Recombination in a plant virus: template-switching in cauliflower mosaic virus

Nigel Grimsley, Thomas Hohn and Barbara Hohn

Friedrich Miescher-Institut, PO Box 2543, CH-4002 Basel, Switzerland

Communicated by B.Hohn

A hybrid plasmid, containing tandemly arranged pieces of two different but well-defined cauliflower mosaic virus (CaMV) genomes, was used to study the mechanism by which infectious viral DNA can escape from transforming DNA. Systemic viral infection followed inoculation of *Brassica* plants with a strain of *Agrobacterium tumefaciens* containing the hybrid plasmid in its T-DNA. Restriction mapping of uncloned viral DNA from these plants, and sequencing of relevant portions of cloned viral DNA, showed that the majority of viral progeny were probably descendants of DNA produced by transcription/reverse transcription of the viral genome, thus providing further evidence for the hypothesis that this process is normally involved in viral replication. The reverse transcriptase enzyme, which is thought to undergo an intramolecular template switch during viral replication, is shown to move very close to the 5' end of the terminal repeat on the 35S RNA molecule before switching templates. The remaining minority of viral genomes can best be explained as arising from products of recombination between homologous regions of CaMV DNA.

Key words: gene conversion/reverse transcription

Introduction

Biochemical evidence and theoretical considerations strongly suggest that cauliflower mosaic virus (CaMV) replicates via a 35S RNA transcript, which spans the entire 8-kb long viral genome and bears a 180-nucleotide long direct repeat at the termini (Guilley *et al.*, 1983; Hull and Covey, 1983; Pfeiffer and Hohn, 1983; reviewed by Hohn *et al.*, 1985). Reverse transcription of this intermediate is thought to proceed by (i) binding of the polymerase to the primer/primer binding site complex, (ii) synthesis of the anti-sense DNA strand, involving either intra- or intermolecular template switching in the region of the terminal repeat, and, subsequently or simultaneously, (iii) the synthesis of the second strand of DNA. To date, the only other well-characterized viral transcript is the 19S RNA (Covey *et al.*, 1981; Odell *et al.*, 1981; Guilley *et al.*, 1982); it encodes open reading frame VI and terminates at the same map position (m.p.) as the 35S transcript.

Oligomers of viral DNA cloned in bacterial vectors are known to give rise to systemic infection in plants inoculated with naked DNA, even though the vector part of the plasmid is present (Lebeurier *et al.*, 1982; Walden and Howell, 1983; Grimsley *et al.*, 1986). The mechanism, however, by which viral DNA escapes from such constructions has not been elucidated. Homologous recombination and/or reverse transcription of an RNA molecule spanning different viral DNA units could be envisaged. This communication describes analyses of viral progeny obtained following infection of plants with a hybrid oligomer composed

of different characterized viral strains, contained in the T-DNA of *Agrobacterium tumefaciens*. Analyses of progeny viruses present in 'agro-infected' plants showed that (i) the majority of progeny viruses probably escaped from the transforming DNA by transcription/reverse transcription, and (ii) a minority of progeny viruses probably arose from homologous recombination events occurring at different map positions. Since a variety of mechanisms are involved in the production of the 'recombinant' progeny described here, the reader is referred to Clark's definition of recombination, 'any of a set of pathways in which elements of nucleic acid interact with a resultant change of linkage of genes or parts of genes' (Clark, 1971), which encompasses all of these events.

Results

The experiments described below were designed to differentiate between the various mechanisms by which infectious viral DNA can be produced from the transforming DNA. A plasmid was constructed (pCa305) which carries one complete genome of CaMV isolate 'CM4-184' (CaMV.CM4-184, Howarth *et al.*, 1981; Dixon *et al.*, 1986) and a 3.3-kb fragment derived from CaMV isolate 'S' (CaMV.S); thus the plasmid has a partial duplication of CaMV DNA sequences (Figure 1). The homologous regions of the partial duplication show differences in their nucleotide sequences and restriction endonuclease sites (Figures 2 and 3). In addition, CaMV.CM4-184 contains a 9-bp insertion relative to CaMV.S. These differences were used to define more precisely the stretches of DNA in which recombination occurred; the genetic maps of progeny viruses give information about the map position(s) of events occurring in the initial stages of escape of infectious DNA.

