Analysis of recombination between viral RNAs and transgene mRNA under conditions of high selection pressure in favour of recombinants

Marco Morroni,1,2 Jeremy R. Thompson1 and Mark Tepfer1

Correspondence
Mark Tepfer
tepfer@icgeb.org

1Plant Virology Group, ICGEB Biosafety Outstation, Via Piovega 23, 31056 Ca’ Tron di Roncade, Italy
2Dipartimento di Produzione Vegetale, Università degli Studi di Milano, via Celoria 2, 20133 Milano, Italy

One possible environmental risk related to the utilization of virus-resistant transgenic plants expressing viral sequences is the emergence of new viruses generated by recombination between the viral transgene mRNA and the RNA of an infecting virus. This hypothesis has been tested recently for cucumber mosaic virus (CMV) by comparing the recombinant populations in transgenic and non-transgenic plants under conditions of minimal selection pressure in favour of the recombinants. Equivalent populations were observed in transgenic and non-transgenic plants but, in both, there was a strongly dominant hotspot recombinant which was shown recently to be nonviable alone in planta, suggesting that its predominance could be reduced by applying an increased selection pressure in favour of viable recombinants. Partially disabled I17F-CMV mutants were created by engineering 6 nt deletions in five sites in the RNA3 3′-non-coding region (3′-NCR). One mutant was used to inoculate transgenic tobacco plants expressing the coat protein and 3′-NCR of R-CMV. A total of 22 different recombinant types were identified, of which 12 were, as expected, between the transgene mRNA and the mutated I17F-CMV RNA3, while 10 resulted from recombination between the mutated RNA3 and I17F-CMV RNA1. Twenty recombinants were of the aberrant type, while two, including the dominant one detected previously under conditions of minimal selection pressure, were homologous recombinants. All recombinants detected were very similar to ones observed in nature, suggesting that the deployment of transgenic lines similar to the one studied here would not lead to the emergence of new viruses.

INTRODUCTION

Recombination plays a critical role in the variation observed in plus-sense RNA viruses, although the diversity of recombinant molecules in any virus population and the possible mechanism(s) underlying their production are only now beginning to be understood. From a safety standpoint, recombinant viruses are a source of concern, as they can lead to virus emergence. It has been known for several years that recombination between viral sequences expressed in plants as transgene mRNAs and the genomic RNAs of a related virus can occur. This raises the question of whether recombination in virus-resistant transgenic plants expressing viral sequences could favour emergence of novel recombinant viruses (for review, see Tepfer, 2002). The preferred strategy to evaluate this potential risk is to compare the populations of recombinants in transgenic and non-transgenic plants under conditions of minimal selection pressure in favour of the recombinants (Aaziz & Tepfer, 1999), but this has proved to be difficult to achieve. In a recent study, Turturo et al. (2008) were the first to succeed, using transgenic plants expressing the coat protein (CP) and the 3′-non-coding region (3′-NCR) of RNA3 of a subgroup II strain of cucumber mosaic virus (R-CMV). The 3′-NCR can serve as an initiation site for the viral replicase (Teycheney et al., 2000), thus making it possible for full-length recombinant RNA3 that can potentially be replicated to be produced by a single crossover. Turturo et al. (2008) showed clearly that the populations of recombinant viral RNAs were similar in the transgenic plant system when infected with a subgroup I CMV (I17F-CMV) and in the non-transgenic plants infected simultaneously with both CMV strains. However, the populations were, in both cases, completely dominated by recombinants at a single hotspot. Moreover, when the hotspot recombinant was tested alone by inoculation onto tobacco plants, it was found to be non-viable (Pierrugues et al., 2007).
Considering that in other experimental systems using cucumoviruses, aberrant recombinants were frequently reported (de Wispelaere et al., 2005; Fernandez-Cuartero et al., 1994; Masuta et al., 1998; Shi et al., 2007; Suzuki et al., 2003), it was surprising that none were observed by Turturo et al. (2008), suggesting that the prevalence of the hotspot masked the presence of other recombinant molecules. In order to test this hypothesis, plants of the same transgenic line were inoculated with an attenuated mutant of I17F-CMV, so that recombinants that restore the viral genome would be favoured by selection pressure. As predicted, when the hotspot recombinant observed under conditions of minimal selection pressure was strongly counter-selected, numerous additional recombinant types were observed.

