Deletions and recombinations with the RNA1 3’ ends of different tobraviruses have created a multitude of tobacco rattle virus TCM-related RNA2 species in *Alstroemeria* and tulip

R. Koenig,1 D.-E. Lesemann,1 E. Pfeilstetter,1 S. Winter1 and C. W. A. Pleij2

1Julius Kühn-Institute, Federal Research Center for Cultivated Plants, Institute for Epidemiology and Pathogen Diagnostics, Institute for National and International Plant Health and DSMZ Plant Virus Department, Messeweg 11, D38104 Braunschweig, Germany
2Leiden Institute of Chemistry, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands

In vegetatively propagated *Alstroemeria* plants that showed pronounced stunting and necrotic leaf spots, a tobravirus infection was diagnosed in which one tobacco rattle virus (TRV, strain AL) RNA1 species was associated with seven different RNA2 species. The latter differed considerably in size and in the types of their 3’ RNA1-related sequences. The 5’ RNA2-specific part of all these RNA2 molecules showed almost 100 % sequence identity with that of RNA2 of the TRV isolate TCM from tulip, but in some of these RNA2 molecules it was shorter than in the TCM isolate, whereas in others it was longer. One of the TRV AL RNA2 molecules, i.e. TC39PE-a, contained the full set of three full-length RNA2-specific ORFs (ORF2a, -2b and -2c), whereas the previously analysed TCM sequence contained only ORF2a and -2b. In four of these TRV AL RNA2 molecules, i.e. those that had a relatively short RNA2-specific part, the 3’ end was identical to that of the cognate TRV AL RNA1, but in the other three, which had a long RNA2-specific part, it was closely related to that of pea early browning virus (PEBV) RNA1, which was not detected in the infected plants. A comparison with previously described TRV/PEBV RNA2 recombinants suggested that the various TRV AL RNA2 molecules may represent various steps and side steps in an evolutionary process, which is apt to open the wide host range of TRV also to PEBV-derived RNA2 species.

INTRODUCTION

Tobravirus genomes have been described to consist of two RNA species. RNA1 contains two large, 5’ overlapping genes for replication-associated proteins of ~135–140 and ~195–200 kDa, and further downstream two smaller ORFs (1a and 1b), which encode the ~30 kDa movement protein and an ~13–16 kDa silencing suppressor, respectively (MacFarlane, 1999; Robinson, 2004b). The translation product of ORF1b has also been found to be necessary for seed transmission of pea early browning virus (PEBV) (Wang et al., 1997). RNA2 consists of a 5’ RNA2-specific part and a 3’ terminal part that shows a high percentage sequence identity to a tobravirus RNA1. This 3’ part will be referred to as the ‘RNA1-related part’ in the following. A similar terminology has been used by Angenent et al. (1989) and Hernandez et al. (1996). Both parts of the RNA2 molecules in various tobravirus isolates show considerable differences in size and composition. The RNA2-specific part contains the coat protein (CP) gene (ORF2a) and in several virus isolates further downstream a second ORF (ORF2b) and even a third one (ORF2c), which is usually followed downstream by an RNA2-specific non-coding region. The gene product of ORF2b is necessary for nematode transmission (MacFarlane et al., 1996; Hernández et al., 1997), whilst the gene product of ORF2c greatly increases the efficiency of PEBV transmission by *Trichodorus primitivus* (MacFarlane et al., 1999), but not that of tobacco rattle virus (TRV) isolate PpK20 by *Paratrichodorus pachydermus* (Hernández et al., 1997). Some isolates contain an additional small ORF for a protein of approximately 9 kDa between ORF2a and -2b.

During a survey on the occurrence of tobraviruses in ornamental plants, we identified a strain of TRV (strain AL) in an *Alstroemeria aurea* plant, which was obtained from a botanical garden and showed pronounced stunting, shoot distortions and necrotic leaf spots (Fig. 1). *Alstroemeria*
species (family Alstroemeriaceae, order Liliales) originate from South America. They have become very popular worldwide as cut flowers as well as garden plants. Because they are propagated mainly vegetatively by their rhizomes, the many viruses that have been detected in them (see Fuji et al., 2007 and references therein) may have persisted in individual plants for a long time. We found that TRV AL contains one genetically distinct RNA1 species and, in addition, seven RNA2 species that differed considerably in their sizes and in the types of their 3’ RNA1-related sequences.

