Recombination in RNA viruses and in virus-resistant transgenic plants

Rachid Aaziz and Mark Tepfer
Laboratoire de Biologie Cellulaire, INRA-Versailles, 78026 Versailles cedex, France

Introduction
Recombination in RNA viruses is a general phenomenon, and is considered to play a major role as a driving force in virus variability and thus in virus evolution. An ever increasing number of RNA viruses has been shown to undergo RNA recombination, whether under natural or experimental conditions. More than 30 years after its discovery, the mechanisms of viral RNA recombination are only now beginning to be elucidated. Recent reports strongly suggest that RNA recombination is linked to virus replication, and that it occurs by a copy-choice mechanism. The detection of recombination between host transcripts and infecting viral RNA genomes has given rise to new concerns about the release of virus-resistant transgenic crops, since recombination could generate viruses with properties different from the infecting strain. We will describe here current knowledge of recombination in RNA viruses, and then will present results showing recombination in transgenic virus-infected plants, and finally will discuss to what extent this may lead to increased variability in plant RNA viruses.

Recombination in RNA viruses
Adaptation of living organisms to a changing environment through evolution requires a compromise between genetic variation and phenotypic selection, and has generated the considerable variability that we encounter every day. In the virus world, and particularly in that of RNA viruses, high variability is observed, and is thought to be due to three main evolutionary forces: mutation, reassortment (for viruses with a segmented genome) and recombination (Roossinck, 1997; Domingo & Holland, 1997). Viral RNA replication is characterized by a high mutation rate, due to the lack of proofreading-repair of viral RNA-dependent RNA polymerases (RdRp). This, in conjunction with short replication times and a high multiplicity, leads towards a dynamic mutant population, a unique feature of RNA viruses termed quasi-species and corresponding to a swarm of sequence variants (Holland & Domingo, 1998). However, genetic divergence is restricted by the necessity to maintain a functional viral RNA genome and by environmental selective pressures. By allowing the spread of new mutations, both reassortment and recombination increase genetic variability and favour the creation of variant viruses that may be best adapted to withstand future environmental selective pressure on the virus population. In addition, RNA recombination is thought to rescue viral genomes by repairing mutation errors in essential viral genes or in structures that could be introduced during RNA replication (Lai, 1992; Carpenter & Simon, 1996).

Until the 1960s, genetic recombination was considered to be a feature confined exclusively to DNA molecules. DNA viruses, prokaryotes and eukaryotes undergo DNA recombination during sexual reproduction, such as in meiosis or in bacterial conjugation. Likewise, DNA recombination has been reported during mitosis of somatic cells in cases of gene conversion. Recombination between DNA molecules will not be treated further here (for review see Osman & Subramani, 1998), nor will we discuss recombination in retroviruses, since their mode of replication is quite different from that of RNA viruses, and recombination is an integral part of this process (for review see Mansky, 1998).

RNA recombination was originally observed in poliovirus by Hirst (1962) and Ledinko (1963). By co-infecting HeLa cells with two distinct strains of type 1 poliovirus carrying single mutations, these authors were able to obtain recombinant viruses that had inherited the two mutations from both parental strains. Since then, recombination in RNA viruses has been abundantly documented, and the list of recombinant viruses continues to grow. Two types of analysis provide evidence of viral RNA recombination. One is based on phylogenetic analysis of RNA viral genomes. In this case, when different areas of a viral genome are apparently more closely related to different viruses, this suggests that recombination has led to a genomic rearrangement at some time during the evolution of these viruses. Indeed, as nucleotide sequences of an increasing number of RNA viruses have been determined, a growing amount of evidence for natural RNA recombination is being reported. Insertions of cellular sequences into viral RNA genomes, presumed to arise through
RNA recombination, have also been cited (Mayo & Jolly, 1991; Masuta et al., 1992; Lai, 1992). In addition, defective interfering RNAs, which are also generated by RNA recombination, have been reported in both plant and animal RNA viruses (Lazzarini et al., 1981; Fields & Winter, 1982; Inoue-Nagata et al., 1997). These examples illustrate the major role of recombination in RNA virus evolution. The second type of evidence is from the abundant direct laboratory experimental data. By making use of either non-replicative (for animal viruses) or movement-defective (for plant viruses) variant viruses, functional recombinant viruses restored through RNA recombination have been obtained in several laboratories. Thus, many RNA viruses have been shown to undergo RNA recombination (Table 1).