**METHODS**

**Plants, viruses and inoculations.** Tobacco (*Nicotiana tabacum* Xanthi XH8D8) plants were maintained in greenhouse conditions of 25 ± 2 °C and 75 ± 10% relative humidity with a 16 h day length. Transgenic line 40.5 expresses the CP and entire 3'-NCR of R-CMV RNA3 (subgroup II) (Turturo et al., 2008), and is not resistant to I17F-CMV (subgroup I). Cloned cDNAs of RNAs 1, 2 and 3 (pI1T7, pI2T7 and pI3T7, respectively) of I17F-CMV (Jacquemond & Lot, 1981) (kindly provided by M. Jacquemond, INRA, Avignon, France) were linearized and plants were inoculated as described previously (Thompson et al., 2008). Leaf samples were harvested at 8, 10 and 15 days post-inoculation (p.i.), frozen in liquid nitrogen and stored at −80 °C.

**Deletion mutagenesis.** Five different mutants were created by deleting six nucleotides at different positions in the 3'-NCR of I17F-CMV RNA3 (Fig. 1). The primers for mutagenesis (see Supplementary Table S1) consisted of 15 nucleotides 5' and 3' of the deleted nucleotides (Sigma Aldrich). During a first PCR, Mut. primers + and – were used in combination with primers M13R (5'-gtacagcatgcaacagg-3') and I1178+ (5'-gattgcttattgtctactg-3') to generate fragments A and B, respectively. Reagents for a reaction volume of 50 µl were: 0.5 µl *Pfu* Polymerase 3 U µl⁻¹ (Promega), 5 µl *Pfu* buffer (10 ×), 4 µl dNTPs (10 µM), 1 ng pI3T7 and 1 µl forward and reverse primers (10 µM). Thermocycler conditions were: 94 °C for 2 min (1 ×); 94 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min (35 ×); and 72 °C for 10 min (1 ×). In a second PCR, fragments A and B were fused in order to obtain a unique fragment containing the deletion. Reagents for a reaction volume of 50 µl were: 0.5 µl *Pfu* Polymerase 3 U µl⁻¹ (Promega), 5 µl *Pfu* buffer (10 ×), 4 µl dNTPs (10 µM), 0.5 µl A and B and 1 µl primer I1178+ and M13R (10 µM). Thermocycler conditions were: 94 °C for 5 min (1 ×); 94 °C for 1 min, 55 °C for 1 min, 72 °C for 2 min (35 ×); and 72 °C for 10 min (1 ×). The PCR results were verified by gel electrophoresis of 3 µl of the reaction products in a 1.2% agarose gel in TAE (40 mM Tris/acetate, 1 mM EDTA). The samples of the second PCR were purified by using the JETQUICK kit (Genomed) and then digested alongside pI3T7 by using *Bst* (20 U µl⁻¹) and *Sal* (10 U µl⁻¹) (New England Biolabs). Digested fragments were separated in a 1% agarose/TAE gel, purified using the QIAquick Gel Extraction kit (Qiagen) and ligated with T4 DNA ligase by a standard ligation procedure described by the manufacturers (Fermentas).

**RNA extraction.** Total RNA was purified from systemically infected leaves by using a two-phase phenol procedure described previously (Aaziz & Tepfer, 1999) with minor changes. Leaf material (0.4 g) was ground in 800 µl buffer solution (200 mM Tris/HCl, 100 mM LiCl, 5 mM EDTA, 1% SDS), the debris was removed by centrifugation and 1 vol. phenol (pH 4.1–4.3) was added. A standard phenol/chloroform procedure was used to re-extract the aqueous phase. Total RNA was precipitated for 2 h at 4 °C in 5 M LiCl, washed twice with ethanol and resuspended in 60 µl sterile deionized water. The RNA extracts were checked by electrophoresis on a 1.2% agarose/TAE gel, quantified by spectrophotometry and stored at −80 °C.

**RT-PCR and cDNA cloning.** Reverse transcription using primer R2191–5' (5'-aacctgagacaccac-3') was carried out as previously described (Turturo et al., 2008). The PCR was carried out using primers I1188+LNA3 (5'-atgctGaaCttgatatataga-3', locked nucleic acids shown in upper case) and R2183–5' (5'-ggagacaccagaggtgattgg-3') with an annealing temperature of 60 °C. Transcripts for the positive control were prepared as described previously (Turturo et al., 2008), while negative controls consisted of: (i) water, (ii) total RNA extracted from transgenic plants and (iii) a mixture of total RNA extracted from transgenic plants and total RNA extracted from a I17F-CMV-infected non-transgenic tobacco. The RT-PCR products were analyzed by electrophoresis on a 1% agarose/TAE gel, ligated with T4 DNA ligase by a standard ligation procedure described by the manufacturers (Fermentas).