Plasmid pCa305, when introduced into a host plant, would allow production of a 35S transcript, beginning in CaMV.CM4-184 and ending in CaMV.S. Reverse transcription of this RNA molecule would yield one specific type of progeny DNA molecule containing regions of CaMV.S and CaMV.CM4-184, as defined by the *SalI* cloning site and by the position at which the reverse transcriptase enzyme switches from the 5' to the 3' end of the template. A schematic summary of these events is shown in the upper part of Figure 1. On the other hand, if homologous DNA recombination were to occur between the duplicated regions of the two viral genomes, progeny DNA molecules would arise in which one junction between CaMV.CM4-184 and CaMV.S sequences would be defined by the *SalI* cloning site and the other by the position at which a recombination event occurred. Progeny viruses escaping by homologous recombination would, therefore, be expected to show a variety of genetic maps reflecting the random positions at which recombination took place. The lower half of Figure 1 shows some examples of recombination events of this kind.

pCa305 was inserted into the T-DNA of a non-oncogenic *A. tumefaciens* strain, and turnip plants were then 'agro-infected' with this strain (Grimsley *et al.*, 1986) by inoculation at the crown of the plant. Viral DNA was extracted from six turnip plants

Fig. 1. Mechanisms of viral escape from transforming DNA. Two different viral isolates, CaMV.CM4-184 and CaMV.S, represented by cross-hatching at different intensities and in opposite directions, were used to construct pCa305, which is shown in linear form. Open reading frames (ORFs) are given roman numerals and numbers relate to standard map coordinates of CaMV.S (Hohn *et al.*, 1982), which are used as a standard throughout the paper. The deletion found in ORF II of CaMV.CM4-184 is shown by a Δ . pCa305 was then introduced into plants by 'agro-infection' (Grimley *et al.*, 1986). Escape of infectious viral progeny from the transforming DNA by transcription/reverse transcription (above) leads to one specific genetic map in progeny viruses, defined by the position of transcriptional start and stop signals. On the other hand, escape by homologous DNA recombination would produce a variety of genetic maps (below).

6 weeks following inoculation. Fine restriction mapping of specific regions in the genomes of the progeny viruses from individual plants was achieved by probing Southern transfers of viral DNA digested with the enzymes *Hae*III, *Hpa*II and *Taq*I, for which there are differences in the region of interest (Figure 2). As probes a 1.1-kb *Bgl*III restriction fragment spanning the region of the postulated template switch [probe (i), Figure 2] and a 0.9-kb *Pst*I–*Pvu*II restriction fragment spanning the promoter region of 19S RNA [probe (ii), Figure 2] were chosen.

Strong 'hot-spot' indicates replicative recombination

Probing of CaMV.CM4-184 and CaMV.S DNA (controls in Figure 2 part C) show that indeed the expected restriction map differences are visualized. For example, a *Hpa*II fragment (1.65 kb) typical of CaMV.CM4-184 and two *Hpa*II fragments (1.03 and 0.85 kb) typical of CaMV.S can be visualized with probe (i). Likewise, probe (ii) reveals differences in the *Taq*I restriction pattern.

