Transgenic Res (2007) 16:587–597
DOI 10.1007/s11248-006-9051-1

ORIGINAL PAPER

Stability of *Potato virus X* expression vectors is related to insert size: implications for replication models and risk assessment

Linda Avesani · Giampiero Marconi · Francesca Morandini · Emidio Albertini · Matteo Bruschetta · Luisa Bortesi · Mario Pezzotti · Andrea Porceddu

Received: 30 May 2006 / Accepted: 17 October 2006 / Published online: 11 January 2007
© Springer Science+Business Media B.V. 2006

Abstract We investigated the stability of expression constructs based on *Potato virus X* (PVX) as a function of insert length. Five different inserts ranging in length from 261 to 1,758 bp (human *proinsulin*, murine *interleukin-10*, HIV-1 *nef*, petunia *expansin-1* and human *gad65*) were expressed using a PVX vector in *Nicotiana benthamiana* plants for three sequential passages. Using a competitive RT-PCR approach we demonstrated that insert–deletion could occur in the first infection cycle for all inserts, but that this was much more likely to be the case for longer ones. This suggested a negative correlation between insert length and vector stability. Sequence analysis of the deleted constructs suggested that recombination usually occurred at sites close to the duplicated sub-genomic promoter, but in a smaller number of cases the foreign gene itself was probably involved, resulting in partially deleted constructs containing transgene fragments. The implications of these results in the context of recombinant protein expression and its risks are discussed.

Keywords *Potato virus X* · Transient expression · Recombinant proteins · Viral replication · Risk assessment

L. Avesani · F. Morandini · M. Bruschetta · L. Bortesi · M. Pezzotti (✉)
Dipartimento Scientifico e Tecnologico, Università degli Studi di Verona, Strada Le Grazie 15, 37134 Verona, Italy
e-mail: mario.pezzotti@univr.it

G. Marconi · E. Albertini
Dipartimento di Biologia Vegetale e Biotecnologie Agroambientali e Zootecniche, Università degli Studi di Perugia, Perugia, Italy

A. Porceddu
Dipartimento di Scienze Agronomiche e Genetica Vegetale Agraria, Università degli Studi di Sassari, Sassari, Italy

Introduction

Plant viruses have been developed as powerful and versatile expression vectors, allowing the production of a wide range of recombinant proteins in plants, including biopharmaceuticals and industrial proteins (Canizares et al. 2005; Yusibov et al. 2006). The most widely used viruses are *Tobacco Mosaic virus* (TMV) (Kumagai et al. 1993; Verch et al. 1998; Karasev et al. 2005), *Cowpea Mosaic virus* (Gopinath et al. 2000; Liu et al. 2005) and *Potato virus X* (PVX) (Ziegler et al. 2000; Franconi et al. 2002; Avesani et al. 2003).

Potato virus X is the type species of the genus Potexvirus. It has a single-stranded RNA genome with a 5′ end cap and a 3′ polyadenylate tail. The genome contains five open reading frames

(ORFs) encoding polymerase, a set of three movement proteins [triple gene block (TGB)] and the coat protein (CP). The virus assembles to form filamentous particles consisting of approximately 1,270 CP subunits encapsidating a single RNA molecule. It reaches high titres in its natural host, potato, but also in the experimental hosts *Nicotiana benthamiana* and *Arabidopsis thaliana*.

Potato virus X-based transient expression vectors are available in three formats. The first uses a duplicated sub-genomic promoter to drive exogenous gene expression (Baulcombe et al. 1995). In the second, foreign proteins are expressed as N-terminal CP fusions. The exogenous protein could be fused directly to the N-terminus of CP (Marusic et al. 2001; Uhde et al. 2005) or it could be linked to the CP via the *Foot and Mouth Disease virus* 2A catalytic peptide (Santa Cruz et al. 1996). The 2A sequence promotes cotranslational cleavage of the fusion protein, and maintains virus infectivity by facilitating the production of both cleaved and recombinant CPs (Santa Cruz et al. 1996; O'Brien et al. 2000). In the third format, the CP sub-genomic promoter drives the transcription of a bicistronic mRNA containing both the foreign gene and CP gene, separated by an internal ribosome entry site (Toth et al. 2001).

All three vectors share some shortcomings: (1) foreign genes are expressed transiently and therefore are not inherited by successive generations of plants; (2) there are size constraints on the inserted sequences; (3) infected plants must be grown under containment to prevent the recombinant virus escaping into the environment; and (4) recombinant viruses tend to be unstable when serially passaged, resulting in the elimination of exogenous genes.

From an environmental perspective, the predominant concern associated with the use of modified plant viruses is the risk of recombinant virus strains escaping into the environment. Recombinant proteins could be expressed in adventitiously infected plants, and this could directly or indirectly affect human health or lead to other non-intended effects. For these reasons, any risk assessment carried out on plant virus vectors should consider in detail the properties of the inserted genetic material and any potential environmental hazard arising from virus modification.

In the context of a PVX vector risk assessment, we studied the relationship between foreign gene insert length and stability using five different ORFs ranging in length from 261 to 1,758 bp. These genes comprised human *proinsulin*, murine *interleukin-10*, HIV-1 *nef*, petunia *expansin-1* and human *gad65*. They were cloned in the Gateway PVX-based destination vector (pPVX201_GAT) under the control of the PVX CP duplicated sub-genomic promoter.

Human proinsulin is a cytosolic 9.4 kDa protein produced by pancreatic beta-cells, and is the precursor of insulin. Murine interleukin-10 is an 18.5 kDa anti-inflammatory secreted cytokine which displays pleiotropic immunoregulatory properties. The HIV-1 *nef* gene product is a 27 kDa cytosolic protein critically important for virus replication and disease progression in vivo. It is highly conserved both in human and simian virus strains (Kestler et al. 1991). However, the precise biological function of NEF remains poorly characterized in vitro, with previous reports suggesting that NEF might be either cytotoxic or cytostatic (Yamada et al. 2003). Up to now there are no immunological studies on the possible use of NEF as a vaccine against HIV because of the difficulties encountered in its expression in animal cells.

