

11-1-2001

Structure and Temporal Dynamics of Populations within Wheat Streak Mosaic Virus Isolates

Jeffrey S. Hall

U.S. Department of Agriculture

Roy C. French

University of Nebraska - Lincoln, rfrench2@unl.edu

Thomas Jack Morris

University of Nebraska - Lincoln, jmorris1@unl.edu

Drake C. Stenger

University of Nebraska - Lincoln

Follow this and additional works at: <http://digitalcommons.unl.edu/plantpathpapers>



Part of the [Plant Pathology Commons](#)

Hall, Jeffrey S.; French, Roy C.; Morris, Thomas Jack; and Stenger, Drake C., "Structure and Temporal Dynamics of Populations within Wheat Streak Mosaic Virus Isolates" (2001). *Papers in Plant Pathology*. Paper 14.

<http://digitalcommons.unl.edu/plantpathpapers/14>

This Article is brought to you for free and open access by the Plant Pathology Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Plant Pathology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Structure and Temporal Dynamics of Populations within Wheat Streak Mosaic Virus Isolates

JEFFREY S. HALL,^{1,2} ROY FRENCH,^{1,3} T. JACK MORRIS,² AND DRAKE C. STENGER^{1,3*}

Agricultural Research Service, U.S. Department of Agriculture,¹ School of Biological Sciences,² and Department of Plant Pathology,³ University of Nebraska, Lincoln, Nebraska 68583

Received 28 March 2001/Accepted 1 August 2001

Variation within the Type and Sidney 81 strains of wheat streak mosaic virus was assessed by single-strand conformation polymorphism (SSCP) analysis and confirmed by nucleotide sequencing. Limiting-dilution subisolates (LDSIs) of each strain were evaluated for polymorphism in the P1, P3, NIa, and CP cistrons. Different SSCP patterns among LDSIs of a strain were associated with single-nucleotide substitutions. Sidney 81 LDSI-S10 was used as founding inoculum to establish three lineages each in wheat, corn, and barley. The P1, HC-Pro, P3, CI, NIa, N1b, and CP cistrons of LDSI-S10 and each lineage at passages 1, 3, 6, and 9 were evaluated for polymorphism. By passage 9, each lineage differed in consensus sequence from LDSI-S10. The majority of substitutions occurred within NIa and CP, although at least one change occurred in each cistron except HC-Pro and P3. Most consensus sequence changes among lineages were independent, with substitutions accumulating over time. However, LDSI-S10 bore a variant nucleotide (G₆₀₁₆) in NIa that was restored to A₆₀₁₆ in eight of nine lineages by passage 6. This near-global reversion is most easily explained by selection. Examination of nonconsensus variation revealed a pool of unique substitutions (singletons) that remained constant in frequency during passage, regardless of the host species examined. These results suggest that mutations arising by viral polymerase error are generated at a constant rate but that most newly generated mutants are sequestered in virions and do not serve as replication templates. Thus, a substantial fraction of variation generated is static and has yet to be tested for relative fitness. In contrast, nonsingleton variation increased upon passage, suggesting that some mutants do serve as replication templates and may become established in a population. Replicated mutants may or may not rise to prominence to become the consensus sequence in a lineage, with the fate of any particular mutant subject to selection and stochastic processes such as genetic drift and population growth factors.

RNA viruses have the common property of high mutation rates due to an error-prone RNA polymerase (10, 12, 13). Because of this, RNA viruses can increase in genetic diversity and undergo rapid evolutionary changes (11, 12, 34, 39). Sequence diversity among isolates of plant RNA viruses has been well documented (1, 2, 9, 15, 24–28, 33, 45, 46, 55–59, 65). However, low levels of variation found within individual isolates of many plant viruses imply that accumulation of fixed differences can be a slow process (1, 22, 23, 25, 27, 28, 36, 44, 45, 52, 56). Thus, evolutionary and population genetic factors favoring genetic stability such as purifying selection or episodes of genetic bottlenecks must be in play for many plant viruses.

The deterministic quasispecies model is applicable when large pools of virus are in direct competition (11–13, 37, 39, 60). However, there are numerous examples where competition among viral genomes appears to be limited to various degrees, resulting in stochastic rather than deterministic outcomes (16, 20, 30, 34, 40, 41, 50). It has been proposed that no sharp boundary exists between the deterministic and stochastic modes of population behavior (48). Random drift cannot occur at the deterministic end of the spectrum but gradually becomes increasingly likely as the effective, or replicating, population size decreases (16, 38, 48). Selection, on the other hand, becomes less effective with a decline in population size. One

mechanism for reducing effective population size is the introduction of genetic bottlenecks (35, 41, 46). Recently it was demonstrated that colonization of wheat by strains of wheat streak mosaic virus (WSMV) is a discontinuous process (17), suggesting that within-plant subdivision may impose bottlenecks during systemic infection. If so, genetic drift should be a prominent factor in WSMV evolution over time.