Six different plants originating from agro-infections with pCa305 were analysed using probes (i) and (ii); Figure 2 part C shows the restriction patterns of virus populations of three of the six plants. Note that in all cases the pattern of major bands follows

the pattern of CaMV.S. Probe (i) detects the 0.50/0.53-kb double band and the 0.32-kb *Hae*III bands, and also the 1.03-kb and 0.85-kb fragments generated by *Hpa*II. Probe (ii) detects the 1.31-kb and 0.34-kb *Hae*III fragments, and the *Taq*I 0.27/0.29-kb double band, which are all diagnostic of CaMV.S, as well as the 0.73-kb *Taq*I band common to both strains (Figure 2, part B). Consequently we can conclude that the majority of the progeny must have CaMV.S sequences between m.p. 5543 (typical *Taq*I site of CaMV.S) and m.p. 7226 (typical *Hae*III site of CaMV.S). This observation is consistent with an escape of these viruses from the T-DNA by transcription/reverse transcription or by homologous recombination in a preferred region, i.e. between the *Hae*III site at m.p. 7226 and the *Bst*EII site at 127, which was used for cloning and thus defines the junction between viral and plasmid DNA sequences (see Figure 1).

To delimit the hot-spot further seven viral clones were isolated from five different plants by insertion into the *Sal*I site of the vector pUC8. Restriction analysis of the clones (data not shown) revealed that they were descendants of the majority populations described above. In addition, the presence of the 9-bp insertion characteristic of CaMV.CM4-184 was revealed on agarose gels by the slower movement of the appropriate restriction enzyme fragments (data not shown). Figure 2 part B shows the observed restriction maps of the majority progeny viruses. 0.8-kb *Hinc*II or 0.3-kb *Eco*RV to *Bgl*III restriction fragments (Figure 2) from the pUC8 clones were sequenced (Figure 3). Comparison with the parental strain revealed that, in all cases, the maps were consistent with a switch within 45 nucleotides of the 5' end of the 35S RNA molecule since sequences typical of CaMV.S were present from m.p. 5543 to m.p. 7226, but the 9-bp insertion and other nucleotide exchanges typical of CaMV.CM4-184 were present in all cases. These observations could also result from the presence of a 'hot-spot' for homologous DNA recombination in the 'unassigned' region (Figure 3, bottom), which constitutes only about 1/13 of the total DNA homology available.

'Recombination' events in minority progeny

A small proportion of virus particles [0–20% (see below), depending on the particular turnip plant analysed] cannot have arisen in the way described above because they possess the restriction pattern characteristic of strain CaMV.CM4-184 upstream of the postulated template switch site. The results depicted (Figure 2) were chosen to represent plants containing different minority progeny virus populations (compare the less intense bands seen in digests from test plants, marked 'O' or arrowed, with the parental maps and controls). These viruses could have been produced by a number of different mechanisms: (i) homologous DNA recombination (Figure 1), (ii) template-switching of the reverse transcriptase enzyme at positions other than the termini of the 35S RNA ('illegal' template-switching), and (iii) mismatch repair of a region of hybrid DNA formed at the position of homologous recombination (somatic gene conversion). A pictorial representation of recombination events occurring in different plants is given in Figure 4.

Some viral genomes are probably products of homologous DNA recombination. In plant 2 (Figure 2), probing with fragments (i) and (ii) reveals several minor bands typical of CaMV.CM4-184. Probe (i) detects the 0.86-kb *Hae*III fragment and the 1.65-kb *Hpa*II band. Probe (ii) shows the 0.47-kb and 1.19-kb *Hae*III bands, and also the 0.55-kb *Taq*I band. Thus minority progeny viruses representing CaMV.CM4-184 sequences between m.p. 5543 (defined by the absence of the CaMV.S *Taq*I site), and m.p. 127 (defined by the *Bst*EII cloning site), are present (Figure 4). An initial escape event must, therefore, have occurred upstream

