

# Virus-Derived ssDNA Vectors for the Expression of Foreign Proteins in Plants

Edward P. Rybicki and Darrin P. Martin

**Abstract** Plant viruses with ssRNA genomes provide a unique opportunity for generating expression vehicles for biopharming in plants, as constructs containing only the replication origin, with the replication-associated protein (Rep) gene provided in *cis* or in *trans*, can be replicationally amplified in vivo by several orders of magnitude, with significant accompanying increases in transcription and expression of gene(s) of interest. Appropriate replicating vectors or replicons may be derived from several different generic geminiviruses (family *Geminiviridae*) or nanoviruses (family *Nanoviridae*), for potential expression of a wide range of single or even multiple products in a wide range of plant families. The use of vacuum or other infiltration of whole plants by *Agrobacterium tumefaciens* suspensions has allowed the development of a set of expression vectors that rival the deconstructed RNA virus vectors in their yield and application, with some potential advantages over the latter that still need to be explored. Several modern applications of ssDNA plant vectors and their future potential will be discussed.

---

E. P. Rybicki · D. P. Martin

Institute of Infectious Disease and Molecular Medicine,  
Cape Town, South Africa

E. P. Rybicki (✉)

Department of Molecular and Cell Biology,  
University of Cape Town, PB Rondebosch, Cape Town,  
Western Cape 7701, South Africa  
e-mail: ed.rybicki@uct.ac.za

## Contents

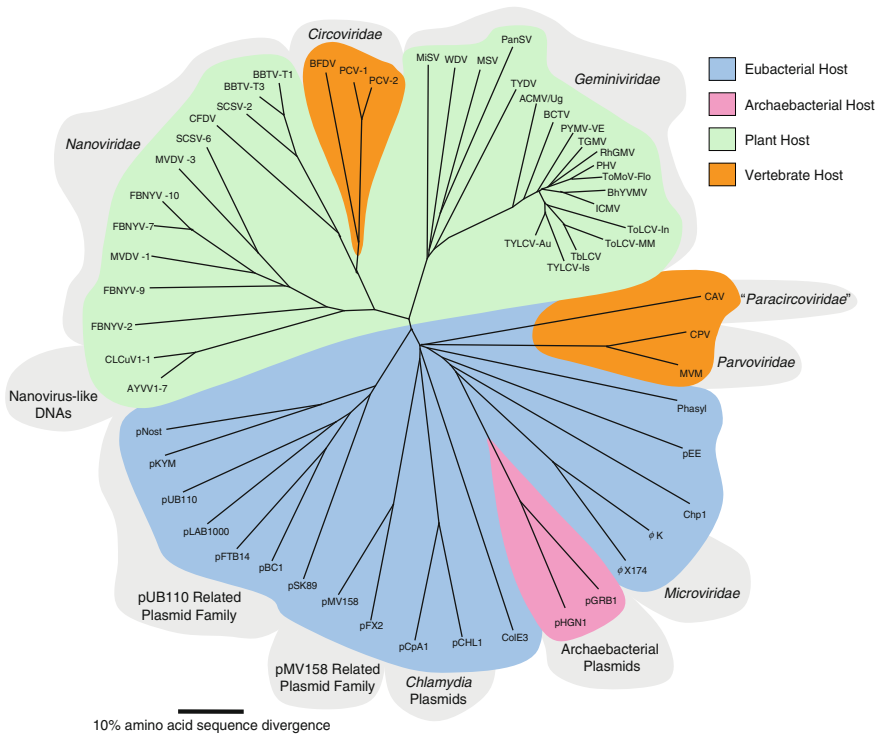
1	Introduction.....	20
1.1	ssDNA Plant Viruses.....	20
1.2	Replication of the Viruses.....	25
2	ssDNA Viruses as Vectors: Historical .....	27
2.1	Geminiviruses .....	27
2.2	Nanoviruses.....	32
3	ssDNA Viruses as Vectors: Current Practice.....	32
4	Conclusion and Future Prospects.....	37
	References.....	40

## 1 Introduction

### 1.1 ssDNA Plant Viruses

Single-stranded (ss-) DNA viruses are spread among all domains of life, infecting prokaryotes (*Microviridae*, *Inoviridae*), animals (*Anelloviridae*, *Circoviridae* and *Parvoviridae*), as well as plants (*Geminiviridae*, *Nanoviridae*). All of these viruses, as well as a number of plasmids of bacteria and phytoplasmas, share a common replication mechanism: this is rolling circle replication (RCR), mediated by a virus or plasmid-encoded replication-associated protein. They also probably share a common origin, which may well be ancient and bacterial (Koonin and Ilyina 1992), as essential elements of these protein sequences appear to have been conserved across all sequenced examples (Fig. 1). The increasing potential of the plant viruses as vectors may well extend to use of other members as well—and in this respect it is interesting that a variety of phytoplasma-derived extrachromosomal DNAs encode geminivirus-like Reps (Nishigawa et al. 2001; Rekab et al. 1999); that a geminivirus-like mycovirus has been very recently discovered (Yu et al. 2010), and that a variety of algal- and otherwise-derived circo- and geminivirus-like sequences have been found in metagenomic screens of oceanic water samples (Nawaz-ul-Rehman and Fauquet 2009; Rosario et al. 2009).

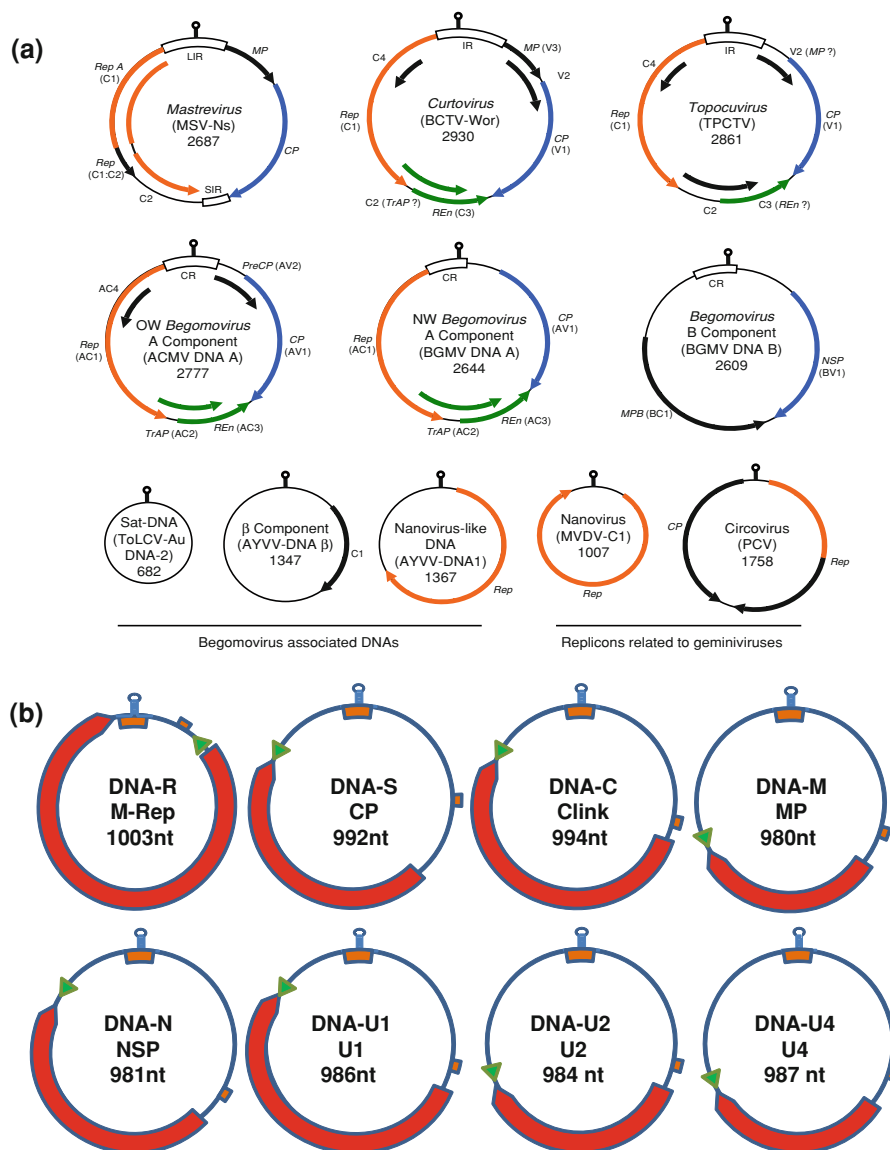
Except for the parvoviruses, which have linear genomes, the ssDNA plant viruses are like all other ssDNA viruses in having circular genomes. They belong to the two taxonomic families *Geminiviridae* and *Nanoviridae*. The following section discusses basics of the molecular biology of the viruses that are important to understand for any subsequent discussion of their use as vectors for high-level expression of proteins in plants.



**Fig. 1** The possible evolutionary relationships of rolling circle replication (RCR) domains within the replication-associated proteins encoded by a number of ssDNA replicons. Based on a published alignment of the ~ 80 amino acids spanning the three RCR motifs from a diverse group of ssDNA replicons (Koonin and Ilyina 1992), this unrooted dendrogram was constructed by the neighbour joining method (Saitou and Nei 1987) using 1,000 bootstrap iterations. While bootstrap support for nodes near the centre of the dendrogram is relatively poor (all nodes with less than 30% support have been collapsed), support for all the presented nanovirus, circovirus and geminivirus groupings was in excess of 50%. There was 48% bootstrap support for the pUB110 plasmid family being more closely related to the geminivirus/circovirus/nanovirus grouping than to other prokaryotic replicons

### 1.1.1 Geminiviruses

Geminiviruses are divided into four genera; namely, *Mastrevirus*, *Begomovirus*, *Curtovirus* and *Topocuvirus*, based on genetic organisation and insect vectors. Mastre-, curto- and *topocuviruses* all have single-component genomes of 2.7–2.8 kb in size, while begomoviruses may have single- or two-component genomes, each of the same size (Stanley et al. 2005). Their genomes all have a characteristic organisation as shown (Fig. 2a), with ORFs diverging in virion (= V) and complementary (= C) senses from an intergenic region which contains elements of both V- and C- sense promoters, and the origin of replication (= *ori*). The *ori* in all cases contains a characteristic stem-loop forming sequence, with the



invariant nonanucleotide sequence TAATATTAC in the loop forming the genomic (+) strand origin of replication. A smaller intergenic region is found in mastreviruses, at the convergence of V and C-sense ORFs.