WSMV is the type species of the newly established genus *Tritimovirus* within the family *Potyviridae* (62). The monopartite genome (9,384 nucleotides) of WSMV is translated as a polyprotein that is subsequently cleaved by virus-encoded proteinases into 8 to 10 mature proteins. Three WSMV strains (Type, Sidney 81, and El Batán 3) have been completely sequenced (9, 62). Type and Sidney 81 have 97.6% nucleotide sequence identity and are representative of American WSMV isolates. El Batán 3, from Mexico (51), has diverged from the American strains and retains ~79% nucleotide sequence identity to Type and Sidney 81 (9). Examination of variation among the three WSMV strains suggests that negative selection operates to conserve certain regions of the genome, while much of the divergence may be explained by genetic drift (9). Field populations of WSMV are complex and consist of numerous genotypes that occur in the same field but rarely (~2%) in the same infected plant (33). Maintenance of genetic diversity within WSMV field populations may be facilitated by three distinct mechanisms of genetic isolation: cross-protection, subdivision of populations within infected plants, and vector transmission bottlenecks (17). Thus, sympatric viral lineages are

* Corresponding author. Mailing address: USDA-ARS, 344 Keim Hall, University of Nebraska, Lincoln, NE 68583. Phone: (402) 472-2710. Fax: (402) 472-2853. E-mail: dstenger@unlnotes.unl.edu.

TABLE 1. Primer sets used for RT-PCR and restriction endonucleases used to prepare PCR products for SSCP analysis

Primer set	Downstream primer ^a	Upstream primer ^a	Restriction endonuclease
RCF1	<u>AGCTGGATCCTTTTTTTTTTTTTTTT</u>		
P1	<u>CCGGATCCCTAATAAGTTGTGATAAAGCC</u>	<u>CCGGATCCCAATGGCAACAGCGAATTGT</u>	<i>Sau3AI</i>
HC-Pro	<u>CCGAATCCTCAGCCAATTTTGTAACTTT</u>	<u>CCGGATCCCCAGTGGCATAACAATAC</u>	<i>HaeIII-NheI</i>
P3	<u>CCGGATCCTATTGGTATTCAACCAATTC</u>	<u>CCGGATCCGCGGGTTCCAAGAGACTGTT</u>	<i>RsaI</i>
CI5'	<u>CAACGACAACGTCTGCATGCAAATTCAC</u>	<u>GTACCCGGGGAAGCGAGAGGAAG</u>	<i>Sau3AI</i>
CI3'	<u>GAGGATCCTCACTCGAACTTGACACTACG</u>	<u>GATGTCGAAGCACTCCGAGCTGAGAAG</u>	<i>HaeIII</i>
NIa	<u>GACTTCTAGATCATTGCCAACTAACCAAG</u>	<u>GTCTAAGCTTGGGCAAAGCAGCAGCAA</u>	<i>TaqI</i>
NIb	<u>GGGATCCTCATTCTGACACGAGTATTG</u>	<u>CCGGATCCCCAGTCTCCAATGACGAA</u>	<i>AluI</i>
CP	<u>CGATTTTTTTTTTTTTTGGCCGTCGCCCT</u>	<u>AGCTTATGAGCGCTATTATTGCAGCAT</u>	<i>AluI</i>

^a Underlines denote 5'-terminal regions of primers that do not correspond to WSMV sequence.

genetically isolated from one another such that stochastic processes may influence divergence.

In this report, variation within WSMV isolates was assessed after passage through a genetic bottleneck imposed by low-multiplicity-of-infection (MOI) inoculation and upon repeated passage at high MOI. These experiments were designed to detect shifts in the consensus sequence of a lineage at various points in time. A single-strand conformation polymorphism (SSCP) assay was employed, making it practical to screen for consensus sequence polymorphism in multiple samples across large genomic regions. We further examined nonconsensus variation existing after low-MOI bottlenecking and after repeated passage at high MOI. Collectively, our results indicate that populations of WSMV change over time and that there are multiple processes affecting variation within plant virus populations.

MATERIALS AND METHODS

Generation of WSMV LDSIs. WSMV-Type (ATCC PV57; GenBank accession no. AF285169) was isolated from wheat collected in Kansas (32). WSMV-Sidney 81 (GenBank accession no. AF057533) was recovered from wheat grown in western Nebraska (6). A limiting-dilution method was used to resolve individual infectious units. Limiting-dilution subisolates (LDSIs) were recovered from sets of wheat (*Triticum aestivum* L. cv. Centurk) seedlings inoculated with 10^{-5} dilutions of sap extract, in which only ~1% of the plants became infected. Altogether, 22 Sidney 81 and 25 Type LDSIs were obtained. The probability of each LDSI being initiated by >1 infectious unit was $\leq 0.5\%$ (5). Systemically infected LDSI leaf samples were harvested 21 days postinoculation and stored at -80°C .

Passage inoculations. Sidney 81 LDSI-S10 was used as the founding inoculum for passage experiments. The first passage with LDSI-S10 inoculum and all subsequent passage inoculations were conducted at high MOI using a 10^{-1} dilution of sap extract. Nine separate lineages (three each in wheat cv. Centurk, corn [*Zea mays* L.] var. N28ht, and barley [*Hordeum vulgare* L.] cv. Black Hullless) were established at passage 1 and maintained for nine serial passages conducted at 3-week intervals. A single passage in one lineage consisted of five plants grown in a common pot. For each lineage at each passage, systemically infected leaf samples were collected 21 days postinoculation (p.i.) and stored at -80°C . To minimize differential selective pressures due to environmental conditions, the passage experiment was conducted in a growth chamber under controlled conditions (20°C with a 16-h photoperiod).