The smaller isoform of the enzyme glutamic acid decarboxylase (GAD65) is a major autoantigen in human Type-1 diabetes mellitus (T1DM) (Sanjeevi et al. 1996). Human GAD65 is a 65 kDa protein associated with the cytosolic surface of GABA-containing vesicles. GAD65 has previously been expressed in transgenic tobacco and carrot plants and the recombinant protein was membrane-anchored and demonstrated appropriate immunoreactivity with T1DM-associated autoantibodies (Porceddu et al. 1999). *Petunia hybrida* expansin-1 is a 27 kDa α-expansin protein involved in cell-wall loosening (Zenoni et al. 2004). The goal of expressing proinsulin, interleukin-10, NEF and GAD65 in *N. benthamiana* plants is to produce a large quantity of each recombinant protein, to determine whether plants can be used to produce pharmaceutical proteins.

Materials and methods

Conversion of pPVX201 into a Gateway destination vector

The pPVX201 vector (kindly provided by Prof David Baulcombe, Sainsbury Laboratory, Norwich, UK) comprises a DNA copy of the PVX genome with a duplicated sub-genomic promoter. To make it compatible with the Gateway cloning system, the plasmid was linearized with *Sal*I, purified and treated with Klenow polymerase to generate blunt ends. The resulting molecule was ligated with reading frame cassette A (RfA) of the Gateway® Vector Conversion System (Invitrogen). The RfA Cassette A contains *att*R1, CmR, *ccd*B and *att*R2 sequences. The resulting plasmid, pPVX201_GAT is a destination vector for gene cloning with the Gateway Technology.

Construction of recombinant viral vectors

The five foreign genes were inserted into the pPVX201_GAT destination vector essentially as described in Karimi et al. (2002). In each case, the forward primer contained four additional bases at the 5′ end, not present in the native sequence, which were required for directional cloning.

The human *proinsulin* cDNA (NCBI accession number NM_000207) was amplified using primers PROINSfor (5′-CAC CAT GTT TGT GAA CCA ACA CC-3′) and PROINSrev (5′-CTA GTT GCA GTA GTT CTC CAG C-3′). The murine *interleukin-10* cDNA (NCBI accession number NM_010548) was amplified using primers INTERLfor (5′-CAC CAT GCC TGG CTC AGC ACT GCT TAT AGC TCA TCT TT-3′) and INTERLrev (5′-TTA TAG CTC ATC TTT GCT TTT CAT TTT GAT CAT CAT GTA TG-3′). A TAA stop codon followed by the sequence encoding the KDEL retention signal were added at the 5′ end of the reverse primer. The HIV-1 *nef* cDNA (NCBI accession number AY_734560) was amplified using primers NEFfor (5′-CAC CAT GGG TGG CAA GTG GTC-3′) and NEFrev (5′-TCA GCA GTT CTT GAA GTA CTC CG-3′). The petunia *expansin-1* cDNA (NCBI accession number AY_487167) was amplified from the signal peptide cutting site using primers EXPfor (5′-CAC CAT GGA TTA CAA GGA TGA CGA CGA TAA GAG AAT TCC AGG TGT TTA TAG TG-3′) and EXPrev (5′-TTA AAT TCT AAA GTT CTT TCC C-3′). The forward primer included the CACC sequence and the FLAG epitope tag sequence at the 5′ end. The chimeric *gad67/65* cDNA mutated within the enzyme catalytic site (*gad67/65mut*) was obtained using the human *gad65* cDNA (Hampe et al. 2001) and following the cloning strategy previously described for *gad67/65* (Avesani et al. 2003). The chimeric *gad67/65mut* cDNA was amplified by using the primers GADfor (5′-CAC CAT GGC ATC TTC CAC GCC TT-3′) and GADrev (5′-TTA TTA TAA ATC TTG TCC AAG GCG TTC TA-3′).

The five amplicons were independently inserted into the pENTR™/D-TOPO® vector (Invitrogen). These vectors were sequenced to confirm the absence of errors and then used for LR recombination reactions performed independently for each pENTR/D-TOPO vector with the Gateway destination vector pPVX201_GAT. This resulted in five final vectors: PVX_GAT.proins, PVX_GAT.interl10, PVX_GAT.nef, PVX_GAT.*flag*exp1 and PVX_GAT.gad67/65mut.

Infection of plants

Two 12-day-old *N. benthamiana* plants were inoculated with each PVX vector. Primary-infected plants were obtained by direct inoculation with 100 μg vector DNA at a concentration of 1 μg/μl on two Carborundum-dusted leaves. Twelve days post-inoculation (d.p.i.) tissue was collected from systemically invaded leaves. The collected material was used both for subsequent infection and for molecular analysis. The collected plant material used for infection was homogenised in 50 mM phosphate buffer (pH 7) at a final dilution 1/10 v/v.

The secondary- and the tertiary-infected plants were obtained by inoculation, respectively, with the primary- and secondary-infected plant sap extracts on Carborundum-dusted leaves. Tissue was collected 12 d.p.i. as described above.

RNA isolation, sequencing of the viral recombination region and semi-quantitative real-time RT-PCR

Total RNA was extracted from the leaves of plants infected with each vector using Trizol Reagent (Invitrogen). RNA was pooled from each pair of primary-, secondary- and tertiary-infected plants, and treated with RQ1 DNaseI (Promega).