Other ssDNAs associated with certain Old World (OW) begomovirus infections (see Fig. 2a) are satellites (Dry et al. 1997; Mansoor et al. 1999; Saunders and

◀**Fig. 2 a** The genomic arrangements of geminiviruses and some other dependent replicons. Origins of (+) strand synthesis are indicated by the stem-loop symbol at 12 o'clock. OW Old World. NW New World. IR intergenic region. LIR long/large intergenic region. SIR short/small intergenic region. CR common region that is nearly identical in the A and B components of bipartite begomoviruses. Genes and parts of genes in the same colour either express or potentially express proteins with detectable sequence homology. *Rep* replication-associated/initiator protein gene found in all geminiviruses, circoviruses and nanoviruses. *RepA* variant of *Rep* potentially expressed from an unspliced complimentary sense transcript and encoding a multifunctional regulatory protein that is unique to the mastreviruses. *MP* movement protein gene of mastreviruses and curtoviruses. *CP* coat protein gene. *REn* replication enhancer gene found in begomoviruses and curtoviruses. *TrAP* transcription activator protein gene found in begomoviruses. *PreCP* pre-coat protein gene. *NSP* nuclear shuttle protein gene. *MPB* movement protein gene found on the B component of bipartite begomoviruses. **b** Genomic arrangement of a nanovirus: Faba bean necrotic stunt virus (FBNSV). Figure adapted from Grigoras et al. (2009). Genome organization of FBNSV. The eight proven genomic DNAs are designated according to current nomenclature (Vetten et al. 2005). The common region and internal stem-loop (CR-SL) sequence is shown in brown, with stem-loop above it. Protein-encoding ORFs are represented by red arrows. TATA boxes (brown) and polyadenylation signals (green arrows) are shown for each component. Sizes of the individual DNA components are shown within the components. M-Rep = Master Rep (33.2 kDa); CP = coat protein (19.2 kDa); Clink = cell cycle link protein (19.9 kDa); MP = movement protein (13.2 kDa); NSP = nuclear shuttle protein (17.6 kDa); U1- U4 = proteins of unknown function

Stanley 1999; Saunders et al. 2000), and one at least is derived from a nanovirus, but the  $\beta$  component DNA (Saunders et al. 2000) is a unique satellite DNA which can massively enhance the virulence of its associated begomovirus.

Geminiviruses are unique in having geminate or doubled particles, each consisting of two partially assembled T = 1 capsids joined at the missing pentameric vertex, and each encapsidating a single genome component (Bottcher et al. 2004; Zhang et al. 2001) (Fig. 3). There is a severe size constraint on the sizes of genomes that can be encapsidated, with ssDNAs of around 1.3–2.8 kb being strongly favoured, although multiples of the normal genome size can be incorporated into multimeric particles (Casado et al. 2004; Frischmuth et al. 2001).

The geminiviruses are obligately transferred between plants by insect vectors in natural infections: while some of the viruses have been mechanically transmitted, this is usually done only with great difficulty and recourse to methods which mimic injection via piercing mouthparts. In some cases it has been shown that cloned dsDNAs are infectious, but this is only true for dicot hosts. The begomoviruses are vectored by one species of whitefly (*Bemisia tabaci*), and may be the largest single group of characterised plant viruses; they have a world-wide distribution, while only infecting dicotyledonous plants. Interestingly, New World (NW) begomoviruses are very predominantly two-component, while Old World (OW) viruses may be one- or two-component and are often associated with satellite DNAs, usually  $\beta$ -components (see Briddon and Stanley 2006; Nawaz-ul-Rehman and Fauquet 2009). Mastreviruses are vectored by different species of cicadellid

leafhoppers, and have an “Old World” distribution only: most characterised members infect grasses, and Maize streak virus (MSV) is important in maize in Africa and Wheat dwarf virus (WDV) in cereals in Eurasia. However, some mastreviruses have adapted to infect dicots, and Tobacco yellow dwarf (TYDV), Bean yellow dwarf (BeYDV) and Chickpea chlorotic dwarf viruses (ChCDV) are increasingly being implicated in crop diseases worldwide (Thomas et al. 2010). Curtoviruses are vectored by leafhoppers, and topocuviruses by planthoppers. Both curto and topocuviruses have a limited distribution, only a few identified members, and infect only dicots (Stanley et al. 2005).

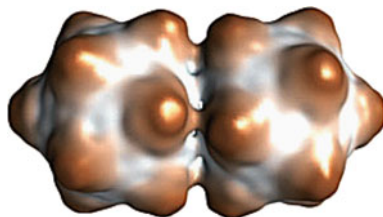
While specific geminiviruses have relatively limited host ranges, experimental host range may be considerably widened: for example, while African cassava mosaic begomovirus (ACMV) does not naturally infect tobacco (*N. tabacum*) or most certainly not maize, the A genome component replicates autonomously in protoplasted cells of both plants after transfection of partially dimeric DNA constructs (Paszukowski et al. 1993). Of interest to its potential use as a vector in heterologous systems, however, was that insertion of 1,165 bp of foreign DNA downstream of the *cp* promoter abolished replication in maize cells, but not in tobacco.

### 1.1.2 Nanoviruses

There are two genera of nanoviruses; namely, *Nanovirus* and *Babuvirus*. The former includes a number of viruses infecting mainly legumes, such as Faba bean necrotic yellows virus (FBNYV), Milk vetch dwarf virus (MDV), and Subterranean clover stunt virus (SCSV). The latter genus consists of one well-described species of virus, namely Banana bunchy top virus (BBTV), and two more putative members in Abaca bunchy top virus (ABTV) and Cardamom bushy dwarf virus (CBDV) ([www.ictvonline.org](http://www.ictvonline.org)). Viruses in both genera are vectored by aphids.

The viruses have genomes consisting of multiple components of around 1 kb in size. These all have the same basic genetic organisation, with a single ORF in the virion (V) sense adjacent to a “common region stem-loop” or CR-SL sequence, very similar to geminiviruses, shared between all components of a given virus (Fig. 2b). Five genome components are common to nano- and babu-viruses: DNA-R encodes a master Rep (MRep) protein; DNA-S encodes the CP; DNA-C encodes Clink, a cell cycle affecting protein which is known to interact with retinoblastoma (Rb) protein analogues (Lageix et al. 2007); DNA-M encodes a movement protein, and DNA-N a nuclear shuttle protein (NSP). There are three other DNAs (DNAs-U1, -U2, and -U4), which encode as yet uncharacterised proteins, found in the nanoviruses FBNYV and MDV. Another DNA (DNA-U3) has been identified from the babuviruses BBTV and ABTV (Grigoras et al. 2009; Gronenborn, 2004).

It is presumed that nanovirus particles are complete T = 1 icosahedra, encapsidating only one genome component each (Harding et al. 1991; Thomas and Dietzgen 1991).



**Fig. 3** Three-dimensional reconstruction of a geminate particle from cryoelectron microscopy data. Geminate particle of Maize streak virus: note characteristic doubled  $T = 1$  particles, paired at missing vertex. [Adapted from Shepherd et al. (2010). Reconstruction by Kyle Dent and Trevor Sewell; copyright by the Electron Microscope Unit, University of Cape Town]

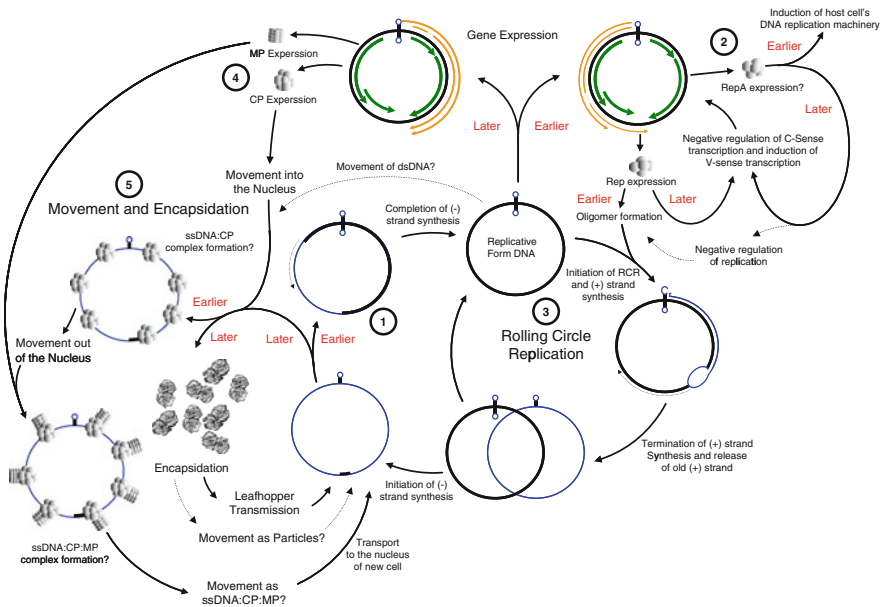
## 1.2 Replication of the Viruses

An understanding of the process of replication of the viruses aids in understanding vector construction, and possibly the uses to which elements of the virus genomes may be put. The following section discusses in some detail aspects of the replication of different gemini- and nanoviruses relevant to the purposes of this review.

The virions are injected via their vector insect's piercing mouthparts directly into cells of the host plant, where they presumably wholly or partially uncoat, and the single-stranded genomes are localised via cell machinery into the nuclei in a process that may be mediated by associated coat protein (CP) molecules. The ssDNA is converted to episomal dsDNA by host polymerases, most probably those involved in repair functions. Early transcription produces mainly mRNA for replication-associated protein (Rep), which serves to nick the double-stranded replication intermediates at the *ori* so as to allow covalent binding of Rep to the free 5' residue. This allows single-strand displacement (rolling circle replication, or RCR) by a processive DNA polymerase adding onto the free 3' end of genomic DNA. The Rep then acts to recircularise the displaced linear ssDNA genomes by ligation, aided by their circularisation due to annealing of the stem-loop region. Encapsidation of genomes is probably due to interaction of transiently sssDNA with accumulating CP, which self-localises to the nucleus via nuclear localisation signals, and subsequent sequestration by assembly (reviewed by Gronenborn 2004; Gutierrez 1999, 2000; Hanley-Bowdoin et al. 2000; Palmer and Rybicki 1998).