RT-PCR. Total nucleic acids were isolated from frozen leaf samples and viral RNA was reverse transcribed as described previously (33). In addition to using the RCF1 primer (Table 1) to initiate cDNA synthesis, random primers were incorporated into reverse transcription (RT) reactions to ensure cDNA representative of the entire genome. Regions of the viral genome (Fig. 1) were amplified by PCR using primer sets listed in Table 1. Variation among LDSIs was estimated by examining four cistrons (P1, P3, NIa, and CP) amplified by RT-PCR. LDSI-S10 and each of the nine lineages derived from this subisolate were further characterized by RT-PCR amplification with all primer sets listed in Table 1, resulting in ~96% coverage of the viral genome. PCR was performed in

a Perkin-Elmer 9600 thermocycler for 35 cycles (except for NIa, for which 30 cycles were employed) using *Taq* DNA polymerase, and the products were analyzed by SSCP and sequencing (see below) to correlate haplotypes with consensus sequences. To examine nonconsensus variation within the CP cistron of LDSI-S10 and derived lineages, the Expand High-Fidelity System (EHFS) PCR (Boehringer Mannheim) was used (30 cycles) to minimize PCR-induced sequence errors.

To determine the intrinsic error rate of each PCR system, plasmids of known sequence served as template donors. Clones of Sidney 81 LDSI-S9 NIa and LDSI-S10 CP served as template sources for PCR with *Taq* polymerase or the EHFS, respectively. Each template donor plasmid was digested with *PvuII*, and the fragments containing each WSMV insert (plus ~350 bp of flanking vector sequence) were gel purified. Each gel-purified *PvuII* fragment (1 ng) was used as a template for PCR with the appropriate CP or NIa primer set (Table 1) under the same conditions described above. The resulting PCR products were gel purified and ligated into pGEM-T (Promega), and the ligation products were used to transform *Escherichia coli* strain DH5 α . A total of 10 (EHFS) or 20 (*Taq* polymerase) clones (defined as reclones) of each PCR product were sequenced, and all substitutions were attributed to polymerase errors introduced during PCR.

SSCP analysis. PCR products were digested with restriction endonucleases to obtain fragments of appropriate lengths (42) for SSCP analysis (Table 1; Fig. 1). Each restriction endonuclease-digested PCR product was denatured in 25% formamide-5 mM NaOH-0.5 mM EDTA at 99°C for 5 min and placed immediately on ice. Single-stranded DNA fragments were separated on 10% polyacrylamide gels for 16 h in 1X Tris-borate-EDTA buffer at 4 and 20°C . SSCP patterns were visualized by silver staining (43). The accuracy of the method in detecting nucleotide substitutions was determined using plasmids of known sequence (10 Type LDSI CP inserts, and 10 Sidney 81 NIa reclone inserts).

Cloning and sequencing. The P1, P3, NIa, and CP cistrons of each LDSI were screened for variation by SSCP. PCR products of predominant and variant SSCP patterns were ligated into pGEM-T (Promega, Madison, Wis.) and transformed into *E. coli* strain DH5 α . PCR products (P1, HC-Pro, P3, CI-5', CI-3', NIa, NIb, and CP) of LDSI-S10 were cloned into pGEM-T, as were PCR products displaying variant SSCP patterns detected in all nine LDSI-S10-derived lineages at passages 1, 3, 6, and 9. Additional PCR products displaying predominant SSCP patterns for each cistron from various lineages and passages also were cloned into pGEM-T. Inserts of two plasmids derived from each PCR product cloned were sequenced in both directions by primer walking using universal and internal sequencing primers. Twenty additional LDSI-S10 NIa clones were partially sequenced (nt 5896 to 6427) to determine if two polymorphic sites appearing in more than one lineage during the passage experiment were resident as minor components of the LDSI-S10 population. Multiple EHFS-PCR CP clones of LDSI-S10 and three lineages at passage 9 also were sequenced on both strands by primer walking. Sequencing was performed at the Iowa State University DNA Synthesis and Sequencing Laboratory, Ames, Iowa. Sequences were compiled using Sequencher 3.1 (Gene Codes Corp., Ann Arbor, Mich.), edited to remove terminal regions in which PCR primers annealed, and aligned using Clustal X (64). A cladogram depicting relationships among consensus sequence variants present at passage 9 was generated by neighbor-joining using Clustal X. Sequence data sets were analyzed using the SITES program (18) and DnaSP v.3.0 (49). The probability that the amount of sequence variation observed in the LDSI-S10 data set was due to PCR error alone was determined by generating 100,000 simulated samples using the computer program MLPCR (67).