First strand cDNA synthesis was carried out using ImProm-II™ Reverse Transcriptase (Promega), to generate a template for PCR amplification. Primary-, secondary- and tertiary-infected plants were tested for the presence of the inserted foreign genes using PVX-specific primers PVX for (5′-ACT GGG GAA TCA ATC ACA GTG-3′) and PVXrev (5′-GTC GAA TGC AGC GAA TTT GTG C-3′). These anneal at sites flanking the foreign gene insertion region and generate one or two products depending on whether the foreign gene is still present or has been eliminated. PCR products representing the virus lacking the foreign gene insert were independently cloned in the pGEM®-T Easy vector (Promega) and sequenced using the standard Sp6 forward primer.

Real-time RT-PCR was carried out using a GeneAmp PCR system 9600 (Perkin Elmer) with the SYBR Green PCR Core Reagent (Applied Biosystem) and the following foreign gene-specific primers: *proinsulin*, Proins1 (5′-GGG TCT TGG GTG TGT AG-3′) and Proins2 (5′-CCG CAG CCT TTG TGA AC-3′); *interleukin-10*, Interl1 (5′-GCT GCC TGC TCT TAC TGA-3′) and Interl2 (5′-TAG GAG CAT GTG GCT CTG-3′); *nef*, Nef1 (5′-CCT GAT TGG CAG AAC TAC AC-3′) and Nef2 (5′-TCT CTG GCT CAA CTG GTA CT-3′); *expansin-1*, Exp1 (5′-GAT TAC AAG GAT GAC GAC GAT AAG-3′) and Exp2 (5′-CCA CCA TAA AAA GTA GCA T-3′); and *gad67/65mut*, Gad1 (5′-GTT TGG AGT TGG CAG AGT AAT-3′) and Gad2 (5′-AGA CAT TTG TGT GCT GAG G-3′).

Potato virus X-specific primers PVX1 (5′-GAC TTT GCT GAT CTA TGG-3′) and PVX2 (5′-GAC ACA AGG TTC AGT CCT C-3′) were also used to amplify a virus-specific sequence upstream the exogenous gene insertion site. These primers anneal at sequences within the TGB of the virus.

Quantitative PCR results for the virus TGB and foreign gene transcripts were normalised against the housekeeping gene actin using primers ACT1 (5′-ATC CCA GTT GCT GAC AAT AC-3′) and ACT2 (5′-GGC CCG CCA TAC TGG TGT GAT-3′). Quantitative PCR results for the foreign gene transcripts were also normalised against the results from the virus TGB amplicon obtained using primers PVX1 and PVX2. The data were organized according to the comparative method described in User Bulletin 2 (Applera).

Protein extraction and quantification

Total plant protein was obtained by homogenising infected leaves under liquid nitrogen using a mortar and pestle. The resulting powder was resuspended in 40 mM HEPES (pH 7.3) containing 5 mM DTT and 1.5% CHAPS, supplemented with complete plant protease inhibitor (Sigma). The homogenate was centrifuged at 15,000 *g* for 30 min at 4°C. Solubilized proteins were quantified using Bradford Reagent (Sigma).

Western blot

Western blots were used to detect interleukin-10, expansin-1 and the GAD67/65mut protein. For each blot, equal amounts of total plant protein were separated by SDS-PAGE and transferred onto Immobilon-P membranes (Millipore). Interleukin-10 was detected using a rat anti-mouse interleukin-10 antibody, JES5-2A5 (BD Pharmingen™), diluted 1:5,000. Horseradish peroxidase (HRP)-conjugated rabbit anti-rat was used as secondary antibody (1:6,000). Expansin-1 was detected with the mouse monoclonal antibody, ANTI-FLAG® M2 peroxidase conjugated (Sigma), diluted 1:5,000. The GAD67/65 protein was detected with an anti-GAD65 mouse monoclonal antibody, GC3108 (BioMol), which recognises a linear epitope at the C-terminus of human GAD65, diluted 1:10,000. HRP rabbit anti-mouse was used as secondary antibody (1:15,000). Signal detection was performed with the ECL Western blotting analysis system (GEHealthcare).

Results

Construction of PVX-based vectors and transient expression of foreign proteins

To facilitate the introduction of multiple foreign genes into our PVX expression system, a Gateway-compatible PVX expression vector was constructed and named pPVX201_GAT. This construct is a derivative of pPVX201 and has the Gateway cloning site inserted between the two copies of the CP sub-genomic promoter.

The human *proinsulin* and HIV-1 *nef* cDNAs were independently inserted in the destination vector pPVX201_GAT by LR recombination to generate the vectors PVX_GAT.proins and PVX_GAT.nef. Mouse *interleukin*-10 cDNA, modified firstly by adding the KDEL endoplasmic reticulum retention signal sequence at the C-terminus, was inserted in the PVX_GAT vector, generating PVX_GAT.interl10. The petunia *expansin-1* cDNA, lacking the endogenous signal peptide sequence and supplemented with the FLAG epitope at the N-terminus, was cloned in the pPVX201_GAT vector, generating PVX_GAT.*flag*exp1. Finally, we constructed a chimeric *gad67/65mut* cDNA in which the region encoding the first 87 amino acids of human GAD65mut was substituted with a homologous region of rat GAD67. The chimeric *gad67/65mut* cDNA was cloned in the PVX_GAT vector, generating PVX_GAT.gad67/65mut.

These five PVX expression vectors were used to infect *N. benthamiana* plants with a serial infection procedure. For each vector and for each infection cycle, symptomatic leaves were collected at 12 d.p.i. from plants showing systemic infection, and cellular extracts were used to propagate the infection and for further molecular analysis.

Detection of parental and recombinant PVX RNAs by competitive RT-PCR

To detect PVX RNA we performed an RT-PCR assay using the primers PVXfor and PVXrev, which flank the gene cloning site, and cDNAs derived from primary-, secondary- and tertiary-infected plants. For each construct, two amplicons were expected, one recombinant product containing the foreign gene insert and one parental-like product corresponding to the empty vector. Amplification products were detected by agarose electrophoresis and ethidium bromide staining.