### 1.2.1 Geminivirus Replication

For all geminiviruses, the 5'-TAATATT↓AC-3' sequence in the mastre-, topocu-, curto- and single component begomovirus large intergenic region (LIR) and two-component begomoviral common region (CR) genomic strand of a dsDNA replicative intermediate is the site of action of Rep: this nicks at the position shown, and covalently binds the 5'-A to allow strand displacement by processive (presumably repair) host polymerase, and subsequent RCR. In all geminiviruses



**Fig. 4** Summary of the geminivirus infection and replication process, with Maize streak mastrevirus as an example. Early during an infection following the synthesis of a dsDNA replicative form; RF (a) RepA is most likely expressed and induces a cellular state in which viral DNA replication can occur (b) Rep is also expressed early and rolling circle replication begins (c) At a later point in the infection process, following genome amplification and possibly Rep and/or RepA induction of the V-sense promoter, MP and CP are expressed (d) and movement and encapsidation occur (e) Represented here is movement of unencapsidated ssDNA but it should be noted that it is possible that dsDNA and/or encapsidated ssDNA may also be moved either cell to cell or systemically within the phloem of plants. Whereas the involvement of MSV CP and MP in movement has been demonstrated (Kotlizky et al. 2000; Liu et al. 1997, 1999), the mechanics of the process are obscure and certain details have been borrowed from a cell to cell movement model proposed for the begomovirus, SqLCV (Lazarowitz and Beachy, 1999; Qin et al. 1998). While the probable timing of events is indicated, it is unlikely, for example, that absolutely no MP and CP expression occurs during the earlier stages of the infection process. ssDNA is represented by blue lines, dsDNA by bold black lines and RNA by orange lines

except the mastreviruses, Rep is expressed from a single (C1) ORF. In mastreviruses, Rep is expressed from the post-translationally spliced transcript of ORFs C1 and C2; RepA is a variant of Rep expressed from the C1 ORF only in an unspliced complementary sense transcript, and encodes a multifunctional regulatory protein that is unique to the mastreviruses. These are the only replication-associated proteins in mastreviruses. However, there is a REn or replication enhancer gene found in begomoviruses and curtoviruses which also has a homologue in topocuviruses, but its function in these viruses is not proven. Similarly, the well-characterised TrAP or transcription activator protein gene found in begomoviruses also has a possible homologue in curtoviruses (Hormuzdi and Bisaro 1995). The PreCP or pre-coat protein gene in OW begomoviruses is apparently involved in ssDNA accumulation



(Wartig et al. 1997), and in curtoviruses the V2 gene is apparently involved in regulation of ssDNA accumulation (Hormuzdi and Bisaro 1993; Stanley et al. 1992).

The actual process of genome replication is illustrated for mastreviruses in Fig. 4 given that these viruses have recently received the most attention as pharming vectors and will be featured in this review. However, the only DNA sequences required for a vector are *rep* and the LIR/CR for begomo, curto and topocuviruses, and *rep* plus the LIR and short intergenic region (SIR) for mastreviruses, if the *rep* expression is driven by a heterologous promoter (Palmer and Rybicki 1997). Host specificity is also less important than for viruses, as expression and replication may occur in a wide range of cell types without a movement constraint: for example, Bean yellow dwarf mastrevirus (BeYDV) with Maize streak virus (MSV) V region genes could replicate in and produce virus particles in *Nicotiana tabacum* protoplasts, but could not infect plants (Liu et al. 1999); the ACMV A genome component replicated successfully in protoplasted maize cells (Paszkowski et al. 1993).

Indeed, a variety of geminivirus-derived replicons also replicate in *A. tumefaciens*, purely as a result of Rep being expressed from cryptic bacterial promoters in the geminiviral LIR in partially redundant clones (Selth et al. 2002). There is also evidence that Ageratum yellow vein begomovirus (AYVV) replicons can be released from monomeric clones in a M13-derived vector in *E. coli* (Wu et al. 2007), and that Indian mung bean yellow mosaic begomovirus (IMBYMV) DNA-A replicates autonomously in budding yeast cells (Raghavan et al. 2004).

### 1.2.2 Nanovirus Replication

The overall process is much the same as with geminiviruses, with all nanoviruses having a “master Rep” which initiates RCR in the CR-SL region in the conserved loop sequence TA(T/G)TATT↓AC (T in babuviruses; G in nanoviruses). However, whereas with geminiviruses the Rep or RepA proteins interact with cell cycle regulatory proteins as well (see Gronenborn 2004), in familial nanoviruses this function is fulfilled by Clink, which has been shown to bind an Rb analogue as well as SKP1, part of the ubiquitin-protein turnover pathway (Aronson et al. 2000; Lageix et al. 2007).

## 2 ssDNA Viruses as Vectors: Historical

### 2.1 Geminiviruses

The use of viruses for the expression of foreign proteins in plants in fact predates the concept of using plants to express pharmaceutically-relevant proteins: while the first “pharming” was done by Hiatt et al. (1989) with the successful expression of a whole monoclonal antibody (mAb) in transgenic tobacco, the first use of a

plant viral vector was by Takamatsu et al. (1987), who expressed a bacterial chloramphenicol acetyltransferase (CAT) gene in tobacco (*Nicotiana tabacum*) plants via inoculation of in vitro transcribed recombinant Tobacco mosaic virus (TMV) genomes. There was at the time, however, a significant problem in using the cloned genomes of plant DNA viruses to infect plants, which retarded their use as gene vectors: this was the general lack of infectivity of the DNA. This was surmounted in an ingenious use of the ability of *A. tumefaciens* to transfer DNA into plant cells by Grimsley et al. (1986), who demonstrated that complete head-to-tail dimeric clones of the dsDNA Cauliflower mosaic virus (CaMV) genomic DNA in *A. tumefaciens* were infectious if inoculated onto turnip stems, a process that became known as “agroinoculation” or sometimes “agroinfection”. The technique relied on the fact that complete or even partially dimeric tandem viral genomic inserts in host cells could recombine to form circular genomes—which were infectious. This was soon applied to geminiviruses, with the demonstration that despite the fact that maize is not a host for *A. tumefaciens*, tandem copies of Maize streak virus (MSV) DNA were infectious when the bacterium was injected into maize seedlings—the first demonstration of the infectivity of cloned MSV (Grimsley et al. 1987). The technology subsequently spread swiftly through the geminivirus research community, with a number of demonstrations of infectivity for various generic geminiviruses (Boulton 1995; Briddon et al. 1989; Donson et al. 1988; Rochester et al. 1990).

Much of this early work was, despite its promise for high-value protein expression, mainly aimed at using geminiviruses as tools for probing viral replication and transport mechanisms, and host functions, and has been the subject of a number of comprehensive reviews (Stanley 1993; Timmermans et al. 1994).

The first use of geminiviruses to vector other genes in plants was by Ward et al. (1988), using African cassava mosaic virus (ACMV) with a coat protein (CP) gene replacement to vector CAT expression in *Nicotiana benthamiana*, and Hayes et al. (1988), who showed that Tomato golden mosaic virus (TGMV) DNA-A with the CP gene replaced could be used to express a bacterial enzyme—neomycin phosphotransferase (NPT)—either via release of a replicating DNA (replicon) from a transgenic tobacco with a partially dimeric DNA-derived insert, or via agroinoculation of transgenic tobacco with a partially dimeric DNA-B insert. In the latter case, the B genome would be mobilised by the self-replicating A genome, and in turn allow systemic spread of the latter by expression of movement proteins. However, the first properly systematic investigation of a recombinant geminivirus for expression of a foreign gene was probably by Hayes et al. (1989). This follow-up study showed that both NPT and  $\beta$ -glucuronidase (GUS) could be stably expressed at elevated levels relative to non-replicating controls, in plants transgenic for partially dimeric coat protein gene replacement TGMV DNA-A genomes, with or without the CP promoter being replaced by the CaMV 35S promoter, even if the DNA-A was significantly increased in size (Hayes et al. 1989). The constructs could also be used as “agroinfectious” vectors—introduced via recombinant *A. tumefaciens* inoculation—in the presence of DNA-B, but with a significant chance of vector deletion back to the wild-type size. The amplification

of gene expression relative to the same constructs in single-gene copy transgenics under the 35S promoter was about 100-fold for both NPTII and GUS.

The potential use of mastreviruses for expressing foreign proteins was demonstrated by Kammann et al. (1991) who first established that it was possible to make a “shuttle vector” containing the *E. coli* ColE1 origin of replication and parts of the Wheat dwarf virus (WDV) genome, that could replicate in both bacteria and wheat cells. This was expanded by Matzeit et al. (1991), who demonstrated that replacement of the *cp* gene in the monomeric WDV shuttle vector by bacterial neomycin phosphotransferase, chloramphenicol acetyltransferase, and  $\beta$ -galactosidase genes did not affect the ability of the genome to replicate in protoplasts derived from *Triticum monococcum* cells, even though the  $\beta$ -galactosidase gene doubled the size of the vector.

A very interesting use for mastreviruses was pioneered by Laufs et al. (1990), who used Wheat dwarf virus (WDV) to vector the maize transposon Activator (Ac) and derivatives into wheat, maize and rice protoplasts. Rapid and efficient excision of Ac was detected only if the virus could replicate; excision of non-autonomous Ds elements was also possible if transposase activity was provided in *trans*. This was followed by evidence that agroinfection of a MSV-derived replicon could be used to vector the Ds element into whole maize plants, with proof of transposition being restoration of a wild-type MSV replicon and subsequent symptom expression (Shen and Hohn 1992). Sugimoto et al. (1994) then showed that Miscanthus streak mastrevirus (MiSV) could be used to vector Ds carrying a payload of a hygromycin phosphotransferase gene into rice protoplasts, which resulted in stable nuclear transformation. Proof that replication of the vector was required for efficient excision and transposition was given by Wirtz et al. (1997), who used WDV to vector Ds into maize protoplasts. The first authors raised the point at the time of the potential of combining a transposon with a geminiviral replicon for plant molecular genetic engineering; however, this potential appears to have been superseded as other techniques such as *A. tumefaciens* transformation improved.

An important milestone for the use of mastreviruses was the demonstration by Shen and Hohn (1994) that MSV-based replicons could be used via agroinfection to vector expression of  $\beta$ -glucuronidase (GUS) in maize plants via agroinfection. Expression of GUS was amplified by up to tenfold compared to use of a non-replicating vector; however, the vectors replicated only in cells of leaf primordia which had been exposed to recombinant *A. tumefaciens* during the infection phase, and did not spread systemically.

Two investigations which stand out from the late 1990s were those of Palmer et al. (1999) using MSV, and by others using the mastrevirus Tobacco yellow dwarf virus; TYDV (Needham et al. 1998). In the first, MSV-derived replicons were generated in suspension-cultured Black Mexican sweetcorn cells after biolistic co-bombardment with partially dimeric DNA constructs encoding a variety of constructs expressing the bialaphos resistance gene (*bar*) for initial selection purposes. The effect on replication of increasing the genomic size was tested by including the CaMV 35S promoter plus *bar* only; CaMV 35S Pr and the 550-bp

maize *adh* I intron and 68-bp TMV  $\Omega$  RNA leader sequences upstream of *bar*; and a fusion between *bar* and *E. coli* glutathione reductase (*gor*) genes. The three recombinant viral vectors ranging in size from 2.7 to 4.8 kb replicated efficiently; replicons were structurally stable, replicated to copy numbers of over 500 per haploid genome, and were detected for up to 3 years after introduction, and 2 years without selection. Expression levels of recombinant proteins were between 4 and 80 times the best level achieved by conventional transformation, depending on the protein. The TYDV work looked at expression of GUS from the *uidA* gene incorporated into an integrated partially dimeric TYDV-derived construct which produced replicons in regenerated plants: this was interesting at the time because of the success in incorporating a complete and normally-expressing *rep* gene in a transgenic plant; it is significant because of later developments involving inducible TYDV-based replicons in transgenic plants.