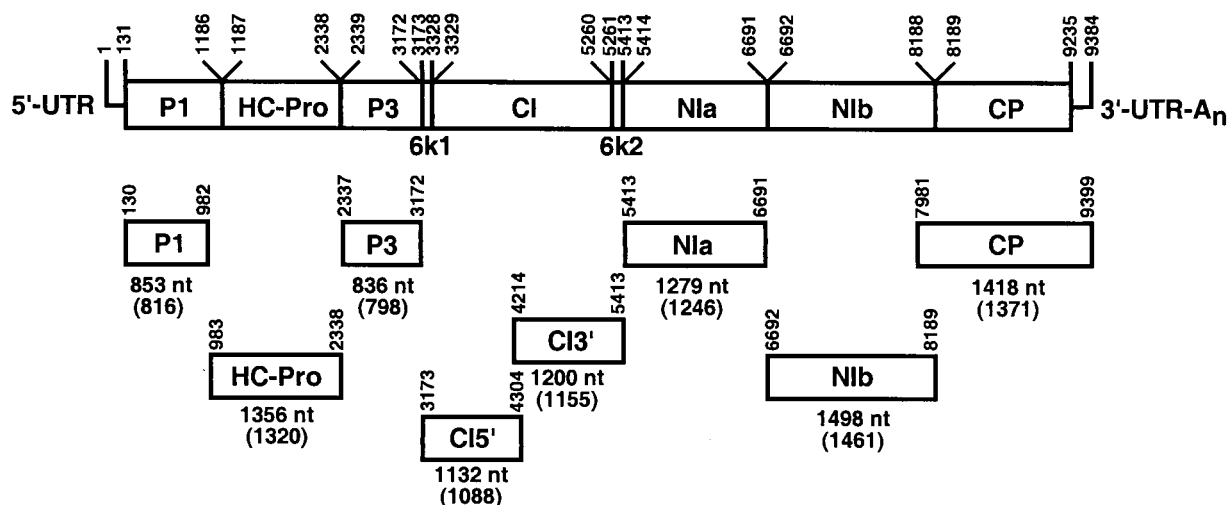


FIG. 1. Physical map of the WSMV genome depicting locations of RT-PCR products amplified for analysis of genotypic variation. Arrangement of cistrons within the polyprotein open reading frame bracketed by upstream (5'-UTR) and downstream (3'-UTR) untranslated regions is presented at the top. The 3'-terminal polyadenylated tail of variable length is designated A_n. Nucleotide coordinates (vertically oriented numbers) correspond to the genomes of both Type and Sidney 81 strains of WSMV, whose genomes are identical in length and may be aligned without gaps. Genomic location and nucleotide lengths of RT-PCR products are presented below the complete genome map. Nucleotide sequence lengths analyzed for nucleotide polymorphism are indicated in parentheses. The difference between total and analyzed sequence lengths represents terminal regions where primers used in PCR annealed.

RESULTS

Accuracy of SSCP for detection of nucleotide substitutions. SSCP has been used to characterize variation in both nuclear genes and virus populations (7, 19, 37, 43, 61, 68). This technique is dependent on a bulk property of PCR products, so that nucleotide variation due to misincorporation by *Taq* DNA polymerase does not interfere with the analysis. Electrophoresis at two different temperatures has been reported to increase the accuracy of this technique in detecting single nucleotide polymorphisms to ~98% (68). To evaluate accuracy as applied to WSMV, SSCP analysis was performed with PCR products amplified from 20 cloned templates of known sequence. Of the 20 plasmids examined, 12 had one or more unique nucleotide changes (defined as sporadics by Smith et al. [58]) whereas 8 did not. Each clone free of sporadics yielded an SSCP pattern identical to the original PCR product from which it was cloned. All but one clone bearing sporadic substitutions yielded novel SSCP patterns, as expected. Based on this analysis, SSCP was estimated to be 95% (19 of 20) accurate for detection of nucleotide substitutions (data not shown). Thus, each unique haplotype identified by SSCP could be correlated with specific sequences for all 47 LDSIs.

Low-MOI inoculation generates LDSIs with different consensus sequences. Four regions of each LDSI genome were amplified by RT-PCR (Fig. 1), with haplotypes detected by SSCP for the CP cistron presented in Fig. 2 and summarized for all four cistrons in Table 2. The predominant haplotypes for each cistron have been designated A (Sidney 81) and D (Type). Three different haplotypes each were detected in the Sidney 81 CP and P1 cistrons. Sidney 81 haplotypes B (CP) and C (P1) were present only together in the same LDSIs. Two Sidney 81 haplotypes were detected in NIa, while no variation was found in P3 of Sidney 81. Type had two variant haplotypes that each

occurred once, one in the CP cistron and one in the P3 cistron. The occurrence of haplotype configurations across the four cistrons is shown in Table 2. For Sidney 81, five distinct linkage patterns were detected, with 55% of the LDSIs bearing at least one cistron with a variant haplotype. In contrast, 92% of Type LDSIs displayed the predominant haplotype for all four cis-

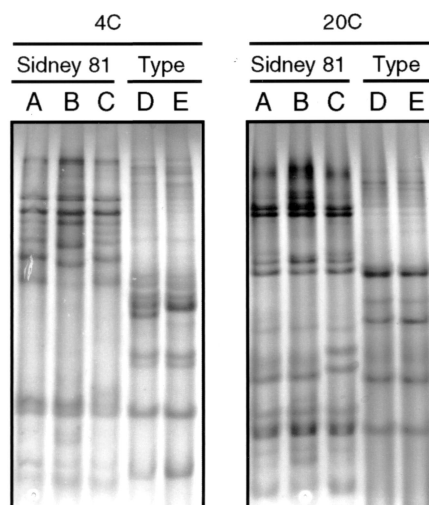


FIG. 2. SSCP analysis of the WSMV coat protein cistron. Presented are SSCP profiles for different haplotypes (A to E) observed among LDSIs of the Sidney 81 and Type strains. Denatured, restriction endonuclease-digested PCR products were electrophoresed under non-denaturing conditions in 10% polyacrylamide gels at 4 and 20°C. Alteration of the SSCP pattern at one or both electrophoresis temperatures corresponded to single-nucleotide substitutions in variant haplotypes (B, C, and E) relative to the predominant SSCP patterns observed among Sidney 81 (A) or Type (D) LDSIs.