The relative intensity of these two bands clearly separated the samples into two groups. The first group represented constructs with shorter inserts (*proinsulin*, *interleukin-10* and *nef*) and in these cases the parental amplicon appeared only in secondary- and tertiary-infected samples (Fig. 1a). The second group represented constructs with longer foreign gene inserts (*gad67/65mut* and *expansin-1*) and the lower molecular weight parental amplicon was detected as early as the first infection cycle (Fig. 1b).

The smaller parental amplicons detected in secondary-infected leaves were isolated, cloned in a standard vector and sequenced. To determine whether these products could be generated as

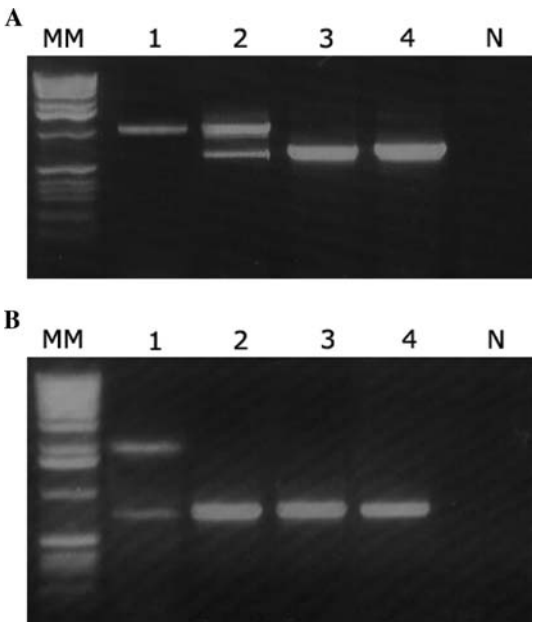

Fig. 1 RT-PCR analysis of cDNA from total RNA extracted from leaves of *proinsulin* (**a**) and *gad67/65mut* (**b**) primary-, secondary- and tertiary-infected plants. Isolated cDNAs were subjected to PCR using the PVX-specific primers PVXfor and PVXrev. *Lane 1*:molecular size marker (*MM*) (1 kb DNA Ladder Promega). *Lanes 2–4*: primary- (*1*), secondary- (*2*) and tertiary-infected (*3*) plants. *Lane 5*: PVX-based vector used for the infection, positive control (*4*). *Lane 6*: healthy plant, negative control (*N*)

PCR artefacts, the same primers were used to amplify DNA directly from the expression constructs using the same thermal profile as the RT-PCR described above. In each case only the larger bands of the expected sizes were produced (data not shown).

Quantitative representation of insert elimination

To produce a quantitative representation of the insert elimination process, the ratio of insert to PVX transcript was studied for the three infection cycles using real-time RT-PCR. DNase I-treated reverse transcribed total RNA which was isolated from two primary-, secondary- and tertiary-infected plants, was amplified using PVX and insert specific primers.

The ratio of *proinsulin* to PVX decreased from 1 in primary-infected plants to 0.6 in secondary-infected plants and 0.002 in tertiary-infected plants (Fig. 2a). A similar profile was seen for *interleukin-10* and *nef* (Fig. 2b, c). The ratio of *expansin-1* to PVX decreased from 1 in primary-infected plants to 0.13 and 0, respectively, in secondary- and tertiary-infected plants (Fig. 2d). The ratio of *gad67/65mut* to PVX fell even more dramatically, from 1 in primary-infected plants to 0.016 in secondary-infected plants and to nearly undetectable levels in the last infection cycle (Fig. 2e).

Expression of interleukin-10, expansin-1 and GAD67/65mut proteins

To correlate the RNA and protein expression profiles, western blots were carried out to detect and quantify the levels of interleukin-10, expansin-1 and GAD67/65mut in infected plants. Interleukin-10 was detectable by western blot analysis in primary- and secondary-infected leaves, while there was no detectable signal in tertiary-infected leaves or in leaves infected with wild-type PVX (Fig. 3a).

A similar profile was obtained for expansin-1 and GAD67/65mut, although in this case there was no detectable protein in the secondary-infected leaves either. These results are shown in Fig. 3b, c.

Sequence analysis of recombinant virus

The process of insert elimination was also investigated by sequencing the region between the PVXfor and PVXrev annealing sites in 20 randomly selected recombinant clones for each construct in secondary- and tertiary-infected plants. Sequence comparisons with the parental construct highlighted the absence of sequences corresponding to the CP duplicated promoter, the *att*B recombination sites and the inserted gene. This structure, resembling the wild-type virus and thus named BW (back to wild type), was probably