Based on the similarity of parental RNA molecules, two types of RNA recombination have been defined (Lai, 1992). Homologous RNA recombination (HR) occurs between two similar or related RNA molecules at precisely comparable (precise HR) or divergent (imprecise or aberrant HR) matched crossover sites. In contrast, non-homologous recombination (NHR) occurs between dissimilar or unrelated RNA molecules. This nomenclature is somewhat unsatisfactory for two reasons. First, the term homologous generates a semantic problem, since it signifies that the parental molecules share common ancestry, which is not necessarily the case, for instance, if recombination occurs in short segments of sequence identity in otherwise unrelated genes. Second, it is not informative about the mechanism by which recombination occurs. In an attempt to alleviate these problems, an alternative classification was proposed recently (Nagy & Simon, 1997). This classification is based both on the mechanism generating recombinants and on the nature of the recombinant products, and defines three classes of RNA recombination. Similarity-essential recombination occurs when sequence similarity between parental RNA molecules is essential in the recombination event. Depending on the end-products, two types are distinguished: precise and imprecise similarity-essential recombination. Similarity-non-essential recombination occurs when sequence similarity of parental RNA molecules is not apparently required; other RNA determinants such as secondary structure, heteroduplex formation, etc. determine the recombination event. Similarity-assisted recombination combines the characteristics of the two above-mentioned classes. This novel classification system may also prove to be temporary, and in particular may need to be reconsidered as new light is shed on the mechanisms of RNA recombination. In spite of its flaws, we will continue to use Lai’s system here, since it is the more widely accepted, and since in many cases the mechanism of recombination is unknown.

Two mechanistic models are generally proposed to account

### Table 1. Direct laboratory evidence of recombination in RNA viruses

<table>
<thead>
<tr>
<th>Virus group</th>
<th>Virus</th>
<th>Reference*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Animal</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picornavirus</td>
<td>Poliovirus</td>
<td>Hirst (1962)</td>
</tr>
<tr>
<td></td>
<td>Foot-and-mouth disease virus</td>
<td>McCahon et al. (1985)</td>
</tr>
<tr>
<td>Coronavirus</td>
<td>Murine hepatitis virus</td>
<td>Lai et al. (1985)</td>
</tr>
<tr>
<td></td>
<td>Infectious bronchitis virus</td>
<td>Kottier et al. (1995)</td>
</tr>
<tr>
<td>Alphavirus</td>
<td>Sindbis virus</td>
<td>Weiss &amp; Schlesinger (1991)</td>
</tr>
<tr>
<td>Orthomyxovirus</td>
<td>Influenza A virus</td>
<td>Bergmann et al. (1992)</td>
</tr>
<tr>
<td>Nodavirus</td>
<td>Flock house virus</td>
<td>Li &amp; Ball (1993)</td>
</tr>
<tr>
<td><strong>Plant</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bromovirus</td>
<td>Brome mosaic virus</td>
<td>Bujarski &amp; Kaesberg (1986)</td>
</tr>
<tr>
<td></td>
<td>Cowpea chlorotic mottle virus</td>
<td>Allison et al. (1990)</td>
</tr>
<tr>
<td>Carmovirus</td>
<td>Turnip crinkle virus</td>
<td>Cascone et al. (1990)</td>
</tr>
<tr>
<td>Tobamovirus</td>
<td>Tobacco mosaic virus</td>
<td>Beck &amp; Dawson (1990)</td>
</tr>
<tr>
<td>Alfamovirus</td>
<td>Alfalfa mosaic virus</td>
<td>van der Kayl et al. (1991)</td>
</tr>
<tr>
<td>Tombusvirus</td>
<td>Cucumber necrotic virus, tomato bushy stunt virus DI-RNA†</td>
<td>White &amp; Morris (1994)</td>
</tr>
<tr>
<td>Cucumovirus</td>
<td>Cucumber mosaic virus, tomato aspermy virus</td>
<td>Fernández-Cuartero et al. (1994)</td>
</tr>
<tr>
<td>Potyvirus</td>
<td>Zucchini yellow mosaic virus</td>
<td>Gal-On et al. (1998)</td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Qβ</td>
<td>Munishkin et al. (1988)</td>
</tr>
<tr>
<td></td>
<td>ϕB</td>
<td>Onodera et al. (1993)</td>
</tr>
</tbody>
</table>