RESULTS

TRV AL RNA1

The complete nucleotide sequence of TRV AL RNA1 excluding small 5’ and 3’ portions derived from primers 179 and 158 (MacFarlane, 1996), respectively, was determined and was found to be clearly distinct from the RNA1 molecules of other TRV isolates, i.e. the potato isolates PpK20 (Ratcliff et al., 2001) and ORY (Sudarshana & Berger, 1998), the spinach isolate SYM (Hamilton et al., 1987) and the somewhat more distantly related potato isolates PpO85M (Robinson, 2004a) and MI (Crosslin et al., 2010). TRV AL RNA1 shared a mean of ~92–93% sequence identity with these RNAs, but was only rather distantly related to the RNA1 molecules of the two tobraviruses, i.e. PEBV (MacFarlane et al., 1989) and Pepper ringspot virus (PepRSV; GenBank accession no. L23972) (Fig. 2a).

The RNA2-specific parts of the various TRV AL RNA2 molecules

Altogether, seven different variants of TRV AL RNA2, named TC3’PE-a, TC3’PE-b, TC3’PE-c, TC3’AL-a, TC3’AL-b, TC3’AL-c and TC3’AL-d, were identified (Fig. 3). The first two letters in these names were chosen to indicate that the sequences of their RNA2-specific parts were found to share between 99.7 and 100% sequence identity with the corresponding RNA2 regions including the 5’ untranslated region (UTR), 2a (CP) gene and 2b gene (in some molecules) of the TRV isolate TCM from tulip (Angenent et al., 1986). There were considerable differences, however, in the sizes of the RNA2-specific and also of the RNA1-related parts of these TRV AL RNA2 molecules. The RNA2-specific parts of TC3’AL-b, -c and -d (Fig. 3e–g) contained only ORF2a, i.e. the CP gene, downstream of the 5’ UTR sequence, whereas those of TC3’AL-a RNA2 (Fig. 3d) and TRV TCM RNA2 contained, in addition, ORF2b. The RNA2-specific sequence of TC3’AL-a RNA2 was slightly shorter than that of TRV TCM RNA2, which contained a small portion of a partially deleted ORF2c that had fused with the 3’ part of ORF1a on its RNA1-derived 3’ end (Uhde et al., 1998). The RNA2-specific sequences of TC3’PE-a, -b and -c (Fig. 3a–c) were longer than that of TRV TCM RNA2 and the former two carried a third ORF (ORF2c). ORF2c on TC3’PE-a (879 nt) was probably complete, because it had a similar size to ORF2c of TRV Sp (861 nt; Schmidt & Koenig, 1999) and TRV PaY4 (858 nt; Vassilakos et al., 2001) with which it shared the highest percentages of sequence identity (Fig. 2f). TC3’PE-b and -c (Fig. 3b and c) appeared to be deletion mutants of TC3’PE-a. TC3’PE-b contained a 306 nt deletion in the 5’-proximal part of ORF2c, causing the deletion of aa 38–139 in its translation product. TC3’PE-c contained an even larger deletion of 648 nt, which started in the intergenic region between ORF2b and -2c and ended in the 3’ part of ORF2c, which, due to this deletion, lacked an AUG start codon. Excluding the deletions, TC3’PE-a, -b and -c shared more than 99.7% sequence identity.

As mentioned above, all our TRV AL RNA2-specific sequences and the TRV TCM RNA2 showed almost 100% sequence identity in those parts that were present in the individual RNAs. One difference was noted in the CP gene sequence. In the published TRV TCM sequence (GenBank accession no. X03955), the CP gene has an additional A at position 589 (probably due to a sequencing error), which was not found in any of our 13 TRV AL RNA2 cDNA clones or PCR products covering this region. This additional base would lead to a frame shift and an altered CP C terminus with two additional amino acids, i.e. -RSWTWNVLV (TRV TCM) versus -GRGPGT (TRV AL).

