Spontaneous Mutagenesis of a Plant Potyvirus Genome after Insertion of a Foreign Gene

VALERIAN V. DOLJA,[1] KERRI L. HERNDON,[1] THOMAS P. PIRONE,[2]
AND JAMES C. CARRINGTON[1]*

*Department of Biology, Texas A&M University, College Station, Texas 77843-3258,[1] and
Department of Plant Pathology, University of Kentucky,
Lexington, Kentucky 40546[2]*

The RNA genome of tobacco etch potyvirus (TEV) was engineered to express bacterial β-glucuronidase (GUS) fused to the virus helper component proteinase (HC-Pro). It was shown previously that prolonged periods (~1 month) of TEV-GUS propagation in plants resulted in the appearance of spontaneous deletion variants. Nine deletion mutants were identified by nucleotide sequence analysis of 40 cDNA clones obtained after polymerase chain reaction amplification. The mutants were missing between 1,741 and 2,074 nucleotides from TEV-GUS, including the sequences coding for most of GUS and the N-terminal region of HC-Pro. This region of HC-Pro contains determinants involved in helper component activity during aphid transmission, as well as a highly conserved series of cysteine residues. The deletion variants were shown to replicate and move systemically without the aid of a helper virus. Infectious viruses harboring the two largest HC-Pro deletions (termed TEV-2del and TEV-7del) were reconstructed by subcloning the corresponding mutated regions into full-length DNA copies of the TEV genome. Characterization of these and additional variants derived by site-directed mutagenesis demonstrated that deletion of sequences coding for the HC-Pro N-terminal domain had a negative effect on accumulation of viral RNA and coat protein. The TEV-2del variant possessed an aphid-nontransmissible phenotype that could be rescued partially by prefeeding of aphids on active HC-Pro from another potyvirus. These data suggest that the N-terminal domain of HC-Pro or its coding sequence enhances virus replication or genome expression but does not provide an activity essential for these processes. The function of this domain, as well as a proposed deletion mechanism involving nonhomologous recombination, is discussed.

The potyviruses belong to the picornavirus-like supergroup of positive-strand RNA viruses and represent the largest and economically most important group of plant viruses (33). The potyvirus genome of ~10 kb codes for several replication-associated proteins, three proteinases, a single capsid protein typical of filamentous viruses, and accessory proteins necessary for processes such as aphid transmission and cell-to-cell movement (Fig. 1A) (20, 43). All potyvirus proteins are derived by virus proteinase-mediated processing of a single polyprotein. The nuclear inclusion protein a (NIa) proteinase is involved in the maturation of replication-associated and capsid proteins (9). The P1 proteinase (also termed the 35-kDa proteinase) originates from the N-terminal region of the polyprotein and catalyzes cleavage only at its own C terminus (47, 48). The helper component proteinase (HC-Pro), which is adjacent to P1 within the polyprotein, also self-cleaves at its own C terminus (8, 11). Each of the proteinases is a multidomain protein with the proteolytic domain residing in the C-terminal region. The N-terminal domain of HC-Pro confers helper component activity, which is required for aphid-mediated transmission of potyviruses (3, 46). Interestingly, both the N- and C-terminal domains of HC-Pro were previously found to have sequence similarities with the p29 protein encoded by the hypovirulence-associated virus (HAV) of *Cryphonectria parasitica*, a filamentous fungus that causes chestnut blight disease (14, 30, 38).

In a previous study, we described the cloning of a full-length DNA copy of the tobacco etch potyvirus (TEV) genome containing an insertion of a reporter gene coding for β-glucuronidase, or GUS (18) (Fig. 1B). Inoculation of plants with the corresponding RNA transcripts resulted in systemic infection and expression of GUS fused to the N terminus of HC-Pro (Fig. 1B). This virus permitted monitoring of TEV replication and movement in planta by sensitive histochemical and fluorometric GUS assays. Although plant-to-plant transfer of TEV-GUS every 4 to 6 days resulted in stable retention of the foreign insert, prolonged propagation for several weeks in individual plants led to spontaneous deletions of the GUS sequence. These mutants also appeared to harbor deletions of the HC-Pro coding sequence (18). The deletion of inserted heterologous genes was also described for three other dissimilar plant viruses (13, 19, 44), although in these cases virus-specific sequences appeared to remain intact. In addition to providing insight into genome recombination and rearrangement, these data highlight a potential problem in the application of RNA viruses as expression vectors (1, 7).

Here we report the sequencing and functional characterization of the several spontaneous TEV-GUS deletion mutants. We found that nearly all deletion variants were missing most of the GUS gene, as well as the region specifying the cysteine-rich, N-terminal domain of HC-Pro. Although these viruses were viable, they exhibited reduced levels of replication and were nontransmissible by aphids. Analysis of the deletion breakpoints implicated nonhomologous recombination as the likely mechanism of mutagenesis.
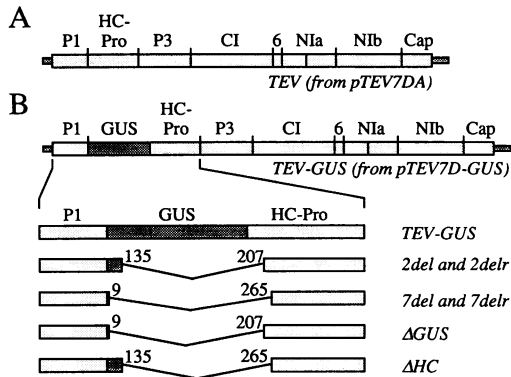
* Corresponding author.

