

Evaluation of potential risks associated with recombination in transgenic plants expressing viral sequences

Camilla Turturo,^{1†} Arianna Friscina,^{1†} Stéphane Gaubert,^{2‡} Mireille Jacquemond,³ Jeremy R. Thompson¹ and Mark Tepfer^{1,2}

Correspondence
Mark Tepfer
tepfer@icgeb.org

¹Plant Virology Group, ICGEB Biosafety Outstation, Via Piovega 23, 31056 Ca' Tron di Roncade, Italy

²INRA, Laboratoire de Biologie Cellulaire, UR501, 78006 Versailles cedex, France

³INRA, Station de Pathologie Végétale, UR407, BP 94, 84143 Montfavet cedex, France

Virus-resistant transgenic plants have been created primarily through the expression of viral sequences. It has been hypothesized that recombination between the viral transgene mRNA and the RNA of an infecting virus could generate novel viruses. As mRNA/viral RNA recombination can occur in virus-resistant transgenic plants, the key to testing this risk hypothesis is to compare the populations of recombinant viruses generated in transgenic and non-transgenic plants. This has been done with two cucumoviral systems, involving either two strains of cucumber mosaic virus (CMV), or CMV and the related tomato aspermy virus (TAV). Although the distribution of the sites of recombination in the CMV/CMV and TAV/CMV systems was quite different, equivalent populations of recombinant viruses were observed in both cases. These results constitute the first comparison of the populations of recombinants in transgenic and non-transgenic plants, and suggest that there is little risk of emergence of recombinant viruses in these plants, other than those that could emerge in non-transgenic plants.

Received 27 July 2007
Accepted 18 September 2007

INTRODUCTION

The first virus-resistant transgenic plants were reported two decades ago (Abel *et al.*, 1986) and, since the early 1990s, various concerns have been expressed regarding the potential impact of such plants, particularly those expressing viral genes (Tepfer, 2002). One of the most intractable questions has been whether viral sequences expressed from the transgene could recombine with the genome of a non-target virus, and thus lead to emergence of a novel virus. As has been reviewed previously (Tepfer, 2002), mRNA/viral RNA recombination has been clearly shown to occur in transgenic plants expressing sequences from several viruses (cauliflower mosaic virus, cowpea chlorotic mottle virus, plum pox virus, tobacco mosaic virus and tomato bushy stunt virus) when they are inoculated with viruses that bear deletions in the sequence expressed from the transgene, but there is no clear evidence that this phenomenon constitutes a risk of virus emergence.

As mRNA/viral RNA recombination can occur in transgenic plants, the critical question is whether this can be the source of recombinants different from those that would be

produced in non-transgenic plants infected with more than one virus, a situation that occurs frequently in the field. Attempts to answer this question should be based on a comparative study of the populations of recombinants that occur under conditions of the lowest possible selection pressure in favour of the recombinants (Tepfer, 2002). This requires development of highly sensitive RT-PCR techniques for specifically amplifying recombinant viral RNAs that occur as a very minor component of the total viral RNA population. This has been shown to be possible with two cucumoviruses, cucumber mosaic virus (CMV) and tomato aspermy virus (TAV) (Aaziz & Tepfer, 1999; de Wispelaere *et al.*, 2005), but was not possible with three other virus groups (Dietrich *et al.*, 2007; Koenig & Buttner, 2004; Meier & Truve, 2006). For this reason, we have developed the cucumovirus recombination system further.

To do this, we have improved the RT-PCR detection of recombinant cucumoviral RNAs and have compared the populations of recombinant RNA3 molecules that appear in transgenic and non-transgenic plants, both between TAV and CMV and also between two isolates representing the major CMV subgroups. In order to increase the possibility of finding differences, these studies were carried out under worst-case conditions, i.e. with transgenic plants that express, in addition to the coat protein (CP), the 3' non-coding region (NCR) of RNA3, as this can serve as an

†These authors contributed equally to this work.

‡Present address: Laboratoire Interactions Plantes–Pathogènes, UMR INRA/INAPG/UnivP6, 16 rue Claude Bernard, 75005 Paris, France.

initiation site for the viral replicase (Teycheney *et al.*, 2000) and thus increase the frequency of appearance of recombinants. Our results show that similar populations of recombinant RNA3 molecules appeared in both experimental systems and suggest that novel recombinants should not be expected in plants expressing this type of CP gene.

METHODS

Viruses, plants and inoculations. R-CMV (Devergne *et al.*, 1969) (subgroup II) and I17F-CMV (Jacquemond & Lot, 1981) (subgroup I) were isolated in the south of France, and P-TAV was isolated in Hungary (Salánki *et al.*, 1994). Cloned cDNAs of R-CMV, I17F-CMV and P-TAV were used initially to establish the viruses (Carrère *et al.*, 1999; Salánki *et al.*, 1997; Fernandez-Delmond *et al.*, 2004), after which they were maintained by regular inoculation of new plants.

All transgenic tobacco plants were derived from *Nicotiana tabacum* 'Xanthi XHFD8'. The transgenic lines were transformed with a vector described previously (Jacquemond *et al.*, 2001), but the transgene included the entire 3' NCR of RNA3 in addition to the CP coding sequence of R-CMV. The transgene is under the control of the cauliflower mosaic virus (CaMV) 35S promoter and the pea E9 RuBisCO small-subunit polyadenylation signal. None of the plant lines studied was resistant to P-TAV or I17F-CMV (M. Jacquemond & M. Tepfer, unpublished).

Plants at the three- to four-leaf stage were dusted lightly with carborundum and rub-inoculated with 20 µl viral crude sap, prepared by grinding symptom-expressing leaves in 0.1 M sodium phosphate buffer, pH 7.0. For recombination experiments, co-infection of non-transgenic plants with P-TAV and R-CMV was carried out by inoculating first with P-TAV, followed by inoculation with R-CMV 5 days post-inoculation (p.i.), as described previously (de Wispelaere *et al.*, 2005). Co-infection with I17F-CMV and R-CMV was done concomitantly, using purified viral RNAs (5 µg each virus ml⁻¹). Homozygous T2 progeny of transgenic plants were inoculated with either P-TAV or I17F-CMV.