**Fig. 1.** The five 6 nt deletions engineered into the 3'-NCR of I17F-CMV RNA3. The similarity plot corresponds to the alignment of the 3'-NCR of all the viral RNAs present in the infected transgenic plants analysed: I17F-CMV RNA1, 2 and 3 and R-CMV RNA3. The grey bars indicate the position of deleted nucleotides on I17F-CMV RNA3. The horizontal axis numbering refers to the nt position on I17F-CMV RNA3. The level of similarity (from 0 to 1) is shown on the vertical axis. The black lines below the horizontal axis represent the secondary structures present in the 3'-NCR. Stem–loops SLM, SLL and SLK have been identified on R-CMV but were not confirmed on I17F-CMV; SLJ, SLG and SLF have been shown experimentally in both the viruses; while SLL and SLH are only predicted (Ahlquist et al., 1981; Felden et al., 1994; Joshi et al., 1983; Rietveld et al., 1983; Thompson et al., 2008).
were separated in a 1% agarose/TAE gel, purified using the QIAEX II Gel extraction kit (Qiagen) and cloned in the pGEM-T easy vector according to the manufacturer’s instructions (Promega). Sequencing of cloned PCR products was done by BMR Genomics and analysed by using the Vector NTI Advance 9 software package (Invitrogen). All samples that gave a recombinant PCR fragment(s) were re-tested in an independent experiment and the resultant sequence(s) was determined for at least five clones per fragment.

RESULTS

Creation of 6 nt deletions in RNA3

Partially disabled I17F-CMV RNA3 molecules that could be restored to virulence by recombination with the mRNA encoded by the R-CMV CP + 3’-NCR transgene were created by engineering 6 nt deletions into the 3’-NCR of I17F-CMV RNA3 (Fig. 1). The primary target for mutation was the 100 nt region located immediately upstream of the 3’-terminal tRNA-like structure (TLS), which is essential for viral replication (Sivakumaran et al., 2000). Only one deletion (Del2095-2100) was engineered in the TLS, at a site distant from the core promoter for minus-strand synthesis. This part of the 3’-NCR was selected in order to have the greatest possible length of sequence in which homologous messenger/viral RNA recombination could restore the mutation, and to include previously reported hotspots (de Wispelaere et al., 2005). Most of the deletions removed sequences shown experimentally to be involved in the formation of stem–loop structures (Thompson et al., 2008).

Kinetics of infection of deletion mutants

Transcripts of each mutant or wild-type RNA3 of I17F-CMV in combination with transcripts of wild-type RNAs 1 and 2 of I17F-CMV were inoculated onto 40 plants (20 transgenic and 20 non-transgenic). The date of the first appearance of systemic symptoms was recorded. The lowest infectivity was observed for mutant Del2095-2100, which infected only 10% of the inoculated plants (transgenic or non-transgenic) (Fig. 2a), while an intermediate level of infectivity was observed in the plants inoculated with mutants Del2015-2020 and Del2047-2052, which showed markedly delayed symptoms and poor infection rates (Fig. 2b, c). In contrast, non-transgenic plants inoculated with mutants Del2000-2005 and Del2080-2085 showed comparable symptom development with wild-type I17F-CMV (Fig. 2d, e), except that in the former, symptoms were delayed for 2 days. With Del2015-2020 and Del2047-2052, approximately twice as many plants developed symptoms at 10–21 days p.i. in the non-transgenic plants as the transgenic ones, suggesting a partial resistance to infection by these attenuated viruses in the transgenic plants that was not observed with wild-type virus. According to their infectivity on non-transgenic plants, the mutants therefore could be classified into three groups: (i) wild-type-like (Del2000-2005 and Del2080-2085), (ii) moderately disabled (Del2015-2020 and Del2047-2052) and (iii) severely disabled (Del2095-2100). Since only the moderately disabled ones were useful for the
purposes of this study, all further studies were carried out with Del2047-2052, since this lengthened the sequence in which restoration by homologous recombination could be detected.

