RNA (1995), *1*:491-500. Cambridge University Press. Printed in the USA.

Mutation analyses of molecularly cloned satellite tobacco mosaic virus during serial passage in plants: Evidence for hotspots of genetic change

GAEL KURATH[1] and J. ALLAN DODDS[2]

[1] National Biological Service, Northwest Biological Science Center, Seattle, Washington 98115, USA
[2] Department of Plant Pathology, University of California, Riverside, California 92521-0122 USA

ABSTRACT

The high level of genetic diversity and rapid evolution of viral RNA genomes are well documented, but few studies have characterized the rate and nature of ongoing genetic change over time under controlled experimental conditions, especially in plant hosts. The RNA genome of satellite tobacco mosaic virus (STMV) was used as an effective model for such studies because of advantageous features of its genome structure and because the extant genetic heterogeneity of STMV has been characterized previously. In the present study, the process of genetic change over time was studied by monitoring multiple serial passage lines of STMV populations for changes in their consensus sequences. A total of 42 passage lines were initiated by inoculation of tobacco plants with a helper tobamovirus and one of four STMV RNA inocula that were transcribed from full-length infectious STMV clones or extracted from purified STMV type strain virions. Ten serial passages were carried out for each line and the consensus genotypes of progeny STMV populations were assessed for genetic change by RNase protection analyses of the entire 1,059-nt STMV genome. Three different types of genetic change were observed, including the fixation of novel mutations in 9 of 42 lines, mutation at the major heterogeneity site near nt 751 in 5 of the 19 lines inoculated with a single genotype, and selection of a single major genotype in 6 of the 23 lines inoculated with mixed genotypes. Sequence analyses showed that the majority of mutations were single base substitutions. The distribution of mutation sites included three clusters in which mutations occurred at or very near the same site, suggesting hot spots of genetic change in the STMV genome. The diversity of genetic changes in sibling lines is clear evidence for the important role of chance and random sampling events in the process of genetic diversification of STMV virus populations.

Keywords: mutation; RNA genome; RNase protection analyses; satellite tobacco mosaic virus; serial passage

INTRODUCTION

The concept that RNA genome populations contain extremely high levels of genetic heterogeneity and evolve approximately 1,000-fold faster than DNA-based populations has been extensively studied and reviewed (Domingo et al., 1985; Steinhauer & Holland, 1987; Domingo & Holland, 1988, 1994, and references therein). Both the heterogeneity and rapid evolution are thought to be due to the replication of these populations by RNA polymerases, which do not have the elaborate proofreading capabilities associated with DNA polymerases. Thus, RNA virus populations exist as distributions of variants, all related to a consensus sequence that represents the average for that population (quasispecies). Although there have been numerous studies documenting these phenomena, nearly all have involved RNA viruses of animal hosts, and there are relatively few characterizing the actual generation of heterogeneity under controlled experimental conditions. In order to do so in a plant host system, the RNA genome of satellite tobacco mosaic virus (STMV) was chosen as a model subject due to several advantageous features explained below.

STMV is a plant satellite virus due to its dependence for replication on co-infection with a helper tobamovirus, tobacco mild green mosaic virus (TMGMV) (Valverde & Dodds, 1986, 1987). The genome of STMV is

Reprint requests to: Gael Kurath, National Biological Service, Northwest Biological Science Center, 6505 NE 65th Street, Seattle, Washington 98115 USA; e-mail: Gael-Kurath@nbs.gov.

a single plus-sense RNA molecule of 1,059 nt. It is comprised of a coding region from base 53 to base 642 that contains two overlapping open reading frames (ORFs), and a long 3′ untranslated region extending from base 643 to base 1059 (Mirkov et al., 1989). Thus, the small size of the RNA facilitates assessment of the entire genome for mutations rather than being limited to a partial region, as is often the case with larger viruses, and the genome structure means that results with this RNA will provide information on the process of genetic change in both coding and noncoding genomic RNA regions. In addition, full-length infectious clones are available to initiate serial passages with an RNA inoculum that is nearly genetically homogeneous (Kurath et al., 1992).

The quasispecies nature of STMV has been demonstrated in a detailed study of the genetic heterogeneity within a type strain population (Kurath et al., 1992). It was found that the type strain contained many sequence microheterogeneity variants and two major genotypes, designated type 5 (T5) and type 6 (T6), that can be represented experimentally by RNA transcribed from full-length infectious STMV clones pSTMV5 and pSTMV6, respectively. T5 and T6 genotypes differ by a single base substitution at nt 751, which is easily detectable by RNase protection analyses (see below). In a further study, a survey of the genetic heterogeneity between the consensus genotypes of 15 STMV field isolates also showed great diversity, including many T5 and T6 genotypes (Kurath et al., 1993a), suggesting the site at or near nt 751 as a major heterogeneity site in natural STMV populations.

In these earlier studies, the use of RNase protection analyses (Meyers et al., 1985; Winter et al., 1985) as a method to follow genetic change in STMV RNA populations was clearly established and very fruitful, and it was therefore also used as a major tool in the present study. The goal of the RNase protection methodology is to detect and map minor sequence differences, including single base substitutions, between closely related RNA populations. As a tool for these analyses a radioactive, minus-sense STMV RNA is transcribed from a clone of the STMV genome. This is used as a hybridization probe that is annealed to various target RNAs, which are RNAs transcribed from clones or genomic RNA of STMV populations. The probe and target RNAs are annealed and the resulting RNA:RNA duplexes are then incubated with RNases A and T1 under conditions where sites of sequence mismatch can be cleaved. Thus, if the target RNA has mismatches relative to the probe, the probe will often be cleaved at those sites. If there are no mismatches, the probe will be protected intact. The cleavage reaction products are denatured, separated by electrophoresis, and the probe fragments are visualized by autoradiography.

The sizes of the probe cleavage products are used to map the sites of sequence difference between the probe and the target RNA. It is well known that RNase protection analyses do not detect and cleave all mismatches to completion, but this phenomenon has been experimentally characterized (Meyers et al., 1985; Winter et al., 1985) and does not detract from the utility of the method as long as it is taken into consideration. RNase protection analyses do detect many single base differences and are very reproducible, making them especially useful in comparative studies such as this one, where progeny are compared with parental RNAs of a known RNase protection pattern. The major advantage to RNase protection analyses is the large number of samples that can be assayed. Forty target RNAs can be assessed simultaneously for sequence variability, allowing much more expansive studies to be undertaken than would be possible if RNA sequencing were the method of identifying differences. The work reported here involved analyses of at least 650 RNA populations by RNase protection. Once populations with known genetic differences were identified and mapped, the amount of RNA sequencing needed to characterize the mutations was feasible.