Fig. 2. Predicted and observed restriction maps of progeny viruses. **(A)** Maps of parental strains. **(B)** Maps of the progeny strain (H35S) predicted by template-switching of the reverse transcriptase enzyme between the terminal repeats of the 35S transcript. **(C)** Observed band patterns of control DNA from cloned viral sequences (see Materials and methods) or mixtures of progeny viruses extracted from individual plants, produced by probing Southern transfers either (i) with the 1.1-kb *Bgl*II fragment or (ii) with the 0.9-kb *Pst*I–*Pvu*II fragment. The scale showing approximate band sizes in kilobase pairs was calculated using *Hind*III-digested bacteriophage lambda DNA and *Hae*III-digested Φ X174 DNA run alongside the viral DNA (data not shown). O, —: these bands indicate the presence of progeny escaping from the T-DNA by mechanisms other than the transcription/reverse transcription replication route observed in normal viral infection. Only one of the two 19S transcripts (see Figure 1) is shown. It should be noted that some of the bands in the *Hae*III digests do not correspond precisely to those shown in parental controls, because: (i) the CaMV construction used for agro-infection is CaMV.CM4-184 left of the *Sal*I site and CaMV.S right of this site; thus a new *Hae*III fragment 1314 bp long is expected and is indeed observed, and (ii) the controls pCa74 and pCa292 were constructed (see Materials and methods) by cloning into the *Sal*I site of plasmid vectors and therefore contain bacterial sequences flanking the *Sal*I site. The expected size for the *Hae*III flanking fragment of pCa74 is 1.16 kb, as observed, and the expected size for the *Hae*III flanking fragment of pCa292 is 1.19 kb, as observed.

Fig. 3. The position of template switching of the reverse transcriptase enzyme. Sequencing of 0.8-kb fragments from seven individual clones of viral DNA isolated from five different turnip plants revealed that template switching occurs between map positions 7434 and 7479 (CaMV.S coordinates). In all cases, sequences to the right of this were like CaMV.CM4-184 and to the left like CaMV.S, as summarized at the bottom. The base marked * could not be clearly identified on sequencing gels, although the 9-bp insertion present in all cases could clearly be seen on restriction mapping gels (data not shown). The presence of the base marked ** was verified by restriction mapping (*Hae*III site).

of m.p. 5543. The underlying mechanism for this event could be either homologous DNA recombination or illegitimate template-switching. However, the region upstream of m.p. 5764 (19S transcript promoter, see Figure 2) in CaMV.CM4-184 is unlikely to be transcribed in plants at the initial stage of escape of infectious viral DNA from transforming DNA, because no further CaMV promoter has been identified between the 19S promoter and the *Sa*II site. Furthermore only bacterial sequences are present upstream of the *Sa*II site. Thus, 'illegal' template-switching is not a plausible explanation for the observed 0.55-kb *Taq*I fragment (arrowed), and we conclude that some progeny genomes were probably produced by homologous recombination within this 0.7-kb region (error bar, Figure 4). Although the occurrence of the documented event might obscure the observation of further events (all bands typical of CaMV.CM4-184 are present), the relative intensities of the minor bands suggest that the minority viral progeny are descendants of a single homologous recombination event.

A variety of recombination events can be detected within individual plants. **Plant 1:** the presence or absence of particular minority bands (Figure 2) allows the map positions of three discrete events to be deduced (Figure 4): (a) The 0.86-kb *Hae*III fragment visualized by probe (i) represents CaMV.CM4-184 sequences at m.p. 7726, but no trace of the 1.65-kb CaMV.CM4-184-specific *Hpa*II band is visible. In the *Hpa*II digestion, only 1.0-kb and 0.85-kb bands are seen, thus indicating the progeny viruses to be like CaMV.S at m.p. 7119. A cross-over event involving all of the detectable minority progeny viruses must therefore have occurred between m.p. 7119 and m.p. 7226. (b) In the same way, a cross-over event can be localized between m.p. 6070 (the presence of the 0.47-kb CaMV.CM4-184 *Hae*III band shows that the *Hae*II site of CaMV.S is absent in a minority of progeny