The foundation for the use of geminiviral replicons for molecular pharming was probably cemented in the early 2000s, primarily with the publication of several papers describing the use of Bean yellow dwarf mastrevirus (BeYDV). In the first, Mor et al. (2003) describe a two-component system tested in biolistically transfected suspension-cultured tobacco NT-1 cells, with Rep supplied in *trans* with a dependent replicon containing the *cis*-acting sequences LIR and SIR (see Fig. 2a), and expressing jellyfish green fluorescent protein (GFP) or GUS genes under control of a CaMV 35S Pr. They noted significant enhancement of indicator gene expression—up to 40-fold—and established that the two-component system worked well, as well as showing that only the mastreviral Rep, and not RepA as well, was necessary for effective expression. In the second, and purportedly the first vaccine-related application, Hefferon and Fan (2004) essentially reiterated the first group's results, and also reported use of a BeYDV-derived replicon for plant codon use optimised *Staphylococcus* enterotoxin B protein (SEB) gene expression as a vaccine model: they reported up to 20-fold increases in SEB concentration for NT-1 cells containing replicating vector compared to cells bombarded with the *rep*-version only. However, the vaccine connection was unconvincing, as yields were not given, nor was the final vector construct described, and no immunological evidence was offered. In their second paper, Hefferon et al. (2004) used a BeYDV replicon in NT-1 tobacco and Ad-2 *Arabidopsis thaliana* suspension-cultured cells to show enhanced production ( $\sim 20\times$ ) of, and apparently correct glycosylation of, a MAP-K protein from *Arabidopsis* and Potato virus X CP, respectively. They also purified sufficient MAP-K from cells for biochemical analyses. In what is possibly the only use of a monopartite begomovirus for the purpose of foreign gene expression, Tamilselvi et al. (2004) described an Ageratum yellow vein virus (AYVV)-derived “*E. coli*—plant shuttle vector”, which was used to electrophoretically transfect *N. benthamiana* mesophyll-derived protoplasts, and biolistically transfect tobacco BY2 suspension-cultured cells. The authors describe the plasmid vector—which contained a monomeric *cp*<sup>−</sup> AYVV genome, and a pUC19 backbone—as allowing extrachromosomal replication in plant cells as well as in *E. coli*, under ampicillin selection.

While the investigations described above laid a useful foundation for use of gemini- and other ssDNA viruses as gene vectors for pharmaceutical production, there were important limitations to their application. First, while whole-plant expression of proteins was occasionally possible via replicating gemini-virus genomes in transgenic plants, it proved very difficult to successfully regenerate plants with integrated replication-competent geminiviral genomes. This is not surprising in retrospect, given the potential for Rep and other viral proteins for interacting with cell cycle regulatory proteins (see Gutierrez et al. 2004; Hanley-Bowdoin et al. 2000, 2004). However, it meant that either only very rare integration events with few viral genomes could be used to express foreign proteins, or that transgenic cell or callus cultures had to be used (see Palmer et al. 1999). Second, delivery of the viruses into plants for transient expression either via agroinfection or as cloned dsDNA was necessarily only into a limited number of sites per plant: this meant that the viral genomes had to be able to spread out of inoculated tissue in order to be effective as vectors, with all of the limitations in payload this necessitated, given size limitations for cell-to-cell transport and the necessity of incorporating movement protein gene(s). Another problem that came to light when *trans*-complementing viral genomes—each encoding a function necessary for the other component—were investigated as infectious vectors, was that recombination back to wild type was very efficient and very quick (Palmer and Rybicki 2001). In such a case, autonomously-replicating and spreading wild-type genomes would quickly come to dominate the infection, reducing transgene expression to zero. Again, this is obvious in retrospect, as it is known that several different begomoviral B-genomes can be associated with a single A-genome in natural infections, after apparently obtaining the requisite CR sequence by recombination so as to allow strain-specific recognition by the A genome-encoded Rep (Briddon et al. 2010).

The potential of using geminiviruses for high-level transgene expression therefore mostly languished, while the use of RNA viruses increased (see Rybicki, 2009a; Yusibov et al. 2006). However, possibly the most successful application of geminivirus-mediated transgene expression technology was reported by Zhang and Mason (2006), who used an alcohol-inducible two-component (*rep* + replicon) system in doubly-transformed tobacco NT-1 cells and in transgenic potato plants, to express GFP and the Norwalk norovirus (NV) CP. Rep expression was driven by the *alcA* Pr, under control of a CaMV 35S Pr-driven *alcR* or alcohol response protein gene, while replicons were flanked by LIRs, with gene expression driven by a 2xCaMV 35S Pr with a Tobacco etch potyvirus (TEV) 5' UTR translational enhancer. In doubly-transformed NT-1 cells induced with 0.1% ethanol, there was significant amplification of GFP and NV CP expression up to 4 days post induction. In potato plants, expression peaked at 7 days and declined thereafter, possibly due to silencing. It was evident that protein expression was not proportional to mRNA expression, indicating that saturation of protein synthetic machinery could result from too high a level of transgene expression.

## 2.2 Nanoviruses

By contrast to the case for geminiviruses, the use of cloned nanovirus genomes was minimal for any purpose other than discovery of genome functions, given that the full genome of any nanovirus was only assembled by 2009 (Grigoras et al. 2009), for a new nanovirus called Faba bean necrotic stunt virus (FBNSV) (see also Fig. 2b). Early work that showed potential, however, was that of Dugdale et al. (1998), who investigated the potential of BBTV promoters for enhancement of gene expression in biolistically bombarded tobacco and banana cell cultures, and in transgenic tobacco and banana. Interestingly, expression was best in tobacco cells, with two promoters (BBTV DNA-2 and -6) being twofold better and as good as the CaMV 35S Pr. In banana cells, activities of Prs from DNAs 4 and 5—corresponding to movement protein (M) and Clink, respectively (Wanitchakorn et al. 2000)—were highest, and DNA 1 (Rep) the least, which was considered appropriate given the requirement for larger amounts of MP and Clink than for Rep.

Possibly the first use of replicons derived from a nanovirus was the use of parts of the Faba bean necrotic yellows virus (FBNYV) genome, for which the eight cloned genome components analogous to the related FBNSV had been shown to be infectious via biolistic but not via aphid transmission (Aronson et al. 2002; Timchenko et al. 2006). This work used agroinoculation of the complete M-Rep-encoding DNA-R and DNA-C segments, together with constructs based on the DNA-C encoding Clink, to express GFP and GUS in *N. benthamiana* leaf disks or *Medicago trunculata* leaves and leaf disks: while M-Rep served to create replicons from and amplify all FBNYV-derived DNAs, Clink co-expression was found to significantly amplify protein expression (Aronson et al. 2002). In the case of GUS, this was  $\sim 6\times$  that driven by CaMV 35S promoter alone; for both GFP and GUS expression levels were correlated with viral DNA levels. This work established the potential of nanoviruses to be used as expression vectors, and also showed that the minimal replicon would need *mrep* in *cis* or in *trans*, releasing a replicon comprising the CR-SL of any cognate genome segment, together with either a heterologous promoter driving just the gene of interest, or *clink* as well with a DNA-C-derived replicon with the native *clink* promoter.

## 3 ssDNA Viruses as Vectors: Current Practice

The modern use of these viruses for expression of pharmaceutically-relevant proteins mostly dates from the application of another technology: this was the application of “agroinfiltration”, or the use of recombinant *Agrobacterium tumefaciens* for transient gene transfer to normal plants, by means of syringe or vacuum infiltration of whole leaves or plants (Fischer et al. 1999; Schob et al. 1997). This process enables the somatic transfer of T-DNA to a high percentage of the cells in the infiltrated leaf or leaves, where it may express its payload whether

it remains episomal, or becomes integrated [see Rybicki (2009a), for a review of its modern use in vaccine protein expression].

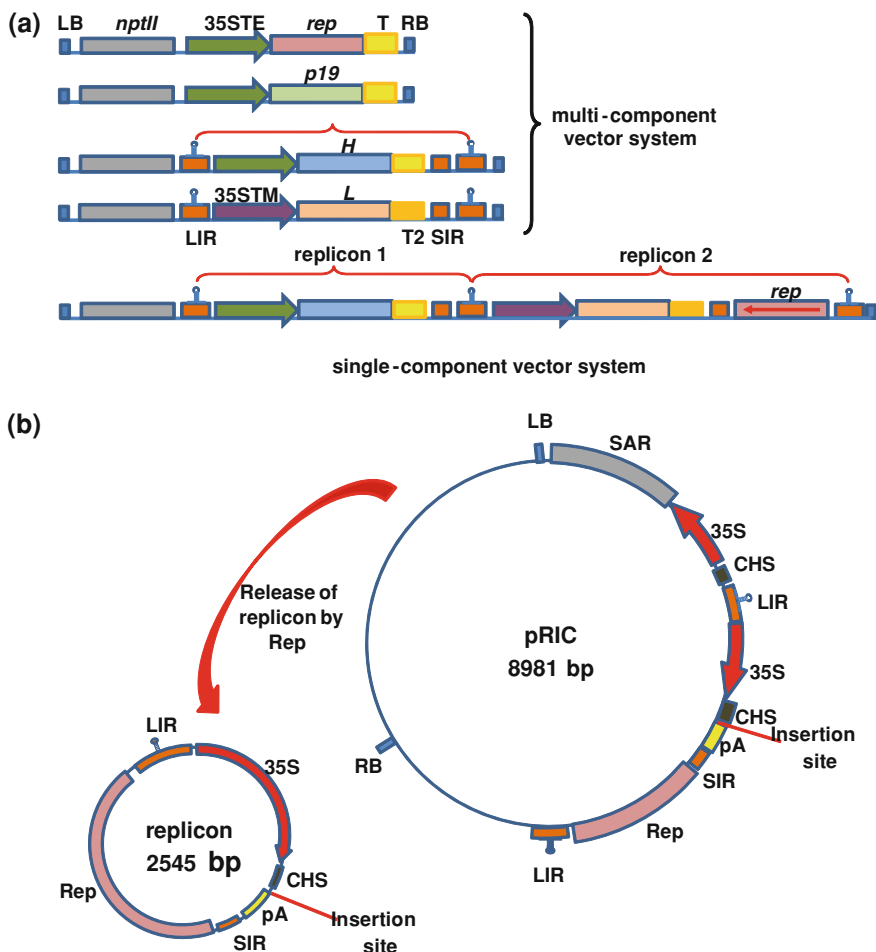
The combination of agroinfiltration with agroinfection effectively turned the majority of dicot-infecting geminiviruses or nanoviruses into potential expression vectors for dicots—not, however, for monocots, as these are generally recalcitrant to agroinfiltration due to leaf morphology and *Agrobacterium*-related host range properties (Heath et al. 1997; Reavy et al. 2007). The take-up by the geminivirus and plant molecular biology research community was swift—but the main application was and is in further demonstrating and investigating the phenomenon of post-transcriptional gene silencing, or what is now known as siRNA-mediated gene silencing (e.g.: Fofana et al. 2004; Muangsang and Robertson, 2004; Rodriguez-Negrete et al. 2009; Turnage et al. 2002).