FIG. 1. Genetic maps of TEV and selected TEV variants used in this study. (A) Wild-type TEV-HAT as derived from transcription of pTEV7DA. The single open reading frame is indicated by the extended, lightly shaded box. Positions encoding polyprotein cleavage sites are shown by the vertical lines. The names of individual proteins are above the map. (B) TEV-GUS and deletion variants. TEV-GUS differs from TEV-HAT in that the former contains the coding region for β-glucuronidase (GUS) inserted between the P1 and HC-Pro sequences. The polyprotein encoded by TEV-GUS contains a functional P1 cleavage site between P1 and GUS, resulting in the formation of a GUS–HC-Pro fusion protein. The sequences missing from the deletion variants of TEV-GUS are shown in the expanded diagram. The numbers indicate the nucleotide positions of the deletion endpoints in the GUS and HC-Pro coding sequences. CI, cylindrical inclusion protein; Cap, capsid protein.

MATERIALS AND METHODS

Propagation and purification of virus and RNA. Routine mechanical inoculation of young *Nicotiana tabacum* cv. Xanthi nc plants with wild-type TEV-HAT (obtained from William G. Dougherty, Oregon State University) and TEV mutants was performed with the aid of Carborundum with homogenates prepared by grinding infected leaves in 2 volumes of 10 mM Tris-HCl–1 mM EDTA (pH 7.6). Homogenates were aliquoted, frozen at −85°C, and thawed prior to inoculation as needed. Virus was purified according to the method of Dougherty and Hiebert (21). In some cases, crude virus preparations from individual infected plants were obtained by the same method but with exclusion of the CsCl gradient step. Viral RNA was purified by proteinase K digestion, phenol extraction, and ethanol precipitation (12).

Cloning and sequencing of spontaneous TEV deletion mutants. cDNA spanning the region of deletion in the GUS–HC-Pro coding sequence was obtained by reverse transcription of RNA from crude virus preparations followed by polymerase chain reaction (PCR). Reverse transcription and PCR were conducted according to established procedures (27) with only minor modifications, including the use of reverse transcriptase buffer from GIBCO BRL rather than PCR buffer in the initial step. The first-strand primer was complementary to TEV nucleotides 1456 to 1477 within the HC-Pro coding region (the numbering was in accord with Allison et al. [2]), while the second-strand primer corresponded to nucleotides 781 to 799 within the P1 region. The first- and second-strand primers contained *Hin*dIII and *Bam*HI recognition sequences, respectively, to facilitate insertion into the vector pTL7SN (39). All plasmids were propagated in *Escherichia coli* HB101. The sequence of the insert DNA spanning the deletion endpoints in recombinant plasmids was determined with the Sequenase kit (U.S.

Biochemicals) and a primer corresponding to TEV nucleotides 1015 to 1032. Each spontaneous deletion variant identified was assigned a code designation (TEV-1del and TEV-2del, etc.).

Construction of full-length cDNA clones of genomes representing selected deletion variants. The sequences between nucleotides 663 and 2416 in the genomes of deletion variants TEV-2del and TEV-7del were reverse transcribed, amplified by PCR, and digested with *Sna*BI and *Eco*RI, which recognize sites within the P1 and HC-Pro coding regions, respectively. The cleaved PCR fragments were cloned into *Sna*BI- and *Eco*RI-digested pTL7SN.3-0027DA, which contains sequences corresponding to the 5′-terminal 2,681 and 3′-terminal 154 nucleotides of TEV genome. This plasmid was derived from pTL7SN.3-0027D (18) by the removal of 2 nonviral nucleotides separating the 5′ terminus of the TEV sequence from the start site of SP6 RNA polymerase transcription (46a). The resulting plasmids, therefore, possessed the sequences surrounding the deletion breakpoints that were present in TEV-2del and TEV-7del. To generate full-length copies of TEV genomes containing these deletions, the *Bst*EII-*Bst*EII fragment corresponding to TEV nucleotides 1430 to 9461 was isolated from pTEV7D (18) and was inserted into *Bst*EII-digested derivatives of pTL7SN.3-0027DA. These plasmids were named pTEV7DA-2del$_r$ and pTEV7DA-7del$_r$ (Fig. 1B). Viruses derived from transcripts of these plasmids were designated TEV-2del$_r$ and TEV-7del$_r$ (the "r" subscript indicates reconstructed virus). As a control, the same manipulations (reverse transcription, PCR, and subcloning) were performed with the wild-type TEV-HAT genome, resulting in plasmid pTEV7DA$_r$.

Site-directed mutagenesis. The deletion mutant TEV-2del (and TEV-2del$_r$) retained nucleotides 1 to 135 from GUS but lost HC-Pro sequence up to nucleotide 207 (Fig. 1B). Two additional deletion variants were generated by site-directed loop-out mutagenesis of pTEV7DA-2del$_r$ (31). The GUS sequence between nucleotides 10 and 135 was removed, resulting in pTEV7DA-ΔGUS. The HC-Pro deletion in pTEV7DA-2del$_r$ was enlarged by the removal of sequence corresponding to HC-Pro nucleotides 208 to 265, yielding pTEV7DA-ΔHC (Fig. 1B).

In vitro transcription and inoculation of plants. RNA transcripts capped with m^{7}GpppG were synthesized with bacteriophage SP6 RNA polymerase (Ambion) and CsCl gradient-purified, *Bgl*II-linearized plasmid DNA as described previously (10). The undiluted transcription mixtures were applied manually to the leaves of young tobacco plants (10 μl per leaf, two leaves per plant) that were dusted with Carborundum.

Immunoblot analysis and Northern (RNA) hybridization. Total sodium dodecyl sulfate (SDS)-soluble proteins were extracted from leaf tissue by homogenization in 4 volumes of protein dissociation buffer (0.625 M Tris-HCl [pH 6.8], 2% SDS, 10% 2-mercaptoethanol, 10% glycerol) and were subjected to immunoblot analysis with anti-capsid or anti-HC-Pro sera as described previously (42). Quantitation of capsid protein levels in leaf extracts was done by using capsid protein standards from purified virus and by densitometry of immunoblots with a Model 620 Video Densitometer (Bio-Rad).