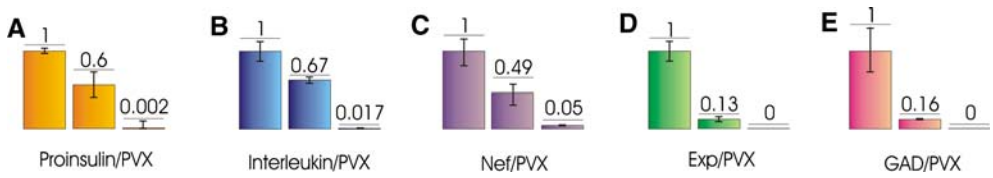

Fig. 2 Real-time RT-PCR analysis. Total RNA was extracted from infected leaves. Each sample represents a pool of two infected plants. Semi-quantitative real-time RT-PCR analysis was performed with specific primers homologous to TGB PVX-specific transcripts for PVX quantification, to actin-specific transcript and to *proinsulin-*, *interleukin-10-*, *nef-*, *expansin-1-* and *gad67/65mut*-specific transcripts. (**a**) Semi-quantitative real-time RT-PCR analysis of *proinsulin* primary- (*bar 1*), secondary- (*bar 2*) and tertiary-infected (*bar 3*) plants. *Proinsulin* expression levels were normalised to PVX (PROINS/PVX). (**b**) Semi-quantitative real-time RT-PCR analysis of *interleukin-10* primary- (*bar 1*), secondary- (*bar 2*) and tertiary (*bar 3*)-infected plants. *Interleukin-10* expression levels were normalised to PVX (INTERLEUKIN/PVX). (**c**) Semi-quantitative real-time RT-PCR analysis of *nef* primary- (*bar 1*), secondary- (*bar 2*) and tertiary-infected (*bar 3*) plants. *Nef* expression levels were normalised to PVX (NEF/PVX). (**d**) Semi-quantitative real-time RT-PCR analysis of *expansin-1* primary- (*bar 1*), secondary- (*bar 2*) and tertiary-infected (*bar 3*) plants. *Expansin-1* expression levels were normalised to PVX (EXP/PVX). (**e**) Semi-quantitative real-time RT-PCR analysis of *gad67/65mut* primary- (*bar 1*), secondary- (*bar 2*) and tertiary-infected (*bar 3*) plants. *Gad67/65mut* expression levels were normalised to PVX (GAD/PVX)

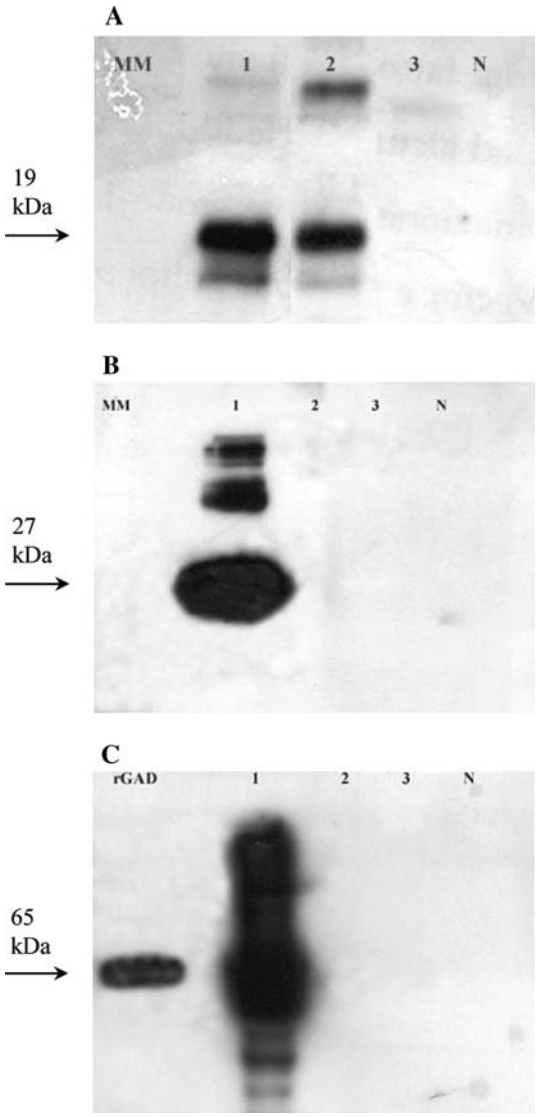

Fig. 3 Western blot analysis of total proteins from PVX-infected plant leaves. (**a**) Interleukin-10 Western blot analysis. In each lane 10 μg of extracted protein was loaded. *Lane 1*: molecular marker. *Lane 2*: primary-infected plants (*1*). *Lane 3*: secondary-infected plants (*2*). *Lane 4*: tertiary-infected plants (*3*). *Lane 5*: negative control (*N*), PVX-infected plant. (**b**) Expansin-1 Western blot analysis. In each lane 10 μg of extracted protein was loaded. *Lane 1*: molecular marker. *Lane 2*: primary-infected plants (*1*). *Lane 3*: secondary-infected plants (*2*). *Lane 4*: tertiary-infected plants (*3*). *Lane 5*: negative control (*N*), PVX-infected plant. (**c**) GAD67/65mut Western blot analysis. In each lane 10 μg of extracted protein was loaded. *Lane 1*: positive control (15 ng of purified recombinant human GAD65, rGAD). *Lane 2*: primary-infected plants (*1*). *Lane 3*: secondary-infected plants (*2*). *Lane 4*: tertiary-infected plants (*3*). *Lane 5*: negative control (*N*), PVX-infected plant

generated by homologous recombination (HR) between the two copies of the sub-genomic CP promoter. Seven clones showed a structure different to that of BW; these structures retaining some fragments of the inserted region, were named RPV (Rearranged Parental Vectors).

Three RPV sequences (two deriving from *gad67/65mut* and one from petunia *expansin-1* contruct) contained an *att*B1 site flanked by two CP promoters (Fig. 4). These structures may be derived from recombination events involving *att*B sites.

The occurrence of BW and RPV structures for each construct are summarized in Table 1. The genetic structure of a BW representative clone and of a representative RPV clone for *interleukin-10*, *nef*, *expansin-1* and *gad67/65mut* are illustrated in Fig. 4. It is interesting to note that the ratio between BW and RPV structures were higher for constructs with short inserts (*proinsulin*, *nef* and *interleukin-10*) than for those with long inserts (*gad67/65mut* and *expansin-1*). An exception to this scenario is represented by *interleukin-10*, which in spite of its length showed an high percentage of RPV structures.

The sequence of one clone, derived from the *expansin-1* secondary-infected plant, revealed a unique structure that could not be classified either as BW or RPV. This contained a 42-bp fragment that was not related to the sequence of the parental vector and whose source is unknown.