RNA extraction. Total RNA was purified from either inoculated or systemically infected leaves. It has been shown previously that equivalent populations of recombinants are observed in inoculated and systemically infected leaves (de Wispelaere *et al.*, 2005). For non-transgenic plants inoculated with P-TAV and R-CMV, total RNA was extracted from samples collected 7 days after the second inoculation (i.e. 12 days after the first inoculation) by using a two-phase phenol procedure (Aaziz & Tepfer, 1999). For non-transgenic plants inoculated with I17F-CMV and R-CMV, total RNA was extracted from samples harvested 7 days p.i. as described previously (Fernandez-Delmond *et al.*, 2004). Samples of transgenic plants were harvested at 10 days p.i. and RNA was purified by a silica-capture procedure (Rott & Jelkmann, 2001). The integrity of all extracted RNA was checked by electrophoresis in a 6.7% formaldehyde, 1% agarose gel buffered with 0.02 M MOPS, pH 7.0.

RT-PCR and cDNA cloning. Primers were given names consisting of three parts: virus strain for which the sequence was specific (R, I or T for R-CMV, I17F-CMV or P-TAV, respectively), nucleotide position on the viral genome of the 5' end of the primer, and sense (+ or -) depending on whether the primer sequence corresponded to the positive or negative viral strand. The presence of locked nucleic acids is indicated by the suffix LNA, and in the primer sequence by upper-case letters.

Primers containing locked nucleic acids were designed by using the tools at two sites (<http://lnatools.com/> and <http://www.genelink.com/>

[newsite/products/LNALockedNucleicAcidDESIGN.asp](http://www.genelink.com/newsite/products/LNALockedNucleicAcidDESIGN.asp)). These sites were particularly useful for determining the proper number and placement of locked nucleic acids within the primers. First-strand cDNAs were synthesized at 52 °C as recommended by the manufacturer of the reverse transcriptase (Promega) with minor modifications, in a 20 µl reaction mixture containing 1 µg total RNA, 140 U Moloney murine leukemia virus reverse transcriptase RNase H- point mutant (Promega) and 0.6 µM primer R2191- (5'-aacctgtggaaaaccaca-3') for I17F-CMV/R-CMV experiments, and R2058-LNA2 (5'-CatCggcgtgtgaaaga-3') for P-TAV/R-CMV ones. The cDNA mixture (3 µl) was submitted to PCR in a 25 µl reaction volume (after 2 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at optimal annealing temperature, 1 min at 68 °C) with 1.5 U Platinum *Taq* (Invitrogen) according to the manufacturer's instructions. Optimal annealing temperatures for primer pairs I1188+LNA3 (5'-atTgtCtaCtgactatata-3') plus R2183- (5'-ggaaaaccacggcgatac-3'), and T1279+ (5'-cgtaataataacaacaactcg-3') plus R2058-LNA2, were 55 and 58 °C, respectively. Because of the variability observed between batches of reverse transcriptase and reverse transcriptase buffer, all experiments included appropriate positive and negative controls.

Positive controls for the detection of recombinant RNA3 molecules consisted of *in vitro* transcripts generated from engineered viral chimeric RNA3 cDNAs. Plasmid pTR3 (Salánki *et al.*, 1997), which contains the 5' part of P-TAV and the 3' part of R-CMV, was used in the P-TAV/R-CMV experiments, and pIR3 (Carrère *et al.*, 1999), which contains the 5' part of I17F-CMV and the 3' part of R-CMV, was used in the I17F-CMV/R-CMV ones. Transcription was carried out by using T7 RNA polymerase (Promega) according to the manufacturer's instructions. Template DNA was subsequently degraded with TURBO DNase (Ambion) and its absence was verified by PCR. RNA was purified by phenol:chloroform extraction and ethanol-precipitated (Maniatis *et al.*, 1982), then quantified spectrophotometrically. The essential negative controls were mixtures of total RNA extracted from transgenic plants or ones infected with only one virus. The absence of amplification from mixtures made either before or after reverse transcription demonstrated the absence of recombination *in vitro* by the reverse transcriptase or *Taq* polymerase, respectively.

Following electrophoresis in a 1% agarose/TAE (40 mM Tris/acetate, 1 mM EDTA) gel, PCR products were purified on MicroSpin S-400 columns (GE Healthcare), cloned in the pGEM-T Easy vector (Promega) and sequenced by BMR Genomics (Padua, Italy). Sequence analyses were carried out by using the Vector NTI Advance 9 software package (Invitrogen). The target was to analyse ten clones from each RT-PCR-positive sample. In some cases, fewer than ten clones were obtained and, for a few samples, more than ten clones were analysed.

RESULTS

Specificity and sensitivity of RT-PCR

Three main aspects of the RT-PCR assay were developed to optimize specificity and sensitivity and to minimize the possibility of recombination *in vitro*. These were: (i) use of a reverse transcriptase with an inactivated RNase H domain, based on reports that RNase H activity is correlated with higher frequencies of template switching both *in vivo* (Hwang *et al.*, 2001; Svarovskaia *et al.*, 2000) and *in vitro* (Fernandez-Delmond *et al.*, 2004); (ii) the optimization of RT-PCR conditions, in particular the incorporation of locked nucleic acids in reverse transcription and PCR

primers (Tolstrup *et al.*, 2003; Wang *et al.*, 1999); and (iii) inclusion in every assay of a robust set of controls. In particular, these included negative controls using combinations of independently extracted plant total RNA containing parental viral or transgene mRNA, in order to eliminate the possibility that the recombinants observed were created *in vitro*. In addition, all plant samples testing positive for recombinants were retested at least once to ensure repeatability of the RT-PCR profile prior to sequence analysis.

Once RNase H and reverse transcriptase(s) had been assessed for efficacy, primers were designed based on standard criteria and, in addition, we also tested primers containing locked nucleic acids for their performance with regards to both specificity and sensitivity. Primers containing locked nucleic acids exhibit enhanced thermal stability when hybridized with their DNA or RNA targets, allowing them to function under more stringent conditions than their counterparts without locked nucleic acids, thereby improving the possibility of discrimination between highly similar RNAs whilst at the same time increasing sensitivity (Tolstrup *et al.*, 2003; Valoczi *et al.*, 2004; Wang *et al.*, 1999).