**Analysis of recombinant molecules of mutant Del2047-2052 infecting transgenic tobacco**

In a second experiment, screening for recombinant RNA3 was carried out in 44 transgenic tobacco plants inoculated with transcripts of I17F-CMV RNA1, RNA2 and mutated RNA3 Del2047-2052. Samples from symptomatic plants were taken at 8, 10 and 15 days p.i. (14, 19 and 28 plants, respectively) and analysed for the detection of recombinant molecules by RT-PCR. Typical results are shown in Fig. 3. Water, total RNA extracted from a transgenic plant and a mixture of total RNA extracted from transgenic plant and total RNA extracted from a non-transgenic plant infected by wild-type I17F-CMV were used as negative controls in order to monitor possible sample contamination and *in vitro* recombination (Fig. 3, negative controls). A chimeric construct recreating a precise homologous recombinant in the CP was used as a positive control (Fig. 3, lane +). The samples illustrated show the different patterns obtained: no recombinants detected (samples AF21, AH29), samples producing a unique band in the range of ~ 850–1200 bp, (samples AD16, AG26, AG27, AH31, AH32, AI33, AI34), samples producing more than one band, indicating recombinants of more than one size (sample AG28), and samples producing a band of the expected size but that was identified later as being derived from a parental molecule (a false-positive) illustrated by sample AE20.

All the fragments amplified that were between ~850 and 1200 bp were then excised from the gel, cloned and sequenced. Those outside this range were either too weakly amplified to clone or consisted of parental sequences (not shown). The total number of clones analysed for each plant was variable and related to different factors. At least five clones were analysed for each band purified after the RT-PCR step, with a maximum of two bands purified per sample (AD14 8 days p.i., AG27 8 and 10 days p.i., AG28 15 days p.i., AI33 15 days p.i. and AM44 8 days p.i., data not shown). In the majority of the cases, additional clones derived from bands purified after repeated RT-PCR were also analysed. This was done mainly for the samples that produced very faint RT-PCR bands (AD14 8 and 10 days p.i., AG27 8 days p.i., AI36 10 days p.i., AL37 10 days p.i., AM44 8 days p.i., data not shown) or to confirm previous observations (AG27 15 days p.i., AG28 15 days p.i., AM41 10 and 15 days p.i., AM44 10 days p.i. and AM43 10 days p.i.).

In total, recombinant fragments were detected in 18% of the samples at 8 days p.i., 25% at 10 days p.i. and 32% at 15 days p.i. [similar to previous observations at 10 days p.i. for I17F-CMV/R-CMV (26%) and Tomato aspermy virus (TAV) P-TAV/R-CMV (32%) (Turturo et al., 2008)]. In total, 22 different types of recombinant molecules were detected.

**Recombinants between Del2047-2052 RNA3 and transgene mRNA RNA3 sequences**

The majority of the types of recombinants detected were, as expected, between the I17F-CMV RNA3 and the R-CMV transgene mRNA (Fig. 4). Of them, ten were aberrant, while just two were of the precise homologous type. Both of the latter had already been detected under conditions of low selection pressure (Turturo et al., 2008): the predominant recombinant in the CP region (3/3l, Fig. 4) and a minor one in a block of sequence identity at R-CMV RNA3 position 1998–2040 (3/3i, Fig. 4). Of the aberrant recombinants, six types can be classified into two different families based on the crossover position on I17F-CMV RNA3. Recombinants 3/3b and 3/3c belong to family *x*

**Fig. 3.** Agarose gel separation of RT-PCR products generated from RNA extracted from R-CMV transgenic plants infected with I17F-CMV mutants using primers specific for the detection of recombinant molecules. Samples from transgenic plants (reference number above each lane) infected with I17F-CMV Del2047-2052 at 15 days p.i. The negative controls were water (lane 1), total RNA from uninfected transgenic leaf tissue (2) and total RNA from uninfected transgenic leaf tissue combined with an equal concentration of total RNA extracted from non-transgenic plants infected with I17F-CMV (3). The positive control (+) was a mixture of total RNA extracted from non-transgenic plants and 20 fg of transcript derived from an engineered I17F/R-CMV precise recombinant construct. Arrows show the position of a 1 kb fragment in the marker lane (M).
<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Plant</th>
<th>Days p.i.: 8</th>
<th>10</th>
<th>15</th>
<th>No. of plants</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/3a</td>
<td>1,2087</td>
<td>A133</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4/16</td>
</tr>
<tr>
<td></td>
<td>R,1884</td>
<td>A136</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A137</td>
<td>/</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AM44</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>3/3b</td>
<td>1,2064</td>
<td>A133</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>1/16</td>
</tr>
<tr>
<td></td>
<td>R,1091</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 3/3c | 1,2067 | A134  | /           | +  | −  | 3/16          |
|      | R,1091-2 | A136  | −           | +  | −  |               |
|      |         | A137  | /           | +  | −  |               |