TABLE 2. Occurrence of haplotype configurations among LDSIs of WSMV Sidney 81 and Type as determined by SSCP

Configuration	Haplotypes ^a				No. of LDSIs ^b
	P1	P3	NIa	CP	
Sidney 81					
1	A	A	A	A	10
2	C	A	A	B	2
3	A	A	A	C	2
4	B	A	B	A	2.5
5	B	A	A	A	5.5
Type					
1	D	D	D	D	23
2	D	D	D	E	1
3	D	E	D	D	1

^a Letters correspond to different SSCP patterns.

^b Fractional frequencies due to Sidney 81 LDSI 9 NIa SSCP being a mixture of haplotypes A and B.

trons. Sidney 81 LDSI-S9 yielded an SSCP pattern for NIa that appeared to be a mixture of haplotypes A and B (Table 2), an interpretation supported by the sequences of two clones (Table 3). To preclude stochastic variation among samples arising during early stages of amplification, RT-PCR-SSCP analysis was repeated for all variant haplotypes as well as several of the predominant haplotypes for each cistron. In every case, the SSCP patterns were consistent (data not shown).

PCR products derived from five LDSIs per cistron (representing all haplotypes) were cloned for both Type and Sidney 81. For all variant SSCP patterns observed among the LDSIs, both clones derived from a PCR product contained the same nucleotide substitution at one position (Table 3). In every case, a shift in SSCP pattern at one or both electrophoresis temperatures resulted from a single-nucleotide change in the consensus sequence. Of seven polymorphic nucleotide sites found in the two LDSI populations, five were transitions and two were transversions. One polymorphism was in the 3'-untranslated region of Sidney 81. Within the polyprotein coding region, four synonymous and two nonsynonymous changes occurred (Table 3).

Nonsporadic substitutions present in the LDSI sequence data set were associated with unique haplotypes as identified by SSCP, whereas sporadic substitutions were not. Among 84,620 nucleotides sequenced from LDSI clones, 79 sporadic

TABLE 3. Consensus sequence nucleotide substitutions among LDSIs of WSMV Sidney 81 and Type

Strain	PCR product	Change from predominate haplotype ^a			
		Haplotype	nt	Codon	Amino acid
Sidney 81	P1	C	928	GAU to GAC	Asp to Asp
Sidney 81	P1	B	961	UUC to UUU	Phe to Phe
Sidney 81	NIa	B	6016	UCA to UCG	Ser to Ser
Sidney 81	CP	B	8246	GCA to UCA	Ala to Ser
Sidney 81	CP	C	9254	C to U	3'-UTR ^b
Type	P3	E	2699	CUG to UUG	Leu to Leu
Type	CP	E	8319	GGA to GUA	Gly to Val

^a The predominant haplotype defined by SSCP is haplotype A (Sidney 81) or D (Type).

^b 3'-UTR, 3' untranslated region.

TABLE 4. Sporadic nucleotide changes in clones derived from LDSIs of WSMV Sidney 81 and Type

Strain	PCR product	No. of sporadic changes/nt sequenced (10 clones each)	No. of sporadic changes/nt/cycle of RT-PCR
Sidney 81	P1	4/8,160	0.14×10^{-4}
Type	P1	8/8,160	0.27×10^{-4}
Sidney 81	P3	5/7,980	0.17×10^{-4}
Type	P3	7/7,980	0.24×10^{-4}
Sidney 81	NIa	14/12,460	0.36×10^{-4}
Type	NIa	16/12,460	0.41×10^{-4}
Sidney 81	CP	13/13,710	0.26×10^{-4}
Type	CP	12/13,710	0.24×10^{-4}
Totals		79 ^a /84,620	0.27×10^{-4b}
NIa reclone		15/12,460	0.40×10^{-4}

^a 69 transitions, 9 transversions, and 1 single-base deletion.

^b Weighted average.

changes were noted (Table 4). All polymerases have intrinsic error rates, and if these are not accounted for, an overestimation of variability is the result (3, 58). Although SSCP and sequencing were sufficient to identify differences in consensus sequences among LDSIs, this procedure was not suitable for discrimination of nonconsensus variations resident within LDSIs from *Taq* polymerase errors introduced during PCR. Because sporadic rates among the LDSI data sets and the NIa reclone experiment were similar (Table 4), partitioning of sporadic substitutions as real (produced in vivo by viral polymerase) or artifactual (produced in vitro by RT or *Taq* polymerase) variation was not attempted.

Passage at high MOI also results in consensus sequence change. The linkage pattern (BABA) determined for the P1, P3, NIa, and CP cistrons of LDSI-S10 was shared by two other LDSIs (Table 2). The LDSI-S10 P3 and CP cistrons displayed the predominant haplotype observed among Sidney 81 LDSIs that corresponded to Sidney 81 sequence AF057533. Although LDSI-S10 P1 was characterized by SSCP as haplotype B, sequencing revealed that haplotype A also was present as a minority of the population. The LDSI-S10 NIa haplotype B represented a variant (G₆₀₁₆) that also occurred in two other LDSIs (Table 3). SSCP analysis performed on HC-Pro, CI5', CI3', and NIB PCR products defined haplotypes for these LDSI-S10 cistrons. Sequence comparisons of LDSI-S10 clones for each of these additional PCR products revealed no nonsporadic differences between LDSI-S10 and Sidney 81 sequence AF057533.