In order to assess the sensitivity of the RT-PCR assay in the detection of recombinant virus molecules, tenfold serial dilutions of artificial chimeric RNA3 transcribed *in vitro* were diluted in total plant RNA extracted from non-infected tobacco leaves. A comparative RT-PCR assay, carried out with primers with or without locked nucleic acids, showed that an increase in sensitivity was achieved with primers containing locked nucleic acids in both recombinant systems, with a greater yield of amplification product at all dilutions tested (Fig. 1). The limits of detection were 800 and 15 fg for the I17F-CMV/R-CMV and P-TAV/R-CMV systems, respectively. By using these primers, no discernible fragment was amplified when tested against negative-control RNA extracted from transgenic plants infected singly with P-TAV, R-CMV or I17F-CMV,

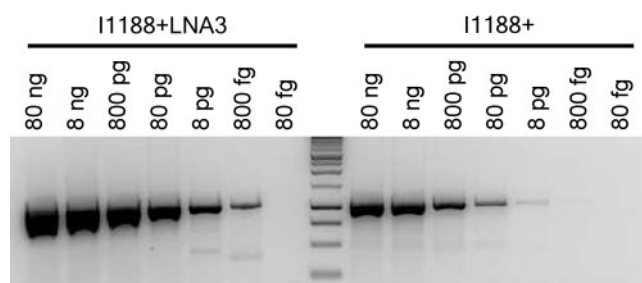


Fig. 1. Comparison of the sensitivity of detection of recombinant viral RNA by RT-PCR using primers with or without locked nucleic acids. Using primer pairs I1188+LNA3/R2189- (left) and I1188+/R2189- (right), RT-PCR analysis was performed with RNA templates synthesized *in vitro*, diluted in a tenfold series (from 80 ng to 80 fg) with 1.0 µg RNA extracted from healthy tobacco plants.

or from combinations of independently extracted plant total RNA containing parental viral or transgenic RNA (Fig. 2). In some experimental samples, occasional additional bands were observed. However, they never appeared in the negative controls and, when sequenced, were found to be non-recombinants derived from one of the parental RNAs or from non-viral sequences (data not shown). For example, this was the case for the upper bands in Fig. 2(b) lanes 13 and 14, and similar non-recombinants also occurred occasionally in the I17F-CMV/R-CMV system described below.

Comparison of recombination sites between P-TAV and R-CMV in non-transgenic and transgenic plants

When 79 non-transgenic plants co-infected with P-TAV and R-CMV were screened by RT-PCR for the presence of recombinant RNA3, seven plants contained putative recombinant molecules (for example, see Fig. 2a). The proportion of plants yielding putative recombinants (8.9%) was comparable with that found in a previous study (9.6%) in which recombinant molecules were detected in the converse direction, i.e. by using a P-TAV specific primer at the 3' NCR and an internal R-CMV primer (de Wispelaere *et al.*, 2005). Putative recombinant RT-PCR products from six plants were cloned and sequenced, which in all cases confirmed their recombinant nature, with crossovers at seven different recombination sites (Table 1). In each plant, one to three crossover sites were observed. Of the 61 cloned recombinant molecules analysed, 60 arose by precise homologous recombination at six sites with 100% identity, with a predominance of crossovers in larger blocks of sequence identity (Fig. 3a). As reflected in both the number of plants in which a given recombination site was observed (Table 1) and the proportion of clones at that site (Fig. 4b), three recombination sites predominated (nt 1940–1965, 1998–2018 and 2020–2038), with a scattering of less frequent crossovers elsewhere. All of the major sites had been identified previously as crossover sites for precise homologous recombination in the converse direction (de Wispelaere *et al.*, 2005), with just one minor site observed in this study that was not observed in the converse, and vice versa. One of the recombination sites that was not in the 3' NCR (one clone at nt 1441–1442) was also the only site at which an imprecise homologous recombinant, in which there was a 2 nt substitution (TC→AG) at the crossover site, was observed.

Transgenic plants expressing a transgene composed of the CP and 3' NCR of R-CMV to high levels were created. They were fully susceptible to I17F-CMV and P-TAV. When 85 homozygous transgenic plants were infected with P-TAV and screened by RT-PCR for the presence of recombinants, 27 plants (32%) contained putative recombinant molecules (for example, see Fig. 2b). RT-PCR products from 12 randomly chosen positive plants were

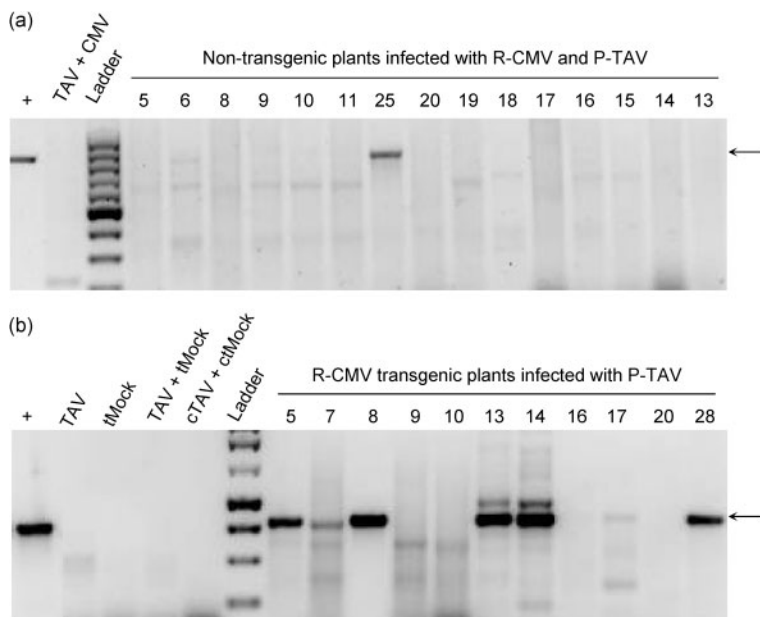


Fig. 2. Detection by RT-PCR of recombinants occurring between P-TAV and R-CMV. (a) Non-transgenic tobacco plants co-infected with P-TAV and R-CMV; (b) R-CMV transgenic plants infected with P-TAV. To the left of the ladders, negative-control lanes include plant total RNA samples from non-transgenic tissue infected with P-TAV (TAV), non-transgenic tissue infected with R-CMV combined with non-transgenic tissue infected with P-TAV (TAV+CMV), mock-inoculated transgenic tissue (tMock), non-transgenic tissue infected with P-TAV combined with mock-inoculated transgenic tissue (TAV+tMock). cDNA produced from non-transgenic tissue infected with P-TAV combined prior to PCR with cDNA produced from mock-inoculated transgenic tissue is also shown (cTAV+ctMock). The positive control (+) was either (a) 15 fg or (b) 150 fg RNA transcribed from the chimeric construct pTR and diluted in plant total RNA taken from non-transgenic tissue. The arrows on the right indicate the expected size of a precisely homologous recombinant (approx. 780 bp).

