 1993, and references therein).

If production of RepA is implicated in down-regulation of replication and activation of virion sense transcription, and Rep is not involved in virion sense promoter transactivation, as suggested by Collin *et al.* (1996), then perhaps complexes in which RepA is involved have slightly different binding specificities. It will be very interesting to see whether heterologous Rep (or RepA) can participate as effectively in effecting Rep-mediated transcriptional activation of virion sense and repression of complementary sense gene expression. The fact that the sequence of the *rep* gene proximal iteron is, in some Subgroup I geminiviruses, not completely homologous to the left-hand loop iteron, and the fact that often only three or four nucleotides of the right-hand loop iteron are conserved in a particular virus (Figure 1.7) would point to the probability that there are other uncharacterised factors involved in Rep recognition of specific DNA binding sites.

identified a bending locus in WDV (see the MSV-Kom LIR in Figure 1.6.). Therefore, an involvement of DNA bending in this region in replication of Subgroup I geminiviruses generally is dubious. This was recently supported by data from Suárez-López and Gutiérrez (1997) who showed that the DNA bending locus has only minimal impact on replication of WDV. However, this does not preclude a role for these sequences in transcription from the virion sense promoter of WDV. It remains possible that DNA bending does occur in the LIR of other Subgroup I geminiviruses, perhaps mediated by binding of certain protein factors.

The negative sense origin of replication of Subgroup I geminiviruses: the SIR

The SIR of all Subgroup I geminiviruses functions as the complementary sense origin of replication. The origin of the DNA primer which binds in this region is unknown. The MSV-N primer molecule associated with virions of this virus has several ribonucleotides at its 5′ end, and a heterogeneous 3′ end (Donson *et al.*, 1984). The presence of ribonucleotides at the 5′ end of the primer molecule suggests that the DNA portion is primed from a longer RNA primer. The heterogeneity which the authors noted in the 3′ end of the MSV primer molecule could be due to interruption of RNA-primed complementary strand synthesis by the encapsidation process. It is interesting to note that Hafner *et al.* (1997) recently showed that banana bunchy top virus (BBTV), a member of the only other (putative) taxonomic family of true plant DNA viruses, also possesses a virion-associated DNA primer molecule, but these authors found no evidence for ribonucleotides linked to the 5′ end of the primer. It is also interesting that Saunders *et al.* (1992) found that complementary strand synthesis of ACMV is RNA-primed, but Subgroup II and III viruses have no virion-associated DNA primer. The issues of the origin of the Subgroup I virion-associated primers and of how formation of complementary sense DNA is regulated in geminiviruses, and whether this is similar in plant BBTV-like viruses, are still far from resolved.

Apart from its role in replication, the SIR also contains consensus transcription termination and polyadenylation signals, and thus could also play a role in control of gene expression. Shen and Hohn (1991) performed the only study of the SIR by mutagenesis published thus far. They found that deletion of a short segment of the MSV genome in the SIR was tolerated and that they could insert small oligonucleotides in the *Asn* I site found in the

SIR, without abolishing viral infectivity. However, infectivity and symptom severity declined as the size of the oligonucleotide inserted in the *Asn*I site was increased. This could be due to some interference in the formation of 3′ ends of viral mRNAs, but this possibility has not been investigated further. The oligonucleotide insertions were also not always stable, and were sometimes found to be deleted from virus genomes in infected plants.

Control of viral gene expression

It is imperative for any virus that there be tight regulation of expression of gene products which are involved in different stages of the viral infection cycle. For geminiviruses, the primary level of control of gene expression must be at initiation of transcription. While it is probable that there may also be control at the post-transcriptional level, by such mechanisms as control of the proportion of spliced to unspliced mRNAs, and even at the level of post-translational modification of viral protein products, these processes are as yet uninvestigated in geminiviruses.

Control of expression of the complementary sense genes

In Subgroup I geminiviruses, the main viral protein which is required at an early stage in the viral life cycle is Rep. The other protein which could be encoded by the C1 ORF, RepA, could be expressed from an unspliced, or 3′ truncated, mRNA. Collin *et al.* (1996) found that mutant WDV which contained a cDNA copy of the *rep* gene in place of the wild type gene replicated to a higher level in tissue cultured cells than did the wild type virus, which retained the potential to express RepA. This would seem to present an attractive mechanism for control of the level of viral replication, with Rep promoting high level replication of the viral genome at an early stage when replication is the primary focus, and the production of RepA mediating some sort of down-regulation of viral replication later on in the viral replication cycle.

In investigations arising from the discovery of the MSV-Nm isolate (Boulton *et al.*, 1991a; b), Boulton and colleagues found that transcription of the complementary sense genes of MSV-N could be initiated from two or three points. Mutation of the TATA box at -101 in MSV-Nm resulted in the initiation of the complementary sense transcript primarily from the more proximal TATA boxes at -57 and -62. Interestingly, this coincided with an

increase in the level of a 3'-truncated complementary sense transcript, relative to the full-length transcript (M.I. Boulton, personal communication). This could well result in a relative increase in the cellular concentration of RepA and decrease in Rep, which may account for part of the MSV-Nm phenotype. Whether this phenomenon indicates some sort control of complementary sense gene expression at the post-initiation of transcription level is not certain, but warrants further investigation.

Cessation of replication at a late stage in the viral life cycle may correspond with repression of the *rep* gene promoter, if Rep and/or RepA is capable of repression of its own promoter, as is the case in Subgroup III geminiviruses (Eagle *et al.*, 1994; Hong and Stanley, 1995; Sunter *et al.*, 1993). The putative Rep-binding sequence in the iteron situated between the TATA box and transcription start site of the *rep* gene would allow Rep to mediate repression of its own promoter, by interfering with initiation of transcription or assembly of a transcription initiation complex (Argüello-Astorga *et al.*, 1994a; b). The SV40 replication initiator protein T-Ag also represses its own promoter at a late stage of the viral replication cycle by binding to the viral control region and sterically interfering with RNA polymerase binding or progression. It may be worthwhile noting that T-Ag also activates viral early genes at an early stage in the viral life cycle. This activation appears to be independent of T-Ag binding to viral DNA, but may instead be a result of interaction with host proteins (reviewed by Fanning and Knippers, 1992). It would be interesting to investigate the effect that Rep expression has on its own promoter at different stages in the viral infection. However, in the absence of experimental data on control of *rep* gene expression in Subgroup I geminiviruses, these hypotheses are purely speculative.

Control of Virion-Sense Gene Expression

The virion sense promoter of MSV was the first geminiviral promoter to be investigated in any detail, by Fenoll *et al.* (1988), with the presumptive aim of this investigation finding a strong constitutive promoter for transgenic cereals, as an alternative to the CaMV 35S RNA promoter. However, the MSV virion sense promoter was only about 20% as strong as the CaMV 35S promoter in Black Mexican sweetcorn (BMS) protoplasts. Perhaps the most important factor which should be taken into account when examining the virion sense promoter of any Subgroup I geminivirus is the possibility that the promoter is transactivated by the early viral protein (Rep). It does not make logical sense for a geminiviral coat protein or movement protein promoter to be maximally active at an early

stage in the viral life cycle, before the copy number of the genome is high enough to efficiently initiate systemic movement. Unfortunately, the study of the MSV rightward promoter by Fenoll *et al.* (1988; 1990) was done in the absence of any viral proteins, so the results must be taken to represent the promoter activity divorced from a normal infection situation.

Fenoll *et al.* (1988) used a 1138-bp fragment upstream of the MSV coat protein gene as the full-length promoter. The marker gene which they used to evaluate promoter strength was the *E.coli* chloramphenicol acetyl transferase (CAT) gene, which they cloned as a transcriptional fusion, at a site 24 bp upstream of the coat protein ATG. These authors were unfortunately unaware of the presence of an intron in the V1 ORF. In this light, it is interesting that they found that deletion of most of the V1 ORF, up to 35 bp 3′ of the V1 ATG, virtually abolished CAT expression. This would suggest that either processing of the intron, or the physical presence of an RNA element within the V1 ORF is required for coat protein expression, at least in the context of this experiment. However, this is in contrast to results reported for WDV, which show that deletion of V1 has no effect on coat protein expression (Hofer *et al.*, 1992). In this light, it may well be significant that Margaret Boulton's group has found no evidence of RNA processing in the WDV virion-sense transcript.

According to the results presented by Fenoll *et al.* (1988), the MSV virion sense promoter extends as far as an *Asu*II site, about 530 bp upstream of the coat protein start codon. They found that a 122-bp *Asu*II-*Apa*I restriction fragment, including the stem-loop origin of replication structure, was essential for full promoter activity. This *upstream activating sequence* (UAS) seemed to have similar, albeit weaker, transcription activating activity to the distal region (DR) segment of the CaMV 35S promoter. *In vitro,* the MSV UAS bound plant nuclear factors, probably distinct from those which bind the DR of the CaMV 35S RNA promoter (Fenoll *et al.*, 1988)

In a subsequent publication, Fenoll *et al.* (1990) examined the MSV UAS further. They found that the stem-loop structure was not necessary for the transcription-activating activity of the UAS. A conserved region consisting of two direct CG-box repeats found at the base of the stem loop structure was important for the transcription activating and maize nuclear factor binding activities of the UAS (see Figure 1.6). This region, which has homology

with the binding site of transcription factor *Sp1* in several animal and animal virus promoters, was called the *rightward promoter element* (*rpe1*). The maize nuclear factors which interacted with *rpe1 in vitro*, bound in a noncooperative fashion to each of the two GC-boxes, and on opposite faces of the DNA helix. It will be very interesting to determine the identity of the transcription factors which bind *rpe1*, as these are also likely to interact in some way with Rep bound at the stem-loop, either in initiation of replication or activation of virion sense gene expression, or both.

The transactivation phenomenon associated with geminivirus virion sense gene expression has attracted a great deal of interest, not least because of implications in exploiting this for different applications in biotechnology (discussed in section 1.2.3). It is common for early viral protein products to *trans*-activate the expression of late genes; adenovirus E1A, SV40 T-Ag and papillomavirus E2 proteins, for example, are involved in these processes. Transactivation of late gene promoters by early gene products could represent an elegant way for the virus to sense the amount of early protein present in an infected cell and to gauge the stage of the viral replication cycle. Although it is established that the complementary sense genes of Subgroup I geminiviruses are required for maximal expression of the virion sense genes, it is not clear whether this phenomenon is mediated by Rep, Rep A, or by a RepA-Rep hetero-oligomeric complex.

There are potentially two ways in which Rep could enhance the expression of the coat protein promoter. The first of these is an indirect mechanism: simply by increasing the gene copy number. Given the data of Fenoll *et al.* (1988) which show that the MSV virion sense promoter is fairly weak, at least in the absence of Rep, then increasing the copy number could have a positive effect on the overall level of expression of the promoter. This mechanism would fit with maximal expression of the virion sense genes being a "late" function, corresponding with the stage in the viral life cycle where the level of RF-DNA is highest. However, one would anticipate by analogy with Subgroup III geminiviruses that there must be more specific activation (or derepression) of virion sense gene expression at a later stage in the virus life cycle.

There have been two reports of Rep-mediated transactivation of the virion sense promoter in Subgroup I geminiviruses: for WDV by Hofer *et al.* (1992) and for CSMV by Zhan *et al.* (1993). Zhan *et al.* (1993) reported modest enhancement of expression of the activity of

the coat protein promoter by Rep: between two and three-fold. They also found that the coat protein also enhanced its own expression slightly. However, when I analysed the sequence of the DNA fragments they claim to have used as potential promoter regions for their transient expression assays, it seems that they were erroneously using fragments of the C1 ORF and *rep* gene proximal regions of the LIR as their virion sense promoter constructs. In this light, the data must unfortunately be considered unreliable, but it remains possible that the authors reported incorrect sequences of PCR primers they used to generate the promoter constructs.

The study by Hofer *et al.* (1992) showed that they were not able to detect any activity of the WDV virion sense promoter in the absence of Rep. This implies the basal coat protein expression level was below the detection limit of their assay, since Matzeit *et al.* (1991) could detect NptII activity from non-replicating WDV coat protein gene replacement constructs. The authors present convincing evidence that the C2 ORF is required for expression of the virion sense promoter of WDV, as constructs which contained frameshift mutations in this ORF, and therefore could not replicate, also did not produce detectable levels of CAT in their expression assay.

The main evidence that Rep of Subgroup I geminiviruses transactivates the virion sense promoter was based on results generated from analysis of a WDV deletion mutant constructed by Hofer *et al.* (1992) which was still able to express the *rep* gene, but was not able to replicate due to deletion of an 84-bp region of the LIR including the left hand side of the stem loop and part of the nonanucleotide sequence. Although the authors were unable to detect replicating WDV DNA, this mutant virus still expressed a fairly high level of CAT. However, the stem-loop deletion mutant constructed by Hofer *et al.* (1992) still had the first iteron sequence (with alternative nicking site) intact, and in using the *Ssp*I site in the WDV nonanucleotide sequence to delete an *Apa*I-*Ssp*I fragment, these authors generated a second, almost identical secondary replication initiation site as that which Heyraud *et al.* (1993a) proved could function to initiate replication of high molecular weight concatamers of the WDV genome. Thus, the deletion construct pAPS84 described by Hofer *et al.* (1992) should be proficient for replication of high molecular weight concatamers of the WDV CAT replacement mutants, even if no monomeric dsDNA replicons were detected. Unfortunately, if this is so, it makes the authors' conclusion that transactivation of the virion sense promoter contributes to about 50% of the observed

virion sense promoter activity in replicating constructs invalid. In this light, it is perhaps significant that the kinetics of CAT expression from the pAPS84 construct closely parallel those of pWDV2CAT, the parental construct with the 84 bp fragment of the LIR which was deleted in pAPS84 still intact. The authors had previously correlated the CAT expression kinetics with the replication levels of pWDV2CAT, so these results would support my speculation that pAPS84 should be able to replicate. The fact that the deletion construct could express 50% of the amount of CAT as the replication proficient construct may merely reflect that pAPS84 cannot replicate to as high a copy number. The issue of virion sense promoter transactivation by Rep in Subgroup I geminiviruses is therefore not yet satisfactorily resolved.

In a recent publication, Collin *et al.* (1996) showed that a replicating WDV construct with a cDNA of the *rep* gene in place of the full-length gene expressed very low levels of a reporter gene (GUS) cloned in place of the coat protein gene. However, the identical construct with the *rep* intron intact was able to express much higher levels of GUS. This indicates that RepA is in some way required for virion sense gene expression. Given these authors' finding that the WDV replicon expressing the *rep* cDNA could replicate to a higher level than the construct which could express RepA, this would suggest that expression of RepA has two major effects: (1) down-regulation of replication, and (2) enhancement of virion sense gene expression. This could mean that expression of RepA induces viral "late" functions. If these results are confirmed, it would be interesting to see if Rep actively suppresses virion-sense expression.

The apparent transactivation of the virion sense promoter by RepA observed by Collin *et al.* (1996) was not confirmed in the context of nonreplicating constructs. In WDV this would be quite difficult to achieve without deleting the *rep* gene proximal iteron and the stem loop to make a *bona fide* replication-deficient construct. When (or if) the phenomenon of virion sense promoter transactivation by RepA in Subgroup I geminiviruses is proved, it will be interesting to determine if the effect is mediated by RepA alone, or by a Rep-RepA complex. Collin *et al.* (1996) noted that RepA does not have a discernible nuclear targeting sequence, whereas Rep does, so RepA may only be able to get to the nucleus as part of a complex with full-length Rep or a host protein.

To conclude my review of the literature on control of viral gene expression in Subgroup I geminiviruses, I note that there is very little known for certain about the control of both complementary sense and virion sense gene expression in these viruses. On the issue of transactivation of virion sense gene expression, the first question which needs to be answered is whether this happens at all in Subgroup I geminiviruses, although there is indication from Collin *et al.* (1996) that it does. In Subgroup III geminiviruses, transactivation of the otherwise virtually silent TGMV virion sense promoters by TrAP was first documented by Sunter and Bisaro (1991). The same group later showed that, in the presence of TrAP, genome amplification by Rep enhanced the expression of the TGMV coat protein promoter a further 60- to 90-fold (Brough *et al.*, 1992). Results of Matzeit *et al.* (1991) for WDV and my results for MSV (presented in Chapter 3) show that genome amplification certainly contributes to the overall level of expression of the coat protein promoter. Divorcing the amplification effect from any Rep/RepA transactivation effect presents the next challenge. As the AC2 gene product is a non-sequence specific DNA binding protein *in vitro* (Noris *et al.*, 1996; Sung and Coutts, 1996), transactivation could well come about from an effect which TrAP has on another protein bound at the viral promoter. If Subgroup I Rep/RepA mediated transactivation occurs, is this as a direct result of Rep binding to the LIR, or indirectly through interaction with host factor/s? If specific DNA binding is required for Rep to activate the coat protein promoter, then one would expect that transactivation of the virion sense promoter would be a more virus-specific phenomenon in Subgroup I geminiviruses than Sunter *et al.* (1993) noticed for TGMV TrAP.

Movement and encapsidation of Subgroup I geminiviruses

Interactions Between Movement Proteins and the Viral Genome

Plant virus movement can be divided into 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 (reviewed recently by Citovsky, 1993; Leisner and Howell, 1993; Lucas and Gilbertson, 1994; and Gilbertson and Lucas, 1996). As DNA viruses which replicate in the nucleus, geminiviruses are presented with an extra hurdle in movement: entry and exit from the nucleus. The specific mechanisms involved in movement of geminiviruses are not fully elucidated yet, and in

some cases the issues are fairly controversial. Most investigation on movement has again centred on Subgroup III and the role of the B component movement proteins, but has resulted in the development of plausible models for the molecular mechanisms involved in bipartite virus movement (reviewed by Sanderfoot and Lazarowitz, 1996). Aspects of these models could also apply to movement of Subgroup I geminiviruses.

Although the molecular mechanisms involved in movement of Subgroup I geminiviruses are less well understood than in Subgroup III, it is well established that both virion sense ORFs of the former are involved in viral movement. Boulton *et al.* (1989b) and Lazarowitz *et al.* (1989) showed that MSV mutants with point mutations, small insertions, deletions and gene replacement mutations in one or other of the V1 or V2 ORFs were not able to move systemically in inoculated plants. These mutants were generally still able to replicate, as both groups were able to detect low levels of the mutant viral genomic DNA replicating in inoculated leaves, but not in leaves which emerged later. These findings were corroborated by Woolston *et al.* (1989), who also found that WDV constructs with mutations in the coat protein gene were not infectious in plants, although they could replicate to wild-type levels in protoplasts.

The V1 ORF probably encodes the movement protein of Subgroup I geminiviruses. This is supported by data of Boulton *et al.* (1993) who showed that mutants of MSV which were not infectious due to mutations in the V1 ORF could still replicate to wild-type levels in protoplasts, could express coat protein and generate ssDNA, and were also encapsidation proficient. Mutation in the movement protein gene therefore did not disrupt expression of the coat protein gene, and did not prevent encapsidation of viral plus strand DNA. The MSV V1 protein is definitely present in infected plants (Mullineaux *et al.*, 1988). Its role as a movement protein is further strengthened by the computer predictions of a hydrophobic potential trans-membrane or membrane-embedded domain in the V1 protein of all Subgroup I geminiviruses (Boulton *et al.*, 1993). More recently, Dickinson *et al.* (1996) found that MSV V1 protein is localised to the cell wall, and like the movement proteins of several plant RNA viruses, is associated with secondary (branched) plasmodesmata.

The coat protein of all geminiviruses is required for insect transmission, and is probably the sole viral genetic determinant of insect vector specificity (Briddon *et al.*, 1990). The role

of the coat protein in geminivirus movement is not well understood; it is, however, indisputably involved in movement of monopartite viruses from all three genera. It is conventionally accepted that coat protein is dispensible for viral movement in bipartite Subgroup III geminiviruses. However, recent data suggest that the role of coat protein in movement of these viruses is probably under-appreciated. For example, Ingham *et al.* (1995) found evidence for functional redundancy between SqLCV BV1 and coat protein, in that the presence of coat protein could mask the effects of some BV1 mutations. In addition, Pooma *et al.* (1996) have found that coat protein is required for movement of bipartite geminiviruses in hosts to which they are not well adapted.

The coat protein of geminiviruses from all subgroups is implicated in the accumulation of single stranded DNA in infected cells. In Subgroup I geminiviruses, coat protein mutants never accumulate detectable levels of ssDNA in transient replication assays, or in inoculated leaves (personal observation; Boulton *et al.*, 1989b; 1993; Lazarowitz *et al.*, 1989). Coat protein mutants of Subgroup II and monopartite Subgroup III geminiviruses are non-infectious and also accumulate reduced amounts of ssDNA. Likewise, bipartite Subgroup III geminiviruses with mutations in the coat protein gene accumulate reduced, but detectable, levels of ssDNA (Azzam *et al.*, 1994; Jeffrey *et al.*, 1996; Ingham *et al.*, 1995 and references therein). It is not yet clear whether the accumulation of ssDNA is mediated by coat protein simply sequestering ssDNA by encapsidation, or whether this is due to a specific genetic regulatory switch from dsDNA replication to production of ssDNA. It is interesting to note that viruses from Subgroups II and III of the *Geminiviridae* have other proteins which are also implicated in ssDNA accumulation: the precoat ORF (V1) of some monopartite and bipartite Subgroup III geminiviruses (Rigden *et al.*, 1993; Paddidam *et al.*, 1996); V3 (R2) of BCTV as well as BV1 (Jeffrey *et al.*, 1996), which accounts for why coat protein mutants of these viruses still accumulate low levels of ssDNA. However, V1 of Subgroup I geminiviruses plays no discernible role in ssDNA accumulation; this is a function specific to coat protein in this genus.

Given that the coat protein of bipartite Subgroup III geminiviruses is dispensible for systemic movement in permissive hosts, it is unlikely that geminiviruses move encapsidated. It is not known if coat protein plays a direct role in movement of Subgroup III viruses, as BV1 does, but both Ingham *et al.* (1995) and Jeffrey *et al.* (1996) have proposed that the role of coat protein in movement of Subgroup III viruses is associated

with its role in stimulation of ssDNA accumulation. These authors propose that increased levels of ssDNA, mediated by coat protein action, can compensate for deficiencies in movement, either due to mutation in BV1 or to host limitations.

Like the Subgroup I geminiviral movement protein (V1), BC1 localises to cell wall and plasma membrane fractions in infected plants (von Arnim *et al.*, 1993; Pascal *et al.*, 1993). The movement protein of MSV is probably postranslationally modified, since its apparent size in infected plants is 14 kDa, compared with the 10.9 kDa predicted from the primary amino acid structure (Dickinson *et al.*, 1996). BC1 also seems to be subject to postranslational modification (von Arnim *et al.*, 1993; Pascal *et al.*, 1993). A study by Noueiry *et al.* (1994) showed that BC1 protein of the Subgroup III geminivirus, bean dwarf mosaic virus (BDMV), as has been shown for movement proteins of several RNA viruses, changed the plasmodesmal size exclusion limit (SEL) when microinjected into tobacco or bean cells. In addition, these authors showed that BC1 was able to spread rapidly to adjacent cells, as is TMV movement protein (Waigman and Zambrysiki, 1995). Transgenic tobacco plants expressing BC1 (BL1) protein of SqLCV exhibit symptoms typical of virus infection, a finding which is consistent with the observation that this protein is a major determinant of pathogenicity in Subgroup III geminiviruses (von Arnim and Stanley, 1992; Ingham *et al.*, 1995). The finding that transgenic tobacco expressing MSV movement protein and tobacco expressing the movement protein of the novel South African Subgroup I virus BeYDV also exhibit symptoms of viral disease (Margaret Boulton, personal communication) strengthens the hypothesis that BC1 and the movement protein of Subgroup I geminiviruses have similar functions.

In contrast to BC1, the BDMV BV1 protein was neither able to modify the plasmodesmal SEL, nor was it able to move from cell to cell (Noueiry *et al.*, 1994). In addition, transgenic plants expressing BV1 were phenotypically normal (Pascal *et al.*, 1993). DNA binding studies with BC1 and BV1 protein showed that BC1 binds ssDNA very weakly and does not bind dsDNA at all, while BV1 appears to have specific affinity for ssDNA, and binds more weakly to dsDNA and RNA (Pascal *et al.*, 1994). BV1 localised to the nucleus in both plant and insect cells (Pascal *et al.*, 1994, Sanderfoot and Lazarowitz, 1995), and microinjected BV1 was able to move both ssDNA and dsDNA out of the nucleus (Noueiry *et al.*, 1994). These results suggest that it is the BV1 movement protein, not BC1, which interacts directly with viral DNA to potentiate movement.

These results are in some ways in conflict with results reported by Nouiery *et al.* (1994), who investigated whether BC1 could promote cell-to-cell movement of BDMV viral DNA, by co-microinjecting BC1 with BDMV DNA purified from *E.coli.* Either single stranded or double stranded forms of the cloned BDMV genome (as ~5.5 kb monomeric clones in pBluescript) were fluorescently labelled and microinjected into mesophyll cells together with wild type or movement-deficient mutant forms of BC1. Interestingly, dsDNA rapidly moved out of injected cells, but labelled ssDNA and RNA remained intracellular; this movement also only occurred when wild type BC1 protein was injected, showing that specific transport of dsDNA was mediated by BC1. However, these authors did not show whether their proteins purified from *E.coli* had DNA binding activity, as the studies of Pascal *et al.* (1994) showed that BV1 produced in insect cells bound ssDNA strongly, and BC1 does not bind dsDNA at all. Similarly, MSV movement protein has no DNA binding activity at all, and coat protein binds both ssDNA and dsDNA (Liu *et al.*, 1997; Boulton, personal communication).

Geminiviruses typically accumulate high levels of dsDNA in infected cells; Noueiry *et al.* (1994) feel that this provides some justification for why geminiviruses move as dsDNA. However, none of these authors has taken into account the fact that their studies on movement and DNA binding have been performed on dsDNA purified from *E.coli*, whereas geminiviral RF-DNA is chromatin-like, that is, it is nucleosome-associated and packaged into minichromosomes (Abouzid *et al.*, 1988; Pilartz and Jeske, 1992). Therefore, the dsDNA which is used in microinjection studies (and in DNA binding studies) is unlikely to be biochemically similar to geminivirus RF-DNA. In this light, it is not easy to extrapolate results obtained with naked plasmid dsDNA purified from *E.coli* to the situation *in planta*. It would be interesting to see if coat protein and/or BV1 have the same binding specificity for geminivirus minichromosome structures. Histones, for obvious reasons, are basic proteins which have strong nuclear localisation signals–which one imagines would be at conflict with a drive to export the RF-DNA from the nucleus. Therefore, if dsDNA exits the nucleus, it must be in a form which is stripped of histones.

Recently, Sanderfoot and Lazarowitz (1995) and Sanderfoot *et al.* (1996) showed that BC1 interacts specifically with BV1 and redirects it from the nucleus to the perifery of the cell. BV1 seems to be a nuclear shuttle protein, which probably binds viral ssDNA and moves it

in and out of the nucleus. They hypothesise that BC1 specifically recognises BV1 (with bound ssDNA) and traps the complex in the cytoplasm, and then directs it to the cell periphery. Given that BC1 can modify plasmodesmal SEL and apparently has only very weak nucleic acid binding properties, it is perhaps the BV1-ssDNA complex which BC1 recognises and facilitates the movement of to adjacent cells. If one extrapolates the functions of BV1 to coat protein, and BC1 to movement protein, one could propose a similar model for cell-to-cell movement of Subgroup I geminiviruses: that is, that coat protein bound to ssDNA is recognised by the movement protein and mobilised to adjacent cells.

Further investigation by Sanderfoot *et al.* (1996) reveals that BV1 has two distinct nuclear localisation signals (NLS) at its N-terminus, while domains essential for interaction with BC1 reside at the C-terminus. With the evidence of Sanderfoot *et al.* (1996) that BV1 and BC1 interact via specific domains, it is not surprising that Schaffer *et al.* (1995) found that B component movement proteins require the homologous partner for optimal function. These authors constructed chimaeric viruses from TGMV and BGMV. They found that TGMV with both BGMV movement proteins in place of the TGMV alleles and the reciprocal hybrid virus were infectious, albeit with reduced virulence in the original optimal host (*N. benthamiana* or bean, respectively). However, when TGMV BC1 and BGMV BV1 were provided with DNA A of either virus, the chimaeric virus was severely defective, and the combination of BV1 from BGMV and BC1 from TGMV was not viable. This provides good evidence that specific interactions between the movement proteins of Subgroup III geminiviruses is required for optimal function. By a similar approach, we have found evidence for specific interaction between movement protein and coat protein of Subgroup I geminiviruses (E. van der Walt, K.E. Palmer and E.P. Rybicki, unpublished results). On analysis of chimaeric viruses with movement and/or coat protein genes from the related viruses, MSV-Kom (moderately severe) and MSV-Set (mild), we found that all hybrid viruses showed reduced virulence, and that viruses with both coat and movement proteins from the same genetic background were more fit than viruses which had combinations of coat protein and movement protein from different viruses. We feel this is good evidence for specific interaction between coat protein and movement protein of Subgroup I geminiviruses in promoting viral movement, and indicates that there are likely to be strong parallels between movement mechanisms of Subgroup III and Subgroup I geminiviruses.

An issue which has not yet been fully explored is that of size limitations in geminiviral movement. Given that coat protein, and by implication, encapsidation, is not absolutely necessary for movement in bipartite geminiviruses it would seem that there should be no major size constraints on the size of the DNA which is moved. On the contrary, however, coat protein gene replacement constructs based on both TGMV and ACMV rapidly revert to close to wild type genome size on movement in infected plants (Etessami *et al.*, 1989; Hayes *et al.*, 1989; Elmer and Rogers, 1990). Likewise, MSV-based constructs which contain insertions in the short intergenic region which do not interfere with replication and gene expression, cannot move unless the inserted DNA is deleted and the genome reverts to wild type size (Shen and Hohn, 1991; 1992; 1994; 1995). However, it is not certain what role encapsidation plays in movement of Subgroup I geminiviruses, so the size limitation for movement could be one imposed by encapsidation constraints.

The controversial study by Noueiry *et al.* (1994) found no size limitation on cell-to-cell movement of dsDNA mediated by BC1, since both cloned virus (~5.5 kb) and pBluescript (~2.95 kb) were moved efficiently. If their finding that geminiviruses move as dsDNA is confirmed, this would imply that the size limitation operates on the level of long distance movement, rather than plasmodesmal transport. If dsDNA is the form in which geminiviruses move, then one would imagine that there must be some mechanism by which BV1 can distinguish between host DNA and viral DNA, perhaps on the level of its circular, supercoiled conformation.

Although the issues in cell-to-cell spread of geminiviruses are far from solved, the next, largely unexplored, frontier to confront is systemic movement. The involvement of coat protein in this process needs to be elucidated. Given that there is an evolutionary relationship between coat protein and BV1, and that coat protein and movement protein of Subgroup I viruses interact specifically (van der Walt *et al.*, unpublished results), it would seem reasonable to propose that Subgroup I geminivirus coat protein has a role as: (1) a ssDNA binding protein; (2) a nuclear shuttle protein; (3) a protein which interacts with movement protein to mediate movement of ssDNA-coat protein complexes from cell-to-cell. However, the role of both coat protein and movement protein in long distance movement is unknown.

It has been conventionally accepted that geminiviruses are phloem limited, although there have been some recent reports of exceptions to this rule. Lucy *et al.* (1996) investigated the spread of MSV in maize by *in situ* hybridisation and immunocytochemical localisation approaches. These authors found that MSV was limited to the vascular tissue in the shoot apex and stem tissues. Virus was not present in the apical meristem, and was only detectable from plastochron five. The virus was able to escape from vascular restriction into photosynthetic tissues of young leaves before these tissues became source tissues, that is when the leaf tissues were still net importers of photoassimilates. This escape from vascular tissue probably coincides with the development of metaphloem elements, which have abundant plasmodesmal connections to facilitate symplastic loading of photoassimilates in mature leaf tissues. The virus most likely uses these plasmodesmal connections to invade photosynthetic cells in the leaf lamina, at about plastochron 12, before the emergence of the leaf from the whorl. Lucy *et al.* (1996) also found that MSV could escape from the vasculature via less abundant plasmodesmal connections in thin walled phloem complexes to gain access to the midvein and lateral bundles from the surrounding parenchyma. It will be of great interest to determine how the viral movement proteins function to bring about the pattern of virus distribution observed by these authors. Streak width is frequently greatly reduced in more resistant maize varieties, and in strains of MSV which cause mild symptoms in maize; it would be important to determine if the virus is more limited to vascular tissues in these cases, as this may indicate that resistance is a manifestation of a reduced ability of the virus to efficiently break its vascular limitation.

Encapsidation and Insect Transmission

The mechanisms involved in encapsidation of geminiviruses are not well understood, or investigated. It is not known if there is a specific encapsidation signal present on viral DNA, or whether coat protein encapsidates any circular ssDNA of approximately the right size. If there is a specific sequence which functions as an encapsidation signal, this must be present in the common region of Subgroup III geminiviruses, in order for both genomic components to be encapsidated. In this light it is interesting that an ACMV which was engineered to have the BCTV coat protein in place of its own was still infectious (Briddon *et al.*, 1990). The virus was transmissible by leafhoppers instead of whitefly, which proved that coat protein is the sole genetic determinant of vector specificity in geminiviruses. The fact that BCTV coat protein was able to encapsidate both DNA A and DNA B of ACMV

would seem to indicate that there is no specific DNA encapsidation signal, unless it is conserved between Subgroup II and III geminiviruses.

The question of whether two Subgroup I geminiviruses could *trans* encapsidate each other was addressed in a study by Margaret Boulton (1991) where she co-infected maize plants with MSV and DSV. These viruses have different leafhopper vectors: MSV is transmissible by *Cicadulina* species, mainly *C. mbila,* while the leafhopper vector of DSV is *Nesoclutha declivata.* Virus transmission by *C. mbila* apparently resulted in only MSV being transmitted to uninfected maize, which would seem to indicate that the MSV coat protein did not encapsidate DSV ssDNA. This issue certainly warrants further investigation.

The mode of leafhopper transmission of MSV, and presumably other Subgroup I geminiviruses, is a persistent, circulative and non-propagative one (Reynaud and Petterschmitt, 1992). As yet, the molecular factors involved in virus attachment at and transport across the leafhopper hindgut wall, entry into the haemocoel and transport to the salivary gland are uncharacterised. Reynaud and Petterschmitt (1992) showed that the titre of the virus maintained in the leafhopper decreases throughout the insect's lifetime, which provides some evidence that the virus does not replicate in the vector.

In conclusion of this section of my review of the litereature, since the last major review of the *Geminiviridae* (Timmermans *et al.*, 1994), research into the molecular biology of these fascinating viruses has expanded at a great rate, but several questions remain unanswered, particularly for Subgroup I. It is quite clear that further investigations will generate more valuable insight into fundamental issues in plant molecular biology, and greater understanding of geminivirus molecular biology will no doubt lead to novel mechanisms to control these potentially devastating pathogens. Apart from their academic interest, geminiviruses also hold great potential for application in biotechnology and as tools for solving various problems in plant molecular biology. These applications will be reviewed in the following section.

1.2.3. The Use of Geminiviruses in Biotechnology and Plant Molecular Biology

The use of geminiviruses as gene vectors has been reviewed by Mullineaux *et al.* (1992), Stanley (1993), and Timmermans *et al.* (1994). As this topic is central to this thesis, I too will review the use of geminiviruses as gene vectors, but will again focus on the applications of Subgroup I geminiviruses in plant biotechnology and plant molecular biology. From the wealth of information which has been obtained recently from basic research on geminiviruses, the potentials and limitations for use of geminiviruses as gene vectors are becoming clearer. There are several general thematic areas of plant molecular biology and biotechnology where geminiviruses find utility. These are:

- as markers and vectors for gene transfer;
- as vectors for transient or infectious expression of foreign genetic sequences in plants or plant cells;
- as episomal vectors in transgenic plants; and
- as a source of genetic elements which could be of use in expression of foreign genes in transgenic plants.

Geminiviruses as Markers for Gene Transfer

Most geminiviruses, apart from a few members of Subgroup III, are not mechanically transmissible as virions, or as cloned DNA. Grimsley *et al.* (1986) found that *Agrobacterium tumefaciens* could be used to efficiently transfer infectious clones of CaMV to plants; they extended this technique, termed "agroinfection" or "agroinoculation", to MSV and became the first authors to report facile transmission of MSV to maize without the use of leafhoppers (Grimsley *et al.*, 1987). This was the first good evidence that *Agrobacterium* could transfer its T-DNA to a graminaceous plant, and raised hopes that *Agrobacterium*–and even MSV–could be used to generate transgenic cereals. From a molecular virology point of view, this was also a major breakthrough in the study of Subgroup I geminiviruses, as it now allowed for genetic manipulation of cloned viruses, and re-introduction of these viruses into plants.

For agroinfection, a completely or at least partially tandem dimeric clone of the geminiviral genome is inserted in the transferred DNA (T-DNA) of a binary or co-integrate vector and

transferred into a strain of *Agrobacterium* which has the virulence genes intact (Grimsley *et al.*, 1989; for agroinfection methods, see Escudero and Hohn, 1994 or Boulton, 1995). A tandem multimer of the viral genome is required to facilitate escape of the cloned monomer from the T-DNA in plant cells, either by homologous recombination between the duplicated genomic sequences, or (more commonly) by a replicative release mechanism (Stenger *et al.*, 1992; Heyraud *et al.*, 1993b). Replicative release is mediated by transient expression of Rep, which recognises the viral origin of replication, nicks at the first stem-loop sequence and displaces a single stranded copy of the viral genome in a process very similar to normal rolling circle replication. Thus, the minimum highly agroinfectious clone would be approximately a monomer of the viral genome flanked by copies of the intergenic region origin of replication.

Agroinfection of Subgroup I geminiviruses represents a highly sensitive assay for transfer of T-DNA from *Agrobacterium* species to cereals, as the success of a rare event (T-DNA transfer) is amplified by viral replication and subsequent systemic spread. The technique has been used to great success to evaluate the efficiency of interaction between different strains and species of *Agrobacterium* and agriculturally important cereals. Generally, cereals are agroinoculated at or around the shoot apical meristem. Boulton *et al.* (1989a) found that nopaline strains of *A. tumefaciens* and agropine strains of *A. rhizogenes* could transfer T-DNA to maize far more efficiently than octopine strains, merely by evaluating the efficiency of agroinfection. These authors also found that nopaline strains could be used to agroinfect several other graminaceous species. Similarly, Marks *et al.* (1989) found that certain *A. rhizogenes* strains and nopaline strains of *A. tumefaciens* were more efficient in agroinfection of wheat with WDV, and Boulton *et al.* (1989a) and Creissen *et al.* (1990) showed that only *A. rhizogenes* was effective in agroinfection of barley. With the recent reports of recovery of transgenic rice (Hiei *et al.*, 1994) and maize (Ishida *et al.*, 1996) by *Agrobacterium*-mediated gene transfer, I predict that there will be increased interest in the use of agroinfection of geminiviruses for evaluating the efficiency of interaction between *Agrobacterium* strains and different cereal genotypes.

Agroinfection studies have also been very useful for investigation of the molecular biology of interaction between *Agrobacterium* and cereals. Schläppi and Hohn (1992) used agroinfection of MSV to evaluate T-DNA transfer from *Agrobacterium* to immature maize embryos, as integration of the T-DNA into totipotent cells would be required to achieve

stable transformation. The authors scored the successful transfer of T-DNA from *Agrobacterium* to the immature embryos by the number of germinated plants infected with MSV. There was a developmentally regulated susceptibility to agroinfection, with the youngest embryos tested (10 days after pollination, d.a.p.) not competent for agroinfection. The authors interpreted this as a requirement for some degree of differentiation in the shoot apex for successful T-DNA transfer. This phenomenon may represent a need for plant derived factors for *vir* gene induction, and/or *Agrobacterium* attachment. Recently, Escudero *et al.* (1996) found that 10 d.a.p. embryos are susceptible to agroinfection by *Agrobacterium* introduced intracellularly by microinjection, but this is plant genotype dependent. This reinforces how agroinfection with MSV as a genetic marker for gene transfer has revealed, and no doubt will continue to yield, important information about the interaction between maize and *Agrobacterium*.

Other studies have shown the value of Subgroup I geminiviruses as sensitive markers for direct gene transfer. Chen and Dale (1992) imbibed dry dissected wheat seeds with exposed shoot apical meristems in an aqueous solution of cloned dimeric WDV DNA, and achieved a 16% efficiency of infection. Although this efficiency seems quite good, they achieved a 79% infection rate by agroinfection by *Agrobacterium*, containing the same WDV dimer, by a similar imbibing protocol. In the same report, introduction of WDV into presoaked dissected seeds by microprojectile bombardment resulted in only 3% infection. Thus, infection by WDV clearly illustrates that *Agrobacterium*-mediated transfer represents the best route for introduction of foreign DNA into the tissue explant examined. Creissen *et al.* (1990) used replication of WDV as a sensitive marker for successful gene transfer to viable cells of barley microspore-derived cultures by microprojectile bombardment. Transient expression of marker genes like GUS or anthocyanin regulatory genes is frequently used to evaluate the success of direct gene transfer. The advantage of using geminivirus replication to assess direct gene transfer lies in the fact that transient gene expression often does not reflect that the cells are viable, or indeed still alive at a time after then gene transfer event. However, cells must presumably be physiologically sound to be undergoing enough DNA replication that one can detect on a Southern blot.

Geminiviruses as vectors for gene transfer

Initially, there was hope that agroinfection with geminiviruses could be used to generate transgenic plants by random integration of viral DNA carrying foreign genes into the plant chromosome. This phenomenon probably does occur at a low level, as suggested by Bejarano *et al.* (1996), who found evidence of ancient integration of geminiviruses into the genome of *Nicotiana.* However, geminiviruses are never seed-transmitted, which indicates that there may be some mechanism which precludes them from invading the germline. The investigation of MSV spread in infected plants by Lucy *et al.* (1996) showed that MSV was not found in the shoot apical meristem, presumably because of high cell division rates and the absence of vasculature routes for infection of this region. Grimsley *et al.* (1988) originally suggested that meristematic cells in the shoot apex were the cells infected with MSV by agroinfection. However, Shen *et al.* (1993) and Shen and Hohn (1994) later found that *Agrobacterium* preferentially transfers T-DNA into cells of leaf primordia which are already formed at the time of inoculation, and not into the meristematic region. Thus, if chromosomal integration of Subgroup I geminivirus-associated genes introduced by agroinfection occurs, it will be only in somatic tissues.

An ingenious way of encouraging insertion of DNA from a geminiviral vector into the chromosome of a cell in which it is replicating, is to promote active transposition of the foreign DNA from the viral vector into the plant genome. Laufs *et al.* (1990) used WDV to introduce the maize transposable element *Activator* (*Ac*) into wheat, rice and maize protoplasts. The *Ac* element was cloned in place of the WDV movement protein and coat protein genes. The investigators found that *Ac* efficiently excised from the viral genome and suggested that the transposition was dependent on DNA replication, which van Schaik and Brink (1959) had shown originally by classical genetic means. Recently, Wirtz *et al.* (1997) used replicating and non-replicating versions WDV replicons to prove definitively that transposition of deletion-derivatives of *Ac* (*Ds,* or *Dissociation* elements) require DNA replication for efficient transposition. Both sets of authors (Laufs *et al.*, 1990 and Wirtz *et al.*, 1997) showed that *Ds* elements, which are not capable of autonomous transposition, were mobilised in *trans* by *Ac* supplied on a non-replicating plasmid. The proven utility of the *Ac-Ds* system for transposon tagging, both in maize and heterologous plant systems (Chasan, 1993) led McElroy *et al.* (1997) to investigate whether it functions in barley. As barley is difficult to transform routinely, these authors used a transient assay for activity of

the *Ac* transposase in barley tissue, and, because of the requirement of DNA replication for *Ds* transposition (Wirtz *et al.* 1997), used engineered WDV replicons bearing *Ds* elements.

None of these authors (Laufs *et al.*, 1990; McElroy *et al.*, 1997; Wirtz *et al.*, 1997) determined whether the transposed *Ac* or *Ds* elements were integrated into the host chromosome. However, Sugimoto *et al.* (1994) constructed a recombinant MiSV genome with an artificial *Ds* element carrying a hygromycin resistance gene in place of V1 and V2. They introduced the recombinant MiSV into rice protoplasts together with *Ac* and found that the *Ds* element efficiently excised from the viral vector. The authors recovered transgenic hygromycin resistant calli, some of which seemed to result from insertion of *Ds* alone, but most contained MiSV-homologous sequences, which suggests random insertion of the viral vector into the host genome.

Experiments which have used geminivirus replicons to investigate transposition have given a good example of how geminiviruses may be used to investigate fundamental issues in plant molecular biology. Researchers from Barbara Hohn's laboratory have used MSV to introduce *Ds* elements into intact maize plants by agroinfection and have generated some useful data on how the *Ac*/*Ds* maize transposable element system works at the molecular level. Shen and Hohn (1991) showed that MSV could tolerate insertions of small oligonucleotides in the *Asn*I site found in the SIR, with little loss of infectivity. They expanded on this research by introducing the smallest known *Ds* element (*Ds1*: 405 bp) into the *Asn*I site, and introduced the recombinant virus into maize seedlings by agroinfection (Shen and Hohn, 1992). The virus containing the *Ds1* insertion was not agroinfectious on plants where there was no active *Ac* elements, but when *Ac* transposase was supplied from the plant genome, an MSV infection followed excision of the defective transposable element from the MSV DNA. This system can therefore be used as an assay for the presence of *Ac* transposase in any maize line. Investigation of the transposition "footprints" left in the MSV SIR showed that these were typical of those associated with *Ds* transposition in the maize genome (Shen and Hohn, 1992). All excisions were imprecise, and most resulted in deletions, complementary transversions and inversions of bases flanking the original *Ds1* insertion site. The value of this investigation was that it generated a larger number of sites to analyse the *Ds* transposition footprints than were available before. Transposition of *Ac*-group transposable elements is frequently associated with duplication of an eight base pair site at the insertion point. Shen *et al.* (1992) showed

from their analysis of the cloned MSVs isolated from infected plants that the duplications were not essential for excision of the transposable element. The same group recently extended their excision assay system to examine the *cis* sequence requirements in *Ds* elements for excision (Bravo-Angel *et al.*, 1995). They found that the *Ac* transposase binding motif in the 5′ subterminal region of *Ds1* is essential for excision. The authors were also able to modify the excision efficiency by mutagenesis at the 3′ end of *Ds1*. These studies have therefore proved very useful for determining *cis* requirements for *Ac*/*Ds* element transposition.

Geminiviruses as transient gene expression vectors

The most obvious advantage of linking genes of interest to geminivirus replicons is to achieve an increase in expression of the foreign gene, which should be associated with the increase in gene copy number. In the absence of routine stable transformation technology for cereals, several authors introduced recombinant Subgroup I geminiviruses into protoplasts to evaluate the effect which genome amplification has on expression of genes linked to the viral replicon. Recombinant WDV-based constructs with marker genes cloned in place of the coat protein gene, under the control of the virion sense promoter, replicated to high copy number in protoplasts isolated from *Triticum monococcum*, maize and rice (Matzeit *et al.*, 1991). The expression of the neomycin phosphotransferase II (*nptII*) gene was enhanced about 20-fold in replicating constructs as compared with non-replicating controls. Constructs containing the CAT or β-galactosidase genes also replicated in transfected protoplasts. The finding that WDV could tolerate large insertions of foreign DNA such as the 2.9-kb *Ac*-based transposable element (Laufs *et al.*, 1990) and the 3.0-kb β-glactosidase gene (Matzeit *et al.*, 1991) with no effect on replication, apart from some decrease in copy number, showed that there is no strict upper size limit imposed on replication of geminivirus constructs, and that size limitation may rather act at the level of viral spread. Suárez-López and Gutiérrez (1997) recently showed that increasing size of WDV-based vectors carrying the *gus* gene resulted in progressively decreased efficiency of replication. These authors also found that replication of the WDV-based vectors was required for maximal expression of the *gus* marker gene linked to the coat protein promoter, but did not quantify the effect of replication on the levels of GUS expression.

A more sophisticated variant of a WDV gene replacement vector was developed by Kamman *et al.* (1991a) and Ugaki *et al.* (1991). These authors both constructed plant-*E. coli* shuttle vectors by including an *E. coli* plasmid origin of replication (ColE1 and p15A, respectively) and a selectable marker which functions in bacteria in the gene replacement construct. Thus, the plasmids which had replicated in plant cells could be rescued by subsequent transformation of *E. coli*. Further modifications of the basic shuttle vectors could provide useful plant transposon traps; geminivirus constructs which are rescuable in *E. coli* might also be useful for studying DNA recombination and rearrangements in plant cells.

Ugaki *et al.* (1991) and Timmermans *et al.* (1992) introduced WDV-based vectors carrying the GUS gene under the control of the CaMV 35S RNA promoter into protoplasts derived from cultured maize endosperm cells. The level of GUS expression increased with replication of the constructs over time, as observed by Matzeit *et al.* (1991) and Hofer *et al.* (1992). However, the levels of GUS expression which Timmermans *et al.* (1992) observed were less than 10-fold greater than those produced from non-replicating control plasmids. This could be attributed to depletion of proteins like transcription factors or RNA capping enzymes which are required for optimal gene expression. The recombinant WDV-GUS shuttle vector replicated to an astounding 30 000 copies per cell, which is much greater than the 70 to 1000 copies reported for TGMV-based constructs (Hayes *et al.*, 1988b; 1989; Kanevski *et al.*, 1992).

Given that Rep protein of Subgroup III geminiviruses can act in *trans* to support the replication of the DNA B component, it is likely that Subgroup I Rep is also able to *trans*-replicate constructs which carry the appropriate origin of replication. This raises interesting possibilities for the design of a binary expression system based on a Subgroup I geminivirus, which would allow for the introduction of two or more viral-based replicons into the same cell. Theoretically, only one of these constructs would need an intact *rep* gene. Kamman *et al.* (1991a) first showed that this was possible with *rep* deletion mutants of WDV which were replicated in *trans* by wild type virus. The WDV-based shuttle vector construct (Ugaki *et al.*, 1991) with a CaMV35S-GUS expression cassette in place of the virion sense open reading frames, was made replication-deficient by the introduction of a frameshift mutation in the *rep* gene. The construct was efficiently replicated by another (replication proficient) WDV-based construct which was co-transfected into maize

endosperm cells (Timmermans *et al.*, 1992). The levels of expression of the GUS gene were similar to those obtained from the autonomously replicating construct in endosperm cells, but were considerably lower in BMS cells. The authors did not investigate this further, but suggested that the reduced levels of GUS expression in BMS cells might be due to some sort of replication interference due to limiting amounts of proteins required for viral replication in these cells compared with endosperm cells. The replication-proficient helper virus genome was considerably smaller than that of the GUS expression construct, so would probably replicate more efficiently in transfected BMS cells. It is interesting that the levels of marker gene expression in autonomously replicating Subgroup I geminivirus expression constructs in protoplasts or isolated plant cells seems to be between 10- and 20-times that from a non-replicating control (Matzeit *et al*, 1991; Timmermans *et al.* 1992), which is considerably lower than the 60- to 90-fold enhancement found for TGMV-based vectors in tobacco protoplasts (Brough *et al.*, 1992b).

Subgroup III geminivirus replicons have found use in analysis of transcription and RNA processing of linked genes in transient expression assays. Although there have been no reports of the use of Subgroup I geminiviruses for this application, there is no reason why these viruses could not also be used to evaluate expression of genes at the RNA level in cereal cells. Coat protein replacement constructs of TGMV cloned as tandem 1.5-mers in an *Agrobacterium* binary vector were used to evaluate transcription of the GUS and CAT genes from the coat protein promoter in tobacco leaf disks (Hanley-Bowdoin *et al.*, 1988). Similarly, McCullough *et al.* (1991) used the same TGMV coat protein transcription fusion vector to examine splicing of various plant introns in tobacco leaf disks. They found that introns from dicotyledonous plants (soybean B-conglycinin and pea rubsico small subunit, *rbc*S) were efficiently spliced in the tobacco leaf disk assay. An intron from wheat *rbcS* was also processed fairly efficiently (73% spliced) in tobacco, showing that this dicotyledonous plant is able to process an intron from a monocotyledonous plant gene. The authors feel that the geminivirus-leaf disk transient expression system has major advantages over the use of protoplasts for evaluating gene expression, as protoplasts are frequently physiologically degenerate, and for this reason expression of genes in protoplasts may not accurately reflect the situation *in planta*. Although it is not possible to use *Agrobacterium*-leaf disk co-cultivation for graminaceous plants, Subgroup I virus-based replicating constructs are easily introduced into tissue cultured cells and leaf tissues by microprojectile bombardment, and into protoplasts by transfection or electroporation.