* Only the first reported references are indicated.
† DI-RNA, Defective interfering RNA.
for RNA recombination: (i) a cleavage–ligation model, which supposes a breakage at two distinct sites of the same or different precursor RNA molecules, followed by ligation of the two cleaved RNA molecules, and (ii) an RdRp-mediated copy-choice model, in which recombination would occur during RNA synthesis if the RdRp pauses (on the donor strand) and switches to another site on the same template or to another one (acceptor strand) to resume nascent RNA synthesis (Cooper et al., 1974). Studying in vitro RNA recombination of 
Qβ bacteriophage, Chetverin et al. (1997) exclusively observed non-homologous recombinants, which they proposed to be generated by a mechanism of ligation–transesterification. However, this is the only example supporting the ligation model. In contrast, in vivo recombination studies of the same bacteriophage disagree with this model (Palasingam & Shaklee, 1992). These authors showed the existence of NHR, but also HR, which would have most likely been obtained by a template-switching mechanism. In addition, despite numerous attempts, Ramaswamy et al. (1995) could not demonstrate possible ligation activity that could be involved in RNA recombination of non-replicative RNA molecules of Sindbis alphavirus. Thus, nearly all the existing evidence supports the RdRp-mediated template-switching model. Consistent with this is the finding that RNA replication is necessary for poliovirus RNA recombination (Kirkegaard & Baltimore, 1986). Furthermore, studies on brome mosaic bromovirus (BMV) have shown that mutations in viral proteins 1a and 2a, which are components of the BMV replicase, affected viral RNA recombination (Nagy et al., 1995; Figlerowicz et al., 1997, 1998). Although RdRp-mediated template-switching is the most accepted model, it has not yet been formally demonstrated. Considering the difficulties, it is unlikely that formal proof of template switching will be obtained in in vivo recombination experiments. There is more likelihood of success with a cell-free system, such as poliovirus (Molla et al., 1991), or Qβ (Biebricher & Luce, 1992) and φ6 (Qiao et al., 1997) bacteriophages. Recent elegant studies with turnip crinkle tombusvirus replicase show that it may be possible to dissect the mechanism of RNA recombination in vitro with this system as well (Nagy & Simon, 1998a,b).

In almost all experimental studies, recombinant viruses were recovered under conditions of strong selection pressure. As a result, only those of notably higher fitness were detected, and those negatively selected were lost. Therefore, the frequency of recombination is nearly always underestimated, and perhaps some mechanisms of recombination still remain unknown. For example, previous studies have identified a recombinational hot-spot region in mouse hepatitis coronavirus (Banner et al., 1990). However, in the absence of selection, initial recombination events were randomly distributed, but after several passages of the recombinant viruses in tissue culture, recombination sites became localized to a restricted region corresponding to the initially identified hot spot (Banner & Lai, 1991). This study suggests that RNA recombination is

more frequent than previously thought, and also that the recombinant viruses obtained may reflect the results of selective pressure rather than the mechanism of recombination itself.

It would appear that different RNA viruses do not recombine with the same frequency or efficiency. Detailed phylogenetic studies on potyviral strains gave evidence of numerous recombination events between strains, whereas when cucumoviruses were analysed, no evidence of recombination was observed (Revers et al., 1996; Candresse et al., 1997). Although this may in part be due to differences in the efficiency of detection in different virus systems, it is tempting to speculate on the existence of processes regulating RNA recombination that could be determined either by the virus itself (the RdRp and/or sequence-specific or structural signals in the viral RNA) or by the host. In particular, nothing is currently known about the role of host factors in RNA recombination.