cloned and sequenced, which in all cases confirmed their recombinant nature (Table 1). Of the 126 recombinant clones analysed, all arose by precise homologous recombination at seven sites with 100% sequence identity, with the major sites of recombination in the larger blocks of sequence identity (Table 1; Fig. 3a). As in the non-transgenic plants, there was good agreement between the number of plants in which a given recombination site was observed (Table 1) and the proportion of clones observed at that site (Fig. 4b). In the transgenic plants, four major sites of recombination predominated (nt 1902–1922, 1940–1965, 1998–2018 and 2020–2038), with minor sites in other blocks of sequence identity. The recombination sites identified were the same as those identified in doubly infected plants, except for two minor crossover sites at nt 1757–1764 and 1977–1981, with one and two clones, respectively. Overall, similar populations of RNA3 recombinants were observed in transgenic and non-transgenic plants (Fig. 4b), and the most important sites were in large blocks of sequence identity (Fig. 3a) and were equivalent to those observed in non-transgenic plants screened for the converse recombinants (de Wispelaere *et al.*, 2005).

Comparison of recombination sites between I17F-CMV and R-CMV in non-transgenic and transgenic plants

Sixty non-transgenic plants co-infected with R- and I17F-CMV were screened by RT-PCR for the presence of recombinants. Five were identified as containing putative recombinant molecules, which was subsequently verified

by cloning and sequencing (Table 1). The proportion of plants containing recombinants (8.3%) was similar to that found in the P-TAV/R-CMV system (8.9%) described above. In each plant, one to six crossover sites were observed (nine in total). Of the 51 cloned recombinant molecules, 47 were of a precise homologous type at eight sites of 100% identity, with one strongly predominant site at nt 1523–1554, as 59% of the total crossovers occurred at this site (Fig. 4c). Imprecise homologous recombination was observed at one site (four clones) located at nt 1680–1695, which included an insert of 25 or 26 nt, corresponding to a duplication of sequence from either parental virus.

Of the 144 transgenic plants inoculated with I17F-CMV screened for recombinant RNA3 by RT-PCR, 38 (26%) yielded putative recombinants. This is similar to what was observed in the P-TAV/R-CMV system (32%). Eight of the positive samples were selected randomly for further characterization; as in the rest of this study, cloning and sequencing confirmed the recombinants. All 80 clones studied were produced by precise homologous recombination in the block of sequence identity at nt 1523–1554 (Table 1; Fig. 3b). This is the same site that predominated among the recombinants observed in non-transgenic plants (Fig. 4c). As in the P-TAV/R-CMV system, the overall pattern of the distribution of recombination sites was equivalent in transgenic and non-transgenic plants. However, it is noteworthy that, although fewer plants and clones were analysed, additional rare recombination sites were observed in non-transgenic plants that were absent in the transgenic ones (Fig. 4c).

(a)

P-TAV 1160
R-CMV 1180 A-CUG--GU-GU-UGUG-UU---UUU-U---UU---GU---GU-U--C-A--G-CA-A---UCU---AU---CCCAA--C--GUA-----G--G-C

P-TAV 1259
R-CMV 1275 G-CG-CC-CGU-G-GGU---UCG-----C-C--CUG-U---GAU--AG---U---GCUUUAC-CAGCA--U-----GCU-A-----

P-TAV 1359
R-CMV 1355 -AUAG---CCUC-----U-GU---CC--C-CUU-A-CA-CC-AC-UU--U-G-UAG--A-A--UGU---C-GGUUA-ACUU-CAC-UC--U-----U--
0/126
1/61

P-TAV 1454
R-CMV 1455 -ACCG-CU--AA-UGAGAA-G-----UUUCGU--AAGUU---U-U-CC-G--C-GU--C-GA-UAU--UAAGAAG---GU-UC--G--U-CAA-U

P-TAV 1554
R-CMV 1555 -AG--U-AA-CC-U--CC-AA-UU--AUUCUAC--U-UGGU--C--UUCG-----U-----A-C--C--UCUU-C---G---C-U-U-U---UG

P-TAV 1654
R-CMV 1655 UU---CGA-GG---U-C--C-GU--U--UUUAUCAG-A-G-U-C--CCGG--UUCA--C-AACAUAAG-U---U--GAC-UGUC--A--U--G-GCUG

P-TAV 1754
R-CMV 1755 A-**AUCGCGA**-AUG-GUAA-UA-GCC-U--U-GU-UAUC-AAAGACGAU--CCU-GA---GA-GA-AU-GU--U-CA--U-GA--U-GAGCA-CA-CG
1/126
0/61

P-TAV 1854
R-CMV 1855 -AUUCCU--C-C---A-GCUCCG---U--U-CG---G--U-C-**CG**-CGUCCGAAGACGUAAACUAC-CU-U-AA-CG---CGAGUG-**CUGAGUUG**
2/126 13/126
3/61 3/61

P-TAV 1953
R-CMV 1948 **GUAGUAUUGCUCUAAACU**-CCUGAAGUC-**CUAAA**-G-----UU--GC-G-**GAACGGUUGUCAUCCAGCU-ACGGCUAAAUGGUCAGUC**-U-UG-U--
14/126 2/126 1/126 82/126 11/126
16/61 0/61 2/61 24/61 12/61

P-TAV 2049
R-CMV 2048 -A-A-GCCG--G-UCUU--A-----G--GAGGU-CC-UUGAA--C-U--UCC--G--UUCUUCGGAAG---G-CUU-----C-U-AGA--CUC---GUG
R2058-