SSCP analysis of all eight PCR products for each of the nine lineages derived from LDSI-S10 was performed at passages 1, 3, 6, and 9 (Table 5). Sequences of cloned PCR products that displayed variant SSCP patterns revealed nonsporadic substitutions, whereas sequences of cloned PCR products with SSCP patterns identical to that of LDSI-S10 did not. Nonsporadic nucleotide substitutions corresponding to variant haplotypes identified by SSCP (Table 5) are specified in Table 6. The passage data set (including LDSI-S10) consisted of 191,299 nucleotides sequenced. The total number of nucleotides screened for substitutions by SSCP was considerably larger.

After nine passages, no WSMV lineage maintained the original LDSI-S10 haplotype pattern (Table 5). Shifts in consensus

TABLE 5. Haplotypes determined by SSCP and verified by sequencing of two clones at select passages for nine lineages derived from LDSI-S10

Gene	Founding inoculum LDSI-S10 (passage 0)	SSCP haplotype ^a for lineage:								
		Wheat			Corn			Barley		
		WA	WB	WC	CA	CB	CC	BA	BB	BC
P1	<u>B</u> ^b	B,B, <u>B</u> ,B	B,B, <u>B</u> ,B	B,B, <u>B</u> ,B	B,B, <u>B</u> ,B	<u>E</u> , <u>E</u> , <u>E</u> , <u>E</u>	B,B, <u>A</u> ,B	B,B, <u>B</u> ,B	B,B, <u>B</u> ,B	B,B, <u>B</u> ,B
HC-Pro	<u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A, <u>A</u> ,A	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>
P3	<u>A</u>	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A
CI5'	<u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>	<u>B</u> , <u>A</u> , <u>B</u> , <u>B</u>	A,A,A, <u>A</u>	A,A,A, <u>A</u>
CI3'	<u>A</u>	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A, <u>B</u> , <u>B</u> , <u>B</u>	A,A, <u>A</u> ,A
N1a	<u>B</u>	B,B, <u>A</u> ,A	B,B, <u>C</u> ,A	B,B, <u>A</u> ,A	B,B, <u>A</u> ,A	B,B, <u>A</u> ,E	F,F,F,F	B,B, <u>A</u> ,A	B,B, <u>G</u> ,G	B,B, <u>A</u> ,A
NIb	<u>A</u>	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,G	A,A, <u>B</u> , <u>B</u>	A,A, <u>C</u> ,A	A,A, <u>A</u> ,A	A,A, <u>A</u> ,A
CP	<u>A</u>	A, <u>A</u> , <u>F</u> , <u>F</u>	A, <u>A</u> , <u>N</u> , <u>K</u>	A, <u>A</u> , <u>G</u> , <u>M</u>	A, <u>A</u> , <u>H</u> , <u>I</u>	A, <u>A</u> , <u>A</u> , <u>J</u>	A, <u>A</u> , <u>A</u> , <u>A</u>	A, <u>A</u> , <u>A</u> , <u>L</u>	A, <u>A</u> , <u>A</u> , <u>A</u>	A, <u>A</u> , <u>A</u> , <u>A</u>

^a Haplotypes are designated by letters; the four letters in each group are for passages 1, 3, 6, and 9, respectively. The underlines indicate haplotypes verified by sequencing of two clones.

^b LDSI-S10 is also known to contain P1 haplotype A as a minor variant.

sequence (i.e., new alleles) were detected in all cistrons except HC-Pro and P3 (Table 6). However, most changes in consensus sequence occurred within two cistrons, with CP yielding 12 allelic shifts at 11 sites and NIa having 15 events at 7 sites. While the barley BC lineage had only one change, most lineages accumulated multiple changes in consensus sequence. The wheat WA and corn CB lineages accumulated the largest number of nucleotide changes (five each) in their respective consensus sequences after nine passages. The total accumulation of nucleotide substitutions in consensus sequences at each passage is illustrated in Fig. 3. Several changes in predominant haplotypes occurred in passage 1 (P1 of corn CB; NIa and CI of corn CC), although most consensus sequence changes were detected in passages 6 and 9. Some variant haplotypes gained transitory prominence, only to be subsequently replaced by the original haplotype sequence (NIa of wheat WB and NIb of barley BA), or disappeared, only to reappear in later passages (CI of corn CC). Several variant SSCP patterns were mixtures of haplotypes (NIa of wheat WB passage 6), in which two alleles at the same polymorphic site were detected by nucleotide sequencing and were probably present at similar frequencies such that both contributed to the novel SSCP pattern observed.

Consensus sequence changes shared by more than one lineage derived from LDSI-S10. Consensus sequence divergence accrued by and surviving in passage 9 is illustrated in Fig. 4. While most lineages accumulated unique changes in consensus sequence, several exceptions were noted. Eight of the nine lines had a common substitution relative to LDSI-S10. This change within NIa ($G_{6016} \rightarrow A$) was at a known polymorphic site in the LDSI data set and restored the consensus sequence of all but one lineage to the allele that predominated among Sidney 81 LDSIs (Tables 2 and 5). Since both alleles were present in the Sidney 81 LDSIs (with one LDSI a known mixture), it is possible that A_{6016} also was present in LDSI-S10 as a minority variant. However, a partial sequence (nt 5896 to 6427) determined for 20 additional clones of the LDSI-S10 NIa PCR product did not detect A_{6016} .