The isolation of total RNA from leaf tissue was as described previously (12). RNA (5 μg) was denatured by glyoxal treatment, separated by electrophoresis through a 1% agarose gel, and blotted onto a GeneScreen Plus membrane (DuPont). Northern hybridization was performed with a randomly primed, ^{32}P-labeled probe prepared from a DNA
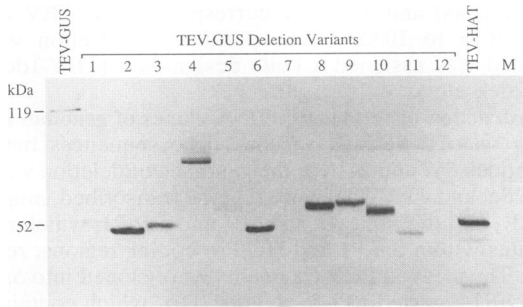
FIG. 2. Immunoblot analysis of extracts from upper leaves of TEV-GUS-inoculated tobacco plants with anti-HC-Pro sera. Each lane contains a sample from an individual plant inoculated with TEV-GUS, TEV-HAT, or buffer (lane M). All plants were propagated 4 to 5 weeks postinoculation except for the plant in lane TEV-GUS, which was propagated for 6 days postinoculation. Each of the samples in lanes 1 to 12 contains a TEV-GUS-derived, spontaneous deletion variant that encoded a truncated GUS–HC-Pro fusion protein. Lanes 4 and 3 contain extracts from plants representing first and second passages, respectively, that derived from a single TEV-GUS-infected plant. All lanes were loaded with equivalent amounts of extract. The weak appearance of HC-Pro-related proteins in lanes 1, 5, 7, and 12 is due to their low yield or rapid turnover. The molecular masses of wild-type HC-Pro and GUS–HC-Pro fusion protein are shown.

TABLE 1. Summary of TEV-GUS deletion mutants

Deletion[a]	Deletion length (nt)	5' end deletion within GUS[b]	3' end deletion within HC-Pro[c]	No. of cDNA clones[d]
TEV-1del	1,787	31 (G)	No deletion	5
TEV-2del	1,890	135 (T)	207 (G)	1
TEV-3del	1,828	80 (A)	90 (A)	1
TEV-4del	1,939	66 (T)	187 (G)	9
TEV-5del	1,741	92 (A)	15 (C)	3
TEV-6del	1,789	81 (T)	52 (T)	2
TEV-7del	2,074	9 (T)	265 (C)	10
TEV-8del	1,804	91 (G)	77 (C)	8
TEV-9del	1,840	84 (G)	106 (A)	1

[a] Variants TEV-4del, TEV-5del, and TEV-6del were isolated from the same plant, as were variants TEV-8del and TEV-9del. All other variants were from individual plants. Note that the numbering of deletions does not correspond to the numbering of lanes in Fig. 2.

[b] Numbering starts with the first nucleotide of the GUS coding sequence. The last GUS-derived nucleotide is shown in parentheses.

[c] Numbering starts with the first nucleotide of the HC-Pro coding sequence, which corresponds to TEV genome nucleotide 1057. The first HC-Pro-derived nucleotide is shown in parentheses.

[d] Number of plasmids sequenced after cloning of PCR products. See the text for details.

fragment corresponding to TEV nucleotides 1430 to 9461. The amount of radioactivity bound to each lane was measured with a Beckman LS 5801 Scintillation Counter.

Aphid transmission assays. Aphids (*Myzus persicae* Sulz) were reared and handled as previously described (4). For transmission directly from infected leaves, aphids were allowed a 5- to 10-min acquisition access period and then transferred to tobacco (*N. tabacum* cv. Ky 14) seedlings. Ten aphids were placed on each of 10 test plants for each treatment in each experiment. In helper component prefeeding experiments, aphids were first allowed a 5- to 10-min acquisition access period to a concentrated preparation of potato virus Y helper component through a Parafilm membrane as detailed elsewhere (6). Aphids were then transferred to a TEV-infected leaf and processed as described above. Infection of test plants was scored visually.

RESULTS

Spontaneous deletions of GUS and HC-Pro sequences in TEV-GUS. As demonstrated previously, the transfer of TEV-GUS from plant to plant at intervals of 4 to 6 days resulted in the stable retention of the GUS insert over seven passages (18). Continuation of this experiment up to 21 passages for approximately 120 days further demonstrated the stability of the foreign insert under these conditions. Immunoblot analysis of total protein from the 21st passage plant revealed the full-size, GUS–HC-Pro fusion product (Fig. 2). The level of GUS activity in this plant was comparable to that in plants from the early passages (data not shown). In contrast, prolonged propagation (3 to 4 weeks) of virtually all TEV-GUS-infected plants led to deletion mutants encoding HC-Pro-related proteins that were considerably smaller than GUS–HC-Pro (Fig. 2, lanes 1 to 12). Several of these variants, all of which were found in systemically infected leaves, expressed HC-Pro forms that were smaller than the wild-type protein (52 kDa), indicating that these mutants contained deletions of GUS and HC-Pro

sequences. In some cases, apparent intermediate deletion mutants that, upon further passage, were replaced by variants containing larger deletions (Fig. 2; compare lanes 3 and 4) were identified, an observation consistent with previous results (18). In control plants infected for 5 weeks by wild-type TEV-HAT, no deletions of HC-Pro were evident (Fig. 2), suggesting that the deletion phenomenon was specific to TEV carrying the foreign gene.