P-TAV 2147
R-CMV 2132 CACGGUA--ACACUGAU--UACC---AGAGUGC-GG-AUCGCCU--GGUUUU-C---G-UC-CC-UA-GG---CCA

(b)

I17F-CMV 1145
R-CMV 1129 GCCCCUUAUUUCUUGGAUGCUUCUCGCG-GAU-GCGUU-U-GU-U--U-AC-----GAGU-GUGUGU-G-U-UGUUUU-U-UUUU---CGU--

I17F-CMV 1244
R-CMV 1219 -A-UG-G-C-A---UGGACAAAUCUG-AUC--CCA-UGCU-GU-G-A-C--CG-CG-CGUCG-CCGCGU-G-GGUUC-CG-UCCGC---UC---GC

I17F-CMV 1341
R-CMV 1312 GGAUGC---UU--G-G--UUG-C-CAGCAG-U---G--ACU-AAUA--AC-U-GC---UGGUCGUCC-ACU-UUAACCACCCAAACUU-GU-GG-AGU
0/80
1/51

I17F-CMV 1441
R-CMV 1412 GAA-**GCUGUA**-ACC-GG-UACAC-UUCACAUCAUUUACCCU-AA-CC-CC---AAAU-GA---GG-UC-UAUU-CGGUA-AAGGUUGU---U-CC-GAUU
0/80
3/51

I17F-CMV 1541
R-CMV 1512 CAGUCACGGA-**UAUGUAAGAAGCUUGUUUCGCGCAUUCAAAU**--G--UUAAUCCUUUGCCGAAAUUGUAUUUAUCCGUGUGGGU-ACAGU-CG-AAAGU
80/80
30/51

I17F-CMV 1641
R-CMV 1612 -CCU-C-UC-UC-GA--U-UCCGU-**GCCGCCAUCUCUGCUAUGUU**-G--GA-GG---UCACCGGU-**UGGUUUUAUCAGUAUGC**-GC-UC--GAGU-CA-
0/80 0/80 0/80
1/51 3/51 4/51

I17F-CMV 1741
R-CMV 1712 GCCAACAA-AA--U--U-UAUGA-CU-UC-G-GAUGCGCGCUGAUUAU-GG-GACAUG-G-AAGUACGCCGUCCU-GU-UAUUC-AAAGACGAU---CUAG

I17F-CMV 1841
R-CMV 1812 AGA-GGACGAG-U-**GUACUUCAGU**-GAC-UCGAGCA-CAACG-AUUC-A--UC--G--UGCUCG-C-GU-**A**-U-CGUGU---U-CC-G-UCC---
0/80
2/51

I17F-CMV 1937
R-CMV 1911 -----CU-C--UCUC--U-GCG-G-GCUGAGUUG-**AGU**--UGCU-UAAACUG-CUGAAGUC-CUAAACGU-UU-----CGG-**GAACGGGUUGUCC**
0/80
1/51

I17F-CMV 2025
R-CMV 2011 **AUCCAGCUUACGGCUAAAUGGUCAGUCGUG**-----C-ACGCC--U-G--UUACAA-AU-UC-GAGG--CC-UUGAAA-CAUCUCCUAG-UUUUCUUG
0/80
6/51

I17F-CMV 2124
R-CMV 2108 GAAGG-CUUCG---G-G-A-CUC--GCAC-----A--UG-UA-U--CA-G-GU-CGGGU--C-CC-----UUUC---GG--CUCCAUAAGGAGAC
R2183-
R2191-

Fig. 3. Position of recombination crossover sites in transgenic and non-transgenic plants. The consensus nucleotide sequence of an alignment between the 3' half of RNA3 of (a) P-TAV and R-CMV and (b) I17F-CMV and R-CMV is shown. Bases that are shown are identical in both viruses. Dashes represent lack of identity. Bases highlighted in grey indicate blocks of identity in which recombination was detected, with the proportion of total recombinants for each system represented as a fraction. For the transgenic and non-transgenic systems, these figures are shown above and below the sequence, respectively. Respective nucleotide positions are shown on the left. The arrows show the position of the primers used for RT-PCR. CP start and stop codons are shown in italics and in bold type without grey highlighting, respectively, with a line above or below the consensus sequence depending on which virus is referred to, based on the alignment.

Table 1. Number and distribution of recombination sites in co-infected non-transgenic plants and singly infected transgenic plants. The upper part of the table concerns P-TAV/R-CMV recombinants, and the lower part I17F-CMV/R-CMV recombinants. On the left are data from co-infected non-transgenic plants, and on the right data for singly infected transgenic ones.

Recombination site*		Co-infected non-transgenic plants						Total no. plants per site	Infected transgenic plants										Total no. plants per site			
P-TAV	R-CMV	25†	30	33	37	68	88		5	8	12	13	14	15	17	19	28	37	38	85		
1441–1442	1442–1443‡	–	–	–	–	–	1‡§	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1756–1763	1757–1764	–	–	–	–	–	–	0	–	–	–	–	–	–	–	–	–	–	1	–	–	1
1900–1901	1899–1900	–	–	–	3	–	–	1	–	1	–	–	–	–	–	–	1	–	–	–	–	2
1903–1923	1902–1922	–	3	–	–	–	–	1	–	–	1	–	2	6	–	–	–	–	–	–	4	4
1945–1970	1940–1965	1	1	2	5	7	–	5	3	4	–	–	1	4	–	1	–	–	1	–	–	6
1982–1986	1977–1981	–	–	–	–	–	–	0	–	–	–	–	–	–	–	1	1	–	–	–	–	2
1994–1995	1993–1994	–	–	–	2	–	–	1	1	–	–	–	–	–	–	–	–	–	–	–	–	1
1999–2019	1998–2018	3	6	4	–	2	9	5	5	5	8	17	17	–	3	–	5	10	12	–	–	9
2021–2039	2020–2038	7	–	4	–	1	–	3	1	–	1	–	1	–	3	5	–	–	–	–	–	5
Total no. clones per plant		11	10	10	10	10	10		10	10	10	17	21	10	6	7	7	10	14	4		
Total no. sites per plant		3	3	3	3	3	2		4	3	3	1	4	2	2	3	3	1	3	1		
I17F-CMV	R-CMV	5	7	13	31	41			8	67	73	78	79	83	86	90						
1401–1409	1372–1380	–	–	–	1	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1445–1450	1416–1421	–	3	–	–	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1552–1583	1523–1554	9	2	10	1	8	–	5	10	10	10	10	10	10	10	10	10	10	10	10	10	8
1666–1685	1637–1656	–	–	–	1	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1699–1706	1670–1677	–	3	–	–	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1709–1724	1680–1695‡	1‡	3‡	–	–	–	–	2	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1855–1865	1826–1836	–	–	–	2	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
1968–1970	1950–1952	–	–	–	1	–	–	1	–	–	–	–	–	–	–	–	–	–	–	–	–	0
2012–2055	1998–2041	–	–	–	4	2	–	2	–	–	–	–	–	–	–	–	–	–	–	–	–	0
Total no. clones per plant		10	11	10	10	10			10	10	10	10	10	10	10	10	10	10	10	10	10	10
Total no. sites per plant		2	4	1	6	2			1	1	1	1	1	1	1	1	1	1	1	1	1	1