More recently, Perriman *et al.* (1995) presented an interesting application of a Subgroup III geminivirus-based RNA expression system. These authors are investigating the application of hammerhead ribozymes in control of plant gene expression. Apparently, high levels of ribozyme expression are required for ribozymes to be effective, so the authors investigated linkage of the ribozyme transcription unit to an ACMV DNA A-based replicon. They had previously found that embedding a ribozyme into a tRNA sequence enhanced the ribozyme stability *in vivo*. As a model system, the authors had designed a hammerhead ribozyme which would cleave the CAT RNA sequence. Their ribozyme-tRNA expression construct, under the control of an RNA polymerase III promoter, was inserted into the ACMV replicon and transfected into tobacco protoplasts together with a CAT expression construct. Expression of the ribozyme from the replicating ACMV construct resulted in cleavage of CAT mRNA and significant reduction in CAT activity. This study represents the first use of an RNA polymerase III promoter in a recombinant geminivirus.

Geminiviruses as Infectious Gene Expression Vectors in Plants

The finding that the coat protein gene of Subgroup III geminiviruses like ACMV and TGMV is not required for systemic infection of tobacco (Stanley and Townsend, 1986; Brough *et al.*, 1988; Gardiner *et al.*, 1988), suggested that this gene could be replaced with a foreign gene, for high level expression in non-transgenic plants mediated by infectious recombinant geminiviruses. This, and several other applications of geminiviruses as gene vectors, was suggested by Rogers *et al.* in a 1986 patent application (EP 0 221 044 B1, granted in 1992). Ward *et al.* (1988) replaced the coat protein of ACMV with the CAT gene, generating a recombinant DNA A about 70 bp larger than the wild type. They manually inoculated *N. benthamiana* with the recombinant DNA A construct, together with DNA B. The recombinant virus was infectious in tobacco, and resulted in symptoms similar to the wild type virus. The authors achieved high levels of CAT enzyme expression (80 U/mg soluble protein) in the infected tobacco plants, but did not comment on how this compared with levels of CAT which may be achieved in conventionally-produced transgenic tobacco plants. In a similar approach, Hayes *et al.* (1988b) replaced the coat protein ORF of TGMV with the *neo* gene, which encodes *nptII*. The movement functions were supplied by DNA B, either as a dimeric clone which was co-agroinfected with the recombinant DNA A, or integrated as a tandem dimer into the genome of tobacco plants

into which the DNA A was introduced by agroinfection. Four out of twenty co-agroinfected plants, and 18 out of 20 agroinfected transgenic plants became infected with TGMV. The average *neo* gene copy number corresponded to about 160 copies per amphidiploid tobacco genome for the plants infected by DNA A-DNA B co-agroinfection. A higher copy number of about 490 per genome was achieved by agroinfecting the transgenic DNA B plants. The authors found that the expression of NptII was enhanced considerably over the level expressed in plants transgenic for a non-replicating DNA A monomer: 23-fold for the co-agroinfected plants, and 69-fold for the agroinfected transgenic DNA B plants. Apart from affording significantly increased expression of foreign genes in infected plants, these studies suggest that this approach might be useful for evaluating the expression of foreign gene constructs and the phenotype conferred by the expression of foreign genes before making transgenic plants.

Further investigation of the use of Subgroup III geminiviruses for infectious expression of foreign genes, however, provided a caveat: this approach will probably only work efficiently if the size of the recombinant DNA A construct is kept close to the wild type. This would limit the size of proteins which can be expressed in this system to about 30 kDa. When Elmer and Rogers (1990) agroinfected *N. benthamiana* plants with a GUS-coat protein replacement mutant of DNA A, together with DNA B, they found that the intact GUS construct did not move systemically in infected plants; rather, all infected plants contained deleted and rearranged versions of the original GUS replacement clone. This clone was structurally stable in agroinfected leaf disks, but not in plants, which suggests that there is a strict upper size limitation on systemic spread in *N. benthamiana* plants. In contrast, Hayes *et al.* (1989) found that recombinant DNA A constructs larger than the wild type genome were more stable than observed by Elmer and Rogers (1990) when they were used to agroinfect transgenic *N. tabaccum* plants with an integrated dimer of DNA B to provide the viral movement functions. Modified TGMV DNA A *neo* vectors with the CaMV 35S promoter inserted in front of the *neo* gene were infectious in transgenic tobacco containing DNA B dimers, but some of the replicating recombinant *neo* vectors did undergo deletions in the foreign gene sequences. The authors still noted increases in NptII activity of about 100-fold in the agroinfected plants relative to transgenic plants containing the *neo* gene under the control of the CaMV 35S promoter. The DNA A-35S *neo* constructs were about 3.5 kb in size. Larger recombinant DNA As with the GUS gene in place of the coat protein ORF were less stable than the *neo* constructs, but nevertheless

mediated significant enhancement of GUS expression: between 18- and 100-times that of a transgenic plant expressing GUS from the coat protein promoter of a monomeric TGMV construct. It is not clear why the two groups found such different behaviours associated with the movement of the GUS gene replacement constructs in infected plants. It is possible that the larger constructs were able to move more easily in plants in which the DNA B was provided from an integrated master copy in the chromosome, or that there are different size stringencies associated with TGMV movement in *N. tabaccum*, used by Hayes *et al.* (1989), compared with *N. benthamiana*.

Monopartite geminiviruses present a problem in the design of infectious gene expression constructs since all genes are essential for systemic infection. This was originally proved by Lazarowitz *et al.* (1989), who, in separate experiments, replaced the MSV coat protein gene, and both virion sense ORFs with the CAT gene, or a hygromycin resistance cassette. Partial multimers of the recombinant MSV genome were used to agroinfect maize seedlings. None developed typical MSV symptoms, but a few plants contained detectable levels of replicating recombinant viral DNA in the first two or three inoculated leaves. In addition, some of the plants agroinfected with constructs containing the CAT gene in place of the MSV coat protein expressed low levels of CAT enzyme.

With the information that MSV constructs with small oligonucleotide insertions in the *Asn*I site of the SIR are sometimes still infectious, Shen and Hohn (1994; 1995) tried to construct infectious MSV-based gene vectors by inserting larger marker genes in the *Asn*I site in the SIR of the MSV genome, which otherwise remained completely intact. To this end, these authors initially inserted a GUS expression cassette into the SIR of an agroinfectious MSV genome. Unfortunately, although replication of the 5.9-kb construct could be detected in the first two or three leaves of some agroinfected plants, there was no evidence of systemic infection. The authors evaluated the success of the agroinfection experiment by histochemical staining for GUS activity in leaves from agroinfected seedlings. The number of blue GUS-positive spots on the leaves of plants agroinfected with the replicating vector was significantly higher than in plants which were agroinoculated with *A. tumefaciens* carrying only the GUS expression cassette. One week after inoculation, the mean number of spots per plant was 4.5 times higher in plants agroinfected with the MSV-GUS construct than those inoculated with the non-replicating control, and 7.3 times the control number after two weeks, suggesting that the GUS

expression from the replicating MSV construct was more stable than the non-replicating control. The authors also agroinfected seedlings with a smaller MSV construct, where they deleted the virion sense ORFs. The number of GUS-positive spots from this construct (MSV-D-GUS) was slightly higher than that obtained from the larger construct which had the movement and coat protein genes intact, which implies that the smaller construct was able to replicate more efficiently than the larger one. However, in both cases, most spots were visible on the first or second leaves, and the highest leaf that they found GUS activity in was leaf 5. This shows that neither of the constructs was able to move significantly, and implies that there is also some size limitation imposed on movement of Subgroup I geminiviruses. This could be due either to encapsidation constraints, if encapsidation is required for long distance movement, or to some sort of limitation on the size of DNA which can be moved systemically as a nucleoprotein complex, as would seem to be the case for Subgroup III geminiviruses.

In a variation on the experiments with the GUS expression cassette, Shen and Hohn (1995) inserted a herbicide resistance (*bar* gene) cassette into the SIR of MSV. The *bar* gene, which encodes phosphinothricin acetyl transferase (PAT), is very widely used as a selectable marker in transgenic cereals and confers tolerance to phosphinothricin-based herbicidal compounds such as bialaphos, Basta™ or Ignite™. The *bar* gene is potentially more suited for application as a marker gene in MSV-based constructs as it is small (the coding region is about 550 bases), and results in herbicide tolerance, an easily screenable phenotype in plants. The MSV-*bar* construct had a *bar* expression cassette inserted into the *Asn*I site in the SIR, with the CaMV35S promoter and terminator providing transcriptional control sequences. The size of this construct was 5.6 kb; the MSV-D-*bar* construct had the virion sense ORFs deleted, as before, and was about 1 kb smaller (Shen and Hohn, 1995). Neither of these constructs was capable of systemic movement even though the movement and coat protein genes were intact in MSV-*bar*. However, some seedlings agroinfected with either of these constructs were tolerant to application of herbicide, which shows that sufficiently high levels of PAT were expressed to inactivate the herbicidal compound. Larger seedlings did not remain herbicide tolerant, confirming that the replication of the recombinant viruses was limited to the first inoculated leaves.

Given that Subgroup I geminiviruses have no genes which are dispensible for infectivity, and that there seems to be a size limitation on spread of Subgroup I virus genomes, the only

option left for generation of infectious gene vectors is to complement movement and/or replication functions in *trans*. This could be achieved in two different ways: (1) by co-infecting a gene replacement mutant with a virus which can supply the deleted functions in *trans*; or (2) by supplying replication or movement functions from transgenic plants. Neither of these options has been explored in any detail for Subgroup I geminiviruses. Boulton *et al.* (1989) found that they could complement point mutations or small oligonucleotide insertions or deletions in the V1 and V2 ORFs by supplying wild-type virus or complementary mutants in *trans*. However, the complementing mutants always eventually recombined to generate wild type virus. Similarly, Lazarowitz *et al.* (1989) tried to complement movement functions between a virus with a point mutation in V1 and a coat protein gene replacement mutant, which was 114 bp larger than the wild type virus. Of over 500 agroinoculated seedlings, they obtained only one infected plant, which, presumably as a result of recombination, contained wild type virus DNA.

Complementation of movement or replication functions in transgenic plants

To remove or reduce the chances of recombination between complementing virus constructs, it is possible that plants expressing the appropriate viral gene or genes from copies integrated into the plant genome could be used to complement movement and/or replication-deficient virus constructs. Transgenic tobacco plants which contain integrated dimers of DNA B can supply DNA B movement functions which obviates the need for co-agroinfection of DNA A (or DNA A gene replacement mutants) with DNA B. This does not strictly constitute complementation by transgenes, as the DNA B is merely rescued from the genome and replicated by DNA A (Rogers *et al.*, 1986a; b; Hayes *et al.*, 1988b; 1989).

Transgenic plants expressing BV1 or BC1 of SqLCV do not complement movement of virus constructs which contain mutations in these genes (Pascal *et al.*, 1993), and tomato plants expressing the TYLCV-Israel coat protein show resistance to virus infection (Kunik *et al.*, 1994). Expression of movement and coat protein genes is probably a virus "late" function, and the timing of expression should thus be tightly controlled, so it is not surprising that expression of these genes in transgenic plants did not complement mutations, and even conferred some resistance to virus infection. In contrast, transgenic tobacco plants expressing all of the complementary sense ORFs of TGMV from a single

promoter could in some cases complement viruses which contained mutations in these genes (Hanley-Bowdoin *et al.*, 1989). Similarly, Hanley-Bowdoin *et al.* (1990) found that transgenic plants expressing TGMV Rep protein, even at a level three times that found in a natural infection, could complement mutations in the *rep* gene of TGMV, such that these plants could support both replication and systemic movement of a replication-deficient TGMV mutant. This finding would seem to indicate that expression of Subgroup I Rep in transgenic plants should allow complementation of replication-deficient gene replacement mutants with a foreign gene in place of the viral *rep* gene. However, transgenic plants expressing TGMV Rep were very difficult to generate and maintain and were not stable for Rep expression through multiple generations (Linda Hanley-Bowdoin, personal communication). Recently, Bendahmane and Gronenborn (1997) reported that they found it impossible to generate transgenic tomato plants which expressed sense constructs of the the TYLCV-Sr *rep* gene. This is probably due to the effect which Rep has on the cell cycle, by inactivation of plant Rb homologues and induction of enzymes involved in DNA synthesis (Nagar *et al.*, 1994; Xie *et al.*, 1995; 1996).

In contrast to the above, the expression of ACMV Rep in transgenic plants resulted in resistance to virus infection, rather than complementation of Rep-deficient mutants (Hong and Stanley, 1996). The results of the latter study suggest the possibility that resistance was the manifestation of some sort of co-suppression phenomenon, but the authors did not rule out the possibility that they may have selected for mutant forms of the *rep* gene in the plants which were successfully regenerated. So, it seems unlikely that either viral movement or replication functions can be efficiently complemented by the expression of the viral genes in transgenic plants.

Gene amplification from integrated vectors

The concept of using geminivirus replicons to amplify foreign gene copy number in transgenic plants was first introduced by Rogers *et al.* (1986a), who reported that transgenic petunia plants which carried tandem dimeric inserts of TGMV DNA A in the genome contained episomes of autonomously replicating DNA A monomers. This finding suggested that dimers of autonomously replicating derivatives of any geminivirus genome could be integrated into the genome of a transgenic plant, where an extrachromosomal copy of the recombinant viral DNA could be mobilised from the integrated template,

probably by replicative release, and amplified by autonomous replication. Hayes *et al.* (1988b) used this approach to generate transgenic plants with their autonomously replicating derivative of TGMV DNA A with the *neo* gene in place of the coat protein ORF. Transgenic plants contained an average of 75 copies of the mutant DNA A per amphidiploid tobacco genome, replicating as episomes. These transgenic plants also expressed about 10-fold more NptII enzyme than control plants with the same *neo* gene construct. An identical coat protein gene replacement construct with the GUS gene in place of the coat protein replicated to a level of between 85 and 140 copies per host genome, and GUS expression was enhanced between 19 and 36 times the level expressed in control plants (Hayes *et al.* 1989). Similarly, transgenic plants with a CaMV35S-*neo* expression cassette linked to the DNA A replicon had an average of 100 copies of the recombinant DNA A episome per tobacco genome, and expressed 24-fold more NptII than control plants.

A major advantage of this approach for expression of genes in transgenic plants is that the extrachromosomal location of the gene of interest would reduce position effects which are frequently responsible for variation in transgene expression levels between lines of transgenic plants. Indeed, Hayes *et al.* (1989) noticed little variation in expression of the GUS gene standardised to copy number between different transgenic lines in which the recombinant DNA A is replicating. This contrasts with findings of Meyer *et al.* (1992) for ACMV DNA A-derived constructs: these authors found significant variation in copy number between different transgenic lines, which might be due to the chromosomal integration position affecting the frequency of episome mobilisation. In all cases, the episome copy number seems very low compared with a wild-type infection. It is not clear whether the low copy number is due to some down-regulation of episome copy number which occurred during regeneration of transgenic plants, or whether this reflects a low frequency of mobilisation of the integrated copy, so that the overall copy number represents an average of widely varying copy numbers in different cells.

Transgenic tobacco cell lines were generated with autonomously replicating TGMV DNA A-based gene replacement constructs by introduction of partial dimers into tobacco suspension cultured cells by microprojectile bombardment (Kanevski *et al.* 1992). Once again, these constructs had the coat protein coding sequence replaced by the kanamycin resistance gene (*nptII*), or an *nptII* expression cassette under the control of the

A.tumefaciens nos gene promoter and terminator sequences. The authors selected for kanamycin resistant callus lines which were screened for the presence of the recombinant TGMV-based vectors as high copy number episomes in the extrachromosomal DNA. About ten percent of the transgenic cell lines contained TGMV-based episomes, in three different copy number classes: I: 10 to 100 copies; II: 100-500 copies and III: 500 to 1000 copies per cell. The expression of NptII by constructs with the *nptII* gene transcribed under the control of the *nos* promoter was roughly correlated with copy number, with three lines in class I expressing between four and nine times the amount of NptII than the control (non-replicating) line, and two class III lines expressing 28 and 31 times the control levels, respectively. The copy number of the replicating episomes decreased substantially over time, so that after six months, class III lines had copy numbers previously found in class I, and lines which initially had lower copy numbers of viral epsiomes no longer contained replicating extrachromosomal DNA. The authors speculated that the reduction in copy number observed in the transgenic tobacco lines might be attributed to methylation of the master copy integrated into the chromosome.

Cytosine methylation can result in reduced replication of geminiviral constructs by two distinct mechanisms: direct interference in replication, and interference in viral gene expression (Brough *et al.*, 1992a; Ermak *et al.*, 1993). The reduction in copy number of episomes relative to that seen in normal viral infections could be due to host control of replication levels by methylation. According to the authors (Rogers *et al.*, 1986; Hayes *et al.*, 1988b; 1989; Meyer *et al.*, 1992), transgenic plants containing replicating virus DNA seemed phenotypically normal, but there was no mention of whether the primary transgenics reported on were fertile, and whether the replication of the DNA A-derived episomes was maintained in subsequent generations. One would anticipate that there must have been some adverse effect on the viability of these plants mediated by Rep expression, as was found by Hanley-Bowdoin *et al.* for their Rep-expressing transgenic plants (personal communication). There have been no published reports of the use of Subgroup I geminivirus replicons as gene amplification vectors in transgenic plants or cell lines.

Geminiviruses sometimes accumulate subgenomic DNAs which are deletion derivatives of their genomic components; some of these act as defective interfering (DI) genomes. Stanley *et al.* (1990) investigated the potential of using a DI DNA derived from ACMV DNA B in a transgenic resistance strategy. A tandem repeat of an ACMV-derived DI DNA

was integrated into the chromosome of transgenic *N. benthamiana* plants, which was rescued and amplified by inoculation of cloned virus. Presumably the rescuing virus Rep recognises the origin of replication in the dimer of the DI DNA, and mediates replicative release of the DI DNA from the chromosome. Smaller viral genomes are logically replicated faster, and potentially at the expense of the full-length viral genome, thus the transgenic plants showed significant symptom amelioration compared with controls. This approach for engineering greater resistance to geminiviruses has also been applied successfully to BCTV (Frischmuth and Stanley, 1994; Stenger, 1994). There is potential for this strategy to be applied to a viral genome amplification strategy for overexpression of genes of interest in transgenic plants: this would simply require the production of transgenic plants with integrated tandem dimers or at least partial multimers with two intergenic regions of a recombinant replication-deficient virus construct. The integrated construct should be rescued and amplified by wild type virus, in the same way as ACMV and BCTV DI DNAs. Theoretically the only viral sequences required in the recombinant construct would be the sequences needed in *cis* for replication: the common region for Subgroup II and III geminiviruses, and the LIR and SIR for Subgroup I viruses. There are no reports as yet on the use of this approach for development of an inducible or rescue-based geminivirus gene amplification system.

Geminivirus genetic elements for expression of foreign genes in plants

Plant pararetroviruses like CaMV, figwort mosaic virus (FMV), rice tungro bacilliform virus (RTBV) and Commelina yellow mottle virus (CoYMV), have been useful sources of strong constitutive and tissue specific promoters and transcription termination and polyadenylation signals (reviewed by Mushegian and Shepherd, 1995). The CaMV 35S promoter is certainly the most widely used in transgenic plants. There was initially hope that geminiviruses would also provide strong constitutive promoters for transgene expression, but geminiviral promoters are apparently tightly regulated according to the stage of the virus life cycle, and hence use of these promoters in transgenic plants divorced from the normal viral regulatory mechanisms may be of little use. However, the phenomenon of TrAP-mediated transactivation of the Subgroup III geminivirus coat protein promoter could be exploited to great effect in enhancing tissue specific gene expression in transgenic plants, and geminiviral transcription termination and polyadenylation signals may also prove useful additions to a gene expression "toolkit".

A patent application by Kridl *et al.* (1994) outlines several uses that TrAP-mediated transactivation could have in plant biotechnology. The inventors propose that the coat protein promoter of a Subgroup III geminivirus could be used to drive the expression of a gene required to be expressed at high level in a specific tissue of a transgenic plant, for example a gene involved in seed development, or fruit ripening, or development of seed coat hairs like cotton fibres. The TrAP gene would be expressed in the same transgenic plants under the control of a tissue-specific promoter; TrAP would then transactivate the viral coat protein promoter, specifically in the tissue in which TrAP is produced. The coat protein promoter construct could be present in the transgenic plants either in the context of a replicating episome, or without Rep-mediated amplification.

The coat protein promoter transactivation concept was recently used by Hong *et al.* (1996) for a different application: in engineering virus-induced cell death, as a means of achieving geminivirus resistance in transgenic plants. These authors first constructed transgenic tobacco plants with the GUS gene under the control of the ACMV coat protein promoter. In some transgenic lines, they showed that when TrAP was supplied by infection of transgenic plants with ACMV, the coat protein promoter was induced up to 100-fold. The authors then constructed transgenic plants with the coat protein promoter driving the expression of dianthin, a ribosome inhibiting protein. The rationale for this experiment was that transactivation of the coat protein-dianthin expression cassette by virus infection would lead to death of infected cells–an engineered hypersensitive response, which would limit virus spread. When challenged with ACMV, transgenic plants produced necrotic lesions on infected leaves, which is not typical of ACMV infection, and indicates dianthin-induced cell death. Transgenic plants generally accumulated reduced levels of ACMV DNA and attenuated symptoms compared with the controls, and frequently recovered from the virus infection.

The question of Rep-mediated transactivation of the virion sense promoter of Subgroup I geminiviruses still remains open. It is logical that this should also occur in this genus of the *Geminiviridae*, and if it does occur, development of an engineered hypersensitive response system analagous to that of Hong *et al.* (1996) for engineered Subgroup I geminivirus resistance is also an attractive possibility. It is also possible that a system for exploiting transactivation of the Subgroup I viral promoter in transgenic cereals could be

designed for controlling heterologous gene expression. A problem one would have in designing a gene expression enhancement system such as that proposed by Kridl *et al.* (1994) for this subgroup is that the Rep transactivation effects would need to be separated from the effects which this protein has on the cell cycle, unless the *rep* gene is expressed under a very tightly regulated promoter. For this application, the domains of Rep involved in virion-sense promoter transactivation should be identified, as it may be possible to separate the transactivation and cell cycle regulatory functions. Even very low levels of Rep expression could prove refractory to regeneration of transgenic plants, and at the least might adversely affect fertility of the plants.

In conclusion of this review of the literature, it is obvious that research into the molecular biology of geminiviruses has yielded, and no doubt will continue to yield, very useful insight into the control of plant gene expression and DNA replication. There is great potential for exploitation of aspects of geminivirus biology in applied plant molecular biology and biotechnology of plants, but before these viruses can be used to their full potential there is a great deal of basic research which needs to be done. Subgroup I geminiviruses remain less well explored than Subgroup III viruses, and therefore present an exciting challenge to molecular virologists. Investigation of the control of viral gene expression in these small and seemingly simple viruses will show how they cause disease with such a small genome, which surely must require many intricate genetic regulatory mechanisms, many of which may find use in biotechnology.

1.3 PROJECT AIMS AND BACKGROUND

MSV is probably the most economically significant Subgroup I geminivirus: it is considered to be the most important pathogen of maize in Africa where it can be responsible for devastating crop losses. As genetic manipulation and investigation of MSV is the subject of this thesis, it is appropriate that I give a short overview of this virus and its potential for use as a gene vector, and the particular isolate of MSV which I used (MSV-Kom), before I present the results of my investigations into the use of maize streak virus as a gene vector.

MSV is the causal agent of maize streak disease, which was first described as "mealie variegation" in 1901 by Fuller, the government entomologist in Natal, now part of South Africa. Symptoms on susceptible maize and other cereal hosts are manifested by long chlorotic streaks which may extend over the entire leaf surface, and severe stunting, depending on the stage at which the plant was infected. MSV is obligately transmitted by leafhoppers of the genus *Cicadulina*; in Southern Africa, the most important leafhopper vector is *C. mbila* Naude ("mbila" is the Zulu word for maize). Cytologically, virus-like particles accumulate in the nuclei of infected cells, typically in large inclusions in mesophyll and bundle sheath cells, and sometimes also as nuclear inclusions in vascular-associated tissues (Pinner *et al.*, 1993). Virus-like particles and genomic DNA are found in all leaf tissues, except the epidermis and xylem (Lucy *et al.*, 1996).

Research on the molecular diversity of cereal infecting geminiviruses has shown that maize streak disease is caused by a heterogeneous, but closely related, group of MSV strains (Briddon *et al.*, 1994; Hughes *et al.*, 1992 and Rybicki, unpublished results). The genomic sequences of four closely related MSV strains are published: MSV-**N**igeria (Mullineaux *et al.*, 1984), MSV-**K**enya (Howell, 1984; 1985); MSV-**S**outh **A**frica (Lazarowitz, 1988); and MSV-**R**eunion (Petterschmitt *et al.*, 1996). The isolate of MSV I used in the investigations presented here is MSV-**Kom**, a moderately severe isolate from Komatipoort, Mpumalanga Province, South Africa, which was sequenced and characterised in this laboratory (W. H. Schnippenkoetter *et al.*, in preparation). This virus's closest sequenced relative is MSV-SA (Lazarowitz, 1988). In sweetcorn, cv. Jubilee, MSV-Kom causes moderately severe disease. The virus infects many maize cultivars, and is also infectious in some wheat and barley cultivars, and in *Digitaria sanguinalis* and *D. setigera* (W.H. Schnippenkoetter, M.B. von Wechmar and E.P. Rybicki, unpublished). The clone of the virus genome, pKom500, is agroinfectious in sweetcorn and a variety of maize cultivars (D.P. Martin and E.P. Rybicki, unpublished). I have shown a representation of the MSV-Kom genomic organisation in Figure 1.8., below.

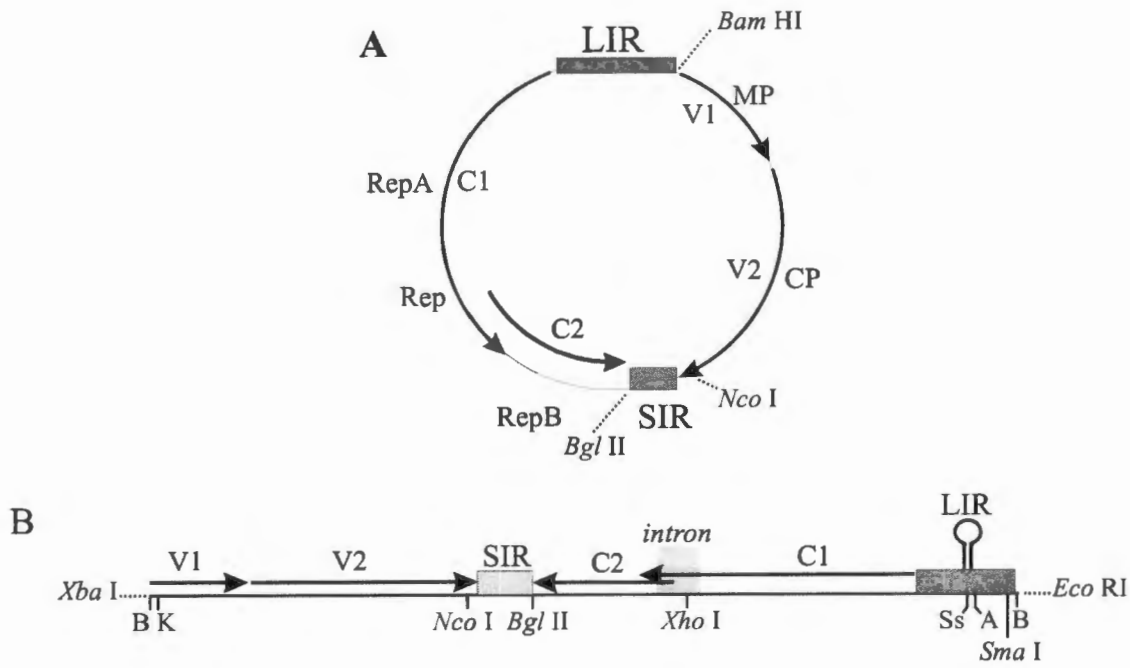


Figure 1.8: Diagrammatic Representations of the MSV-Kom Genome (2690 base pairs).

Abbreviations: A: *Apa*I; B: *Bam*HI; K: *Kpn*I; Ss: *Ssp*I
A. The circular RF-DNA of MSV-Kom, with long and short intergenic regions (LIR and SIR) and four open reading frames. V1 encodes the movement protein; V2 is the coat protein ORF, C1 and C2 together constitute the *rep* gene. B. The genome of MSV-Kom linearised at the unique *Bam*HI site which overlaps the second amino acid of the movement protein gene (the sequence in this region is ATGGATCC, where GGATCC is the *Bam*HI recognition sequence). The 5′ G of the *Bam*HI site is nucleotide 1 of the genome. Restriction sites often used in manipulation of the MSV-Kom genome in this study are shown. The *Nco*I site is 42 bases before the stop codon of V2, at position 1005; *Bgl*II overlaps the stop codon of C2; *Xho*I occurs in the intron sequence, at position 1686; *Sma*I is at position -34 relative to the presumed transcription start site of V1. The genome of MSV-Kom was cloned as a *Bam*HI fragment in the *Bam*HI site of pUC19 to give pKom500. The insert is cloned such that the vector *Eco*RI site is at the LIR side and the *Xba*I site is on the V1 side.

It should be clear from my review of the literature that, apart from their well-characterised use in agroinfection, the use of Subgroup I geminiviruses as gene vectors for plant molecular biology and biotechnology applications has not been very intensively explored as yet. The main aim of my research was to investigate whether Subgroup I geminiviruses could be used in a gene amplification system in transgenic cereals. When I took up this challenge in 1993, the technology for transforming cereals was still in its infancy. Of the cereal crops of economic importance in South Africa, only maize had been successfully transformed, by groups in two American biotechnology companies (Fromm *et al.* and Gordon-Kamm *et al.*, reported in September 1990). Several authors had previously shown that Subgroup III geminiviruses, which infect dicotyledonous plants, could be used as gene amplification systems in transgenic plants, but there is still no published report of the use

of a cereal infecting geminivirus (Subgroup I) as a gene amplification system in transgenic cereals. This probably reflects the recalcitrance of cereals to transformation, which has only recently been overcome, but which is still a fairly difficult and tedious process. The first major challenge was therefore the establishment of a maize tissue culture and transformation facility at the University of Cape Town. This I achieved with Sandy Lennox, also of the Microbiology Department.

The major focus of the research reported in this thesis was to establish whether MSV replicons could be used to amplify genes, both transiently and in stably transformed maize cells; my investigations in this regard are presented in Chapters 2 and 3. I have described a separate, but not exclusive, series of experiments on my investigations into whether MSV could be used as an infectious gene expression system in agroinfected maize seedlings in Chapter 4. As presented, this certainly does not represent a chronological progression of my Ph.D. research, but rather the cohesion of diverse parallel investigations into the general topic.

CHAPTER 2

GENERATION OF TRANSGENIC CELL LINES CONTAINING MSV-BASED AUTONOMOUSLY REPLICATING GENE VECTORS

TABLE OF CONTENTS

SUMMARY

This chapter examines the feasibility of obtaining stably transformed maize cells containing autonomously replicating gene vectors based on MSV. I constructed gene replacement derivatives of MSV-Kom by replacing the virion sense genes with selectable marker expression cassettes which contained the bialaphos resistance gene (*bar*) under the transcriptional control of the CaMV 35S RNA promoter. Sequences within the right hand side of the LIR, between the conserved stem-loop sequence and the V1 start codon, were necessary for function of the MSV origin of replication, but the sequences between the virion sense promoter TATA box and the V1 start codon 34 bp downstream were dispensable for replication. I tested the effect of increasing genomic size on replication by modifying the *bar* expression cassette by including: (1) the 550-bp maize *adh* I intron and 68-bp TMV Ω RNA leader sequences in the 5′ untranslated leader of the *bar* gene; and (2) a gene fusion between the *bar* ORF and the *E. coli* glutathione reductase (*gor*) gene. Recombinant viral vectors ranging in size from 2.69 kb to 4.8 kb replicated efficiently in cells transfected by microprojectile bombardment. I generated transgenic maize cell lines in two different model systems: regenerable HiII callus and suspension cultured Black Mexican sweetcorn (BMS) cells. Although bialaphos-resistant transgenic HiII plants were regenerated from control lines which did not contain MSV DNA, I was unable to regenerate plants from transgenic callus containing high copy number viral episomes carrying the *bar* gene. Transgenic bialaphos-resistant BMS cell lines were generated with MSV-based vectors containing the *bar* gene (with and without the maize *adh*I intron and TMV Ω RNA leader), and with a *bar-gor* gene fusion. Between 38 and 60% of transgenic cell lines contained replicating viral episomes carrying the *bar* gene. The replicons seemed structurally stable, and replicated to copy numbers of over 500 per haploid genome in transgenic cell lines for more than one year after they were initially introduced. I noted significant enhancement of *bar* gene expression, both at the protein and RNA levels, which was associated with the presence of episomal vector DNA. The maize *adh*I intron and TMV Ω RNA leader sequences are known to enhance transient gene expression, but these elements did not seem to have a significant effect on expression of the *bar* gene from replicating viral constructs. The results suggest that MSV may provide a useful system for gene amplification in transgenic cereal cells.

2.1. INTRODUCTION

Given the extremely broad range of plant species which different members of the *Geminiviridae* infect, geminivirus-based extrachromosomal gene amplification systems could probably be used to enhance foreign gene expression in virtually any agronomically important crop species. However, to date, the only reports of recombinant geminivirus-based gene amplification systems in transgenic plants are of TGMV- or ACMV-based vectors in tobacco plants (Hayes *et al.*, 1988b; 1989; Meyer *et al.*, 1992) or suspension-cultured cells (Kanevski *et al.*, 1992). Apart from the successes demonstrated with Subgroup III geminiviruses, there have been no reports of the use of Subgroup I geminiviruses as extrachromosomal gene amplification systems in transgenic plants. This probably reflects the technical difficulties associated with transformation of cereals, which have only recently been overcome (reviewed by Christou, 1996).

The aim of the research I will present in this chapter was to investigate whether MSV, as a representative Subgroup I geminivirus, could be used to significantly amplify the copy number and expression level of genes in transgenic maize. In this regard, it would be instructive to review the history of maize transformation. The first report of stable transformation of maize was by Fromm *et al.* in 1986. They transformed the non-regenerable Black Mexican sweetcorn (BMS) cell line (Sherridan, 1982) by protoplast electroporation. Subsequently, Klein *et al.* (1989a) and Spencer *et al.* (1990) showed that transformation of suspension cultured BMS by microprojectile bombardment was straightforward and efficient, and probably far more convenient than electroporation since particle bombardment obviated the need to generate protoplasts. The facility with which BMS cells can be transformed has made this cell line very useful as a model system for transformation and gene expression studies in maize.

However, in contrast to BMS, production of transgenic maize plants has proved more technically challenging. The first report of regeneration of transgenic maize from protoplasts was by Rhodes *et al.* (1988); these plants were unfortunately sterile, probably because of the large amount of tissue culture manipulations required for their regeneration. Two sets of researchers then simultaneously reported the production of fertile transgenic maize plants in September 1990 (Fromm *et al.*, 1990 and Gordon-Kamm *et al.*, 1990).

Both used friable embryogenic callus cultures ("Type II") of a non-elite hybrid (A188 x B73) for delivery of marker genes by microprojectile bombardment. Armstrong *et al.* (1991) reported the development, from A188 x B73 crosses, of maize genetic stocks called HiII, because of the high frequency of initiation of type II callus formation from immature embryos (8 to 12 days after pollination, d.a.p.) seen in this germplasm. Due to the relative ease of culturing the type II callus and of transformation and regeneration of plants, HiII germplasm now forms the basis of most routine maize transformation experiments; laboratories commonly use either embryogenic callus or immature embryos derived from HiII for "biolistic" gene delivery. Transformation of other maize lines which generate more structured "type I" callus is considerably more difficult. Although there have recently been reports of the use of alternative gene delivery methods, most significantly by *Agrobacterium*-mediated gene transfer (Ishida *et al.*, 1996), microprojectile bombardment of callus or immature embryos remains the most broadly applicable method for transformation of maize.

The methods routinely used for transformation of the model transgenic maize systems, BMS and HiII, are quite similar. In both cases, friable callus is used for biolistic delivery of the desired genes. The BMS callus is almost always grown in suspension culture, and is non-embryogenic and not regenerable. One needs to select embryogenic HiII callus for microprojectile bombardment; this callus is not usually grown in suspension culture, but is grown on solid media, and friable embryogenic callus is selected at each subculture. In both cases, a thin layer of callus tissue is spread on filter paper disks or on the surface of solid media before bombardment, usually with 1 μm gold particles carrying plasmids with genes of interest. For transient expression studies, GUS genes or genes encoding anthocyanin regulatory proteins C1 and B (or Lc and R) expressed from constitutive promoters, are commonly used. The efficiency of a particle bombardment experiment can then be evaluated by counting the number of blue GUS-expressing cells after histochemical staining, or simply by counting the number of red, anthocyanin-expressing cells.

For stable transformation experiments, a selectable marker gene is introduced with the desired gene construct and transient marker genes. The selectable gene is usually cloned in a transcriptional fusion with a constitutive promoter, such as the CaMV 35S or the maize ubiquitin promoters. The selectable marker gene cassette may be linked to the gene construct of interest, or these can be co-bombarded on separate plasmids. After a short

recovery period, generally between three days and two weeks, bombarded calli are transferred to selective media containing the selective herbicidal or antibiotic compound. After a period of selection which ranges from about six weeks to several months, resistant calli are selected. Transgenic BMS callus is maintained on solid media, or re-introduced into suspension culture. HiII calli are transferred to regenerative media and regenerated, usually also in the presence of the selective agent.

Several different selectable marker genes are currently used in maize transformation: these include genes encoding enzymes which selectively detoxify herbicides such as phosphinothricin (PPT), or mutant enzymes which are insensitive to herbicides, as well as those which detoxify aminoglycoside antibiotics like kanamycin or hygromycin. Of these, the most efficient and by far the most commonly used selectable markers encode phosphinothricin acetyltransferases (PAT) which acetylate the glutamine synthetase inhibitor, PPT (De Bock *et al.*, 1987; D'Halluin *et al.*, 1992; Wehrmann *et al.*, 1996). PPT causes ammonia accumulation and impairment of photosynthesis in plants by inhibition of glutamine synthetase. The herbicidal compound bialaphos is a natural tripeptide consisting of L-PPT and two L-alanine residues (Wehrmann *et al.*, 1996), and is produced by at least two *Streptomyces* species: *S. hygroscopicus* and *S. viridochromogenes*. Bialaphos-detoxifying PAT enzyme genes have been cloned from both species: *bar* (bialaphos resistance) from *S. hygroscopicus* and *pat* from *S. viridochromogenes*. Both genes are widely used in maize transformation, and have almost identical activities, conferring tolerance to bialaphos and synthetic PPT-containing herbicides, such as glufosinate, Basta TX™ or Ignite™ (Wehrmann *et al.*, 1996).

Both groups reporting the use of TGMV as a gene amplification system in transgenic tobacco plants (Hayes *et al.*, 1988b; 1989) and suspension-cultured cells (Kanevski *et al.*, 1992) used the kanamycin resistance gene, *nptII*, as both a selectable and screenable marker. The *nptII* gene is not very widely used in maize transformation, however, as there is a high level of inherent resistance to kanamycin in maize. In the experiments detailed in this chapter, I therefore used the *bar* gene as a dual-purpose selectable and screenable marker, for several reasons. As explained above, this gene is the most convenient for use in production of transgenic maize, and as such, protocols for the production of transgenic maize with the *bar* gene as a selectable marker were available. The *bar* gene coding sequence is small (only 549 bp) and consequently would not increase the size of a

recombinant virus too drastically. Botterman *et al.* (1991) showed that PAT tolerates C-terminal fusions, so I was able to make recombinant MSV constructs which expressed a PAT fusion protein. An assay for PAT activity in transformed cells, based on measurement of enzyme kinetics, is also available (Thompson *et al.*, 1987; D'Halluin *et al.*, 1992).

The strategy I used for evaluating whether MSV-based replicons could be used to develop a gene amplification system for transgenic maize was similar to that employed by Kanevski *et al.* (1992) in their investigation of the use of recombinant TGMV constructs carrying the *nptII* gene in transgenic tobacco cells. I replaced the virion-sense ORFs of MSV with a chimaeric *bar* gene, the expression of which was driven by the 790-bp CaMV 35S RNA promoter. Transcription termination and polyadenylation signals were supplied by the MSV SIR. I dimerised the recombinant MSV35S-*bar* construct to facilitate the escape of the monomeric circular viral construct by replicative release and introduced the dimeric construct into both HiII callus and BMS cells.

In the course of these initial investigations, I also attempted to establish whether constructs based on the MSV genome could be used to amplify the expression of genes considerably larger than the MSV virion sense genes. In addition, I was interested in determining whether one could incorporate genetic elements which have been shown to significantly enhance the expression of genes in transient and transgenic systems; this would at the same time help to answer questions on increasing genomic size on the replication of recombinant MSV replicons in transgenic maize.

In genetic manipulation of cereals, several authors have shown that an intron included in the 5'-untranslated leader sequence of gene expression constructs can significantly enhance the expression of genes in transient expression assays in cereals (e.g. Callis *et al.*, 1987; McElroy *et al.*, 1991; reviewed by McElroy and Brettell, 1994). Callis *et al.* (1987) found that the first intron from the maize alcohol dehydrogenase I (*adh* I) gene enhanced expression from the CaMV 35S RNA promoter by about 8-fold. Plant RNA virus untranslated leader sequences, like the 68-nt A/C-rich TMV "Ω" sequence function as translational enhancers and can mediate significant increases in the translation rate of mRNAs in transient assays, up to 8-fold in maize (Gallie *et al.*, 1987; 1989; Dowson Day *et al.*, 1993). I used a gene expression cassette with the shortened version of the CaMV 35S promoter, TMV Ω RNA leader and maize *adh* I first intron to drive the expression of

the *bar* gene in recombinant MSV constructs. The level of expression of the *bar* gene should theoretically be considerably higher from this promoter-intron cassette than from the CaMV 35S promoter alone. To start to test whether larger MSV constructs could be used in gene amplification strategies, I used a *bar* gene fusion with the *E.coli* glutathione reductase (*gor*) gene to increase the size of the *bar* expression cassette: this was a convenient gene to test the effect of increasing the genome size of recombinant MSV construct.

2.2 MATERIALS AND METHODS

2.2.1 Plasmid construction

All standard DNA manipulations were performed as described by Sambrook *et al.* (1989), and according to specifications of the manufacturers of the DNA modifying enzymes (Boehringer Mannheim, Amersham or Promega). Large scale plasmid purifications from *E.coli* were by equilibrium centrifugation in caesium chloride-ethidium bromide density gradients, or by anion-exchange chromatography, using a Nucleobond™ kit (Machery-Nagel). I used a clone of the genome of MSV-Kom as the backbone for construction of all V1/V2 gene replacement mutants. MSV-Kom is cloned as a *Bam*HI fragment in pUC19 [pKom500]. See Figure 1.8 for a genomic restriction map of pKom500.

Plasmids containing maize expression cassettes, pMF6 (Callis *et al.*, 1987) and pPHP7502 were obtained from Dr Vicki Chandler (Oregon State University) and Dr Brad Roth (Pioneer Hi-Bred, International, Inc., Des Moines, IA), respectively. Plasmids containing *bar*-based selectable marker genes in pDPG165 (Spencer *et al.*, 1990) and pPHP7503 were from Dekalb Plant Genetics, Inc. (Mystic, Connecticut) and Pioneer Hi-Bred, respectively. The plasmid pAHC25 (Christensen and Quail, 1996), which contains a *bar* gene and a GUS gene, each driven by the maize ubiquitin gene promoter, was supplied by Dr P. Quail (USDA Plant Gene Expression Center, Berkley, CA). pDPG208, containing a GUS gene (Gordon-Kamm *et al.*, 1990), was provided by Michael Spencer of Dekalb Plant Genetics. The plasmids containing anthocyanin regulatory genes, p35SC1 and p35SB-Peru (Goff *et al.*, 1990) were supplied by Dr Vicki Chandler. See Appendix B for genetic maps of expression plasmids not available in the literature.

For gel purification of restriction fragments, agarose gels were stained with methylene blue (Flores *et al.*, 1992). DNA was extracted from gel slices using a Geneclean™ kit (Bio101). Preparation and transformation of competent *E. coli* cells was by the method of Armitage *et al.* (1988).

pKEP116

The basic CaMV35S-*bar* V1/V2 gene replacement construct, pKEP115, was made by replacing the V1 ORF and most of the V2 ORF of MSV-Kom with the CaMV 35S promoter-*bar* expression cassette from pDPG165 (Spencer *et al.*, 1990). A tandem dimer of this plasmid was constructed, and named pKEP116.

pKEP141

The monomeric construct containing the *bar* expression cassette from pPHP7503 (CaMV35S RNA promoter fused to the Ω RNA leader from TMV, the first intron from the maize *adh*IS gene and the *bar* gene; Appendix B) in place of the V1 and V2 ORFs, was called pKEP139. The plasmid pKEP139 was then cloned as a tandem dimer to create pKEP141.

pKEP142

The *E.coli* glutathione reductase gene (*gor*) was amplified from pGR (Kunert *et al.*, 1990) by the polymerase chain reaction (PCR). The PCR primer GOR-F was designed so that the *gor* gene could be fused in frame with the *bar* gene. There is a *Bgl*II site over the stop codon of the *bar* gene in pDPG165: this is shown in the sequence ACCGAGATC*TGA*TGA (White *et al.*, 1989), where the *Bgl*II site is underlined and the TGA stop codon is italicised. I therefore included a *Bgl*II site in the primer, and removed the ATG start codon; the *gor* sequence would then be amplified without the ATG start and with the first amino acid after the *bar* sequence being a threonine, not serine as in the native *gor* sequence. The sequence of the primer GOR-F (with *Bgl*II site underlined) was:

5′-CCCAGATCTCTAAACACTATGATTAC-3′

The primer GOR-R was designed to be complementary to the 3′ region of the *gor* gene, just after the stop codon. An *Nco*I site was included in the primer sequence to facilitate insertion of the PCR product into the MSV construct at the *Nco*I site adjacent to the MSV

short intergenic region. The sequence of the GOR-R primer (with the *Nco*I site underlined) was:

5′-CCCCCATGGTTTAACGATTGTCACG-3′

The PCR reaction mix was set up with a high template DNA concentration (100 ng of pGR per reaction) to minimise the accumulation of PCR-derived mutations. The PCR cycle conditions were as follows: initial template denaturation was at 94°C for 60 seconds, followed by 30 cycles of denaturation at 94°C for 60 seconds, primer annealing at 50°C for 30 seconds and product extension at 72°C for 120 seconds. The 1.5-kb PCR fragment was digested with *Bgl*II and *Nco*I, gel purified and ligated with pPHP7502 (Appendix B) which had been digested with the same enzymes and gel purified. This plasmid was named pKEP131. The *Bgl*II-*Eco*RI fragment of pKEP131, containing the *gor* gene and 3′ region of the potato proteinase inhibitor gene II (*pin*II) was excised from the plasmid and gel purified. This fragment was ligated with the *bar* gene in pDPG165 (Spencer *et al.* 1990, Appendix B), which had been cut with *Bgl*II and *Eco*RI to facilitate directional cloning of the *gor* gene. The resulting plasmid was named pKEP135. I then cloned the *bargor* coding sequence and *pinII* transcription termination sequences behind the CaMV35S promoter, TMV Ω RNA leader sequence and maize *adhIS* first intron in the plasmid pPHP7503 by replacing the *Bam*HI-*Eco*RI fragment with the *bar* gene and *pinII* sequences of pPHP7503 with the *bargor-pinII* unit from pKEP135. This resulted in the *bargor* expression construct, pKEP138.

The CaMV35S-Ω-*adh*I intron 1-*bargor* expression cassette on a 3822 base pair *Xba*I-*Nco*I restriction fragment was linked to the basic MSV replicon from pKom500, again in place of the virion sense ORFs, to create pKEP140. I also dimerised pKEP140 to create pKEP142.

2.2.2 Production of transgenic HiII callus by microprojectile bombardment and regeneration of transgenic maize plants

For the production of transgenic bialaphos resistant HiII callus lines, I followed the methods of Fromm (1994). HiII callus was obtained from DeKalb Plant Genetics (Mystic, Connecticut). Highly embryogenic friable callus was selected and maintained according to Armstrong (1994). Transgenic maize plants were regenerated essentially as described by Fromm (1994) and Armstrong (1994) [see Appendix A for tissue culture protocols].

Embryogenic callus was selected under a dissecting microscope and a thin layer of cells was placed on a filter paper disk on HiII medium containing 0.4 M mannitol to minimise tissue damage during bombardment (Vain *et al.*, 1993; Dunder *et al.*, 1995). The tissue was placed on high osmoticum-containing medium (HiII plus 0.4 M mannitol) with 10 μg/ml of silver nitrate (Armstrong and Songstad, 1993) for four hours of pre-treatment before bombardment.

Particle preparation

I used the protocol described by Dunder *et al.* (1995) to prepare gold particles for bombardment of cells. One microgram of each of the constructs carrying the *bar* gene was used for each precipitation reaction (six shots). In precipitations with the *bar* expression cassette pDPG165, I also included 100 ng of each of the anthocyanin reporter genes p35SC1 and p35SB-Peru in each precipitation, along with 500 ng of pDPG208: thus, I could evaluate the efficiency of each bombardment from the relative number of red pigmented cells two days after the bombardment experiment, and could also evaluate bombardment efficiency by histochemical staining for GUS activity.

Microprojectile bombardment procedure

I used a Biorad PDS-1000/He particle gun device to deliver plasmid DNA to maize cells. The operation of this particle gun is described extensively in the Biorad PDS-1000/He User's Manual and has been reviewed recently by Kikkert (1993). High osmoticum medium plates with cells on filter paper disks were placed in the bombardment chamber. Generally, rupture disks rated to burst at 900 psi were used, with a gap distance of 6 mm, a macrocarrier flight distance of 10 mm and a particle flight distance of 6 cm. The bombarded tissue was left on high osmoticum medium overnight; the filter paper disks with tissue were then transferred to HiII medium.

Transformation of Hi II callus with pKEP116 and bialaphos resistance control plasmids.

It is well established that the transgene or transgenes of interest need not necessarily be linked to the selectable marker gene on the same plasmid construct (Gordon-Kamm *et al.*, 1990; Fromm *et al.*, 1990). I used the anthocyanin regulator gene constructs and β-glucuronidase expression cassettes to evaluate the efficiency of bombardment and therefore included these in precipitation of the plasmid DNAs of interest onto gold microcarriers (Dunder *et al.*, 1995). The basic autonomously replicating MSV construct, pKEP116, was used as the MSV replicon in these earlier transformation experiments; as I had derived the *bar* gene construct which I inserted in place of the V1 and V2 ORFs in pKom500 from pDPG165 (Spencer *et al.*, 1990; Gordon-Kamm *et al.*, 1990), I used pDPG165 and pAHC25, a plasmid which contains the *bar* and GUS genes under the control of the maize ubiquitin promoter (Christensen and Quail, 1996) as the "non-replicating" control plasmids.

The bombarded tissue was maintained on HiII medium without selection for one to two weeks before the tissue was transferred to selective medium (HiII medium with 3mg/l of bialaphos). Tissue was routinely subcultured on fresh HiII with 3 mg/l bialaphos every two weeks for 8 to 12 weeks (Fromm, 1994), until callus sectors clearly resistant to the selective herbicide became visible. Resistant callus sectors were transferred to fresh medium and maintained as separate transgenic lines: I presumed that isolated resistant clumps of callus represented independent transformation "events". Table 2.1 shows the genetic constructs and bombardment experiments carried out towards the production of transgenic HiII callus lines and plants.

Regeneration of transgenic maize plants

I regenerated transgenic maize plants from HiII callus using previously described protocols (Armstrong, 1994; Fromm, 1994). All regenerative media contained 3 mg/l of bialaphos. Regenerated plants were tested for herbicide tolerance by spraying with a commercial herbicide formulation (Ignite™, from Hoechst) at a concentration of 15 ml per litre of water, as recommended by the manufacturer.

Table 2.1: Particle bombardment experiments done to generate transgenic HiII callus and plants.