Discussion

The use of plant viruses for the expression of proteins in plants has increased significantly over the last decade. Advantages over transgenic plants include higher yields, rapid production timescales and the absence of permanent genetic modification. However, there are also risks associated with the use of virus vectors, and more work needs to be done to characterise these risks in greater detail and develop strategies to deal with them.

Viruses generally have compact genomes and insertions of large genes that are not required for replication and propagation are often lost by natural selection over several replication cycles. Although this observation is cited in many risk assessments, very few studies have thus far

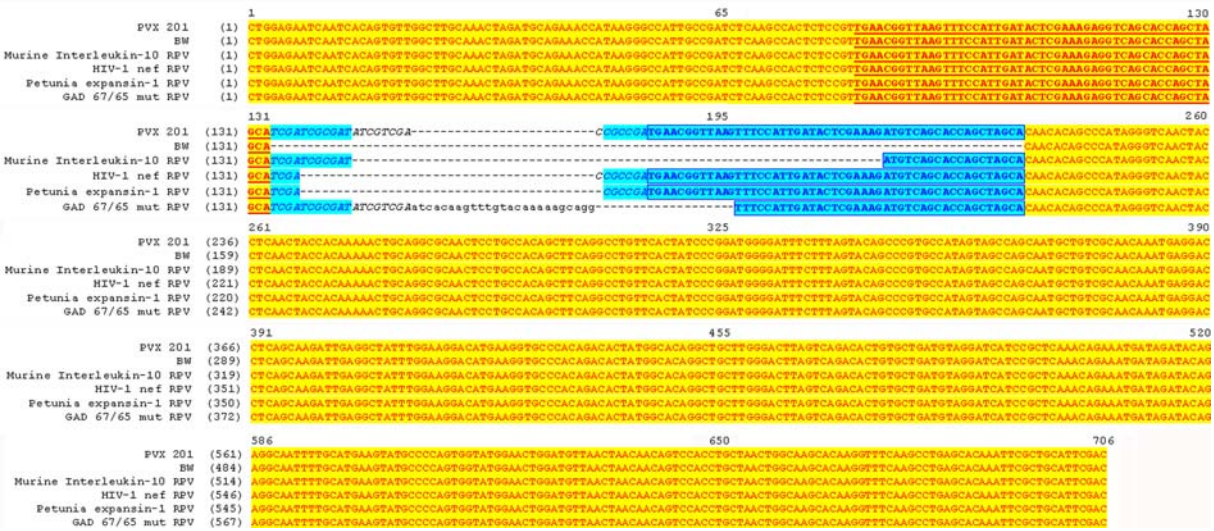

Fig. 4 pPVX201 and recombinant virions sequence comparison. Analysed regions are comprised between PVXfor and PVXrev primer annealing sites. The CP duplicated promoter sequence is *red*, *underlined* and in *bold* whereas CP promoter sequence is *blue* and in *bold uppercase*. The PVX vector sequence with original cloning sites is in *capital letters in cursive*. The *att*B sequences are in *italics*. Comparison included PVX201, BW (represent all virions expressing genes with recombination back to wild type), murine *interleukin-10* RPV, HIV-1 *nef* RPV, petunia *expansin-1* RPV and *gad67/65mut* RPV (all genes RPV represent different types of virion recombination in rearranged parental vector)

examined the relationship between the insert length or sequence and the time taken for its elimination. We have analysed in detail the stability of a PVX vector, pPVX201, harbouring five heterologous sequences—human *proinsulin*, murine *interleukin-10*, HIV-1 *nef*, petunia *expansin-1* and chimeric *gad67/65mut*—ranging in length from 261 to 1,758 bp.

To simplify gene cloning we converted pPVX201 into a Gateway destination vector. Gateway cloning technology facilitates high-throughput cloning of target sequences by making use of the bacteriophage lambda site-specific recombination system.

Table 1 BW and RPV values in percentage for each PVX-based construct used

Gene	Length (bp)	Recombinant virus	
		BW (%)	RPV (%)
Human *proinsulin*	261	100	0
Murine *interleukin-10*	546	17	83
HIV-1 *nef*	620	88	12
Petunia *expansin-1*	794	77	23
gad67/65mut	1,758	37.5	62.5

To express the five proteins in *N. benthamiana* plants we used an experimental design based on three serial infections, thus obtaining primary-, secondary- and tertiary-infected plants. The primary-infected plants were obtained by manual inoculation using vector DNA. The secondary- and tertiary-infected plants were obtained by infections with sap from primary- and secondary-infected leaves, respectively.

Using RT-PCR and real-time RT-PCR, we demonstrated that insert elimination could occur as early as the first infection passage for all five constructs. Furthermore, there was a positive correlation between insert length and the elimination rate. The long inserts *gad67/65mut* (1,758 bp) and *expansin-1* (794 bp) were rapidly lost, being nearly undetectable in the second infection cycle. On the contrary, the *proinsulin*, *interleukin-10* and *nef* sequences were more stable and could be detected in the third infection cycle, although a clear increase of empty vectors was also evident from the second infection cycle.

Experimental evidences, collected from several systems (Bromoviruses, Carmoviruses and *Alfalfa*

Mosaic virus), have shown that recombination in positive sense RNA viruses occurs at high frequency, being a driving force behind virus variability and thus virus evolution (Lai 1992). Viral recombination is classified on the basis of the similarity of parental RNA molecules. RNA recombination is considered homologous (HR) when it occurs between two similar or related RNA molecules at comparable-matched (precise HR) or divergent-matched (imprecise or aberrant HR) cross-over sites. It is considered non-homologous when it involves dissimilar or unrelated RNA molecules (Nagy and Simon 1997).

The most accepted model of RNA recombination is based on template-switching by viral RNA-dependent RNA polymerase (RdRp). It predicts that RdRp switches templates during RNA synthesis, jumping from donor to acceptor RNA, and continues RNA synthesis on the acceptor template, using the nascent RNA as a primer (Lai 1992; Nagy and Simon 1997).