*RNA3 nucleotide positions of the block of sequence identity in which recombination occurred.

†Numbers refer to individual plants.

‡Imprecise homologous recombination site.

§Numbers refer to the number of recombinant clones. – indicates no recombinant clones with a crossover at this site.

DISCUSSION

During genome replication, RNA viruses produce numerous variants due to misincorporation and recombination, upon which negative and positive selection act to eliminate less fit molecules and to favour the best-adapted ones. The variability created by recombination is a key element of virus evolution, as it is at the origin of numerous evolutionary branch points. However, as the dominant molecular types are generally well adapted to their usual host plants, the constant creation and elimination of recombinants remains relatively invisible. Nonetheless, it is precisely this population of rare recombinants that must be compared in transgenic and non-transgenic plants, to determine whether the former could be the source of novel variants that could lead to emergence of new viral diseases. In a broader context, to understand the role of recombination in virus adaptation and evolution, the most

important questions concern the structure of populations of recombinant viral genomes that appear in infected plants and the fitness of individual recombinant types in appropriate hosts. The cucumovirus systems exploited here are, at least to our knowledge, the only ones available so far where this type of study is possible for plant viruses.

The overall pattern of P-TAV/R-CMV recombinants observed, with a predominance of homologous precise recombination in the larger blocks of sequence identity, is similar to that observed previously in the converse direction (de Wispelaere *et al.*, 2005). However, there were two salient differences between the results of the two studies. For R-CMV/P-TAV recombinants, an extremely strong hot spot for recombination was observed by de Wispelaere *et al.* (2005) at the GC dinucleotide at positions 1899–1900, which was not a hot spot for P-TAV/R-CMV recombination in either the transgenic or non-transgenic

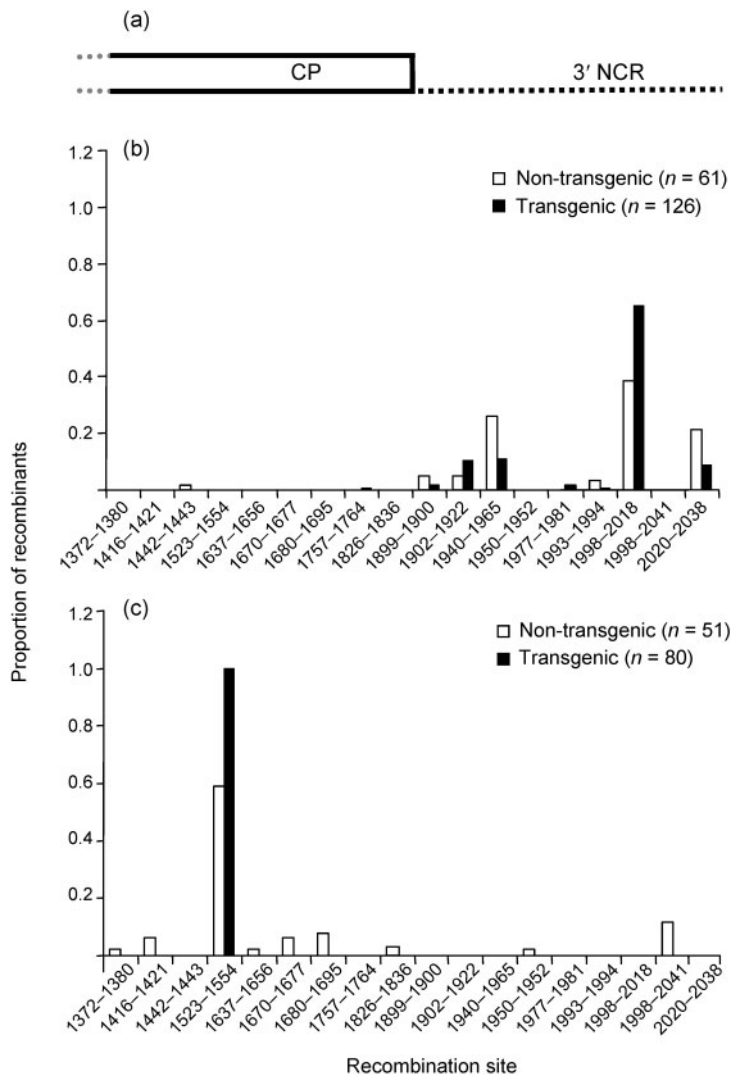


Fig. 4. Comparison of sites of recombination in transgenic and non-transgenic plants. (a) Schematic diagram of the part of RNA3 studied (not drawn to scale), showing the CP ORF, and the 3' NCR. (b, c) Location and proportion of recombination events at specific sites relative to the total number of recombinants in infected transgenic and co-infected non-transgenic plants between (b) P-TAV and R-CMV and (c) I17F-CMV and R-CMV. The numbers shown on the *x*-axis are the R-CMV coordinates of the blocks of sequence identity in which recombination was observed between R-CMV and either P-TAV (b) or I17F-CMV (c). Using the Pearson product-moment correlation method, *r* values for transgenic vs non-transgenic were 0.84 and 0.98 for R-CMV/P-TAV and R-CMV/I17F-CMV, respectively.

plants studied in the present work. In addition, more than one-third of the R-CMV/P-TAV recombinants were of an aberrant type, resulting in a 160–180 nt duplication in the 3' NCR (de Wispelaere *et al.*, 2005). As similar recombinants have also been observed in certain natural CMV and TAV isolates (Chen *et al.*, 2002; Moreno *et al.*, 1997), it is surprising that this type of recombinant was not detected in the present study, but this may be due to the instability in tobacco plants of I17F-CMV/R-CMV recombinants of this type (Pierrugues *et al.*, 2007).