The CP genes, i.e. ORF2a (Fig. 2d), and also ORF2b (Fig. 2e) of TRV AL and TRV TCM showed their highest percentages of sequence identities with the corresponding sequences of the Dutch PEBV isolate E116. Somewhat more distant relationships were found with the corresponding sequences of the rather closely interrelated TRV PaY4, TRV On and TRV Sp, and even more distant ones with those of TRV TpO1, PEBV SP5, TRV ORY and TRV PpK20. For the TRV TCM sequence, this has already been noted previously by MacFarlane (1999), Schmidt & Koenig

Fig. 1. A. aurea showing stunting and distortion of the upper parts of the shoots (left side) and irregular yellow and necrotic leaf spots (right side).
(1999) and Swanson & MacFarlane (1999). A similar pattern of relationships, especially with the more closely related viruses, was seen with ORF2c (Fig. 2f), which is present in only a few tobravirus RNA2 sequences, including that of the newly described TC3\textsuperscript{9}PE-a sequence, but not in the TRV TCM sequence (Fig. 3).

The RNA1-related parts of the TRV AL RNA2 molecules

TRV TCM RNA2 and the different variants of TRV AL RNA2 differed not only in the sizes of their RNA2-specific 5' parts, but also in the sizes of their RNA1-related 3' parts (Fig. 3). The shortest RNA1-related 3' end was found in TC3\textsuperscript{9}PE-a, -b and -c (Fig. 3a–c). In these RNA2 molecules, it consisted of only ~230 nt upstream of the 3' terminal region of 19 nt, which is recognized in tobravirus RNAs by primer 158 (MacFarlane, 1996). The longest RNA1-related sequence was found in TC3\textsuperscript{9}AL-b (Fig. 3e) where it consisted of 1227 nt upstream of the 19 nt 3'-terminal primer region and was larger than the RNA2-specific part. It contained the entire ORF1b and a large 3' portion of ORF1a. The entire ORF1b and various 3' portions of ORF1a were also found in the RNA1-related 3' portions of TC3\textsuperscript{9}AL-c and -d (Fig. 3f and g) and of TRV TCM RNA2.

In view of the almost complete sequence identity between the RNA2-specific parts of the TRV AL and TRV TCM RNA2 molecules, it was surprising to find that the 3' RNA1-related portions of these RNAs differed considerably, not only in their sizes but also in their nucleotide sequences. The RNA1-related parts of TC3\textsuperscript{9}AL-a, -b, -c and -d (Fig. 3d–g) all showed 100% sequence identity with the corresponding portions of TRV AL RNA1, but only about 94% with those of TRV TCM RNA1 and RNA2, which share 100% sequence identity with each other (Angenent et al., 1986) (shown for the 230 nt upstream of the 19 nt region of primer 158 in Fig. 2g). The most surprising observation, however, was that the 3' terminal RNA1-related parts of the 3'UTRs of TC3\textsuperscript{9}PE-a, -b and -c (Fig. 3a–c) showed little sequence identity with the 3' end of
TRV AL RNA1. They did, however, show a very close relationship to the 3' end of RNA1 of the British PEBV isolate SP5 (~99% sequence identity) and to the 3' ends of the RNA2 molecules of other isolates, i.e. the British PEBV isolate SP5, the Dutch PEBV isolate E116 and the recombinant tobravirus isolate I6 (Robinson, 1994) (Fig. 2g). The last two letters in the names of the various TRV AL RNA2 molecules were chosen to designate the tobravirus RNA1 to which the RNA1-related part in the respective RNA2 has its closest relationship, i.e. the 3' ends of the TC3'PE RNA2 molecules showed a very close relationship to the 3' end of RNA1 of the British PEBV SP5 isolate, whereas the 3' ends of the TC3'AL RNA2 molecules shared 100% sequence identity with the 3' end of TRV AL RNA1. If this type of nomenclature were to be applied in a more general way, the acronym for TRV TCM RNA2 molecules would be TC3'TC, because its RNA1-related 3' end shares 100% sequence identity with TRV TCM RNA1 (Angenent et al., 1986).

The presence of a 3' end closely related to that of PEBV RNA1 in some of the TRV AL RNA2 molecules raised the question of whether a PEBV-like RNA1 would also be present in infected Alstroemeria. We were unable, however,
to detect any further PEBV RNA1 sequences in these plants using several RNA1-specific primers. Thus, apparently only TRV AL RNA1, and not PEBV RNA1, was present in the infected A. aurea plants. It has been shown previously that the specific recognition of tobavivirus RNA2 molecules by the RNA1-encoded replicase proteins is determined by sequences at the 5’ end but not by those at the 3’ end of these RNAs (Robinson, 1994; Mueller et al., 1997). The replicase proteins encoded on TRV AL RNA1 can, therefore, apparently also amplify the TC3’PE RNA2 molecules present in A. aurea.