Mapping of TEV-GUS deletion endpoints. Virions and corresponding RNA were purified from six individual plants infected by TEV-GUS deletion variants expressing HC-Pro-related proteins of various sizes. The regions overlapping the deletion sites were copied by reverse transcription, amplified by PCR, cloned in a plasmid, and sequenced. Analysis of 40 cDNA clones revealed nine different deletion variants (Table 1). The deletions ranged from 1,741 to 2,074 nucleotides in length, showing almost complete loss of the GUS gene in all mutants. Each mutant, however, retained a short segment of between 9 and 135 nucleotides from the 5' end of the GUS sequence. All but one of the variants also was lacking between 15 and 265 nucleotides from the 5'-terminal region of the HC-Pro coding sequence. Interestingly, in no cases were the 5' and 3' deletion endpoints entirely within the GUS coding region. All deletions that were mapped resulted in preservation of the open reading frame and in some examples led to the formation of new hybrid codons at the junction sites. Two of the six plants analyzed contained multiple deletion mutants, although in both instances one type predominated among the population recovered (Table 1). No mutants contained multiple deletions, nor were non-GUS or non-HC-Pro sequences located at the junctions, suggesting that the deletions were generated by the removal of contiguous genome segments.

The relative positions of the 5' and 3' deletion endpoints are presented diagrammatically in Fig. 3. It is striking that all of the 5' endpoints cluster within a 130-nucleotide region of the GUS gene, whereas the 3' endpoints all map to the first 265 nucleotides of the HC-Pro sequence. The alignment of sequences flanking the endpoints of deletions failed to reveal any common elements of primary structure (data not
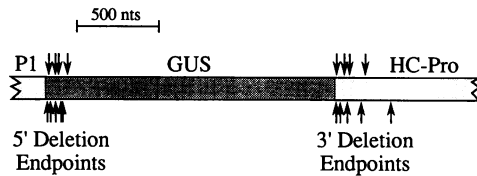
FIG. 3. Relative positions of deletion endpoints within the GUS and HC-Pro coding sequences in spontaneous mutants of TEV-GUS. Each arrow indicates a unique endpoint identified by sequence analysis.
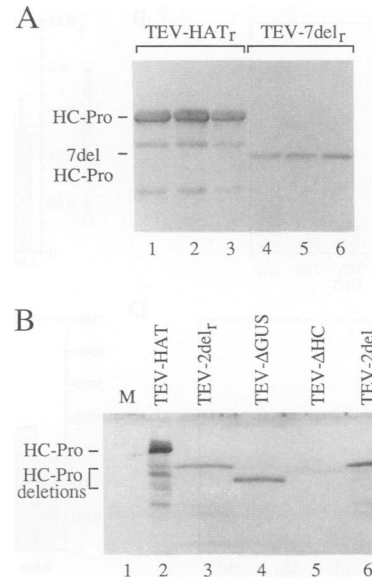

FIG. 4. Immunoblot analysis with anti-HC-Pro sera of extracts from plants inoculated with synthetic RNA transcripts. (A) Extracts were from systemically infected leaves of three plants inoculated with transcripts from pTEV7DA-HAT$_r$ (lanes 1 to 3) and three plants inoculated with transcripts from pTEV7DA-7del$_r$ (lanes 4 to 6). (B) Extracts were from a mock-inoculated plant (M) (lane 1), plants inoculated with virus preparations of wild-type TEV-HAT (lane 2) and TEV-2del (lane 6), or synthetic RNA transcripts from pTEV7DA-2del$_r$ (lane 3), pTEV7DA-ΔGUS (lane 4), and pTEV7DA-ΔHC (lane 5). Each lane was loaded with an equivalent amount of extract. The positions of HC-Pro and truncated HC-Pro derivatives are shown at the left of each panel.

shown). In fact, the extent of homology of sequences surrounding the junction sites did not exceed 2 nucleotides in any of the mutants. Similarly, no nucleotide preferences at the deletion endpoints were evident (Table 1).

Replication of virus derived from DNA copies of TEV genomes containing the largest HC-Pro deletions. The fact that the deletion mutants were able to spread systemically suggested that the HC-Pro region lacking in the mutants was not essential for virus replication and intraplant movement. Since it was possible that these plants also contained a low level of nondeleted TEV-GUS providing functions necessary for these processes, uniform virus populations were prepared for the two variants (TEV-2del and TEV-7del) containing the largest HC-Pro deletions. An ~350-nucleotide segment covering the deletion site in both genomes was amplified by PCR and inserted into pTEV7DA, generating pTEV7DA-2del$_r$ and pTEV7DA-7del$_r$. The resulting constructs harbored the deletion in a wild-type TEV-HAT background and lacked any other sequence changes potentially present in the original mutants. As a control, the corresponding genome segment from wild-type TEV-HAT was amplified and cloned in pTEV7DA to produce pTEV7DA$_r$.

Between 60 and 100% of plants inoculated with synthetic transcripts from pTEV7DA$_r$, pTEV7DA-2del$_r$, and pTEV7DA-7del$_r$ became infected systemically, demonstrating conclusively that the mutants were capable of replication and movement independent of a helper virus. The appearance of systemic symptoms at 4 days postinoculation was observed with plants infected by TEV-HAT and the deletion mutants. The appearance of systemic symptoms in plants inoculated with TEV-GUS, on the other hand, required 5 postinoculation days. Immunoblot analysis using anti-HC-Pro serum indicated that the two reconstructed deletion variants encoded truncated HC-Pro-related proteins, whereas the reconstructed wild-type virus expressed a normal-size product (Fig. 4, lanes 3).

Although the N-terminal HC-Pro sequences clearly were not essential for virus viability, it was noticed that the levels of HC-Pro appeared to be considerably less in plants infected by the mutants than in those infected by wild-type TEV-HAT (Fig. 4). Also, the level of GUS-HC-Pro fusion protein in TEV-GUS-infected plants appeared less than that of HC-Pro in plants infected by TEV-HAT (Fig. 2). To determine more precisely the levels of replication of wild-type and mutant viruses in systemically infected tissue, the relative amounts of capsid protein and virus RNA were measured by immunoblot and Northern blot analyses, respectively.