The resulting recombinant pool will be subjected to environmental selective pressure, hence the best-adapted recombinant will replicate and proliferate more efficiently and will become better represented in the recombinant pool.

Chapman et al. (1992) hypothesized that insert elimination in a PVX-based vector expressing the *gusA* gene may occur by both HR and non-HR. The most prevalent deleted form showed a length compatible with unmodified PVX. This suggested that insert–deletion occurred mainly by HR between the duplicated CP promoters. However, there has been no detailed sequence analysis of recombined structures, so doubts could be raised over the generation of new variants.

Tobacco mosaic virus-based vectors have been previously tested for feasibility in mediating recombinant protein expression. The use of a duplicated sub-genomic promoter to drive foreign gene expression was unsuccessful. Upon plant infection the vectors had their foreign sequences deleted, failing to be transported systemically (Dawson et al. 1989). The observed deletion was exact, resulting in wild-type TMV progeny. To avoid viral recombination a hybrid viral RNA vector was constructed containing an heterologous sub-genomic promoter to drive foreign gene expression (Donson et al. 1991).

Our sequence analysis of recombined structures suggests that the recombination sites were situated within the sub-genomic promoter, which could act as an hot spot for HR.

Most of the recombined sequences were very similar to a wild-type PVX, having lost one copy of the CP promoter and all exogenous sequences. Moreover, the moderate variability observed among the PCR products suggested that HR was very precise or that strong selective pressure generated the empty-vector version.

Only three sequences out of 100 analysed contained an *att*B1 site flanked by two CP promoters. These structures may have been formed by an RNA polymerase template switch in correspondence with the *att*B sites. The very low incidence of these structures suggests that, in this context, *att*B sites could be considered as a recombination hot spot of secondary importance.

Two types of structural elements may have played a role in inducing recombination at *att*B sites. Secondary structures at the *att*B site could have halted RNA polymerase, inducing the switch to another molecule in correspondence with the 3′ CP promoter. Alternatively, the partial homology between the *att*B1 and *att*B2 sites could have induced an imperfect template switch which resulted in the deletion of the *att*B2 site. Further constructs with the *att*B sites located in different positions would help in discriminating between the two hypotheses. It is worth mentioning that the *att*B sites employed by the Gateway technology have been mutagenized to minimize secondary structures (Hartley et al. 2000).

Taken together all these observations may suggest that in case of accidental environmental release, recombinant viral vectors would rapidly evolve into wild-type PVX and therefore the risk of heterologous sequence carry-over would be minimal. The low percentage of recombined structures containing fragments of the foreign gene insert indicates that in most cases, the wild-type PVX form is generated by a single HR event which probably involves a template switch during positive strand replication.

Alternatively, it is possible that longer inserts could allow the formation of occasional heteroduplexes, which can promote non-HR. This recombinant form could lead to complete insert

elimination by HR. A time course analysis of the recombination events may provide useful information to test this hypothesis and therefore provide a more complete picture of the dynamics of viral populations derived from vector constructs. An interesting exception to this scenario is represented by *interleukin-10*, which shows an high number of RPV structures. This finding suggests that the PVX *interleukin-10* construct recombined mostly through an heterologous mechanism. Studies conducted in the *Brome Mosaic virus* system have demonstrated that heterologous recombination can be very high for constructs containing recombinant active sites such as inverted repeats, short sequences with high AU contents or regions homologous to viral sequences (Alejska et al. 2005; Shapka and Nagy 2004). We have not identified at first glance such sequences in the primary structure of *interleukin-10* gene and therefore a more detailed study is required.

The vector harbouring the *expansin-1* sequence displayed anomalous recombination. One out of 20 clones presented a sequence of unknown origin. There could be several mechanisms at the basis of the observed phenomenon, mediated by the virus itself, by the host plant or by both.

Assuming a viral role, the mechanism generating these sequences could be the same as that of mammalian Oncoretroviruses. These viruses capture host sequences by recombination, joining viral and host cellular sequences (Vennstrom et al. 1994). Furthermore, evidence suggests that host genes can affect the composition of oncogenic retrovirus populations (Marchlik et al. 2005).

Assuming a plant (host) role, the sequences could have originated from plant mRNA captured by accident. The origin of these sequences could be similar to those that generate ‘‘filler sequences’’ in T-DNA genome integration events (Windels et al. 2003).

References

Alejska M, Figlerowicz M, Malinowska N, Urbanowicz A, Figlerowicz M (2005) A universal BMV-based RNA recombination system—how to search for general rules in RNA recombination. Nucleic Acid Res 33(12):e105

Avesani L, Falorni A, Tornielli GB, Marusic C, Porceddu A, Polverari A, Faleri C, Calcinaro F, Pezzotti M (2003) Improved *in planta* expression of human islet autoantigen glutamic acid decarboxylase (GAD65). Transgenic Res 12:203–212

Baulcombe D, Chapman S, Santa Cruz S (1995) Jellyfish green fluorescent protein as a reporter for virus infections. Plant J 7(6):1045–1053

Canizares MC, Nicholson L, Lomonossoff GP (2005) Use of viral vectors for vaccine production in plants. Immunol Cell Biol 83:263–270

Chapman S, Kavanagh T, Baulcombe D (1992) Potato virus X as a vector for gene expression in plants. Plant J 2(4):549–557

Dawson WO, Lewandowski DJ, Hilf ME, Bubrick P, Raffo AJ, Shaw JJ, Grantham GL, Desjardins PR (1989) A tobacco mosaic virus-hybrid expresses and loses an added gene. Virology 172(1):285–292

Donson J, Kearney CM, Hilf ME, Dawson WO (1991) Systemic expression of a bacterial gene by a tobacco mosaic virus-based vector. Proc Natl Acad Sci USA 88(16):7204–7208