The ability of STMV populations to evolve rapidly in response to altered selection pressure has been demonstrated and characterized at the genetic level in experimental co-infections with helper viruses other than TMGMV (Kurath et al., 1993b). In the current study, we address the process of genetic change in STMV populations over time by analyzing the progeny of STMV replication in multiple tobacco passage lines. The conditions for these passages were the standard conditions used to propagate the STMV type strain for several years, so that these experiments, as much as possible, assessed genetic change in the absence of any shift in selection pressure. Passage lines inoculated with STMV RNA transcribed from infectious clones provided information about both mutation and subsequent competition between variants. Lines inoculated with purified type strain STMV RNA were included to see if the same phenomena occurred in naturally heterogeneous populations.

Our overall goal was to observe and characterize the process of genetic change, i.e., mutation and selection, with regard to frequency, timing, sites, and individual mutation types. There is no basis for assuming that these aspects of genetic change are random, but in the absence of experimental data for most systems, there is very little detail to our understanding of how the process occurs. With this study we attempt to generate a picture of the nature of the genetic change process in RNA genome populations in plant hosts, using the STMV genome as a model. Because genetic change is the fundamental basis of evolution, this is a first step toward understanding in molecular detail how an RNA genome population evolves in plants. Future studies with other viruses will be necessary to determine whether the evolutionary phenomena described here

for STMV are characteristic of evolution of other plant RNA viruses or of RNA viruses in general.

RESULTS

STMV serial passage lines

In a preliminary study to see if serial passage of STMV would generate detectable genetic changes in an experimentally feasible time frame, a small number of passage lines were originally inoculated with TMGMV helper virus and RNA transcribed from infectious STMV cDNA clones. This study included two independent lines inoculated with RNA transcribed from the type 5 clone, pSTMV5 (hereafter called T5 RNA), two lines with RNA transcribed from the type 6 clone, pSTMV6 (hereafter called T6 RNA), and three lines in which T5 and T6 RNAs were inoculated in an equimolar mixture (hereafter called T56 RNA). The passage lines from this study were denoted with a prime, and were designated T5-A′ and T5-B′; T6-C′ and T6-D′; and T56-E′, T56-F′, and T56-G′, respectively. STMV progeny populations from these lines were purified from total systemically infected leaf tissue and assessed for genetic changes by RNase protection assays after passages 1, 2, 3, 4, 5, 7, and 10. The results of these assays, which will be described below, showed that genetic changes had occurred in six of the seven passage lines, but that sibling lines did not show the same changes. Thus, it became clear that changes could easily be detected within a reasonable number of passages, but larger numbers of passage lines would be necessary to see trends or to assess frequencies of various kinds of changes.

The study was therefore expanded by initiating several more passage lines so that there were approximately 10 sibling lines initiated with each of the inocula described above. For comparison with passage lines of molecularly cloned STMV inocula, 12 lines were also initiated with genomic RNA from the STMV type strain (wt RNA) (Kurath et al., 1992). The new passage lines had no prime in the designations, and were T5-A through T5-H, T6-J through T6-P, T56-Q through T56-X, and wt-A through wt-L. The complete study thus involved 42 independent passage lines of STMV RNA (Table 1). In order to handle the larger number of samples, the new lines were passaged as pooled sap rather than purified virions and progeny populations were assessed by RNase protection assays using total nucleic acid extracts from pooled leaf tissues after passages 1, 4, 7, and 10.

Detection of genetic changes in STMV progeny

Analyses of progeny STMV populations by RNase protection assays were done using two overlapping, minus-sense probes that together assessed the entire length of any target STMV genome for genetic change (Kurath et al., 1992). Probe A was complementary to bases 602–1059 at the 3′ end, and probe B was complementary to bases 1–791 at the 5′ end of the STMV sequence in clone pSTMV6, as shown in Figure 1. When these two probes were used in assays with STMV progeny populations from each passage line, genetic changes were detected at various times, most often between the first and seventh passages, and the changes generally remained stable from the seventh through the tenth passages.

RNase protection patterns of the 10th passage progeny from each line are shown in Figures 2 and 3. In these figures, groups of sibling passage lines assayed with a single probe are shown on individual panels. Immediately adjacent to the marker lane on the left of each panel is a control assay of the original RNA inoculum for the group of passage lines in that panel. The control lanes labeled "5" in panels A and B of Figure 2 show the characteristic patterns of type 5 STMV RNA, in which probe A is cleaved into two major fragments approximately 300 and 144 bases long, and probe B is cleaved to produce a small fragment of approximately 24 bases and a long fragment indistinguishable on

TABLE 1. Detection of mutations in STMV serial passage lines by RNase protection analyses.[a]

STMV RNA inoculum[b]	# Passage lines	Non-MHS mutations	Conversion at MHS[c]	Selection at MHS[d]
T5	10	2/10	3/10	NA
T6	9	1/9	2/9	NA
T56	11	3/11	NA	4/11
wt	12	3/12	NA	2/12
Totals	42	9/42	5/19	6/23

[a] Ratios in the last three columns indicate the number of passage lines with the indicated change/total number of lines in each group. NA indicates not applicable.

[b] Tobacco plants were inoculated with TMGMV helper virus and STMV inocula as follows: T5, RNA transcribed from type 5 clone pSTMV5; T6, RNA transcribed from type 6 clone pSTMV6; T56, an equimolar mixture of T5 and T6 RNAs; wt, genomic RNA from STMV type strain known to be a natural mixture of T5 and T6 genotypes.

[c] Conversion between T5 and T6 genome types at the major heterogeneity site (MHS) of STMV (Kurath et al., 1992) in lines inoculated with a single type.

[d] Selection at the MHS of STMV of a single type, T5 or T6, in passage lines inoculated with a mixture of T5 and T6 genomes.