The results presented here were obtained under conditions of low selection pressure, using wild-type virus inocula that are not targeted by the virus-resistance transgene, in order to observe the maximum number of types of recombinant produced. The recombination sites were extremely different for I17F-CMV/R-CMV and P-TAV/R-CMV, with, in the latter case, a single predominant hot spot for recombination in a large block of sequence identity within the CP gene (nt 1523–1554). As no crossovers were observed in nearby blocks of sequence identity of similar

length, this hot spot may be determined by the presence of strong RNA secondary-structural features. However, with both experimental systems, we have observed equivalent populations of recombinant cucumoviruses in transgenic and non-transgenic plants, nearly all of which were of the precise homologous type.

As the recombinants described here have never been observed in the field, many of them may have reduced fitness. In a parallel study, the fitness of certain I17F-CMV/R-CMV and R-CMV/I17F-CMV recombinants was evaluated (Pierrugues *et al.*, 2007). It is particularly interesting to note that the predominant I17F-CMV/R-CMV recombinant observed here under conditions of low selection pressure was in fact non-viable when inoculated alone into several CMV host plants, including tobacco.

In both experimental systems, recombinants were detected in a higher proportion of transgenic plants than in non-transgenic ones. In addition, there was a tendency to observe stronger RT-PCR amplification from samples from

infected transgenic plants. Neither of these features is surprising because, as has been pointed out previously (Tepfer, 2002), recombinants can only occur in the small proportion of cells of non-transgenic plants in which both viruses are replicating, whereas the transgene mRNA and viral RNAs will be present in all infected cells. As explained in the Introduction, these experiments were carried out with transgenic plants designed to favour recombination, as the transgene includes the viral 3' NCR. These results give further support to the broadly supported suggestion that virus-resistant transgenic plants for commercial release should not include the 3' NCR, in order to prevent initiation on the transgene mRNA by the viral replicase and thus strongly reduce the appearance of recombinants; this has been shown to be highly effective in two other viral systems (Tepfer, 2002).

It is important to consider the limits to impose on the interpretation of the results of this type of study. First, as there are no equivalent results with other virus groups, it is impossible to know to what extent the conclusions can be generalized. Second, these experiments were carried out under highly controlled conditions, so it would be of great interest to determine whether they can be validated under the more variable conditions encountered in the field. Third, as it is impossible to sample the molecules present exhaustively, we cannot formally exclude the possibility that extremely rare recombinants could be present in transgenic plants, but absent in non-transgenic ones. Here, it should be noted that we have observed the opposite tendency, as we detected I17F-CMV/R-CMV recombinants in non-transgenic plants that were absent in the transgenic ones. In any case, in addition to exclusion from the transgene of the 3' NCR, there is an additional feature of the virus cycle that would further reduce the likelihood of emergence of rare recombinants, even if their intrinsic fitness were high. This is the presence of at least two narrow population bottlenecks, one at the stage of systemic movement of the virus and the other at vector transmission, which will strongly favour the most abundant RNAs (Ali *et al.*, 2006; Li & Roossinck, 2004).

We have observed equivalent populations of recombinant cucumoviral RNAs in transgenic and non-transgenic plants, which argues strongly in favour of the same mechanism(s) of recombination being at work in both types of plant. Thus, our results are clearly not compatible with the hypothesis that novel recombinant viral genomes would be produced in transgenic plants, and thus are contrary to the hypothesis that these plants would increase the likelihood of emergence of new viruses, which would in addition require that any novel recombinant be able to compete favourably with the parental viruses.

Transgenes expressing a viral CP sequence can confer resistance via two types of mechanisms (reviewed by Goldbach *et al.*, 2003). In some systems, including CMV-resistant tobacco (Jacquemond *et al.*, 2001), the CP itself is involved in resistance, in which case, high levels of transgene mRNA accumulate. In other systems, resistance

is mediated by post-transcriptional gene silencing (PTGS) and, in this case, only low levels of mRNA are observed. Clearly, the issue of potential impacts of mRNA/viral RNA recombination is more pertinent in the former cases and, indeed, it has been proposed that it would be preferable to use PTGS-mediated transgenic resistance in order to reduce the probability of appearance of recombinants (Tepfer, 2002). However, when PTGS-mediated resistance breaks down, this may create circumstances particularly favourable to recombination in transgenic plants (Jakab *et al.*, 1997).

As our results showed no appreciable risk due to recombination in plants that accumulated high levels of transgene mRNA, we suggest that this should not be a major criterion for selecting a virus-resistance strategy and that far more important criteria would include the efficacy, breadth and stability of the resistance obtained.

ACKNOWLEDGEMENTS

We thank Laura Chiappetta for growing the plants used. This research was supported by grants from the Fondazione Cassamarca, the EC (contract QLK3-CT-2000-00361) and the French ANR (contract ANR-05-POGM-001).