Analyses of primary and secondary structures around the recombination sites

Table 1 shows an overview of the sequences surrounding the recombination sites in the various TRV AL RNA2 deletion and deletion/recombination mutants. In some of the recombinants, there were short overlaps of the 3’ and 5’ sequence portions derived from the respective parental molecules (indicated in Table 1 in bold), whereas in others such overlaps were not observed.

<table>
<thead>
<tr>
<th>Recombinant RNA2 molecule</th>
<th>Recombination sites and surrounding sequences of the 20 5’ and 20 3’ nt*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental molecule 1</td>
<td>Parental molecule 2</td>
</tr>
<tr>
<td>TC3’PE-b</td>
<td>(nt 2304–2346)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 2304–2346)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 2610–2652)</td>
</tr>
<tr>
<td>TC3’PE-c</td>
<td>(nt 2183–2222)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 2183–2222)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 2820–2859)</td>
</tr>
<tr>
<td>TC3’AL-a</td>
<td>(nt 2081–2125)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 2081–2125)</td>
</tr>
<tr>
<td>TRV AL RNA1</td>
<td>(nt 6186–6230)</td>
</tr>
<tr>
<td>TC3’AL-b</td>
<td>(nt 1136–1180)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 1136–1180)</td>
</tr>
<tr>
<td>TRV AL RNA1</td>
<td>(nt 5488–5532)</td>
</tr>
<tr>
<td>TC3’AL-c</td>
<td>(nt 1158–1197)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 1158–1197)</td>
</tr>
<tr>
<td>TRV AL RNA1</td>
<td>(nt 5583–5622)</td>
</tr>
<tr>
<td>TC3’AL-d</td>
<td>(nt 1123–1165)</td>
</tr>
<tr>
<td>TC3’PE-a</td>
<td>(nt 1123–1165)</td>
</tr>
<tr>
<td>TRV AL RNA1</td>
<td>(nt 6033–6077)</td>
</tr>
</tbody>
</table>

*Sequence regions in the parental molecules that are also present in the recombinants are printed in upper case and are underlined. Sequence regions of the parental molecules immediately downstream or upstream of the recombination sites that are not present in the recombinants are shown in lower case. In some recombinants, there was an overlap of the parental sequences in the recombinant molecule; the overlapping nucleotides have been highlighted in bold.
In the AL3’PE-a, -b and -c and TCM RNA2 sequences, however, STAR consistently predicted a series of pseudoknots, which the program Mfold could not produce. The region containing these pseudoknots started close to the 3’ end of ORF2b and extended through almost the entire intergenic region between ORF2b and -2c. It contained a series of seven hairpins, of which six were able to form pseudoknots. This highly pseudoknotted region, the so-called ‘stalk region’, has been recognized previously in TCM RNA2 (Pleij et al. 1987; C. W. A. Pleij, unpublished observations). Pleij et al. (1987) suggested previously that this ‘stalk region’ might be stimulatory for recombination, for instance by presenting an obstacle for the RNA replicase. The recombination site in TRV TCM RNA2 was located 65 nt downstream of this stalk region. Interestingly, the recombination site spanning the deletion region in TC3’PE-c was found immediately downstream of this pseudoknot stalk. TC3’AL-a contained only a 5’ portion of this stalk region. Possibly, the three pseudoknots that were still present might be sufficient to produce a similar effect, as this effect might also be exerted by the above-mentioned hairpin structure in TC3’AL-b, -c and -d. 

It has been proposed by Ohshima et al. (2007) that recombination sites are characterized by an AU-rich region downstream and a GC-rich region upstream of the recombination site. Taking the same window of 50 nt on both sides of the recombination sites in all the deletion and deletion/recombination mutants depicted in Fig. 3, we indeed found AU values of >50% (up to 78%) downstream of the recombination sites. The same was not seen, however, for the GC content of the 50 nt stretch of the upstream part.