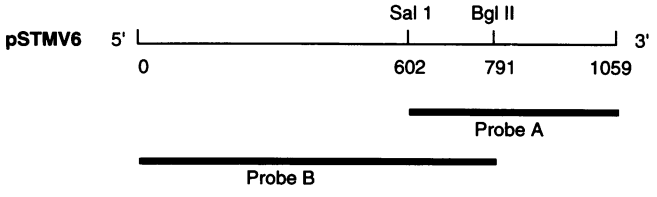

FIGURE 1. RNase protection assay probes produced by transcription of portions of the STMV genome sequence in cDNA clone pSTMV6.

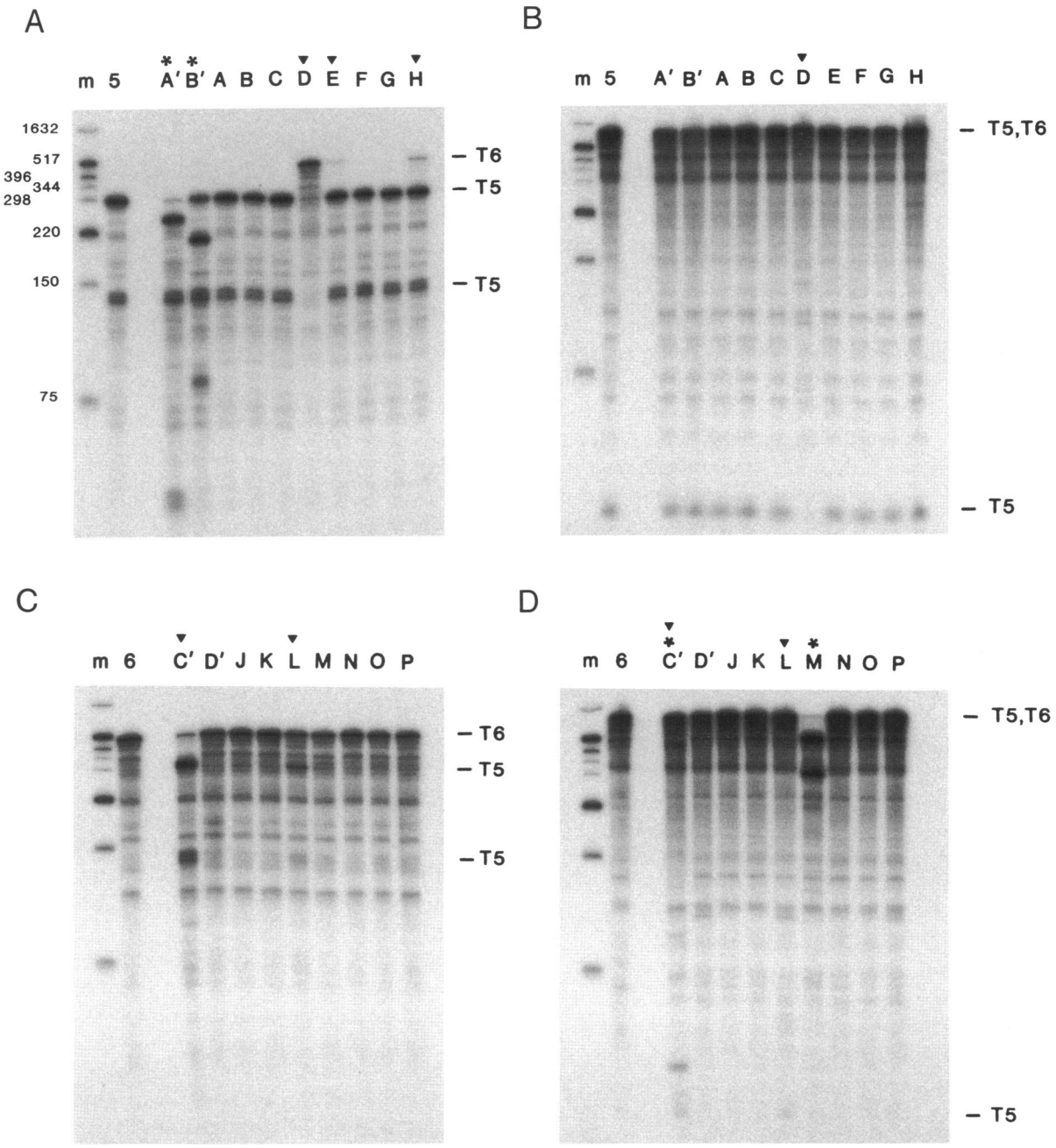

FIGURE 2. RNase protection assay patterns of progeny populations of single inoculum STMV passage lines after 10 passages. Progeny of T5 passage lines are shown in **A** (probe A) and **B** (probe B), and progeny of T6 passage lines are shown in **C** (probe A) and **D** (probe B). Individual passage line designations are shown as upper case letters above the lanes, control assays of T5 and T6 RNA inocula are marked 5 and 6, respectively, and m indicates DNA marker fragments with sizes (in nucleotides) as shown to the left of A. Fragments diagnostic for T5 and T6 patterns are noted to the right of each panel. *, passage lines in which a novel mutation was detected; ▼, lines showing conversion between T5 and T6 patterns at the STMV major heterogeneity site (Kurath et al., 1992).

these gels from the full-length probe B (Kurath et al., 1992). These type 5 patterns are due to cleavage of the type 6 probes at the major heterogeneity site of STMV, which lies at base 751, in the overlapping region present in both probes A and B. Control lanes labeled "6" in panels C and D of Figure 2 show the characteristic type 6 patterns, in which the majority of each probe is protected intact at approximately 457 bases for probe A and 791 bases for probe B, with minor amounts of nonspecific digestion as expected (Kurath et al., 1992). The control lanes in Figure 3 show that the inocula for these groups of passage lines were mixtures of the type 5 and type 6 patterns as expected.