Plasmids Used in Bombardment	Experiment Description [a]
pDPG165 p35SC1 p35SB-Peru pDPG208	Non-replicating bialaphos resistance control, with non-selected marker genes (anthocyanin pigment and GUS)
pAHC25	Non-replicating bialaphos resistance control with *bar* and GUS genes each under the control of the strong maize ubiquitin promoter
pKEP116	Autonomously replicating MSV construct containing the *bar* gene in place of the virion sense ORFs

[a] four precipitations were done and a total of 12 plates bombarded for each experiment.

2.2.3 Production of transgenic Black Mexican sweetcorn cell lines

Friable, rapidly growing, non-regenerable Black Mexican sweetcorn suspension culture cells were obtained from Dr Brad Roth at Pioneer HiBred, Inc. (Des Moines, IA.) and cultured as described in Appendix A. Transgenic cell lines were produced by microprojectile bombardment and selection as described by Kirihara (1994), except that selective media contained 3 mg/l of bialaphos.

Preparation of tissue for bombardment

I routinely placed 1 ml of suspended BMS cells on the centre of a sterile 5.5 cm diameter Whatmans # 4 filter disk in a Buchner funnel, with a slight vacuum applied. The cells were spread evenly in an approximately 3 cm diameter circle in the centre of the filter paper disk. The filter paper was then transferred to BMS solid media (Appendix A) containing 0.2 M mannitol as osmoticum to reduce damage to the bombarded cells (Vain *et al.*, 1993), and 10 μg/ml of $AgNO_3$. The filter paper disks with BMS cells were kept on high osmoticum media for four hours before microprojectile bombardment, and for 16 hours post-bombardment, whereafter the disks were transferred to solid media without mannitol.

Particle preparation and microprojectile bombardment parameters

Precipitation of plasmids onto gold particles was according to the protocol of Dunder *et al.* (1995). Table 2.2 shows the plasmids used in different precipitations. For microprojectile bombardment, 650 psi rupture disks, a gap distance of 6 mm, macrocarrier flight distance of 5 mm and particle flight distance of 6 cm were the standard parameters used. Each plate of tissue was bombarded twice. One microgram of plasmid DNA was used per shot (six micrograms per precipitation). In the non-replicating bialaphos resistance control experiments (Table 2.2), each shot contained 500 ng of pDPG165, 50 ng each of p35SC1 and p35SB-Peru and 400 ng of pDPG208.

Table 2.2: Particle bombardment experiments for generation of transgenic BMS cell lines.

Plasmids Used in Precipitations	Experiment Description[a]
pDPG165 p35SC1 p35SB-Peru pDPG208	Non-replicating bialaphos resistance control: CaMV35S-*bar*
pPHP7503	Non-replicating bialaphos resistance control: CaMV35S-Ω-intron-*bar*
pKEP138	Non-replicating bialaphos resistance control: CaMV35S-Ω-intron-*bargor*
pKEP116	Dimer of MSV replicon containing CaMV35S-*bar*
pKEP141	Dimer of MSV replicon containing CaMV35S-Ω-intron-*bar*
pKEP142	Dimer of MSV replicon containing CaMV35S-Ω-intron-*bargor*

[a] two precipitations were done, and six plates were bombarded for each experiment

2.2.4 Transient Replication Assay

Transient replication assays were done essentially as described by Xie *et al.* (1995). Precipitations contained 1.2 μg of MSV replicon plasmids (200 ng per shot), and 200 ng each of anthocyanin regulatory genes p35SB-Peru and p35SC1 to evaluate the efficiency of

the bombardment experiment. Preparation of BMS tissue, precipitation and microprojectile bombardment parameters were the same as in section 2.2.3.

2.2.5 DNA analysis

Isolation of DNA from transgenic callus

DNA isolation from HiII or BMS callus was done according to the method of Chen and Dellaporta (1994), with some minor modifications. After the first precipitation step, the DNA solution was resuspended in TE buffer with 10 μg/ml of RNase A. After incubation for several hours at 4°C to ensure complete resuspension of the DNA and digestion of RNA, the DNA solution was extracted once with phenol:chloroform (1:1) and once with chloroform. The aqueous phase was precipitated with ethanol after the addition of 0.1 volumes of 3M sodium acetate, pH 5.2. The DNA pellet was washed with 70% ethanol and resuspended in TE buffer. DNA concentrations were determined by measuring the absorbance at 260 nm. For small amounts of callus tissue, the buffer volumes specified by Chen and Dellaporta were Scaled down proportionately.

DNA isolation and detection of replicating viral DNA in transfected BMS cells

Three days after bombardment the BMS cells were scraped off the filter paper disks and placed in a 1.5 ml microcentrifuge tube with a small amount of diatomaceous earth (celite) to facilitate grinding; the cells were then frozen in liquid nitrogen and ground to a fine powder with a microcentrifuge tube pestle. I then isolated DNA, enriched for low molecular weight DNA, by the method of Anat and Subramanian (1992). The DNA yield from each isolation was estimated by measuring the absorbance at 260 nm.

Southern hybridization

Standardised amounts of DNA from each sample were separated by electrophoresis in a 0.8% TBE agarose slab gel. Southern blotting of the DNA from the gel to positively charged nylon membrane (Boehringer Mannheim or Amersham) was done by standard procedures (Sambrook *et al.*, 1989). The probe used for detection of *bar* gene-homologous sequences was generated by PCR in the presence of digoxygenin-labelled dUTP, with primers BARP1 (5′-CGTCAACCACTACATCGAG-3′) and BARP2 (5′-GAAACCCACGTCATGCCAG-3′) and 1 ng of pDPG165 as template DNA. The PCR

cycling conditions were: 94° C for 60 seconds, followed by 30 cycles of 94° C for 45 seconds; 53° C for 30 seconds and 72° C for 45 seconds. The final elongation step was at 72° C for 300 seconds. This reaction generated a digoxygenin-labelled hybridisation probe of 418 bp.

Screening of transgenic callus lines for episomal vector DNA

I employed two different DNA hybridisation methods for screening transgenic callus lines for the presence of extrachromosomal DNA homologous to the *bar* coding sequence: Southern hybridisation of 2 μg of total genomic DNA run on 0.8% TBE agarose gels and blotted onto nylon membranes; and 500 ng of total genomic DNA slot-blotted onto positively-charged nylon membranes (Sambrook *et al.*, 1989). Samples which showed strong hybridisation signals on slot blots were taken to contain high copy number DNA if the hybridisation signal exceeded that of a blotted plasmid DNA sample which contained the equivalent of 100 copies per maize genome of *bar* in pDPG165.

Estimation of DNA copy number

Samples of DNA isolated from cell lines shown to contain episomal MSV-derived DNA between 6 and 12 months post-bombardment were digested with *Sau*3aI, a restriction enzyme with a four-base recognition sequence which conveniently cuts a few nucleotides upstream of the *bar* gene and over the stop codon, generating a 550-bp fragment. Given the approximate mass of one copy of the maize genome (4 pg; Bennet *et al.*, 1982), I determined what amounts of pDPG165 DNA would need to be digested with *Sau*3aI to correspond with 100, 250, 500, 750, and 1000 copies of *bar* per maize haploid genome, in a total amount of 1 μg of DNA. One microgram of each of the DNA samples containing MSV-*bar* vector DNA was digested with *Sau*3aI and the 550-bp DNA fragments separated on a 1.2% TBE agarose gel; standard amounts of digested pDPG165 DNA were run on the same gel. Southern blotting and hybridisation with a digoxygenin-labelled probe homologous to the *bar* gene were done as described above.

2.2.6 Quantification of PAT in transgenic callus

The functional activity of PAT in transgenic cell lines was quantified by following the rate of 5,5-dithio-bis-2-nitrobenzoic acid (DNTB) reduction at 412nm, as described by D'Halluin *et al.* (1992). Protein concentrations were determined according to the method

of Bradford (1976). PAT activity in each sample was determined in duplicate. Because no temperature-controlled recording spectrophotometer was available, the assays were done at room temperature (approximately 25° C), using a Shimadzu recording spectrophotometer.

2.2.7 RNA isolation and northern analysis

RNA was isolated from callus cells with the Trizol™ Reagent, according to the protocol supplied by the manufacturers (Life Technologies/BRL). RNA was suspended in sterile diethylpyrocarbonate (DEPC)-treated water, with 0.1% SDS. The concentration of RNA in each sample was determined by reading the absorbance at 260 nm (Sambrook *et al*, 1989). Ten micrograms of each RNA sample was run on a 1.5% non-denaturing TBE gel, and RNA was transferred to positively charged nylon membrane (Boehringer Mannheim) by capillary transfer in 10x SSC. Hybridisation with a *bar* DNA probe labelled by PCR (see above) was as described in the DIG Users' Guide to Filter Hybridization (Boehringer Mannheim).

2.3 RESULTS

2.3.1 Construction of *bar* gene replacement derivatives of MSV-Kom for introduction into tissue cultured maize cells

Construction of monomeric clones derived from pKom500

Various *bar* expression cassettes were used to replace the virion sense ORFs of MSV-Kom in pKom500, as outlined in Figure 2.1. Two of the plasmids carrying a *bar* expression cassette were constructed by me (most of the plasmids I constructed have pKEP as prefixes). The plasmid pKEP7503 was a derivative of pPHP7503 (Figure 2.1; Appendix B), which was obtained by simply changing the orientation of pJW7503 (Janet Willment, UCT Microbiology Department), derived by deletion of the TMV Ω and *adh*I intron sequences from pPHP7503. Plasmids pPHP7503 and pKEP7503 contain a shorter (450-bp) version of the CaMV 35S promoter than does pDPG165 (790 bp). The *bar* expression cassette, pPHP7503, contains the TMV Ω RNA leader and maize *adh*I intron sequences, which should enhance expression of the *bar* gene cloned downstream. The *bargor* expression cassette, pKEP138, contains the *bar* fusion gene cloned in exactly the same context as the *bar* gene in pPHP7503.

I engineered monomers of recombinant MSV constructs by deleting the entire MSV V1 ORF and most of the V2 ORF up to the *Nco*I site, which at position 1005 on the MSV-Kom genome, is just short of the V2 stop codon (nt 1047). I inserted the *bar* expression cassettes with promoters and upstream regulatory elements in place of the virion sense genes; in all four constructs (pKEP151; pKEP115; pKEP139 and pKEP140), transcription termination and polyadenylation signals for expression of *bar* were supplied by sequences in the MSV SIR. Thus, I engineered recombinant MSV constructs of varying sizes for testing the effect of increasing genomic size on replication as well as for testing the effects of TMV Ω and the maize *adh*I intron on the expression of *bar*. The insert in pKEP151, the recombinant MSV construct which carries the *bar* expression cassette from pKEP7503 was only 33 bp larger than the MSV-Kom genome (2723 bp); pKEP115 contained a 3070-bp insert; and pKEP139 and pKEP140 contained 3323- and 4827-bp recombinant viral inserts, respectively (Figure 2.1).

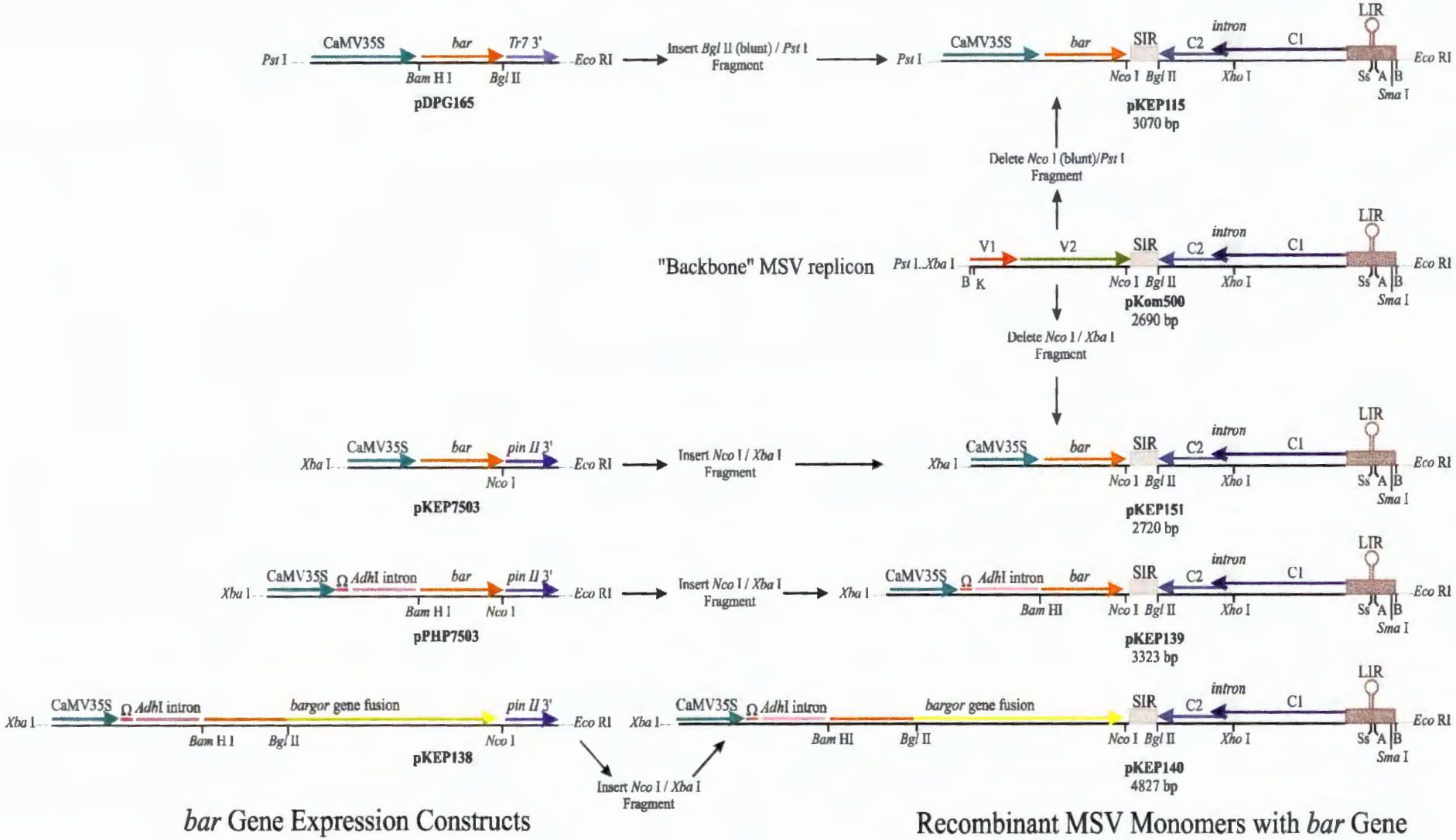


Figure 2.1: Construction of monomeric *bar* gene replacement constructs of MSV.

The figure illustrates the cloning strategy used in the construction of V1/V2 gene replacment constructs of MSV, and the structure of *bar* expression cassettes used in the generation of transgenic maize cell lines. Abbreviations: A: *Apa*I; B: *Bam*HI; K: *Kpn*I ; Ss: *Ssp*I; Ω: TMV RNA leader.

Dimerisation of inserts in recombinant viral constructs

Dimers or partial multimers of geminivirus genomes are required to facilitate escape of cloned genomes from the cloning vector sequences. I used the strategy outlined in Figure 2.2 to dimerise the *bar*-containing monomers described above; the Figure describes the dimerisation of pKEP151 to produce pKEP152. The principle was the same for construction of the other dimeric constructs. As there are no data on what specific sequences in the MSV LIR are required for replication, in dimerisation of the constructs described in Figure 2.1, I used three different restriction sites within, or adjacent to the LIR. The *Bam*HI site I used for dimerisation of pKEP115 overlaps nt 1 of the V1; it is therefore at the very edge of the LIR. Thus, when the monomeric recombinant virus escapes from the dimer, the LIR will be entirely intact. I used a *Sma*I site within the LIR for dimerisation of pKEP151, pKEP139 and pKEP140. In MSV-Kom, the *Sma*I site is at nt 2653. This is 6 bp upstream of the TATA box which forms part of the MSV virion sense promoter (Morris-Krsinich, 1985; see Figure 1.6 for the sequence of the MSV-Kom LIR). Thus, in constructs dimerised using the MSV *Sma*I site, recombinant virus constructs which result from release from the cloning vector will contain a small (39-bp) deletion in the LIR. This results in removal of the virion sense promoter TATA box, and presumably functional inactivation of the virion sense promoter, which incidentally should still be functional in constructs dimerised using *Bam*HI.

In Subgroup III geminiviruses, no sequences on the right hand side of the stem-loop structure in the intergenic region are required for viral replication (see Chapter 1 for a thorough discussion of geminivirus replication). By analogy, therefore, deletion of the sequences between the *Sma*I site and the V1 ATG should not interfere with MSV replication. I also constructed a dimer of pKEP151 using the *Apa*I site which occurs at the base of the right hand side of the stem-loop in MSV-Kom (nt 2567), but which does not form part of this structure. Table 2.3 outlines cloning of dimers of the various *bar* gene replacements of MSV-Kom, and the exact size of the recombinant monomeric virus which would result from escape from the dimeric construct by replicative release or homologous recombination

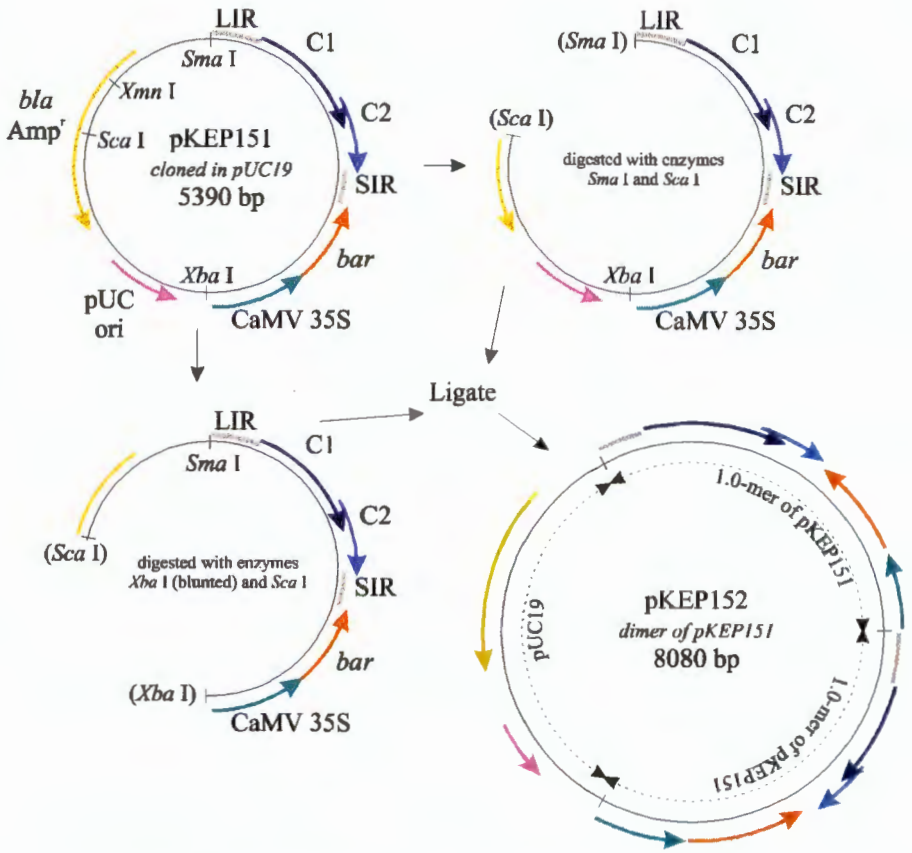


Figure 2.2: Strategy used in dimerisation of recombinant MSV-*bar* gene replacement constructs.

For illustrating the concept, I have shown how pKEP151 was dimerised to produce pKEP152. The same principles apply for cloning of all other dimeric constructs.

Table 2.3: Construction of plasmids containing dimers of *bar* gene replacement constructs.

Name of plasmid monomer	LIR-associated restriction site used in dimerisation	Name of plasmid dimer	Expected size of released circular viral genome
pKEP115	*Bam*HI	pKEP116	3074 bp
pKEP151	*Sma*I	pKEP152	2690 bp
pKEP151	*Apa*I	pKEP167	2567 bp
pKEP139	*Sma*I	pKEP141	3284 bp
pKEP140	*Sma*I	pKEP142	4788 bp

2.3.2 Transient replication assay for size constraints on replication, and sequence requirements in the LIR for replication

Before making transgenic cell lines, I needed to determine whether the recombinant MSV constructs with small deletions in the LIR would still be replication competent, and whether the largest and smallest *bar* gene replacement constructs could replicate. To this end I used a transient replication assay similar to that used by Collin *et al.* (1996) and Xie *et al.* (1995; 1996). BMS suspension-cultured cells were bombarded with the construct which would release a wild-type sized viral genome (pKEP152), and with pKEP142, which would release the largest recombinant genome, 4788 bp, nearly double the wild-type size. Both pKEP152 and pKEP142 contain a *Sma*I to *Bam*HI deletion in the LIR. To determine whether the sequences between the V1 ORF and the *Apa*I site closer to the stem-loop sequence were required for replication, I bombarded cells with pKEP167 (Table 2.3).

Figure 2.3 clearly shows that both pKEP152 and pKEP142 release viable recombinant viral constructs which replicate to high copy number in BMS cells. Thus the sequences between the *Sma*I site in the LIR and the V1 ATG, including the V1-proximal TATA box, are not required for replication. In addition, the 4788-bp recombinant MSV genome containing a chimaeric *bargor* fusion gene could replicate well, indicating that viral constructs almost double the genome size are still replication competent and that the distance between the LIR and SIR is not critical for replication of MSV. However, the construct pKEP167, which would release a recombinant virus with virtually the entire right hand side of the LIR deleted, did not release a viable replicon. This implies that there may be sequences between the *Apa*I site and the *Sma*I site in the MSV-Kom LIR which are required for normal functioning of the origin of replication. High molecular weight DNA could be seen (indicated in Figure 2.3), indicating that the input dimeric plasmid construct was probably replicating autonomously, but that the deletion in the first LIR generated in the dimerisation process did not allow release of a viable monomeric replicon.

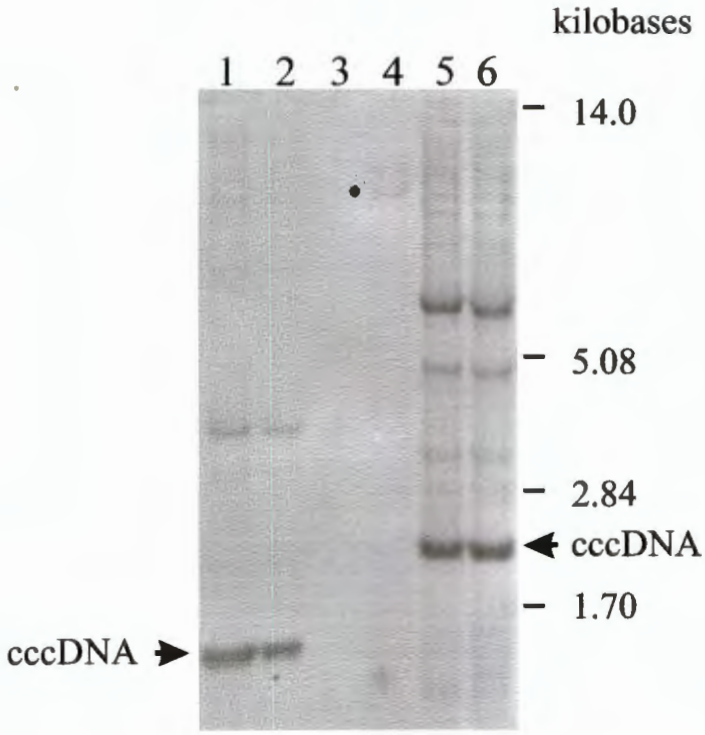


Figure 2.3: Transient assay for replication of recombinant virus constructs contained in plasmids pKEP152, pKEP167 and pKEP142.

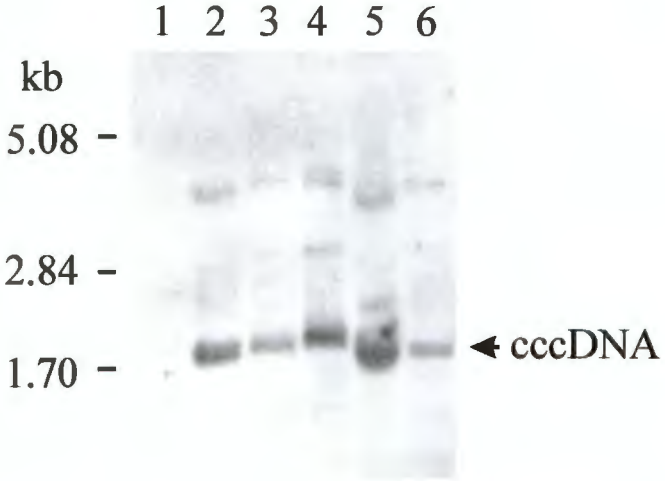
BMS cells were bombarded with microprojectiles carrying dimeric clones of the recombinant viruses. DNA was isolated three days post-bombardment. The Southern blot was probed with a digoxygenin-labelled probe homologous to the *bar* gene. Each bombardment was done in duplicate. Lanes 1 & 2: DNA isolated from cells bombarded with pKEP152; lanes 3 & 4: pKEP167; lanes 5 & 6: pKEP142. High molecular weight DNA forms probably corresponding to the entire replicating plasmid (pKEP167) are indicated (•).

2.3.3 Production of transgenic bialaphos-resistant HiII callus and attempts at regeneration of transgenic plants

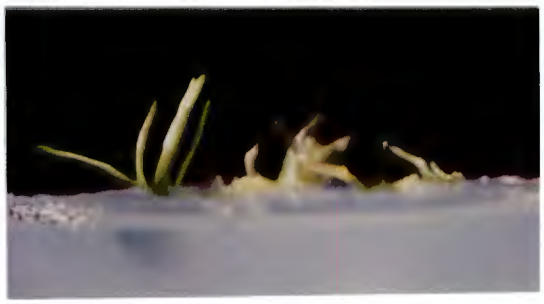
From bombardment experiments with pKEP116 (Table 2.1), I recovered a total of six independently transformed callus lines after 12 weeks on selective media containing 3 mg/l of bialaphos, a level which does not require especially high levels of *bar* gene expression to recover transgenic callus. The bombardment experiments with *bar* expression cassettes pDPG165 and pAHC25 yielded four and three transgenic lines, respectively. One of the pDPG165 transgenic lines expressed red anthocyanin pigment, presumably as a result of expression of the anthocyanin regulator genes which were included in the bombardment experiment (not shown).

I screened the six transgenic lines obtained by bombardment with the autonomously replicating construct (pKEP116) by extraction of total DNA from the callus and Southern hybridisation with a *bar*-specific probe. Figure 2.4 shows that all six lines contained extrachromosomal DNA homologous to the *bar* gene, with all the DNA forms commonly seen in virus infections, except for ssDNA. There was some variation in the apparent copy number of the episomal DNA present in each line, as there was an equal amount of total genomic DNA (5 μg) loaded in each lane.

Attempts at regeneration of transgenic callus lines containing episomal DNA were unsuccessful. However, transgenic plants from five bialaphos-resistant transgenic callus lines were regenerated: from all three pAHC25 lines, and two from the four pDPG165 lines. Transgenic callus containing the MSV-based constructs turned green on exposure to light on regenerative media, and some small abnormal-looking shoots developed (Figure 2.4), but the development of the callus halted at this stage and it eventually turned brown and stopped growing. A second attempt at regeneration of these six transgenic lines was also unsuccessful.



A.



B.

C.

Figure 2.4: Attempts at production of transgenic maize plants with high copy number viral replicons.

A. Southern blot of undigested total DNA isolated from six transgenic HiII lines bombarded with pKEP116.

Viral DNA forms are clearly visible in low molecular weight DNA, except in lane 1, where the episomal DNA is only faintly visible. Covalently closed circular DNA is indicated.

B. An example of transgenic HiII (pKEP116) plantlets starting to regenerate on selective media.

Trangenic callus regeneratedafter bombardment with pKEP116 never regenerated any further than these plants did; after several weeks at this stage, plants turned brown and died.

C. Herbicide tolerant transgenic HiII maize plant.

The transgenic plant (large pot), containing pAHC25 DNA, and non-transgenic control plants (small pot) were sprayed with a commercial PPT-based herbicide (Ignite, from Hoecsht-Agrevo). The transgenic plant is clearly far more tolerant to the herbicide than the control plants.

2.3.4 Production of bialaphos-resistant transgenic BMS cell lines

Table 2.4 shows the number of transgenic callus lines obtained from bombardment experiments detailed in Table 2.2.

Table 2.4: Production of bialaphos resistant transgenic BMS cell lines with non-replicating control plasmids and autonomous MSV-based replicons

Plasmids used in precipitations	Experiment description	Number of transformed callus lines recovered
pDPG165 p35SC1 p35SB-Peru pDPG208	Bialaphos resistance control plasmid pDPG165 with anthocyanin regulatory genes and GUS	53
pKEP116	Autonomously replicating MSV construct carrying CaMV35S promoter and *bar* (from pDPG165)	155
pPHP7503	Bialaphos resistance control plasmid	35
pKEP141	Autonomously replicating MSV construct carrying CaMV 35S promoter, TMV Ω and *bar* from pPHP7503	27
pKEP138	Bialaphos resistance control plasmid carrying *bar-gor* gene fusion	15
pKEP142	Autonomously replicating MSV construct carrying *bargor* expression cassette from pKEP138.	9

The variation in the numbers of transformed BMS cell lines recovered after microprojectile bombardment experiments reflects considerable variability in the efficiency of different bombardment experiments. I performed the first two bombardment experiments shown in Table 2.4 on the same day. The next four experiments were carried out on two successive days several months later. The first experiment was clearly far more successful, but the factors responsible for this are unfortunately not easily identified. Of the 53 transgenic lines which resulted from bombardment of BMS with pDPG165, GUS and anthocyanin regulatory genes, 21 (39,6%) displayed varying amounts of anthocyanin pigmentation, indicating successful co-transformation and co-expression of these unselected genes.

Transgenic cell lines were named according to the number of the selection plate they were isolated from.

2.3.5 Screening of transgenic BMS cell lines for the presence of extrachromosomal vector DNA

The first series of transgenic BMS cell lines with autonomously replicating MSV vectors generated contained pKEP116. Four months after the bombardment experiment, I randomly selected 10 transgenic callus lines which appeared to be healthy and were growing vigorously on media containing 3 mg/l of bialaphos. I isolated total genomic DNA from these lines, and separated any possible extrachromosomal DNA forms from the high molecular weight DNA by electrophoresis in a 0.8% TBE agarose gel. No extrachromosomal DNA forms were visible in the agarose gel when stained with ethidium bromide and visualised on a UV transilluminator (results not shown). DNA was blotted onto a nylon membrane, and hybridised with a digoxygenin-labelled *bar* probe, which should detect homologous sequences in the DNA. Of the 10 transgenic lines, six contained very high copy number extrachromosomal DNA forms, as illustrated in Figure 2.5. These six transgenic lines were selected for further analysis of vector stability and PAT expression levels.

I used a different strategy to screen for the presence of high copy number *bar*-homologous DNA in transgenic cell lines containing pKEP141 and pKEP142. Total genomic DNA was isolated from several randomly selected transgenic cell lines. Five hundred nanograms of DNA from each of these lines were transferred to a nylon membrane by slot blotting. Defined amounts of pDPG165 DNA which corresponded to between 100 and 1000 copies per haploid maize genome were included on the membrane as standards. The membrane was hybridised with a digoxygenin-labelled *bar* probe, and lines which contained high copy numbers of *bar* sequences were identified from the intensity of the hybridisation signal, as shown in Figure 2.5. Of the 25 lines containing pKEP141, 11 contained copy numbers of approximately equal to or greater than 100 copies per haploid maize genome (cf. slots A2 to A8) , and of the eight lines containing pKEP142, three contained high copy number *bar*-homologous DNA (greater than 100 copies/maize genome).

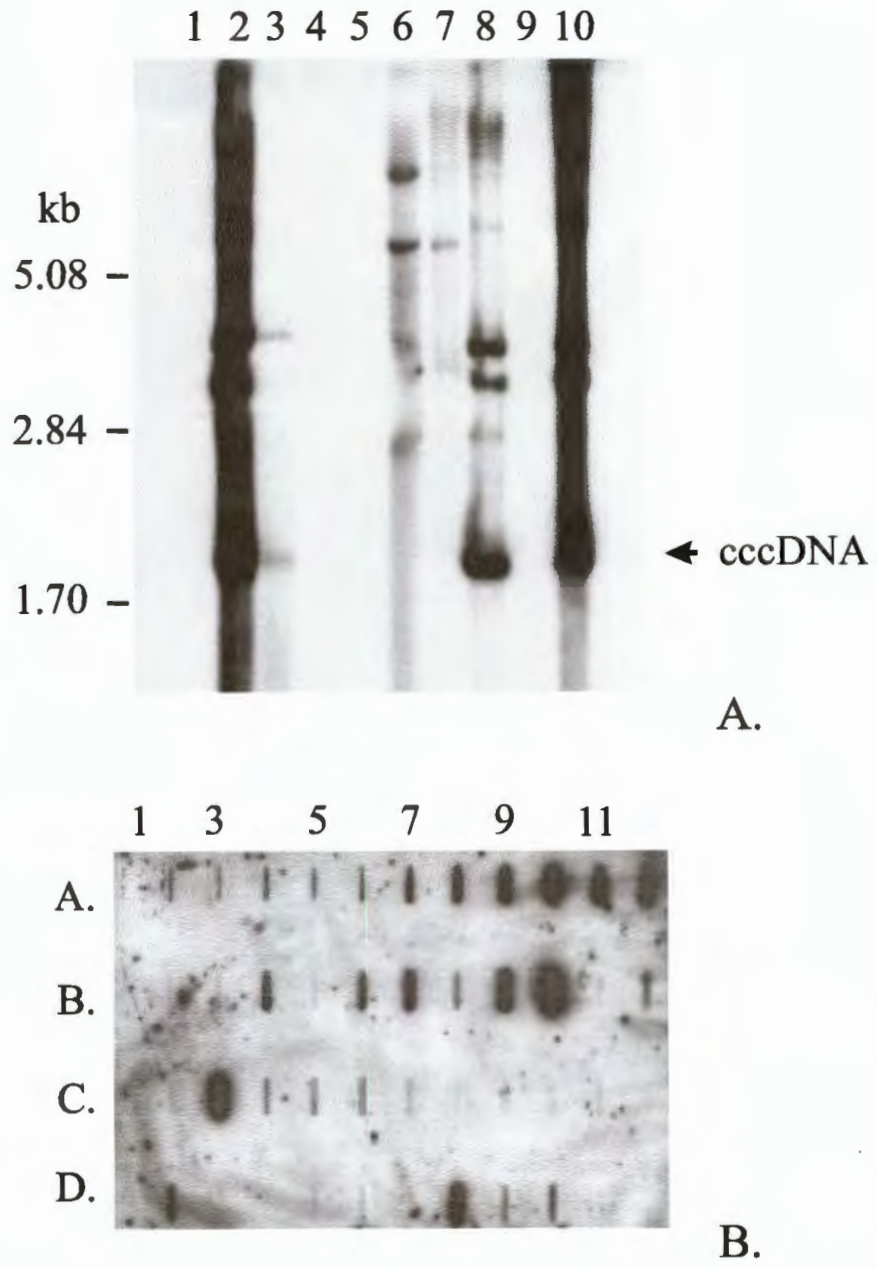


Figure 2.5: Screening of transgenic BMS lines for the presence of high copy number episome DNA.

A: Southern blot of DNA isolated from 10 transgenic BMS lines (lanes 1-10) generated by bombardment with pKEP116. Five micrograms of total genomic DNA was loaded in each lane. The Southern blot was probed with a digoxygenin labelled *bar* probe. Six out of the 10 lines contain replicating vector DNA.
B: Slot blots containing known amounts of DNA from transgenic BMS lines generated by bombardment with pKEP141 and pKEP142. Known concentrations of pDPG165 DNA corresponding to approximately 100, 250, 500, 750 and 1000 copies per haploid maize genome were loaded in slots in row A (A1: no DNA; A2 to A6: 100 copies; A7: 250; A8: 500; A9: 750; A10, A11 and A12: 1000). 500 ng of total genomic DNA from transgenic pKEP141 lines were loaded in slots B1-B12; C1-C12 and D1. DNA from transgenic pKEP142 lines was loaded in slots D5 to D12.

2.3.6 Copy number determination

The approximate copy number of *bar*-carrying extrachromosomal replicons was determined in twelve selected transgenic cell lines which prescreening experiments (2.3.5) showed contained the highest copy number of the *bar* episomes. Six lines were generated from bombardment with pKEP116 and three each with pKEP141 and pKEP142. I determined the *bar* copy number by digesting one microgram of total genomic DNA with *Sau*3aI, which excises the *bar* gene from all three constructs. Standards of *Sau*3aI-digested pDPG165 DNA corresponding to 100, 240, 500, 750 and 1000 copies per genome were run on the same gel, which was blotted and hybridised with a digoxygenin-labelled probe. The computer program GelTrak (Dennis Maeder, Biochemistry Department, University of Cape Town) was used to quantify the DNA concentrations from the resulting chemiluminogram (not shown), relative to the standards. Because I had observed some variation in copy number between DNA isolations from the same samples at different times, I did not attempt to assign exact *bar* copy numbers to each cell line, but rather preferred to express copy number as an approximate range. In all twelve cell lines, the *bar* copy number exceeded 500 copies per maize haploid genome (Table 2.5). The copy number determinations were done 10 months after bombardment with pKEP116, and 6 months after bombardment with pKEP141 and pKEP142.

Table 2.5: Copy number of episomal MSV-derived replicons in selected transgenic BMS cell lines

Plasmid Construct	Cell Line Number	Copy Number per Maize Genome	Time After Bombardment
pKEP116	AB4	~1000	10 months
	AB11	750-1000	10 months
	AB12	~1000	10 months
	AB16	500-750	10 months
	AC1	>1000	10 months
	AC2	>1000	10 months
pKEP141	BB2	>1000	6 months
	BB3	500-750	6 months
	BB4	500-750	6 months
pKEP142	CE1	500-750	6 months
	CE3	~1000	6 months
	CE4	750-1000	6 months

Having determined that all twelve of the selected transgenic BMS cell lines contained high copy number *bar*-homologous DNA, it was important to determine whether the full-size recombinant viral DNA was replicating in these cell lines, or whether there was accumulation of subgenomic DNA forms which contained the *bar* gene. One microgram of the undigested total genomic DNA from two randomly selected lines of each class was run on a 0.8% TBE agarose gel and transferred to a nylon membrane by Southern transfer. Hybridisation with a digoxygenin-labelled probe showed that episomal DNA was indeed still present in the low molecular weight DNA, and that most genomic forms associated with geminiviruses were present, except for single stranded DNA (Figure 2.6).

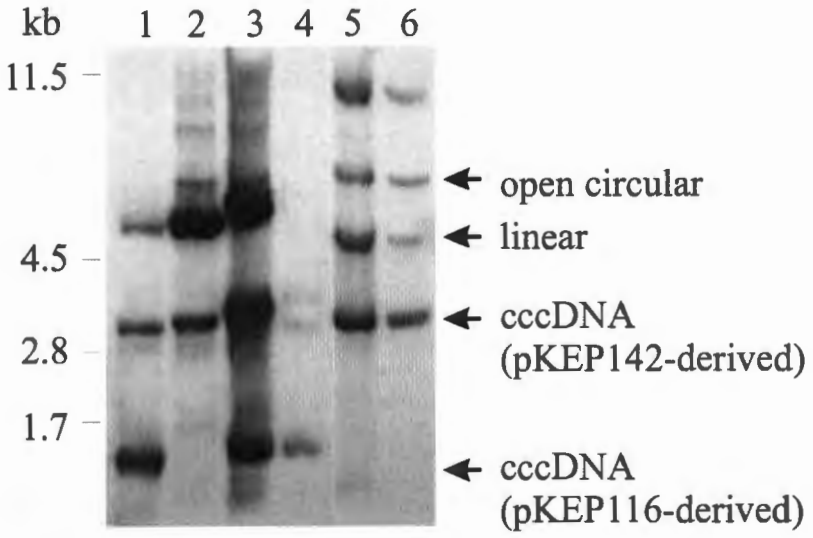


Figure 2.6: Vector DNA forms present in low molecular weight DNA from selected transformed cell lines.

One microgram of total DNA from selected transformed cell lines was isolated and electrophoresed undigested in a 0.8% TBE agarose gel. The Southern blot was probed with a digoxygenin probe homologous to the *bar* gene. No significant amount of vector-specific ssDNA was present in the transformed cell lines. Lane 1: transgenic cell line AB4; lane 2: cell line AC1; lane 3: cell line BB3; lane 4: cell line BB4; lane 5: cell line CE4; lane 6: cell line CE2.

One line, pKEP116 AC1, showed no covalently closed circular DNA band and increased levels of the band which probably represents open circular DNA. It is possible that a concatamer of pKEP116 DNA was replicating in this line, or that the sample had been more roughly treated, and only very low levels of intact cccDNA were present in this DNA isolation. As will be seen in Figure 2.7, the DNA was probably nicked.

2.3.7 Stability of MSV-derived episomes in transgenic BMS cell lines

To assess whether subgenomic DNA forms were accumulating in the extrachromosomal DNA, I digested one microgram samples of a representative example of transgenic cell lines generated from bombardment with each different plasmid with *Xho*I, an enzyme which cuts each construct once, within the *rep* intron sequence. Southern hybridisation revealed that in each case the predominant DNA forms corresponded with the intact, full length size of the recombinant viral construct. Some minor bands which hybridised with the MSV probe (indicated in Figure 2.6) could have represented very low levels of subgenomic DNA forms, or residual undigested DNA. The restriction enzyme used to digest these DNA samples, *Xho*I, is sensitive to methylation on cytosine in the enzyme's recognition sequence (C/T**C***GAC). The results implied that the extrachromosomal DNA forms were predominantly unmethylated. Methylation of geminiviral DNA has been implicated in interference in viral replication in Subgroup III geminiviruses (Brough *et al.*1992; Ermak *et al.*, 1993). The transgenic cell line AC1 which showed no ccDNA forms showed mainly full-length DNA forms, indicating that the viral DNA forms were intact in this line and the DNA was nicked in the original sample.

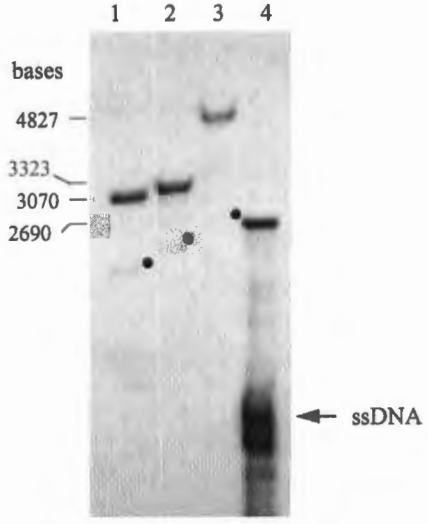


Figure 2.7: Structural stability of replicating viral constructs in transgenic BMS.

One microgram of total genomic DNA isolated from transgenic cell lines was digested with *Xho*I, which cuts all constructs once, within the MSV-Kom *rep* gene intron sequence. One hundred nanograms of MSV-Kom DNA isolated from an agroinfected sweetcorn plant was used to illustrate size differences between the wild type virus and recombinant constructs. Digested DNA was electrophoresed in a 0.8% TBE agarose gel. The Southern blot was probed with a digoxygenin-labelled probe homologous to the MSV *rep* gene. The Southern blot clearly shows the presence of ssDNA in DNA from the infected plant. No ssDNA was present in the transgenic cell lines. Lane 1: transgenic cell line AC1; lane 2: BB3; lane 3: CE4. Lane 4: maize plant infected with MSV-Kom. Spurious bands which may represent residual undigested DNA or subgenomic DNAs are indicated (•).

2.3.8 Assay for PAT activity in transgenic BMS cell lines

The levels of active PAT enzyme in transgenic cell lines were determined by measuring enzyme kinetics. For each of the three different autonomously replicating MSV-based vectors, three individuals from the corresponding transgenic bialaphos resistant "control" lines were tested to determine what effect the amplification of the *bar* gene linked to the viral replicon had on PAT levels in the transgenic cell lines. Figure 2.8 shows the PAT expression levels measured in transgenic BMS lines. For each of the three control transgenic lines (pDPG165, pPHP7503 and pKEP138), I have indicated the **highest** expressor of the three lines tested to give an idea of how much the viral replicon contributed to the expression of the *bar* gene, over transgenic cell lines which may represent the best expressors when generated by conventional means. The PAT expression levels I observed in all nine non-replicating control lines in Table 2.6.

Table 2.6: PAT Expression in non-replicating control transgenic BMS cell lines

Transgenic Cell Line	PAT Activity (10^{-6} units/μg protein)
pDPG165, No. 1	0.0
pDPG165, No. 2	0.3
pDPG165, No. 3	0.0
pPHP7503, No. 1	30.7
pPHP7503, No. 2	0.0
pPHP7503, No. 3	6.5
pKEP138, No. 1	2.2
pKEP138, No. 2	0.9
pKEP138, No. 3	6.1

As can be seen in Figure 2.8 and Table 2.6, PAT activity was barely detectable in the three transgenic BMS lines which had been generated by bombardment with pDPG165. It is obvious that only extremely low levels of *bar* expression are necessary to confer bialaphos resistance. All of the transgenic cell lines which contained replicating constructs expressed higher levels of PAT than their cognate non-replicating control line. In the case

of the pKEP116 transgenic BMS lines (AB4 - 16, AC1 & AC2), this level was between 74 and 355 times the level seen in the best expressor of the three pDPG165 control lines, however, the three control lines tested expressed exceptionally low levels of PAT. The replicating versions of the pPHP7503 and pKEP138 lines expressed between 2.4 and 5.1 times the amount of PAT.

Transgenic lines containing the *bar* gene with the *adh*I intron and TMV Ω sequence in the 5′ untranslated leader of the *bar* transcription unit appeared to be expressing higher levels of PAT, but with the low numbers of transgenic lines tested in this experiment, no firm conclusions can be drawn about the effect of these sequences on *bar* expression in transgenic BMS cells. The data presented in Figure 2.8 show that, although there may have been enhancement of PAT expression in transgenic cell lines expressing PAT from chromosomally integrated copies of *bar* genes with TMVΩ and maize *adh*I intron sequences, this phenomenon was not seen in lines in which the *bar* expression unit was present on an episome (compare hatched bars with each other and solid-filled bars of BB2, BB3 and BB4 with solid blue bars of AB lines).

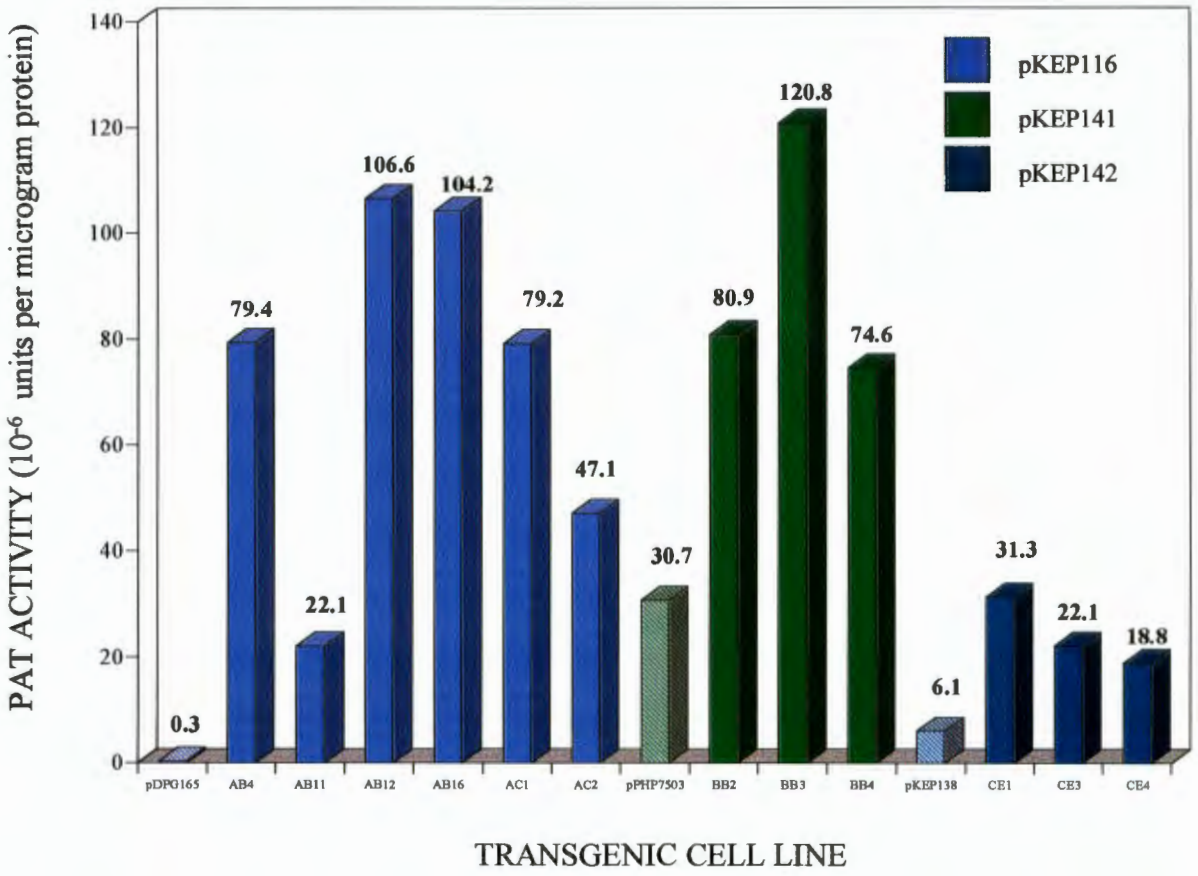


Figure 2.8: PAT Activity in transgenic BMS cell lines with autonomously replicating MSV-based replicons (solid bars) and non-replicating control plasmids (hatched bars)

As can be seen in Figure 2.8, the PAT expression levels were significantly higher in transgenic cell lines where the recombinant virus cloned in pKEP142 was replicating (lines CE1, CE3 and CE4) than in the best non-replicating expressor generated from transformation with pKEP138. However, the PAT activity in transgenic cell lines CE1, CE3 and CE4 was in all cases equal to, or less than that seen in the best pPHP7503 expressor. This would seem to imply that, although the PAT-GR gene fusion was clearly functional (since the transgenic lines are resistant to bialaphos), the fusion protein may not be as active as the native protein.

2.3.9 Northern blot analysis of *bar* RNA expression in BMS cell lines

In the light of the observations that the transgenic cell lines which contained high copy numbers of *bar* genes expressed only 2 to 5-fold greater levels of PAT enzyme, I decided to examine the expression of the *bar* mRNA. Unfortunately at this stage only one of the lines generated by bombardment with pKEP141 and six pKEP116-derived lines were available. I was able to determine the *bar* RNA expression levels in the three transgenic cell lines which still contained episomal DNA two years after their generation (AB4; AB11; and AC2), and one transgenic pKEP141 line (BB4), as well as in the two transgenic cell lines which no longer contained replicating *bar* constructs (AB16 and AC1). These results are presented in Figure 2.9. RNA from two pDPG165-transformed control lines was loaded in lanes 1 and 2; RNA from lines which contained replicating pKEP116-derived episomal DNA was loaded in lanes 3, 5 and 7, and a line with pKEP141-derived episomes in lane 4. RNA from lines where the MSV vector was no longer replicating is loaded in lanes 6 and 8.

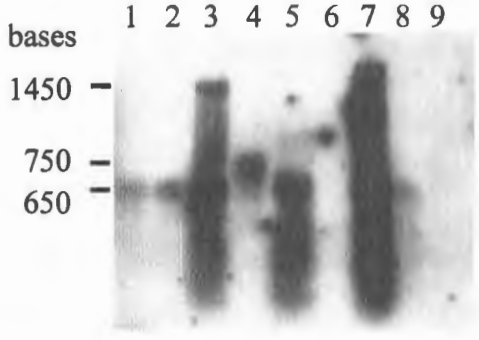


Figure 2.9: Northern blot analysis of *bar* transcripts in bialaphos resistant transformed cell lines.

Total RNA was isolated from cell lines transformed with pKEP116, pKEP141 or pDPG165. Lanes 1 & 2: Non-replicating *bar* control lines (pDPG165); lane 3: line AB4; lane 4: line BB4; lane 5: line AB11; lane 6: line AB16; lane 7: line AC2; lane 8: line AC1. Lane 9: nontransgenic BMS. Approximate sizes of transcripts are indicated.