REFERENCES

- Aaziz, R. & Tepfer, M. (1999). Recombination between genomic RNAs of two cucumoviruses under conditions of minimal selection pressure. *Virology* **263**, 282–289.
- Abel, P. P., Nelson, R. S., De, B., Hoffmann, N., Rogers, S. G., Fraley, R. T. & Beachy, R. N. (1986). Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science* **232**, 738–743.
- Ali, A., Li, H., Schneider, W. L., Sherman, D. J., Gray, S., Smith, D. & Roossinck, M. J. (2006). Analysis of genetic bottlenecks during horizontal transmission of cucumber mosaic virus. *J Virol* **80**, 8345–8350.
- Carrère, I., Tepfer, M. & Jacquemond, M. (1999). Recombinants of cucumber mosaic virus (CMV): determinants of host range and symptomatology. *Arch Virol* **144**, 365–379.
- Chen, Y. K., Goldbach, R. & Prins, M. (2002). Inter- and intramolecular recombinations in the cucumber mosaic virus genome related to adaptation to *Alstroemeria*. *J Virol* **76**, 4119–4124.
- de Wispelaere, M., Gaubert, S., Trouilloud, S., Belin, C. & Tepfer, M. (2005). A map of the diversity of RNA3 recombinants appearing in plants infected with cucumber mosaic virus and tomato aspermy virus. *Virology* **331**, 117–127.
- Devergne, J. C., Cardin, L. & Marais, A. (1969). Etude d'une souche du virus de la mosaïque du concombre isolée en France de *Ranunculus asiaticus* L. *Ann Phytopathol* **1**, 329–337 (in French).
- Dietrich, C., Miller, J., McKenzie, G., Palkovics, L., Balázs, E., Palukaitis, P. & Maiss, E. (2007). No recombination detected in artificial potyvirus mixed infections and between potyvirus derived transgenes and heterologous challenging potyviruses. *Environ Biosafety Res* **6**, 207–218.
- Fernandez-Delmond, I., Pierrugues, O., de Wispelaere, M., Guilbaud, L., Gaubert, S., Divéki, Z., Godon, C., Tepfer, M. &

- Jacquemond, M. (2004).** A novel strategy for creating recombinant infectious RNA virus genomes. *J Virol Methods* **121**, 247–257.
- Goldbach, R., Bucher, E. & Prins, M. (2003).** Resistance mechanisms to plant viruses: an overview. *Virus Res* **92**, 207–212.
- Hwang, C. K., Svarovskaia, E. S. & Pathak, V. K. (2001).** Dynamic copy choice: steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of *in vivo* template switching. *Proc Natl Acad Sci U S A* **98**, 12209–12214.
- Jacquemond, M. & Lot, H. (1981).** L'ARN satellite du virus de la mosaïque du concombre. I. Comparaison de l'aptitude à induire la nécrose létale de la tomate d'ARN satellites isolés de plusieurs souches de virus. *Agronomie* **1**, 927–932 (in French).
- Jacquemond, M., Teycheney, P.-Y., Carrère, I., Navas-Castillo, J. & Tepfer, M. (2001).** Resistance phenotypes of transgenic tobacco plants expressing different cucumber mosaic virus (CMV) coat protein genes. *Mol Breed* **8**, 85–94.
- Jakab, G., Vaistij, F. E., Droz, E. & Malnoë, P. (1997).** Transgenic plants expressing viral sequences create a favourable environment for recombination between viral sequences. In *Virus-Resistant Transgenic Plants: Potential Ecological Impact*, pp. 45–51. Edited by M. Tepfer & E. Balazs. Heidelberg: Springer.
- Koenig, R. & Buttner, G. (2004).** Strategies for the detection of potential beet necrotic yellow vein virus genome recombinations which might arise as a result of growing A type coat protein gene-expressing sugarbeets in soil containing B type virus. *Transgenic Res* **13**, 21–28.
- Li, H. & Roossinck, M. J. (2004).** Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J Virol* **78**, 10582–10587.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982).** *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Meier, M. & Truve, E. (2006).** An attempt to identify recombinants between two sobemoviruses in doubly infected oat plants. *Environ Biosafety Res* **5**, 47–56.
- Moreno, I. M., Bernal, J. J., Garcia de Blas, B., Rodriguez-Cerezo, E. & Garcia-Arenal, F. (1997).** The expression level of the 3a movement protein determines differences in severity of symptoms between two strains of tomato aspermy cucumovirus. *Mol Plant Microbe Interact* **10**, 171–179.
- Pierrugues, O., Guilbaud, L., Fernandez-Delmond, I., Fabre, F., Tepfer, M. & Jacquemond, M. (2007).** Biological properties and relative fitness of inter-subgroup cucumber mosaic virus RNA 3 recombinants produced *in vitro*. *J Gen Virol* **88**, 2852–2861.
- Rott, M. E. & Jelkmann, W. (2001).** Characterization and detection of several filamentous viruses on cherry: adaptation of an alternative cloning method (DOP-PCR), and modification of an RNA extraction protocol. *Eur J Plant Pathol* **107**, 411–420.
- Salánki, K., Balázs, E. & Burgyán, J. (1994).** Nucleotide sequence and infectious *in vitro* transcripts of RNA 3 of tomato aspermy virus pepper isolate. *Virus Res* **33**, 281–289.
- Salánki, K., Carrère, I., Jacquemond, M., Balázs, E. & Tepfer, M. (1997).** Biological properties of pseudorecombinant and recombinant strains created with cucumber mosaic virus and tomato aspermy virus. *J Virol* **71**, 3597–3602.
- Svarovskaia, E. S., Delviks, K. A., Hwang, C. K. & Pathak, V. K. (2000).** Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *J Virol* **74**, 7171–7178.
- Tepfer, M. (2002).** Risk assessment of virus-resistant transgenic plants. *Annu Rev Phytopathol* **40**, 467–491.
- Teycheney, P.-Y., Aaziz, R., Dinant, S., Salánki, K., Tourneur, C., Balázs, E., Jacquemond, M. & Tepfer, M. (2000).** Synthesis of (–)-strand RNA from the 3' untranslated region of plant viral genomes expressed in transgenic plants upon infection with related viruses. *J Gen Virol* **81**, 1121–1126.
- Tolstrup, N., Nielsen, P. S., Kolberg, J. G., Frankel, A. M., Vissing, H. & Kauppinen, S. (2003).** OligoDesign: optimal design of LNA (locked nucleic acid) oligonucleotide capture probes for gene expression profiling. *Nucleic Acids Res* **31**, 3758–3762.
- Valoczi, A., Hornyik, C., Varga, N., Burgyan, J., Kauppinen, S. & Havelda, Z. (2004).** Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. *Nucleic Acids Res* **32**, e175.
- Wang, G., Gunic, E., Girardet, J. L. & Stoisavljevic, V. (1999).** Conformationally locked nucleosides. Synthesis and hybridization properties of oligodeoxynucleotides containing 2',4'-C-bridged 2'-deoxynucleosides. *Bioorg Med Chem Lett* **9**, 1147–1150.