Franconi R, Di Bonito P, Dibello F (2002) Plant derived-human papillomavirus 16 E7 oncoprotein induces immune response and specific tumor protection. Cancer Res 62:3654–3658

Gopinath K, Wellink J, Porta C, Taylor KM, Lomonossoff GP, van Kammen A (2000) Engineering cowpea mosaic virus RNA-2 into a vector to express heterologous protein in plants. Virology 267:159–173

Hampe C, Hammerle L, Falorni A, Robertson S (2001) Site-directed mutagenesis of K396R of the 65 kDa glutamic acid decarboxylase active site obliterates enzyme activity but not antibody binding. FEBS Lett 488:185–189

Hartley JL, Temple GF, Brasch A (2000) DNA cloning using in vitro site-specific recombination. Genome Res 10(11):1788–1795

Karasev AV, Foulke S, Wellens C, Rich A, Shon KJ, Zwierzynski I, Hone D, Koprowski H, Reitz M (2005) Plant based HIV-1 vaccine candidate: Tat protein produced in spinach. Vaccine 23:1875–1880

Karimi M, Inze D, Depicker A (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci 7(5):193–195

Kestler AV, Ringler DJ, Mori K, Panicali DL, Sehgal PK, Daniel MD, Desrosiers RC (1991) Importance of the Nef gene for maintenance of high virus loads for development of AIDS. Cell 65:651–662

Kumagai MH, Turpen TH, Weinzettl N (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

Lai MM (1992) RNA recombination in animal and plant viruses. Microbiol Rev 56(1):61–79

Liu L, Canizares MC, Monger W, Perrin Y, Tsakiris E, Porta C, Shariat N, Nicholson L, Lomonossoff GP (2005) Cowpea mosaic virus-based systems for the production of antigens and antibodies in plants. Vaccine 23:1788–1792

Marchlik E, Kalman R, Rosenberg N (2005) Decreased virus population diversity in p53-null mice infected with weakly oncogenic Abelson virus. J Virol 79(18):11618–11626

Marusic C, Rizza P, Lattanzi L, Mancini C, Spada M, Belardelli F, Benvenuto E, Capone I (2001) Chimeric plant virus particles as immunogens for inducing murine and human immune responses against human immunodeficiency virus type 1. J Virol 75(18):8434–8439

Nagy PD, Simon AE (1997) New insights into the mechanisms of RNA recombination. Virology 235:1–9

O'Brien GJ, Bryant CJ, Voogd C, Greenberg HB, Gardner RC, Bellamy AR (2000) Rotarvirus VP6 expressed by PVX vectors in *Nicotiana benthamiana* coats PVX rods and also assembles into virus like particles. Virology 270:444–453

Porceddu A, Falorni A, Ferradini N, Casentino A, Calicinaro F, Faleri C, Cresti M, Lorenzetti F, Brunetti P, Pezzotti M (1999) Transgenic plants expressing human glutamic acid decarboxylase (GAD65), a major autoantigen in insulin-dependent diabetes mellitus. Mol Breed 5:553–560

Sanjeevi CB, Falorni A, Robertson J, Lernmark A (1996) Glutamic acid decarboxylase (GAD) in insulin-dependent diabetes mellitus. Diabetes Nutr Metab 9:167–182

Santa Cruz S, Chapman S, Roberts AG, Roberts IM, Prior DAM, Oparka KJ (1996) Assembly and movement of a plant virus carrying a green fluorescent protein overcoat. Proc Natl Acad Sci USA 93:6286–6290

Shapka N, Nagy P (2004) The AU-rich RNA recombination hot spot sequence of Brome mosaic virus is functional in tombusviruses: implications for the mechanism of RNA recombination. J Virol 78(5):2288–2300

Toth RL, Chapman S, Carr F, Santa Cruz S (2001) A novel strategy for the expression of foreign genes from plant virus vectors. FEBS Lett 489(2–3):215–219

Uhde K, Fischer R, Commandeur U (2005) Expression of multiple foreign epitopes presented as synthetic antigens on the surface of Potato virus X particles. Arch Virol 150(2):327–340

Vennstrom B, Raynoscheck C, Jansson L, Doederlein G, Lhotak V, Johnson A, Beug H (1994) Retroviral capture of c-erb B proto-oncogene sequences: rapid evolution of distinct viral genomes carrying mutant v-erbB genes with different transforming capacities. Oncogene 9:1307–1320

Verch T, Yusibov V, Koprowski H (1998) Expression and assembly of a full-length monoclonal antibody in plants using a plant virus vector. J immunol Methods 220:69–75

Windels P, De Buck S, Van Bockstaele E, De Loose M, Depicker A (2003) T-DNA integration in Arabidopsis chromosome presence and origin of filler DNA sequences. Plant Physiol 133:2061–2068

Yamada T, Kaji N, Odawara T, Chiba J, Iwamoto A, Kitamura Y (2003) Proline 78 is crucial for human immunodeficiency virus type Nef to down-regulate class I human leukocyte antigen. J Virol 77:1589–1594

Yusibov V, Rabindran S, Commandeur U, Twyman RM, Fischer R (2006) The potential of plant virus vectors for vaccine production. Drugs R D 7(4):203–217

Zenoni S, Reale L, Tornielli GB, Lanfaloni L, Porceddu A, Ferrarini A, Moretti C, Zamboni A, Speghini A, Ferranti F, Pezzotti M (2004) Downregulation of the *Petunia hybrida* alpha-expansin gene PhEXP1 reduces the amount of crystalline cellulose in cell walls and leads to phenotypic changes in petal limbs. Plant Cell 16(2):295–308

Ziegler A, Cowan GH, Torrance L, Ross HA, Davies HV (2000) Facile assessment of cDNA constructs for expression of functional antibodies in plants using the Potato virus X vector. Mol Breed 6:327–335