| 3/3d | 1,2061 | AG28  | /           | −  | +  | 1/16          |
|      | R,1862 |       |             |    |    |               |

| 3/3e | 1,2059 | AH32  | /           | /  | +  | 1/16          |
|      | R,1091 |       |             |    |    |               |

| 3/3f | 1,2058 | AG26  | −           | −  | +  | 7/16          |
|      | R,1091 | AG27  | +           | −  | −  |               |
|      |         | AH31  | +           | +  | −  |               |
|      |         | AH32  | /           | /  | +  |               |
|      |         | A136  | +           | +  | −  |               |
|      |         | A137  | /           | +  | −  |               |
|      |         | AM44  | +           | −  | −  |               |

| 3/3g | 1,2053 | AM41  | +           | −  | −  | 2/16          |
|      | R,1090 | AM44  | +           | −  | −  |               |

| 3/3h | 1,2058 | AM41  | −           | +  | −  | 3/16          |
|      | R,1091 | AM43  | /           | +  | −  |               |
|      |         | AM44  | −           | +  | −  |               |

| 3/3i | 1,2012-5a | A133  | −           | +  | −  | 1/16          |
|       | R,1998- |       |             |    |    |               |
|       | 2046    |       |             |    |    |               |

| 3/3j | 1,2014 | AH6   | −           | +  | −  | 4/16          |
|      | R,1091 | AD14  | +           | +  | −  |               |
|      |         | AG26  | −           | +  | −  |               |
|      |         | AG27  | −           | −  | +  |               |

| 3/3k | 1,1993 | AL37  | /           | +  | −  | 1/16          |
|      | R,1899 |       |             |    |    |               |

| 3/3l | 1,1552-82 | AG27  | +           | −  | −  | 2/16          |
|       | R,1523-54 | AM44  | +           | −  | −  |               |

**Fig. 4.** Type and distribution of recombinants between I17F-CMV RNA3 and R-CMV transgene RNA. Group designations (α or β) are indicated for each recombinant. The names indicate the nucleotide position of the crossover site on I17F-CMV RNA3 and on R-CMV RNA3. Recombinants that were (+) or were not (−) detected at 8, 10 and 15 days p.i. are indicated. Asymptomatic plants at the time of analysis (†) were not tested. The schematic diagrams show the parts of the I17F-/R- CMV CP sequence; 3′-NCR are schematically represented as aligned molecules, with the upper part corresponding to I17F-CMV RNA3 and the lower part to R-CMV RNA3. The recombinant molecule is indicated by a dotted line, where the parallel stretches show the parental affinities. The diagonal dotted line serves only to schematically link the two parental strands and contains no sequence. The nucleotide positions of crossover sites on the two RNAs are shown. The deletion engineered at nt 2047–2052 on I17F-CMV RNA3 is indicated by a transverse vertical line through the line corresponding to the I17F 3′-NCR. Point mutations on the I17F-CMV RNA3 sequence of each recombinant are indicated: substitutions, C to U at nt 2064 (●), U to C at nt 1894 (†), U to C at nt 2022 (†), U to G at nt 2054 (○), G to A at nt 2055 (◇), GG to AA at nt 2055–2056 (◇); and deletions, A at nt 2043 (†), GG at nt 2055–2056 (■).
(I17F-CMV RNA3 nt 2067–2068), and recombinants 3/3e, 3/3f, 3/3g and 3/3h to family β (I17F-CMV RNA3 nt 2058–2059). The recombinants of this latter family were present in 56% of the plants in which recombinants were detected and were the most common recombinant types observed (Table 1). Regarding the other I17F-CMV RNA3/transgene recombinants, the members of family α and recombinants 3/3a and 3/3j were present in 25% of the positive samples, while recombinant 3/3l and recombinants 3/3d, 3/3i and 3/3k were present in just two and one plant, respectively (Fig. 4).

Several stable point mutations on the I17F-CMV RNA3 part were always associated with certain recombinants in all the plants where they were observed. The most prominent was a U to C substitution at position 2022 that was always present on RNA3 recombinants 3/3b, 3/3d and 3/3f, followed by a series of different mutations at nt 2054–2056 on RNA3 recombinants 3/3c, 3/3e, 3/3g and 3/3h (Fig. 4).