One of the first applications of the new technology for geminivirus-based expression of a foreign as opposed to a vaccine-relevant gene was in fact done in suspension cell or hairy root cultures on *Nicotiana glutinosa*, using BeYDV. Collens et al. (2007) used the two-component complementing vectors previously described, delivered via auxotrophic *A. tumefaciens*, to express GUS. While the work was proof that a transient expression system was compatible with a geminivirus vector, it was a curiosity rather than a milestone, given only modest (2×) increase in recombinant protein yield.

An interesting excursion in the field was by Kim et al. (2007), with a curtovirus-derived vector: they used Beet severe curly top virus (BSCTV)-derived replicons to express GFP in *N. benthamiana* leaf disks infiltrated with recombinant *A. tumefaciens*. While their best GFP expression was not impressive—they claim a maximum of 1.5% of total soluble protein (TSP)—they did achieve a 3× higher expression level with a 35S Pr in BCTV compared to 35S control, a 3.2× enhancement of that expression if they used a Cassava vein mosaic caulimovirus (CaVMV) Pr instead, and a further 2.4× enhancement if they co-infiltrated leaf disks with *A. tumefaciens* expressing the silencing suppressor protein p19 from Tomato bushy stunt tombusvirus (TBSV). They also showed that whole leaf infiltration resulted in higher levels of GFP than leaf disk infiltration, that it was far easier than leaf disk inoculations, and concluded that “...we have demonstrated the feasibility of a BCTV replicon-based vector system, which could potentially be used in a wide variety of plant cells to express proteins of commercial or medical interest, or proteins for biochemical analysis.” However, apart from one further report of a Beet curly top curtovirus (BCTV)-derived virus-induced gene silencing or VIGS vector (Golenberg et al. 2009), no further developments have occurred with curtovirus vectors.

The first vaccine-relevant development of the modern age, then, was the investigation by Huang et al. (2009) of the potential of an agroinfiltration-mediated BeYDV-based transient expression replicon system to express large amounts of GFP, Hepatitis B virus (HBV) core antigen (HBc), and Norwalk norovirus (NV) capsid protein (NVCP). They used very similar vectors to the earlier work with transgenic tobacco cells—that is, *rep*<sup>−</sup> replicons with a 35S Pr co-transformed with Rep-expressing constructs—but added single-component replicons, with *rep*





in its normal orientation between the right-hand LIR and the SIR at the 3' end of the construct, as well as co-infiltration with p19-expressing vectors for silencing suppression. After rigorous optimisation, the authors succeeded with the single-vector system in producing HBc virus-like particles (VLPs) at 0.8 mg/g leaf fresh weight within 4 days of agroinfiltration—without p19. This was significant, as the previous highest plant yield of HBc was achieved using a two-component non-replicating system which incorporated p19 (Sainsbury and Lomonosoff 2008).

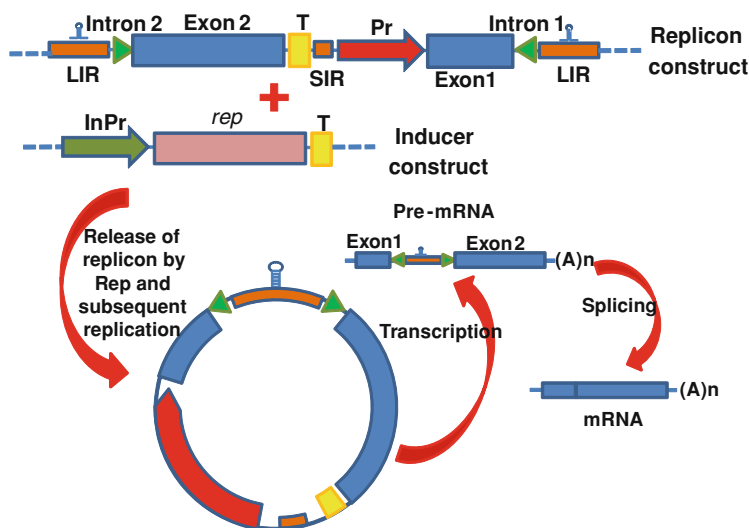
Further successes with the BeYDV system were achieved by the same group with full-size antibody expression (Huang et al. 2010), and by our group with expression of two human virus proteins (Regnard et al. 2010). In the case of antibody expression, it was possible to use either the three-component (*rep* + replicon + p19) system or a single-component, self-replicating vector system that released two replicons, to efficiently express both heavy and light



◀ **Fig. 5** Bean yellow dwarf mastrevirus (BeYDV)-derived replicon vectors. **a** Multiple and single vector multiple replicon constructs used for IgG expression and assembly. The schematic—adapted from the vectors created by Huang et al. (2010)—shows essential details of multigene and replicon systems suitable for the expression of several genes in one cell. The multicomponent system includes four separate plasmids, cointroduced via *A. tumefaciens* infiltration. All constructs encode a neomycin phosphotransferase II gene cassette for kanamycin resistance (*npt II*), and left and right border regions from *A. tumefaciens* T-DNA (LB + RB), transcription termination sequences (T + T2). 35STE CaMV 35S promoter with Tobacco etch potyvirus 5' untranslated region (UTR) as translational enhancer; 35STM 35S Pr with Tobacco mosaic virus 5' UTR. *rep* Rep gene (C1/C2 ORF from BeYDV). *p19* p19 silencing suppressor protein gene from *H* IgG heavy chain gene. *L* IgG light chain gene. LIR long intergenic region of BeYDV genome, shown with stem-loop region; SIR short intergenic region. **b** Single vector and replicon constructs used for vaccine protein expression. The schematic shows the single-replicon, single-component vector described by Regnard et al. (2010) for expression of single genes. T-DNA is the sequence contained within left (LB) and right (RB) borders of the *A. tumefaciens* vector plasmid pTraC described by Maclean et al. (2007); in this system this also contains a CaMV 35S promoter driving a scaffold attachment region (SAR) protein. The replicon includes similar elements as described for Fig. 5a between two LIR regions. pA polyadenylation sequence; CHS chalcone synthase gene 5'-UTR. In addition, the system uses a Tomato spotted wilt tospovirus (TSWV) NSs silencing suppressor protein supplied in *trans* by a co-agroinfiltrated construct

chains of a full IgG molecule specific for Ebola virus, at levels up to 0.5 mg/g leaf tissue, or about 10× the standard transgenic plant yield of monoclonal antibodies (mAbs) (Rybicki 2009a) (see Fig. 5a). The authors rightly point out that this represents a significant advance in transient expression technology, as the otherwise state-of-the-art ICON Genetics RNA virus-derived vectors cannot be used similarly as single-gene-expressing replicon constructs compete, so that multiple component proteins cannot be expressed in the same cell. The authors speculate that “Our success in producing the fully assembled tetrameric functional IgG (two hetero-oligomeric subunits) with a two replicon single vector strongly suggests that simultaneous expression of as many as four hetero-subunits can be easily achieved using two of such vectors, or by creating single vectors with three or more tandem linked replicons”.

The other recent success with BeYDV-based replicons was with another single-component self-replicating vector. This was a shuttle vector based on the mild strain of BeYDV sequenced by our laboratory (BeYDV-m; Halley-Stott et al. 2007). The vector comprises the native genome situated between two LIR copies, with the BeYDV *cp* and *mp* genes replaced with a 2xCaMV 35S promoter and transgene (Fig. 5b). The vector was tested by comparing it with a high-expressing non-replicating vector, pTraC, using transient cytoplasmic expression in *N. benthamiana* of the expression marker EGFP, and the genes encoding the subunit vaccine antigens, HPV-16 major CP L1 (HL1) and the HIV-1C p24 capsid protein (CA) antigen, all of which had previously been successfully expressed in plants using pTraC as well as other vectors (Maclean et al. 2007; Meyers et al. 2008). Expression vectors were co-infiltrated with *A. tumefaciens* expressing the NSs protein from Tomato spotted wilt tospovirus (TSWV), given that this markedly increases expression in our hands using these vectors. It was interesting at the outset that the HL1 vector replicationally



**Fig. 6** InPact—inducible in-plant activation technology—transgenic vector schema. The diagram illustrates the broad concept of the InPact system: this relies on doubly transgenic plants with two independent insertions of the constructs shown. The inducer construct consists of the following components: InPr = inducible promoter (e.g.: alcohol-responsive transcriptional unit); *rep* = Rep gene of a ssDNA virus. The replicon construct has the payload gene divided into two exons (Exon 1/2), separated as shown, with a carefully-designed intron acceptor and intron donor sequence (Intron 1/2) flanking the LIR. *T* transcription terminator; SIR is shown for a mastrevirus-derived construct, but would not be required for other geminivirus-derived or nanovirus-derived constructs. Replicational release of the replicon by expression in trans of Rep, and subsequent circularisation, allows reconstitution of a payload gene with a LIR-containing intron. Transcription and splicing restores a mRNA with an undivided ORF as shown. Details taken from material obtained from JL Dale, Farmacule Bioindustries Pty Ltd, Victoria, Australia

released in *E. coli* and plants, but not in *A. tumefaciens*, given earlier experience with geminiviruses replicating in bacteria (see earlier, and Selth et al. 2002). Gene copy number in plants was assayed by quantitative PCR, and shown to increase for replicons by around 100x (HL1 gene) to 1000x (EGFP and p24 genes) by 3 days post-infiltration relative to non-replicating pTraC versions, and to remain steady for 7 days or more. Protein production peaked around 5 days for HL1, at around 1.5x that of pTraC; at 3 days for p24 at >3x, and was still increasing at 7 days for EGFP at >7x the pTraC level. These results are significant given that we have previously achieved high expression levels for EGFP and levels of HL1 of 0.5 g/kg wet weight using pTraC alone (Maclean et al. 2007), and 16 mg/kg for p24 targeted to the ER (Meyers et al. 2008).