A second example of a substitution shared among lineages was $U_{6418} \rightarrow C$ in NIa. This change occurred in passage 1 of corn CC and remained stable in this lineage through passage 9. C_{6418} also was detected as a mixture with U_{6418} , resulting in a

novel SSCP pattern seen only in wheat WB at passage 6. However, C_{6418} was not detected in wheat WB at passage 9 and the SSCP pattern corresponded to haplotype A (bearing only the near-global substitution $G_{6016} \rightarrow A$). All 20 additional NIa clones of LDSI-S10 partially sequenced (nt 5896 to 6427) had U_{6418} and not C_{6418} .

While wheat lineages WB and WC both contained the adjacent substitutions $C_{8764} \rightarrow U$ and $A_{8765} \rightarrow C$, first detected at passage 6, these two shared substitutions became the consensus sequence only transiently at passage 6 in wheat WC. At passage 9, both alleles at positions 8764 and 8765 were detected in both lineages, such that the haplotypes observed were due to mixtures. Neither U_{8764} or C_{8765} was detected among 17 LDSI-S10 CP clones of an EHFS-PCR product.

Polymorphism revealed by sequencing but not detected by SSCP. Although SSCP is a powerful tool for detecting nucleotide substitutions, the procedure is not 100% accurate. Therefore, some consensus sequence nucleotide substitutions may have escaped notice. To investigate this possibility, a number of PCR products displaying the predominant haplotype in the LDSI data set or the same haplotype as LDSI-S10 in the passage data set were cloned and sequenced for each cistron. Within the LDSI data set, one discrepancy was noted. The LDSI-S10 P1 PCR product was characterized as haplotype B by SSCP (Table 2). However, the sequences of two clones indicated that both haplotypes A (C_{961}) and B (U_{961}) were present. The simplest explanation is that both alleles were present in LDSI-S10 but that U_{961} comprised the majority such that haplotype B dominated the SSCP pattern.

Examination of nonconsensus sequence variation. In addition to consensus sequence shifts, the amount and pattern of sequence variation within lineages may have changed over time. To examine sequence heterogeneity present within viral populations before and after nine passages, the CP cistrons of LDSI-S10 and lineages wheat WB, barley BA, and corn CC at passage 9 were amplified using EHFS-PCR, cloned, and sequenced. After 30 cycles of EHFS-PCR with a DNA template of known sequence, the observed mutation frequency was 2.2×10^{-4} /nt. The error rate calculated for EHFS-PCR (0.72×10^{-5} /nt/cycle) was 5.6-fold lower than the error rate obtained with *Taq* polymerase (0.4×10^{-4} nt/cycle). These PCR error rates were similar to those reported by others (3, 57,

TABLE 6. Nucleotide substitutions detected among lineages passed from LDSI-S10

PCR product	Nucleotide substitution ^a at:				Amino acid change
	Passage 1	Passage 3	Passage 6	Passage 9	
Wheat lineage WA					
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
	X	X	G* ₅₇₄₃ →A*	G ₅₇₄₃ →A	Met→Ile
	X	X	G* ₆₂₉₈ →A*	G ₆₂₉₈ →A	Arg→Arg
CP	X	X	U* ₈₁₉₄ →C*	U ₈₁₉₄ →C	Ser→Ser
	X	X	A ₈₉₆₄ →G	A ₈₉₆₄ →G	Lys→Arg
Wheat lineage WB					
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
	X	X	U* ₆₄₁₈ →C*	X	His→His
CP	X	X	X	C ₈₅₅₆ →U	Ala→Val
	X	X	C* ₈₇₆₄ →U*	C* ₈₇₆₄ →U*	Cys→Cys
	X	X	A* ₈₇₆₅ →C*	A* ₈₇₆₅ →C*	Ile→Leu
Wheat lineage WC					
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
CP	X	X	A ₈₂₄₈ →U	A* ₈₂₄₈ →U*	Ala→Ala
	X	X	C ₈₇₆₄ →U	C* ₈₇₆₄ →U*	Cys→Cys
	X	X	A ₈₇₆₅ →C	A* ₈₇₆₅ →C*	Ile→Leu
Corn lineage CA					
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
CP	X	X	G ₈₄₂₃ →A*	G ₈₄₂₃ →A	Asn→Asn
			C ₉₂₅₉ →U	C ₉₂₅₉ →U	3'-UTR
Corn lineage CB					
P1	C ₃₂₅ →U	C ₃₂₅ →U	C ₃₂₅ →U	C ₃₂₅ →U	Asn→Asn
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
	X	X	X	G ₆₁₂₇ →A	Met→Ile
NIb	X	X	X	C ₇₃₁₈ →U	Phe→Phe
CP	X	X	X	C ₈₈₈₀ →U	Ala→Val
Corn lineage CC					
P1	X	X	U ₉₆₁ →C	X	Phe→Phe
CI	C ₃₄₆₆ →U	X	C ₃₄₆₆ →U	C ₃₄₆₆ →U	Ile→Ile
NIa	U ₆₄₁₈ →C	U ₆₄₁₈ →C	U ₆₄₁₈ →C	U ₆₄₁₈ →C	His→His
NIb	X	X	C ₇₆₂₁ →U	C ₇₆₂₁ →U	Gly→Gly
Barley lineage BA					
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
NIb	X	X	C ₈₀₁₄ →U	X	Phe→Phe
CP	X	X	X	A ₈₅₀₂ →T	Gly→Gly
Barley lineage BB					
CI	X	A ₄₃₈₄ →G	A ₄₃₈₄ →G	A ₄₃₈₄ →G	Ser→Ser
NIa	X	X	U ₅₆₆₈ →C	U ₅₆₆₈ →C	Gln→Leu
	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser
Barley lineage BC					
NIa	X	X	G ₆₀₁₆ →A	G ₆₀₁₆ →A	Ser→Ser

^a X, no change from LDSI-S10; *, both nucleotides detected in the same lineage and passage.