For each group of sibling passage lines, genetic changes that occurred during serial passages were detected as visible differences between the RNase pro-

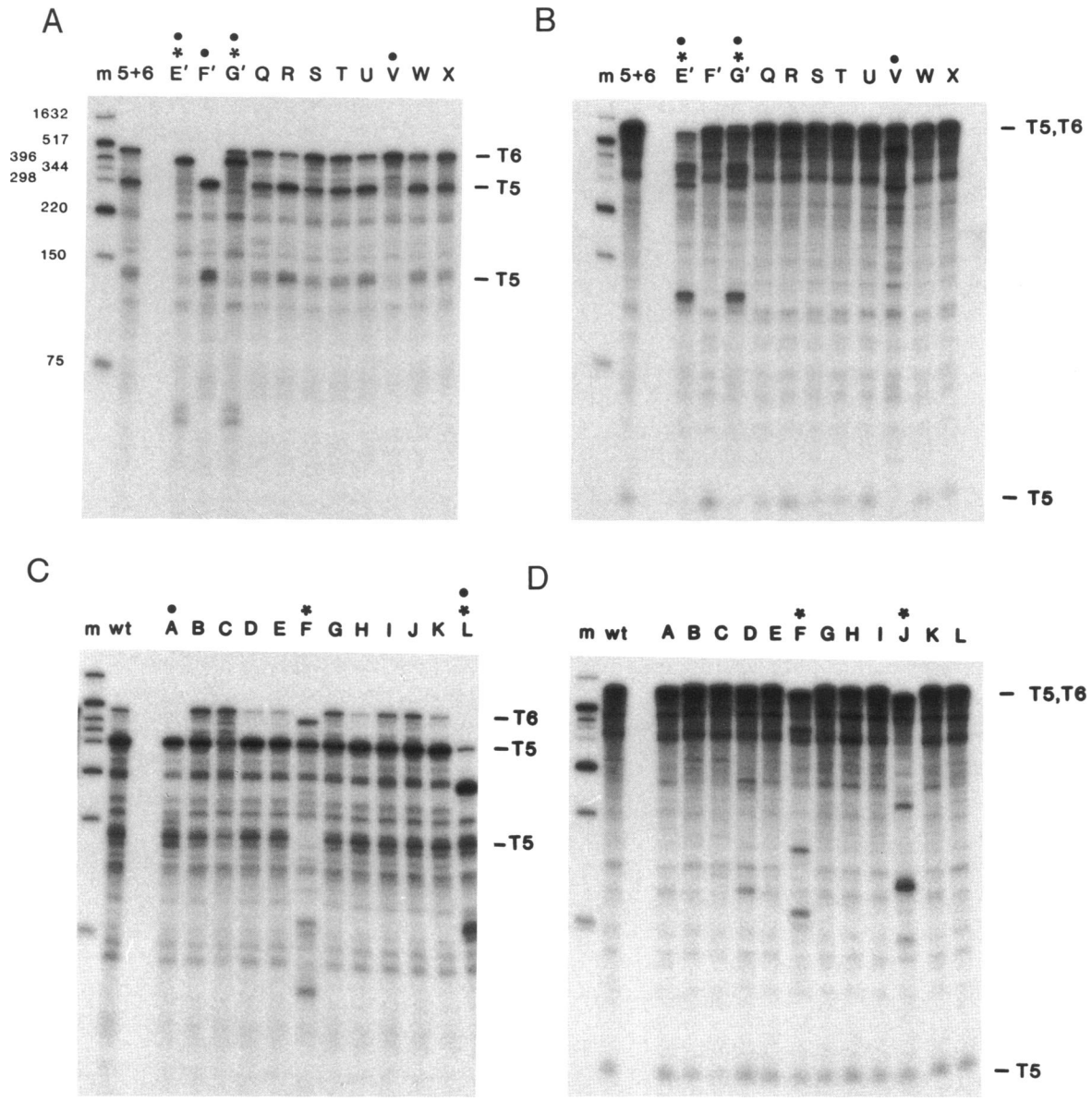

FIGURE 3. RNase protection assay patterns of progeny populations of mixed inoculum STMV passage lines after 10 passages. Progeny of T56 passage lines (mixture of T5 and T6 RNA) are shown in **A** (probe A) and **B** (probe B), and progeny of wtRNA passage lines are shown in **C** (probe A) and **D** (probe B). Individual passage line designations are shown as upper case letters above the lanes, and control assays of T56 and wt RNA inocula are marked 5+6 and wt respectively. Marker lanes (m) and T5 and T6 diagnostic fragments are as in Figure 2. *, passage lines in which a novel mutation was detected; ●, lines showing selection of a single genotype from the original mixture.

tection patterns of the original inoculum RNA in the control lanes and the patterns of the progeny populations. Three different kinds of genetic changes were detected in the various passage lines, and they are summarized in Table 1. Nine of the 42 passage lines showed novel mutations (at sites other than the major heterogeneity site) evident as new cleavage fragments. This occurred in the patterns of passage lines from each different group, and included lines T5-A′, T5-B′, T6-C′, T6-M, T56-E′, T56-G′, wt-F, wt-J, and wt-L (Figs. 2, 3). In some cases, the cleavage generating the new fragments went to completion (e.g., T56-E′) and, in other cases, cleavage was only partial (e.g., T56-G′), indicating either a mixed population of parental and mutant genotypes or incomplete digestion of the mismatch in the assay (Meyers et al., 1985; Winter et al., 1985).

The second kind of change detected was mutations resulting in apparent conversion at the major heterogeneity site of STMV, indicated by the appearance of diagnostic T6 fragments in T5 passage lines and diagnostic T5 fragments in T6 passage lines. This occurred at a clearly detectable level in 5 of the 19 passage lines where it was possible, i.e., those inoculated with a single genotype. In two lines (T5-D and T6-C′) the major-

ity of the population showed conversion and, in the other three lines (T5-E, T5-H, and T6-L), conversion was evident in only a minor portion of the population (Fig. 2). The progeny of lines T5-E and T5-H most closely resemble the original balance of T5 and T6 genotypes in the type strain STMV (lanes wt, Fig. 3).

The third kind of change observed did not involve mutation, but rather apparently complete selection of a single genotype in lines inoculated with mixtures of T5 and T6 genotypes. This occurred in 6 of the 23 lines in the T56 RNA and wt-RNA groups, including three lines that had additional novel mutations (T56-E', T56-G', and wt-L), and three lines in which the only detectable difference was the selection (T56-F', T56-V, and wt-A) (Fig. 3). The selection showed no bias, in that three lines selected T5 genotypes and three selected T6 genotypes.

Timing of the genetic changes

The timing of the genetic changes detected in various passage lines could be estimated by looking at the RNase protection assay patterns of progeny from intermediate passages. For lines with mutations, the first passage at which the mutation was detected is listed in Table 2, illustrating that the mutations occurred at a wide range of different times. After the first appearance of the mutation most of these lines showed an increase in the proportion of the population exhibiting the mutation in subsequent passages, and then an apparent stabilization of the new genotype as the master sequence. Passage lines T5-A', T5-B', T6-C', and T6-D' were carried out to 25 passages and showed no further change relative to the progeny after 10 passages (data not shown).