Transgenic lines AB4, AB11, AC2 and BB4, which contained high copy number episome DNA, also showed the highest levels of *bar*-specific transcripts. In contrast, in lines which no longer contained replicating vector DNA there were either only low levels of *bar* transcripts or no detectable *bar* RNA (AB16 and AC1; Figure 2.9, lanes 6 and 8). Two bialaphos resistant transgenic BMS lines which contained pDPG165 also expressed detectable amounts of *bar* gene transcripts. The *bar* transcript expressed in the transgenic cell line which contained pKEP141 (line BB4; Figure 2.9, lane 4) was significantly larger than the approximately 650 bp transcript seen in control lines, and lines which contained pKEP116. The *bar* transcription unit present in pKEP141 contains the maize *adh*I intron (550-bp) and TMV Ω RNA leader (68 bp). I had therefore expected this RNA to be at least 68 bp longer than that produced in the other transgenic lines, if the intron had been removed. The size shift observed in this *bar* transcript was fairly small (estimated at 100 bp), which suggests that the intron sequence was efficiently processed. Two of the lines which contained pKEP116-derived replicons (AB4 and AC2; Figure 2.9, lanes 3 and 7) also expressed significantly larger transcripts (approximately 1450 bp) which hybridised with the *bar* probe. The origin of these transcripts is uncertain, but it is possible that they arose from initiation of transcription from the MSV virion-sense promoter. The promoter sequences in the LIR were left intact in the construction of pKEP116. The three transgenic lines which contained replicating DNA from the pKEP116 constructs also showed a smear of RNA without discrete bands, which may indicate inefficient transcript termination or transcript instability. This phenomenon was not investigated further.

2.4 DISCUSSION

The results presented in this chapter show that there is good potential for the development of MSV-based extrachromosomal gene amplification systems for transgenic cereals. I was able to use autonomously replicating gene replacement constructs of the MSV genome to select transgenic cell lines. A significant proportion of these contained high copy number viral replicons. This is the first report of the use of a Subgroup I geminivirus as a transgene amplification system, and is the second report of the use of a plant virus to select clonal cell lines, the first being that of *nptII*-bearing TGMV used for selection of transgenic tobacco cells (Kanevski *et al.*, 1992).

For these investigations, I used recombinant viral constructs which contained various versions of the *bar* gene as a dual-purpose selectable and screenable marker gene in transgenic maize. All constructs contained the *bar* ORF under the transcriptional control of the nominally constitutive CaMV35S RNA promoter. I had decided not to use the MSV virion sense promoter in these exploratory investigations for two main reasons. Fenoll *et al.* (1988) had found that this promoter is considerably weaker than the CaMV35S promoter in maize cells, and although genome amplification should theoretically increase the expression levels of any gene cloned downstream as a transcriptional or translational fusion, I doubted that the low levels of expression from the MSV promoter would be sufficient to confer bialaphos resistance in non-replicating control transgenic lines. Secondly, the possibility that the Rep protein transactivates the expression of the coat protein or virion sense promoter in MSV, as has been suggested for other Subgroup I geminiviruses (Hofer *et al.*, 1992; Zhan *et al.*, 1993), would complicate determination of the effect of genome amplification on the expression of the *bar* gene. For this reason, it was preferable to use a promoter which should not be affected by Rep.

Inactivation of the MSV virion sense promoter, by deletion of 39 bp upstream of the V1 ORF, had no discernible effect on replication of viral constructs, both in a transient replication assay and in stably transformed cells. However, a more substantial deletion of most of the LIR proximal to the V1 ORF, but not including the stem-loop structure or associated "iteron" sequences (Argüello-Astorga *et al.*, 1994a) abolished replication. In Subgroup III geminiviruses, no sequences on the right-hand side of the stem-loop form part of the viral origin of replication. In contrast, these results suggest that there may be sequences near the stem-loop sequence of Subgroup I viruses which are involved in replication. These could be vital elements of the *rep* gene promoter, but may also represent Rep binding sites or sites where host-derived ancilliary replication factors bind. These possibilities are being explored by Janet Willment in this laboratory. Northern blot analysis of *bar* gene transcripts produced in transgenic cell lines containing pKEP116-derived replicating vector DNA showed that, in addition to the expected *bar* transcript, a considerably larger transcript of about 1400 bp was produced in two out of three lines. In the pKEP116 construct, the viral virion sense promoter remained intact, so the larger transcript may have been produced from this promoter.

The relative proportion of the total of transgenic cell lines containing high copy number MSV-derived replicons was much greater than that which Kanevski *et al.* (1992) found for a similar TGMV-based system: the proportion of transgenic BMS cell lines containing episomal DNA varied between 38% and 60%, and all six transgenic HiII lines I recovered contained replicating viral DNA. I did not detect ssDNA forms; this is quite normal for geminiviral coat protein mutants (Boulton *et al.*, 1989b; 1993; Kanevski *et al.*, 1992; Woolston *et al.*, 1989).

I was able to regenerate transgenic herbicide tolerant HiII plants from embryogenic callus, but could not regenerate transgenic plants from the bialaphos resistant callus lines which contained pKEP116-derived vector DNA. In the light of the highly labour intensive nature of the work required to produce transgenic maize, and given that it was fairly difficult to obtain fresh HiII callus or immature embryos, I decided not to pursue this approach any further, but preferred to concentrate on the easier BMS system to establish the feasibility of using MSV to amplify transgenes. It is possible that expression of the *rep* gene in the transgenic cell lines interfered with the regeneration process. There are data accumulating which suggest that geminivirus Rep proteins are able to modify the cell cycle in their host plants. The Subgroup III viral Rep protein appears to induce the expression of proliferating cell nuclear antigen (PCNA), a DNA replication processivity factor which is produced in cells in S-phase (Nagar *et al.*, 1995). Although Hanley-Bowdoin *et al.* (1990) were able to generate transgenic tobacco plants expressing functional TGMV Rep, Dr Hanley-Bowdoin disclosed that she found it extremely difficult to generate these plants, and that the expression of Rep is not stable in subsequent generations (personal communication). In addition, Bendahmane and Gronenborn (1997) were unable to generate transgenic plants expressing TYLCV Rep. Xie *et al.* (1995) and Collin *et al.* (1996) found that WDV Rep protein interacts with Rb, a protein which is implicated in cell cycle control in vertebrate cells. A plant Rb homologue has been identified which interacts with WDV Rep (Grafi *et al.*, 1996; Xie *et al.*, 1996), so it appears that, like DNA tumour viruses, geminiviral Rep protein may control the host cell cycle in some way. In this light it was perhaps not surprising that I experienced some difficulty in regenerating transgenic plants from callus lines which contained *rep*-expressing replicons. However, without data from a much larger number of transgenic cell lines, I cannot be certain that regeneration of transgenic maize plants with high copy number MSV replicons is impossible.

Although it appeared from the experiments done with pKEP116 and its parental bialaphos resistance "control" plasmid (pDPG165) in BMS cells that transformation with replicating vector DNA is more efficient than with non-replicating constructs, given that 155 transgenic lines were obtained from bombardment with pKEP116 compared with only 53 from pDPG165, the results with pKEP141 and pKEP142 (Table 2.4) would not seem to support this hypothesis. Without further data, the differences in numbers of transgenic cell lines recovered from each bombardment experiment could only be ascribed to chance variation between the experiments.

The replication of MSV-based gene vectors in transgenic BMS cells appeared to have no detrimental effect on callus growth. The copy number of the viral episome in these transgenic lines six to ten months after introduction was high, in all cases exceeding 500 copies per haploid genome, or 1000 copies per cell. These copy numbers are, on the whole, higher than those observed by Kanevski *et al.* (1992) in transgenic tobacco cells containing TGMV-based vectors. Like the authors of the TGMV study (Kanevski *et al.*, 1992), I noticed no significant differences in copy number of the different sized recombinant MSV vectors, but there was less variability in the range of copy numbers observed between cell lines with replicating MSV vectors, in comparison with 10-fold variation seen with TGMV. Some transgenic BMS lines still contained high copy number episomal DNA more than a year after the MSV-derived vector plasmids were introduced. Kanevski *et al.* (1992) found, however, that the *nptII* gene copy number on TGMV vectors generally declined to 10 copies or less per cell after only 6 months. The MSV vector copy number did also decline, but after a much longer period. The drop in copy number could be due to methylation of the viral construct, which has been implicated in reducing or abolishing geminivirus replication (Brough *et al.*, 1992; Ermak *et al.*, 1992).

The recombinant MSV-based replicons appeared to be particularly stable in transgenic cell lines, given that I was able to detect copy numbers of greater than 500 per haploid maize genome in transgenic cell lines generated more than 10 months previously (Table 2.5). In addition, the fact that I detected no significant levels of subgenomic DNAs accumulated in the transgenic lines up to 10 months after introduction of the vector indicates that these recombinant vectors are also structurally stable in transgenic cell lines. Unfortunately, the transgenic cell lines containing pKEP141 and pKEP142 were lost 8 months after the initial bombardment experiments, but three of the original six pKEP116 lines (AB4, AB11 and

AC2) still contained high copy number episomal DNA two years after the transgenic lines were originally produced. In January 1997, nearly three years after pKEP116 was introduced into the BMS cells, line AB4 still contained some episomal DNA forms, but the copy number appeared much reduced (data not shown).

The vector copy number observed in transgenic BMS, although generally higher than that observed in Subgroup III viral systems, is still considerably lower than the 30 000 copies per cell of WDV-derived vectors observed by Timmermans *et al.* (1992) in transfected endosperm protoplasts. It is possible that there is some selection acting to down-regulate replication of the MSV constructs in the transgenic cells. Neither I nor Kanevski *et al.* (1992) determined whether a master copy of the viral replicon was integrated into the genome of the transgenic lines. However, the facts that the vectors: (1) replicated for such long periods of time, and (2) appeared structurally stable would argue in favour of the idea that a partial multimer integrated into the genome provided a template from which circular extrachromosomal forms of the recombinant virus escaped by replicative release. In addition, transgenic cell lines remained bialaphos resistant after the vector had stopped replicating, so at least a single copy of the *bar* gene must have integrated into the cell genome. The lower copy number may be accounted for by infrequent mobilisation or release of the viral replicon from the master copy, such that only a small proportion of the cells in the transgenic line contain high copy number vectors at any one time. However, it should be noted that the copy number of 30 000 per cell reported by Timmermans *et al.* (1992) may be artificially high. Maize endosperm cells undergo a natural DNA amplification process called endoreduplication, so it is possible that these cells provide higher levels of DNA replication proteins necessary for generation of these exceptionally high vector copy numbers. However, MSV replicates to such high copy number in infected maize plants that viral DNA forms are often visible in relatively low amounts of total cellular DNA run on agarose gels and stained with ethidium bromide (Palmer *et al.*, 1997). The overall copy number of MSV-derived replicons in transgenic maize cells was certainly lower than that observed in natural MSV infections. Increased genomic size can not account completely for the reduction in copy number, as the MSV-35S-*bar* replicon derived from pKEP116 is only about 300 bp larger than MSV-Kom, while the *bargor* expression replicon is 1.7 kb larger than the pKEP116 one, but replicates to about the same level. The *bargor* containing recombinant virus, at 4.78 kb, is the largest geminiviral-

based replicon used in transgene amplification to date, and is still capable of replication to 1000 copies per haploid genome in maize cells.

There is generally no linear relationship between gene copy number and gene expression levels in geminiviral-based transgene amplification systems. However, Hayes *et al.* (1988b; 1989) and Kanevski *et al.* (1992) found that the transgenic lines which expressed the highest level of marker genes were also those that contained the highest copy number of the viral replicon. While there were very substantial increases in expression of *bar* associated with replication of pKEP116, relative to those in transgenic lines expressing PAT from pDPG165, I noted only modest increases in PAT activity associated with the presence of MSV-based replicons carrying *bar* constructs with the *adh*I intron and TMV Ω RNA leader.

In transient gene expression assays, the TMV Ω sequence and the maize *adh* I intron each enhance gene expression levels about 8-fold in maize protoplasts (Callis *et al.*, 1987; Gallie *et al.*, 1989). It is thought that intron sequences may enhance transcript stability, and ensure that RNA processing events like 5′ capping and polyadenylation occur more efficiently. The TMV Ω sequence acts as a translational enhancer. There are no published data on what effect the presence of both elements has on the expression levels of mRNAs, but it is likely that it is additive, at the very least. Thus, the presence of the intron and RNA leader sequences could probably account for the higher levels of PAT activity I observed in the non-replicating control lines which contained the *bar* gene with these elements, relative to lines which contained pDPG165. I am not aware of reports on comparison of expression levels between constructs with and without the TMV Ω sequence or maize *adh*I intron in transgenic cereals, but RNA leaders are certainly known to enhance expression of foreign genes in transgenic dicots (e.g. Mason *et al.*, 1996).

Given the wide variation in expression of the *bar* gene in the small number of transgenic lines carrying non-replicating *bar* expression constructs, I would be inclined to predict that the maximum PAT activity I found in lines containing pDPG165 is artificially low. I am therefore reluctant to claim that linkage to the MSV replicon enhanced the expression of the CaMV 35S-*bar* transcription unit as significantly over the highest non-replicating expressor as the chart in Figure 2.7 suggests, i.e. between 74- and 355-fold. In contrast, the viral replicon derived from pKEP141 seemed to enhance expression of PAT between

2.4- and 3.9-fold over the level seen in the non replicating construct (pPHP7503). The PAT-GR fusion protein was apparently less active than the native protein, as evidenced from the lower level of expression seen in transgenic BMS containing both replicating and non-replicating plasmids. Nonetheless, the viral replicon afforded between 3.1 and 5 times enhanced PAT activity in transgenic cell lines. There is interest in increasing the levels of glutathione reductase (GR) for engineering resistance to oxidative stress in transgenic plants (Foyer *et al.*, 1991; Broadbent *et al.*, 1995; reviewed by Foyer *et al.*, 1994). While an investigation of the effect of high level expression of GR from a replicating MSV vector would have been interesting, it fell outside the scope of this investigation. For the purposes of this study, the *bargor* gene fusion served merely to provide a functionally active PAT protein in the context of a significantly larger gene expression construct.

The significantly higher PAT activity in transgenic lines containing non-replicating *bar* genes with genetic elements shown to enhance gene expression levels in transient expression systems suggests that, in some cases at least, RNA leaders and introns can enhance gene expression levels as much as the linkage to a viral replicon. However, if real, the expression enhancement afforded by these elements did not extend to the genes expressed from replicating viral constructs. From the limited number of samples available for analysis of *bar* mRNA produced in transgenic cells with viral replicons, it was clear that linkage to the MSV replicon afforded significant increases in transcription of the *bar* gene over lines where the vector was not replicating. It is possible that there is an upper limit of PAT activity which is tolerated or allowed in maize cells, and that this is approached with the genes expressed from the viral replicons. This may explain why I noticed only two- to five-fold enhancement of PAT activity in lines where RNA expression elements already provided significant expression enhancement in the non-replicating controls. I can only speculate that high levels of PAT are toxic to the cell in this case.

Another alternative is that the level of PAT expression obtained represents the highest possible level which is achievable in BMS cells. The codon usage in the *bar* gene is not optimised for expression in maize, and it is possible that some sort of limit imposed on the translation of the *bar* mRNA restricts the amount of the enzyme which can accumulate. For example, the fourth codon in the *bar* gene (CGA: arginine) is only used about 6% of the time in maize, and 3% in cereals overall (Murray *et al.*, 1989; Nakamura *et al.*, 1997). Optimisation of codon usage to conform more closely to that used in highly expressed plant

genes has been shown to greatly increase expression of other bacterial proteins such as Bt crystal toxins (e.g. Adang *et al.*, 1993) and *E. coli* heat-labile enterotoxin β subunit (Hugh S. Mason, personal communication) in transgenic plants, probably by increasing translation rates and enhancing RNA stability. Moreover, rare codons close to the translational start point can halt translation and increase mRNA instability (Gallie, 1993; Jacobson and Peltz, 1996). It is also important to note that the assay I used to measure the amount of PAT in transgenic maize cells is useful for determining the amount of **active** PAT in cells; an immunoassay would be more appropriate for determining the total amount of PAT protein produced, which may differ from the amount of active protein.

To provide definite answers to these gene expression questions requires that the experiments be repeated with a different marker gene, preferably one which has optimal codon usage for plant expression. While the *bar* gene is convenient to use for maize transformation, its utility as a reporter gene for gene expression studies is as yet unproven. To determine the effect which genome amplification has on expression of genes linked to a MSV replicon in transgenic cells, it would be advisable to use a reporter gene which has proven useful for analysis of gene expression such as GUS, firefly luciferase, or even *nptII*.

In conclusion, the main value of the research reported in this chapter has been proof of the concept that MSV can be used to amplify genes in transgenic cereal cells. Generation of transgenic maize cell lines containing high copy number MSV-derived episomes was straightforward, and apparently more efficient than for constructs based on Subgroup III geminiviruses. The vector constructs described replicated to very high copy number, for long periods of time in transgenic BMS. Vectors were apparently also structurally stable. The question of exactly how much viral-mediated amplification contributes to the expression of linked genes remains open. While it seems that regeneration of transgenic maize plants with constitutively replicating constructs may be difficult, it could be possible to express the *rep* gene in a more regulated fashion to promote tissue specific or inducible gene amplification and consequent increases in gene expression levels.

CHAPTER 3

THE EFFECT OF GENOME AMPLIFICATION ON THE ACTIVITY OF THE MSV-Kom COAT PROTEIN PROMOTER IN TRANSIENT ASSAYS AND TRANSGENIC CELL LINES

TABLE OF CONTENTS

SUMMARY

Mutations were introduced into the MSV-Kom genome by site-directed mutagenesis to generate constructs which would be useful for determining the effect of genome amplification on the expression of the MSV coat protein promoter. An *Nco*I site introduced over the start codon of the coat protein ORF allowed the *gus* gene to be cloned in a transcriptional and translational fusion with the MSV coat protein upstream expression regulatory elements. In certain *gus*-containing constructs, the *rep* gene was inactivated by introduction of a frameshift near the start of the RepA ORF; in others, the geminivirus origin of replication was inactivated by insertion of a 4-bp oligonucleotide. A microprojectile bombardment-mediated transient gene expression assay was used to evaluate the expression of the *gus* gene under the control of the coat protein promoter, in the context of replicating and non-replicating genomes. I found low, constitutive expression of GUS under the control of the coat protein promoter. Genome amplification enhanced GUS expression levels 45-fold. However, I observed no significant Rep-mediated transactivation of GUS expression from either a co-bombarded viral replicon or from a Rep expression cassette. However, increased expression of Rep from another replicon or from a Rep expression cassette further enhanced coat protein activity in autonomously replicating constructs.

I generated stably transformed bialaphos resistant cell lines which contained two autonomously replicating MSV constructs: one containing a CaMV35S promoter-*bar* gene expression cassette, and the other a *gus* gene, under the control of the MSV coat protein promoter. There was significant enhancement of GUS expression levels in lines which contained GUS episomes. There was also evidence for accumulation of deletion derivatives of the GUS, but not the *bar* replicons. I observed efficient recombination between co-bombarded constructs: mutations which inactivated the *rep* gene or the plus strand origin of replication in GUS plasmids were lost, presumably by recombination with the replicating *bar* construct.

3.1 INTRODUCTION

In order to evaluate the potential of MSV as a gene amplification vector, the question of how much genome amplification contributes to the expression of viral genes, or marker genes linked to the MSV replicon, needs to be answered. Data from investigations into the effect of genome amplification on the expression of genes linked to Subgroup III geminiviral replicons suggest that viral DNA replication to high copy number is required for high level expression of the coat protein gene. Although the transactivation effect of the AC2 (AL2) gene product on transcription of the TGMV virion sense promoters is substantial (Sunter and Bisaro, 1991; 1992), this effect is almost eclipsed by the 60- to 90-fold enhancement of coat protein expression by genome amplification (Brough *et al.*, 1992b). Recently, Suárez-López and Gutiérrez (1997) presented convincing evidence that WDV replication is required for maximal expression of GUS from the coat protein promoter, although these authors did not quantify the effect that genome amplification has on the expression of genes linked to this promoter. Collin *et al.* (1996) found that RepA is necessary for coat protein activity in WDV: constructs which contained a *rep* cDNA in place of the native sequence, and consequently could not express RepA, replicated more efficiently than the wild type replicon, but expression of GUS from the virion sense promoter was negligible. Thus, RepA may be involved in transactivation of the Subgroup I virion sense promoter.

The only published analysis of the MSV coat protein promoter was by Fenoll *et al.* (1988; 1990). These authors showed low constitutive activity of the virion sense promoter in transfected BMS protoplasts: this was only about a third of that from the CaMV 35S RNA promoter. They did not, however, evaluate what effect genome amplification had on expression of genes linked to the MSV coat protein promoter. The aim of the investigations presented in this chapter, therefore, was to determine what effect(s) MSV replication and Rep-mediated transactivation have on the expression of the coat protein promoter, both in transient expression assays and in transgenic cell lines.

Most investigations of the activities of geminivirus promoters have employed protoplasts isolated from suspension cultured cells. However, the current trend in molecular analyses of promoter structure and function in plants is towards using intact cells transfected by microprojectile bombardment for transient expression assays. Physiologically, tissue

explants, callus, or suspension-cultured cells approximate the natural situation for gene expression more closely than protoplasts.

The technology for biolistic transient assays was pioneered by maize molecular geneticists interested in transcription factors which induce expression of anthocyanin genes in maize (Klein *et al.*, 1989b), and has since been extended to analyses of many other promoters. Two different reporter gene constructs are commonly used in the assays; one is an experimental reporter gene, cloned behind the DNA sequences of interest, and the other serves as an internal control for the efficiency of the precipitation onto microprojectiles, and of the bombardment itself (for good examples, see: Klein *et al.*, 1989b; Goff *et al.*, 1990; 1992; Radicella *et al.*, 1992; Scheffler *et al.*, 1994 and Unger *et al.*, 1993). The expression of the experimental reporter gene is standardised to the expression of the internal control reporter, and expressed as a ratio of the internal control expression level. The rationale for this approach is that the expression of the internal control plasmid should be constant in every experiment. Any variation will be due to experimental variability, and so expression of the experimental reporter gene should be adjusted accordingly.

The *E. coli* β-glucuronidase (*gusA* or *uidA*) gene is very widely used as a reporter gene in plant molecular biology. In this chapter, I report on the construction of *gus*-coat protein replacement plasmids which I used in transient assays to determine: (1) the basal expression level of the coat protein promoter in BMS cells in the absence of replication and possible Rep transactivation; (2) the effect of viral DNA amplification on the expression of GUS; and (3) the effect of *trans* replication and Rep transactivation on the expression of the MSV virion sense promoter. For the latter experiments, Rep was supplied from either another viral replicon, or from an expression cassette. The firefly luciferase gene (Ow *et al.*, 1986) was chosen as a sensitive marker gene for use as an internal standard in the transient assays.

In the light of the data presented in Chapter 2, where I found that bialaphos resistant transformed BMS cell lines contained high copy number replicating constructs, but lower than anticipated expression of PAT, I decided to generate transgenic bialaphos resistant BMS cell lines containing the MSV-GUS constructs together with a replicating *bar* construct. These experiments would answer several different questions, in the first instance whether the low levels of PAT expression would extend to other genes linked to

an MSV replicon. For practical use of the system it was also important to know whether a selectable marker was required in *cis* for generation of transgenic cell lines containing high copy number episomal DNA. To date, all reports of the generation of transgenic plants or transformed cell lines containing geminivirus-derived episomes have used only one replicon. Development of a possible "binary" amplification system where two or more genes could be amplified in the same cell is an attractive possibility, which was also explored here.

3.2 MATERIALS AND METHODS

All standard cloning and DNA manipulation techniques were as described by Sambrook *et al.* (1989), and according to the specifications of enzyme manufacturers (Boehringer Mannheim, Promega or Amersham). Large scale plasmid isolations were done by ion exchange chromatography using a Nucleobond™ kit (Machery-Nagel).

3.2.1 Mutagenesis of the MSV-Kom genome

3.2.1.1 *Oligonucleotide-directed mutagenesis*

I used the Chameleon™ double-stranded site-directed mutagenesis kit (Stratagene, La Jolla, California) to introduce an *Nco*I site into the genome of MSV-Kom at the start codon of the coat protein gene, and a *Pst*I site and shift in the open reading frame near the start of the C1 ORF of the *rep* gene. The Chameleon™ kit employs two mutagenic oligonucleotides in each mutagenesis reaction: the selection oligonucleotide is homologous to the *bla* (ampicillin resistance) gene commonly found in pUC-based plasmids, while the second oligonucleotide is designed to introduce the desired mutation in a specific sequence which is cloned in the same vector.

Design of mutagenic oligonucleotides

I designed two oligonucleotides according to the criteria outlined in the kit, to introduce each of the desired mutations in the MSV-Kom genome. Each had as little secondary structure as possible, was greater than 25 nucleotides in length, contained the mutagenic sequence alteration in the middle of the oligonucleotide sequence where the mutation was

flanked by at least 10 nucleotides, was phosphorylated on the 5′ end, and gel purified from polyacrylamide gels. The *Sca*I→*Mlu*I and *Mlu*I→*Sca*I selection primers, homologous to sequences within the *bla* (ampicillin resistance) gene of the cloning vector (pUC19), were designed by Stratagene. The *Nco*I and *Pst*I mutagenic primers were designed to anneal to the same strand of the plasmid DNA as the *Sca*I→*Mlu*I and *Mlu*I→*Sca*I selection primers, as required by the mutagenesis protocol. Both were homologous to the complementary sense of the MSV-Kom genome. Oligonucleotides were obtained from Genosys (Cambridge, UK).

The *Nco*I mutagenic primer incorporates the ATG initiation codon of the coat protein gene in the recognition sequence of the *Nco*I restriction endonuclease (CCATGG). In the sequence of the *Nco*I mutagenic primer, the sequence of the MSV-Kom genome was altered at one nucleotide only: in the mutated sequence, this mutation would result in the second amino acid in the coat protein being changed from serine to alanine, a conservative amino acid substitution. The sequence of the mutagenic oligonucleotide was as follows:

5′-CTTGGACGTGG***C***CCATGGCTGATTGC-3′

The new *Nco*I site is italicised and the position of the A→C transversion at nucleotide position 318 of the negative sense of the MSV-Kom sequence is indicated in boldface, and underlined.

The purpose of the introduction of the *Pst*I mutation was to induce a +2 shift in the ORF and premature termination of translation of the Rep protein very near its amino terminus. The sequence of this oligonucleotide was as follows:

5′-CTATCCAAAGTGTC***TG***CAGAAAATCCTG-3′

I have indicated the position of the insertion of the TG dinucleotide (underlined and in boldface) which forms part of the new *Pst*I site, which I have shown in italics within the oligonucleotide sequence. This mutation would alter the MSV genomic sequence at position 2304, and induce a +2 shift in the RepA ORF.

Site-directed mutagenesis

The monomeric clone of the MSV-Kom genome in pUC19 (pKom500) was used as the basis for the mutagenesis experiments. The methods followed for site directed mutagenesis were as described in the Chameleon™ Double-Stranded, Site-Directed Mutagenesis Kit Instruction Manual (Stratagene). For introduction of an *Nco*I site over the start codon of the coat protein ORF, I heat-denatured 0.25 pmol of pKom500 and annealed 25 pmol each of the *Sca*I→*Mlu* I selection primer and the *Nco*I mutagenic primer. The second strand was synthesised with T7 DNA polymerase in the presence of dNTPs, T4 ligase and single-stranded DNA binding protein. All of the plasmid DNA was then digested with *Sca*I to eliminate parental plasmids, and then used to transform the repair-deficient *E. coli* strain, XL*mutS* (genotype: *Δ(mcrA)198 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 gyrA96 relA1 lac mutS::*Tn*10* (Tet^r) [F′ *proAB lacI*^{q}*ZΔM15* Tn*5* (Kan^r)]), which cannot distinguish between the parental and mutagenised strand, and so half of the plasmids propagated in this strain should contain the mutation. The transformed *E. coli* were grown in liquid medium overnight, in the presence of ampicillin. Plasmid DNA was isolated using Qiagen™-tip miniprep columns, and digested with *Sca*I. This step again selects for mutant plasmids, which should not contain the *Sca*I restriction site. The digested plasmid DNA was used to transform *E. coli* strain XL1-blue (Stratagene). Plasmid DNA was isolated from ampicillin-resistant colonies by the standard small-scale alkaline lysis method (Sambrook *et al.*, 1989), and restriction with *Nco*I was used to identify colonies containing mutant plasmids. The plasmid containing the correct additional *Nco*I site was called pKEP170.

The *Nco*I mutant plasmid pKEP170 was used as the basis for the mutagenesis experiment for introduction of the *Pst*I site near the 5′ end of the *rep* gene in MSV-Kom. The *Sca*I site within the *bla* gene of this plasmid had been converted into an *Mlu* I site in the previous round of mutagenesis (see above). The selection primer used in this case was therefore the *Mlu* I→*Sca*I primer (Stratagene). The mutagenesis protocol was identical to that followed for the construction of pKEP170, except that *Mlu* I was used to enrich for mutant plasmids, and *Pst*I was used to screen for putative mutant plasmids. The resultant mutant plasmid was called pKEP171.

3.2.1.2 Site directed mutagenesis by "restriction site mobilisation"

Repeated efforts at using double-stranded site-directed mutagenesis (as described above) for introducing a mutation into the loop sequence of pKom170 and pKEP171 to generate origin of replication-deficient viral constructs were unsuccessful. I therefore decided to use a different approach for generation of loop mutants. This involved a restriction site mobilisation strategy: CAT19, a chloramphenicol resistance gene cassette flanked by *Xba*I sites (Fuqua, 1992) was used to introduce an *Xba*I site into the loop sequence of pKEP170 and pKEP171. Both pKEP170 and pKEP171 were partially digested with *Ssp*I, the recognition site of which occurs three times in these plasmids: once within the pUC sequence, once in the loop sequence at nt 2537 and once within the *rep* gene at nt 2117. Linearised plasmids which had been cut only once with *Ssp*I were purified from agarose gels (Geneclean™) which had been stained with methylene blue (Flores *et al.*, 1992). The CAT19 cassette was liberated from the parental plasmid by digestion with *Xba*I. The ends of the restriction fragment were made blunt with Klenow polymerase and dNTPs (Sambrook *et al.*, 1989). CAT19 was also gel purified, and ligated with *Ssp*I-linearised pKEP170 and pKEP171. *E. coli* transformants which contained the CAT19 insert in one of the three *Ssp*I sites were selected by plating on media containing ampicillin (100 μg/ml) and chloramphenicol (30 μg/ml). Recombinant plasmids which contained the CAT19 cassette inserted into the loop-associated *Ssp*I site were identified by limited restriction mapping. Insertion of the blunted CAT19 cassette into an *Ssp*I site should regenerate both *Xba*I sites. Plasmids containing the CAT19 cassette in the loop sequence of pKEP170 were named pKEP172, and those with CAT19 in the loop of pKEP171 were named pKEP173.

To remove the CAT19 cassette from pKEP172 and pKEP173, I digested the plasmids partially with *Xba*I, as they now each contained three *Xba*I sites (one in the pUC19 multiple cloning site, and sites flanking CAT19). Linearised 5.4 kb plasmids which remained intact, except for the removal of the CAT19 fragment, were gel purified, self ligated and used to transform *E.coli*. Mutant plasmids which contained an *Xba*I site in the loop sequence of pKEP170 and pKEP171 were named pKEP174 and pKEP175, respectively.

All mutated plasmids were sequenced to verify that mutations were correct. DNA sequencing reactions were done using a Thermosequenase™ kit (Amersham) with Cy^5 Far

Red-end labelled primers. Sequence was determined by running labelled products on a Pharmacia ALF-Express™ automated DNA sequencer, according to the manufacturers' instructions. Sequence data was processed with the Pharmacia software package AM V 3.02, which controls and evaluates the sequence data generated by the ALF-Express automated sequencer.

3.2.2 Plasmid construction

Luciferase expression cassette: pKEPLuci

A plasmid containing the firefly luciferase gene (Ow *et al.*, 1986), under the control of an oestrogen-inducible promoter, pPHP5947, was supplied by Dr Brad Roth (Pioneer Hi-Bred). I constructed a luciferase expression cassette, with the expression of the luciferase gene driven by the CaMV 35S promoter, by linking the smaller *Sca*I-*Nco*I restriction fragment of pPHP7502 to the larger *Sca*I-*Nco*I fragment of pPHP5947. The resulting plasmid was called pKEPLuci (see Appendix B for the restriction maps of pPHP5947 and pPHP7502).

pKEP*gusNco*

The *gusA* gene contained in plasmid pGUSN358→S (Clontech, Palo Alto, CA.) was used as a convenient source of the promoterless *gusA* sequence as it contains an *Nco*I site over its ATG start codon. However, the sequence had been modified to destroy a putative N-linked glycosylation site (N358→S), with a small decrease in activity of the enzyme (Farrel and Beachy, 1990). I was interested in comparing the MSV promoter strength with that of the CaMV35S promoter in pDPG208 (Gordon-Kamm *et al.*, 1990), where the sequence is wild-type at amino acid 358. I therefore substituted the 3′ region of the *gusA* gene in pGUSN358→S with the 3′ end of the *gusA* gene, including the *A. tumefaciens nos* 3′ region with transcription termination and polyadenylation signals from pDPG208. To this end, I digested pGUSN358→S with *Sna*BI and *Eco*RI and replaced this fragment with the *Sna*BI-*Eco*RI fragment from pDPG208. The resulting plasmid was called pKEP*gusNco*.

Recombinant MSV constructs

The *gus-nos* 3′ insert from pKEP*gusNco* was cloned behind the V1 ORF in pKEP170 by first digesting both pKEP*gusNco* and pKEP170 with *Nco*I and *Eco*RI. This excises the *gus-nos* cassette from pKEP*gusNco* and the entire V2-SIR-C2-C1-LIR fragment from pKEP170, leaving the V1 ORF linked to the pUC19 backbone (see the restriction map of the parental plasmid pKom500 in Figure 1.8). I then ligated the *gus-nos* fragment with the V1 ORF in pUC19 to create pKEP176. The V1-*gus-nos* fragment from pKEP176 was then linked to each of all four mutant MSV-Kom plasmids in the following way: pKEP176 was digested with *Eco*RI and the sticky ends made blunt with T4 polymerase and dNTPs. Similarly, each of pKEP170; pKEP171; pKEP174 and pKEP175 was digested with *Nco*I (near the end of the coat protein ORF), which was made blunt in the same way. The plasmids were then all restricted with *Sal*I, which cuts within the pUC19 multiple cloning site proximal to the V1 gene. The "vector" fragments corresponding to mutant MSV-Kom constructs cut with *Sal*I and blunted *Nco*I were then ligated with the *Sal*I-blunt *Eco*RI "insert" fragment to create plasmids pKEP170*gus*, pKEP171*gus*, pKEP174*gus* and pKEP175*gus*.

To facilitate cloning tandem dimers of the *gus* constructs using the *Sca*I dimerisation strategy (Chapter 2, Figure 2.2), I had to subclone the inserts in each plasmid into another plasmid because some of these had lost the *Sca*I site in the mutagenesis procedure (section 3.2.1). The recombinant viral genomes in the pKEP*gus* plasmids are all cloned as *Bam*HI fragments; I subcloned these into pBluescriptSK. The resulting plasmids all retained their previous names, with "SK" added as a suffix. I dimerised all plasmids by digesting them partially with *Bam*HI and to completion with *Sca*I. The appropriate linear fragments were gel purified and ligated to create tandem dimeric clones of the *gus* coat protein replacement constructs. These were named pKEP170*gus*d; pKEP171*gus*d; pKEP174*gus*d and pKEP175*gus*d, respectively.

Rep expression cassette

The *rep* coding sequence was amplified from pKom500 by PCR with the following primers: C1KEP (TTAGGATCCCTCAGCCTCAACCTCC), homologous to a region just upstream of the C1 ATG, and C2KEP (ACGCAAACAATACAGGGGGGTAGTC), which anneals within the SIR. A *Bam*HI site was included in the 5′ end of C1KEP to facilitate

cloning. The PCR cycle conditions were: 30 cycles of 94° C for 60 seconds, 65° C for 60 seconds and 72° C for 120 seconds, followed by a final step at 72° C for 300 seconds. The PCR product was polished with the Klenow fragment of DNA polymerase I and dNTPs, then digested with *Bam*HI. It was cloned into pBluescript KS, which had been digested with *Bam*HI and *Eco*RV. The plasmid was named pKEP104. The DNA sequence of the entire 1.3 kb fragment was determined, as described above.

The *rep* coding sequence was excised from pKEP104 as a *Bam*HI-*Bgl*II fragment and cloned into the *Bam*HI site of pMF6 (Callis *et al.*, 1987), behind the CaMV35S promoter and maize *adh*I intron, with the 3′ region of the *nos* gene providing transcription termination and polyadenylation signals. This plasmid was called pKEP106. The *Bam*HI-*Not*I fragment of pKEP106 carrying the *rep* gene and *nos* terminator was then cloned behind the strong rice *actin 1* promoter in pCOR113 (McElroy *et al.*, 1991) to make pKEP129. This plasmid was subsequently linked to the chlorsulfuron resistance cassette from pPHP7504 (Appendix B). The rice actin-*rep* expression cassette linked to the chlorsulfuron resistance gene was called pKEP132.

3.2.3 Reporter gene expression assays

Histochemical detection of GUS activity

Detection of the expression of the *gusA* gene in both bombarded and transgenic BMS cells was done using the GUS substrate 5-bromo-4-chloro-3-indolyl glucuronide (X-Gluc), according to standard protocols (Jefferson *et al.*, 1987; Hull and Devic, 1995).

Chemiluminescent detection of GUS activity

Sensitive and quantitative detection of expression of the *gus* gene by chemiluminescence was done using the GUS-Light™ kit from Tropix, Inc. (Bedford, Massachusetts), according to the manufacturer's specifications, except that the assay volumes were halved. For detection of GUS in transgenic cells, the GUS assay buffer from the GUS-Light kit was used for protein extraction. For detection of GUS in transient expression assays, the total cellular proteins were extracted in the luciferase extraction buffer so that the same protein extract could be used for both GUS and luciferase assays. Ten microlitres of cell extract were used in the assay. A Picolite™ (Pierce Instruments) luminometer was used to read the chemiluminescence. I used a three second delay after injection of the Light Emission

Accelerator; the samples were counted for 5 seconds. For quantitation of GUS activity in transgenic cell lines, the protein concentration in individual extracts was determined by the method of Bradford (1976), and the GUS activity standardised according to the protein concentration.

Bioluminescent assay for firefly luciferase activity

For detection of the firefly luciferase gene product in BMS cells, I used the Luciferase Assay Kit (Tropix, Bedford, Massachusetts) according to the manufacturer's specifications. A Picolite luminometer was used with a delay time of one second following injection of substrate B, and measurement for ten seconds.

3.2.4 Transient expression assays

Tissue preparation

For transient expression assays, I routinely placed 1 ml of suspended BMS cells on the centre of a sterile 5.5 cm diameter Whatmans # 4 filter disk in a Buchner funnel, with a slight vacuum applied. The cells were spread evenly in an approximately 3 cm diameter circle in the centre of the filter paper disk. The filter paper was then transferred to BMS solid media (Appendix A) containing 0.2 M mannitol as osmoticum to reduce damage to the bombarded cells (Vain *et al.*, 1993), and 10 μg/ml of $AgNO_3$. The filter paper disks with BMS cells were kept on high osmoticum media for four hours before microprojectile bombardment, and for 16 hours post-bombardment, whereafter the disks were transferred to solid media without mannitol.

Particle preparation and microprojectile bombardment parameters

Precipitation of plasmids onto gold particles was done according to the protocol of Dunder *et al.* (1995). For microprojectile bombardment, 650 psi rupture disks, a gap distance of 6 mm, macrocarrier flight distance of 5 mm and particle flight distance of 6 cm were the standard parameters used. To reduce the effect of variation between individual bombardment experiments or DNA precipitations, the activity of the *gus* gene from the MSV virion sense promoter or from pDPG208 was corrected to the expression of a luciferase internal standard, from a standard amount of co-precipitated pKEPLuci. Each precipitation of DNA onto gold microcarriers contained 6μg of one of the pKEP*gus* dimeric plasmids, 1200 ng of internal control (pKEPLuci). For transactivation and *trans*-

replication experiments, 1200 ng of pKEP116 (Chapter 2) or pKEP132 were included in plasmid precipitations. Each individual DNA precipitation was used for six bombardments and each plate of BMS tissue was bombarded twice. Every bombardment was repeated at least nine times, i.e. at least three precipitations were done.

Data analysis

For data analysis, the same enzyme extract was used for measuring both luciferase and GUS activities. Both enzyme assays were done on the same day, three days after bombardment; luciferase assays were always done first. The GUS activity was standardised to the luciferase internal control; thus GUS activity was expressed as a ratio of GUS expression to luciferase expression. Microsoft's Excel™, version 5.0a, was used to perform statistical analyses (standard deviations and 95% confidence intervals).

3.2.5 Production of transgenic BMS cell lines

The protocols used for production of transgenic BMS cell lines were as described in Chapter 2 (section 2.2.3). Plasmid DNA precipitations contained 3 μg of pKEP116 (Chapter 2, section 2.2.1) and 3 μg of the appropriate pKEP*gus* dimeric construct, i.e. 500ng of each plasmid per shot. Transgenic cell lines were selected on BMS media containing 3 mg/l of bialaphos.

3.2.6 DNA analysis

DNA isolations and Southern hybridisation were done as described previously (Chapter 2, section 2.2.5). A probe homologous to the *gus* gene was prepared by random-primed labelling of pGUSN358→S with digoxygenin-dUTP, as described in the DIG User's Guide to Filter Hybridisation (Boehringer Mannheim). Probes were removed from membranes by incubation twice for 20 minutes in 0.2 M NaOH, 0.1% (w/v) SDS at 37° C. Membranes were then rinsed in 2 x SSC, and re-used for hybridisation with a different probe.

PCR amplification of MSV sequences from transgenic cell lines was with degenerate primers designed to amplify the region corresponding to nucleotides 1595-209 in the MSV-Kom genome, from near the 5′ end of the C2 ORF to within the V1 ORF, including the LIR (E.P. Rybicki, unpublished). The sequences of these primers are:

MSV 1770-1792: 5′-TTGGVCCGMVGATGTASAG-3′ ; and

MSV 215-234: 5′-CCAAKDTCAGCTCCTCCG-3′,

where the ambiguity codes are M = A or C; K = G or T; S = C or G; V = A or C or G; and D = A or G or T. The PCR cycle conditions used were: initial denaturation at 95° C for 12 seconds, followed by 30 cycles of 94° C for 45 seconds, 54° C for 45 seconds and 72 °C for 90 seconds. The final cycle had an elongation time of 180 seconds.

3.3 RESULTS

3.3.1 Mutagenesis of the MSV-Kom genome

The introduction of an *Nco*I restriction site over the start codon of the coat protein gene in pKEP170 (Figure 3.1) facilitated fusion of the *gusA* sequence to the coat protein promoter. A single nucleotide change was required to introduce the new *Nco*I site. This T→G transversion in the MSV-Kom sequence induced an amino acid substitution of the second residue of the coat protein, from serine to alanine in the mutant virus clone, pKEP170 (Figure 3.1).

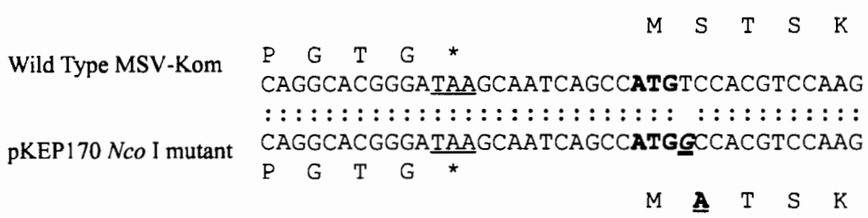


Figure 3.1: Introduction of an *Nco*I restriction site over the start of the MSV-Kom coat protein ORF.

The sequence of MSV-Kom in the vicinity of the introduced *Nco*I mutation is shown above the sequence of the pKEP170 *Nco*I mutant in the same region. The 3′ end of the V1 ORF and the 5′ end of the V2 ORF are shown. The TAA stop codon of the V1 ORF is underlined, and the 4 C-terminal amino acids of the movement protein are indicated above and below both sequences. The start codon of the coat protein gene is in boldface in both sequences. The T→G transversion introduced by site directed mutagenesis is shown in the lower nucleotide sequence in pKEP170. This is the only nucleotide change which distinguishes the wild type MSV-Kom sequence from that in pKEP170. As indicated, the mutation causes a single amino acid substitution in the MSV-Kom coat protein: the second amino acid is changed from serine (S) to alanine (A).

In order to quantify the basal activity of the virion sense promoter in the context of a non-replicating genome in the absence of Rep activity, I inactivated the *rep* gene by inducing a frame shift at position 2304. In the mutant virus construct, pKEP171, the insertion of a TG

dinucleotide resulted in the introduction of a new *Pst*I site and a +2 shift in the C1 reading frame from this site (Figure 3.2). Translation of a *rep* mRNA transcribed from this construct would thus yield a peptide which contains the first 24 amino acids of Rep and 10 missense amino acids, after which translation would terminate at an in-frame stop codon. Thus, the *rep* gene was effectively inactivated in this construct, which also contained the *Nco*I site introduced into pKEP170 (Figure 3.1).

MSV-Kom Rep

```
 M   A   S   S   S   S   N   R   Q   F   S   H   R   N   A   N
ATG GCC TCC TCC TCA TCC AAC CGT CAG TTC TCA CAC AGG AAC GCT AAC

 T   F   L   T   Y   P   K   C   P   E   N   P   E   I   A   C
ACG TTC CTA ACC TAT CCA AAG TGT CCA GAA AAT CCT GAA ATC GCC TGT

 Q   M   I   W   E   L   V   V   R
CAG ATG ATC TGG GAG CTC GTC GTT CGT ... ... ...
```

pKEP171 mutated Rep (+2 frameshift and premature stop)

```
 M   A   S   S   S   S   N   R   Q   F   S   H   R   N   A   N
ATG GCC TCC TCC TCA TCC AAC CGT CAG TTC TCA CAC AGG AAC GCT AAC

 T   F   L   T   Y   P   K   C   L   Q   K   I   L   K   S   P
ACG TTC CTA ACC TAT CCA AAG TGT CTG CAG AAA ATC CTG AAA TCG CCT

 V   R   stop
GTC AGA TGA TCT GGG AGC TCG TCG TTC GT
```

Figure 3.2: Sequence of the 5′ end of the C1 ORF and translation of the N-terminus of the protein product of the C1 ORF in wild type MSV-Kom (top) and pKEP171 (bottom).

The sequence of the MSV-Kom C1 ORF and N-terminus of the Rep protein is shown in comparison with the mutated sequence in pKEP171. The TG dinucleotide insertion which results in the introduction of a *PstI* site and +2 frameshift at position 2304 in the MSV-Kom genome is indicated underlined, and in boldface. The truncated translation product of the C1 ORF in pKEP171 is shown; 10 missense amino acids are underlined.

To divorce possible genome amplification effects from Rep transactivation of the virion sense promoter, I needed to inactivate the MSV plus-strand origin of replication with as little effect on the overall genome structure as possible. Origin of replication-deficient mutants would be useful to determine what effect Rep expressed off its own promoter, in the absence of replication, has on virion sense expression compared with double mutants which cannot express their own Rep protein. Double *rep* and origin mutants would therefore allow determination of Rep-mediated transactivation in the absence of replication, if Rep were supplied from another replicating virus, or from a Rep expression cassette. Small insertions into the loop sequence of MSV and other geminiviruses effectively abolish replication (Schneider *et al.*, 1992; Roberts and Stanley, 1994). A

convenient *Ssp*I restriction site (AAT↓ATT) occurs in the nonanucleotide sequence found in all geminivirus origins of replication (TAATATTAC). Restriction with *Ssp*I therefore generated a useful blunt-ended site in which to insert a short sequence to inactivate the MSV origin of replication. I inserted four nucleotides (CTAG) using a "restriction site mobilisation" strategy, outlined in Figure 3.3. This insertion generated a new *Xba*I site in the loop sequence of MSV-Kom in plasmid pKEP174, which contains the loop and coat protein ATG changes, and in pKEP175, the "double" loop and Rep mutant, which also bears the new *Nco*I site over the coat protein start codon.

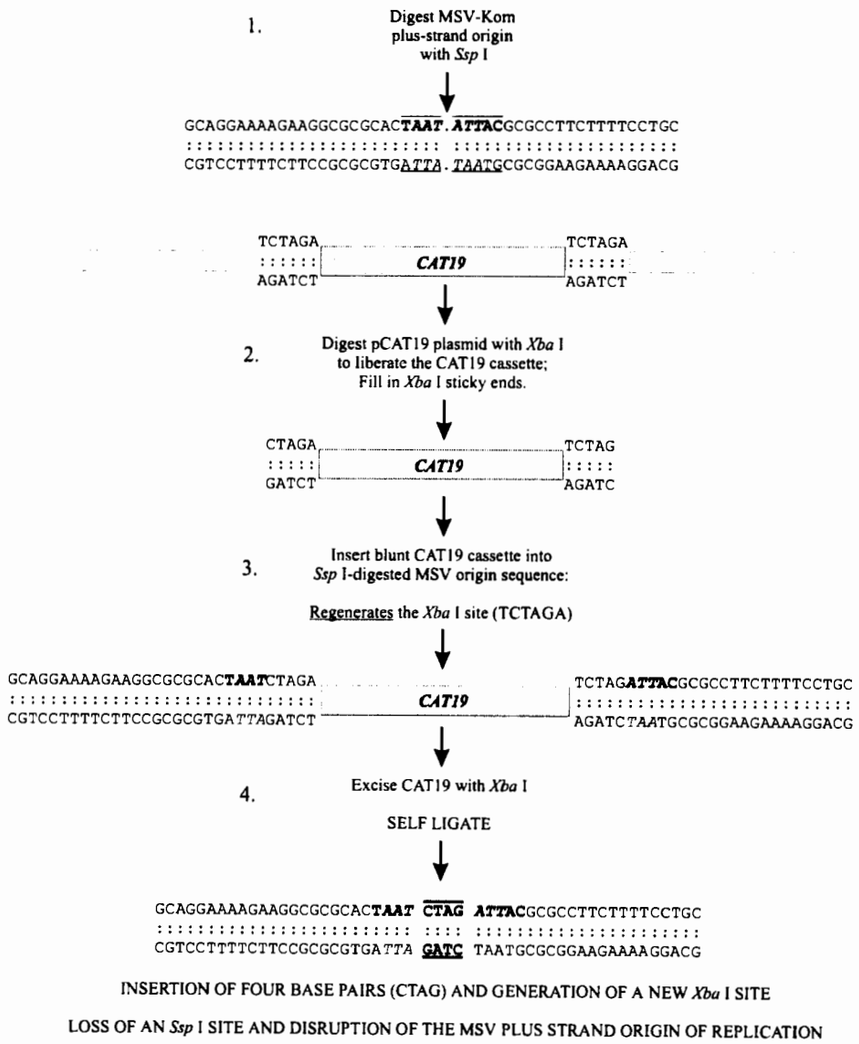


Figure 3.3: Strategy used to functionally inactivate the MSV-Kom plus strand origin of replication by introduction an *Xba*I site in the conserved loop sequence.

A chloramphenicol resistance marker gene flanked by *Xba*I sites was used to introduce four nucleotides into the *Ssp*I site in the loop sequence of the MSV-Kom stem-loop origin of replication, as outlined in the Figure.

3.3.2 Construction of reporter and transactivator plasmids

I constructed a *gusA* gene with an *Nco*I site over the GUS start codon suitable for fusion to the virion sense promoter sequences of the MSV mutants described above. The *gus* gene from pKEP*gusNco* was cloned downstream of the V1 ORF, such that the *gus* gene was cloned in an exact transcriptional and translational fusion with the coat protein expression regulatory sequences. This GUS construct was then linked to all four MSV mutant constructs to generate monomeric clones of approximately 4.0 kb. These constructs were all cloned as tandem dimers to facilitate escape of a circular monomeric copy of the recombinant virus genome by replicative release (if appropriate) or by homologous recombination. I have represented the genomic structure of the four mutant monomeric viruses in Figure 3.4B.