While these developments are interesting and hold promise, possibly the best development in ssDNA virus-based expression technology has come from a company that is using it to make high-value proteins in plants via an inducible transgenic system. The company is Farmacule BioIndustries Pty Ltd of Victoria,

Australia ([www.farmacule.com](http://www.farmacule.com)); the patented technology is called InPact, for “In-Plant Activation Technology”, and relies on induced expression of Rep from one transgene to activate an independently inserted replicon construct. However, this replicon is constructed in such a way that, in the linear inserted form, the promoter is downstream from the gene of interest, which is itself divided so that expression from the gene is impossible unless the replicon is reconstituted into a circular form—and then only after splicing, as the whole LIR has been made into an intron (see Fig. 6). The technology allows tight control of expression of Rep—usually via alcohol induction—and near-absolute control over replicon release, and therefore of protein expression. Thus, transgenic plants potentially expressing even proteins that are highly toxic to them, or which would normally drastically affect development, can be grown normally until biomass is suitable for harvest, at which time the inducer is used to allow production of Rep and subsequently highly amplified expression of the replicon and protein (see Rybicki 2009b for a report). The system has been used successfully to produce high yields of vitronectin, a high-value protein used in tissue culture which is normally produced from bovine blood as well as enzymes like trypsin. The technology could potentially make use of any Rep-replicon pairing, but most development has been done using the indigenous mastrevirus TYDV. While development so far has apparently used the two-component system, it is of course possible to have plants transgenic for a replicon only, and introduce Rep via agroinfiltration if desired.

## 4 Conclusion and Future Prospects

The future use of geminivirus- or nanovirus-derived replicons for pharming is almost certainly very bright: the expression systems have been tested for a variety of geminiviruses in particular, and for a variety of plants, cells and types of proteins, and have generally been found to perform at least as well as conventional agroinfiltration, and considerably better in the recent examples shown above. Vector systems may be single component or be constituted as two or more components; while agroinfiltration may be used for whole plants, it is also possible to use it for leaves and even cell cultures; the use of transgenic plants are possible, especially if these are part of the InPact or similar system with expression only after induction, and so can transgenic cell lines be used over several years of culture. We can predict that the application of ssDNA vector and expression systems will take off exponentially in coming years, much as the use of deconstructed RNA virus vectors such as the MagnICON system of Icon Genetics has recently (Gleba et al. 2005, 2007).

The two main advantages of the technology for high-level pharmaceutical protein production over systems like the deconstructed RNA virus vectors are first, host range of the potential vectors, and second, the fact that it is not necessary to use two different viruses to express more than one polypeptide in the same cell,

where heterodimeric or greater associations are needed to produce a functional protein.

In the first place, the host range of ssDNA plant viruses is vast: if one only considers crop plants, the range includes such staples as maize, wheat, rice, banana, tomato, potato, cassava, sweet potato, sugar beets, cotton and beans. In one justification for the establishment of BCTV curtovirus as a silencing vector, Golenberg et al. (2009) point out that it infects 41 families of dicotyledonous plants in six subclasses—including several useful and well-characterised production hosts such as Fabaceae and Solanaceae. This is far greater than the essentially limited Tobacco mosaic virus strain normally used in “Magnification”, which is pretty much limited to *N. tabacum* and preferentially *N. benthamiana*—and while the developers are introducing other viruses with a wider range of hosts, such as Potato virus X (PVX) and potyviruses (see Rybicki 2009b), each of these is another challenge and another complicated design and sequence engineering job, whereas geminivirus vector development in particular so far has required little of this for good yields.

The second point was dealt with to some extent above in discussion of the work of Huang et al. (2010): they successfully managed to use co-agroinfiltration of two vectors to get individual leaf cells to produce both heavy and light chains of an IgG molecule, which was about as good as using a single vector in terms of yield. A TMV-based vector cannot produce more than one protein, and attempts to co-infiltrate more than one have failed due to competition between genomes—necessitating the use of non-interfering viruses (Giritch et al. 2006), with all of the immediate complexities that entails in terms of cloning strategies, and the added burden of vector design for maximal yield which will be beyond most laboratories.

Another useful feature of the DNA viruses is that they replicate and produce mRNA in the nucleus using cellular machinery, with all of the possible advantages that entails—such as the possibility of taking advantage of splicing, to limit expression of the transgene in the *A. tumefaciens*, if this is used; addition of polyA tails to limit mRNA degradation; the possibility of getting expression in transgenic plants at the time of choosing by means of a simple induction mechanism, and so on.

One possible drawback of using agroinfiltrated ssDNA replicons compared to MagnICON vectors is that presently, expression with the former is essentially limited to cells in which T-DNA transfer has occurred, while the latter maintain transport functions and can spread to untransfected or new cells. The advantage of this property is that, first, it serves to spread expression into effectively all leaf tissue in the plant; second, that it allows a significant reduction in the amount of *A. tumefaciens* that must be used for infiltration, given that the whole plant need not be infiltrated for maximal expression. This is not a trivial barrier to surmount, given that geminivirus and possibly nanovirus genome movement is complicated by there being a number of different proteins involved in the different genera of geminiviruses, let alone between the two families. For example, two-component NW begomoviruses exploit the B-encoded MP or BC1 ORF-encoded protein as a plasmodesmatal gate protein and the NSP or BV1 as a nuclear shuttle protein (Noueiry et al. 1994; Pascal et al. 1994). However, single-component begomo-

curto- and topocu- and mastreviruses have neither of these, and instead use A-type genome component encoded functions, such as the MP or V1 ORF and CP or V2 ORF proteins of mastreviruses, for movement between cells and nuclear transport respectively (Liu et al. 1997, 2001), or the C4 and/or V2 “pre-coat” ORF protein(s) for movement and CP for nuclear localisation in begomoviruses (Jupinet et al. 1994; Padidam et al. 1996). Moreover, begomovirus genome movement between cells exhibits a strong size-dependency (Gilbertson et al. 2003; Rojas et al. 1998), and the inability of MSV-derived vectors with inserts to spread in plants has already been noted. Additionally, the thought of using transgenic plants expressing movement proteins to complement replicon spread is itself a problem, as the plants may have abnormal phenotypes and in fact be partially or substantially resistant to cognate virus infection and spread (Pascal et al. 1993; von Arnim and Stanley 1992).

Taken together these observations add up to a complex problem to be solved if one wishes ssDNA virus-derived replicon vectors to spread in agroinfiltrated or -inoculated plants. However, there are some obvious solutions if this is regarded as a problem, such as using transgenic plants incorporating the InPact technology, where expression is limited only by which tissue has been exposed to the inducing agent, or *mp*-transgenic plants which are not symptomatic, or in which MP expression is inducible, with agroinfiltration of replicons not exceeding genome size so that spread will occur.

It is also feasible to engineer ssDNA replicons in the same way that the deconstructed TMV-based vectors have been; that is, with codon optimisation, removal of cryptic splice sites, molecular breeding of various of their proteins to either remove unwanted properties, or enhance desired characteristics. Thus, it is possible that expression could be enhanced without adding movement functions, to the point where these are not considered necessary.

It is worth mentioning that while we have treated geminiviruses and nanoviruses as being essentially equivalent for purposes of this review, in terms of expression at least, they are probably not. For example, while considerable protein expression amplification was demonstrated using the mastreviruses MSV in cells, and BeYDV in cells and plants, the same was not true of the curtovirus BSCTV; nor have the begomoviruses that have been tested been particularly impressive in terms of yield of recombinant protein (see Palmer and Rybicki 1997). This may have to do with the fact that mastreviruses in both cereal plants and isolated cells appear to replicate to very high copy number compared to begomoviruses, and probably also curto- and topocoviruses: indeed, early workers found that WDV in protoplasts could replicate up to 30,000 genome copies per cell (Matzeit et al. 1991), a finding essentially replicated for MSV by Palmer et al. (1999), compared to values for begomoviruses of the order of 1,000 genomes/cell. This difference in cell and whole plant DNA content can be very simply demonstrated in protocols for extraction of replicative form dsDNA from infected plants: it is far easier to purify this from mastrevirus-infected maize, wheat or even beans, than it is to purify begomovirus-derived dsDNA. In fact, MSV dsDNA obtained from maize using a bacterial plasmid miniprep protocol is abundant and pure enough to be

sequenced directly by the Sanger dideoxy method (Palmer et al. 1998). Thus, while it is possible to use any ssDNA plant virus as a vector, in practice the choice may be constrained by the simple fact of degree of genome amplification: the levels of mRNA and probably to a lesser extent of protein, given the capacity of geminiviral replicons to saturate the cell synthetic machinery, will inevitably be greater for a greater copy number.

In closing, it should be evident that there is ample room for further optimisation of ssDNA virus-based expression systems in plants: a few representatives of only two generic sets of geminiviruses have been tested out of four sets and hundreds of sequenced genomes available; additionally, familial nanoviruses are almost completely unexploited in terms of their potential. Lessons learned with the plant viruses could also be applied to the new ssDNA viruses being found almost daily which infect fungi, algae and animals—and some of these or elements derived from them may even be found to be useful in plant systems, and vice versa.

**Acknowledgments** We gratefully acknowledge financial assistance from the University of Cape Town, the National Research Foundation and the South African AIDS Vaccine Initiative, Guy Regnard and James Dale for illustrations, and lab members past and present for their sterling work.