**DISCUSSION**

The finding of seven different TRV RNA2 molecules in *Alstroemeria* was a surprise because, so far, for most tobraviruses studied at the molecular level either only one RNA1 and one RNA2 species or a single RNA1 species only has been described (Robinson, 2004b). The isolates that had been studied at the molecular level were all propagated on an experimental host, usually a tobacco species. Additional RNA2 molecules had possibly not been looked for, because they were not expected. The ease of detection of a multitude of different TRV RNA2 molecules in *Alstroemeria* might be due to the fact that novel PCR techniques, which were not available at the time of earlier studies, allowed the amplification of cDNA sequences directly from naturally infected plants.

There has been one report (Hernandez et al., 1996) that the RNA2 composition of a tobravirus (TRV isolate PpK20) may change rapidly during serial passages on tobacco. The original PpK20 isolate contained only one RNA2 species with apparently intact ORF2a, -2b and -2c and a short RNA1-related 3’ end. However, in some experiments, as early as after one or a few passages on tobacco, several smaller RNA2 mutants were detected that carried large, different-sized deletions in their RNA2-specific part. In addition, in some of these mutants, the size of the RNA1-related 3’ part had increased and contained 3’ portions of the RNA1-derived 16K gene, which was not present in the original PpK20 RNA2 and must therefore have been acquired by recombination from the cognate RNA1. Deletions of genomic areas that are no longer necessary when a virus is transmitted manually to an experimental host or when it is propagated for prolonged periods of time without vector transmission in its natural host have also been observed with other normally vector-transmitted viruses (e.g. Reddy & Black, 1974; Shirako & Brakke, 1984; Koenig et al., 1986; Bouzoubaa et al., 1991; Reavy et al., 1998; Koenig, 2000; Sandgren et al., 2001). The frequent acquisition of the 3’ end of one RNA (RNA1) by another RNA (RNA2) of the same virus or of another virus in the same genus may, however, be a unique feature of tobraviruses.

Most, if not all, TRV AL RNA2 molecules as well as the TRV TCM RNA2, which all start at their 5’ ends with practically identical 5’UTRs and coding sequences, are apparently also the result of recombinations and/or deletions. TC3’PE-a (Fig. 3a) seems to be the most ancient form. TC3’PE-b and TC3’PE-c are apparently merely deletion mutants of it, whereas TC3’AL-a, -b, -c and -d and also TRV TCM RNA2 are, in addition, recombination mutants, which have acquired 3’ ends of different lengths from various TRV rather than PEBV RNA1 molecules (Fig. 3).

TC3’PE-a fulfils the requirements of being non-recombinant as defined by MacFarlane (1999), because it contains the full set of three intact RNA2-specific ORFs (CP gene, 2b and 2c) and its RNA1-related 3’ end lacks RNA1-specific coding sequences. However, its PEBV RNA1-related rather than TRV RNA1-related 3’ end and the composition of its RNA2-specific 5’ part (Table 2) suggest that even this RNA2 may be the product of complex recombination events. Angelent et al. (1986) and Goulden et al. (1991) have already pointed out that the TRV TCM RNA2 sequence (and the same would apply to its extended versions, TC3’PE-a, -b and -c) contains a complex mixture of TRV- and PEBV-related sequence elements.

Mixtures of various TRV- and PEBV-derived sequence elements that suggest complex evolutionary histories have also been detected previously in some other tobravirus RNA2 molecules (summarized in Table 2). The prototype of a non-recombinant full-length PEBV RNA2 is represented by the RNA2 in the almost identical British isolates SP5 and TpA56 (PEBV-B) (Goulden et al., 1990). The RNA2 of the Dutch PEBV isolate E116 (PEBV-D), however, has been recognized to be recombinant (Swanson & MacFarlane, 1999). It has PEBV-B RNA-related sequences at its 5’ and 3’ ends (Table 2, Fig. 2b, c, g). However, its RNA2-specific coding sequences (ORF2a and -2b) are TRV-like; they share the highest percentages of sequence identities with those of TRV TCM (Swanson & MacFarlane, 1999).
TRV PaY4 RNA2 has been included for comparison. Data from Angenent et al. (1986), Goulden et al. (1991), MacFarlane (1999), Robinson et al. (1987), Robinson (1994), Schmidt & Koenig (1999), Swanson & MacFarlane (1999), Uhde et al. (1998) and the present report.