I constructed pKEPLuci, a luciferase reporter gene suitable for efficient expression in maize cells, for use as an internal standard for transient assays (Figure 3.4A). The Rep protein required for transactivation experiments was supplied from pKEP116, a dimer of a CaMV35S-*bar* gene replacement of the virion sense ORFs (see Chapter 2) or from a chimaeric *rep* gene expression cassette, which I constructed by cloning the MSV-Kom *rep* gene under the control of the strong rice *actin 1* upstream regulatory sequences (Figure 3.4C).

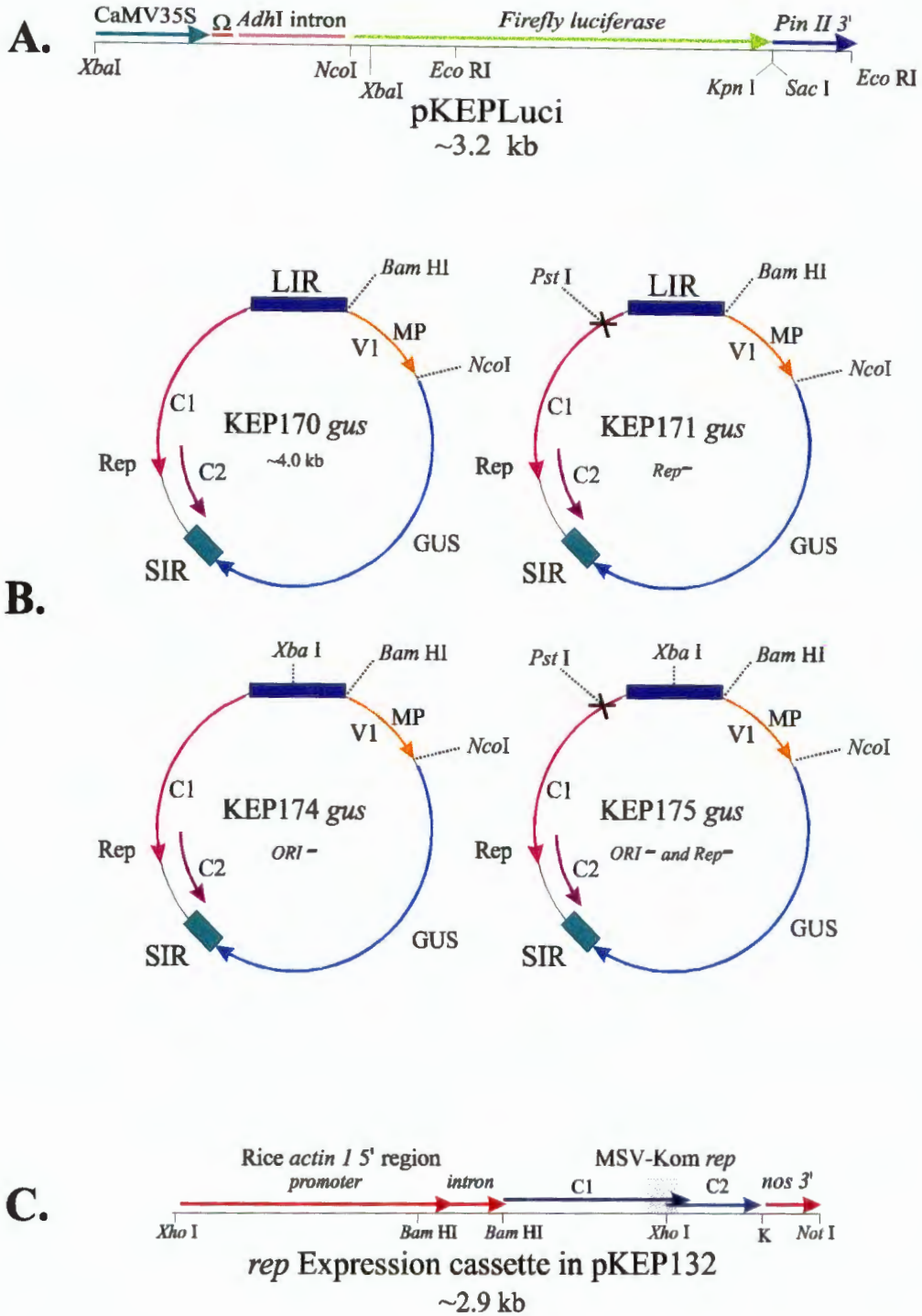


Figure 3.4: Constructs used to determine the effect of genome amplification and transactivation on the expression of the MSV virion sense promoter.

A. Luciferase expression cassette used as an internal control in transient assays. The plasmid insert is shown.

B. MSV *gus* coat protein replacement constructs. The diagram shows the expected circular monomer which could be released from the cloning vector by replicative release or homologous recombination. All constructs were cloned as dimers, linearised at the *Bam*HI site.

C. Rep expression cassette, pKEP132, which is linked to the chlorsulfuron resistance cassette from pPHP7504 (Appendix B).

3.3.3 The effect of genome amplification on expression of GUS under the control of the MSV virion sense promoter.

The plasmids which contained dimers of the four *gus* coat protein replacement constructs were introduced into BMS suspension cultured cells by microprojectile bombardment. Assays for GUS activity were done three days after bombardment, as this is the stage when replication levels appear to be highest in BMS bombarded with MSV-based as well as WDV-based, constructs (J.A. Willment, K.E. Palmer and E.P. Rybicki, unpublished; Suárez-López and Gutiérrez, 1997). Histochemical staining for GUS activity in samples of bombarded tissues showed clearly that there was significant enhancement of GUS expression from the autonomously replicating construct, pKEP170*gus*d (Figure 3.5). The small number of blue-stained cell clusters present in tissue bombarded with the three replication deficient constructs showed that there is also a low level of constitutive activity of the MSV coat protein promoter in BMS cells. This basal promoter activity was not dependent on expression of Rep, as expression-deficient mutants pKEP171*gus*d and pKEP175*gus*d both showed some promoter activity, as did pKEP174*gus*d which is capable of *rep* expression, but not replication.

Figure 3.5: Expression of GUS in cells bombarded with MSV constructs carrying the *gusA* sequence in place of the coat protein ORF.

Well 1: BMS cells bombarded with pKEP170*gus*d, which yields an autonomously replicating GUS construct; well 2: BMS bombarded with pKEP171*gus*d which cannot replicate or express functional Rep protein due to a frameshift introduced in *rep*; well 3: BMS bombarded with pKEP174*gus*d, which can express Rep from its own promoter, but which cannot replicate due to insertion of 4 bp in the loop sequence of the stem-loop origin of replication in the LIR; well 4: BMS bombarded with pKEP175*gus*d, which can neither replicate nor express Rep, as it carries both *rep* gene and loop mutations.

A quantitative analysis of GUS expression in bombarded cells showed that replicating constructs indeed expressed approximately 45-fold more GUS than cells which were expressing GUS from the coat protein promoter in the absence of replication, and 5.6-fold

more GUS than a non-replicating construct with the *gus* sequence under the control of the CaMV35S promoter and maize *adh*I first intron in pDPG208 (Gordon-Kamm *et al.*, 1990). The results of the transient assay experiments are summarised in Figure 3.6 and tabulated in Table 3.1. The constitutive expression of GUS from the MSV coat protein promoter in the absence of replication (mean of 19334.8 relative light units, RLU) was significantly higher than background in non-bombarded cell extracts, with a mean GUS activity of 1314 RLU. The data in Table 3.1 show that, although there was large variation in GUS readings observed in different replicates of the same bombardment experiment, adjustment of the results according to the expression of the luciferase internal control yielded fairly uniform readings, with acceptable standard deviations and 95% confidence intervals for the mean. The GUS/luciferase ratio for the non-replicating construct pKEP174*gus*d, which could express Rep from its own promoter was not higher than either of the constructs which could not express Rep (pKEP171*gus*d and pKEP175*gus*d). Therefore, I detected no Rep-mediated transactivation of the virion-sense promoter in this experiment.

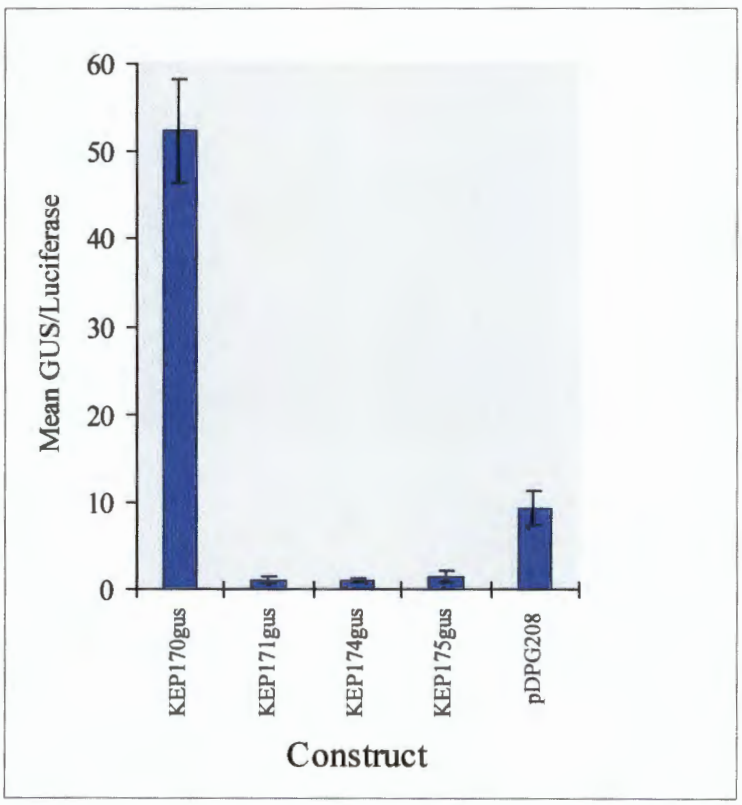


Figure 3.6: Expression of GUS in BMS cells after bombardment with replicating (pKEP170*gus*d) and non-replicating MSV-based constructs, and pDPG208.

GUS activity was standardised to luciferase expression from a standard amount (200 ng per shot) of co-bombarded pKEPLuci, and is thus expressed as a GUS/luciferase ratio. Error bars represent 95% confidence intervals.

Table 3.1: Transient Bombardment Assay Results

Plasmid/ sample no.	GUS (RLU)*	Luciferase (RLU)*	Ratio GUS/Lucif.	Mean Gus/Lucif.	Std Dev Gus/Lucif.	95% CI Gus/Lucif.
Unbomb/1	1140	1224	N/A	N/A	N/A	N/A
Unbomb/2	1866	1418				
Unbomb/3	937	1393				
KEP170*gus*/1	1722610	32028	53.785005	52.405205	9.1176326	5.956735
KEP170*gus*/2	1189020	28459	41.779958			
KEP170*gus*/3	676930	9946	68.060527			
KEP170*gus*/4	242822	5229	46.43756			
KEP170*gus*/5	911846	19359	47.101916			
KEP170*gus*/6	625639	10875	57.530023			
KEP170*gus*/7	782594	17073	45.838107			
KEP170*gus*/8	523479	8144	64.279452			
KEP170*gus*/9	681439	14550	46.834296			
KEP171*gus*/1	62237	44949	1.3846137	1.0797271	0.6400903	0.418184
KEP171*gus*/2	64819	75163	0.8623791			
KEP171*gus*/3	113838	46399	2.453458			
KEP171*gus*/4	6607	5694	1.1603442			
KEP171*gus*/5	9047	9749	0.9279926			
KEP171*gus*/6	6421	4407	1.4570002			
KEP171*gus*/7	12674	19571	0.6475908			
KEP171*gus*/8	11081	27707	0.399935			
KEP171*gus*/9	10722	25274	0.4242304			
KEP174*gus*/1	19310	25315	0.7627889	1.0252524	0.3273254	0.2138484
KEP174*gus*/2	6185	10193	0.606789			
KEP174*gus*/3	8713	13658	0.6379411			
KEP174*gus*/4	4070	3876	1.0500516			
KEP174*gus*/5	7318	6466	1.1317662			
KEP174*gus*/6	15742	9514	1.6546143			
KEP174*gus*/7	6230	5904	1.0552168			
KEP174*gus*/8	5665	4523	1.2524873			
KEP174*gus*/9	6017	5594	1.0756167			
KEP175*gus*/1	24340	9357	2.6012611	1.5242947	1.0036745	0.655721
KEP175*gus*/2	28246	8321	3.3945439			
KEP175*gus*/3	13485	10812	1.2472253			
KEP175*gus*/4	11847	8567	1.3828645			
KEP175*gus*/5	21183	10996	1.9264278			
KEP175*gus*/6	17816	10787	1.6516177			
KEP175*gus*/7	5730	13294	0.4310215			
KEP175*gus*/8	10684	20502	0.5211199			
KEP175*gus*/9	12012	21352	0.5625703			
pDPG208/1	134701	17374	7.7530218	9.2856253	2.417146	1.9340812
pDPG208/2	62125	7271	8.5442168			
pDPG208/3	126652	17871	7.0870125			
pDPG208/4	175628	19222	9.1368224			
pDPG208/5	220544	16647	13.248273			
pDPG208/6	96072.9	9661	9.9444053			

* = Relative Light Units; Std Dev = standard deviation; CI = confidence interval; N/A: not applicable

3.3.4 The effect of expression of Rep in *trans* on the expression of the GUS gene under the control of the MSV virion sense promoter

The results of the investigations into the effects of expression of Rep on the activity of the virion sense promoter are tabulated below (Table 3.2). These results showed large standard deviations, as reflected in the 95% confidence interval for the mean. This probably reflected an increased degree of experimental variability introduced by co-bombardment with three plasmids. In addition, there may also have been some instrument error at the very high levels of GUS activity observed at the upper spectrum of these readings, some of which exceeded 6 x 10^6 RLU in a 5 second count period, which may have saturated the luminometer photomultiplier tube.

Table 3.2: Evaluation of the effect of *trans*-replication and possible transactivation on the expression of the virion sense promoter

Constructs Introduced into BMS Cells	Experiment Description	Number of Replicates	Mean GUS/Luciferase Reading	95% Confidence Interval for Mean
pKEP170*gus*d	Autonomously replicating GUS construct	9	52.4	5.96
pKEP170*gus*d pKEP116	Autonomously replicating GUS and autonomously replicating *bar*	9	91.4	22.4
pKEP170*gus*d pKEP132	Autonomously replicating GUS and Rep expression cassette	9	156.0	29.3
pKEP171*gus*d	Rep gene-deficient GUS dimer	9	1.08	0.42
pKEP171*gus*d pKEP116	Rep gene-deficient GUS dimer and autonomously replicating *bar*	9	24.1	13.1
pKEP171*gus*d pKEP174*gus*d pKEP175*gus*d	Non-replicating GUS dimers (individually bombarded)	27	1.00	0.2
pKEP175*gus*d pKEP116	Unreplicatable GUS dimer (loop mutation) and autonomously replicating *bar*	9	1.30	0.32
pKEP175*gus*d pKEP116	Unreplicatable GUS dimer and Rep expression cassette	9	1.44	0.35

BMS cells which were bombarded with pKEP116 or pKEP132 in addition to the autonomously replicating GUS construct (pKEP170*gus*d) showed significant enhancement of the mean GUS expression (1.7-fold and 3-fold, respectively) in the presence of constructs which would provide extra Rep protein. The GUS expression in the *rep*-gene mutant plasmid (pKEP171*gus*d) was enhanced to approximately half the level seen in the autonomously replicating construct by provision of the Rep protein required for its replication in *trans* from pKEP116 (Table 3.2). However, expression of Rep from neither the replicating construct nor the Rep expression cassette caused any significant increase in the expression of GUS from the coat protein promoter in the construct which was incapable of replication due to the inactivation of its origin of replication (pKEP175*gus*d). Unfortunately, there was insufficient time to test the effect of Rep expression from pKEP132 on GUS expression in cells bombarded with pKEP171*gus*d.

3.3.5 Co-transformation of BMS with autonomously replicating *bar*-constructs and GUS-coat protein replacement constructs

To generate transgenic cell lines, I bombarded BMS cells with equal amounts of pKEP116 and each of the dimeric GUS-coat protein replacement constructs described above (sections 3.3.1 and 3.3.2). Two precipitations were done for each GUS construct, and a total of 24 plates of BMS cells were bombarded, six plates for each treatment. This was a particularly efficient transformation experiment: I was able to recover well over 200 independently transformed bialaphos resistant cell lines for each of the GUS plasmids. The initial stage of screening of the bialaphos resistant cell lines was by histochemical staining with X-Gluc. Table 3.3 shows the number of lines screened for each construct and the proportions which showed GUS activity. Figure 3.7 shows photographs of cell lines stained for GUS activity. Most GUS-positive cell lines showed varied intensity and distribution of blue stain throughout the callus samples.

Table 3.3: Screening of bialaphos resistant transgenic cell lines for GUS activity

pKEP170-series Plasmids used in bombardment	Number of bialaphos resistant cell lines screened	Number of resistant cell lines displaying blue pigment	Percentage of GUS positive cell lines
pKEP170*gus*d	54	22	41
pKEP171*gus*d	24	6	25
pKEP174*gus*d	24	4	17
pKEP175*gus*d	42	0	0

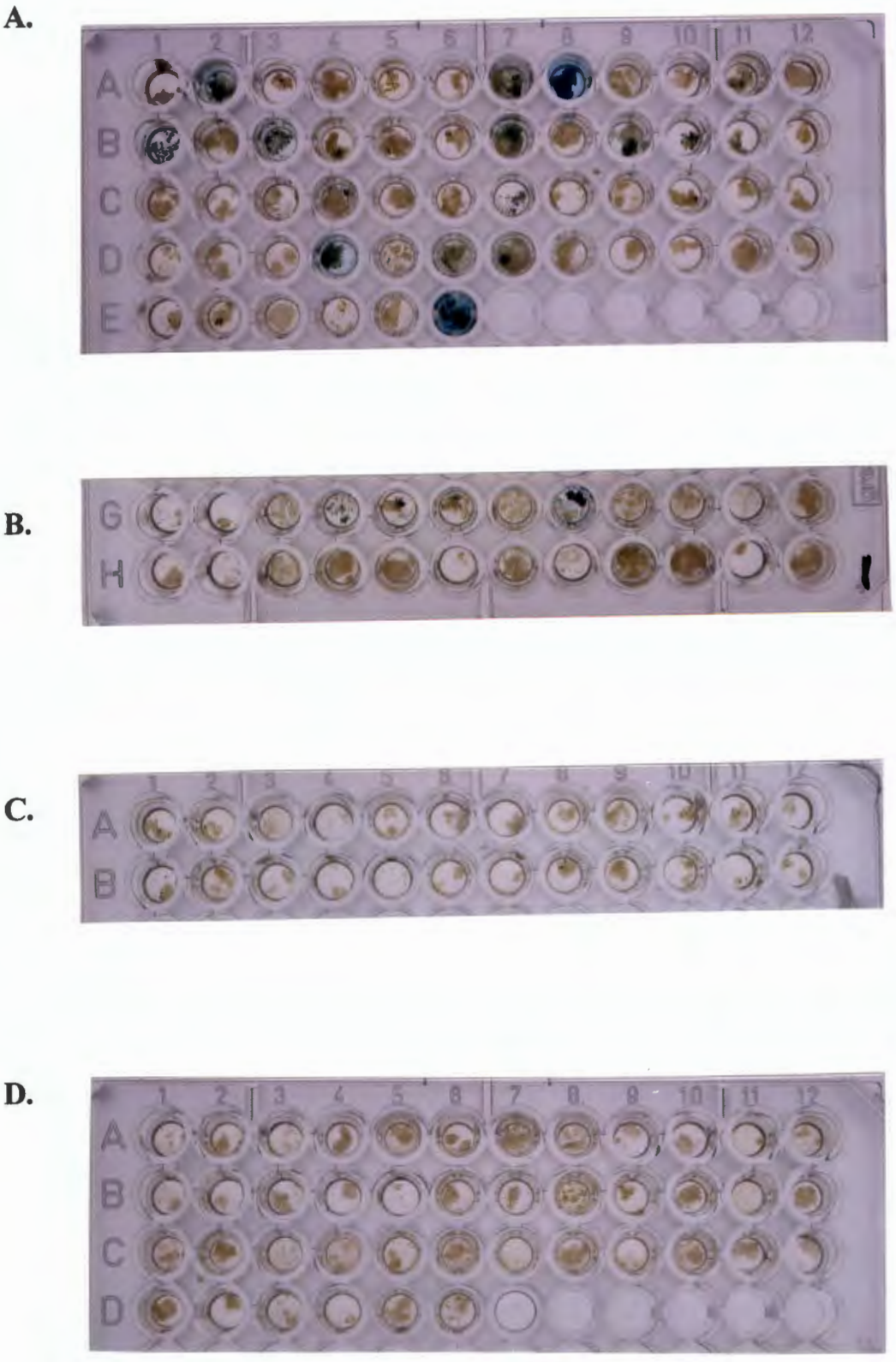


Figure 3.7: Expression of GUS in bialaphos resistant BMS cell lines.

Cell lines were generated by bombardment with pKEP116 and the following dimeric GUS constructs:

A: pKEP170*gus*d, autonomously replicating GUS replacement of the coat protein ORF.

B: pKEP171*gus*d, replication-deficient GUS construct, with a frameshift in the *rep* gene.

C: pKEP174*gus*d, replication-deficient GUS construct, with an insertion in the plus strand origin of replication (loop mutation).

D: pKEP175*gus*d, replication-deficient GUS construct carrying two mutations: a frameshift in the *rep* gene and the loop mutation.

3.3.6 Quantitation of GUS expression levels in bialaphos resistant cell lines

The GUS expression levels in all transformed cell lines which showed blue staining with X-Gluc were quantified by the GUS-Light™ assay (Tropix); eight bialaphos resistant lines which had been generated by bombardment with pKEP175*gus*d and pKEP116 were also included in the assay to quantify the GUS activity typically found from the coat protein promoter in the absence of replication. There was a large amount of variation in the GUS expression levels between different lines, from a level not significantly above background (100-200 light units/μg of protein) to close to 100 000 LU/μg. These results are summarised in Table 3.4. I classified the GUS expression levels into four different categories: (I) between 100 and 200 LU/μg protein, which falls within the range measured for 5 unbombarded samples; (II) 200 to 1000 LU/μg; (III) 1000 to 10000 LU/μg; and (IV) 10000 to 100000 LU/μg of protein.

Table 3.4: Classification of bialaphos resistant cell lines according to GUS expression levels.

Plasmid used to generate transgenic cell lines	Total number of cell lines	Number of cell lines in each expression category			
		I	II	III	IV
pKEP170*gus*d	22	1	8	8	5
pKEP171*gus*d	6	-	1	3	2
pKEP174*gus*d	4	-	2	1	1
pKEP175*gus*d	8	1	6	1	-

The bar chart in Figure 3.8 shows the range of GUS expression levels seen in bialaphos resistant cell lines. There seemed to be a great deal of variation in the amount of GUS expressed in different transgenic lines, with 95% of the lines expressing GUS levels greater than background, and 21 out of 40 lines expressing levels higher than expected from the cell lines which should not contain any episomal DNA, i.e. those generated with pKEP175*gus*. Significantly, those cell lines which showed the most intense GUS staining in the histochemical assay were not in all cases the same lines which showed the highest GUS expression in this quantitative assay.

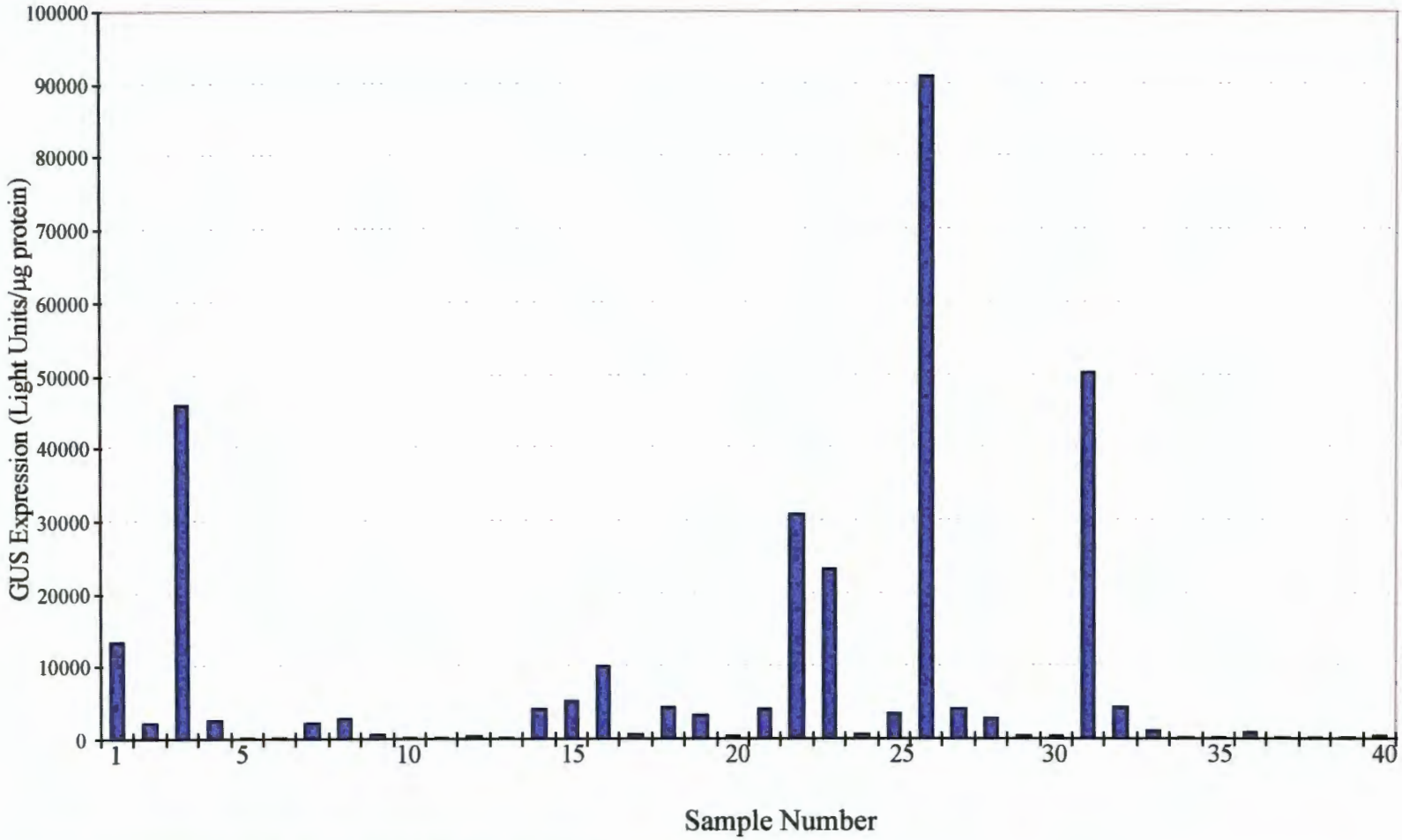


Figure 3.8: GUS expression levels in 40 independently transformed BMS cell lines.

GUS activity was determined by chemiluminescence. Sample numbers 1 to 22 contained pKEP116 (autonomously replicating MSV construct with bialaphos resistance marker) and pKEP170*gus*d; samples 23 to 28 contained pKEP116 and pKEP171*gus*d; 29 to 32 contained pKEP116 and pKEP174*gus*d; and samples 33 to 40 contained pKEP116 and pKEP175*gus*d.

3.3.7 Screening of bialaphos resistant cell lines for the presence of extrachromosomal DNA.

All of the bialaphos resistant BMS lines which showed GUS activity greater than 1000LU/μg of total protein, and some which showed lower levels, were analysed for the presence of high copy number viral replicons in their low molecular weight DNA, by Southern hybridisation. The membrane was first hybridised with a GUS-homologous probe to determine which lines contained high copy number pKEP170*gus*-series replicons, and then stripped and re-probed with a *bar* probe (shown in Figure 3.9).

Of the 36 bialaphos resistant cell lines tested, 30 contained *bar*-homologous episomal DNA (Figure 3.9) three months after bombardment. All of the *bar* replicons appeared to be of the same size, and there was no evidence for the accumulation of any DNA forms which might represent subgenomic-sized DNA. There was very good correlation between the presence of pKEP116-derived replicons and the various GUS-derived replicons in transformed cells: 25 out of 36 lines contained extrachromosomal GUS-homologous DNA forms. There was one cell line (Figure 3.9 lane A/C 18) which contained some extrachromosomal DNA homologous to the GUS probe and no *bar*-homologous episomes, but all other *bar*-negative cell lines also contained no GUS replicons.

In contrast to the 3.15-kb *bar* replicons, many of the GUS construct-derived episomes appeared to have accumulated DNA forms which were smaller than the full genome size (4.0 kb). Some of these putative subgenomic DNAs are indicated in Figure 3.9. There was good correlation between the GUS expression levels and copy number and intactness of the GUS replicons, with the highest GUS expressors also containing the highest amounts of intact GUS replicons. I have indicated the band which I presume represents the 4.0 kb GUS replicon cccDNA form in Figure 3.9. Transgenic cell lines which expressed the highest GUS levels, i.e. in class IV, (section 3.3.6) are indicated in the Figure. In every case, the lines which showed GUS activity less than 1000 LU/μg of protein were those which showed no GUS episomes. There were obviously GUS replicons present in lines generated by bombardment with pKEP171*gus*d, which has a frameshift in the *repA* ORF, and pKEP174*gus*d, with a mutated loop sequence, and in one pKEP175*gus*d line which contains both mutations (lane D13). The GUS replicons in the pKEP171*gus*d lines may have been *trans*-replicated by the *bar* replicon, but those in the pKEP174*gus*d lines must represent recombinants.

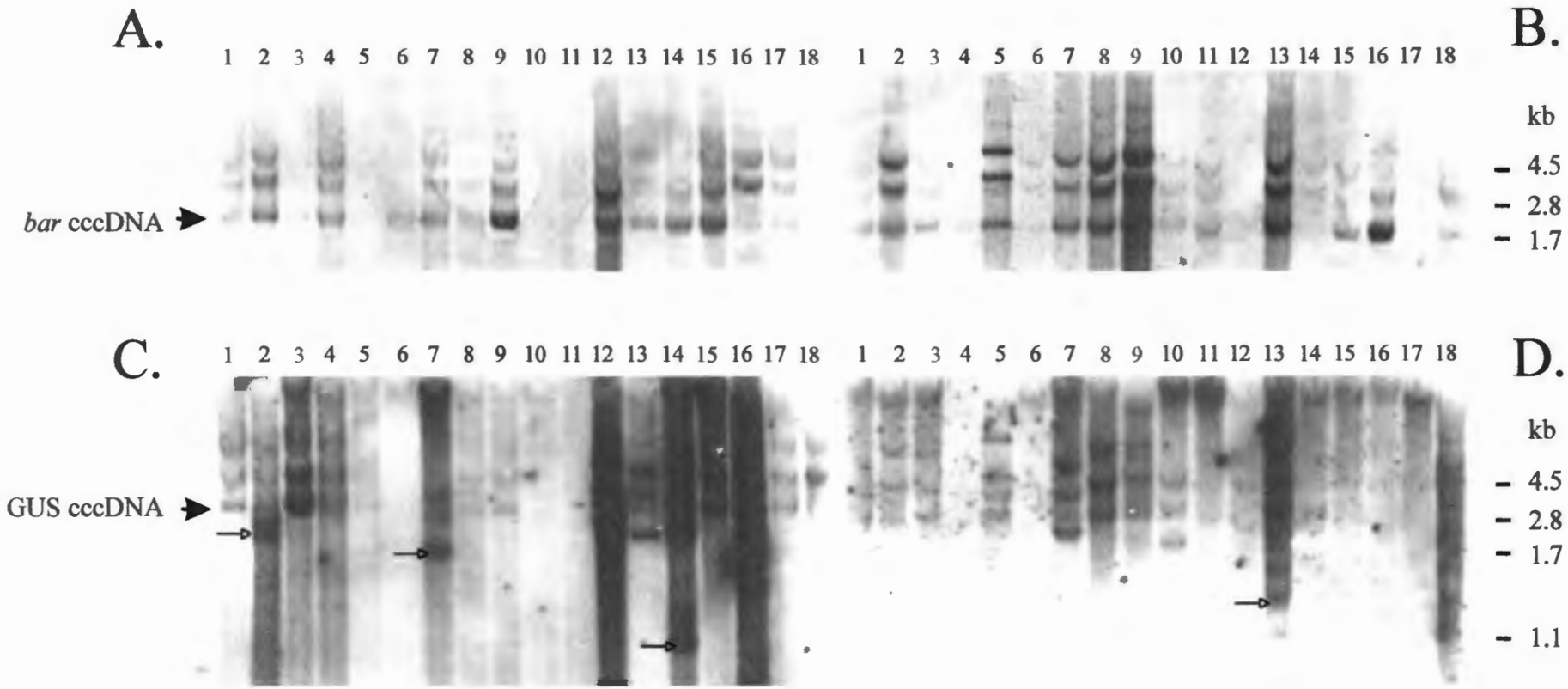


Figure 3.9: Replication of *bar* replicons derived from pKEP116 and GUS replicons derived from pKEP170gusd-series plasmids in bialaphos resistant cell lines.

The blots were probed first with a digoxygenin-labelled GUS probe (C & D), then stripped and re-probed with a digoxygenin-labelled *bar* probe (A & B). I have indicated the cccDNA forms of both the *bar* replicon and the GUS replicon (closed arrows). The open arrows indicate possible subgenomic DNAs derived from the GUS replicons. Transformed lines which expressed the highest levels of GUS, i.e. in class IV, are in lanes A/C 1, 3, 7, 15 & 18 and B/D 3, 8 & 9.

3.3.8 Detection of recombination between MSV-derived replicons in transformed BMS

Results described in the previous section (3.3.7), suggested that some of the GUS replicons may have arisen out of a recombination event between the *bar* replicon, derived from pKEP116, and the replication-deficient inserts contained in plasmids pKEP171*gus*d, pKEP174*gus*d and pKEP175*gus*d. Accordingly, I examined whether the mutagenesis marker restriction sites were present in the replicon DNA. PCR primers MSV 215-234 and MSV 1770-1792 amplify a 1154 base pair fragment encompassing the RepA and part of the RepB ORFs, the LIR and part of the movement protein ORF from circular viral DNA. The regions where these primers anneal were both present in the GUS replicon DNA, but the annealing site for MSV 215-234 was absent in the *bar* replicons, which have no V1 ORF. Thus, PCR would amplify only those regions present in the GUS constructs; restriction with *Xba*I would identify the loop mutation present in pKEP174*gus*d and pKEP175*gus*d, and restriction with *Pst*I would identify constructs which contained the frameshift mutation in the Rep A ORF. The results of the analysis of 9 lines are shown in Figure 3.10.

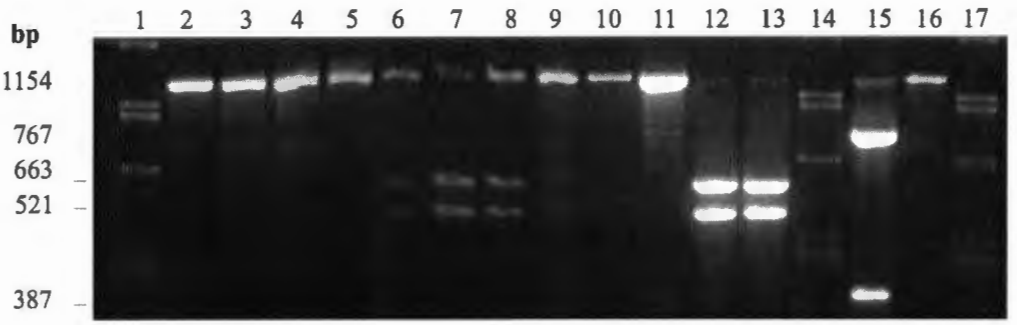


Figure 3.10: PCR and restriction digest analysis of replicon DNA in bialaphos resistant cell lines generated by bombardment with pKEP116 and various mutant GUS constructs.

Lanes 1, 14 and 17 contain molecular weight marker DNA (λ DNA digested with *Pst*I). The size markers (in kb) are, from top to bottom: 1.70 ; 1.16; 1.09; 0.81; 0.51; 0.47. Lanes 2 to 5 contained PCR products amplified from cell lines containing pKEP170*gus*d-derived replicons. Lanes 6 to 9 contained PCR products amplified from cell lines containing pKEP171*gus*d-derived replicons and lane 10 contained a PCR product from a cell line which contained DNA from pKEP175*gus*d. Lanes 11, 12 & 13 contained PCR products amplified from pKEP170*gus*d, pKEP171*gus*d and pKEP175*gus*d plasmid DNA. The PCR products in lanes 2 to 13 were all digested with *Pst*I. Lane 15 contained the PCR product amplified from pKEP175*gus*d plasmid DNA, digested with *Xba*I and lane 16 contained the PCR product amplified from the same bialaphos resistant cell line as in lane 10, digested with *Xba*I. The 1154 bp band present in the digested products from plasmid-amplified DNA represents residual undigested DNA.

In the first instance, these results showed that, if subgenomic forms of the GUS replicons accumulated in the extrachromosomal DNA, there were no detectable deletions in the region which spanned the PCR product, as no products smaller than the full-length 1154-bp were amplified. In cell lines generated with the *rep* gene mutant (pKEP171*gus*d), a substantial proportion of the replicons present still contained the *Pst*I mutation, but there appeared to be some which had lost the *Pst*I site, presumably by recombination with the *bar* replicon (Figure 3.10, lanes 6 to 9). Thus, there were both autonomously replicating GUS constructs, and constructs which were being replicated in *trans* present in these cell lines. The PCR product amplified from the transgenic line generated by bombardment with the double mutant, pKEP175*gus*d did not contain either mutation, which indicated that the GUS replicon in this cell line had arisen by recombination between the input plasmid DNA and the *bar* replicon. This was the only pKEP175*gus*d-derived cell line which showed GUS activity higher than 1000 LU/μg of protein (Figure 3.8) and which contained episomal GUS-homologous DNA (Figure 3.9, lane D13).

3.4 DISCUSSION

The results presented in this Chapter show conclusively that amplification of the GUS gene by linkage to an MSV replicon can significantly enhance its expression from the coat protein promoter, both in transient assays and in stably transformed cells. In transient assays, I introduced MSV constructs with the GUS gene, cloned as a transcriptional and translational fusion to the coat protein regulatory sequences, into BMS cells by microprojectile bombardment. I found that genome amplification enhanced GUS expression approximately 45-fold over the expression level seen from non-replicating constructs. This enhancement is less than the 60- to 90-fold noted by Brough *et al.* (1992) for replicating constructs with GUS replacements of the coat protein sequences of TGMV, but significantly higher than the 20-fold enhancement found for constructs which contained the *nptII* gene in place of the WDV coat protein sequence (Matzeit *et al.*, 1991). The data reported in this Chapter concur with those of Suárez-López and Gutiérrez (1997), who found that WDV replication was required for maximal expression of the *gus* gene cloned under the control of the WDV virion sense promoter. These authors did not, however, provide any quantitative analyses of the GUS expression levels from these virus constructs.

It is important to note that both WDV studies (Matzeit *et al.*, 1991 and Suárez-López and Gutiérrez, 1997) used constructs where the *gus* gene was cloned as only a transcriptional fusion to the virion sense promoter, at a site upstream of the coat protein start codon. In contrast, I designed my constructs such that the *gus* gene used the coat protein ATG, so any upstream translational control sequences which may be important for the coat protein expression remain intact. The *Nco*I site which I introduced over the coat protein ATG start codon to facilitate cloning of the *gus* gene did not change any nucleotides upstream of the ATG. However, the T→G transversion mutation incidentally changed the mRNA translation initiation context to one that more closely resembles the consensus sequence for optimum translation initiation in plants (Gallie, 1993; Joshi, 1987; Lutcke *et al.*, 1987), where a G at +4 has been shown to be important (Kozak, 1989). While this may have enhanced the coat protein expression levels compared with those found in the wild type sequence context, the change does not seem to alter symptom phenotype in sweetcorn agroinfected with a dimer of pKEP170 (K.E. Palmer, E. van der Walt and E.P Rybicki, unpublished; Chapter 4).

It is difficult to reconcile my results on the effect of genome amplification on coat protein expression with those of Hofer *et al.* (1992), who postulated that genome amplification only doubles the coat protein expression, and that Rep-mediated transactivation of the virion sense promoter is primarily responsible for the enhanced levels of coat protein expression associated with replication. I was unable to find any transactivation of MSV coat protein promoter expression, with Rep supplied from its own promoter in non-replicating constructs, or in *trans* from another viral replicon, or from a Rep expression cassette. However, as I pointed out in Chapter 1, these authors may have misinterpreted their results, as Heyraud *et al.* (1993a) subsequently showed that a similar stem-loop deletion mutant as used by Hofer *et al.* (1992) was still capable of replication of concatemeric genome forms.

Increasing levels of Rep protein from another, smaller replicon which should replicate to higher levels (Suárez-López and Gutiérrez, 1997), or from a Rep expression cassette, enhanced the GUS expression levels from replicating constructs even further. Expression of Rep in *trans* also increased the GUS expression in constructs which carried a frameshift mutation in the *rep* gene, albeit to a lower level than autonomously replicating constructs, which implies that these constructs were being replicated in *trans*. Together, these results

suggest that, either the concentration of Rep is the factor which limits the copy number of these constructs, or alternatively that the coat protein promoter is more efficiently transactivated when the concentration of Rep is increased, or both. My studies did not address the evidence of Collin *et al.* (1996) that production of RepA is required for coat protein expression, given that constructs with *rep* cDNAs replicate to higher copy numbers than viruses with the *rep* intron intact, but do not express very much GUS from the virion sense promoter.

There are several possible explanations for the fact that I did not observe transactivation of the virion sense promoter in constructs not able to replicate because of a mutation in the plus strand origin of replication. These include the possibilities that: (1) transactivation of the coat protein promoter by Rep does not occur in MSV; (2) Rep-mediated transactivation of the coat protein promoter requires a host factor (or the coat protein itself) which is not present in BMS cells; or (3) that the insertion of four nucleotides (CTAG) into the loop sequence of the MSV plus strand origin of replication destroys a recognition site vital for the transactivation phenomenon. It is difficult to believe that there is no specific viral control of coat protein promoter activity: transactivation of the promoter is an elegant way of a viral "early" protein activating "late" functions. In the light of suggestions that Subgroup I Rep binds to specific sequences within the stem-loop structure (Argüello-Astorga *et al.*, 1994a) and speculations that Rep may recognise the stem loop structure specifically in these viruses (Chapter 1), it is possible that the CTAG insertion in the loop sequence disrupts Rep binding at the stem-loop sequences, and therefore abolishes transactivation. Control of coat protein expression therefore still requires further investigation in Subgroup I geminiviruses.

I was able to show, for the first time, that transgenic cell lines can be produced which contain two separate replicating geminivirus-based replicons. Although there was a great deal of variation in the GUS expression levels between different transformed cell lines, the results showed that one can achieve up to about 90-fold enhancement of GUS expression in transgenic cell lines which contain replicating episomal GUS constructs over lines where the GUS construct is present, but not replicating.

The variability observed in GUS expression levels between different transgenic lines may be ascribed to one or more of several factors. I noticed that the distribution of GUS-

positive sectors in stained calli was often not uniform, and also that the best expressors in the chemiluminescent GUS assay were not in every case the lines which showed most GUS staining in the histochemical assay. This would suggest that the replication level of the GUS replicon is not constant throughout the callus tissue, and that the copy number observed reflects the average of varied copy numbers in different regions of the callus, perhaps dependent on the cell cycle. Alternatively, as suggested by Meyer *et al.* (1989; 1992), mobilisation of an integrated replicon from a dimeric copy might be only an infrequent event, which occurs rarely and in only a small proportion of the transformed cells. Another factor which might account for variation in GUS expression might be the accumulation of subgenomic deletion derivatives of the replicating GUS constructs, which might function as defective interfering constructs.

Smaller than unit-sized DNAs seemed to be present in some lines with replicating GUS constructs. This implies that these are not completely structurally stable, although I detected no deletions in the virus sequences amplified by PCR. It is possible that if deletions occurred, these were within the foreign insert sequences. It is possible that selection for smaller constructs was imposed by the presence of the smaller *bar* replicon in the same cell. Alternatively, it is possible that high levels of GUS expression put a physiological load on the cells, which resulted in selection for lines in which deletions had occurred in the GUS replicons; in this case, one could speculate that the *bar* replicon was structurally stable because the low expression level of this gene did not result in stress to the cell line. It is clear that the viral constructs in the transformed cells efficiently recombine with each other, given that lines with GUS constructs carrying mutations in the *rep* gene in some cases lost this mutation, and that the GUS replicon in a line which had been bombarded with a double *rep*-loop mutant had lost both mutations in the presence of an autonomously replicating *bar* replicon.

In conclusion, the results presented in this Chapter show that the gene amplification afforded by linkage to an MSV replicon is dependent to some extent on the gene which is amplified. The expression of a gene which has proven utility as a marker gene, *gusA*, was significantly enhanced over that found for the *bar* gene, in Chapter 2. The fact that GUS constructs could be used in the generation of stably transformed cell lines which contained extrachromosomally replicating virus constructs shows that a selectable marker gene is not a vital requirement for use of MSV as a useful gene amplification system. It is important to

note that the larger GUS replicons may not be completely structurally stable in the presence of the *bar* construct, so the development of a binary gene amplification system might be fraught with problems involving homologous recombination. This might therefore require that two different viruses with limited sequence homology be used.

CHAPTER 4

INVESTIGATIONS INTO THE USE OF MAIZE STREAK VIRUS AS AN INFECTIOUS GENE VECTOR

TABLE OF CONTENTS

SUMMARY

The potential for use of MSV as an infectious gene vector by complementation of movement and encapsidation functions in *trans* was investigated. When vectors of identical size to the wild type virus, carrying the CaMV 35S promoter and *bar* genes in place of the virion sense ORFs were introduced into BMS cells together with the wild type virus, ssDNA forms of the *bar* construct were visible. This showed that wild type virus can complement the ssDNA-negative phenotype of coat protein deletion constructs. Agroinfectious plasmids were constructed which contained a dimer of the *bar* gene replacement construct cloned in tandem with: (1) the wild type virus genome and (2) a replication-deficient version of the wild type genome which contained a frameshift mutation associated with a *Pst*I site in the *rep* gene and an *Nco*I restriction site over the coat protein start codon. Both constructs were agroinfectious in sweetcorn cv. Jubilee, but *Agrobacterium* carrying binary vectors with dimers of the *bar* gene replacement or the replication-deficient *Pst*I-mutant did not induce infections alone. A mixture of separate *Agrobacterium* strains carrying the *bar* gene replacement and the *Pst*I mutant virus resulted in 6% of inoculated seedlings developing symptoms of MSV infection. Analysis of the viral DNA present in plants agroinfected with all of the plasmids which contained a dimer of the *bar* gene replacement construct showed that the mutant virus was present in up to the first four leaves of the seedlings, but that the presence of wild-type or replication deficient virus (*Pst*I mutant) did not complement movement of the *bar*-containing replicon. Neither the *bar* replicon nor the *Pst*I mutant was separately transmitted by leafhoppers (*Cicadulina mbila*). When the *bar* replicon was present in the same cells as the *Pst*I mutant, replication- and movement-competent virus resulted from homologous recombination between the two constructs. This recombinant virus still contained the *Nco*I mutation over the coat protein ORF, and was still transmissible by *C. mbila*, despite a single amino acid substitution (serine to alanine) present at the N-terminus of the coat protein.

A clone of the genome of Digitaria streak virus (DSV), which is not transmissible by *C. mbila*, was dimerised and was shown to be agroinfectious in sweetcorn. An infectious 1.1-mer of the MSV-Kom genome was cloned in tandem with the agroinfectious DSV construct, so that sweetcorn seedlings could be simultaneously agroinfected with both DSV and MSV-Kom simultaneously. DSV could be transmitted by *C. mbila* from these doubly infected plants to uninfected sweetcorn. These results showed that MSV-Kom can *trans*-encapsidate the DSV genome, indicating either that there is no specific encapsidation signal on the genome of Subgroup I geminiviruses, or that this in conserved between MSV and DSV.

4.1 INTRODUCTION

As cereal-infecting Subgroup I geminiviruses are not mechanically transmissible to plants as cloned DNA, agroinfection (Grimsley *et al.*, 1986; 1987) is currently the only reliable method to introduce cloned Subgroup I geminivirus genomes, and recombinant viral genomes functioning as vectors for foreign genes, into plants. Agroinfection relies on the natural interaction between *Agrobacterium* and plant cells, where, in a process which resembles bacterial conjugation, even to the level of pilus formation, the bacterium transfers a piece of its DNA (the T-DNA) into the plant cell. The molecular processes involved in transfer of the T-DNA from *Agrobacterium* to plant cells, and integration of the T-DNA into the plant genome, were reviewed recently by Tinland (1996) and Zupan and Zambryski (1995). The T-DNA is delimited by two (imperfect) 25-bp direct repeats, termed the right and left borders. The border sequences are the only *cis* requirements for T-DNA transfer, so the internal T-DNA sequences may be replaced with foreign genes.

In constructing agroinfectious clones of geminiviral genomes, or derivatives thereof, the viral DNA is inserted within the T-DNA borders as a tandem dimer or partial dimer, to facilitate escape from the T-DNA by replicative release or by homologous recombination between the duplicated genomic sequences (Stenger *et al.*, 1991; Heyraud *et al.*, 1993b). The bacterial virulence proteins VirD1 and VirD2 recognise the T-DNA border sequences and induce nicks at these points; VirD2 protein becomes covalently linked to the 5′ end of the T-DNA (at the right border) by a phosphotyrosine linkage. A ssDNA copy of the T-DNA with the multimeric viral genome is released, presumably by displacement of the nicked strand in a process analagous to rolling circle replication. The T-strand is then coated with the ssDNA binding protein, VirE2, which protects it against nuclease attack, thus ensuring its integrity during transfer from the bacterium to the plant nucleus (Tinland *et al.*, 1994; Rossi *et al.*, 1996). The products of the *vir*B operon form a membrane pore and pilus structure through which the T-strand is transferred into the plant cell (Fullner *et al.*, 1996), piloted by the VirD2 protein, which guides the T-DNA to the nucleus by virtue of the fact that it has strong nuclear localisation signals (Tinland *et al.*, 1992; Rossi *et al.*, 1993). The transfer of the T-DNA from *Agrobacterium* to the plant cell is thus polar, with the DNA proximal to the right border leading. Sometimes truncated versions of the T-DNA are found integrated into the plant genome, but these almost always contain the 5′

end of the T-DNA, as this is the first part to exit the bacterial cell and to enter the plant nucleus (Tinland, 1996).

Bipartite Subgroup III geminiviruses, such as ACMV and TGMV, generally do not require the coat protein for infectivity. This fact facilitated the development of infectious coat-protein replacement constructs for use as gene vectors by Ward *et al.* (1988) and Hayes *et al.* (1988; 1989). In contrast, one is faced with a major obstacle in the development of Subgroup I geminiviruses as gene vectors: the viral coat protein and movement protein genes are essential for infectivity (Boulton *et al.*, 1989b; Lazarowitz *et al.*, 1989; Woolston *et al.*, 1989). In addition, there seems to be a restriction on the size of the viral DNA which may be moved systemically in agroinfected plants, although this does not extend to the ability of the viral construct to replicate (Shen and Hohn, 1992; 1994; 1995). Boulton *et al.* (1989b) found that mutations in the movement protein and coat protein genes could be complemented in *trans* by co-infection with complementing MSV mutants. I therefore decided to attempt to complement the movement and encapsidation-defective phenotype of a wild type-sized MSV construct carrying a chimaeric CaMV35S-*bar* expression cassette in place of both virion sense ORFs. To this end, I co-agroinfected one set of maize plants with this gene-replacement construct and the wild type genome, and another with a replication-deficient mutant genome which would require the Rep protein from the co-infected *bar* construct for its replication.

It is not clear whether the requirement of monopartite geminiviruses for the coat protein to initiate systemic infection represents a requirement for encapsidation, or whether the coat protein functions in movement of nucleoprotein complexes as the bipartite virus BV1 movement protein does. The requirements for encapsidation of geminivirus genomes also remain uninvestigated, apart from an elegant study by Briddon *et al.* (1990). These authors showed that the viral coat protein is the sole determinant of insect vector specificity by replacing the coat protein ORF of the whitefly-transmitted ACMV with that of the leafhopper transmitted Subgroup II virus, BCTV. Recombinant viruses were infectious, and could be transmitted by leafhoppers after intrahaemocoelic injection. This showed that both the DNA A and DNA B components of the chimaeric ACMV were encapsidated by the BCTV coat protein, and implies that there is no specific encapsidation signal present on the virus genome, unless this is conserved between ACMV and BCTV, and is present in the ACMV common region, since this is the only sequence conserved between the two

genomic components of a Subgroup III geminivirus. Given that coat protein and movement protein are both required for infectivity of Subgroup I geminiviruses, construction of infectious gene vectors should not interfere with sequences which are implicated in encapsidation, or at least in interaction between the viral genome and the coat and movement proteins.