## References

- Aronson MN, Complainville A, Clerot D et al (2002) In planta protein–protein interactions assessed using a nanovirus-based replication and expression system. *Plant J* 31:767–775
- Aronson MN, Meyer AD, Gyorgyey J, Katul L, Vetten HJ, Gronenborn B, Timchenko T (2000) *Clink*, a nanovirus-encoded protein, binds both pRB and SKP1. *J Virol* 74:2967–2972
- Bottcher B, Unseld S, Ceulemans H, Russell RB, Jeske H (2004) Geminite structures of African cassava mosaic virus. *J Virol* 78:6758–6765
- Boulton MI (1995) Agrobacterium-mediated transfer of geminiviruses to plant tissues. *Methods Mol Biol* 49:77–93
- Briddon RW, Patil BL, Bagewadi B et al (2010) Distinct evolutionary histories of the DNA-A and DNA-B components of bipartite begomoviruses. *BMC Evol Biol* 10:97
- Briddon RW, Stanley J (2006) Subviral agents associated with plant single-stranded DNA viruses. *Virology* 344:198–210
- Briddon RW, Watts J, Markham PG et al (1989) The coat protein of beet curly top virus is essential for infectivity. *Virology* 172:628–633
- Casado CG, Javier OG, Padron E, Bean SJ, McKenna R, Agbandje-McKenna M, Boulton MI (2004) Isolation and characterization of subgenomic DNAs encapsidated in “single” T = 1 isometric particles of maize streak virus. *Virology* 323:164–171
- Collens JJ, Mason HS, Curtis WR (2007) Agrobacterium-mediated viral vector-amplified transient gene expression in *Nicotiana glutinosa* plant tissue culture. *Biotechnol Prog* 23:570–576
- Donson J, Gunn HV, Woolston CJ et al (1988) Agrobacterium-mediated infectivity of cloned digitaria streak virus DNA. *Virology* 162:248–250
- Dry IB, Krake LR, Rigden JE, Rezaian MA (1997) A novel subviral agent associated with a geminivirus: the first report of a DNA satellite. *Proc Natl Acad Sci USA* 94:7088–7093
- Dugdale B, Beetham PR, Becker DK et al (1998) Promoter activity associated with the intergenic regions of banana bunchy top virus DNA-1 to -6 in transgenic tobacco and banana cells. *J Gen Virol* 79(Pt 10):2301–2311

- Fischer R, Vaquero-Martin C, Sack M et al (1999) Towards molecular farming in the future: transient protein expression in plants. *Biotechnol Appl Biochem* 30(Pt 2):113–116
- Fofana IB, Sangare A, Collier R, Taylor C, Fauquet CM (2004) A geminivirus-induced gene silencing system for gene function validation in cassava. *Plant Mol Biol* 56:613–624
- Frischmuth T, Ringel M, Kocher C (2001) The size of encapsidated single-stranded DNA determines the multiplicity of African cassava mosaic virus particles. *J Gen Virol* 82:673–676
- Gilbertson RL, Sudarshana M, Jiang H et al (2003) Limitations on geminivirus genome size imposed by plasmodesmata and virus-encoded movement protein: insights into DNA trafficking. *Plant Cell* 15:2578–2591
- Giritch A, Marillonnet S, Engler C et al (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfecting with noncompeting viral vectors. *Proc Natl Acad Sci USA* 103:14701–14706
- Gleba Y, Klimyuk V, Marillonnet S (2005) Magnification—a new platform for expressing recombinant vaccines in plants. *Vaccine* 23:2042–2048
- Gleba Y, Klimyuk V, Marillonnet S (2007) Viral vectors for the expression of proteins in plants. *Curr Opin Biotechnol* 18:134–141
- Golenberg EM, Sather DN, Hancock LC et al (2009) Development of a gene silencing DNA vector derived from a broad host range geminivirus. *Plant Methods* 5:9
- Grigoras I, Timchenko T, Katul L et al (2009) Reconstitution of authentic nanovirus from multiple cloned DNAs. *J Virol* 83:10778–10787
- Grimsley N, Hohn B, Hohn T et al (1986) “Agroinfection,” an alternative route for viral infection of plants by using the Ti plasmid. *Proc Natl Acad Sci USA* 83:3282–3286
- Grimsley N, Hohn T, Davies JW et al (1987) *Agrobacterium*-mediated delivery of infectious maize streak virus into maize plants. *Nature* 325:177–179
- Gronenborn B (2004) Nanoviruses: genome organisation and protein function. *Vet Microbiol* 98:103–109
- Gutierrez C (1999) Geminivirus DNA replication. *Cell Mol Life Sci* 56:313–329
- Gutierrez C (2000) DNA replication and cell cycle in plants: learning from geminiviruses. *EMBO J* 19:792–799
- Gutierrez C, Ramirez-Parra E, Mar CM et al (2004) Geminivirus DNA replication and cell cycle interactions. *Vet Microbiol* 98:111–119
- Halley-Stott RP, Tanzer F, Martin DP et al (2007) The complete nucleotide sequence of a mild strain of Bean yellow dwarf virus. *Arch Virol* 152:1237–1240
- Hanley-Bowdoin L, Settledge SB, Orozco BM, Nagar S, Robertson D (2000) Geminiviruses: models for plant DNA replication, transcription, and cell cycle regulation. *Crit Rev Biochem Mol Biol* 35:105–140
- Hanley-Bowdoin L, Settledge SB, Robertson D (2004) Reprogramming plant gene expression: a prerequisite to geminivirus DNA replication. *Mol Plant Pathol* 5:149–156
- Harding RM, Burns TM, Dale JL (1991) Virus-like particles associated with banana bunchy top disease contain small single-stranded DNA. *J Gen Virol* 72(Pt 2):225–230
- Hayes RJ, Coutts RH, Buck KW (1989) Stability and expression of bacterial genes in replicating geminivirus vectors in plants. *Nucleic Acids Res* 17:2391–2403
- Hayes RJ, Petty IT, Coutts RH et al (1988) Gene amplification and expression in plants by a replicating geminivirus vector. *Nature* 334:179–182
- Heath JD, Boulton MI, Raineri DM et al (1997) Discrete regions of the sensor protein *virA* determine the strain-specific ability of *Agrobacterium* to agroinfect maize. *Mol Plant Microbe Interact* 10:221–227
- Hefferon KL, Fan Y (2004) Expression of a vaccine protein in a plant cell line using a geminivirus-based replicon system. *Vaccine* 23:404–410
- Hefferon KL, Kipp P, Moon YS (2004) Expression and purification of heterologous proteins in plant tissue using a geminivirus vector system. *J Mol Microbiol Biotechnol* 7:109–114
- Hiatt A, Cafferkey R, Bowditch K (1989) Production of antibodies in transgenic plants. *Nature* 342:76–78

- Hormuzdi SG, Bisaro DM (1993) Genetic analysis of beet curly top virus: evidence for three virion sense genes involved in movement and regulation of single- and double-stranded DNA levels. *Virology* 193:900–909
- Hormuzdi SG, Bisaro DM (1995) Genetic analysis of beet curly top virus: examination of the roles of L2 and L3 genes in viral pathogenesis. *Virology* 206:1044–1054
- Huang Z, Chen Q, Hjelm B, Arntzen C, Mason H (2009) A DNA replicon system for rapid high-level production of virus-like particles in plants. *Biotechnol Bioeng* 103:706–714
- Huang Z, Phoolcharoen W, Lai H, Piensook K, Cardineau G, Zeitlin L, Whaley KJ, Arntzen CJ, Mason HS, Chen Q (2010) High-level rapid production of full-size monoclonal antibodies in plants by a single-vector DNA replicon system. *Biotechnol Bioeng* 106:9–17
- Jupin I, de Kouchkovsky F, Gronenborn B (1994) Movement of tomato yellow leaf curl geminivirus (TYLCV): involvement of the protein encoded by ORF C4. *Virology* 204:82–90
- Kammann M, Matzeit V, Schmidt B, Schell J, Walden R, Gronenborn B (1991) Geminivirus-based shuttle vectors capable of replication in *Escherichia coli* and monocotyledonous plant cells. *Gene* 104:247–252
- Kim KI, Sunter G, Bisaro DM, Chung IS (2007) Improved expression of recombinant GFP using a replicating vector based on Beet curly top virus in leaf-disks and infiltrated *Nicotiana benthamiana* leaves. *Plant Mol Biol* 64:103–112
- Koonin EV, Ilyina TV (1992) Geminivirus replication proteins are related to prokaryotic plasmid rolling circle DNA replication initiator proteins. *J Gen Virol* 73(Pt 10):2763–2766
- Kotlizky G, Boulton MI, Pitaksutheepong C, Davies JW, Epel BL (2000) Intracellular and intercellular movement of maize streak geminivirus V1 and V2 proteins transiently expressed as green fluorescent protein fusions. *Virology* 274:32–38
- Lageix S, Catrice O, Deragon JM, Gronenborn B, Pelissier T, Ramirez BC (2007) The nanovirus-encoded *Clink* protein affects plant cell cycle regulation through interaction with the retinoblastoma-related protein. *J Virol* 81:4177–4185
- Laufs J, Wirtz U, Kammann M, Matzeit V, Schaefer S, Schell J, Czernilofsky AP, Baker B, 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 SG, Beachy RN (1999) Viral movement proteins as probes for intracellular and intercellular trafficking in plants. *The Plant Cell* 11:535–548
- Liu H, Boulton MI, Davies JW (1997) Maize streak virus coat protein binds single- and double-stranded DNA in vitro. *J Gen Virol* 78(Pt 6):1265–1270
- Liu H, Boulton MI, Oparka KJ, Davies JW (2001) Interaction of the movement and coat proteins of Maize streak virus: implications for the transport of viral DNA. *J Gen Virol* 82:35–44
- Liu H, Boulton MI, Thomas CL, Prior DA, Oparka KJ, Davies JW (1999) Maize streak virus coat protein is karyophilic and facilitates nuclear transport of viral DNA. *Mol Plant Microbe Interact* 12:894–900
- Maclean J, Koekemoer M, Olivier AJ, Stewart D, Hitzeroth II, Rademacher T, Fischer R, Williamson AL, Rybicki EP (2007) Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *J Gen Virol* 88:1460–1469
- Mansoor S, Khan SH, Bashir A, Saeed M, Zafar Y, Malik KA, Briddon R, Stanley J, Markham PG (1999) Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology* 259:190–199
- Matzeit V, Schaefer S, Kammann M, Schalk HJ, Schell J, Gronenborn B (1991) Wheat dwarf virus vectors replicate and express foreign genes in cells of monocotyledonous plants. *Plant Cell* 3:247–258
- Meyers A, Chakauya E, Shephard E, Tanzer FL, Maclean J, Lynch A, Williamson AL, Rybicki EP (2008) Expression of HIV-1 antigens in plants as potential subunit vaccines. *BMC Biotechnol* 8:53
- Mor TS, Moon YS, Palmer KE, Mason HS (2003) Geminivirus vectors for high-level expression of foreign proteins in plant cells. *Biotechnol Bioeng* 81:430–437