The insertion of GUS into the viral genome in TEV-GUS decreased the yields of capsid protein and RNA to levels of 15 and 9%, respectively, of those in the wild type (Fig. 5A and C). The presence of GUS, therefore, had a debilitating

effect on the accumulation of virus. TEV-7del, which specified 3 amino acid residues from GUS but possessed an HC-Pro deletion to residue 89, yielded even less viral protein and RNA, at 8 and 6% of the wild-type levels, respectively (Fig. 5A and C). In contrast, plants infected by TEV-2del, which coded for 45 residues of GUS but lacked 66 HC-Pro residues, accumulated capsid protein and RNA at 25 and 17% of the wild-type levels, respectively (Fig. 5A and C). TEV-2del, therefore, replicated to a degree approximately three times higher than that of TEV-7del.

The differences between the two spontaneous mutants may have been due to the presence of the short GUS sequence having a stabilizing effect in TEV-2del or to the larger deletion within TEV-7del affecting a critical function of HC-Pro. To distinguish between these two possibilities, two site-directed mutations were introduced into pTEV7DA-2del$_r$. In the first, pTEV7DA-ΔGUS, the GUS sequence between nucleotides 10 and 135 was deleted, leaving a short GUS region equivalent to that present in TEV-7del. In the second, pTEV7DA-ΔHC, the HC-Pro deletion was extended from nucleotides 207 to 265, corresponding to the sequence missing in TEV-7del (Fig. 1B). Immunoblot analysis with HC-Pro-specific antibodies revealed the truncated HC-Pro-related products of the expected sizes in plants inoculated with transcripts from the two mutants (Fig. 4B, lanes 4 and 5). On the basis of the apparent quantities of HC-Pro deletion products in these plants, TEV-ΔGUS replicated better than TEV-2del$_r$ and TEV-ΔHC. This was confirmed by quantitation of the levels of capsid protein and RNA in systemic leaves. Compared with TEV-2del$_r$, TEV-ΔGUS
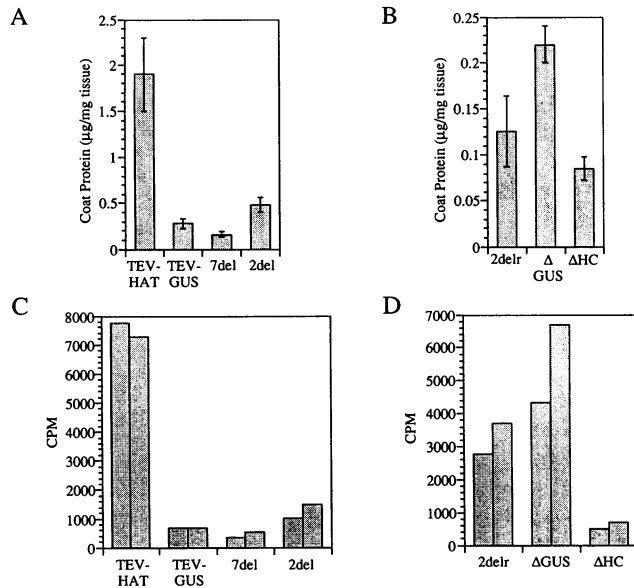
A

B

C

D

FIG. 5. Quantitation of TEV capsid protein (A and B) and RNA (C and D) present in systemically infected leaves of plants inoculated with wild-type and mutant viruses. (A and B) Capsid protein levels in total protein extracts were determined by reflective densitometry of immunoblots with capsid protein standards from purified virions. Each bar represents the average (and standard deviation) of three independent samples. (C and D) Relative virus RNA levels in total RNA extracts were determined by measuring radioactivity from a ^{32}P-labeled probe bound to Northern blots. Two independent samples are shown in each case. For comparative purposes, note that data shown in panels A and C were collected from contemporaneous samples in one experiment, while data in panels B and D were from another set of contemporaneous samples.

directed an average of approximately 70% more capsid protein and RNA, while TEV-ΔHC directed less of these products (Fig. 5B and D). This demonstrated that the presence of the GUS fragment and the absence of additional HC-Pro sequence both conferred a debilitating effect on virus accumulation.

Transmission of TEV mutants by aphids. The abilities of aphids to transmit TEV-HAT, TEV-GUS, TEV-2del, and TEV-7del were tested. Transmission assays demonstrated that only TEV-HAT possessed an aphid-transmissible phenotype (Table 2). The lack of transmission of TEV-GUS, TEV-2del, and TEV-7del could have been due to functional defects in the mutant HC-Pro proteins and/or to the decrease in mutant virus titers and corresponding low concentrations

TABLE 2. Aphid transmission of wild-type and mutant TEV variants

Virus	No. of infected plants/total no. of plants	
	Without HC-Pro prefeeding	With HC-Pro prefeeding[a]
TEV-HAT	34/40	10/10
TEV-GUS	0/40	0/20
TEV-2del	0/30	4/30
TEV-7del	0/30	0/20

[a] Aphids were allowed access to active HC-Pro from potato virus Y prior to transmission assays.

of HC-Pro (3). As shown previously, prefeeding of aphids on concentrated preparations of HC-Pro can facilitate the transmission of mutant viruses expressing defective HC-Pro proteins (3). With highly active preparations of potato virus Y HC-Pro, aphid transmission of TEV-2del was restored, although at low efficiency (Table 2). The inability to transmit TEV-GUS or TEV-7del with the aid of prefeeding was due most likely to low levels of virus accumulation (Fig. 5A and C).

DISCUSSION

Possible mechanism of generation of the spontaneous deletion mutants. The sequence analysis of several deletion mutants demonstrated a random distribution of deletion endpoints inside relatively short segments near the 5' ends of the GUS and HC-Pro coding regions (Table 1 and Fig. 3). Inspection of the sequences of TEV-GUS RNA in and around the deleted regions and computer-assisted analysis of putative secondary structure did not reveal any common structural determinants (data not shown). Considering these observations, and by analogy with other positive-strand RNA viruses (24, 32), we propose illegitimate or nonhomologous recombination as the likely deletion mechanism.