I addressed the question of whether there is a specific sequence requirement for encapsidation of MSV in two different ways. The first was by examining whether the ssDNA formation-negative phenotype of an MSV coat protein replacement mutant exactly the same size as the wild type genome could be complemented in *trans* by the wild type virus. The second assay was a biological one, involving co-agroinfecting sweetcorn seedlings with MSV and DSV and subsequent attempted leafhopper transmission. MSV is transmitted by *Cicadulina* spp., mainly *C. mbila* Naude (Rose, 1978), while DSV is transmitted by *Nesoclutha declivata* Linnavuori and cannot be transmitted by *C. mbila* (Julia and Dollet, 1989; Pinner *et al.*, 1988). Thus, in plants doubly infected with both viruses, DSV may be transmitted by *C. mbila,* if there is *trans*-encapsidation of DSV by the MSV coat protein and the coat protein is the sole determinant of vector specificity.

4.2 MATERIALS AND METHODS

All standard molecular biological techniques were performed as described in Sambrook *et al.* (1989), and according to the specifications of enzyme manufacturers (Boehringer Mannheim, Promega and Amersham).

4.2.1 Clones in *Agrobacterium* binary vectors for agroinfection

MSV, DSV and mutant or recombinant viral constructs were routinely cloned as tandem dimers for introduction into *Agrobacterium* binary vectors pBI121 (Clontech, Palo Alto, CA.) or pBin19 (Bevan, 1984). When I required two constructs to be transferred into maize plants, these were usually cloned on the same T-DNA.

Two different agroinfectious clones of the MSV-Kom genome were used in these studies: pKom504 contains a 1.4-mer of the MSV-Kom genome which contains only one LIR, thus

the virus escapes from the T-DNA by homologous recombination only; pKom603 contains a 1.1-mer of the MSV-Kom genome consisting of the full-length genome cloned in tandem with a copy of the LIR; in this case MSV can escape by replicative release or homologous recombination. These plasmids were supplied by Dr F.L. Hughes and W.H. Schnippenkoetter, of this laboratory.

All of the MSV-Kom-based recombinant viral clones constructed by me were named with the same "pKEP" designation as the multimer of the clone in pUC, followed by "BI" to indicate that the insert from the pUC plasmid had been transferred into the binary vector pBI121.

pKEP177BI

The mutant form of the MSV-Kom genome in pKEP171, with a frameshift and *Pst*I site introduced in the *repA* ORF and *Nco*I site over the start codon of the coat protein gene, was described in Chapter 3. A tandem dimer of pKEP171 was cloned by digesting the plasmid completely with *Sca*I and partially with *Bam*HI; the appropriate fragments consisting of the full-length genome linked to part of the pUC19 plasmid were gel purified and ligated to create the tandem dimer, pKEP177. See Figure 2.2 for a graphic representation of the cloning principle. The dimeric insert in pKEP177 was excised with *Xba*I and *Eco*RI, and cloned into pBI121 which had been digested with the same enzymes, to yield pKEP177BI.

pKEP152BI

The construction of this tandem dimer of a gene replacement construct of MSV, with the movement and coat protein ORFs replaced by a CaMV35S-*bar* expression cassette, was described in Chapter 2, section 2.2.1. The dimeric insert was also excised with *Xba*I and *Eco*RI and inserted into pBI121 to create pKEP152BI.

pKEP153BI

This clone contained both wild-type MSV-Kom and the insert from pKEP152, cloned in a tandem "trimer" or, more accurately, a "hetero-trimer". It was constructed by digesting pKEP152 with *Bam*HI and *Sca*I and pKom500 (wild-type monomer) with *Sca*I and partially with *Bam*HI. The desired fragments were purified and ligated to create pKEP153. The 8.1-kb insert was excised from pKEP153 with *Xba*I and *Eco*RI and inserted into *Xba*I and *Eco*RI-digested pBI121 to generate pKEP153BI.

pKEP154BI

This clone consisted of a dimer made by digesting pKEP151 (a monomeric clone with the CaMV 35S-*bar* cassette in place of both virion sense ORFs; Chapter 2) with *Bam*HI and *Sca*I and pKom500 with *Sca*I and partially with *Bam*HI. The appropriate fragments were purified and ligated to create pKEP154, which is effectively a 1.6-mer of the wild-type MSV-Kom genome. The insert from pKEP154 was cloned into pBI121 to create pKEP154BI.

pKEP178BI

This clone contained both the replication-deficient clone of MSV-Kom, pKEP171, and the insert from the pKEP152 dimer: i.e. another "trimeric" construct. As for pKEP153BI, pKEP178BI was constructed by linking the insert of pKEP152 to pKEP171 to create pKEP178, the insert of which was then cloned in pBI121.

pDSV200Bin

The RF genome of DSV from *Digitaria setigera* from Vanuatu was cloned as a *Hin*dIII fragment in pUC18 by Dr F.L. Hughes, independently of Donson *et al.* (1987); this clone was called pDSV100 (Hughes, 1990; Hughes *et al.*, 1992). I cloned a tandem dimer of this genome by digesting pDSV100 with *Xmn*I and partially with *Xba*I. The appropriate linear fragments were purified and ligated to yield pDSV200. The insert from this clone was digested partially with *Hin*dIII and the 5.4 kb fragment cloned in the *Hin*dIII site of pBin19, generating pDSV200Bin.

pDSV-MSVBin

This plasmid contained infectious genomes of both DSV and MSV cloned on the same T-DNA. To construct this plasmid, I digested the pUC18-based pDSV200 with *Xmn*I and partially with *Hin*dIII and blunted the sticky ends with the Klenow fragment of DNA polymerase I, in the presence of dNTPs. I also digested pKom602 (the precursor pUC19-based plasmid of pKom603) with *Xba*I and *Xmn*I, and polished the sticky ends with Klenow, as before. The appropriate linear fragments were gel purified and ligated to create pDSV-MSV, where the insert present in the hybrid pUC18/pUC19 multiple cloning site was conveniently flanked by *Eco*RI sites. The 8.4-kb *Eco*RI fragment from pDSV-MSV was cloned into the *Eco*RI site of pBin19. Recombinant plasmids were screened for the

presence of the correct insert, and one with the DSV dimer proximal to the right border of the T-DNA was selected, and named pDSV-MSVBin.

4.2.2 Transformation of *Agrobacterium tumefaciens*

The nopaline strain of *Agrobacterium tumefaciens*, C58C1(pMP90) (Koncz and Schell, 1986), was used for agroinfections. This strain carries a cryptic plasmid (C1), a chromosomal Rif^R (rifampicin resistance) marker and a gentamycin resistance (Gm^R) gene cassette which was used to delete the wild type T-DNA of pTiC58 by homologous recombination in the generation of a disarmed Ti plasmid (pMP90). Transformation of *A. tumefaciens* C58C1 (pMP90) was by the freeze-thaw method of Holsters *et al.* (1978). Transformants were selected and maintained on LA plates containing rifampicin (100 μg/ml), gentamycin (40 μg/ml) and kanamycin (100 μg/ml).

4.2.3 Agroinfection

An overnight culture of *A. tumefaciens,* grown in LB media with appropriate antibiotic selection, was concentrated by centrifugation and resuspended in one fifth of the original volume of sterile distilled water. When two strains were used in co-agroinfection experiments, the resuspended cultures were mixed in a 1:1 ratio prior to inoculation. Maize kernels of the sweetcorn cv. "Jubilee" were germinated at 30° C in damp sterile vermiculite. Agroinfection of three day-old maize seedlings was by injection from the side at the coleoptilar node, with 2 to 3 μl of the *Agrobacterium* suspension loaded into a 10 μl Hamilton syringe, essentially as described by Escudero and Hohn (1994). Injected seedlings were transferred to soil and grown in a plant growth room, at about 25° C with a 16: 8 hour day/night cycle. Infection was scored from 5 days post-inoculation.

4.2.4 Transient replication assay

An assay for replication of recombinant virus DNA by microprojectile bombardment was done as previously described (Chapter 2, sections 2.2.3 and 2.2.4). BMS cells were bombarded with microprojectiles carrying pKEP152 or pKEP153.

4.2.5 DNA manipulations

Total plant DNA or DNA from bombarded callus was isolated for Southern hybridisation experiments as described previously (Chapters 2 and 3). For isolation of RF-DNA, I used the method of Palmer *et al.* (1997). Essentially, total nucleic acids were extracted in a standard plant DNA extraction buffer, and purified by phenol/chloroform extraction and alcohol precipitation. The viral RF-DNA was then purified by ion exchange chromatography on commercially available Qiagen plasmid purification columns.

DNA for use in PCR was isolated from individual *Cicadulina mbila* leafhoppers by the method of Zeidan and Czosnek (1991).

Probes homologous to the *bar* gene for Southern hybridisation were labelled by PCR as described in previous chapters. MSV-specific probes were labelled with dig-dUTP by PCR (Dig User's Guide to Filter Hybridization, Boehringer Mannheim) with primers MSV1770-1792 and MSV215-234, using PCR cycling conditions as described in Chapter 3.

4.2.6 Polymerase chain reaction for detection of viral and *bar* gene-homologous sequences in agroinfected plants and *Cicadulina mbila*

Approximately 10 ng of DNA from plants, or 1 µl of a total of 100 µl of nucleic acid extracts from individual insects was used in each PCR. For detection of MSV-Kom sequences in nucleic acid extracts from plants or *C. mbila*, PCR primers MSV1770-1792 and MSV215-234 were used, with reaction conditions as described in Chapter 3. For detection of DSV sequences in infected plants and *C. mbila*, the following PCR primers were used: MSV1770-1792 (E. P. Rybicki, unpublished; Chapter 3), which anneals to the DSV genome at the same position as in the MSV genome, and primer DSVrev (5′-CTCGGCGGGACCAAATTCAA-3′) which anneals specifically to DSV DNA in approximately the same region in the genome as MSV215-234 anneals in MSV. The PCR cycle conditions were: initial denaturation at 95° C for 2 minutes, followed by 30 cycles of 94° C for 45 seconds, annealing at 54 °C for 30 seconds and extension at 72° C for 75 seconds. The final PCR cycle had an extension time of 5 minutes. PCRs for detection of *bar* gene-homologous sequences in agroinfected plants were with primers BARP1 and BARP2, as described in Chapter 2.

4.2.7 Electron microscopy

BMS cells transfected with pKEP153 by microprojectile bombardment were scraped off filters with a spatula and placed in 1.5 ml microcentrifuge tubes. Approximately two volumes of 0.05 M sodium phosphate buffer, pH 7.0 were added, together with a small amount of carborundum. Cells were ground with an Eppendorf micro-pestle. Cellular debris was pelleted by centrifugation at 4°C. The supernatant was collected in a new microcentrifuge tube. Immune electron microscopy (IEM) (Milne and Leseman, 1984) with MSV antiserum was done by Mohamed Jaffer (Electron Microscope Unit, UCT).

4.2.8 Leafhopper transmissions

Maintenance of leafhopper colonies

Leafhoppers (*C. mbila* Naude) were reared on maize plants (cv. Witplat) in wooden cages covered with cotton gauze, at room temperature. Light was provided by natural sunlight. Viruliferous leafhoppers were kept in cages on trolleys with time-controlled artificial lighting (16 hours day/night) in isolated rooms. When necessary, leafhoppers were culled by spraying with insecticide (Kombat™, from Efekto).

Acquisition of virus by leafhoppers

Non-viruliferous leafhoppers, at all stages in the insect life cycle, were transferred on maize leaf material from the stock colonies into temporary "cages" made from gauze-covered hurricane lamp glasses. The leafhoppers were then placed on infected plant material in larger cages and allowed to acquire virus by feeding on the infected material for two to three days.

Leafhopper-mediated inoculation

After the two to three day acquisition period, uninfected sweetcorn cv. Jubilee seedlings (seven to ten days old) were placed into the cage with the viruliferous leafhoppers. Infected plants were cut down, so that leafhoppers would move from the wilted acquisition material to the uninfected sweetcorn seedlings. Viruliferous leafhoppers were allowed to feed on the uninfected sweetcorn for five to seven days, after which time the experiment was terminated by spraying the colony with insecticide and transferring the plants to a plant growth room.

4.3 RESULTS

4.3.1 Construction of *Agrobacterium* binary vectors for agroinfection with MSV-Kom derived constructs.

Diagrams of the various constructs used in agroinfection experiments to establish whether movement and encapsidation functions of MSV could be complemented in *trans* are shown in Figure 4.1. All constructs were flanked by *Xba*I and *Eco*RI restriction sites. The binary vector pBI121 was used as a convenient vector for cloning, as the *gus* ORF in this vector is flanked by *Xba*I and *Eco*RI, and so provided a useful segmented multiple cloning site. The *gus* ORF with *nos* terminator sequences was excised and replaced with the MSV construct of interest. When "trimeric" constructs were cloned the *bar* construct dimer was always cloned at the right border-proximal side of the T-DNA. I could thus be certain that the *bar* construct was transferred into infected maize cells **before** the MSV construct.

4.3.2 Transient replication assay and virus encapsidation in BMS cells

Plasmids containing a dimer of the CaMV35S-*bar* gene replacement of MSV-Kom (pKEP152), the *bar* dimer cloned in tandem with the wild type MSV-Kom genome (pKEP153) and the infectious 1.1-mer of the MSV-Kom genome (pKom602) were each introduced into three plates of BMS cells by microprojectile bombardment. Low molecular weight nucleic acids were isolated from two of the plates from each set, run on an agarose gel, blotted and hybridised with a probe homologous to the *bar* gene. As can be seen in Figure 4.2, the cells bombarded with pKEP152 contained replicating recombinant viral DNA, but no ssDNA forms were present, as one would expect for a coat protein mutant. In contrast, cells bombarded with pKEP153 clearly contained ssDNA forms homologous with the *bar* probe, indicating complementation of the ssDNA-negative phenotype of the coat protein deletion mutant. There were no *bar*-homologous DNA forms present in cells bombarded with pKom602. Extracts of cells bombarded with pKom602 and with pKEP153 were examined for the presence of geminate particles by immuno-electron microscopy (IEM) with an MSV-specific antiserum. There were low numbers of geminate particles present in these cells (Figure 4.2), showing that BMS cells transfected with cloned viral DNA by microprojectile bombardment can support the full MSV life cycle: from replication to formation of ssDNA and encapsidation. There were no geminate particles present in cells bombarded with pKEP152 or in unbombarded BMS (not shown).

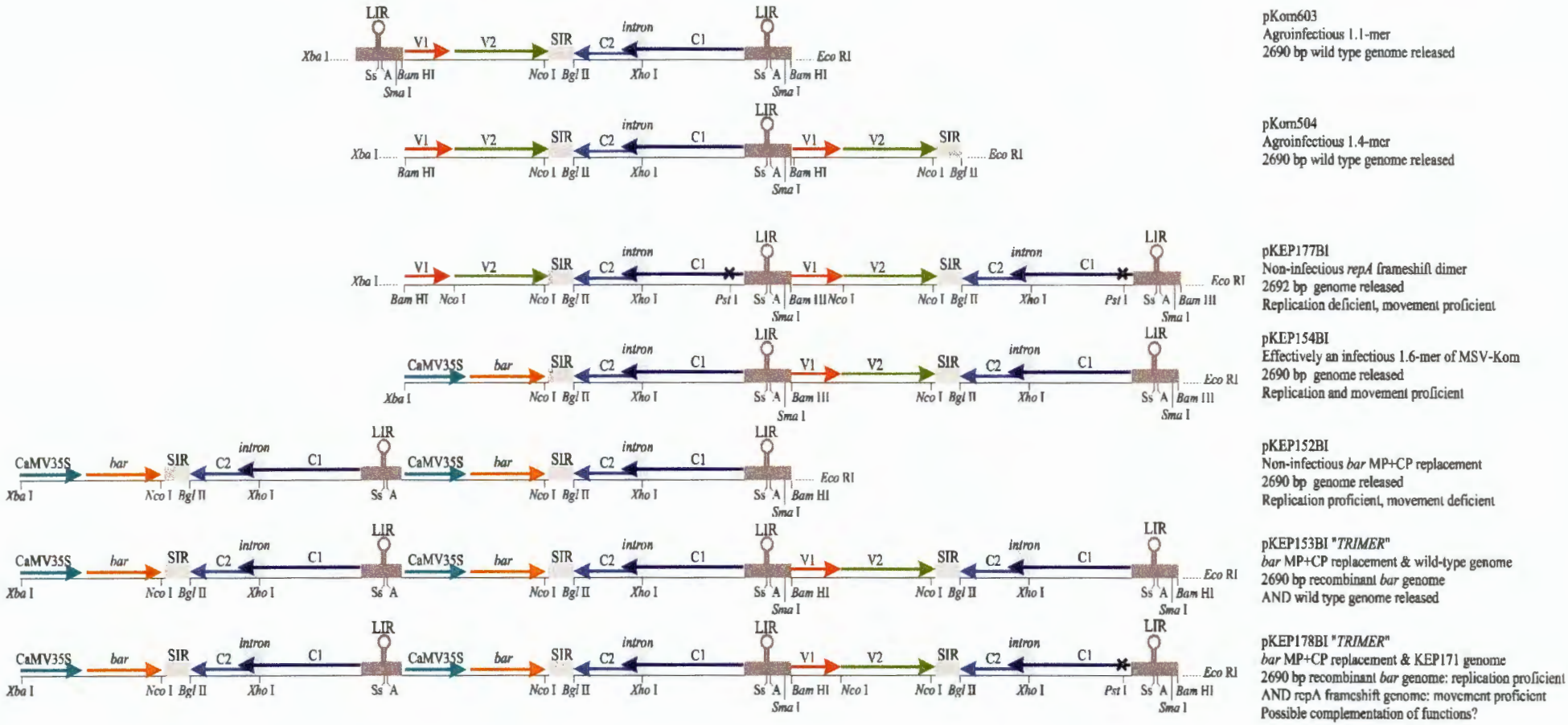


Figure 4.1: MSV constructs in binary vectors for agroinfection experiments.

The construction of recombinant viral multimers in the binary vector pBI121 is shown. In this vector, the T-DNA right border is on the same side as the *Xba*I restriction site, in all cases. The positions of new restriction sites for *Nco*I and *Pst*I are indicated in pKEP177BI and pKEP178BI. Abbreviations: A: *Apa*I; Ss: *Ssp*I.

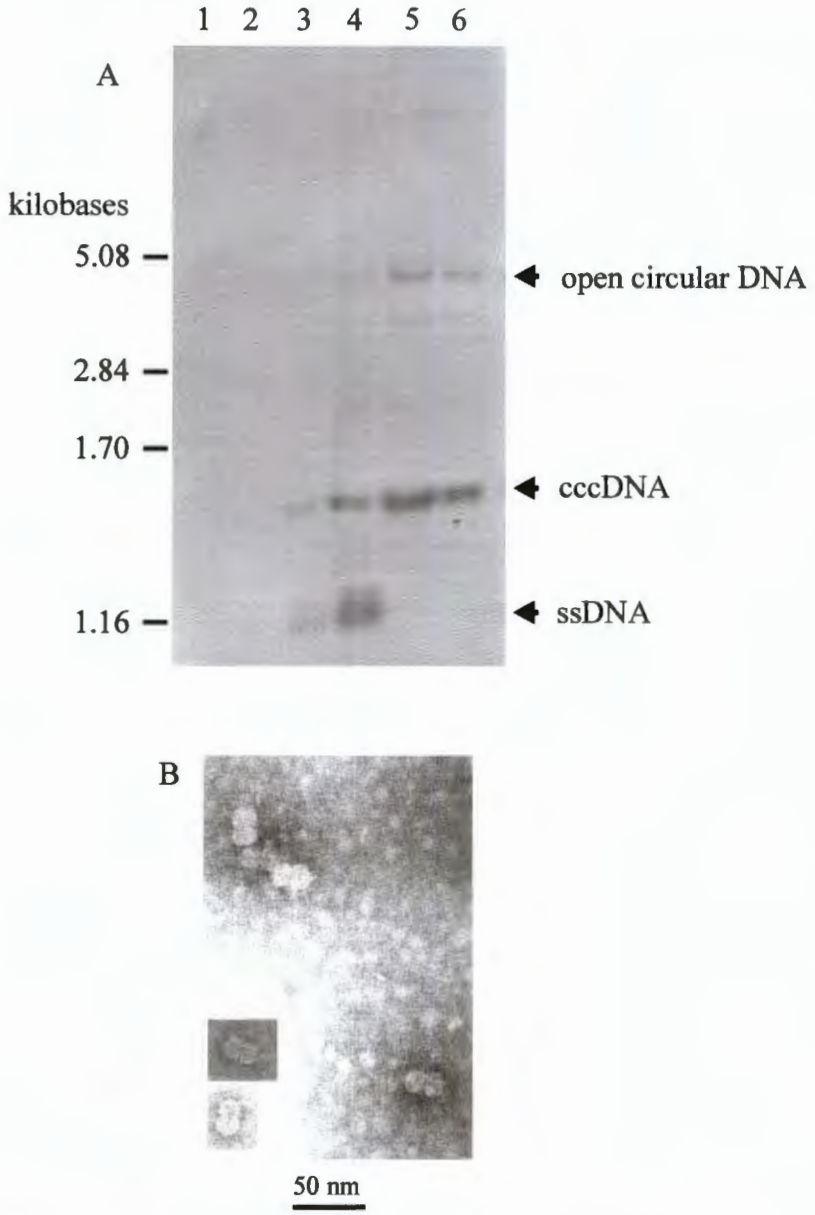


Figure 4.2: Complementation of ssDNA formation of a *bar* movement and coat protein replacement construct by wild type MSV-Kom, and formation of MSV virions in bombarded BMS cells.

BMS cells were bombarded with plasmid DNA containing MSV replicons. Low molecular weight DNA was isolated and run undigested on a 0.8% agarose gel and blotted onto nylon membrane. The Southern blot shown in (A) was probed with a digoxygenin-labelled *bar* probe. Each bombardment was done in triplicate. Two samples of each were run on the gel, and one was examined for the presence of MSV virions by immuno-electron microscopy (IEM) (B).

A: Lanes 1 & 2: BMS cells bombarded with pKom602 (infectious MSV 1.1-mer); lanes 3 & 4: BMS cells bombarded with pKEP153 ("trimer" with a dimer of a coat protein and movement protein replacement construct with CaMV 35S-*bar* cassette, and the wild type MSV genome); lanes 5 & 6: BMS cells bombarded with pKEP152 (dimer of CaMV35S-bar gene replacement of MSV-Kom movement and coat protein genes).

B: Electron micrograph of selected fields showing MSV particles from BMS cells bombarded with pKEP153. The scale bar represents 50 nm.

4.3.3 Agroinfection of sweetcorn cv. Jubilee

Sweetcorn seedlings with *A. tumefaciens* C58C1(pMP90) carrying the binary vectors illustrated in Figure 4.1 were scored for infection from 5 days after injection to 21 days after injection. Injection with *Agrobacterium* carrying any of pKom504, pKom603, pKEP153BI and pKEP178BI yielded plants presumably infected with MSV, as typical symptoms were produced. The symptoms in sweetcorn generated after agroinfection with these four strains were generally not distinguishable from each other; that is, all symptoms resembled wild-type MSV-Kom (Figure 4.3). Plants inoculated with pKEP152BI and pKEP177BI remained asymptomatic, and presumably non-infected, as expected. In three separate experiments a total of five out of 90 plants inoculated became symptomatic after injection with a mixture of *Agrobacterium* strains carrying pKEP152BI and pKEP177BI. Both constructs would have to have been delivered to the same cell in this case, as the *bar* construct in pKEP152BI did not contain the movement or coat protein genes to move and pKEP177BI contained a frameshift in the *rep* gene. Table 4.1 summarises the results of three separate agroinfection experiments with *Agrobacterium* strains carrying the seven binary vector constructs described in Figure 4.1.

Table 4.1: Summary of the results of agroinfection experiments

Binary vector in *A. tumefaciens*	**Experiment I** no. symptomatic/ no. surviving	**Experiment II** no. symptomatic/ no. surviving	**Experiment III** no. symptomatic no. surviving	**Total percentage symptomatic**
pKom504	17/28	8/12	N.D.[b]	63%
pKom603	13/13	10/10	N.D.[b]	100%
pKEP152BI	0/26	0/22	0/11	0
pKEP153BI	33/35	49/52	14/14	95%
pKEP154BI	21/24	12/12	N.D.[b]	92%
pKEP177BI	0/24	0/10	0/12	0
pKEP178BI	46/58	22/25	12/13	83%
pKEP152BI ***with*** pKEP177BI[a]	1/51	3/27	1/12	6%

[a] *Agrobacterium* cultures were mixed in a 1:1 ratio before injection.
[b] Only two repeats of this experiment were done.
N.D. = not determined

Figure 4.3: Typical MSV-Kom symptoms present on the third leaf of agroinfected sweetcorn cv. Jubilee.

Leaves were cut from 16 day-old seedlings which had been agroinfected with *Agrobacterium* carrying binary vectors with agroinfectious constructs. From top to bottom: pKom603, pKom504, pKEP177BI (non-infectious), pKEP154BI, pKEP153BI, pKEP178BI.

The time course of infection of plants injected with each of these agroinfectious constructs is presented in Figure 4.4. This shows that there was a clear difference in the kinetics of the rate of infection associated with injection of *Agrobacterium* strains carrying plasmids where the virus can escape by replicative release (pKom603; pKEP153BI and pKEP154BI) compared with where the virus can escape only by recombination (pKom504 and pKEP178BI). The main difference between infections with pKom504 and pKEP178BI was that injection with pKEP178BI usually resulted in a higher percentage of infected plants. The first symptoms on plants infected with pKom603, pKEP154BI and pKEP153BI were observed very occasionally on the first leaf of the germinated seedlings, but most agroinfected seedlings first showed symptoms on the second leaves, and all infected seedlings showed symptoms on their third leaves. In contrast, plants agroinfected with pKom504 and pKEP178BI never showed symptoms on the first leaf, and only 10 to 20% showed symptoms on the second leaf; symptoms were usually first seen on the third leaves of these plants, but occasionally a plant only showed symptoms starting on the fourth leaf. Of the five plants which showed symptoms after co-agroinfection with a mixture of pKEP152BI and pKEP177BI, one plant showed symptoms from its third leaf, and four plants showed symptoms first on their fourth leaves (results not shown).

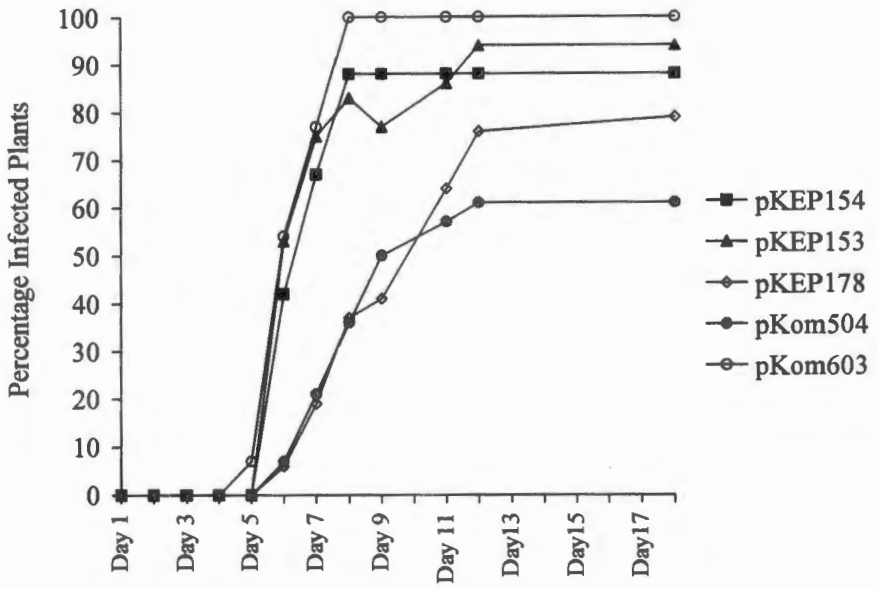


Figure 4.4: Infection kinetics after agroinfection of sweetcorn seedlings with five different agroinfectious constructs.

The rate of symptom development in sweetcorn seedlings agroinfected with *Agrobacterium* containing wild-type agroinfectious constructs pKom504 and pKom603, and constructs carrying MSV 35S*bar* coat protein and movement protein gene replacement constructs. Day 0 is taken to be the day of injection. The results shown are for one experiment series (the second column in Table 4.1), but other repeats of the same experiment showed virtually identical results. Data for agroinfection with pKEP152BI and pKEP177BI are not shown, as these constructs did not induce symptoms in agroinfected seedlings.

4.3.4 Detection of *bar* replicons in agroinfected sweetcorn by Southern hybridisation

Figure 4.5 shows a Southern blot of undigested total DNA which was isolated from the first three leaves of selected plants which had been agroinfected with pKEP152BI and pKEP178BI. The membrane was hybridised first with a *bar*-homologous probe to detect the MSV-CaMV35S*bar* construct and then stripped and hybridised with an MSV-specific probe which would detect both the *bar* gene replacement mutant and MSV sequences derived from pKEP178BI. Low molecular weight *bar*-homologous viral DNA forms could be seen in four of the six plants which had been agroinfected with pKEP152BI (Figure 4.5, plants 1, 2, 3 and 6). The recombinant viral DNA was present in the second leaves of all four plants, and was generally at its highest concentration in the first or second leaves of these plants. Only one plant (plant 1; Figure 4.5) had detectable *bar*-homologous DNA present in its third leaf. Of the six plants which had been agroinfected with pKEP178BI, DNA homologous to the *bar* gene was clearly present in the second leaf of only one plant, but DNA forms corresponding to replicating viral DNA were easily detectable with an MSV probe (Figure 4.5).

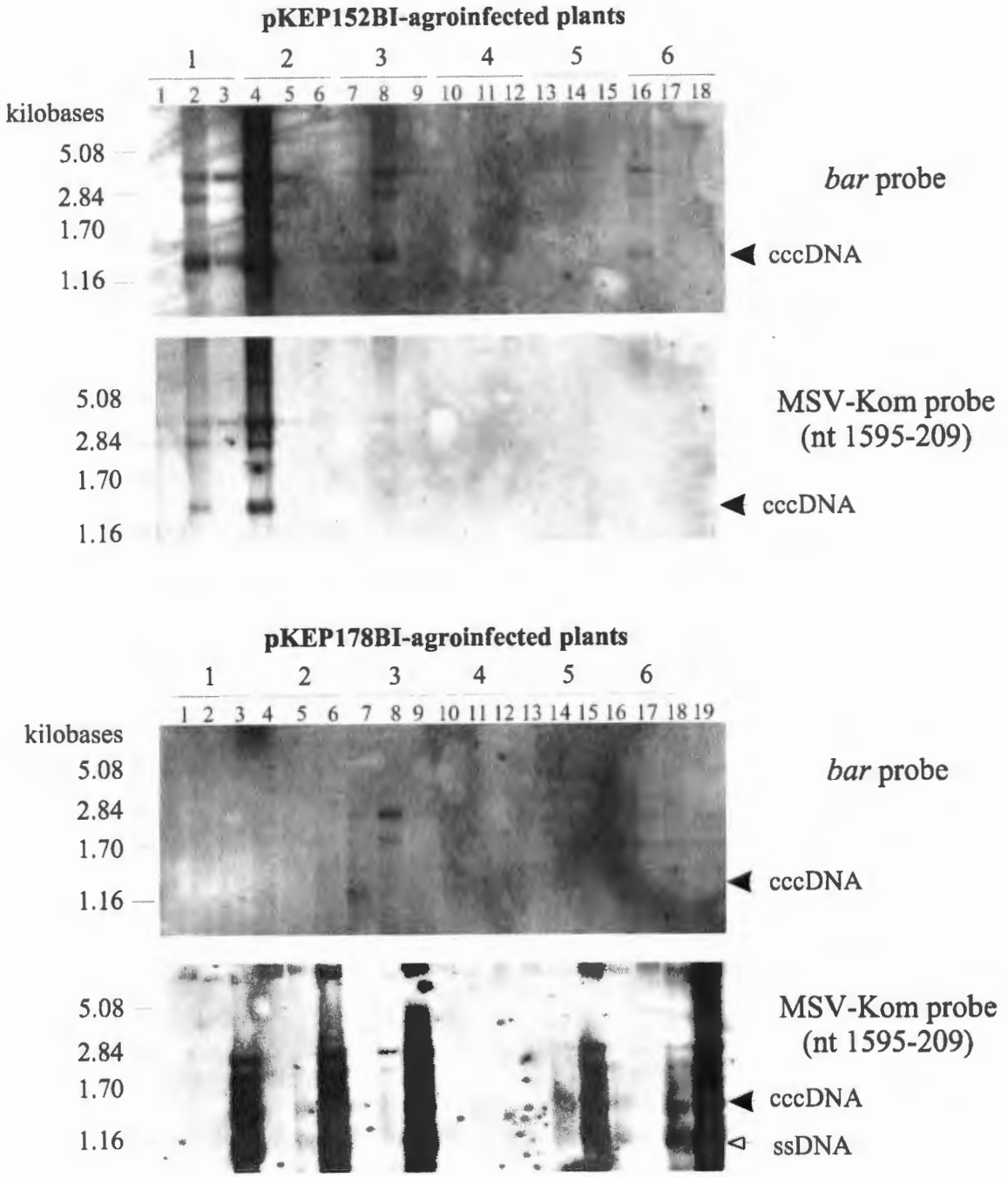


Figure 4.5: Southern blot analysis of the spread of *bar*-gene replacement constructs and MSV in plants agroinfected with pKEP152BI and pKEP178BI.

DNA isolated from the first three leaves of six randomly selected plants (indicated above the lane numbers in the Figure) agroinfected with pKEP152BI and pKEP178BI was electrophoresed in a 0.8% agarose gel, and transferred to nylon membranes. In both cases, the top blot was probed with a digoxygenin-labelled *bar* probe, then stripped and probed with an MSV-homologous probe (bottom). Plant number 4, which was injected with *Agrobacterium* carrying pKEP178BI did not show MSV symptoms. Lane 19 in the lower two blots contained DNA isolated from a plant agroinfected with pKom603 (wild type virus only).

4.3.5 Detection of recombinant virus constructs in agroinfected plants by PCR

I chose to use PCR to detect *bar*-homologous sequences in agroinfected plants for two main reasons: (1) PCR is more sensitive than Southern hybridisation, so this would allow me to detect replicating *bar* constructs when they were present at only low levels; and (2) I found that I had to use extremely high stringency washes after Southern hybridisation to distinguish *bar*-homologous sequences in replicons present at relatively low copy numbers relative to wild type virus, which was present at such high copy number in infected leaves that it was easily visible on the agarose gels blotted to nylon membranes. Even at high stringency, the long exposure times required to detect *bar* sequences usually also resulted in detection of low-level non-specific hybridisation to the much higher copy number MSV sequences (not shown), unless such extremely high stringency washes were done that the detection of the *bar* sequences became relatively insensitive (Figure 4.5). In Chapter 3, PCR with primers MSV1770-1792 and MSV215-234 proved to be a good method to detect the presence of replicating virus carrying the *Pst*I mutation present in pKEP177BI and pKEP178BI. These primers amplify a segment of DNA which spans the part of the RepB and the RepA ORFs, the LIR and part of the movement protein ORF. The *Pst*I mutation introduced into pKEP171 (Chapter 3) was near the start of the RepA ORF, so digestion of PCR products with *Pst*I would provide a sensitive and potentially quantitative assay for the presence of both the wild type and *Pst*-mutant viral genomes in infected plants.

Total DNA was isolated from the first five to seven leaves of plants agroinfected with pKEP152BI, pKEP153BI and pKEP178BI, as well as from plants which became infected after co-agroinfection with pKEP152BI and pKEP178BI. DNA isolated from the third leaves of plants infected with MSV-Kom by agroinfection with pKom603, and uninfected plants injected with pKEP177BI and pKEP152BI served as controls for the PCR. Figure 4.6 shows the results of PCR amplification of a 418-bp fragment of the *bar* gene from individual leaves of plants agroinfected with binary vectors carrying constructs from which an autonomously replicating gene vector exactly the same size as the wild type viral genome could be released. In the relevant samples, I have shown the results of PCR amplification of a 1304 bp fragment of the MSV genome from the same leaf samples of plants agroinfected with pKEP178BI or with pKEP152BI and pKEP177BI. These constructs both carried the mutant MSV genome with the frameshift mutation in the *repA* gene.

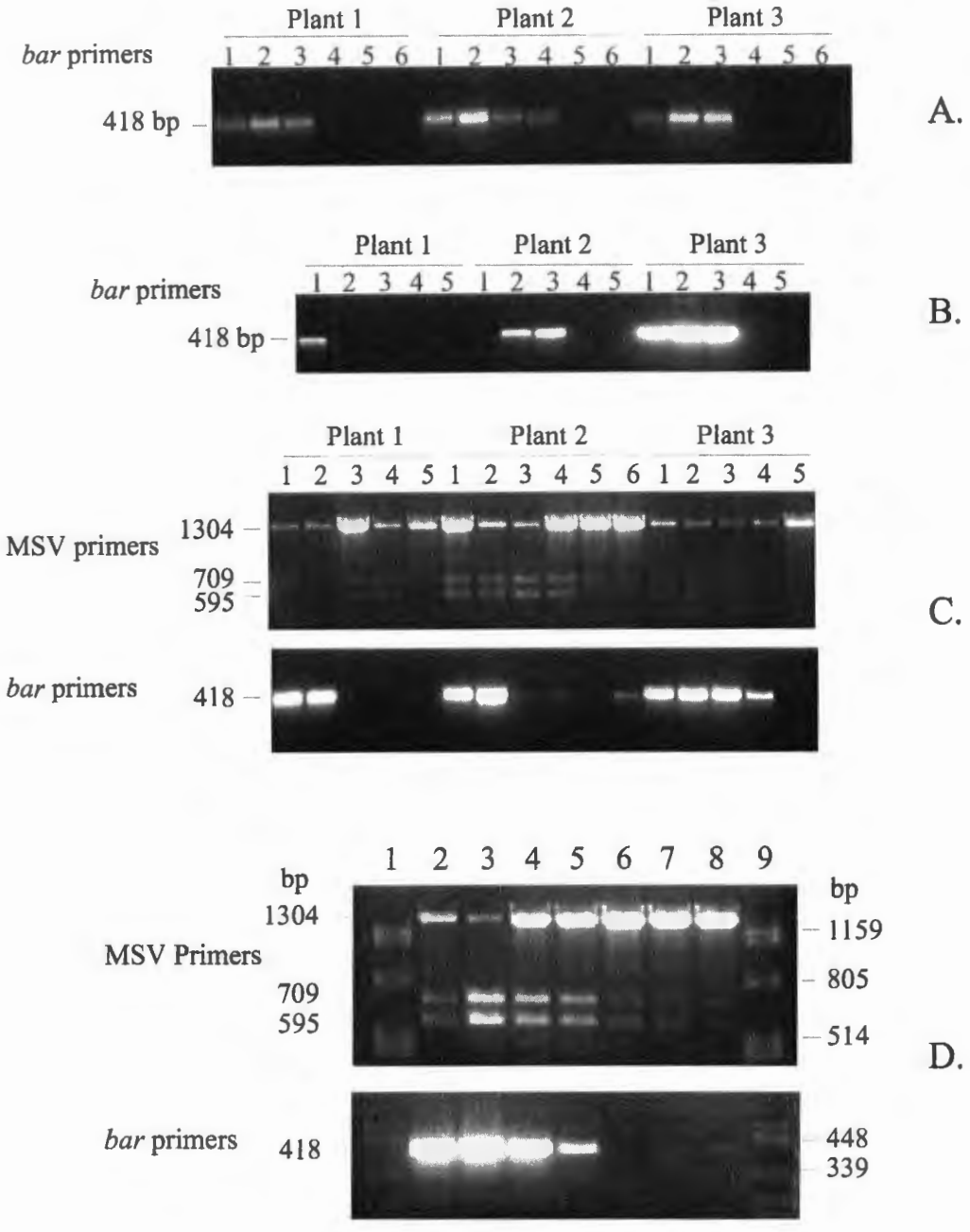


Figure 4.6: Spread of *bar* gene replacement construct and MSV *rep* frameshift mutants in agroinfected plants.

Gel A contains PCR products amplified with *bar* primers from six leaves from each of three plants agroinfected with pKEP152BI. Gel B contains PCR products amplified with *bar* primers from five leaves from each of three plants agroinfected with pKEP153BI. The gels indicated by "C" contain PCR products amplified from five or six leaves of each of three plants agroinfected with pKEP178BI; the upper gel contains products resulting from PCR with MSV-specific primers. The products were digested with *Pst*I to distinguish between non-mutant and *Pst*I mutant replicating virus. The lower gel contains *bar* PCR products amplified from the same DNA samples. The gels in "D" show PCR products from amplification with MSV primers and digested with *Pst*I (upper) and *bar* primers (lower). Lanes 1 & 9 contain molecular weight marker DNA (phage λ digested with *Pst*I), lanes 2 to 8 contain DNA isolated from each of the first seven individual leaves from one plant which was agroinfected with a mixture of pKEP152BI and pKEP178BI.

The results in Figure 4.6 show that the *bar* gene fragment was present in all but one of the plants agroinfected with *Agrobacterium* strains carrying derivatives of pKEP152. The *bar* sequence was present in the first two or three leaves, and occasionally in the fourth leaf of agroinfected plants. With one exception (plant 2, gel C), the *bar* construct did not seem to spread higher than the fourth leaf in plants agroinfected with pKEP153BI or pKEP178BI. In the one plant in which the *bar* construct moved higher than the fourth leaf, it was apparently present at only very low copy number.

In plants agroinfected with pKEP178BI, or with both pKEP177BI and pKEP152BI, the MSV construct carrying the *Pst*I-associated frameshift was always present, but the concentration of this replication-defective virus dropped steadily in higher leaves, often paralleling the decline in concentration of the *bar* construct. However, the *Pst*I mutant virus usually persisted longer than the autonomously replicating *bar* construct. In all samples, the presence of a full length 1304-bp PCR product in addition to those carrying the *Pst*I mutation was thought to represent a virus which had lost the *Pst*I mutation, presumably through recombination. In all cases, the PCR product generated from PCR amplification of the plasmid DNA controls: (1) pKEP153BI, did not digest to any extent with *Pst*I, and (2) pKEP177BI digested completely with *Pst*I (not shown).

4.3.6 Leafhopper transmissions from agroinfected plants

Three plants which had been agroinfected with pKEP153BI and three with pKEP178BI were used in leafhopper transmission experiments to determine whether the 2.69-kb gene replacement mutant virus construct, in the presence of wild type virus which may *trans*-encapsidate it, was transmissible by leafhoppers. Each plant, at approximately the five leaf stage, was placed in a separate cage with non-viruliferous leafhoppers. After three days of acquisition feeding, a sweetcorn seedling was added to each cage. All plants became infected, and symptoms appeared to be typical for MSV-Kom. DNA was isolated from the first infected leaf of each newly infected plant. PCR with *bar*-specific primers was negative, which proved that the *bar* construct was not transmitted by the leafhoppers (results not shown). The replication defective *Pst*I mutant was also not transmitted at any significant level by leafhoppers, as evidenced by the absence of *Pst*I sites in any of the amplification products from the three infected plants (Figure 4.7).

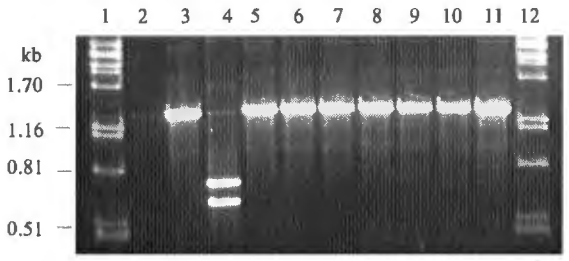


Figure 4.7: PCR analysis of MSV DNA amplified from plants infected by leafhopper transmission from plants agroinfected with pKEP153BI or pKEP178BI.

All PCR products were digested with *PstI*. Lanes 1 & 12: molecular weight marker (λ DNA digested with *Pst*I); lane 2: uninfected DNA control; lane 3: amplification product from 1 ng of pKEP153BI; lane 4: amplification product from 1 ng of pKEP177BI; lanes 5, 6 & 7: amplification products from plants infected by leafhopper transmission from pKEP153BI agroinfected plants; lanes 8, 9 & 10: amplification products from plants infected by leafhopper transmission from pKEP178BI agroinfected plants; lane 11: amplification product from plant agroinfected with pKom603.

Replicative form DNA isolated from plants infected by leafhopper transmission from the pKEP178BI agroinfected plants showed that the virus genome, although it had lost the *Pst*I site, still retained the *Nco*I site which was present over the coat protein start codon in the original construct, as RF-DNA isolated from plants infected by leafhopper transmission of virus from plants originally agroinfected with pKEP178BI showed two bands (692 bp and 1998 bp) on digestion with *Nco*I, but the wild-type virus was only linearised (Figure 4.8; lanes 2 & 3, respectively. See Figure 4.1 for the restriction map of pKEP178BI). This suggested that the serine to alanine substitution at the N-terminus of the coat protein did not interfere with viral encapsidation or with insect transmission.

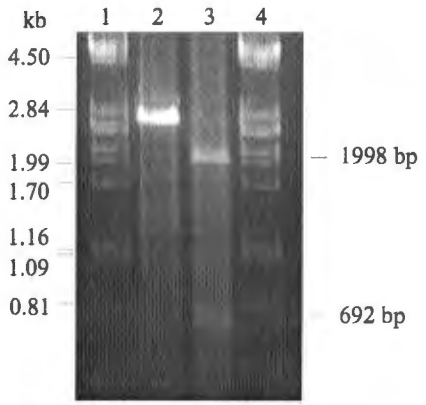


Figure 4.8: Gel showing the presence of the *Nco*I mutation in the genome of the virus transmitted by *C. mbila* to sweetcorn from plants agroinfected with pKEP178BI.

Lanes 1 & 4: molecular weight markers; lane 2: RF-DNA digested with *Nco*I, isolated from a plant infected by leafhopper transmission with virus originally from a plant agroinfected with pKEP153BI. Lane 3: RF-DNA digested with *Nco*I, isolated from a plant infected with virus transmitted from a plant agroinfected with pKEP178BI.

4.3.7 Construction of infectious DSV-containing plasmids

In experiments aimed at establishing whether MSV-Kom could *trans*-encapsidate another viable virus, I cloned a tandem dimer of the distantly related DSV in the *Agrobacterium* binary vector pBin19 (pDSV200Bin; Figure 4.9). I also constructed a binary vector which contained both the infectious clone of MSV-Kom, pKom602 (W.H. Schnippenkoetter and E. P. Rybicki, unpublished) and my dimeric clone of DSV, pDSV200. This construct (pDSV-MSVBin) contained the DSV dimer proximal to the right border of the T-DNA in pBin19 (Figure 4.9), thus ensuring that when MSV symptoms were observed in agroinfected plants, the DSV dimer was already guaranteed to be present in the infected cells.

4.3.8 Agroinfection of sweetcorn cv. Jubilee with DSV-containing constructs

Three day old seedlings of the sweetcorn cv. Jubilee were agroinfected with pDSV200Bin and pDSV-MSVBin. Out of 25 seedlings which survived injection with pDSV200Bin, 10 (40%) showed definite symptoms. Agroinfection with pDSV-MSVBin resulted in 17 out of 21 seedlings (81%) exhibiting symptoms. Figure 4.10 shows a photograph of leaves taken from plants which were agroinfected with pDSV200Bin and pDSV-MSVBin. The symptoms caused by the agroinfectious DSV constructs were extremely mild on sweetcorn, consisting of very few and widely dispersed streaks, in contrast with those typically associated with MSV-Kom infection, visible on leaves of plants agroinfected with pDSV-MSVBin (Figure 4.10).

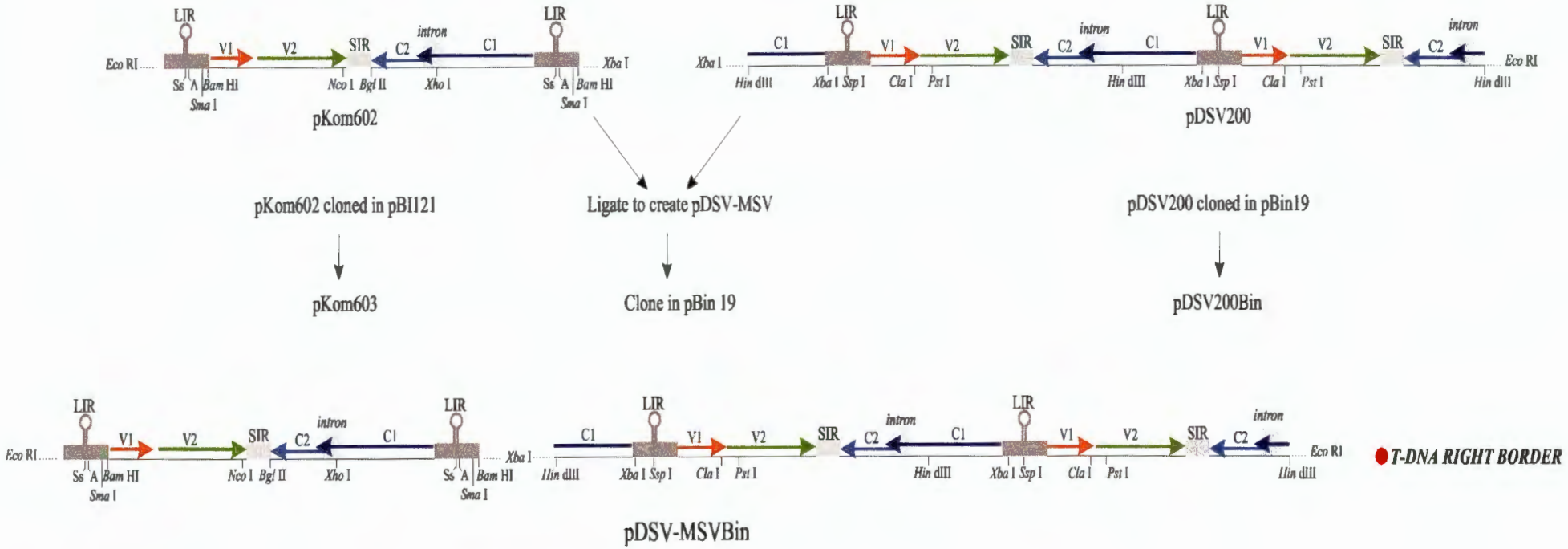


Figure 4.9: Construction of infectious DSV-containing plasmids

The precursor plasmid to pKom603, cloned in pUC19, was linked to a tandem dimer of DSV, pDSV200, in pUC18. The plasmid containing the DSV dimer linked to the infectious MSV-Kom construct was called pDSV-MSV, where the insert was flanked by *Eco*RI sites. The insert in pDSV-MSV was excised as an *Eco*RI fragment and inserted into pBin19. A plasmid with the DSV genome on the same side of the T-DNA as the right border sequence was selected and named pDSV-MSVBin. Abbreviations: Ss: *Ssp*I; A: *Apa*I; LIR: long intergenic region; SIR: short intergenic region.