Table 2. Tobraviruses with RNA2 molecules carrying mixtures of PEBV-B- (British isolate SP5) and TRV-derived sequence elements arranged in order of increasing content of TRV-derived elements

<table>
<thead>
<tr>
<th>Tobravirus RNA2</th>
<th>Supporting RNA1</th>
<th>Closest relationships/other properties of various RNA2 sequence elements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5'UTR upstream of common region (see Fig. 2b)</td>
<td>5'UTR downstream of common region (see Fig. 2c)</td>
</tr>
<tr>
<td>PEBV-D (Dutch isolate E116)</td>
<td>PEBV-B-like†</td>
<td>PEBV-B</td>
</tr>
<tr>
<td>I6</td>
<td>TRV</td>
<td>TRV</td>
</tr>
<tr>
<td>TC3³PE-a</td>
<td>TRV (AL)</td>
<td>TRV</td>
</tr>
<tr>
<td>TC3³PE-b and -c</td>
<td>TRV (AL)</td>
<td>TRV</td>
</tr>
<tr>
<td>TRV TCM</td>
<td>TRV (TCM)</td>
<td>TRV</td>
</tr>
<tr>
<td>TC3AL-a, -b, -c and -d</td>
<td>TRV (AL)</td>
<td>TRV</td>
</tr>
<tr>
<td>TRV PaY4</td>
<td>No information</td>
<td>TRV</td>
</tr>
</tbody>
</table>

*In Fig. 3 only shown for TRV-AL and TRV-TCM.
†Evidence from hybridization experiments (D. Robinson, personal communication).
‡TRV TCM-related viruses are TRV PaY4, TRV Sp, TRV On and PEBV D.
§Evidence from hybridization experiments and serology (Robinson et al., 1987; D. Robinson, personal communication) and the 5' end of the CP gene in the published partial sequence of I6 RNA2 (ten codons) (GenBank accession no. S72876).

1999) and somewhat lower ones with those of TRV ON (Uhde et al., 1998), TRV PaY4 (Swanson & MacFarlane, 1999) and TRV SP (Schmidt & Koenig, 1999) (Fig. 2d, e). Hybridization experiments suggest that the RNA1 of PEBV-D is PEBV-like (D. Robinson, personal communication).

Another recombinant tobavirus RNA2, which unfortunately has only been partially sequenced, is that of the tobacco isolate I6 (Robinson et al., 1987; Robinson, 1994) (Table 2). Its 3' end and the portion of its 5'UTR downstream of the 'common region, identified by Goulden et al. (1991) are PEBV-B RNA2-like (Fig. 2c, g) and probably also so are its RNA2-specific coding sequences, as suggested by the results of hybridization experiments and serological tests and the ten codons of the CP gene that are recognized in the published sequence (GenBank accession no. S72876S1) (Robinson, 1994). The region of the 5'UTR upstream of the common region, however, is TRV-like (Fig. 2b), and replication of the I6 RNA2 depends on TRV RNA1 (Robinson, 1994).

A TRV I6-like RNA2, i.e. a molecule that consists mainly of PEBV-B RNA2-like sequences except for its 5'UTR upstream of the common region, could have been imagined to be a progenitor of a TC3³PE-a-like RNA2. The dependency of its replication on TRV RNA1-encoded enzymes would have enabled a close contact with TRV-like RNA2 molecules. These could have acted in further recombination events as donors for the TRV RNA2-like coding sequences present in TC3³PE-a.

The possible evolutionary origin of PEBV-D RNA2 might be explained by different hypotheses. In mixed infections with TRV, a PEBV-B-like RNA2 might have acquired TRV-like RNA2-specific coding sequences. Alternatively, a TC3³PE-like RNA2 could have regained, in mixed infections with PEBV-B, a 5' end that is PEBV-B RNA2-like also upstream of the common region. The lack of ORF2c in its RNA2-specific part and the presence of RNA1-related coding sequences in its 3' end (Table 2) indicate that it is a deletion/recombination mutant and that it cannot be a progenitor of TC3³PE-a and the other TRV AL RNA2 molecules.
Whilst we can only speculate on possible evolutionary pathways that have led to TC3′PE-like molecules, the following developments that have occurred in Alstroemeria and tulip or possibly other hosts seem to be relatively clear. In different hosts, TC3′PE-a RNA2 has obviously become associated with two slightly different TRV RNA1 species, i.e. TRV TCM RNA1 (Angenent et al., 1986) and TRV AL RNA1 (this paper). Recombinations with different-sized 3′ ends of these TRV RNA1 molecules have replaced the original PEBV RNA1-like 3′ end to yield TRV TCM RNA2 or the various TC3′AL RNA2 molecules, respectively. The new TRV RNA1-derived 3′ ends in these RNA2 molecules all contain RNA1-specific coding sequences, i.e. either a 3′ portion or the whole ORF1b and sometimes, in addition, a 3′ portion of ORF1a (Fig. 3). The RNA1-related 3′ ends of these recombiant RNA2 molecules are 100% identical to the 3′ ends of the supporting TRV RNA1 molecules in tulip or Alstroemeria, respectively (Fig. 2g). All these TRV RNA2 molecules have lost part of their RNA2-specific coding sequences, especially ORF2c and often also ORF2b (Fig. 3). A partial loss of ORF2c has also occurred in TC3′PE-b and -c in which the original PEBV RNA1-like 3′ end is still retained.