In the cases of apparent selection of a single T5 or T6 genotype from a mixture, the selection was most often complete after the first passage, with only one of the six lines showing selection at a later passage (P5-7 for line wt-L) (data not shown). In lines with both selection and mutation, the two phenomena did not coincide in timing (data not shown). Lines T56-E' and T56-G' both showed apparently complete selection of T6 genomes after the first passage, whereas the mutations in those lines did not appear until later passages (Table 2).

Sequence analyses of mutations

All 10th-passage progeny populations showing mutations were further analyzed by RNase protection assays using partial probes C and D, which are complementary to bases 1–602 and 791–1059 of pSTMV6, respectively. These assays provided orientation and mapping data to estimate the locations of the mutation sites in the progeny RNA (data not shown) (Kurath et al., 1992). The five passage lines showing apparent conversion between T5 and T6 genotypes all had mutations that mapped to the major heterogeneity site near base 751.

TABLE 2. Characterization of mutations in STMV serial passage lines.[a]

Passage line	Timing of mutation[b]	Mutation site[c]	Mutation
T5-A'	P1	1012	A to U
T5-B'	P5	849	U to C
T6-M	P2-4	305	A to U
		321	U to C
T56-E'	P2-3	662	G to A
T56-G'	P4-7	662	G to A
T56-V	P5-9	298	U to A
wt-F	P2-3	660	U to G
wt-J	P5-7	112–145	34-base deletion
		76–79[d]	Deletion of an A
		81	A to G, partial
wt-L	P1-4	849	U to C
T5-D	P2-4	751	A to C[e]
T6-C'	P4-6	748	U to C
T5-E	P5-7	MHS	ND
T5-H	P8-10	MHS	ND
T6-L	P2-4	MHS	ND

[a] The first 12 rows are passage lines with novel mutations and the last 5 rows are lines showing conversion between T5 and T6 patterns. ND indicates "not determined."

[b] Numbers in this column indicate the first passage (P) in which the mutation was detectable by RNase protection assay. Entries showing a range of passages are those lines in which only selected passage progeny were assayed (e.g., for line T5-D the mutation was not detected at passage 1, but was present at passage 4, so it appeared sometime during passages 2–4).

[c] Mutation sites shown as nucleotide numbers were determined by direct RNA sequencing of genomic RNA from STMV populations purified after the 10th passage. MHS indicates mutations at or near the major heterogeneity site of STMV at base 751 (Kurath et al., 1992), as determined by RNase protection assay.

[d] This is the deletion of a single A from four A residues at bases 76–79.

[e] The A to C mutation in line T5-D is exact conversion of T5 to T6 at the MHS of STMV (Kurath et al., 1992).

For all lines showing novel mutations, the mutations were characterized by direct sequence analyses of the progeny genomic RNA in the region predicted to contain the mutation. This was also done for two of the conversion lines, T5-D and T6-C', in which the major portion of the population showed apparent conversion at the major heterogeneity site near base 751. In all cases a mutation was found, relative to the known sequence of the inoculum RNA, very near the estimated mutation site (within 5 nt for 9 of the 11 mutations characterized here).

The results of the sequencing studies are compiled in Table 2. Seven of the nine passage lines with novel mutations had single-base substitutions, and one line (T6-M) had two separate single-base substitutions in close proximity to the site estimated by RNase protection analyses. The last line (wt-J), which had been difficult to map from RNase protection data, had three sequence differences near the 5' end of the genome, in-

cluding one single-base substitution, one single-base deletion, and one 34-base deletion (Table 2). In all cases, the mutated genomes comprised the great majority of the populations, as shown by the clarity of the sequence banding pattern at and beyond the mutation(s) (data not shown).

Sequence analyses of two lines that had apparent conversion at the major heterogeneity site showed that one of the lines, T5-D, had undergone true conversion from T5 to T6 due to an A to C mutation at base 751 (Kurath et al., 1992). The other line, T6-C', had a U to C mutation 3 nt away at base 748, making it essentially a pseudoconversion.

Among the total of 12 single-base substitutions characterized by RNA sequencing, there were seven transitions and five transversions (Table 2). The most frequent substitution was U to C transitions (4 of the 12 mutations), but overall there was no evidence of significant bias in the substitution types, and no high prevalence of G to A substitutions as seen within the STMV type strain quasispecies (Kurath et al., 1992). Nine of the 12 substitutions occurred at positions where the parental sequence had an A or a U. Examination of the nucleotide sequences surrounding the mutation sites did not reveal any common context features within 15 bases on either side of the mutations, other than possibly an elevated A+U frequency surrounding some mutation sites. This is most notable for the major heterogeneity site, which is 82% A+U (9/11) in the region extending 5 bases upstream and downstream from nt 751, and 76% A+U (16/21) in the larger region 10 bases upstream and downstream. However, many mutation sites are not significantly above the average A+U frequency for the entire STMV genome, which is 54.3%.

The mutations characterized in these passage lines were distributed along the length of the STMV genome (Fig. 4). The majority of changes (11 of 17 individual mutations, representing 11 of the 14 passage lines that showed change) occurred within the long 3' untranslated region (bases 643–1059). This region has been shown to fold into a 3' tRNA-like structure that is conserved among several tobamovirus genomes, and is hypothesized to be of significance in viral RNA replication and encapsidation (Felden et al., 1994; Gultyaev et al., 1994). Only three passage lines had mutations within the STMV coding regions (bases 53–642) that would have altered the amino acid sequence of viral proteins. In passage line T56-V, the mutation at base 298 would result in a conservative change of serine to threonine at amino acid 46 of the viral coat protein. Passage line T6-M had two mutations in the coat protein, one of which would result in a change from glutamine to leucine at amino acid 48, whereas the other was a silent third codon position change. The progeny of passage line wt-J had three mutations, including a 34-base deletion, in the 6.8-K ORF. The translation of this modified ORF would result in a frame shift after the first eight amino acids and premature termination after a total of 19 amino acids, producing an extremely modified, truncated 6.8-K protein. The fact that this drastically modified variant was obtained in a passage line initiated with wild-type genomic STMV RNA means that it could have arisen by amplification of a previously existing trace variant rather than by novel mutation during this experimental period. This variant is similar to an unusual field isolate of STMV that was found to have a 71-nt deletion within the 6.8-K ORF (D.M. Mathews & J.A. Dodds, unpubl. data), showing precedence for the viability of STMV variants with nonfunctional 6.8-K ORFs.