**Recombinants between Del2047-2052 RNA3 and I17F-CMV RNA1**

In addition to the 12 I17F-CMV RNA3/R-CMV transgene recombinants, 10 types of I17F-CMV RNA3/RNA1 recombinant were found, which were all of the aberrant type (Fig. 5). Eight of these have been classified into three families based on the crossover position on I17F-CMV RNA3: family γ (I17F-CMV RNA3 nt 2071–2072) includes recombinants 3/1a and 3/1b; family δ (I17F-CMV RNA3 nt 2060–2061) includes recombinants 3/1d and 3/1e; and family ε (I17F-CMV RNA3 nt 2058) includes recombinants 3/1f, 3/1g, 3/1h and 3/1i. Each RNA3/RNA1 recombinant type was found in just one plant. The members of family ε were the most common RNA3/RNA1 recombinant type, being present in 25% of the plants in which recombinants were detected. As with the RNA3/RNA3 recombinants, a U to C substitution at position 2022 on the I17F-CMV RNA3 part was observed frequently (recombinants 3/1b, 3/1c, 3/1f, 3/1g and 3/1h), as well as mutations around nt 2054–2056 (recombinants 3/1d and 3/1i).

**Within-plant changes in recombinant populations**

The changes in recombinant populations in individual plants during the sampling period is illustrated in Table 1. As is visible for plants AB6, AD14, AG27, AH31 and AI33, certain recombinant types (3/3a, 3/3j, 3/3g, 3/3f and 3/3h) were detected throughout the testing period, while others were found just occasionally and at a certain times p.i. (i.e. recombinants 3/1b and 3/1h on plant AA4 or the recombinant 3/3i on plant AI33). Moreover, some recombinants were only detected in earlier sampling (recombinants 3/3g, 3/3i and 3/1j, Figs 4 and 5), while two appeared only around 10–15 days p.i. (recombinants 3/3c and 3/3h) (Fig. 4).

Of the recombinants detected more than once (3/3a, 3/3c, 3/3f, 3/3g, 3/3h, 3/3j, and 3/3l), all were RNA3/RNA3 recombinants. The remaining recombinants that were only found once were the RNA3/RNA3 recombinants.

**Table 1. Recombinants (see Figs 4 and 5) detected by RT-PCR at 8, 10 and 15 days p.i. for each plant analysed**

The RNA3/RNA3 (shaded) and RNA3/RNA1 (not shaded) recombinants are indicated. Note, when the sum of the relative number of clones is different from the total number of clones analysed for the plant, the non-counted clones contained just parental derived sequences (i.e. plant AG28, 15 days p.i.). –, No recombinants were detected; /, plant was asymptomatic and therefore not analysed.

*n*, Total number of clones analysed for each plant.
†The number of clones containing that particular recombinant out of the total number of clones analysed at that particular time p.i. is given in parentheses. The family that a recombinant belongs to (where relevant) is given (α–ε).
3/3b, 3/3d, 3/3e, 3/3i, 3/3k and all of the RNA3/RNA1 recombinants (Figs 4 and 5).

Overall, the majority of the recombinant molecules could be grouped in just five families: two for the RNA3/RNA3 recombinants (families α and β, Fig. 4) and three for the RNA3/RNA1 recombinants (families γ, δ and ε, Fig. 5), which demonstrates a general tendency to produce a specific limited set of recombinants.

**DISCUSSION**

**Effects of 6 nt deletions in 117F-CMV RNA3**

The approach used to increase selection pressure in favour of recombinants was to disable the virus in a region 3' of the previously described crossover sites. In this way, it was expected that recombinants that restored full viral activity would be favourably selected. Not surprisingly, the mutant containing the deletion at the beginning of the TLS (mutant Del2095-2100, Fig. 1) was the most debilitated, infecting just 10% of the plants tested (Fig. 2e). As has been previously reported (Boccard & Baulcombe, 1993), a deletion in the TLS can almost abolish minus-strand RNA synthesis by the CMV replicase. At the other extreme, mutants Del2000-2005 and Del2080-2085, affecting stem–loops SLJ and SLF, respectively (Fig. 1), produced kinetics of infection on wild-type plants similar to those of the wild-type virus (Fig. 2a, d). Conversely, mutants Del2015-2020 and Del2047-2052, located in a very highly conserved region shared by all CMV RNAs of all CMV isolates, showed slightly slower kinetics of infection than the wild-type virus (Fig. 2b, c), thus plausibly indicating a partial disabling of the virus or a delayed restoration of the viral activity due to a recombination event.