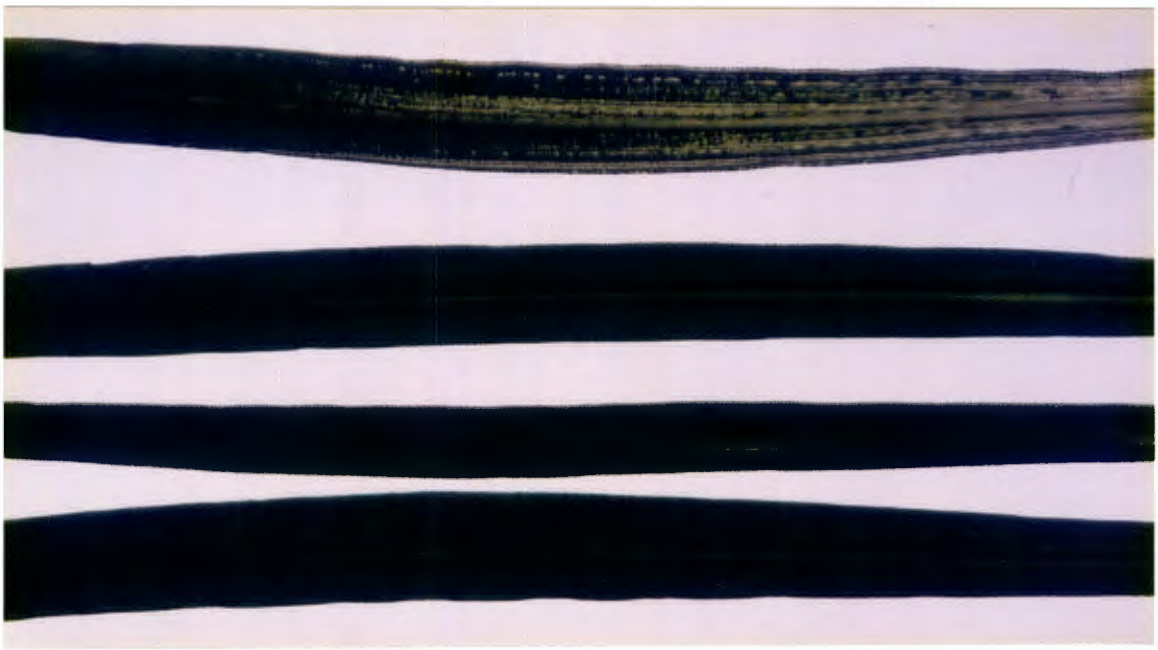


Figure 4.10: Symptoms on leaves of plants agroinfected with binary vectors pDSV200Bin and pDSV-MSVBin.

The top leaf is from a plant agroinfected with pDSV-MSVBin, and shows typical MSV-Kom symptoms. The second and third leaves are from plants agroinfected with pDSV200Bin, and show very mild streak symptoms. The bottom leaf is from a mock-infected plant.

4.3.9 Transmission of DSV with *C. mbila* from a plant infected with both DSV and MSV-Kom

In experiments designed to test whether the MSV coat protein will specifically encapsidate only the MSV genome, I agroinfected plants with both DSV and MSV, and used *C. mbila* to transmit virus from one of these plants to sweetcorn seedlings. Under normal conditions, *C. mbila* can transmit MSV, and not DSV (E. P. Rybicki, personal communication; Julia and Dollet, 1989). The results of PCR amplification of DSV DNA from one of three *C. mbila* which had been feeding on DSV-infected plant material (lane 2) and from two or three of the four leafhoppers (lanes 7, 8 & 9) which had been feeding on the doubly infected plant, showed that *C. mbila* can probably take up DSV virions, but cannot transmit the virus. However, the results shown in Figure 4.11 show that DSV was transmitted from a plant doubly infected with MSV and DSV (lane 14) to each of four individual plants (lanes 10 to 13). Variation in the intensity of amplification may have reflected differing concentrations of DSV in plants.

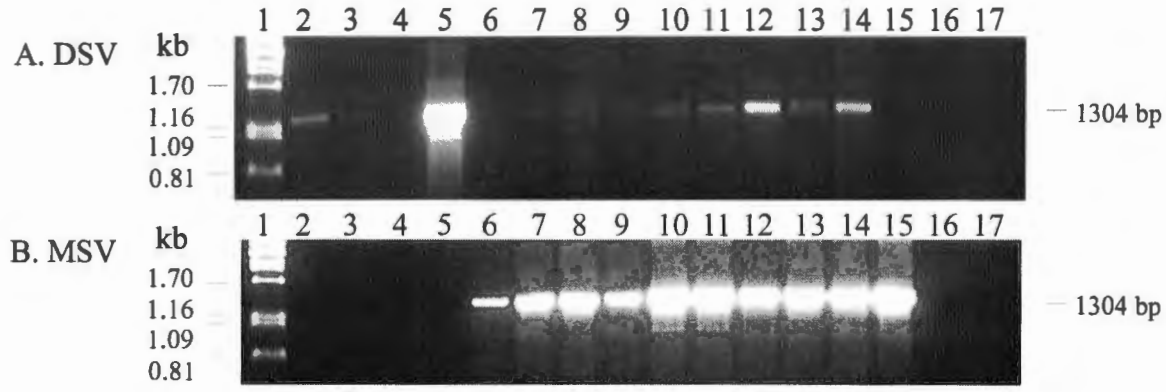


Figure 4.11: Detection of DSV and MSV-Kom sequences in plants agroinfected with DSV-containing constructs, in plants infected by leafhopper transmission from a plant doubly infected with DSV and MSV-Kom and in *C. mbila* which had been feeding on agroinfected plants.

The same DNA extracts were amplified in separate PCRs. Gel A shows the results of PCR with primers MSV 1770-1792 (general Subgroup I virus primer) and DSVrev (specific for DSV). Gel B shows the results of PCR with primers MSV 1770-1792 (general primer) and MSV 213-234 (MSV-specific primer). Lane 1: molecular weight marker (λ DNA digested with *PstI*); lanes 2, 3 & 4: DNA amplified from *C. mbila* which had been feeding on plant A. agroinfected with pDSV200BI; lane 5: acquisition plant A. - infected with DSV by agroinfection with pDSV200Bin; lanes 6 to 9: DNA amplified from *C. mbila* which had been feeding on plant B agroinfected with pDSV-MSVBin; lanes 10 to 13: DNA amplified from the first infected leaf of plants which had been infected by leafhopper transmission from plant B.; lane 14: DNA amplified from acquisition plant B.; lane 15: DNA amplified from a plant agroinfected with pKom602; lane 17: negative control - PCR amplification from a plant exposed to *C. mbila* which had been feeding on plant A.

4.4 DISCUSSION

The potential uses of Subgroup III geminivirus gene vectors which are capable of systemic movement in infected plants are many, but somewhat limited by the small size of the gene which can be inserted in the place of the ostensibly dispensible coat protein ORF. Nevertheless, the system developed by Hayes *et al.* (1988b; 1989) and Ward *et al.* (1988) presents attractive opportunities for rapid and facile testing of genes which encode proteins up to approximately the size of the coat protein (23 kDa), which may be expressed at high levels in infected plants. Several studies by Shen and Hohn (1991; 1992; 1994; 1995) showed that MSV constructs which contained insertions within the SIR, but which otherwise remained intact and replication proficient, still could not move systemically in agroinfected plants. This implied that there was a size limitation on movement of the viral genome, with constructs much larger than the wild type genome being unable to move. Given that the virion sense genes of Subgroup I geminiviruses are both essential for infectivity (Boulton *et al.,* 1989b; Lazarowitz *et al.*, 1989; Woolston *et al.*, 1989; Shen and Hohn, 1994; 1995), the only option remaining for using these viruses as infectious gene vectors is to complement movement and/or replication functions in *trans*. This could either be from transgenes expressed from the plant genome, or from a co-infected viral construct. However, achieving transgenic complementation is likely to be difficult, as several authors have found that transgenic plants expressing geminiviral proteins do not complement mutant virus genomes, and in the cases of ACMV Rep and TYLCV-Is coat protein, can actually confer good resistance to viral infection (Hong and Stanley, 1996; Kunik *et al.*, 1994). The investigations described in this Chapter were therefore aimed primarily at establishing whether wild-type or mutant MSV could complement the movement-deficient phenotype of an MSV gene replacement construct, exactly the same size as the wild type viral genome. This recombinant virus contained a CaMV35S promoter-*bar* expression cassette in place of the virion sense ORFs, which was chosen primarily for convenience of size.

The requirement of coat protein for movement of monopartite geminiviruses remains largely uninvestigated, but implies that these viruses either move encapsidated, or as a coat protein-DNA complex. It is well established that the formation of ssDNA is linked to a major extent to the expression of the coat protein gene, in all three genera of the *Geminiviridae*. Subgroup I viruses with null mutations or deletions of the coat protein

gene do not accumulate detectable levels of ssDNA (Boulton *et al.*, 1989b; 1993 Lazarowitz *et al.*, 1989; Woolston *et al.*, 1989; this study). The results shown in Figure 4.2 suggested that the ssDNA negative phenotype of coat protein replacement mutants of MSV can be complemented in *trans* by wild type virus replicating in the same cell. The formation of *bar*-homologous ssDNA could either have represented sequestration of the recombinant ssDNA genome by encapsidation, or the manifestation of a coat protein-induced switch from accumulation of predominantly dsDNA forms of the viral genome, to production of ssDNA forms.

To further address whether there is a specific encapsidation signal present on the MSV genome which may limit the movement of recombinant viruses if it is deleted, I investigated whether the MSV coat protein could *trans*-encapsidate the genome of the distantly related DSV if both viruses were present in the same cells. I showed that DSV could be detected in plants which were infected by leafhopper transmission from a doubly infected plant, which implies that the MSV-Kom coat protein encapsidated the DSV genome, since DSV is not usually transmissible by *C. mbila*, and further implies that there is no specific encapsidation signal sequence present on Subgroup I viral genomes, unless it is conserved between these two viral species. By analogy with Subgroup III geminiviruses, it is certain that the viral coat protein, and therefore encapsidation of the viral genome, is required for insect transmission (Azzam *et al.*, 1994). The results presented in this Chapter concur with those of Briddon *et al.* (1990), who found that the coat protein of a monopartite Subgroup II geminivirus, BCTV, could encapsidate both genome components of ACMV, a bipartite Subgroup III geminivirus. Sweetcorn is not, however, a good host for DSV, and the DSV DNA could only be detected by PCR. Boulton (1991) attempted a similar experiment and found no transmission of DSV by *C. mbila* in the presence of MSV-N by a Southern blot assay. However, the plant used in her insect transmission study had been co-infected with separate *Agrobacterium* cultures containing DSV and MSV, and it is not certain that both viral genomes were present in the same cells. This experiment may best be repeated with a virus transmissible by *C. mbila* which does not cause such severe symptoms in maize, such as Panicum streak virus (PanSV). In addition, rather than maize, a more permissive host for DSV, such as *Digitaria,* should be used as the plant to which both PanSV and DSV are transmitted. Leafhopper transmission to *Digitaria* would not have been useful for the experiments reported here, as MSV-Kom also causes extremely severe disease in this host (personal observation). Alternatively, more dramatic

evidence for *trans*-encapsidation would have been obtained if the DSV vector *Nesoclutha declivata* had been available to transmit the severe virus *trans*-encapsidated by DSV, to a susceptible host such as maize or *Digitaria.*

There have been no previous reports of natural *trans*-encapsidation of viral genomes of two different geminivirus species in infected plants. If *trans*-encapsidation happens in nature, for example between leafhopper transmitted Subgroup I or II geminiviruses and whitefly transmitted Subgroup III geminiviruses, the process may play an important role in the evolution of new geminivirus species, as it may help the virus to expand its host range to species which are not normally favoured by their primary insect vector. Recombination is likely to play an important role in evolution of geminiviruses: in fact, Subgroup II viruses are probably the result of recombination between a whitefly-transmitted geminivirus and a leafhopper-transmitted virus (Rybicki, 1994). This event would have to have taken place in plants where there were mixed infections, and it is possible that *trans*-encapsidation of viral genomes played a role in the evolution process. This would have ensured the persistence of the mixed infection until a viable recombinant evolved, since both whiteflies and leafhoppers might have been able to transmit both virus species, albeit at low efficiency.

Given the evidence that coat protein complemented the formation of ssDNA in the *bar* gene replacement mutant of MSV, and that MSV can probably *trans*-encapsidate a distantly related virus, it becomes difficult to explain why the *bar* gene replacement mutant was not moved to any significant degree by the wild type virus. The strategy I used in cloning the recombinant virus adjacent to the right border of the T-DNA ensured that the *bar* construct was present in infected cells **before** the wild type or *Pst*I frameshift mutant viruses: that is, that every cell which became infected with MSV had to have the *bar* construct present in addition. The agroinfectious constructs contained the dimeric recombinant *bar* construct cloned in tandem with a single copy of the wild type genome (pKEP153BI) or the *Pst*I frameshift mutant (pKEP178BI; Figure 4.1). It is well established that, in the presence of two origins of replication, replicative release of a monomeric circular copy of the virus construct is far more efficient than release by homologous recombination (Stenger *et al.*, 1991; Heyraud *et al.*, 1993b). From a replicative release perspective, the constructs in pKEP153BI and pKEP178BI were effectively "trimers", with two of the genomes flanked by LIRs. Thus on infection of cells, both the *bar* replacement

construct and the wild type (or *PstI* mutant) genome should escape by replicative release. However, the graph of the rate of infection with pKEP178BI (Figure 4.4) closely resembled infection kinetics typically associated with agroinfection with pKom504, a construct from which the wild type virus can only escape by homologous recombination, and shows a slower rate of symptom development than in plants infected with pKEP153BI, where both virus genomes escape by replicative release. Thus, it can be assumed that either: (1) complementation of functions between the movement-deficient *bar* construct and replication-deficient *PstI* mutant did not occur, and that release of an infection competent virus only arose by homologous recombination in the original T-DNA; or (2) that replicative release of both virus constructs occurred and that homologous recombination between the replicating *bar* construct and the *trans*-replicated *PstI* construct occurred subsequently.

I would favour the second possibility because replicative release is a more efficient process, and the *PstI* frameshift mutant was usually present in infected plants where it spread and was *trans*-replicated by the replication-competent virus. In addition, the proportion of plants infected with MSV after inoculation with *A. tumefaciens* C58C1 (pMP90)(pKEP178BI) was higher than usually obtained with constructs where the virus can only escape by homologous recombination (W.H. Schnippenkoetter and D. P. Martin, personal communication).

These results suggest that, irrespective of whether the replication-competent virus arose from homologous recombination in the original T-DNA or between the *PstI* mutant and the *bar* mutant, the formation of a virus competent for both replication and movement is a prerequisite for symptomatic infection. Boulton *et al.* (1989b) found that they could achieve infection by complementation of virus movement functions in *trans*, but in the context of genomes which were otherwise intact. However, these authors also found that wild type virus eventually arose by homologous recombination between complementing mutants, in every case. In the light of the results presented here, it is not unreasonable to propose that genome size and formation of ssDNA are not the sole factors which determine whether a virus genome is capable of systemic movement with coat protein and movement protein provided in *trans*. It is possible that formation of ssDNA and encapsidation can be uncoupled, and that sequences essential for encapsidation, or for interaction of the recombinant viral genome with a coat protein-movement protein complex, were deleted. It

also remains possible that geminivirus encapsidation signals are not specific sequences, but rather that the secondary or tertiary structure of the virus ssDNA is important. If this were the case, these structures must be conserved between DSV and MSV.

The fact that the virus carrying the *PstI* frameshift persisted in some plants shows that replication functions can be *trans*-complemented in MSV infections. There are no published reports of subgenomic or replication-deficient virus genomes in MSV, so these results provide the first evidence for *trans* replication of MSV mutants *in planta*. Plants carrying the replicating virus mutant did not show any associated symptom amelioration, so the mutant did not function as a DI-DNA, at least in sweetcorn which is a highly susceptible host. I hope that in the future there will be investigations into whether subgenomic virus constructs with deletions in the viral genome function as DI-DNAs and further whether these can be encapsidated. My results on leafhopper transmission from plants infected with pKEP178BI showed that the replication-deficient virus was not transmitted to any significant level. This probably reflects the fact that this virus was the same size as the wild type virus and so had no selective advantage on transmission to uninfected plants by leafhoppers, and so was not maintained to any significant level.

CHAPTER 5

CONCLUSIONS AND GENERAL DISCUSSION

The broad aim of the research presented in this thesis was to examine the potential of MSV as a gene vector. In this concluding chapter, I present my evaluation of the utility that this virus might have in plant molecular biology and biotechnology, within the framework of some of the general themes identified in Chapter 1.

MSV as a transient gene expression vector

Both Subgroup I and III geminiviruses have shown good potential as transient expression vectors in transfected cells or isolated plant tissues, but as yet have not been widely exploited in plant molecular biology. My investigation of the effect of genome amplification on the expression of the MSV coat protein promoter showed that replication of the virus in BMS cells enhanced coat protein promoter expression 45-fold over the level seen from a non-replicating vector. These results concur with, and expand upon those recently reported by Suárez-López and Gutiérrez (1997), in that these authors did not quantify the effect that replication had on the expression of the WDV coat protein promoter. The level of enhancement of coat protein expression in MSV was significantly higher than the 20-fold observed by Matzeit *et al.* (1991), but may reflect differences in our experimental protocols. My investigations used intact cells, transfected by particle bombardment, rather than protoplasts. In addition, in my virus constructs, the reporter gene (*gus*) was cloned as an exact transcriptional and translational fusion with the coat protein promoter, rather than a simple transcriptional fusion.

There is no doubt that there are upstream sequences which are important in controlling the translation of the coat protein, given that it is apparently expressed off a bicistronic mRNA. It will be of great interest to investigate what effect these sequences have on the expression of the GUS gene in my expression constructs in the future. As yet, Subgroup I geminivirus replicons have not been used in transient analyses of the functions of other promoters, but

in the light of the improvement in gene expression levels which are associated with replication of the viral reporter constructs, I predict that these could be of considerable use in plant molecular investigations. The use of MSV to enhance transient gene expression in tissues transfected by particle bombardment certainly need not be limited to tissue cultured cells.

MSV as an infectious gene expression vector in plants

In the late 1980s, two groups (Hayes *et al.*, 1988b; 1989 and Ward *et al.*, 1988) reported exciting data which suggested that bipartite Subgroup III geminiviruses may be useful for expression of small genes in infected plants. However, there have been no subsequent reports in the literature on the use of these viruses for high-level infectious gene expression. The only Subgroup I geminivirus which has been investigated as an infectious gene expression system is MSV, by Shen and Hohn (1994; 1995). These authors showed that intact viral genomes bearing reporter genes inserted in the SIR, and gene replacement constructs with the virion sense ORFs deleted, could express foreign genes at enhanced levels in agroinfected seedlings, but only in the first two or three leaves. This was interpreted as an indication that exceeding the maximum size requirements for encapsidation eliminated systemic movement of the recombinant viruses. My investigations, reported in Chapter 4, expanded on this research.

I attempted to refine the agroinfection system developed by Shen and Hohn (1992; 1994; 1995) for using MSV as a gene vector in plants, by decreasing the size of the recombinant virus construct to exactly the same as the wild type virus. I then investigated whether the wild type virus, and a mutant virus which would rely on the recombinant virus for its replication functions, could complement the movement and encapsidation functions which had been deleted in the vector. While wild type MSV could induce ssDNA formation of the gene replacement mutant, neither wild type nor replication deficient MSV could complement its movement functions. This implied that the gene replacement mutant was lacking sequences important in *cis* for encapsidation and/or systemic movement. The issue of an encapsidation signal was further investigated by co-infecting maize plants with DSV and MSV. Ordinarily DSV is not transmitted by the main insect vector of MSV, *Cicadulina mbila.* However, these leafhoppers were apparently able to transmit DSV from doubly infected plants, which suggests that there is no specific sequence necessary for virus

encapsidation. This unfortunately disproved my supposition that deletion of an encapsidation signal could account for lack of virus movement, but it remains possible that there is an encapsidation signal which is conserved between MSV and DSV, or that there is a specific "movement signal" which was lacking in the gene replacement construct. The molecular mechanisms involved in movement of Subgroup I geminiviruses, and encapsidation of geminiviruses in general, are very poorly investigated. Certainly a great deal more research needs to be done on geminivirus movement before one could propose ways to "force" MSV to function as an infectious gene vector.

A major issue in the development of a "binary" Subgroup I geminivirus infectious gene expression system is homologous recombination between the two components. The results I presented in Chapter 4 show that this is likely to be a major limitation on the successful application of MSV-based vectors. One possible solution would be to develop a binary system based on two distantly related viruses which might provide complementary functions, and show less propensity to recombine. However, Margaret Boulton (1991) found that DSV could not complement MSV movement functions. Ongoing research in the MSV laboratory at UCT suggests that even fairly closely related strains of MSV do not efficiently complement each other's movement functions, and that there may even be some as yet undefined mechanism for specific recognition of the movement and coat proteins' cognate genome sequences (E. van der Walt, K.E. Palmer and E. P. Rybicki, unpublished). Thus, I would predict that MSV has limited application as an infectious gene expression vector, which extends only to transient expression of foreign genes in the first few leaves of inoculated seedlings.

MSV as an episomal gene amplification vector in transgenic plant cells

This topic formed the bulk of the research presented in this thesis. I was able to show that MSV-based vectors could be used to select transgenic cell lines, and that the episomal sequences were maintained at high copy numbers for long periods of time in tissue cultured cells. I also showed, for the first time, that transformed cell lines could be recovered with two separate geminivirus vectors. However, I was not able to regenerate transgenic plants containing replicating viral constructs. In the light of recent investigations that show that geminiviral Rep proteins are able to modify the host cell cycle, it is likely that expression of Rep will be a complicating factor in regeneration of transgenic plants with viral

episomes replicating constitutively. Despite the early promise shown, Subgroup III geminivirus-based episomal gene amplification systems (Hayes *et al.*, 1988b; 1989; Kanevski *et al.*, 1992; Meyer *et al.*, 1992) have yet to be shown to have practical application in expression of proteins of more commercial value than NptII and GUS. These authors also found fairly low copy numbers for the viral vectors in transgenic plants. I would predict that this reflects incompatibility between high-level expression of Rep and regeneration of plants.

The differences observed between expression levels of PAT protein from the *bar* gene in replicating episomes in transformed cell lines (Chapter 2), and the subsequent demonstration of high maximum levels of GUS expression (Chapter 3) reinforce the idea that every protein has its own optimal expression system, and that some proteins can be expressed at higher levels in plants than others. It is important to note that I also found that the expression of GUS varied quite considerably between different transgenic lines containing GUS replicons. Thus, the proposal of using geminiviral episomes to alleviate position-effects might not find solution here. It is likely that variability in the frequency of episome mobilisation, even within the same transgenic line, accounts for temporal and within-tissue variation in marker expression. Nevertheless, the results presented in Chapter 3 showed conclusively that linkage to an MSV replicon can indeed enhance transgene expression significantly: up to 90-fold in certain transformed lines. These data suggest that Subgroup I viral "amplicons" could have a multitude of uses for expressing high levels of proteins in plants, if there is a way to avoid toxicity of Rep protein. There is currently great interest in using transgenic plants to express proteins which have pharmaceutical value (reviewed by May *et al.* 1996 and Mason and Arntzen, 1995). Given that the main consumers of maize are livestock, it would seem that an MSV-based transgene expression-enhancement system may be useful in the veterinary pharmaceutical industry, where the plant material would be consumed "raw". Such a system could also be useful for expression of industrial enzymes which are needed in large quantities, and do not need sophisticated purification protocols.

I envisage that an ideal system for using MSV to improve foreign gene expression in transgenic cereals would involve expression of Rep under the control of an inducible or tissue specific promoter. Here, a multimer of an amplification construct, which need only contain the viral origins of replication, would be integrated into the plant chromosome.

The Rep protein required for mobilisation and amplification of the construct would then only be expressed from a chimaeric gene, under the control of a tightly regulated tissue-specific or chemically inducible promoter. Thus, high level expression of the foreign gene would be achieved under controlled circumstances, and expression of Rep would not complicate regeneration, or select for low frequency mobilisation of the episome. My results in Chapter 3 suggested that levels of Rep expression may be one of the factors that limit episome copy number, since supplementing Rep from an expression cassette further enhanced the high levels of expression from a replicating construct.

MSV genetic elements for expression of foreign genes in transgenic plants

At first glance, it is difficult to see what elements of the MSV genome might be useful for controlling the expression of genes in transgenic plants: according to the literature, and my work, the coat protein promoter is relatively weak, and requires genome amplification for high-level expression, and the *rep* promoter produces transcripts which are barely detectable in infected plants (Morris-Krsinich *et al.*, 1985). However, it is myopic to discount the potential for use of ostensibly weak promoters which might have interesting tissue specificities in transgenic systems. As I pointed out in Chapter 1, it is possible that the *rep* promoter is considerably stronger in specific tissues: the analysis of the transcript levels in infected plants was from older, expanded leaf tissues (Morris-Krsinich *et al*, 1985); the promoter may well be stronger in younger tissues where viral replication is highly active. This issue certainly warrants further investigation in transgenic plants. Other elements of the viral genome which may be useful in plant molecular biology include the V1 intron, which may be implicated in intron-enhancement of coat protein expression, as well as 5′ leaders and 3′ untranslated regions of MSV transcripts. Apart from the coat protein promoters of Subgroup III geminiviruses, the promoters, polyadenylation signals and other expression regulatory sequences from other genes and from Subgroup I and II viruses, are ill-characterised. There is no doubt that they will yield other interesting and useful elements to add to the plant gene expression toolkit.

Hong *et al.* (1996) showed that the phenomenon of coat protein promoter transcriptional activation by the viral transcription activator protein (TrAP) of Subgroup III geminiviruses could be used for engineering an artificial hypersensitive response in transgenic plants.

This finding has exciting potential in engineering resistance to whitefly transmitted geminiviruses, which are to be considered emerging pathogens with potentially devastating impact on human health and well-being. A patent application (Kridl *et al.*, 1994) highlighted several other areas where the geminivirus transactivation phenomenon might find utility in plant biotechnology, and I hope that this augurs well for further enhancing the profile of these exciting viruses in the plant biotechnology world.

It is interesting that I was unable to show transactivation of the MSV coat protein promoter. It could be that the activation of this promoter is merely linked to genome amplification, but transactivation could also be a tissue specific phenomenon, and should be investigated further. An interesting refinement of the induced cell death system of Hong *et al.* (1996) would be to expand on the system I proposed above by linking the activatable gene to a viral replicon, where it would remain silently integrated in the plant chromosome until "rescued" and both transactivated and amplified by invading virus. This would result in even more efficient cell death, and may be applied to transgenic protection against MSV infection if the lethal gene is not sufficiently toxic as to kill the cell when expressed under the coat protein promoter, which I showed has low-level constitutive activity.

In conclusion, the results of this research have indicated that there is definite potential for use of replication of MSV and related Subgroup I geminiviruses to enhance foreign gene expression, both in transient and transgenic systems. It is important to note that many, if not all, of the current postulates on the use of geminiviruses and their genetic elements in biotechnology are based on laboratory data. I hope that the next phase in geminivirus research leads to their application in resolution of practical issues for full proof of concept.

APPENDIX A

PLANT TISSUE CULTURE METHODS

TABLE OF CONTENTS

A.1 HiII CALLUS MAINTENANCE

Established HiII callus was obtained from Dekalb Plant Genetics (Mystic, Connecticut, USA). Callus was maintained on N6 media, originally developed for rice anther culture by Chu *et al.* (1975) and adapted for maize by Armstrong and Green (1985). N6 salts were purchased from Sigma. N6 medium was solidified with 0.7% tissue culture-grade agar (Sigma). Hi II callus was subcultured every week. Friable, embryogenic callus was selected at every subculture, under a dissecting microscope. Embryogenic callus of transgenic cell lines was not, however, specifically selected. Transgenic cell lines were maintained on N6 medium with 3 mg/l of bialaphos (filter sterilised).

A.2 REGENERATION OF HiII PLANTS

Regeneration of plants from transgenic HiII callus was according to the three-step method of Armstrong (1994). Essentially, bialaphos resistant callus was transferred onto each medium for about 2 weeks. The composition of each of the three media is given in Table A.1, below (from Armstrong, 1994). Murashige and Skoog (MS) salts were obtained from Highveld Biologicals (South Africa). All regeneration of transgenic plants was done on media containing bialaphos at 3 mg/l.

Table A.1: Composition of HiII regenerative media

Regen 1 (per litre)	Regen 2 (per litre)	Regen 3 (per litre)
4.4 g MS salts 1.30 mg nicotinic acid[a] 0.25 mg pyridoxine-HCl[a] 0.25 mg thiamine HCl[a] 0.25 mg Ca-pantothenate 100 mg myo-inositol 1 mM asparagine 0.1 mg 2,4-D[c] 0.1 µM ABA[d] 20 g sucrose 7.0 g tissue culture agar pH 5.8 with potassium hydroxide.	4.0 g N6 salts 0.5 mg nicotinic acid[a] 0.5 mg pyridoxine HCl[a] 2.0 mg glycine[b] 60 g sucrose 7.0 g tissue culture agar pH 5.8 with potassium hydroxide	4.4 g MS salts 1.30 mg nicotinic acid[a] 0.25 mg pyridoxine HCl[a] 0.25 mg thiamine HCl[a] 0.25 mg Ca-pantothenate 100 mg myo-inositol 1 mM asparagine 20 g sucrose 7.0 g tissue culture agar pH 5.8 with potassium hydroxide.

[a] Vitamins were prepared as a 100x stock solution in distilled water
[b] Prepared as a 1000x aqueous stock solution
[c] 2,4-D was prepared as a 1000x stock solution in dilute potassium hydroxide.
[d] ABA was prepared as a 1000x aqueous stock solution.

Callus was left on Regen. 1 medium for two weeks in the dark, then transferred to Regen. 2 for a further two weeks in the dark at approximately 27° C. Callus transferred to Regen. 3 was maintained under low light conditions for one to two weeks, and then transferred to full artificial light. All initial regeneration steps were done with tissue on media in standard 13 cm petri dishes. When plantlets started regenerating, these were transferred into sterilised glass jam jars with transparent lids.

Regenerated plantlets were hardened off a 1:1 bark and polystyrene mix, under plastic. Hardened-off plants were then transferred to large pots and grown in a 1:1:1 ratio of river sand, peat and compost, in the UCT Botany greenhouse. Access to this facility was restricted from the public.

A.3 MAINTENANCE OF BLACK MEXICAN SWEETCORN CULTURES

Black Mexican sweetcorn callus was obtained from Pioneer Hi-Bred (Des Moines, Iowa) on solid media BMS media. Liquid BMS medium was essentially MS medium (Murashige and Skoog, 1962; obtained as a lyophilised powder from Highveld Biologicals, South Africa) with 2 mg/l of 2,4-D. Solid BMS media contained 8 g/l of tissue culture grade agar (Sigma). The liquid cultures were established by placing a small amount of vigorously growing callus into 25 ml of liquid BMS medium in a 250 ml DeLong flask. The flask was shaken continuously in a controlled-temperature (27° C) rotary shaker, at 150 rpm in the dark. When the culture was growing vigorously, the volume was increased to 50 ml. The suspension culture was subcultured once a week by transferring approximately 5 ml of suspended cells into 50 ml of fresh BMS medium.

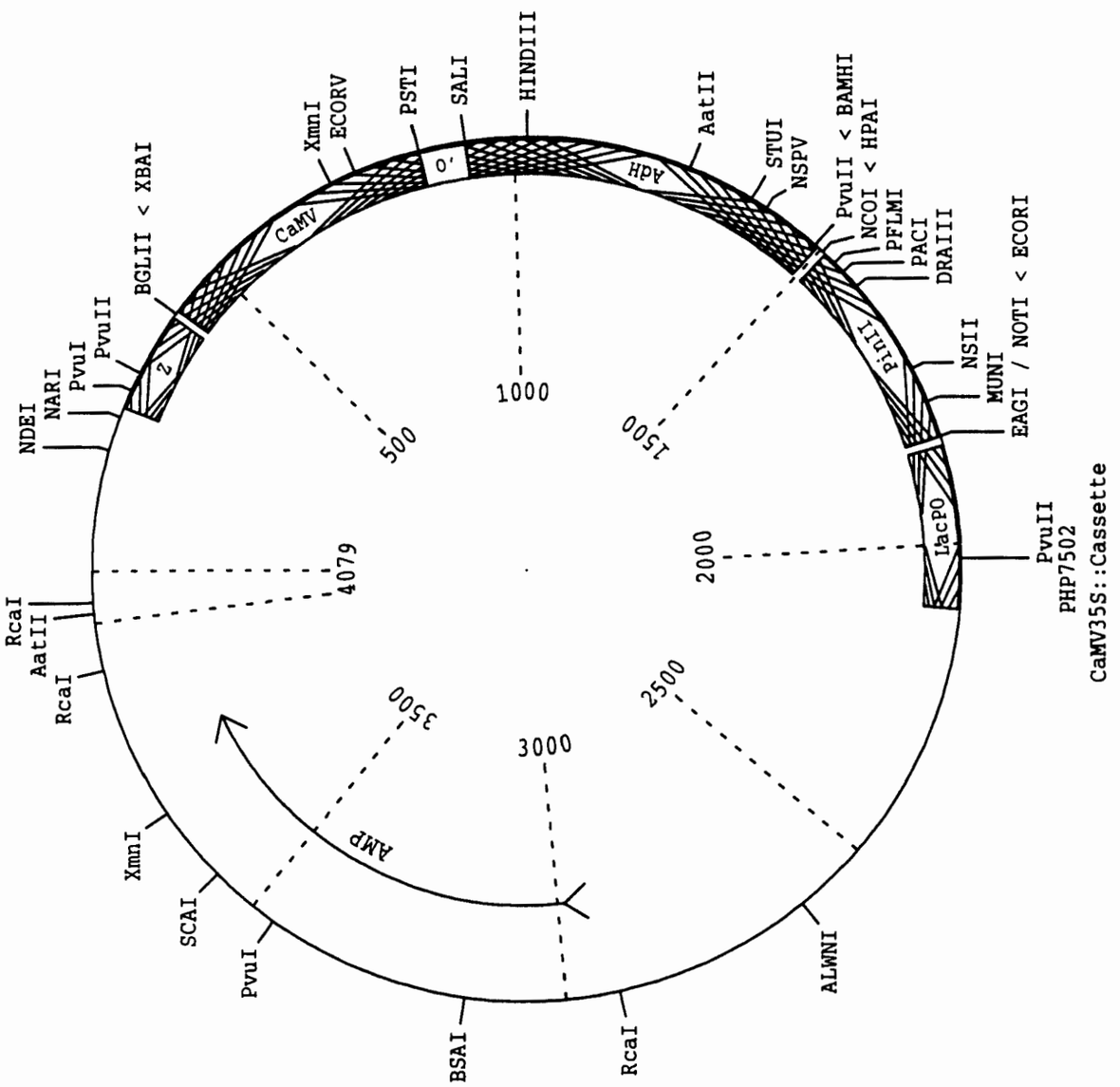
Bialaphos resistant transgenic cell lines were selected and maintained on solidified BMS media containing 3 mg/l of bialaphos. Callus was transferred onto fresh medium every two weeks and maintained at approximately 25° C in the dark.

APPENDIX B

RESTRICTION MAPS OF UNPUBLISHED PLASMIDS AND EXPRESSION CASSETTES OBTAINED FROM PIONEER HIBRED, Inc.

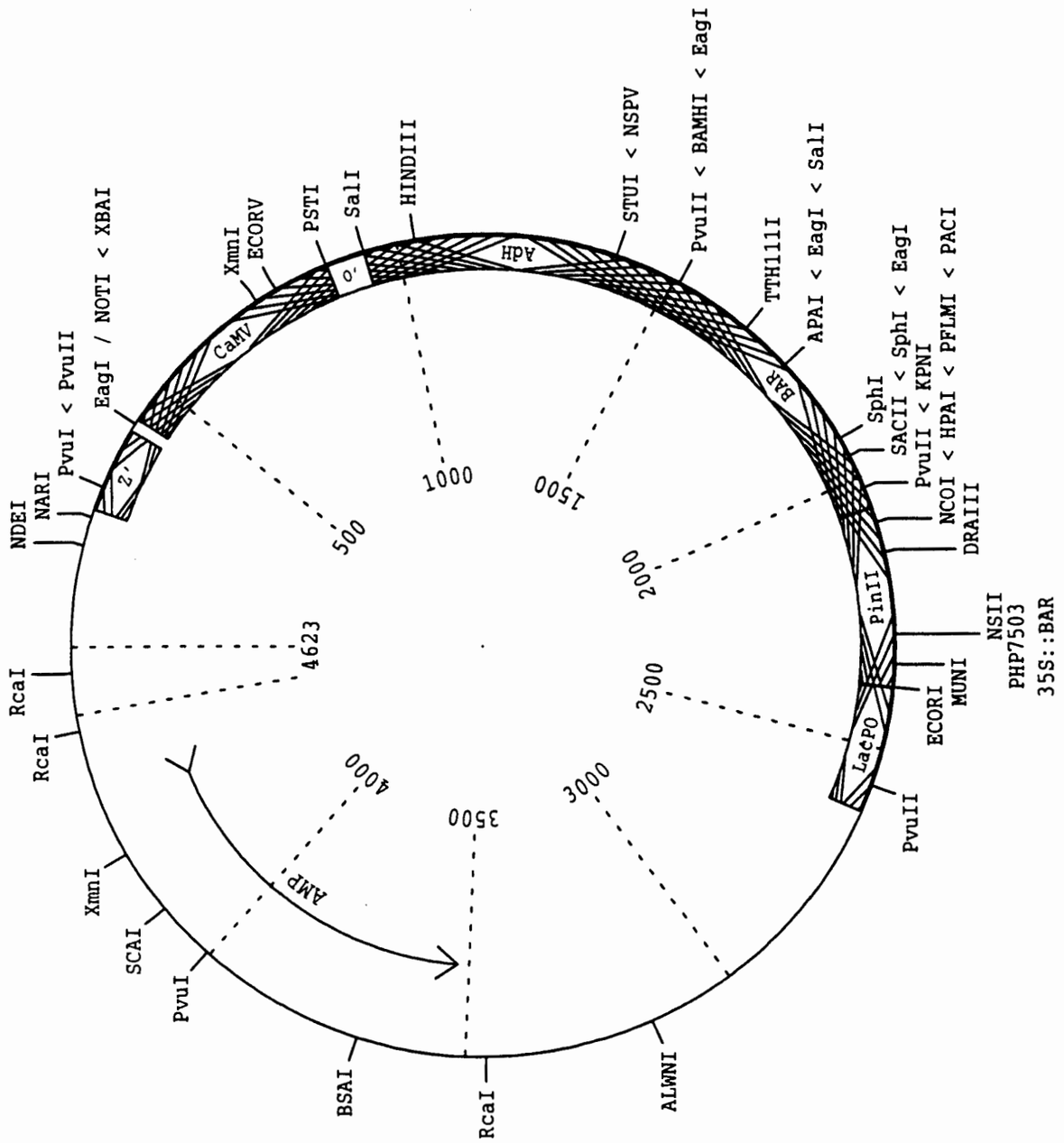
TABLE OF CONTENTS

Note: Restriction maps and construction details of all other plasmids used for obtaining genes and transcription regulatory elements are available in the literature, and referenced in the text.



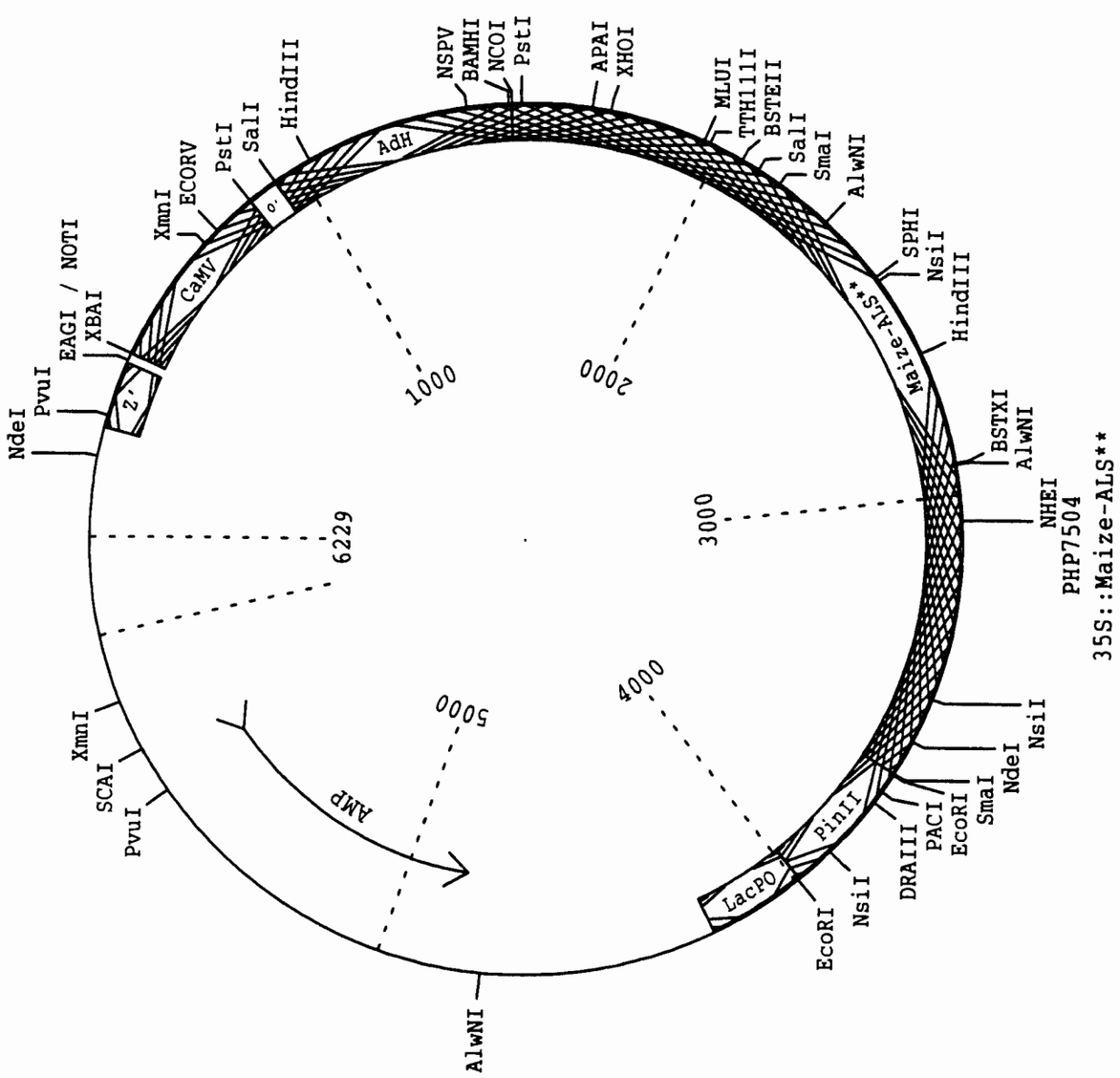
PLASMIDMAP of: Php7502.Seq check: 9565 from: 1 to: 4079

Rothba November 11, 1994 16:30



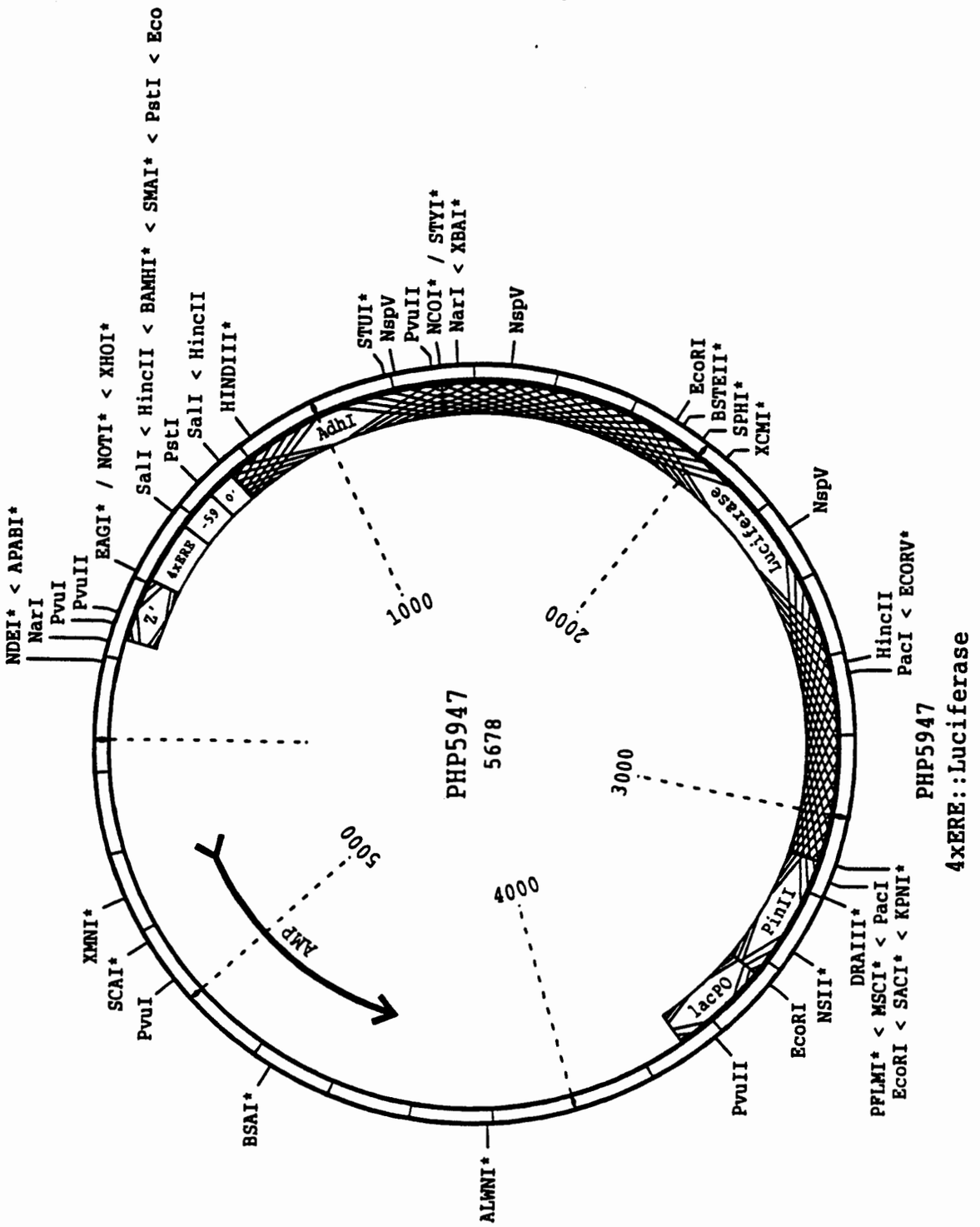
PLASMIDMAP of: Php7503.Seq check: 3896 from: 1 to: 4623

Rothba November 11, 1994 16:30



PLASMIDMAP of: Php7504.Seq check: 3488 from: 1 to: 6229

Rothba November 11, 1994 16:30



PLASMIDMAP of: PHP5947.Seq check: 1029 from: 1 to: 5678

Rothba January 10, 1996 15:03

LITERATURE CITED

Abouzid, A. M., Frischmuth, T., and Jeske, H. (1988). A putative replicative form of abutilon mosaic virus (gemini group) in a chromatin-like structure. *Mol. Gen. Genet.* **212**: 252-258.

Accotto, G. P., Donson, J., and Mullineaux, P. M. (1989). Mapping of Digitaria streak virus transcripts reveals different RNA species from the same transcription unit. *EMBO J.* **8**: 1033-1039.

Accotto, G. P., Mullineaux, P. M., Brown, S. C., and Marie, D. (1993). Digitaria streak geminivirus replicative forms are abundant in S-phase nuclei of infected cells. *Virology* **195**: 257-259.

Adang, M. J., Brody, M. S., Cardineau, G., Eagan, N., Roush, R. T., Shewmaker, C. K., Jones, A., Oakes, J. V., and McBride, K. E. (1993). The reconstruction and expression of a *Bacillus thuringiensis cryIIIA* gene in protoplasts and in potato plants. *Plant Mol. Biol.* **21**: 1131-1145.

Anat, S. and Subramanian, K. N. (1992). Isolation of low molecular weight DNA from bacteria and animal cells. *Meth. Enzymol.* **216**: 20-29.

Andersen, M. T., Richardson, K. A., Harbison, S. -A., and Morris, B. A. M. (1988). Nucleotide sequence of the geminivirus chloris striate mosaic virus. *Virology* **164**: 443-449.

Argüello-Astorga, G. R., Guevara-González, 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.

Argüello-Astorga, G., Hererra-Estrella, L., and Rivera-Bustamante, R. (1994b). Experimental and theoretical definition of geminivirus origin of replication. *Plant Mol. Biol.* **26**: 553-556.

Armitage, P., Walden, R. and Draper, J. (1988). Vectors for the transformation of plant cells using *Agrobacterium*. In: *Plant Genetic Transformation and Gene Expression.* (J. Draper, R. Scott, P. Armitage and R. Walden, Eds.) Blackwell Scientific Publications, Oxford. pp 3-67.

Armstrong, C.L. (1994). Regeneration of plants from somatic cell cultures: applications for in vitro genetic manipulation. In: *The Maize Handbook*, (M. Freeling and V. Walbot, eds.) Springer-Verlag, New York. pp 663-671.

Armstrong, C. L. and Green, C. E. (1985). Establishment of friable,embryogenic maize callus and the involvement of L-proline. *Planta* **164**: 207-214.

Armstrong, C. L., Green, C. E., and Phillips, R. L. (1991). Development and availability of germplasm with high Type II culture formation response. *Maize Genetics Cooperation Newsletter* **65**: 92-93.

Armstrong, C. L. and Songstad, D. D. (1993). Method for transforming monocotyledonous plants. European Patent Application 93870173.7

Azzam, O., Frazer, J., de la Rosa, D., Beaver, J. S., Ahlquist, P., and Maxwell, D. P. (1994). Whitefly transmission and efficient ssDNA accumulation of bean golden mosaic geminivirus require functional coat protein. *Virology* **204**: 289-296.

Bejarano, E. R., Khashoggi, A., Witty, M., and Lichtenstein, C. (1996). Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. *Proc. Natl. Acad. Sci. USA* **93**: 759-764.

Bendahmane, M., Schalk, H. -J., and Gronenborn, B. (1995). Identification and characterization of wheat dwarf virus from France using a rapid method for geminivirus DNA preparation. *Phytopathology* **85**: 1449-1455.

Bendahmane, M. and Gronenborn, B. (1997). Engineering resistance to tomato yellow leaf curl virus (TYLCV) using antisense RNA. *Plant Mol. Biol.* **33**: 351-357.

Bennet, M. D., Smith, J. B., and Heslop-Harrison, J. S. (1982). Nuclear DNA amounts in angiosperms. *Proc. R. Soc. Lond. B* **216**: 179-199.

Bevan, M. (1984). Binary *Agrobacterium* vectors for plant transformation. *Nucl. Acids Res.* **12**: 8711-8721.

Bisaro, D. M. (1994). Recombination in geminiviruses: mechanisms for maintaining genome size and generating genomic diversity. In: *Homologous recombination in plants* (Paszkowski, J., Ed.) Kluwer Academic Publishers, pp 39-60.

Bisaro, D. M. (1996). Geminivirus replication. In: *DNA replication in eukaryotic cells* (De Pamphilis, M. L., Ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Botterman, J., Gosselé, V., Thoen, C., and Lauwereys, M. (1991). Characterization of phosphinothricin acetyltransferase and C-terminal enzymatically active fusion proteins. *Gene* **102**: 33-37.

Boulton, M. I. (1991). *Molecular pathology of maize streak virus*. Ph.D. Thesis, University of East Anglia, Norwich, United Kingdom.

Boulton, M. I. (1995). *Agrobacterium*-mediated transfer of geminiviruses to plant tissues. In: *Plant Gene Transfer and Expression Protocols* (Series: Methods in Molecular Biology Volume 49) (Jones, H., Ed.) Humana Press, Totowa, NJ. pp 77-93.

Boulton, M. I., Buchholz, W. G., Marks, M. S., Markham, P. G., and Davies, J. W. (1989a). Specificity of *Agrobacterium*-mediated delivery of maize streak virus DNA to members of the Gramineae. *Plant Mol. Biol.* **12**: 31-40.

Boulton, M. I. and Davies, J. W. (1990). Monopartite geminiviruses: markers for gene transfer to cereals. *Aspects of Applied Biology* **24**: 79-86.

Boulton, M. I., King, D. I., Donson, J., and Davies, J. W. (1991a). Point substitutions in a promoter-like region and the V1 gene affect the host range and symptoms of maize streak virus. *Virology* **183**: 114-121.

Boulton, M. I., King, D. I., Markham, P. G., Pinner, M. S., and Davies, J. W. (1991b). Host range and symptoms are determined by specific domains of the maize streak virus genome. *Virology* **181**: 312-318.

Boulton, M. I., Pallaghy, C. K., Chatani, M., MacFarlane, S., and Davies, J. W. (1993). Replication of maize streak virus mutants in maize protoplasts: evidence for a movement protein. *Virology* **192**: 85-93.

Boulton, M. I., Steinkellner, H., Donson, J., Markham, P. G., King, D. I., and Davies, J. W. (1989b). Mutational analysis of the virion-sense genes of maize streak virus. *J. Gen. Virol.* **70**: 2309-2323.

Bradford, M. M. (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.

Bravo-Angel, A. M., Becker, H.-A., Kunze, R., Hohn, B. and Shen, W. (1995). The binding motifs for *Ac* transposase are absolutely required for excision of *Ds1* in maize. *Mol. Gen. Genet.* **248**: 527-534.