DISCUSSION

In the passage lines initiated with cloned inocula, it must be recognized that the original mutations leading to the variants observed could have occurred during the single round of transcription by T7 RNA polymerase to produce the inocula RNAs. However, the multiple rounds of replication in plants, the delayed appearance of many of the changes, and the observation of the same types of changes in the wtRNA passage lines suggest that the majority of the change seen involved mutation during replication in serial passage.

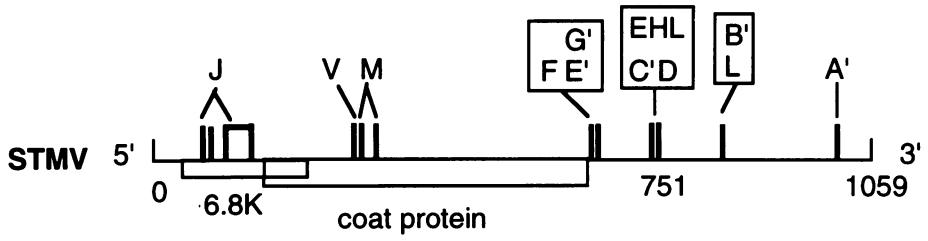

FIGURE 4. Composite map showing the distribution along the STMV genome of the genetic changes observed in all serial passages. The two ORFs in the STMV genome (Mirkov et al., 1989) are shown as rectangles below the horizontal line representing the STMV RNA genome. Letters above the genome are abbreviated notations for the passage lines showing genetic change (see Table 2), the bold rectangle under line J indicates the deletion of bases 112–145, and boxes show groups of passage lines with clusters of genetic change.

In the passage lines inoculated with wtRNA, the inoculum was already a quasispecies mixture with a variant frequency of approximately 36% (Kurath et al., 1992). The observation of similar types of change in the clonal inocula passage lines and the wtRNA passage lines is a valuable demonstration that the phenomena observed were not artifacts due to the use of an artificially uniform inoculum, but represent phenomena that are likely to occur in STMV populations in nature.

The most significant results from this work are the characterization of the nonrandom or "clustered" nature of the sites of several of the genetic changes (Fig. 4). Considering that the STMV genome is 1,059-nt long, and that 17 individual mutations were identified in this study, the occurrence of any two changes at or near the same site should be very unlikely. Nevertheless, three, and possibly four, clusters of mutations were observed within these passages. The most striking was the mutation in 5 of 19 single inoculum passage lines (26%) at or near the major heterogeneity site near position 751. This site has been shown to be variable within (Kurath et al., 1992) and between (Kurath et al., 1993a) STMV populations, and the results described here demonstrate that it also mutates rapidly and with high frequency. Of the two lines sequenced, T5-D showed conversion by the exact same mutation at nt 751 that distinguished the T5 and T6 genotypes in the STMV type strain (Kurath et al., 1992), and line T6-C′ had a mutation nearby at nt 748. Three other lines that showed conversion at or near the same site by RNase protection mapping, suggesting a prevalence of mutation within a small region near nt 751. A similar case of inexact mutation at a major heterogeneity site defining a small region of variability was seen during serial passage of cucumber mosaic virus satellite RNA clone transcripts in tobacco (Kurath & Palukaitis, 1990). High frequency of mutation at a site known to be heterogeneous in natural populations has also been reported in serial passage studies with animal viruses such as hepatitis delta virus (Luo et al., 1990) and simian immunodeficiency virus (Baier et al., 1991).

The second mutation cluster involved identical G to A mutations at nt 662 in passage lines T56-E′ and T56-G′, and a U to G mutation at nt 660 in line wt-F. This again suggests a prevalence of mutation not at an exact site, but within a closely defined region. This is further supported by the fact that, in the original study of the quasispecies nature of the STMV type strain, 1 of the 14 variants identified of 42 genomic clones (variant pSTMV17) differed from the consensus sequence by a G to U mutation exactly at nt 662 (Kurath et al., 1992).

The third case of clustered mutations is the identical U to C mutations at nt 849 in passage lines T5-B′ and wt-L. A possible fourth cluster is the mutations at nt 298 and nt 305 in passage lines T56-V and T6-M, respectively, but these are slightly farther apart than the other clusters.

For plant viruses, there are relatively few studies of genetic change during serial plant passages, and cases of both low and high rates of genetic change have been reported (van Vloten-Doting & Bol, 1988; Skotniki et al., 1992; Keese & Gibbs, 1993; Rao & Hall, 1993). In one study of local lesion-cloned tobacco mild green mosaic virus, no variability was detected by RNase T1 fingerprints in the consensus sequences of four tobacco serial passage lines (Rodriguez-Cerezo & Garcia-Arenal, 1989). A very low mutation rate was also reported for foreign gene sequences inserted into an infectious clone of tobacco mosaic virus (TMV) when the transcript RNA was passaged serially in plants (Kearney et al., 1993). However, serial passages of RNA transcribed from the parental TMV clone without the foreign gene insert resulted in rapid generation of phenotypic mutants at frequencies that varied with different selection pressures and from plant to plant (Aldaoud et al., 1989), reminiscent of the STMV results described here.

Several studies with animal viruses have shown rapid genetic change when cloned inocula were serially passaged in vivo (e.g., Luo et al., 1990; Baier et al., 1991; Domingo et al., 1992). In particular, Domingo et al. (1993) have documented the rapid emergence of antigenic variants in the absence of immune selection pressure for foot and mouth disease virus and other animal viruses, and they proposed a random change model to explain this phenomenon. The STMV passages in the present work were also done in the absence of any apparent shift in selection pressure, and the same model may well apply to explain the diversity of changes observed here. In addition, it should be remembered that, due to the satellite nature of STMV, a potential source of unintentionally modified selection pressure would be genetic change of the helper virus populations. It has been demonstrated experimentally that the helper tobamovirus can exert selective pressure on STMV populations (Kurath et al., 1993b). Although changes in helper virus populations were not assessed in this work, the large number of independent replicate passage lines would preclude any random mutation of the helper virus in a single line from being the major source of variation observed.