Fig. 4. A summary of recombination events observed in three different plants, as determined by restriction mapping. Part of pCa305 is shown, with the homologous CaMV sequences opposite to each other. 'R' indicates the position encoding the 180-nucleotide long terminal repeat of the 35S RNA (but no transcripts are shown). On the standard CaMV.S scale '+' represents a restriction site difference between the two CaMV strains and '*' represents the 9-bp insertion typical of CaMV.CM4-184. The switch position seen in the majority of progeny viruses in all plants is indicated by a thick line, and events seen in minority viruses in individual plants by thinner lines. Error bars indicate the regions of DNA to which particular events have been delimited by restriction map differences. '*': this line is broken because these minority events could indicate the presence of either one or two populations of minority progeny (see Results).

viruses) and m.p. 6512 [this *Hpa*II site typical of CaMV.CM4-184 must be absent, because a minor band of 0.61 kb, detectable by probe (i), which would be produced by a cross-over event between CaMV.S *Hpa*II site at m.p. 7119 and the CaMV.CM4-184 *Hpa*II site at m.p. 6512, is not observed]. The recombination event(s) observed in (b) may or may not involve those viral genomes observed to recombine at the position described in (a), but might involve a population of progeny viruses which have recombined at the 5' end of the 35S RNA in the same way as the majority progeny. Thus this line in Figure 4 is broken with an asterisk. (c) All detectable minority progeny viruses show recombination between m.p. 5543 (CaMV.S-specific *Taq*I site; no CaMV.CM4-184-specific 0.55-kb band visible) and m.p. 5941 [1.2-kb CaMV.CM4-184 *Hae*III band detected by probe (ii), marked 'O' in Figure 2]. Any of the three mechanisms previously referred to could be proposed to account for these observations. All of the recombination events took place in regions transcribed in plants (Figure 2), and perhaps 'illegal' template-switching between the CaMV.CM4-184 19S transcript and the 35S transcript occurred. In this case the third (c) event may have involved switching from the 5' end of the 19S transcript at m.p. 5764 (Guilley *et al.*, 1982) to the 35S transcript. However, restriction mapping can only delimit this event to within ~0.5 kb (Figure 4).

Plant 3: probe (i) does not detect any minor bands (Figure 2C), whereas probe (ii) shows minor bands typical of CaMV.CM4-184. These data can be interpreted, using similar arguments to those used above, to indicate recombination events between m.p. 6070 and m.p. 6512, and between m.p. 4837 and m.p. 5543. However, a third event must have occurred between m.p. 6512 and m.p. 127 to produce viable viral progeny, since plasmid sequences border m.p. 127 (*Bst*EII site). The third event in minority progeny could have occurred either at the same map position as the switch point seen in the majority population, or at another position which cannot be identified due to the similarity of restriction maps between m.p. 7226 and m.p. 127.

Frequency of different events in different plants. A survey of the various recombination events occurring in the minor population of CaMV.CM4-184 molecules was obtained by densitometer

Table 1. Analysis of minor bands in six different plants

Plant	Percentage of viral DNA representing CaMV.CM4-184 sequences present at the given map positions			
	5543	6070	7119	7226
1	0-2	2-10	0-2	10-20
2	10-20	10-20	2-10	10-20
3	2-10	2-10	0-2	0-2
4	< 10 ^a	2-10	0-2	0-2
5	2-10	10-20	0-2	2-10
6	10-20	10-20	10-20	0-2
Totals	4/5 ^b	6/6	2/6	3/6

^aCould not be estimated more accurately.

^bPlant 4 not included.

The percentage of CaMV.CM4-184 sequences present in six plants was estimated at four different map positions in each plant by densitometric comparison of major and minor bands on autoradiographs, with stoichiometric correction for mol. wts. Thus at m.p. 7226, for example, using the *Hae*III digest, the minor band at 0.86 kb (characteristic of CaMV.CM4-184) was compared with the major band at 0.32 kb (CaMV.S), and likewise for other map positions using appropriately chosen bands (see Figure 2). The number of plants containing >2% of CaMV.CM4-184 sequences at particular map positions is given in the 'Totals' row.

measurements from autoradiographs made as previously described. Although all six plants analysed contained a variable proportion of minor bands, the distribution of these bands with respect to map position was not random (Table 1).