- Muangsan N, Robertson D (2004) Geminivirus vectors for transient gene silencing in plants. *Methods Mol Biol* 265:101–115
- Nawaz-ul-Rehman MS, Fauquet CM (2009) Evolution of geminiviruses and their satellites. *FEBS Lett* 583:1825–1832
- Needham PD, Atkinson RG, Morris BAM, Gardner RC, Gleave AP (1998) GUS expression patterns from a tobacco yellow dwarf virus-based episomal vector. *Plant Cell Rep* 17:631–639
- Nishigawa H, Miyata S, Oshima K, Sawayanagi T, Komoto A, Kuboyama T, Matsuda I, Tsuchizaki T, Namba S (2001) In planta expression of a protein encoded by the extrachromosomal DNA of a phytoplasma and related to geminivirus replication proteins. *Microbiology* 147:507–513
- Noueiry AO, Lucas WJ, Gilbertson RL (1994) Two proteins of a plant DNA virus coordinate nuclear and plasmodesmal transport. *Cell* 76:925–932
- Padidam M, Beachy RN, Fauquet CM (1996) The role of AV2 (“precoat”) and coat protein in viral replication and movement in tomato leaf curl geminivirus. *Virology* 224:390–404
- Palmer KE, Rybicki EP (1997) The use of geminiviruses in biotechnology and plant molecular biology, with particular focus on mastreviruses. *Plant Sci* 129:115–130
- Palmer KE, Rybicki EP (1998) The molecular biology of mastreviruses. *Adv Virus Res* 50:183–234
- Palmer KE, Rybicki EP (2001) Investigation of the potential of maize streak virus to act as an infectious gene vector in maize plants. *Arch Virol* 146:1089–1104
- Palmer KE, Schnippenkoetter WH, Rybicki EP (1998) Geminivirus isolation and DNA extraction. *Methods Mol Biol* 81:41–52
- Palmer KE, Thomson JA, Rybicki EP (1999) Generation of maize cell lines containing autonomously replicating maize streak virus-based gene vectors. *Arch Virol* 144:1345–1360
- Pascal E, Goodlove PE, Wu LC, Lazarowitz SG (1993) Transgenic tobacco plants expressing the geminivirus BL1 protein exhibit symptoms of viral disease. *Plant Cell* 5:795–807
- Pascal E, Sanderfoot AA, Ward BM, Medville R, Turgeon R, Lazarowitz SG (1994) The geminivirus BR1 movement protein binds single-stranded DNA and localizes to the cell nucleus. *Plant Cell* 6:995–1006
- Paszkowski U, Zhang SB, Potrykus I, Paszkowski J (1993) Replication of the DNA A component of African cassava mosaic virus in a heterologous system. *J Gen Virol* 74(Pt 12):2725–2729
- Qin S, Ward BM, Lazarowitz SG (1998) The bipartite geminivirus coat protein aids BR1 function in viral movement by affecting the accumulation of viral single-stranded DNA. *J Virol* 72:9247–9256
- Raghavan V, Malik PS, Choudhury NR, Mukherjee SK (2004) The DNA-A component of a plant geminivirus (Indian mung bean yellow mosaic virus) replicates in budding yeast cells. *J Virol* 78:2405–2413
- Reavy B, Bagirova S, Chichkova NV, Fedoseeva SV, Kim SH, Vartapetian AB, Taliansky ME (2007) Caspase-resistant VirD2 protein provides enhanced gene delivery and expression in plants. *Plant Cell Rep* 26:1215–1219
- Regnard GL, Halley-Stott RP, Tanzer FL, Hitzeroth II, Rybicki EP (2010) High level protein expression in plants through the use of a novel autonomously replicating geminivirus shuttle vector. *Plant Biotechnol J* 8:38–46
- Rekab D, Carraro L, Schneider B, Seemuller E, Chen J, Chang CJ, Locci R, Firrao G (1999) Geminivirus-related extrachromosomal DNAs of the X-clade phytoplasmas share high sequence similarity. *Microbiology* 145(Pt 6):1453–1459
- Rochester DE, Kosiratana W, Beachy RN (1990) Systemic movement and symptom production following agroinoculation with a single DNA of tomato yellow leaf curl geminivirus (Thailand). *Virology* 178:520–526
- Rodriguez-Negrete EA, Carrillo-Tripp J, Rivera-Bustamante RF (2009) RNA silencing against geminivirus: complementary action of posttranscriptional gene silencing and transcriptional gene silencing in host recovery. *J Virol* 83:1332–1340
- Rojas MR, Noueiry AO, Lucas WJ, Gilbertson RL (1998) Bean Dwarf mosaic geminivirus movement proteins recognize DNA in a form- and size-specific manner. *Cell* 95:105–113

- Rosario K, Duffy S, Breitbart M (2009) Diverse circovirus-like genome architectures revealed by environmental metagenomics. *J Gen Virol* 90:2418–2424
- Rybicki EP (2009a) Plant-produced vaccines: promise and reality. *Drug Discov Today* 14:16–24
- Rybicki EP (2009b) Third international conference on plant-based vaccines and antibodies. *Expert Rev Vaccines* 8:1151–1155
- Sainsbury F, Lomonosoff GP (2008) Extremely high-level and rapid transient protein production in plants without the use of viral replication. *Plant Physiol* 148:1212–1218
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Saunders K, Bedford ID, Briddon RW, Markham PG, Wong SM, Stanley J (2000) A unique virus complex causes *Ageratum* yellow vein disease. *Proc Natl Acad Sci USA* 97:6890–6895
- Saunders K, Stanley J (1999) A nanovirus-like DNA component associated with yellow vein disease of *Ageratum conyzoides*: evidence for interfamilial recombination between plant DNA viruses. *Virology* 264:142–152
- Schob H, Kunz C, Meins F Jr (1997) Silencing of transgenes introduced into leaves by agroinfiltration: a simple, rapid method for investigating sequence requirements for gene silencing. *Mol Gen Genet* 256:581–585
- Selth LA, Randles JW, Rezaian MA (2002) *Agrobacterium tumefaciens* supports DNA replication of diverse geminivirus types. *FEBS Lett* 516:179–182
- Shen WH, Hohn B (1992) Excision of a transposable element from a viral vector introduced into maize plants by agroinfection. *Plant J* 2:35–42
- Shen WH, Hohn B (1994) Amplification and expression of the beta-glucuronidase gene in maize plants by vectors based on maize streak virus. *Plant J* 5:227–236
- Shepherd DN, Martin DP, van der Walt E, Dent K, Varsani A, Rybicki EP (2010) Maize streak virus: an old and complex ‘emerging’ pathogen. *Mol Plant Pathol* 11:1–12
- Stanley J (1993) Geminiviruses: plant viral vectors. *Curr Opin Genet Dev* 3:91–96
- Stanley J, Bisaro DM, Briddon RW, Brown JK, Fauquet CM, Harrison BD, Rybicki EP, Stenger DC (2005) Geminiviridae. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) *International Committee on the Taxonomy of Viruses (VIIIth Report)*. Elsevier/Academic Press, London, pp 301–326
- Stanley J, Latham JR, Pinner MS, Bedford I, Markham PG (1992) Mutational analysis of the monopartite geminivirus beet curly top virus. *Virology* 191:396–405
- Sugimoto K, Otsuki Y, Saji S, Hirochika H (1994) Transposition of the maize Ds element from a viral vector to the rice genome. *Plant J* 5:863–871
- Takamatsu N, Ishikawa M, Meshi T, Okada Y (1987) Expression of bacterial chloramphenicol acetyltransferase gene in tobacco plants mediated by TMV-RNA. *EMBO J* 6:307–311
- Tamilselvi D, Anand G, Swarup S (2004) A geminivirus AYVV-derived shuttle vector for tobacco BY2 cells. *Plant Cell Rep* 23:81–90
- Thomas JE, Dietzgen RG (1991) Purification, characterization and serological detection of virus-like particles associated with banana bunchy top disease in Australia. *J Gen Virol* 72(Pt 2): 217–224
- Thomas JE et al (2010) Two novel mastreviruses from chickpea (*Cicer arietinum*) in Australia. *Arch Virol* 155(11):1777–1788 (Epub 2010 Aug 24)
- Timchenko T, Katul L, Aronson M, Vega-Arrequin JC, Ramirez BC, Vetten HJ, Gronenborn B (2006) Infectivity of nanovirus DNAs: induction of disease by cloned genome components of Faba bean necrotic yellows virus. *J Gen Virol* 87:1735–1743
- Timmermans MCP, Das OP, Messing J (1994) Geminiviruses and their uses as extrachromosomal replicons. *Annu Rev Plant Physiol Plant Mol Biol* 45:79–112
- Turnage MA, Muangsang N, Peele CG, Robertson D (2002) Geminivirus-based vectors for gene silencing in *Arabidopsis*. *Plant J* 30:107–114
- Vetten HJ, Chu PWG, Dale JL, Harding R, Hu J, Katul L, Kojima M, Randles JW, Sano Y, Thomas JE (2005) *Nanoviridae*. In: Fauquet CM, Mayo MA, Maniloff J, Desselberger U, Ball LA (eds) *Virus taxonomy*, 8th report of the international committee on taxonomy of viruses, Elsevier/Academic Press, p 1259

- von Arnim A, 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
- Wanitchakorn R, Hafner GJ, Harding RM, Dale JL (2000) Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6. *J Gen Virol* 81:299–306
- Ward A, Etesami P, Stanley J (1988) Expression of a bacterial gene in plants mediated by infectious geminivirus DNA. *EMBO J* 7:1583–1587
- Wartig L, Kheyr-Pour A, Noris E, de Kouchovsky F, Jouanneau F, Gronenborn B, Jupin I (1997) Genetic analysis of the monopartite tomato yellow leaf curl geminivirus: roles of V1, V2, and C2 ORFs in viral pathogenesis. *Virology* 228:132–140
- Wirtz U, Osborne B, Baker B (1997) Ds excision from extrachromosomal geminivirus vector DNA is coupled to vector DNA replication in maize. *Plant J* 11:125–135
- Wu CY, Yang SH, Lai YC, Lin NS, Hsu YH, Hu CC (2007) Unit-length, single-stranded circular DNAs of both polarity of begomoviruses are generated in *Escherichia coli* harboring phage M13-cloned begomovirus genome with single copy of replication origin. *Virus Res* 125:14–28
- Yu X, Li B, Fu Y, Jiang D, Ghabrial SA, Li G, Peng Y, Xie J, Cheng J, Huang J, Yi X (2010) A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc Natl Acad Sci USA* 107:8387–8392
- Yusibov V, Rabindran S, Commandeur U, Twyman RM, Fischer R (2006) The potential of plant virus vectors for vaccine production. *Drugs RD* 7:203–217
- Zhang W, Olson NH, Baker TS, Faulkner L, gbandje-McKenna M, Boulton MI, Davies JW, McKenna R (2001) Structure of the Maize streak virus geminate particle. *Virology* 279:471–477
- Zhang X, Mason H (2006) Bean Yellow Dwarf Virus replicons for high-level transgene expression in transgenic plants and cell cultures. *Biotechnol Bioeng* 93:271–279

Plant Viral Vectors

Palmer, K.E.; Gleba, Y. (Eds.)

2014, X, 194 p. 33 illus., 19 illus. in color., Hardcover

ISBN: 978-3-642-40828-1