Recombination in virus-resistant transgenic (VRT) plants

The first VRT plants were obtained 13 years ago (Powell-Abel et al., 1986). They represent the first application of the concept of pathogen-derived resistance (Sanford & Johnston, 1985). Since then, numerous transgenic crops tolerant or resistant to a wide range of viruses have been developed (Beachy, 1997). Different genomic sequences were used, including viral genes encoding coat proteins (CP), replicases, defective movement proteins, proteases or helper components (Lomonossoff, 1995). CP-mediated resistance, which is the most frequently used strategy, has been used to confer resistance to viruses in at least 13 RNA virus genera, including 23 distinct virus species (Grumet, 1995). Three CP-mediated resistant plants have recently been approved for commercial release in the USA and more VRT crops will soon be deregulated (White, 1999). The increasing interest in VRT plants is understandable, since they offer numerous potential agronomical and ecological benefits. This is clearly the case when no corresponding natural host resistance genes have been identified, and also when introgression of such a natural resistance gene into commercial crops would be too lengthy a process, or when it would have undesirable effects on genetically complex plants that are vegetatively propagated. VRT plants should also allow reduction of the amounts of pesticide used to eliminate virus vectors.

Nevertheless, in order to take advantage of these benefits under the best conditions, it is appropriate to examine VRT plants carefully from the point of view of biosafety. Due to possible interactions in transgenic plants between products of the viral transgene, whether RNA or protein, and an incoming virus, three main potential ecological risks should be considered: synergism, heteroencapsidation and recombination (Tepfer, 1993; Robinson, 1996). For some species, there is also
a risk of transgene flow from VRT plants to neighbouring wild relatives (Fuchs & Gonsalves, 1997). Two types of alterations in populations of viruses or neighbouring plants can be distinguished. The first involves phenotypic alterations of VRT plants or infecting viruses, such as an increase in symptom severity due to a virus/transgene synergism or the spread of a non-transmissible virus after heteroencapsidation. This kind of risk, of course, remains restricted to the vicinity of the VRT crop, and will disappear without persistent effects on the environment if VRT plant use is halted. The second involves genotypic modifications, which in essence are expected to be durable, and may involve either plant (transgene flow) or viral (RNA recombination) genomes (for review see Tepfer & Balázs, 1997).

RNA recombination in virus-infected transgenic plants could generate novel viruses with biological properties distinct from those of the parental strain. To investigate this phenomenon, one strategy is to inoculate a movement-defective virus onto transgenic plants constitutively expressing a gene encoding the corresponding protein in non-defective form, and then look for potential recombinant systemic viruses. To date, four experimental systems concerning this phenomenon have been reported, including only three plant viruses, cauliflower mosaic caulimovirus (CaMV), cowpea chlorotic mottle bromovirus (CCMV) and tomato bushy stunt tobravirus (TBSV).

Recombinant CaMVs were first obtained after transgenic Brassica napus expressing CaMV ORF VI was agroinfected with a CaMV isolate lacking ORF VI (Gal et al., 1992). ORF VI of CaMV D4 strain determines systemic movement in Nicotiana bigelovii (Schoelz et al., 1986). Transgenic N. bigelovii expressing the D4 ORF VI inoculated with CM1841-CaMV, a strain naturally non-systemic in solanaceous hosts, also generated recombinant CaMVs (Schoelz & Wintermantel, 1993). In these two examples, strong evidence suggests that recombinant viruses arose by two template-switching events between viral RNA and transcripts from the virus-derived transgene. However, since CaMV is a pararetrovirus (with a DNA step in its virus cycle), homologous DNA–DNA recombination cannot be excluded, since it has already been shown to occur in CaMV (Vaden & Melcher, 1990; Gal et al., 1991). It should also be noted that copy-choice recombination in CaMV as described above is due to the viral reverse transcriptase, and that as a consequence this system may be rather dissimilar to what occurs in recombination mediated by an RNA virus RdRp.