Discussion

Following infection of host plants with constructions of CaMV oligomers sandwiched between T-DNA border sequences, symptoms of systemic viral infection appeared (Grimsley *et al.*, 1986). Agro-infection of plants was used in preference to inoculation with naked DNA because (i) agro-infection is much more efficient, producing a larger number of transformation events, and (ii) only sequences flanked by T-DNA borders are transferred, eliminating the possibility of inoculation of whole viral genomes which may have recombined out of the plasmid during its replication in the bacterial cells. Escape of infectious viral DNA in plants from the transforming DNA could have occurred by two different, but not mutually exclusive, routes: transcription/reverse transcription or DNA recombination. To be able to distinguish between these possibilities a hybrid CaMV construction was used consisting of pieces of known and differing sequences. Using this approach, a molecule containing more than a unit viral genome provides a good substrate for recombination, since only one event is required to produce infectious progeny viruses.

Of the 3.3 kb 'available' for homologous recombination 0.24 kb or less was shown to have been used in five out of five independent viral isolates arising from infection with the CaMV construction. This might point to a very active recombinational hot spot, or to a special mechanism resulting in localized recombinational activity. Increased virulence of a virus arising from a particular kind of recombination event cannot be disregarded because spread of viral infection is, of course, a highly selective process. Although the observed 'hot-spot' occurs in an intergenic region, comparisons of DNA sequences between different strains of CaMV show that this region is more highly conserved than adjacent regions, except for the 9-bp insertion present in some strains (Dixon *et al.*, 1986). It is, therefore, unlikely that the 'hot-spot' is due to increased tolerance to alterations in this region, which might arise at the site of recombination. Since the 'hot-

spot' coincides with the region of the CaMV terminal repeat we conclude that the escape of an RNA genomic form was responsible for the observed recombination event. Template switching of reverse transcriptase on this RNA produces a hybrid CaMV molecule. Theoretically, it is also possible for reverse transcription to begin on the RNA molecule transcribing into plasmid sequences (Figure 1), since both the 35S transcript promoter and primer binding site are present, and then switch to the complete 35S RNA transcript, producing progeny viruses which lack the 9-bp insertion characteristic of CaMV.CM4-184. However, sequencing data show that this cannot be a common event, probably because (i) the CaMV.S sequences downstream of the *Bst*EII site (Figure 1) are absent and the transcript cannot be efficiently processed and (ii) intermolecular template switching of the reverse transcriptase may be less efficient than intramolecular template switching.

The position within the 180-nucleotide terminal repeat of the 35S RNA at which reverse transcriptase changes templates has been determined to be closer than 45 nucleotides from the 5' end. It could well be at the very end, thereby assuring maximal base pairing between the nascent DNA strand and the new template. In addition, it seems rather likely that an enzyme working on a template would not leave it unless forced to do so. Our genetic data therefore support the biochemical observation of 'strong stop' DNA (Turner and Covey, 1984). In one related examination (B. Gronenborn, unpublished data) a DNA sequence corresponding to within one nucleotide from the 5' end of the 35S RNA transcript has been found linked to a region further upstream in the genome, suggesting that an 'illegal' template switch occurred, after movement of the reverse transcriptase to the end of its template.

In every single plant analysed minor bands were observed, with sizes characteristic of strain CaMV.CM4-184, indicating the escape of progeny viruses by a different route(s) to that shown above. A proportion of these molecules was most likely formed by homologous recombination. This is shown, in particular, by the presence of the *Taq*I fragment (arrowed, Figure 2) present in plants 2 and 3, which is probably not a product of 'illegal' template-switching. Since, in this case, homologous DNA recombination is the most likely explanation, at least part of the minority populations of progeny viruses in other plants may have arisen in this way.

Analysis of the restriction maps produced from the presence or absence of minor bands in individual plants has allowed, in some cases, the position of the different recombination events to be delimited to a certain stretch of DNA and, in one case, to within ~100 bp (see Figure 4).