Briddon, R. W., Bedford, I. D., Tsai, J. H., and Markham, P. G. (1996). Analysis of the nucleotide sequence of the treehopper-transmitted geminivirus, tomato pseudo-curly top virus, suggests a recombinant origin. *Virology* **219**: 387-394.

Briddon, R. W., Lunness, P., Chamberlin, L. C. L., and Markham, P. G. (1994). Analysis of the genetic variability of maize streak virus. *Virus Genes* **9**: 93-100.

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

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.

Broadbent, P., Creissen, G., Kular, B., Wellburn, A. R. and Mullineaux, P. M. (1995). Oxidative stress responses in transgenic tobacco containing altered levels of glutathione reductase activity. *Plant J.* **8**: 247-255.

Brough, C. L., Gardiner, W. E., Inamdar, N. M., Zhang, X. -Y., Ehrlich, M., and Bisaro, D. M. (1992a). DNA methylation inhibits propogation of tomato golden mosaic virus DNA in transfected protoplasts. *Plant Mol. Biol.* **18**: 703-712.

Brough, C. L., Hayes, R. J., Morgan, A. J., Coutts, R. H. A. and Buck, K. W. (1988). Effects of mutagenesis *in vitro* on the ability of cloned tomato golden mosaic virus DNA to infect *Nicotiana benthamiana* plants. *J. Gen. Virol.* **69**: 503-514.

Brough, C. L., Sunter, G., Gardiner, W. E., and Bisaro, D. M. (1992b). Kinetics of tomato golden mosaic virus DNA replication and coat protein promoter activity in Nicotiana tabacum protoplasts. *Virology* **187**: 1-9.

Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. *Genes & Development* **1**: 1183-1200.

Chapman, S., Kavanagh, T. and Baulcombe, D. (1992). Potato virus X as a vector for gene expression in plants. *Plant J.* **2**: 549-547.

Chasan, R. (1993). *Ac* tagging moves beyond maize. *Plant Cell* **5**: 361-363.

Chen, D.F. and Dale, P.J. (1992). A comparision of methods for delivering DNA to wheat: the application of wheat dwarf virus DNA to seeds with exposed apical meristems. *Transgenic Research* **1**: 93-100.

Chen, J. and Dellaporta, S. (1994). Urea-based plant DNA miniprep. In: *The Maize Handbook* (M. Freeling and V. Walbot, eds.) Springer-Verlag, New York. pp 523-533.

Choi, I. -R. and Stenger, D. C. (1995). Strain-specific determinants of beet curly top geminivirus DNA replication. *Virology* **206**: 904-912.

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.

Christensen, 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.

Christou, P. (1996). Transformation technology. *Trends in Plant Science* **1**: 423-431.

Chu, C. C., Wang, C. C., Sun, C. S., Hus, C., Yin, K. C. and Chu, C. Y. (1975). Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources. *Scientica Sinica* **18**: 659-668.

Citovsky, V. (1993). Probing plasmodesmal transport with plant viruses. *Plant Physiol.* **102**: 1071-1076.

Collin, S., Fernández-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.

Coutts, R. H. A., Buck, K. W., and Hayes, R. J. (1990). Development of geminivirus-based gene vectors for dicotyledonous plants. *Austr. J. Plant Physiol.* **17**: 365-375.

Creissen, G., Smith, C., Francis, R., Reynolds, H., and Mullineaux, P. M. (1990). *Agrobacterium* - and microprojectile - mediated viral DNA delivery into barley microspore-derived cultures. *Plant Cell Reports* **8**: 680-683.

Dale, P.J., Marks, M.S., Brown, M.M., Woolston, C.J., Gunn, H.V., Mullineaux, P.M., Lewis, D.M., Kemp, J.M., Chen, D.F., Gilmour, D.M. and Flavell, R.B. (1988). Agroinfection of wheat: inoculation of in vitro grown seedlings and embryos. *Plant Science* **63**: 237-245.

De Bock, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gosselé, V., Rao Movva, N., Thompson, C., van Montagu, M. and Leemans, J. (1987). Engineering herbicide resistance in plants by expression of a detoxifying enzyme. *EMBO J.* **6**: 2513-2518.

Dekker, E. L., Woolston, C. J., Xue, Y., Cox, B., and Mullineaux, P. M. (1991). Transcript mapping reveals different expression strategies for the bicistronic RNAs of the geminivirus wheat dwarf virus. *Nucl. Acids Res.* **19**: 4075-4081.

del Solar, G., Moscoso, M., and Espinosa, M. (1993). Rolling circle-replicating plasmids from Gram-positive and Gram-negative bacteria: a wall falls. *Mol. Microbiol.* **8**: 789-796.

DePamphilis, M.L. (1993). Eukaryotic DNA replication: anatomy of an origin. *Annu. Rev. Biochem.* **62**: 29-63.

Desbiez, C., David, C., Mettouchi, A., Laufs, J., and 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.

D'Halluin, K., De Block, M., Denecke, J., Janssens, J., Leemans, J., and Botterman, J. (1992). The *bar* gene as a selectable and screenable marker in plant engineering. *Methods Enzymol.* **216**: 415-426.

Dickinson, V. J., Halder, J., and Woolston, C. J. (1996). The product of maize streak virus ORF V1 is associated with secondary plasmodesmata and is first detected with the onset of viral lesions. *Virology* **220**: 51-59.

Dixon, L. K., and Hohn, T. (1984). Initiation of translation of the cauliflower mosaic virus genome from a polycistronic mRNA: evidence from deletion mutagenesis. *EMBO J.* **3**: 2731-2736.

Dollet, M., Accotto, G. P., Lisa, V., Menissier, J., and Boccardo, G. (1986). A geminivirus, serologically related to maize streak virus, from *Digitaria sanguinalis* from Vanuatu. *J. Gen. Virol.* **67**: 933-937.

Donson, J., Accotto, G. P., Boulton, M. I., Mullineaux, P. M., and Davies, J. W. (1987). The nucleotide sequence of a geminivirus from *Digitaria sanguinalis*. *Virology* **161**: 160-169.

Donson, J., Dawson, W. O., Grantham, G. L., Turpen, T. H., Turpen, A. M., Garger, S. J., and Grill, L. K. (1993). Recombinant plant viral nucleic acids. World Intellectual Property Organization WO 93/03161

Donson, J., Gunn, H. V., Woolston, C. J., Pinner, M. S., Boulton, M. I., Mullineaux, P. M., and Davies, J. W. (1988). *Agrobacterium*-mediated infectivity of cloned digitaria streak virus DNA. *Virology* **162**: 248-250.

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 J.* **3**: 3069-3073.

Dowson Day, M. J., Ashurst, J. L., Mathias, S. F., Watts, J. W., Wilson, T. M. A., and Dixon, R. A. (1993). Plant viral leaders influence expression of a reporter gene in tobacco. *Plant Mol. Biol.* **23**: 97-109.

Dunder, E., Dawson, J., Suttie, J. and Pace, G. (1995). Maize transformation by microprojectile bombardment of immature embryos. In: *Gene Transfer to Plants*. (I. Potrykus and G. Spangenberg, eds.) Springer-Verlag, Berlin and Heidelberg. pp126-138.

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.

Elmer, S. and Rogers, S. G. (1990). Selection for wild type size derivatives of tomato golden mosaic virus during systemic infection. *Nucl. Acids Res.* **18**: 2001-2006.

Ermak, G., Paszkowski, U., Wohlmuth, M., Mittelsten Scheid, O. and Paszkowski, J. (1993). Cytosine methylation inhibits replication of African cassava mosaic virus by two distinct mechanisms. *Nucl. Acids. Res.* **21**: 3445-3450.

Escudero, J. and Hohn, B. (1994). Agroinfection. In: *The Maize Handbook* (Freeling, M. and Walbot, V., Eds.) Springer-Verlag, New York. pp 599-603.

Escudero, J., Neuhaus, G., Schlappi, M., and Hohn, B. (1996). T-DNA transfer in meristematic cells of maize provided with intracellular *Agrobacterium*. *Plant J.* **10**: 355-360.

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

Etessami, P., Watts, J., and Stanley, J. (1989). Size reversion of African cassava mosaic virus coat protein gene deletion mutants during infection of Nicotiana benthamiana. *J. Gen. Virol.* **70**: 277-289.

Evans, D. and Jeske, H. (1993). Complementation and recombination between mutants of complementary sense genes of DNA A of abutilon mosaic virus. *Virology* **197**: 492-496.

Evans, D. and Jeske, H. (1993). DNA B facilitates, but is not essential for, the spread of abutilon mosaic virus in agroinoculated Nicotiana benthamiana. *Virology* **194**: 752-757.

Fanning, E. and Knippers, R. (1992). Structure and function of simian virus 40 large tumor antigen. *Annu. Rev. Biochem.* **61**: 55-85.

Farrell, L. and Beachy, R. N. (1990). Manipulation of β-glucuronidase for use as a reporter in vacuolar targeting studies. *Plant Mol. Biol.* **15**: 821-825.

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., Schwartz, 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.

Finnegan, J. and McElroy, D. (1994). Transgene inactivation: plants fight back!. *Bio/Technology* **12**: 883-888.

Flores, N., Valle, F., Bolivar, F. and Morino, E. (1992). Recovery of DNA from agarose gels stained with methylene blue. *BioTechniques* **13**: 204-205.

Fontes, E. P. B., Eagle, P. A., Sipe, P. S., 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 organisation. *Plant Cell* **6**: 405-416.

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.

Foyer, C., Descourvières, P. and Kunert, K. J. (1994) Protection against oxygen radicals: an important defence mechanism studied in transgenic plants. *Plant Cell and Environment* **17**: 507-523.

Foyer, C., Lelandais, M., Galap, C. and Kunert, K. J. (1991). Effects of elevated cytosolic glutathione reductase activity on the cellular glutathione pool and photosynthesis in leaves under normal and stress conditions. *Plant Physiol.* **97**: 863-872.

French, R., Janda, M., and Ahlquist, P. (1986). Bacterial gene inserted in an engineered RNA virus: efficient expression in monocotyledonous plant cells. *Science* **231**: 1294-1299.

Frischmuth, S., Frischmuth, T., Latham, J. R., and Stanley, J. (1993). Transcriptional analysis of the virion-sense genes of the geminivirus beet curly top virus. *Virology* **197**: 312-319.

Frischmuth, T., Roberts, S., von Arnim, A., and Stanley, J. (1993). Specificity of bipartite geminivirus movement proteins. *Virology* **196**: 666-673.

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.

Fromm, M. E. (1994). Production of transgenic maize via microprojectile-mediated gene transfer. In: *The Maize Handbook* (M. Freeling and V. Walbot, eds.) Springer-Verlag, New York. pp 677-684.

Fromm, M. E., Morrish, F., Armstrong, C., Williams, R., Thomas, J., and Klein, T. M. (1990). Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Bio/Technology* **8**: 833-839.

Fromm, M. E., Taylor, L. P. and Walbot, V. (1986). Stable transformation of maize after gene transfer by electroporation. *Nature* **319**: 791-793.

Fuqua, W. C. (1992). An improved chloramphenicol resistance gene cassette for site-directed marker replacement mutagenesis. *BioTechniques* **12**(2): pp?

Fullner, K. J., Lara, J. C. and Nester, E. W. (1996). Pilus assembly by *Agrobacterium* T-DNA transfer genes. *Science* **273**: 1107-1109.

Gallie, D. R. (1993). Posttranscriptional regulation of gene expression in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44**: 77-105.

Gallie, D. R., Lucas, W. J. and Walbot, V. (1989). Visualizing mRNA expression in plant protoplasts: factors influencing efficient mRNA uptake and translation. *Plant Cell* **1**: 301-311.

Gallie, D. R., Sleat, D. E., Watts, J. W., Turner, P. C. and Wilson, T. M. A. (1987). The 5′-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. *Nucl. Acids. Res.* **15**: 3257-3273.

Gardiner, W. E., Sunter, G., Brand, L., Elmer, J. S., Rogers, S. G., and Bisaro, D. M. (1988). Genetic analysis of tomato golden mosaic virus: the coat protein is not required for systemic spread of symptom development. *EMBO J.* **7**: 899-904.

Gilbertson, R. L., Hidayat, S. H., Paplomatas, E. J., Rojas, M. R., Hou, Y. -M., and Maxwell, D. P. (1993). Pseudorecombination between infectious cloned DNA components of tomato mottle and bean dwarf mosaic geminiviruses. *J. Gen. Virol.* **74**: 23-31.

Gilbertson, R.L. and Lucas, W.J. (1996). How do plant viruses traffic on the 'vascular highway'? *Trends in Plant Science* **1**: 260-268.

Goff, S. A., Cone, K. C. and Chandler, V. L. (1992). Functional analysis of the transcription activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes & Development* **6**: 864-875.

Goff, S.A., Klein, T.M., Roth, B.A., Fromm, M.E., Cone, K. C., Radicella, J. P., and Chandler, V. L. (1990). Transactivation of anthocyanin biosynthetic genes following transfer of *B* regulatory genes into maize tissues. *EMBO J.* **9**: 2517-2522.

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. J., Spencer, T. M., Mangano, M. L., Adams, T. R., Daines, R. J., Start, W. G., O'Brien, J. V., Chambers, S. A., Adams, W. R., Willets, N. G., Rice, T. B., Mackey, C. J., Krueger, R. W., Kausch, A. P., and Lemaux, P. G. (1990). Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell* **2**: 603-618.

Grafi, G., Burnett, R. J., Helentjaris, T., Larkins, B. A., DeCaprio, J. A., Sellers, W. R., and Kaelin, W. G. (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.H., Hohn, B., Hohn, T. and Walden, R. (1986). "Agroinfection", an alternative route for viral infection of plants using the Ti plasmid. *Proc. Natl. Acad. Sci. USA* **83**: 3282-3286.

Grimsley, N., Hohn, B., Ramos, C., Kado, C. and Rogowsky, P. (1989). DNA transfer from *Agrobacterium* to *Zea mays* or *Brassica* by agroinfection is dependent on bacterial virulence functions. *Mol. Gen. Genet.* **217**: 309-316.

Grimsley, N. H., Hohn, T., Davies, J.W. and Hohn, B. (1987). *Agrobacterium*-mediated delivery of infectious maize streak virus into maize plants. *Nature* **325**: 177-179.

Grimsley, N. H., Ramos, C., Hein, T. and Hohn B. (1988). Meristematic tissues of maize plants are most susceptible to agroinfection with maize streak virus. *Bio/Technology* **6**: 185-188.

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.

Hackland, A.F. (1994). The development of transgenic plants resistant to cucumber mosaic virus and tobacco necrosis virus. Ph.D. thesis, University of Cape Town.

Hafner, G. J., Harding, R. M. and Dale, J. L. (1997). A DNA primer associated with banana bunchy top virus. *J. Gen. Virol.* **78**: 479-486.

Hamamoto, H., Sugiyama, Y., Nakagawa, N., Hashida, E., Matsunaga, Y., Takemoto, S., Watanabe, Y., and Okada, Y. (1993). A new tobacco mosaic virus vector and its use for the systemic production of angiotensin-I-converting enzyme inhibitor in transgenic tobacco and tomato. *Bio/Technology* **11**: 930-932.

Hanley-Bowdoin, L., Eagle, P. A., Orozco, B. M., Robertson, D., and Settlage, S. B. (1997). Geminivirus replication. *MPMI* (in press)

Hanley-Bowdoin, L., Elmer, J. S., and Rogers, S. G. (1988). Transient expression of heterologous RNAs using tomato golden mosaic virus. *Nucl. Acids Res.* **16**: 10511-10528.

Hanley-Bowdoin, L., Elmer, J. S., and Rogers, S. G. (1989). Functional expression of the leftward open reading fames of the A component of tomato golden mosaic virus in transgenic tobacco plants. *Plant Cell* **1**: 1057-1067.

Hanley-Bowdoin, L., Elmer, J. S., and Rogers, S. G. (1990). Expression of functional replication protein from tomato golden mosaic virus in transgenic tobacco plants. *Proc. Natl. Acad. Sci. USA* **87**: 1446-1450.

Hanson, S. F., Hoogstraten, R. A., Ahlquist, P., Gilbertson, R. L., Russell, D. R., and Maxwell, D. P. (1995). Mutational analysis of a putative NTP-binding domain in the replication-associated protein (AC1) of bean golden mosaic geminivirus. *Virology* **211**: 1-9.

Hayes, R. J., Coutts, R. H. A., and Buck, K. W. (1989). Stability and expression of bacterial genes in replicating geminivirus vectors in plants. *Nucl. Acids Res.* **17**: 2391-2403.

Hayes, R. J., MacDonald, H., Coutts, R. H. A., and Buck, K. W. (1988a). 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.

Hayes, R. J., Petty, I. T. D., Coutts, R. H. A., and Buck, K. W. (1988b). Gene amplification and expression in plants by a replicating geminivirus vector. *Nature* **334**: 179-182.

Heckel, T. (1996). *Pathogenicity determinants and gene expression of maize streak virus.* Ph.D. Thesis, University of East Anglia, Norwich, United Kingdom.

Heyraud, F., Matzeit, V., Kamman, M., Schaefer, S., Schell, J., and Gronenborn, B. (1993a). 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. (1993b). The conserved nonanucleotide motif of the geminivirus stem-loop sequence promotes replicational release of virus molecules from redundant copies. *Biochimie* **75**: 605-615.

Heyraud-Nitschke, F., Schumacher, S., Laufs, J., Schaefer, S., Schell, J., and Gronenborn, B. (1995). Determination of the origin cleavage and joining domain of geminivirus Rep proteins. *Nucl. Acids Res.* **23**: 910-916.

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994). Efficient transformation of rice (Oryza sativa L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.* **6**: 271-282.

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.

Holsters, M., de Waele, D., Depicker, A., Messens, E., van Montagu, M. and Schell, J. (1978). Transfection and transformation of *Agrobacterium tumefaciens*. *Mol. Gen. Genet.* **163:** 181-187.

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.

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

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.

Hoogstraten, R. A., Hanson, S. F., and Maxwell, D. P. (1996). Mutational analysis of the putative nicking motif in the replication-associated protein (AC1) of bean golden mosaic geminivirus. *MPMI.* **9**: 594-599.

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**: 900-904.

Hormuzdi, S. G. and Bisaro, D. M. (1995). Genetic analysis of beet curly top virus: examination of the roles of L2 and L3 genes in viral pathogenesis. *Virology* **206**: 1044-1054.

Hou, Y. -M. and Gilbertson, R. L. (1996). Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. *J. Virol.* **70**: 5430-5436.

Howell, S. H. (1984). Physical structure and genetic organisation of maize streak virus (Kenyan isolate). *Nucl. Acids Res.* **12**: 7359-7375.

Hughes, F. L. (1990). Molecular investigations of Subgroup I geminiviruses. Ph.D. thesis, University of Cape Town.

Hughes, F.L., Rybicki, E.P. and Kirby, R. (1993). Complete nucleotide sequence of sugarcane streak monogeminivirus. *Arch. Virol.* **132**: 171-182.

Hughes, F. L., Rybicki, E. P., and von Wechmar, M. B. (1992). Genome typing of southern African subgroup I geminiviruses. *J. Gen. Virol.* **73**: 1031-1040.

Hull, G. A. and Devic, M. (1995). The β-glucuronidase (*gus*) reporter gene system: gene fusions, spectrophotometric, fluorometric and histochemical detection. In: *Plant Gene Transfer and Expression Protocols* (Series: *Methods In Molecular Biology*, volume **49**). H. Jones, ed. Humana Press, Totowa, NJ. pp125-142.

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

Ingham, D. J. and Lazarowitz, S. G. (1993). A single missense mutation in the BR1 movement protein alters the host range of the squash leaf curl geminivirus. *Virology* **196**: 694-702.

Ingham, D. J., Pascal, E., and Lazarowitz, S. G. (1995). Both bipartite geminivirus movement proteins define viral host range, but only BL1 determines viral pathogenicity. *Virology* **207**: 191-204.

Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T. and Kumashiro, T. (1996). High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*. *Nature Biotechnology* **14**: 745-750.

Jacobson, A. and Peltz, S.W. (1996). Inerrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annu. Rev. Biochem.* **65**: 693-739.

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.

Jeffrey, J. L., Pooma, W., and Petty, I. T. D. (1996). Genetic requirements for local and systemic movement of tomato golden mosaic virus in infected plants. *Virology* **223**: 208-218.

Joshi, C. P. (1987). An inspection of the domain between putative TATA box and translation start in 79 plant genes. *Nucl. Acids Res.* **16**: 6643-6653.

Joshi, R. L., Joshi, V., and Ow, D. W. (1990). BSMV genome mediated expression of a foreign gene in dicot and monocot plant cells. *EMBO J.* **9**: 2663-2669.

Julia, J. F., and Dollet, M. (1989). *Nesoclutha declivata* Homoptera Cicadellidae: vector of digitaria streak virus (geminivirus) in Vanuatu. *J. Phytopathol.* **127**: 42-28.

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

Jupin, I., Hericourt, F., Benz, B., and Gronenborn, B. (1995). DNA replication specificity of TYLCV geminivirus is mediated by the amino-terminal 116 amino acids of the Rep protein. *FEBS Lett.* **362**: 116-120.

Kamman, M., Matzeit, V., Schmidt, B., Schell, J., Walden, R., and Gronenborn, B. (1991a). Geminivirus-based shuttle vectors capable of replication in Escherichia coli and monocotyledonous plant cells. *Gene* **104**: 247-252.

Kamman, M., Schalk, H. -J., Matzeit, V., Schaefer, S., Schell, J., and Gronenborn, B. (1991b). DNA replication of wheat dwarf virus, a geminivirus, requires two cis-acting signals. *Virology* **184**: 786-790.

Kanevski, I. F., Thakur, S., Cosowsky, L., Sunter, G., Brough, C., Bisaro, D., and Maliga, P. (1992). Tobacco lines with high copy numbers of replicating recombinant geminivirus vectors after biolistic DNA delivery. *Plant J.* **2**: 457-463.

Kikkert, J.R. (1993). The biolistic PDS-100/He device. *Plant Cell Tissue and Organ Culture* **33**: 221-226.

Kikuno, R., Toh, H., Hayashida, H. and Miyata, T. (1984). Sequence similarity between putative gene products of geminiviral DNAs. *Nature* **308**: 562-563.

Kirihara, J. A. (1994). Selection of stable transformants from Black Mexican sweet maize suspension cultures. In: *The Maize Handbook* (M. Freeling and V. Walbot, eds.) Springer-Verlag, New York. pp 690-694.

Klein, T. M., Kronstein, L., Sanford, J. C. and Fromm, M. E. (1989a). Genetic transformation of maize cells by particle bombardment. *Plant Physiol.* **91**: 440-444.

Klein, T. M., Roth, B. A. and Fromm, M. E. (1989b). Regulation of anthocyanin biosynthetic genes introduced into intact maize tissues by microprojectiles. *Proc. Natl. Acad. Sci. USA* **86**: 6681-6685.

Klinkenberg, F. A., Ellwood, S., and Stanley, J. (1989). Fate of African cassava mosaic virus coat protein deletion mutants after agroinoculation. *J. Gen. Virol.* **70**: 1837-1844.

Koncz, C. and Schell, J. (1986). The promoter of T_L-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, 1986.

Koonin, E. V. and Ilyina, T. V. (1992). Geminivirus replication proteins are related to prokaryotic plasmid rolling circle 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.

Kozak, M. (1989). Context effects and inefficient initiation at non-AUG codons in eukaryotic cell-free translation systems. *Mol. Cell Biol.* **9**: 5073-5080.

Kridl, J. C., Knauf, V., and Bruening, G. (1994). Geminivirus-based gene expression system. World Intellectual Property Organisation WO 94/19477.

Kumagai, M. H., Turpen, T. H., Weinzettl, N., della-Cioppa, G., Turpen, A. M., Donson, J., Hilf, M. E., Grantham, G. L., Dawson, W. O., and Chow, T. P. (1993). Rapid, high-level expression of biologically active alpha-trichosanthin in transfected plants by an RNA viral vector. *Proc. Natl. Acad. Sci. USA* **90**: 427-430.

Kunert, K. J., Cresswell, C. F., Mullineaux, P. M. and Foyer, C. H. (1990). Variations in the activity of glutathione reductase and the cellular glutathione content in relation to sensitivity to methylviologen in *Escherichia coli. Arch. Biochem. Biophys.* **282**: 233-238.

Kunik, T., Salomon, R., Zamir, D., Navot, N., Zeidan, M., Michelson, I., Gafni, Y., and Czosnek, H. (1994). Transgenic tobacco plants expressing the tomato yellow leaf curl virus capsid protein are resistant to the virus. *Bio/Technology* **12**: 500-504.

Laufs, J., Jupin, I., Schumacher, S., Heyraud-Nitschke, F., and Gronenborn, B. (1995a). Geminivirus replication: Genetic and biochemical characterization of Rep protein function, a review. *Biochimie* 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 Lett.* **377**: 258-262.

Laufs, J., Traut, W., Heyraud, F., Matzeit, 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.

Laufs, J., Wirtz, U., Kammann, M., Matzeit, V., Schaefer, S., Schell, J., Czernilofsky, A. P., Baker, B., and Gronenborn, B. (1990). Wheat dwarf virus Ac/Ds vectors: expression and excision of transposable elements introduced into various cereals by a viral replicon. *Proc. Natl. Acad. Sci. USA* **87**: 7752-7756.

Lazarowitz, S. G. (1988). Infectivity and complete nucleotide sequence of the genome of a South African isolate of maize streak virus. *Nucl. Acids Res.* **16**: 229-249.

Lazarowitz, S. G. (1992). Geminiviruses: genome structure and gene function. *CRC Crit. Rev. Plant Sci.* **11**: 327-349.

Lazarowitz, S. G., Pinder, A. J., Damsteegt, V. D., and Rogers, S. G. (1989). Maize streak virus genes essential for systemic spread and symptom development. *EMBO J.* **8**: 1023-1032.

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.

Leisner, S.M. and Howell, S.H. (1993). Long distance movement of viruses in plants. *Trends in Microbiology* **1**: 314-317.

Liu, H., Boulton, M. I., and Davies, J. W. (1997). *J. Gen. Virol.* (in press).

Lomonossoff, G. P. and Johnson, E. (1992). Modified plant viruses as vectors. World Intellectual Property Organisation WO 92/18618.

Lowe, K., Bowen, B., Hoerster, G., Ross, M., Bond, D., Pierce, D., and Gordon-Kamm, B. (1995). Germline transformation of maize following manipulation of chimeric shoot meristems. *Bio/Technology* **13**: 677-682.

Lucas, W.J., Bouché-Pillon, S., Jackson, D.P., Nguyen, L., Baker, L., Ding, B. and Hake, S. (1995). Selective trafficking of KNOTTED1 homeodomain protein and its mRNA through plasmodesmata. *Science* **270**: 1980-1983.

Lucas, W. J. and Gilbertson, R. L. (1994). Plasmodesmata in relation to viral movement within leaf tissues. *Annu. Rev. Phytopathol.* **32**: 387-411.

Lucy, A. P., Boulton, M. I., Davies, J. W., and Maule, A. J. (1996). Tissue specificity of Zea mays infection by maize streak virus. *MPMI* **9**: 22-31.

Luehrsen, K. R. and Walbot, V. (1991). Intron enhancement of gene expression and the splicing efficiency of introns in maize cells. *Mol. Gen. Genet.* **225**: 81-93.

Lusardi, M.C., Neuhaus-Url, G., Potrykus, I. and Neuhaus, G. (1994). An approach towards genetically engineered cell fate mapping in maize using the *Lc* gene as a visible marker: transactivation capacity of *Lc* vectors in diffferentiated maize cells and microinjection of *Lc* vectors into somatic embryos and shoot apical meristems. *Plant J.* **5**: 571-582.

Lutcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern H. F., and Scheele, G. A. (1987). Selection of AUG codons differs in plants and animals. *EMBO J.* **6**: 43-48.

MacDonald, H., Coutts, R. H. A., and Buck, K. W. (1988). Characterization of a subgenomic DNA isolated from Triticum aestivum plants infected with wheat dwarf virus. *J. Gen. Virol.* **69**: 1339-1334.

Mason, H. S. and Arntzen, C. J. (1995). Transgenic plants as vaccine production systems. *Trends in Biotechnology* **13**: 388-392.

Mason, H. S., Ball, J. M., Shi, J.-J., Jiang, X., Estes, M. K. and Arntzen C. J. (1996). Expression of Norwalk virus capsid protein in transgenic tobacco and potato and its oral immunogenicity in mice. *Proc. Natl. Acad. Sci. USA* **93**: 5335-5340.

Marks, M.S., Kemp, J.M., Woolston, C.J. and Dale, P.J. (1989). Agroinfection of wheat: a comparison of different *Agrobacterium* strains. *Plant Science* **63**: 247-256.

Matzeit, V., Schaefer, S., Kammann, M., Schalk, H. -J., Schell, J., and Gronenborn, B. (1991). Wheat dwarf virus vectors replicate and express foreign genes in cells of monocotyledonous plants. *Plant Cell* **3**: 247-258.

Matzke, M. A. and Matzke, A. J. M. (1995). How and why do plants inactivate homologous (trans)genes? *Plant Physiol.* **107**: 679-685.

May, G. D., Mason, H. S. and Lyons, P. (1996). Application of transgenic plants as production systems for pharmaceuticals. In: Fullner, G., Bills, D. and McKeon, T. (eds) ACS Symposium Series: *Agriculture as a Renewable Source of Raw Materials*. Washington, DC. ACS Books. (in press).

McCullough, A. J., Lou, H., and Schuler, M. A. (1991). In vivo analysis of plant pre-mRNA splicing using an autonomously replicating vector. *Nucl. Acids Res.* **19**: 3001-3009.

McElroy, D., Blowers, A. D., Jenes, B. and Wu, R. (1991). Construction of expression vectors based on the rice actin 1 (*Act1*) 5′ region for use in monocot transformation. *Mol. Gen. Genet.* **231**: 150-160.

McElroy, D. and Brettell, R. I. S. (1994). Foreign gene expression in transgenic cereals. *Trends in Biotechnology* **12**: 62-68.

McElroy, D., Louwerse, J. D., McElroy, S. M., and Lemaux, P. G. (1997). Development of a simple transient assay for *Ac/Ds* activity in cells of intact barley tissue. *Plant Journal* **11**: 157-165.

Meyer, P., Heidman, I., and Niedenhof, I. (1992). The use of African cassava mosaic virus as a vector system for plants. *Gene* **110**: 213-217.

Meyer, P., Niedenhof, I., Heidmann, I., and Saedler, H. (1989). Extrachromosomal forms of CLV *DNA1* in transgenic plants are inherited by symptom-free progeny. *Plant Science* **65**: 207-216.

Milne, R. G. and Leseman, D. (1984). Immunosorbent electron microscopy in plant virus studies. *Methods in Virology* **3**: 85-101.

Mori, M., Zhang, G. -H., Kaido, M., Okuno, T., and Furusawa, I. (1993). Efficient production of human gamma interferon in tobacco protoplasts by genetically engineered brome mosaic virus RNAs. *J. Gen. Virol.* **74**: 1255-1260.

Morris, B. A. M., Richardson, K. A., Haley, A., Zhan, X., and Thomas, J. E. (1992). The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous hosts. *Virology* **187**: 633-642.

Morris-Krsinich, B. A. M., Mullineaux, P. M., Donson, J., Boulton, M. I., Markham, P. G., Short, M. N., and Davies, J. W. (1985). Bidirectional transcription of maize streak virus DNA and identification of the coat protein gene. *Nucl. Acids Res.* **13**: 7237-7256.

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., Davies, J. W., and Woolston, C. J. (1992). Geminiviruses as gene vectors. In: *Genetic Engineering with Plant Viruses* (Wilson, T. M. A. and Davies, J. W., Eds.) CRC Press, Boca Raton, FL. pp 187-215.

Mullineaux, P. M., Donson, J., Morris-Krsinich, B. A. M., Boulton, M. I., and Davies, J. W. (1984). The nucleotide sequence of maize streak virus DNA. *EMBO J.* **3**: 3063-3068.

Mullineaux, P. M., Guerineau, F., and Accotto, G. P. (1990). Processing of the complementary sense RNAs of Digitaria streak virus in its host and in transgenic tobacco. *Nucl. Acids Res.* **18**: 7259-7265.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bio-assays with tobacco tissue cultures. *Physiol. Plant.* **15**: 473-497.

Murray, E. E., Lotzer, J., and Eberele, M. (1989). Codon usage in plant genes. *Nucl. Acids Res.* **17**: 477-498.

Mushegian, A. R. and Shepherd, R. J. (1995). Genetic elements of plant viruses as tools for genetic engineering. *Microbiological Reviews* **59**: 548-578.

Nagar, S., Pedersen, T. J., Carrick, K. M., Hanley-Bowdoin, L., and Robertson, D. (1995). A geminivirus induces expression of a host DNA synthesis protein in terminally differentiated plant cells. *Plant Cell* **7**: 705-719.

Nakamura, Y., Gojobori, T. and Ikemura, T. (1997). Codon usage tabulated from the international DNA sequence databases. *Nucl. Acids Res.* **25**: 244-245. World Wide Web URL: **http://www.dna.affrc.jp/~nakamura/codon.html**

Navot, N., Zeidan, M., Pichersky, E., Zamir, D. and Czosnek, H. (1992). Use of the polymerase chain reaction to amplify tomato yellow leaf curl virus DNA from infected plants and viruliferous whiteflies. *Phytopathology* **82**: 1199-1202.

Noirot, P., Bagonetti, J. and Novick, R. (1990). Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. *Proc. Natl. Acad. Sci.* **87**: 8560-8564.

Noirot-Gros, M. F., Bidnenko, V., and Ehrlich, S. D. (1994). Active site of the replication protein of the rolling circle plasmid pC194. *EMBO J.* **13**: 4412-4420.

Noris, E., Jupin, I., Accotto, G. P., and Gronenborn, B. (1996). DNA-binding activity of the C2 protein of tomato yellow leaf curl geminivirus. *Virology* **217**: 607-612.

Noueiry, A. O., Lucas, W. J., and Gilbertson, R. L. (1994). Two proteins of a plant DNA virus coordinate nuclear and plasmodesmatal transport. *Cell* **76**: 925-932.

Orozco, B. M. and Hanley-Bowdoin, L. (1996). A DNA structure is required for geminivirus replication origin function. *J. Virol.* **70**: 148-158.

Ow, D. W., Wood, K. V., De Luca, M., de Wet, J. R., Helinski, D. R., and Howell, S. H. (1986). Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. *Science* **234**: 856-859.

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., Beachy, R.N. and Fauquet, C.M. (1996). The role of AV2 ("precoat") and coat protein in viral replication and movement in tomato leaf curl geminivirus. *J. Gen. Virol.* **224**: 390-404.

Palmer, K.E., Schnippenkoetter, W.H. and Rybicki, E.P. (1997). Geminivirus isolation and DNA Extraction. In: G.D. Foster and S. Taylor (Eds) *Plant Virology Protocols*. Series: *Methods in Molecular Biology*. Humana Press. Totowa, NJ. (In Press)

Pascal, E., Goodlove, P. E., Wu, L. C., and Lazarowitz, S. G. (1993). Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. *Plant Cell* **5**: 795-807.

Pascal, E., Sanderfoot, A. A., Ward, B. M., Medville, R., Turgeon, R., and Lazarowitz, S. G. (1994). The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the plant nucleus. *Plant Cell* **6**: 995-1006.

Perl, A., Kless, H., Blumenthal, A., Galili, G. and Galun, E. (1992). Improvement of plant regeneration and GUS expression in scutellar wheat calli by optimization of culture conditions and DNA-microprojectile delivery procedures. *Mol. Gen. Genet.* **235**: 279-284.

Perriman, R., Bruening, G., Dennis, E. S., and Peacock, W. J. (1995). Effective ribozyme delivery in plant cells. *Proc. Natl. Acad. Sci. USA* **92**: 6175-6179.

Peterschmitt, M., Granier, M., Frutos, R. and Reynaud, B. (1996). Infectivity and complete nucleotide sequence of the genome of a genetically distinct strain of maize streak virus from Reunion Island. *Arch. Virol.* **141**: 1637-1650.

Pilartz, M. and Jeske, H. (1992). Abutilon mosaic geminivirus double-stranded DNA is packaged into minichromosomes. *Virology* **189**: 800-802.

Pinner, M. S., Markham, P. G., Markham, R. H. and Dekker, E. L. (1988). Characteristics of maize streak virus: descriptions of strains, symptoms. *Plant Pathol.* **37**: 74-87.

Pinner, M. S., Medina, V., Plaskitt, K. A. and Markham, P. G. (1993). Viral inclusions in monocotyledons infected by maize streak and related geminiviruses. *Plant Pathol.* **42**: 75-87.

Pooma, W. and Petty, I. T. D. (1996). Tomato golden mosaic virus open reading frame AL4 is genetically distinct from its C4 analogue in monopartite geminiviruses. *J. Gen. Virol.* **77**: 1947-1951.

Pooma, W., Gillette, W.K., Jeffrey, J.L. and Petty, I.T.D. (1996). Host and viral factors determine the dispensibility of coat protein for bipartite geminivirus movement. *Virology* **218**: 264-268.

Radicella, J.P., Brown, D., Tolar, L.A. and Chandler, V.L. (1992). Allelic diversity of the maize *B* regulatory gene: different leader and promoter sequences of two *B* alleles determine distinct tissue specificities. *Genes & Development* **6**: 2152-2164.

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.

Rhodes, C. A., Pierce, D. A., Mettler, I. J., Mascarenhas, D. and Detmer, J. J. (1988). Genetically transformed maize plants from protoplasts. *Science* **240**: 204-240.

Rigden, J. E., Dry, I. B., Krake, L. R., and Rezaian, M. A. (1996). Plant virus DNA replication in *Agrobacterium*: Insight into the origins of geminiviruses?. *Proc. Natl. Acad. Sci. USA* **93**: 10280-10284.

Rigden, J. E., Dry, I. B., Mullineaux, P. M. and Rezaian, M. A. (1993). Mutagenesis of the virion sense ORFs of tomato leaf curl geminivirus. *Virology* **193**: 1001-1005.

Roberts, S. and Stanley, J. (1994). Lethal mutations within the conserved stem-loop of African cassava mosaic virus DNA are rapidly corrected by genomic recombination. *J. Gen. Virol.* **75**: 3203-3209.

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

Rochester, D.E., de Paulo, J.J., Fauquet, C.M. and Beachy, R.N. (1994). Complete nucleotide sequence of the geminivirus tomato yellow leaf curl virus, Thailand isolate. *J. Gen. Virol.* **75**: 477-485.

Rogers, S. G., Bisaro, D. M., Horsch, R. B., Fraley, R. T., Hoffmann, N. L., Brand, L., Elmer, J. S., and Lloyd, A. M. (1986a). Tomato golden mosaic virus A component DNA replicates autonomously in transgenic plants. *Cell* **45**: 593-600.

Rogers, S. G., Brand, L. A., Elmer, J. S., Horsch, R. B., and Bisaro, D. M. (1986b). Novel plant vectors. European Patent Specification 86870152.5

Rose, D. J. W. (1978). Epidemiology of maize streak disease. *Ann. Rev. Entomol.* **23**: 259-282.

Rossi, L., Hohn, B., and Tinland, B. (1996). Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. USA* **93**: 126-130.

Rossi, L., Hohn, B. and Tinland, B. (1993). The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localisation signals important for transfer of T-DNA to plants. *Mol. Gen. Genet.* **239**: 345-353.

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

Rybicki, E. P. and Hughes, F. L. (1990). Detection and typing of maize streak virus and other distantly related geminiviruses of grasses by polymerase chain reaction amplification of a conserved viral sequence. *J. Gen. Virol.* **71**: 2519-2526.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanderfoot, A. A., Ingham, D. J., and Lazarowitz, S. G. (1996). A viral movement protein as a nuclear shuttle. The geminivirus BR1 movement protein contains domains essential for interaction with BR1 and nuclear localization. *Plant Physiol.* **110**: 23-33.

Sanderfoot, A. A. and Lazarowitz, S. G. (1995). Cooperation in viral movement: the geminivirus BL1 movement protein interacts with BR1 and redirects it from the nucleus to the cell perifery. *Plant Cell* **7**: 1185-1194.

Sanderfoot, A. A. and Lazarowitz, S. G. (1996). Getting it together in plant virus movement: cooperative interactions between bipartite geminivirus movement proteins. *Trends in Cell Biology* **6**: 353-358.

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.

Saunders, K., Lucy, A., and Stanley, J. (1992). RNA-primed complementary-sense DNA synthesis of the geminivirus African cassava mosaic virus. *Nucl. Acids Res.* **20**: 6311-6315.

Saunders, K. and Stanley, J. (1995). Complementation of African cassava mosaic virus AC2 gene function in a mixed bipartite geminivirus infection. *J. Gen. Virol.* **76**: 2287-2292.

Schalk, H. -J., Matzeit, V., Schiller, B., Schell, J., and Gronenborn, B. (1989). Wheat dwarf virus, a geminivirus of graminaceous plants, needs splicing for replication. *EMBO J.* **8**: 359-364.

Schlappi, M. and Hohn, B. (1992). Competence of immature maize embryos for *Agrobacterium*-mediated gene transfer. *Plant Cell* **4**: 7-16.

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.

Settlage, S. B., Miller, A. B., and Hanley-Bowdoin, L. (1996). Interactions between geminivirus replication proteins. *J. Virol.* **70**: 6790-6795.

Schaffer, R.L., Miller, C.G. and Petty, I.T.D. (1995). Virus and host-specific adaptations in the *BL1* and *BR1* genes of bipartite geminiviruses. *Virology* **214**: 330-338.

Scheffler, B., Franken, P., Schutt, E., Schrell, A., Saedler, H. and Wienand, U. (1994). Molecular analysis of *C1* alleles in *Zea mays* defines regions involved in the expression of this regulatory gene. *Mol. Gen. Genet.* **242**: 40-48.

Shen, W. -H., Das, S. and Hohn, B. (1992). Mechanism of *Ds1* excision from the genome of maize streak virus. *Mol. Gen. Genet.* **233**: 388-394.

Shen, W. -H. and Hohn, B. (1991). Mutational analysis of the small intergenic region of maize streak virus. *Virology* **183**: 721-730.

Shen, W. -H. and Hohn, B. (1992). Excision of a transposable element from a viral vector introduced into maize plants by agroinfection. *Plant J.* **2**: 35-42.

Shen, W. -H. and Hohn, B. (1994). Amplification and expression of the β-glucuronidase gene in maize plants by vectors based on maize streak virus. *Plant J.* **5**: 227-236.

Shen, W. -H. and Hohn, B. (1995). Vectors based on maize streak virus can replicate to high copy numbers in maize plants. *J. Gen. Virol.* **76**: 965-969.

Sherridan, W. F. (1982). Black Mexican sweet corn: its use for tissue cultures. In Sheridan, W. F. (ed.) *Maize for Biological Research.* Plant Molecular Biology Association, Charlottesville, VA. pp 385-388.

Spencer, T. M., Gordon-Kamm, W. J., Daines, R. J., Start, W. G. and Lemaux, P. G. (1990). Bialaphos selection of stable transfromants from maize cell culture. *Theor. Appl. Genet.* **79**: 625-631.

Stanley, J. (1993). Geminiviruses: plant viral vectors. *Curr. Opin. Genet. Dev.* **3**: 91-96.

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., 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.

Stanley, J. and Latham, J. R. (1992). A symptom variant of beet curly top geminivirus produced by mutation of open reading frame C4. *Virology* **190**: 506-509.

Stanley, J., Latham, J. R., Pinner, M. S., Bedford, I., and Markham, P. G. (1992). Mutational analysis of the monopartite geminivirus beet curly top virus. *Virology* **191**: 396-405.

Stanley, J. and Townsend, R. (1986). Infectious mutants of cassava latent virus generated in vivo from intact recombinant clones containing single copies of the genome. *Nucl. Acids Res.* **14**: 5981-5998.

Stanley, J., Townsend, R. and Curson, S. J. (1985). Pseudorecombinants between cloned DNAs of two isolates of cassava latent virus. *J. Gen. Virol.* **66**: 1055-1061.

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., Davis, K. R., and Bisaro, D. M. (1994). Recombinant beet curly top virus genomes exhibit both parental and novel pathogenic phenotypes. *Virology* **200**: 677-685.

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. USA* **88**: 8029-8033.

Stenger, D. C., Stevenson, M. C., Hormuzdi, S. G., and Bisaro, D. M. (1992). A number of subgenomic DNAs are produced following agroinoculation of plants with beet curly top virus. *J. Gen. Virol.* **73**: 237-242.

Suárez-López, P., Martínez-Salas, E., Hernández, P., and Gutiérrez, C. (1995). Bent DNA in the large intergenic region of wheat dwarf geminivirus. *Virology* **208**: 303-311.

Suárez-López, P., and Gutiérrez, C. (1997). DNA replication of wheat dwarf geminivirus vectors: effects of origin structure and size. *Virology* **227**: 389-399.

Sugimoto, K., Otsuki, Y., Saji, S., and Hirochika, H. (1994). Transposition of the maize Ds element from a viral vector into the rice genome. *Plant J.* **5**: 863-871.

Sung, Y. K. and Coutts, R. H. A. (1996). Potato yellow mosaic geminivirus AC2 protein is a sequence non-specific DNA binding protein. *FEBS Letters* **383**: 51-54.

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., Gardiner, W. E., Rushing, A. E., Rogers, S. G., and Bisaro, D. M. (1987). Independent encapsidation of tomato golden mosaic virus A component DNA in transgenic plants. *Plant Mol. Biol.* **8**: 477-484.

Sunter, G., Hartitz, M. D., and Bisaro, D. M. (1993). Tomato golden mosaic virs leftward gene expression: autoregulation of geminivirus replication protein. *Virology* **195**: 275-280.

Sunter, G., Hartitz, M. D., Hormuzdi, S. G., Brough, C. L., and Bisaro, D. M. (1990). Genetic analysis of tomato golden mosaic virus: ORF AL2 is required for coat protein accumulation while ORF AL3 is necessary for efficient DNA replication. *Virology* **179**: 69-77.

Sunter, G., Stenger, D. C., and Bisaro, D. M. (1994). Heterologous complementation by geminivirus AL2 and AL3 genes. *Virology* **203**: 203-210.

Thompson, C. J., Movva, N. R., Tizard, R., Crameri, R., Davies, J. E., Lauwereys, M. and Botterman, J. (1987). Characterization of the herbicide-resistance gene *bar* from *Streptomyces hygroscopicus*. *EMBO J.* **6**: 2519-2523.

Timmermans, M. C. P., Das, O. P., and Messing, J. (1992). Trans replication and high copy numbers of wheat dwarf virus vectors in maize cells. *Nucl. Acids Res.* **20**: 4047-4054.

Timmermans, M. C. P., Das, O. P., and Messing, J. (1994). Geminiviruses and their uses as extrachromosomal replicons. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**: 79-112.

Tinland, B. (1996). The integration of T-DNA into plant genomes. *Trends in Plant Science* **1**: 178-184.

Tinland, B., Hohn, B. and Puchta, H. (1994). *Agrobacterium tumefaciens* transfers single-stranded DNA into the plant cell nucleus. *Proc. Natl. Acad. Sci. USA* **91**: 8000-8004.

Tinland, B., Kouklolíková-Nicola, Z., Hall, M. N. and Hohn, B. (1992). The T-DNA linked-VirD2 protein contains two distinct functional nuclear localization signals. *Proc. Natl. Acad. Sci. USA* **89**: 7442-7446.

Topfer, R., Gronenborn, B., Schell, J., and Steinbiss, H. -H. (1989). Uptake and transient expression of chimeric genes in seed-derived embryos. *Plant Cell* **1**: 133-139.

Turpen, T. H., Reinl, S. J., Charoenvit, Y., Hoffman, S. L., Fallarme, V., and Grill, L. K. (1995). Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. *Bio/Technology* **13**: 53-57.

Ugaki, M., Ueda, T., Timmermans, M. C. P., Viera, J., Elliston, K. O., and Messing, J. (1991). Replication of a geminivirus derived shuttle vector in maize endosperm cells. *Nucl. Acids Res.* **19**: 371-377.

Unger, E., Parsons, R.L., Schmidt, R.J., Bowen, B. and Roth, B.A. (1993). Dominant negative mutants of Opaque2 suppress transactivation of a 22-kD zein promoter by Opaque2 in maize endosperm cells. *Plant Cell* **5**: 831-841.

Usha, R., Rohll, J. B., Spall, V. E., Shanks, M., Maule, A. J., Johnson, J. E., and Lomonossoff, G. P. (1993). Expression of an animal virus antigenic site on the surface of a plant virus particle. *Virology* **197**: 366-374.

Vain, P., McMullen, M. D. and Finer, J. J. (1993). Osmotic pretreatment enhances particle bombardment-mediated transient and stable transformation of maize. *Plant Cell Rep.* **12**: 84-88.

van Schaik, N. W. and Brink, R. A. (1959). Transpositions of *Modulator*, a component of the variegated pericarp allele in maize. *Genetics* **44**: 725-738.

Vasil, V., Marcotte, W.R., Rosenkrans, L., Cocciolone, S.M., Vasil, I.K., Quatrano, R.S. and McCarty, D.R. (1995). Overlap of Viviparous1 (VP1) and abscisic acid response elements of the *Em* promoter: G-box elements are sufficient but not necessary for VP1 transactivation. *Plant Cell* **7**: 1511-1518.

von Arnim, A., Frischmuth, T., and Stanley, J. (1993). Detection and possible functions of African cassava mosaic virus DNA B gene products. *Virology* **192**: 264-272.

von Arnim, A. and Stanley, J. (1992). Inhibition of African cassava mosaic virus systemic infection by a movement protein from the related geminivirus tomato golden mosaic virus. *Virology* **187**: 555-564.

Waigman, E. and Zambryski, P. (1995) Tobacco mosaic virus movement protein-mediated protein transport between trichome cells. *Plant Cell* **7**: 2069-2079.

Walters, D. A., Vetsch, C.S., Potts, D. E., and Lundquist, R. C. (1992). Transformation and inheritance of a hygromycin phosphotransferase gene in maize plants. *Plant Mol. Biol.* **18**: 189-200.

Ward, A., Etessami, P., and Stanley, J. (1988). Expression of a bacterial gene in plants mediated by infectious geminivirus DNA. *EMBO J.* **7**: 1583-1587.

Waters, V. L. and Guiney, D. G. (1993). Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. *Mol. Microbiol.* **9**: 1123-1130.

Wehrmann, A., Van Vliet, A., Opsomer, C., Botterman, J. and Schulz, A. (1996). The similarities of *bar* and *pat* gene products make them equally applicable for plant engineers. *Nature Biotechnology* **14**: 1274-1278.

White, J., Chang, S.-Y. P., Bibb, M.J. and Bibb, M. J. (1989). A cassette containing the *bar* gene of *Streptomyces hygroscopicus*: a selectable marker for plant transformation. *Nucl. Acids. Res.* **18**: 1062.

Wirtz, U., Osborne, B. and Baker, B. (1997). *Ds* excision from extrachromosomal geminivirus vector DNA is coupled to vector DNA replication in maize. *Plant Journal* **11**: 125-135.

Woolston, C. J., Barker, R., Gunn, H., Boulton, M. I. and Mullineaux, P. M. (1988). Agroinfection and nucleotide sequence of cloned wheat dwarf virus DNA. *Plant Mol. Biol.* **11**: 35-43.

Woolston, C. J., Reynolds, H. V., Stacey, N. J., and Mullineaux, P. M. (1989). Replication of wheat dwarf virus DNA in protoplasts and analysis of coat protein mutants in protoplasts and plants. *Nucl. Acids Res.* **17**: 6029-6041.

Wright, E. A. (1995). *Transcription of the maize streak virus genome.* Ph.D. Thesis, University of East Anglia, Norwich, United Kingdom.

Xie, Q., Sanz-Burgos, A. P., Hannon, G. J., and Gutiérrez, C. (1996). Plant cells contain a novel member of the retinoblastoma family of growth regulatory proteins. *EMBO J.* **15**: 4900-4908.

Xie, Q., Suárez-López, P., and Gutiérrez, C. (1995). Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant DNA virus: requirement for efficient viral DNA replication. *EMBO J.* **14**: 4073-4082.

Zeidan, M. and Czosnek, H. (1991). Acquisition of tomato yellow leaf curl virus by the whitefly *Bemisia tabaci*. *J. Gen. Virol.* **72**: 2607-2614.

Zhan, X., Richardson, K., Haley, A., and Morris, B. A. M. (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.

Zupan, J. R. and Zambryski, P. (1995). Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiol.* **107**: 1041-1047.