**Recombinants between 117F-CMV RNA3 and mRNA containing R-CMV RNA3 sequences**

Under conditions of low selection pressure (Turturo *et al.*, 2008), the homologous recombinant with its crossover site in the CP gene at nt 1523–1554, corresponding to recombinant type 3/3l, was clearly predominant both in...
non-transgenic plants doubly infected with R-CMV and I17F-CMV and in transgenic plants containing the CP and the entire 3'-NCR of R-CMV infected with I17F-CMV. In the present study, where there was an increase in selection pressure in favour of recombinants, the same recombinant was also detected, but did not predominate, being present in only two of 16 plants in which recombinants were detected. Recombinant 3/3l has been shown to be non-viable alone in planta (Pierrugues et al., 2007), and our results support this finding, since it was only detected in the earliest samples and always in the presence of other recombinants (Table 1, plant AG27 8 days p.i. and plant AM44 8 days p.i.). Thus, counterselecting recombinant 3/3l indeed made it possible to detect additional recombinant types.

The populations of aberrant recombinants between I17F-CMV RNA3 and the R-CMV transgene mRNA were strongly dominated by the hotspot located at R-CMV nt 1899–1902. This hotspot is exactly at the 5′-end of the conserved motif Box-1 (Blanchard et al., 1996; Thompson et al., 2008), which is common to all subgroup II CMV strains and to TAV but is absent in subgroup I CMV strains (Palukaitis et al., 1992; Shi et al., 1997), and plays a key role in the production of subgenomic RNA5 (de Wispelaere & Rao, 2009; Thompson et al., 2008). Moreover, it has been reported as a hotspot for viral recombination between CMV RNA3 and TAV RNAs 1 and 2 (Suzuki et al., 2003) or TAV RNA3 (de Wispelaere et al., 2005), between TAV RNA3 and CMV RNAs 1 and 2 (Shi et al., 2007), and in Q-CMV (subgroup II) intragenomic recombinants (de Wispelaere & Rao, 2009). Finally, this was also the crossover site observed in the natural recombinant Alstroemeria-CMV (Chen et al., 2002), and in certain TAV strains (Moreno et al., 1997; Raj et al., 2007, 2009).

Concerning the I17F-CMV RNA3 crossover sites on the recombinants detected in the present study, it is interesting to note that they are very similar to some of those previously observed for aberrant recombinants between Y-CMV (subgroup I) RNA3 and TAV RNA1 (Suzuki et al., 2003), between Y-CMV RNA2 and V-TAV RNA2 (Masuta et al., 1998) and between R-CMV RNA3 and P-TAV RNA3 (de Wispelaere et al., 2005). Similarly, recombination in the same region generated the duplications identified by Moreno et al. (1997) and Raj et al. (2007, 2009) in some TAV isolates. In all of these cases, the crossover sites fell in a region of highly conserved secondary structures (Ahlquist et al., 1981), in particular on stem–loops SLG and SLF (Fig. 6).
Recombinants between I17F-CMV RNA3 and I17F-CMV RNA1

Intragenomic recombination (i.e. between genome segments) is known to occur in CMV-infected plants (Canto et al., 2001). Nevertheless, the RT-PCR primers used in the present work were not designed to detect intragenomic recombinants, but just those derived from I17F-CMV RNA3 and the transgene mRNA. Surprisingly, of the 22 types of recombinant molecule observed, ten (all of them aberrant) resulted from recombination between I17F-CMV RNAs 1 and 3 (Fig. 5). This unexpected result can be explained by the fact that the six 3′-terminal nucleotides of primer R2191−, used for the reverse-transcriptase step, can also anneal with nt 3150–3155 of I17F-CMV RNA1, thereby permitting the amplification of such intragenomic recombinants. There is no equivalent site on RNA2, which explains why no RNA3/RNA2 recombinants were detected.

As mentioned above, the vast majority of the crossover sites on I17F-CMV RNA3 in RNA3/RNA3 recombinants were located in a structurally rich region (from nt 1990–2084 of I17F-CMV RNA3) including stem–loop SLG. This was also a hotspot for RNA3/RNA1 recombinants (Fig. 6). In contrast, unlike the RNA3/RNA3 recombinants, no precise RNA1 hotspot was observed in the RNA3/RNA1 recombinants, which may be related to the absence of Box-1 or RNA5 in the RNAs of I17F-CMV. Of the eight different crossover sites on RNA1, one was located within the 1aa protein coding sequence (RNA1 nt 3027, recombinant 3/1e, Fig. 5), two in the region corresponding (in subgroup II CMV) to Box-1 (RNA1 nt 3079 in case of recombinants 3/1c, 3/1f and nt 3083/3084 for recombinant 3/1a, Fig. 5) and five (recombinants 3/1b, 3/1d, 3/1g, 3/1h, and 3/1j, Fig. 5) were dispersed across a region containing several zones of high sequence identity between I17F-CMV RNA3 and RNA1 (nt 3091–3143 of I17F-CMV RNA1).