Greene & Allison (1994, 1996) have elegantly demonstrated recombination between an isolate of CCMV, crippled by deletion of the 3’ one-third of the CP gene, and transgenic N. benthamiana containing the 3’ two-thirds of the CP gene and the 3’ non-coding region of the CCMV genome. Recombinant progeny became systemic on N. benthamiana, following the restoration of a functional CCMV CP gene. The presence of different mutation markers in the systemic virus unequivocally demonstrated that RNA recombination took place between the transgene mRNA and the challenging movement-defective viral genome. Sequence analysis of recombinant viral RNAs revealed an imprecise HR type, with numerous modifications flanking the putative crossover site. When the recombinant strains were tested on a range of host plants, four out of seven displayed novel symptoms (Allison et al., 1997), but when co-inoculated with wild-type virus, none of them was more fit than the parental strain (Allison et al., 1999).

All the above examples correspond to conditions of high selection pressure, i.e. the viral inoculum was movement-defective, and therefore only the systemic recombinant viruses were detected and isolated. However, since defective viruses are at most minor components of natural virus strains, searching for recombinant viruses in VRT plants infected with wild-type viruses under conditions of little or no selection pressure is more pertinent to risk assessment. In more recent work, recombinant CaMV has been isolated from transgenic plants under conditions of moderate selection pressure. Here inoculation of transgenic N. bigelovii plants expressing D4-CaMV ORF VI with W260-CaMV, which is systemic in solanaceous hosts, generated recombinant viruses that had a distinct competitive advantage in N. bigelovii compared to the parental W260 strain (Wintermantel & Schoelz, 1996).

More recently, recombinant viruses have also been recovered under conditions of moderate selection pressure from plants expressing RNA virus sequences. When plants expressing a gene encoding the TBSV CP were inoculated with virus defective for the CP gene, Borja et al. (1999) observed high frequency recovery of virus with a restored CP gene, which must have been generated by at least two crossover events. The sequence of the recombinant viruses was identical to the wild-type, and thus resulted from precise HR. In this case, the CP-defective viral inoculum was systemic, but was at a selective disadvantage relative to the restored wild-type recombinant viruses.

Potential risks associated with recombination in VRT plants

If a strategy could be developed to eliminate the possibility of recombination, while taking full advantage of virus-derived resistance, this would make any detailed evaluation of potential risks associated with recombination unnecessary. One such strategy has been proposed in which the resistance observed in certain plants expressing viral sequences is due to induction of degradation of both the transgene mRNA and the homologous viral RNA by a mechanism similar to that of post-transcriptional gene silencing (for review see Baulcombe, 1996). This type of RNA-mediated resistance could present certain advantages from a biosafety perspective. In particular, since little or no transgene-derived RNA accumulates, the likelihood of its recombination with the RNA of infecting viruses would be expected to be extremely low. There may, however, also be flaws in this type of resistance. In particular, it has been shown...
recently that infection by certain viruses blocks post-transcriptional gene silencing, and thus could be expected to cause breakdown of RNA-mediated resistance (Anandalakshmi et al., 1998; Béclin et al., 1998; Brigneti et al., 1998; Kasschau & Carrington, 1998). It is assumed that when plants are inoculated with a potato potyvirus Y strain, one of which is targeted by RNA-mediated resistance, this provides selection pressure in favour of recombination between the viral genomes, resulting in replacement of the targeted sequence. Thus, in any case, the implications of recombination in plants expressing viral sequences clearly merit close examination.

In discussions of risk assessment, two distinct but interdependent elements are generally distinguished, namely frequency and hazard (Hull, 1994). In the case of recombinant virus infection of VRT plants, evaluation of the frequency of recombination is quite difficult, in particular since the frequency is expected to be low. As discussed in the previous section, almost all recombinant viruses that have been observed reflect the post-recombination selection process, rather than recombination alone. Therefore, recombination frequency in the absolute sense is generally underestimated. Another important point is that infection of VRT plants by a related incoming virus corresponds in many ways to a frequently encountered natural situation, viz. simultaneous infection of non-transgenic plants by two related viruses. Consequently, to evaluate risks of RNA recombination in virus-infected VRT plants, one must compare the frequency of its occurrence in VRT plants to that in doubly infected non-transgenic plants. In addition, determination of these frequencies must be carried out under conditions of little or no selection pressure.