In all plants analysed, the involvement of multiple template switches during reverse transcription, from a 35S RNA template to a 19S RNA molecule, for example, and from there back to a 35S template ('illegal' template-switching), is not excluded by the experimental results, but its existence would be obscured by homologous recombination prevailing in the minority events. Homologous recombination is now being investigated using constructions in which reverse transcription of the large transcript would not result in infectious viral DNA.

The minority progeny present in plants 1 and 3 (Figures 2 and 4, see Results) cannot have arisen by a single recombination event at one map position. Plant 1 minority progeny could have been produced either by a number of recombination events occurring simultaneously at the time of production of infectious viral DNA, or by recombination between two different independently arising progeny, but our data cannot distinguish between these possi-

bilities. Perhaps somatic gene conversion (or mismatch repair) can also occur in regions of hybrid DNA.

Our interpretation of the minor bands is necessarily limited to simple events occurring within the homologous regions of DNA available for recombination at the time of viral escape; the actual population of progeny viruses within one plant may contain recombinants at undetectable levels, viruses which have undergone many recombination events, and non-viable virions. Thus, the observation of inexplicable bands [for example, Figure 2, (ii), *Hae*III band at 0.7 kb (all plants)] is not very surprising.

As previously pointed out (Grimmsley *et al.*, 1986) the presented data give no information about the stage at which the transforming DNA serves as a substrate for recombination. A non-oncogenic strain of *Agrobacterium* was used, and there was little or no proliferation of plant cells with integrated T-DNA at the inoculation site. Thus a continuous and increasing stock of integrated oligomeric viral genomes, which would be produced by a growing crown gall tumour, was not available for the production of viral progeny. Early events might therefore produce most of the observed recombinants. In addition, T-DNA intermediates (Koukolíková-Nicola *et al.*, 1985) might be more accessible to enzymes involved in homologous recombination and/or transcription *en route* between entry into the plant cell and integration into the nuclear DNA than they are following integration; in mammalian systems there is ample evidence that transforming DNA molecules can recombine efficiently (Song *et al.*, 1985), whereas homologous recombination involving chromosomal DNA occurs with low efficiency (Smithies *et al.*, 1985).

The observation of recombination due to template-switching of reverse transcriptase on a hybrid RNA molecule is complemented by the DNA sequence analysis of CaMV.CM4-184, which shows that it probably evolved from intermolecular template-switching events between the two ancestors 'pre-CaMV.S' and 'pre-CaMV.CM4-1841' (Dixon *et al.*, 1986). Such events may have been responsible, at least in part, for the diversification of caulimoviruses in nature. 'Illegal' template-switching, either anywhere in the genome or using the 19S RNA as a template, could also provide a means of rapid evolutionary viral adaptation since the product of gene VI, which is encoded on the 19S RNA, is thought to be important in determining viral symptoms and host range (Daubert *et al.*, 1984). Perhaps this is why the virus produces two major transcripts.

Materials and methods

Standard techniques for restriction enzyme digestions, agarose slab gel electrophoresis, ligations, bacterial transformations, Southern transfers, synthesis of radioactive probes and nucleic acid hybridizations were used (Maniatis *et al.*, 1982).

Plasmid constructions, bacterial strains and inoculations of plants

pCa74 was constructed by insertion of a complete genome of CaMV.CM4-184 in the *Sal*I site of pBR322 (Howarth *et al.*, 1981, clone pCaMV20). Construction of pCa305, pCa292, bacterial strains and agro-infection of plants have been described previously (Grimmsley *et al.*, 1986).

Analysis of viral progeny in individual plants

Viral DNA, containing a mixture of *in vivo* recombined DNA molecules, was extracted (Gardner and Shepherd, 1980) from individual turnip plants, *Brassica rapa* var. 'Just Right', 6 weeks following agro-infection with pCa305, then digested and Southern transfers probed as shown. Controls: pCa74 or pCa292 (see above) were digested with restriction enzymes as shown and transferred to nitrocellulose.