Mechanism of recombination that produces aberrant cucumoviral RNA3

There have been two successive models proposed to explain the structure of aberrant RNA3 recombinants. First, from their observation of aberrant recombinants similar to the majority of the aberrant recombinants described here, Suzuki et al. (2003) hypothesized that the minus-strand equivalents of stem–loops SLG and SLF (Fig. 6) could have an important role in causing the replicase to pause and dissociate from the template during plus-strand synthesis and then subsequently bind on an acceptor site, which they proposed to be the Box-1 motif on minus-strand RNA3. This is compatible with the recombinant found in the pseudorecombinant viruses Y-CMV/v-TAV (Masuta et al., 1998) and P-TAV/TrK7-CMV (subgroup II) (Fernandez-Cuartero et al., 1994), with the results of de Wispelaere et al. (2005) and with the crossovers in certain Spanish and Indian TAV isolates (Moreno et al., 1997; Raj et al., 2007, 2009). In contrast, de Wispelaere & Rao (2009) proposed that the aberrant RNA3 recombinants were produced during minus-strand synthesis. They suggested that after synthesis of a complete copy of minus-strand RNA5, the replicase would switch to the minus-strand promoter located in the TLS at the 3′ end of the CMV RNAs. This is consistent with the recombinant Alstroemeria-CMV isolates (Chen et al., 2002), but is not compatible with the recombinants observed here or in other publications (de Wispelaere et al., 2005; Fernandez-Cuartero et al., 1994; Masuta et al., 1998; Suzuki et al., 2003) since none contained a duplication of the TLS, as was reported by de Wispelaere & Rao (2009). Instead, all crossover sites on I17F-CMV (RNA3 or RNA1) were located 5′ of the TLS, primarily at the hotspot at or near SLG. In addition, the only possible source of RNA5 in the experimental system studied here was the R-CMV transgene mRNA, but no RNA5 was observed by RNA blotting (J. R. Thompson, unpublished data), suggesting that RNA5 was either not produced or produced at very low levels. Clearly, it would be of interest in the future to carry out experiments that are specifically designed to determine basic features of recombination in CMV, including on which strand (plus or minus) recombination occurs.

Conclusions

Unlike previous observations under conditions of low selection pressure in favour of recombinants (Turturo et al., 2008), in the present study, we were able to eliminate the predominance of the homologous recombinant 3/3l, and therefore identify a larger number of types of recombinant molecules. The deletion in the region containing the highly conserved stem–loops SLG and SLF, probably facilitated the detection of aberrant recombinants more frequently, but did not generate the recombinants as such, since similar types of recombinants have been detected previously on non-mutated CMV and TAV (de Wispelaere et al., 2005; Suzuki et al., 2003). Finally, the identification of the same recombinants at different times following inoculation (recombinants 3/3a, 3/3c, 3/3f, 3/3h, 3/3j, 3/1a, 3/1f, 3/1g) in the same plant implies the firm establishment of these molecules within the plant. Indeed, RNA blotting showed that in plant AM41, the recombinant was the predominant virus present, whereas in others, the parental mutant was more abundant (data not shown).

As hypothesized initially, when the recombinant that was predominant under conditions of minimal selection pressure (Turturo et al., 2008) was counter-selected, numerous other recombinant types were observed. Since the recombinants described here all correspond to types that were observed in non-transgenic plants under other experimental conditions, this suggests that they were also present under conditions of low selection pressure. This also supports the idea that the additional types of recombinants observed here do not present an additional risk of emergence of novel recombinant strains. We believe that it should be possible to obtain a coherent picture of the recombinant populations by high-throughput sequen-
ACKNOWLEDGEMENTS

We thank the members of the Plant Virology Group, and particularly Arianna Friscina, for their invaluable help in the laboratory, and Laura Chiappetta for growing the plants. This work was supported by grants from the Fondazione Cassamarca and the French ANR (contract ANR-05-POGM-001).

REFERENCES