Can a nonhomologous recombination mechanism explain the clustering of 5' and 3' deletion endpoints observed in this study? It is clear that in single-stranded RNA viruses this mechanism involves template switching of RNA polymerase, or copy choice, during RNA synthesis and that for poliovirus this occurs during the production of minus-strand RNA (24, 28, 32). The apparent bias for deletion endpoints within short stretches of sequence may be due simply to random dissociation or reassociation of template and replicase (bound to the nascent strand) followed by selection of viable and/or advantageous variants. Such a mechanism has been also proposed by others on the basis of studies of recombination with coronaviruses (5) and pestiviruses (35, 36). The recombinants selected would be those that have lost only elements nonessential for replication and movement. Given that no mutants with deletion endpoints upstream of the GUS gene or downstream of codon 89 in HC-Pro were identified, the P1 protein and the central and/or C-terminal segment of HC-Pro may be indispensable for virus viability in plants. Alternatively, the clustering of deletion endpoints may be due to the increased dissociation frequency of virus RNA polymerase during transcription of the foreign sequence, perhaps depending on the presence of unusual primary or secondary structure. Reassociation of the replicative complex may occur preferentially on authentic TEV sequences, explaining our inability to find 3' deletion endpoints within the GUS gene. The credibility of this explanation suffers, however, in that it requires template dissociation or reassociation during positive-strand synthesis, which contradicts the established recombination mechanism of poliovirus.

Curiously, tobacco vein mottling potyvirus (TVMV) mutants lacking 45 or 225 nucleotides from the 5' end of the HC-Pro sequence were nonviable (41a). These TVMV variants were generated by site-directed deletion techniques rather than by spontaneous mutagenesis. The differences in viability for similar types of TEV and TVMV mutants may be due to actual differences in the roles of HC-Pro during virus replication. On the other hand, the replicative abilities of the TEV mutants may reflect the selection of rare tolerable variants from a large pool of predominantly inactive recombinant genomes.

Given that the presence of GUS has a negative effect on virus accumulation (Fig. 5), deletion mutants lacking the GUS sequence should have a replicative advantage over parental TEV-GUS. The inhibitory influence of GUS may be a consequence of several factors, including interference with HC-Pro function, the increased length of the viral genome, and the introduction of unusual sequence elements that are not favored by the viral polymerase. Some deletion variants may be selected even though they replicate more poorly than the parental virus. For TEV-7del, which accumulated to a level lower than that of TEV-GUS in systemic tissue, the loss of GUS may have increased the rate of virus movement, an idea supported by the more rapid appearance of virus-induced symptoms on systemic leaves.

No deletions were observed in TEV-GUS during 21 rapid plant-to-plant transfers when passages were made every 4 to 6 days (Fig. 2). In contrast, plants infected for 4 to 5 weeks with TEV-GUS consistently spawned deletion variants (Fig. 2) (18). Over a short infection period, the probability of the appearance of advantageous deletion variants may be low. Furthermore, new competitive variants would have little time to multiply and increase within the replicating pool. The lack of deletion mutants appearing in each subsequent transfer may be explained by a genetic bottleneck effect in which relatively few infectious units are passed from one plant to another, thereby excluding the transfer of low-titer variants. Generation and domination of deletion mutants after prolonged periods may reflect an increase in the virus population in the whole plant and therefore an increased probability for recombination.

The instability of TEV-GUS after prolonged infection periods also has relevance for the use of TEV as an RNA-based expression vector, a proposition made previously. The deletion of large foreign genes from several recombinant RNA plant virus genomes has been documented previously (13, 19, 44). The utility of TEV as a short-term (>21-day), high-level expression vehicle may be high, while its use as a vector during the entire vegetative life of a host plant is doubtful.

Functions of the N-terminal domain of HC-Pro. Analyses of sequence conservation between HC-Pro of potyviruses; HC-Pro-related protein of barley yellow mosaic bymovirus, a bipartite relative of the potyviruses; and p29 of HAV, the hypovirulence-associated virus-like agent of *C. parasitica*, revealed three potential domains. These include a cysteine-rich, N-terminal domain present in the potyviral and HAV proteins, a central domain specific for potyviruses, and a C-terminal proteinase domain conserved among all three groups (16, 26, 30, 46). The sequence deleted from TEV-2del and TEV-7del composed the N-terminal domain, including the cysteine-rich motif $CX_8CX_{13}CX_4CX_2C$ (Fig. 6). The facts that these mutants could replicate, move from cell to cell, move systemically, and assemble into virions indicate that the N-terminal domain of HC-Pro is not a factor essential for these processes. This region is required, on the other hand, for aphid-mediated transmission of potyviruses. Analysis of naturally occurring variants and in vitro-generated mutants has shown that lysine at position 54, which is within the cysteine-rich motif, is critical for helper component function (3, 41, 46). The lack of aphid transmissibility of TEV-2del and the partial restoration of transmission by prefeeding of aphids on active HC-Pro from a heterologous potyvirus (Table 2) further confirm the helper activity of the N-terminal domain.