The quasispecies nature of STMV populations has been proven and characterized for the STMV type strain (Kurath et al., 1992), and genetic variability of the consensus sequences for multiple field isolates of STMV has also been shown (Kurath et al., 1993a). The current work did not assess the regeneration of quasispecies variation from clonal inocula, but rather the changes in the consensus sequences of STMV populations during replication in plants. This may represent the process by which longer term genetic diversification of STMV occurs. This is a complex process involving, for the passage lines initiated with clonal inocula, novel mutation to create variants, subsequent competition between variants during continued replication,

and random sampling events that may occur during spread throughout each host plant, and that certainly occur during host-to-host transmission (Domingo et al., 1993; Domingo & Holland, 1994). Due to the complexity of this process, the genetic changes observed in our data cannot be used to estimate frequencies of any specific step, such as rates of mutation. Thus, we do not claim that the clusters described represent hot spots of mutation, but of genetic change. This data demonstrates that genetic diversification of STMV can occur relatively rapidly, and the fact that sibling lines showed different changes is evidence for the important role of chance and random sampling events in determining the consensus sequence of replicating populations (Domingo et al., 1993; Domingo & Holland, 1994). At present we do not have any information on the fitness of the newly generated genotypes relative to the parental inoculum sequences, so the possible role of selection in the changes observed cannot be assessed.

In summary, this was a study in experimental evolution, in which the process of genetic change in RNA genome populations could be observed and characterized under controlled conditions, with enough duplication to discern trends standing out above the chance and random events known to function in RNA evolution (Domingo et al., 1993; Domingo & Holland, 1994). Using STMV RNA populations replicating in plants as a system, and RNase protection analyses and RNA sequencing as methods to assess genetic change, we were able to characterize several phenomena regarding genetic change in STMV RNA populations. Novel mutations (not at the major heterogeneity site) were detected fixed in the progeny STMV populations in 9 of the 42 independent passage lines (21%) within the experimental time frame of 10 serial passages (approximately 6 months of replication). In addition, mutation at the major heterogeneity site near nt 751 occurred in 5 of the 19 lines (26%) inoculated with a single T5 or T6 STMV genotype. Because RNase protection analyses do not detect all mismatches, this is a minimum estimate of the amount of genetic change that actually occurred. In fact, previous comparisons of RNase protection data with direct RNA sequence data determined that, in assays of STMV RNAs, the RNase protection method detected approximately 50% of the actual number of mutations present (Kurath et al., 1992). Therefore, we would estimate that the actual number of genetic changes that occurred in our 42 passage lines was approximately twice the number detected and characterized in this report. This confirms that replicating STMV populations generate heterogeneity and diversify relatively quickly and with high frequency.

As for timing, genetic changes were detected in different lines at various passage numbers throughout the study, indicating a roughly constant probability of mutations occurring and/or becoming fixed. The distribution of the sites of change showed some preference for the 3′ noncoding region of the genome, and were not completely random in that 10 of the 14 changes observed occurred within one of three tightly defined regions (within 3 nt) identified in this study as "hot spots" of genetic change. Sequence analyses showed that there was no significant bias in the substitution types observed. This comprises our first detailed picture of the nature of the genetic change processes occurring in replicating STMV populations, and provides insight into the both the generation of RNA heterogeneity and the fundamental steps of RNA evolution.

MATERIALS AND METHODS

Serial passages

All serial passages were done in tobacco plants (*Nicotiana tabacum* L. cultivar Xanthi nn) using one plant per line for each passage. Passage lines were initiated by inoculating two leaves of small plants (2–4-leaf stage) as described with purified TMGMV helper virions and STMV RNA, which was either transcript RNA from full-length STMV cDNA clones pSTMV5 or pSTMV6 or both, or genomic RNA from the STMV type strain (Kurath et al., 1992). The satellite-free nature of the TMGMV helper preparation was confirmed by rigorous biological assay as described previously (Kurath et al., 1992). Inoculated plants were maintained in the greenhouse with automatic watering and care was taken that nothing touched any of the plants as they grew for 2–3 weeks to approximately the 15–20-leaf stage. At that time tissue was harvested for inoculation of the next passage and for assessment of the systemic STMV population by RNase protection analyses. Inocula for further passages of the original set of passage lines (labeled prime, see text) were virions purified from total systemically infected leaf tissue (Mirkov et al., 1989). For all subsequent lines, the inoculum was sap from pooled tissue samples from four different systemically infected leaves near the top and middle of the plant. This latter method of sampling has been found to accurately reflect the genetic composition of an STMV population in the systemically infected tissues of a tobacco plant under the conditions used here and within the limits of RNase protection assay assessment (G. Kurath & J.A. Dodds, unpubl. data).

RNase protection assays

Synthesis of the ^{32}P-labeled minus-sense probes were as described previously (Kurath et al., 1992). Probes A, B, C, and D were complementary to bases 602–1059, 1–791, 1–602, and 791–1059, respectively, of the full-length STMV cDNA clone pSTMV6 (Kurath et al., 1992). Target STMV RNA populations were either STMV genomic RNA extracted from virions purified from total systemically infected tissue (for prime lines) (Mirkov et al., 1989), or total nucleic acid preparations (Kurath et al., 1992) from pooled leaf tissue samples (approximately 100 mg total) from the four systemically infected leaves used to inoculate the subsequent passage. Control assays were done using plus-sense RNA transcribed from clones pSTMV5 and pSTMV6 (Kurath et al., 1992).

The RNase protection assay protocol was essentially as described by Winter et al. (1985), except that the RNase diges-

tion step was for 1 h at 37 °C rather than 34 °C. All assay reagents were exactly as in Winter et al. (1985). Briefly, each assay began by combining 1×10^5 cpm of radioactive probe with target RNA in hybridization solution. The quantities of the target RNAs used in each assay were 100 ng of purified genomic RNA, approximately 1 μg of total nucleic acid from leaf tissue, or 100 ng of transcript RNA from clones. After annealing overnight to form RNA:RNA duplexes, a salt solution with RNases A and T1 was added and assays were incubated for 1 h at 37 °C. An SDS-proteinase K solution was then added to digest the RNases for 15 min at 37 °C. After extraction and precipitation, the samples were resuspended in a formamide loading solution, heat denatured, run on a 7 M urea, 6% polyacrylamide gel, and autoradiographed as described in Kurath et al. (1992).