Mixed infections have not been extensively studied, except regarding the synergy that can occur between two distantly related viruses (Vance et al., 1995), and co-infection of a given host plant by two closely related viruses can be difficult to achieve. For example, in the case of two cucumoviruses, cucumber mosaic virus (CMV) and tomato aspermy virus (TAV), successful co-infection of Nicotiana hosts is the exception, and generally CMV excludes TAV (Stackey & Franck, 1990; our unpublished data). It is thus not surprising that reports of RNA recombination in mixed infections of wild-type viruses are rare, if not non-existent. Indeed, RNA recombination between two different viruses requires that they undergo RNA replication at the same time and at the same place, even down to the level of subcellular localization. In contrast, in VRT plants the virus-derived transgene is constitutively expressed in all cells, which could favour the occurrence of RNA recombination between a viral transgene mRNA and an invading viral genome. One of the prerequisites for study of recombination during mixed infections is to develop sensitive techniques for detection of recombinant molecules in the presence of a large excess of non-recombinant parental RNAs. Recently, using a sensitive RT–PCR assay, recombination events between wild-type CMV and TAV have been observed in doubly infected plants (our unpublished data).

In spite of the generally perceived importance of the frequency factor in risk assessment, it can be argued cogently that in the case of recombination in VRT plants, considering the enormous scale of the proposed future commercial releases, even extremely rare recombination events could have a significant impact. If this is the case, then it is particularly important to examine closely the other element of the risk equation, namely potential hazard. In this case, the essential question is: will recombination between transgene mRNAs and infecting viruses lead to the creation of strains with novel and more damaging properties?

To investigate the hazard element of the potential risk associated with recombination in VRT plants, one approach is to create recombinant viruses in vitro by exchanging RNA segments between different strains or between different related viruses and then to evaluate their biological properties, such as symptom severity or expansion of host range. Several groups have created recombinant cucumoviruses, and have reported often striking changes in symptomatology (Ding et al., 1996; Salánki et al., 1997; Carrière et al., 1999). However, it is premature to conclude that more virulent recombinant cucumoviruses are likely to appear in transgenic plants, since in none of the above cases was the fitness of the recombinants tested relative to that of the parental viruses. In this regard it is significant that, as mentioned above, the recombinant CCMV strains recovered in transgenic plants under conditions of high selection pressure, although expressing novel symptoms, were less fit than the parental virus (Allison et al., 1999). There is, however, one report of a recombinant cucumovirus created in the laboratory that proved to have a selective advantage relative to the parental strains (Fernández-Cuartero et al., 1994). However, this may be considered to be something of a laboratory freak, since it appeared spontaneously during several years of maintenance in the greenhouse of a pseudo-recombinant strain composed of CMV RNAs 1 and 2 and TAV RNA3, and in fact several recombination events have apparently occurred.

For a thorough evaluation of potential risks associated with recombination between viral sequences transcribed from a transgene and the genome of an infecting virus, several areas of uncertainty remain. The easiest experimental results would be obtained from a detailed comparison of doubly infected non-transgenic plants and singly infected transgenic ones under conditions of minimal selection pressure, although such an experiment is difficult to perform. Not only is this the only way to approach an evaluation of the frequency of recombination under these two situations, but it would also clarify whether the nature of the recombinants created is similar. This latter point is particularly important, since if a similar array of recombinants is generated initially in these two situations, then one would expect that selection pressure, due for instance to virus movement or to transmission to other
hosts, would lead to equivalent virus populations. In this case, one could conclude that indeed plants expressing viral sequences do not present risks different from those already present naturally in a pre-transgenic world. On the other hand, if there are distinct differences in the types of recombinants that occur under these two situations, then an evaluation of the fitness of the recombinants occurring in transgenic plants would be appropriate.

R. Aaziz was supported by an INRA doctoral fellowship. This research was in part financed by a grant from the French Ministry of Research.

References