The recombinant TC3′PE, TC3′AL and TRV TCM RNA2 forms may represent various steps and side steps in a host adaptation process. Whilst the host range of PEBV is restricted mainly to legumes, TRV has one of the widest host ranges known for a plant virus (MacFarlane, 1999). The acquisition of a TRV RNA2 5′ end that enables replication by TRV RNA1-encoded enzymes and the replacement of PEBV-related coding sequences by TRV-related ones are apt to open the wide host range of TRV also to RNA2 molecules originally derived from PEBV.

METHODS

Total RNAs were extracted from infected A. aurea leaves using a RNeasy Plant Mini kit (Qiagen). cDNAs were produced with an Expand reverse transcriptase (Roche) or SuperScript III (Invitrogen) system using primer 158 (MacFarlane, 1996), which is complementary to the 3′-terminal 19 nt of tobravirus RNA2 molecules. For PCRs, the Taq Plus Long (Stratagene), Elongase (Invitrogen), Phusion (NEB) or Extensor (ABgene) system was used. The first PCRs were carried out with primer 158 as an antisense primer and either primer 179, which corresponds to the 5′ termini of several TRV RNA2 molecules (MacFarlane, 1996), or primer TR1 (Udhe et al., 1998), which corresponds to nt 75–92 of the TRV TCM RNA2 sequence as sense primers. Further primers for internal sequences were designed on the basis of the partial sequences obtained. The forward primers (names in parentheses) corresponded to the TR3′PE-a sequence to nt 805–823 in the central part of CP gene (TR113), nt 1053–1071 in the 3′ part of the CP gene (TR53), nt 1368–1387 in the central part of ORF2b (TR95), nt 2007–2027 in the 3′ part of ORF2c (TR89) and nt 2789–2817 in the central part of ORF2c (TR91). The reverse primers were complementary in the TR3′PE-a sequence to nt 2919–2938 in the 3′ part of ORF2c (TR114) and nt 3269–3296 in the PEBV RNA1-like part (PE2). In the TC3′AL-b sequence, the reverse primers were complementary to nt 1606–1624 in the 3′ part of ORF1a (TR32) and nt 2094–2117 in the 3′ part of ORF1b (TR64). Primers for obtaining overlapping PCR products for Alstroemeria TRV RNA1 were designed in conserved regions of various known TRV RNA1 sequences. PCR products were purified using a Jetsorb Gel Extraction kit (Genomed) and sequenced either directly or after cloning into the pGEM-T vector (Promega). Sequencing was carried out by a commercial company (MWG-Biotech). Sequences were assembled and analysed using Invitrogen Vector NTI Advance 10 and DNAMAN (Lynnon Bio/Soft) software.

ACKNOWLEDGEMENTS

We are greatly indebted to Drs S. A. MacFarlane and D. J. Robinson (Scottish Crop Research Institute, Invergowrie, Dundee, UK), Dr W. Burgermeister (Julius Kühn-Institute, Braunschweig, Germany) and Dr A. Guityaev (Leiden University, The Netherlands) for helpful discussions, and to Petra Lüddeke and Marianne Koerbler for expert technical assistance.

REFERENCES


Deletions in the KTER-encoding domain, which is needed for *Polymyxa* transmission, in manually transmitted isolates of *Beet necrotic yellow vein benyvirus*. *Arch Virol* 145, 165–170.