RNA sequence analyses

Direct RNA sequence analysis using the dideoxynucleotide chain termination method was as described (Kurath et al., 1993b), with the exception that 0.3–3 μg of substrate RNA was necessary depending on the specific template used. Primers for sequencing were oligonucleotides complementary to the STMV consensus sequence (Mirkov et al., 1989) at appropriate regions downstream from the sites of mutations estimated by RNase protection assay mapping. Primers were synthesized by the Bioinstrumentation Facility at the University of California, Riverside, California.

ACKNOWLEDGMENTS

We thank Jim Heick for technical assistance with a flair, Deborah Mathews for the use of unpublished results, and several colleagues for critical readings of the manuscript.

Received April 5, 1995; returned for revision May 10, 1995; revised manuscript received June 27, 1995

REFERENCES

Aldaoud R, Dawson WO, Jones GE. 1989. Rapid, random evolution of the genetic structure of replicating tobacco mosaic virus populations. *Intervirology 30*:227–233.

Baier M, Dittmar MT, Cichutek K, Kurth R. 1991. Development in vivo of genetic variability of simian immunodeficiency virus. *Proc Natl Acad Sci USA 88*:8126–8130.

Domingo E, Diez J, Martinez MA, Hernandez J, Holguin A, Borrego B, Mateu MG. 1993. New observations on antigenic diversification of RNA viruses. Antigenic variation is not dependent on immune selection. *J Gen Virol 74*:2039–2045.

Domingo E, Escarmis C, Martinez MA, Martinez-Salas E, Mateu MG. 1992. Foot-and-mouth disease virus populations are quasispecies. In: Holland JJ, ed. *Genetic diversity of RNA viruses. Current Topics in Microbiology and Immunology 176*. Berlin/New York: Springer-Verlag. pp 33–45.

Domingo E, Holland JJ. 1988. High error rates, population equilibrium and evolution of RNA replication systems. In: Holland JJ, Domingo E, Ahlquist P, eds. *RNA genetics, vol 3*. Boca Raton, Florida: CRC Press. pp 3–35.

Domingo E, Holland JJ. 1994. Mutation rates and rapid evolution of RNA viruses. In: Morse SS, ed. *The evolutionary biology of viruses*. New York: Raven Press Ltd. pp 161–184.

Domingo E, Martinez-Salas E, Sobrino F, de la Torre JC, Portela A, Ortin J, Lopez-Galindez C, Perez-Brena P, Villanueva N, Najera R, VandePol S, Steinhauer D, DePolo N, Holland JJ. 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: Biological relevance—A review. *Gene 40*:1–8.

Felden B, Florentz C, McPherson A, Giege R. 1994. A histidine accepting tRNA-like fold at the 3′-end of satellite tobacco mosaic virus RNA. *Nucleic Acids Res 22*:2882–2886.

Gultyaev AP, van Batenburg E, Pleij CWA. 1994. Similarities between the secondary structure of satellite tobacco mosaic virus and tobamovirus RNAs. *J Gen Virol 75*:2851–2856.

Kearney CM, Donson J, Jones GE, Dawson WO. 1993. Low level of genetic drift in foreign sequences replicating in an RNA virus in plants. *Virology 192*:11–17.

Keese P, Gibbs A. 1993. Plant viruses: Master explorers of evolutionary space. *Curr Opin Genet Dev 3*:873–877.

Kurath G, Heick JA, Dodds JA. 1993a. RNase protection analyses show high genetic diversity among field isolates of satellite tobacco mosaic virus. *Virology 194*:414–418.

Kurath G, Palukaitis P. 1990. Serial passage of infectious transcripts of a cucumber mosaic virus satellite RNA clone results in sequence heterogeneity. *Virology 176*:8–15.

Kurath G, Rey MEC, Dodds JA. 1992. Analysis of genetic heterogeneity within the type strain of satellite tobacco mosaic virus reveals several variants and a strong bias for G to A substitution mutations. *Virology 189*:233–244.

Kurath G, Rey MEC, Dodds JA. 1993b. Tobamovirus helper specificity of satellite tobacco mosaic virus involves a domain near the 5′ end of the satellite genome. *J Gen Virol 74*:1233–1243.

Luo G, Chao M, Hsieh S, Sureau C, Nishikura K, Taylor J. 1990. A specific base transition occurs on replicating hepatitis delta virus RNA. *J Virol 64*:1021–1027.

Meyers RM, Larin Z, Maniatis T. 1985. Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science 230*:1242–1246.

Mirkov TE, Mathews DM, Du Plessis DH, Dodds JA. 1989. Nucleotide sequence and translation of satellite tobacco mosaic virus RNA. *Virology 170*:139–146.

Rao ALN, Hall TC. 1993. Recombination and polymerase error facilitate restoration of infectivity in brome mosaic virus. *J Virol 67*:969–979.

Rodriguez-Cerezo E, Garcia-Arenal F. 1989. Genetic heterogeneity of the RNA genome population of the plant virus U5-TMV. *Virology 170*:418–423.

Skotniki ML, Mackenzie AM, Gibbs AJ. 1992. Turnip yellow mosaic virus variants produced from DNA clones encoding their genomes. *Arch Virol 127*:25–35.

Steinhauer DA, Holland JJ. 1987. Rapid evolution of RNA viruses. *Annu Rev Microbiol 41*:409–433.

Valverde RA, Dodds JA. 1986. Evidence for a satellite RNA associated naturally with the U5 strain and experimentally with the U1 strain of tobacco mosaic virus. *J Gen Virol 67*:1875–1884.

Valverde RA, Dodds JA. 1987. Some properties of isometric virus particles which contain the satellite RNA of tobacco mosaic virus. *J Gen Virol 68*:965–972.

van Vloten-Doting L, Bol JF. 1988. Variability, mutant selection, mutant stability in plant RNA viruses. In: Holland J, Domingo E, Ahlquist P, eds. *RNA genetics, vol 3*. Boca Raton, Florida: CRC Press. pp 37–51.

Winter E, Yamamoto F, Almoguera C, Perucho M. 1985. A method to detect and characterize point mutations in transcribed genes: Amplification and over expression of the mutant c-Ki-ras allele in human tumor cells. *Proc Natl Acad Sci USA 82*:7575–7579.