Cloning and sequencing of progeny viruses

Individual viral genomes from test plants were cloned into the *Sal*I site of pUC8 (Vieira and Messing, 1982). 0.8-kb *Hinc*II fragments, and 0.3-kb *Eco*RV to *Bgl*II fragments (see Figure 2), from these clones were then sequenced following insertion into the *Hinc*II, and *Sma*I to *Bam*HI sites, respectively, of M13 mp8 (Messing *et al.*, 1981).

Acknowledgements

We especially acknowledge the help of Thomas Nyffenegger in cloning and sequencing some of the viral isolates, we thank Z.Koukolíková-Nicola and M.Saul for discussions and useful comments on the manuscript, G.Delley for help in preliminary experiments, and B.Gronenborn for providing unpublished results. Part of this work was supported by an EMBO long-term fellowship to N.G.

References

- Clark, A.J. (1971) *Annu. Rev. Microbiol.*, **25**, 436–461.
- Covey, S.N., Lomonosoff, G.P. and Hull, R. (1981) *Nucleic Acids Res.*, **9**, 6735–6747.
- Daubert, S.D., Schoelz, J., Debaio, L. and Shepherd, R.J. (1984) *J. Mol. Appl. Genet.*, **2**, 537–547.
- Dixon, L., Nyffenegger, T., Delley, G., Martinez-Izquierdo, J. and Hohn, T. (1986) *Virology*, in press.
- Gardner, R.C. and Shepherd, R.J. (1980) *Virology*, **106**, 159–161.
- Grimmsley, N.H., Hohn, B., Hohn, T. and Walden, R.M. (1986) *Proc. Natl. Acad. Sci. USA*, in press.
- Guilley, H., Dudley, R.K., Jonard, G., Balasz, E. and Richards, K.E. (1982) *Cell*, **30**, 763–773.
- Guilley, H., Richards, K.E. and Jonard, G. (1983) *EMBO J.*, **2**, 277–282.
- Hohn, T., Richards, K. and Lebeurier, G. (1982) *Curr. Top. Microbiol.*, **96**, 193–236.
- Hohn, T., Hohn, B. and Pfeiffer, P. (1985) *Trends Biochem. Sci.*, **5**, 205–209.
- Howarth, A.J., Gardner, R.C., Messing, J. and Shepherd, R.J. (1981) *Virology*, **112**, 678–685.
- Hull, R. and Covey, S.N. (1983) *Trends Biochem. Sci.*, **8**, 119–121.
- Koukolíková-Nicola, Z., Shillito, R.D., Hohn, B.H., Wang, K., Van Montagu, M. and Zambryski, P. (1985) *Nature*, **313**, 191–196.
- Lebeurier, G., Hirth, L., Hohn, B. and Hohn, T. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2932–2936.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY.
- Messing, J., Crea, R. and Seeburg, P.H. (1981) *Nucleic Acids Res.*, **9**, 309–321.
- Odell, J.T., Dudley, R.K. and Howell, S.H. (1981) *Virology*, **111**, 377–385.
- Pfeiffer, P. and Hohn, T. (1983) *Cell*, **33**, 781–789.
- Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. and Kucherlapati, R.S. (1985) *Nature*, **317**, 230–234.
- Song, K.-Y., Chekuri, S.R., Ehrlich, S. and Kucherlapati, R. (1985) *Mol. Cell. Biol.*, **5**, 3331–3336.
- Turner, D.S. and Covey, S.N. (1984) *FEBS Lett.*, **165**, 285–289.
- Vieira, J. and Messing, J. (1982) *Gene*, **19**, 259–268.
- Walden, R.M. and Howell, S.H. (1983) *Plant Mol. Biol.*, **2**, 27–31.

Received on 30 September 1985; revised on 20 January 1986