It is a good possibility that the N-terminal region of HC-Pro provides another function besides helper component
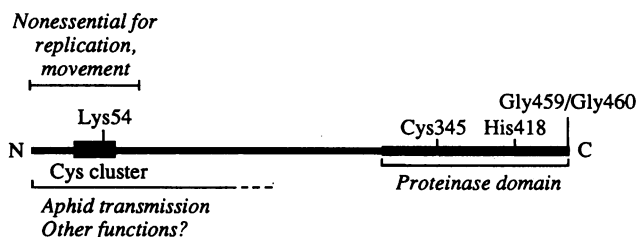

FIG. 6. Schematic representation of functional domains in HC-Pro. Given that viable virus with HC-Pro deletions up to amino acid residue 89 could be recovered, the N-terminal region indicated above the map is considered nonessential for replication and movement. The lysine residue at position 54 is required for helper component activity during aphid-mediated transmission, according to studies of the related potyvirus TVMV (3). The cysteine-rich motif, which is highly conserved among potyviruses, is indicated by the black box on the left. Essential residues in the proteolytic domain, including the catalytic residues Cys-345 and His-418 (39) and the residues around the cleavage site, Gly-459 and Gly-460 (8), are shown. Numbering of amino acid residues is from serine at position 1 in HC-Pro (34, 47).

activity. Although both TEV-2del and TEV-7del were viable in plants, their yields of virus RNA and protein were significantly reduced in comparison to that of TEV-HAT, suggesting a potential accessory function of HC-Pro in viral replication. A similar reduction of virus accumulation was observed when either of two point mutations affecting residues within the N-terminal domain was introduced into the TVMV genome (3). We cannot exclude, however, the possibility that the TEV deletion and TVMV point mutations disrupted or destabilized a *cis*-acting RNA element important for replication.

Interestingly, several other unrelated viruses, including barley stripe mosaic virus and tobacco rattle virus, encode small, cysteine-rich proteins that are nonessential to virus replication, cell-to-cell movement, and encapsidation in plants (22, 40). Further, the cysteine-rich domain in p29 of HAV is dispensable for replication (15). In the barley stripe mosaic virus example, deletion of part of the γb gene, which codes for a 17-kDa cysteine-rich protein, had a debilitating effect on virus protein accumulation, similar to deletion of the HC-Pro N-terminal sequence. A regulatory role involving RNA binding has been proposed as a potential function for barley stripe mosaic virus γb protein, as well as for cysteine-rich proteins of several other plant positive-strand RNA viruses (23, 25, 29, 37). The ability of HC-Pro to bind RNA is not well-characterized, although HC-Pro has affinity for several polyribonucleotides (45). Note that, regardless of their specific biochemical functions, the cysteine-rich proteins are encoded by plant viruses with relatively large genomes (>8,000 nucleotides). Through the capture of accessory protein genes such as these during virus evolution, additional levels of regulation may have been achieved along with the potential to acquire novel viral functions like vector transmissibility (17).

REFERENCES

1. **Ahlquist, P., and R. F. Pacha.** 1990. Gene amplification and expression by RNA viruses and the potential for further application to plant gene transfer. Physiol. Plant. **79:**163–167.
2. **Allison, R., R. E. Johnston, and W. G. Dougherty.** 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology **154:**9–20.
3. **Atreya, C. D., P. L. Atreya, D. W. Thornbury, and T. P. Pirone.** 1992. Site-directed mutations in the potyvirus HC-PRO gene affect helper component activity, virus accumulation, and symptom expression in infected tobacco plants. Virology **191:**106–111.
4. **Atreya, P. L., C. D. Atreya, and T. P. Pirone.** 1991. Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. Proc. Natl. Acad. Sci. USA **88:**7887–7891.
5. **Banner, L. R., and M. M. C. Lai.** 1991. Random nature of coronavirus RNA recombination in the absence of selection pressure. Virology **185:**441–445.
6. **Berger, P. H., A. G. Hunt, L. L. Domier, G. M. Hellmann, Y. Stram, D. W. Thornbury, and T. P. Pirone.** 1989. Expression in transgenic plants of a viral gene product that mediates insect transmission of potyviruses. Proc. Natl. Acad. Sci. USA **86:**8402–8406.
7. **Bredenbeek, P. J., and C. M. Rice.** 1992. Animal RNA virus expression systems. Semin. Virol. **3:**297–310.
8. **Carrington, J. C., S. M. Cary, T. D. Parks, and W. G. Dougherty.** 1989. A second proteinase encoded by a plant potyvirus genome. EMBO J. **8:**365–370.
9. **Carrington, J. C., and W. G. Dougherty.** 1987. Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. J. Virol. **61:**2540–2548.
10. **Carrington, J. C., and D. D. Freed.** 1990. Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. J. Virol. **64:**1590–1597.
11. **Carrington, J. C., D. D. Freed, and T. C. Sanders.** 1989. Autocatalytic processing of the potyvirus helper component proteinase in *Escherichia coli* and in vitro. J. Virol. **63:**4459–4463.
12. **Carrington, J. C., and T. J. Morris.** 1984. Complementary DNA cloning and analysis of carnation mottle virus RNA. Virology **139:**22–31.
13. **Chapman, S., T. Kavanagh, and D. Baulcombe.** 1992. Potato virus X as a vector for gene expression in plants. Plant J. **2:**549–557.
14. **Choi, G. H., D. M. Pawlyk, and D. L. Nuss.** 1991. The autocatalytic protease p29 encoded by a hypovirulence-associated virus of the chestnut blight fungus resembles the potyvirus-encoded protease HC-Pro. Virology **183:**747–752.
15. **Craven, M. G., D. M. Pawlyk, G. H. Choi, and D. L. Nuss.** Papain-like protease p29 as a symptom determinant encoded by a hypovirulence-associated virus of the chestnut blight fungus. Submitted for publication.
16. **Davidson, A. D., M. Prols, J. Schell, and H. Steinbiss.** 1991. The nucleotide sequence of RNA 2 of barley yellow mosaic virus. J. Gen. Virol. **72:**989–993.
17. **Dolja, V. V., and J. C. Carrington.** 1992. Evolution of positive-strand RNA viruses. Semin. Virol. **3:**315–326.
18. **Dolja, V. V., H. J. McBride, and J. C. Carrington.** 1992. Tagging of plant potyvirus replication and movement by insertion of β-glucuronidase into the viral polyprotein. Proc. Natl. Acad. Sci. USA **89:**10208–10212.
19. **Donson, J., C. M. Kearney, M. E. Hilf, and W. O. Dawson.** 1991. Systemic expression of a bacterial gene by a tobacco mosaic virus-based vector. Proc. Natl. Acad. Sci. USA **88:**7204–7208.
20. **Dougherty, W. G., and J. C. Carrington.** 1988. Expression and function of potyviral gene products. Annu. Rev. Phytopathol. **26:**123–143.
21. **Dougherty, W. G., and E. Hiebert.** 1980. Translation of potyvirus RNA in a rabbit reticulocyte lysate: reaction conditions and identification of capsid protein as one of the products of *in vitro* translation of tobacco etch and pepper mottle viral RNAs. Virology **101:**466–474.
22. **Guilford, P. J., V. Ziegler-Graff, and D. C. Baulcombe.** 1991. Mutation and replacement of the 16-kDa protein gene in RNA-1 of tobacco rattle virus. Virology **182:**607–614.
23. **Jackson, A. O., I. T. D. Petty, R. W. Jones, M. C. Edwards, and R. French.** 1991. Analysis of barley stripe mosaic virus pathogenicity. Semin. Virol. **2:**107–119.
24. **Jarvis, T. C., and K. Kirkegaard.** 1991. The polymerase in its labyrinth: mechanisms and implications of RNA recombination. Trends Genet. **7:**186–191.
25. **Jupin, I., T. Tamada, and K. Richards.** 1991. Pathogenesis of beet necrotic yellow vein virus. Semin. Virol. **2:**121–129.
26. **Kashiwazaki, S., Y. Minobe, and H. Hibino.** 1991. Nucleotide sequence of barley yellow mosaic virus RNA 2. J. Gen. Virol. **72:**995–999.
27. **Kawasaki, E. S.** 1990. Amplification of RNA, p. 21–27. *In* M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (ed.), PCR protocols. A guide to methods and applications. Academic Press, Inc., San Diego, Calif.
28. **Kirkegaard, K., and D. Baltimore.** 1986. The mechanism of RNA recombination in poliovirus. Cell **47:**433–443.
29. **Koonin, E. V., V. P. Boyko, and V. V. Dolja.** 1991. Small cysteine-rich proteins of different groups of plant RNA viruses are related to different families of nucleic acid-binding proteins. Virology **181:**395–398.
30. **Koonin, E. V., G. H. Choi, D. L. Nuss, R. Shapira, and J. C. Carrington.** 1991. Evidence for common ancestry of a chestnut blight hypovirulence-associated double-stranded RNA and a group of positive-strand RNA plant viruses. Proc. Natl. Acad. Sci. USA **88:**10647–10651.
31. **Kunkel, T. A., J. D. Roberts, and R. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
32. **Lai, M. M. C.** 1992. RNA recombination in animal and plant viruses. Microbiol. Rev. **56:**61–79.
33. **Matthews, R. E. F.** 1991. Plant virology. Academic Press, Inc., San Diego, Calif.
34. **Mavankal, G., and R. E. Rhoads.** 1991. In vitro cleavage at or near the N-terminus of the helper component protein in the tobacco vein mottling virus polyprotein. Virology **185:**721–731.
35. **Meyers, G., N. Tautz, E. J. Dubovi, and H.-J. Thiel.** 1991. Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. Virology **180:**602–616.
36. **Meyers, G., N. Tautz, R. Stark, J. Brownlie, E. J. Dubovi, M. S. Collet, and H. J. Thiel.** 1992. Rearrangement of viral sequences in cytopathogenic pestiviruses. Virology **191:**368–386.
37. **Niesbach-Klösgen, U., H. Guilley, G. Jonard, and K. Richards.** 1990. Immunodetection in vivo of beet necrotic yellow vein virus encoded proteins. Virology **178:**52–61.
38. **Nuss, D. L.** 1992. Biological control of chestnut blight: an example of virus-mediated attenuation of fungal pathogenesis. Microbiol. Rev. **56:**561–576.
39. **Oh, C.-S., and J. C. Carrington.** 1989. Identification of essential residues in potyvirus proteinase HC-Pro by site-directed mutagenesis. Virology **173:**692–699.
40. **Petty, I. T. D., R. French, R. W. Jones, and A. O. Jackson.** 1990. Identification of barley stripe mosaic virus genes involved in viral RNA replication and systemic movement. EMBO J. **9:**3453–3457.
41. **Pirone, T. P.** 1991. Viral genes and gene products that determine insect transmissibility. Semin. Virol. **2:**81–87.
41a.**Pirone, T. P.** Unpublished observations.
42. **Restrepo, M. A., D. D. Freed, and J. C. Carrington.** 1990. Nuclear transport of plant potyviral proteins. Plant Cell **2:**987–998.
43. **Riechmann, J. L., S. Laín, and J. A. García.** 1992. Highlights and prospects of potyvirus molecular biology. J. Gen. Virol. **73:**1–16.
44. **Scholthof, H. B., T. J. Morris, and A. O. Jackson.** 1993. The

capsid protein gene of tomato bushy stunt virus is dispensable for systemic movement and can be replaced for localized expression of foreign genes. Mol. Plant-Microbe Interact. 6:309–322.

45. **Thornbury, D. W., G. M. Hellmann, R. E. Rhoads, and T. P. Pirone.** 1985. Purification and characterization of potyvirus helper component. Virology **144:**260–267.

46. **Thornbury, D. W., C. A. Patterson, J. T. Dessens, and T. P. Pirone.** 1990. Comparative sequences of the helper component (HC) region of potato virus Y and an HC-defective strain,

potato virus C. Virology **178:**573–578.

46a.**Verchot, J., and J. C. Carrington.** Unpublished data.

47. **Verchot, J., K. L. Herndon, and J. C. Carrington.** 1992. Mutational analysis of the tobacco etch potyviral 35-kDa proteinase: identification of essential residues and requirements for autoproteolysis. Virology **190:**298–306.

48. **Verchot, J., E. V. Koonin, and J. C. Carrington.** 1991. The 35-kDa protein from the N-terminus of a potyviral polyprotein functions as a third virus-encoded proteinase. Virology **185:**527–535.