66). Error due to RT was accounted for by assuming a maximum misincorporation rate of $0.36 \times 10^{-4}/nt$ (21), to give a combined in vitro mutation frequency estimated to be $2.6 \times 10^{-4}/nt$. Note that although error introduced by RT occurs during the first stage of amplification, RT errors are independent and therefore still yield primarily sporadic substitutions after PCR amplification, as exemplified by the absence of non-sporadic substitutions in the LDSI-S10 population (Fig. 5) and the entire data set of Schneider and Roossinck (52). Sixteen sites of nucleotide substitutions were found in the LDSI-S10 population sample. Employing a novel maximum-likelihood

test of template homogeneity devised by Weiss and von Haeseler (67), the probability that these were all due to PCR misincorporation was estimated to be less than 0.02%. Thus, 63 to 76% of sporadic substitutions present in sequence data sets obtained using RT followed by EHFS-PCR could not be accounted for by in vitro polymerase error and therefore probably represented substitutions (here termed singletons) due to viral polymerase error (Table 7; see Fig. 6).

None of the singletons observed in LDSI-S10 EHFS-PCR-derived CP clones were present in any of the three passage 9 populations examined (Fig. 5 and data not shown). Nonethe-

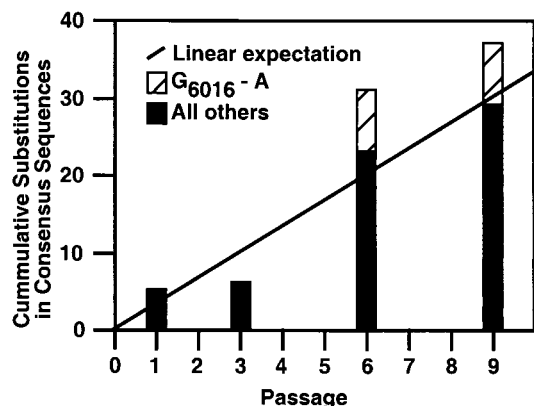


FIG. 3. Histogram of cumulative substitutions in consensus sequence summed for nine lineages derived from an LDSI of WSMV Sidney 81. The total bar height corresponds to the frequency of all substitutions in consensus sequences sampled after 1, 3, 6, and 9 passages. Black regions of bars correspond to all consensus sequence substitutions other than the near-global reversion ($G_{6016} \rightarrow A$) which occurred in eight of nine lineages (hatched region of bars). The line denotes the linear-expectation best fit to the observed data (excluding $G_{6016} \rightarrow A$) that intersects the origin (no mutations at time of limiting dilution inoculation).

less, after nine passages, corrected mutation frequencies in three lineages passaged in different hosts did not significantly differ from that of LDSI-S10 (Table 7). There was, however, a qualitative difference in the composition of mutants before and after passage, as exemplified by Fig. 5. All polymorphic sites in the LDSI-S10 sample were singletons whereas nucleotide substitutions after nine passages included both singletons and non-singletons shared among several clones (Table 7; Fig. 5). Non-singletons were shared among two independent RT-PCR samples from a single lineage at passage 9 (Fig. 5), demonstrating that they are not in vitro artifacts. Although the proportion of unique substitutions (sporadic substitutions) in a

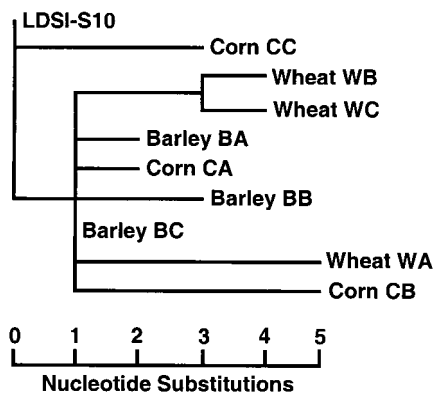


FIG. 4. Cladogram reflecting consensus sequence substitutions accrued by and surviving in nine WSMV lineages after nine passages in wheat, corn, or barley. The founding inoculum for all lineages was derived from LDSI-S10. The tree was drawn manually, with common nodes reflecting consensus sequence substitutions shared by more than one lineage at passage 9. Branch lengths correspond to the number of nucleotide substitutions (1 through 5) in the consensus sequence of each lineage relative to LDSI-S10 after nine passages.

population may be designated as resulting from viral polymerase (singletons) or PCR error, the origin of any individual sporadic mutant cannot be determined. However, polymorphism shared among clones (nonsingletons) is unlikely to be an artifact of PCR (58), particularly if the same substitution appears in independent RT-PCR products from a sample (Fig. 5). Thus, the shared substitutions probably represent viral polymerase-induced errors that have replicated and reached detectable frequencies. Among the shared substitutions, five were transitions (four C to U and one A to G) and four were transversions (two each A to U and A to C). Three of these were silent, and four resulted in amino acid substitutions (data not shown).

Allelic frequencies of polymorphic sites within sampled clones of LDSI-S10 and each passage 9 population are presented in Fig. 6. For example, if a nucleotide substitution was found in three clones of a sample, it was placed in allele size class 3. The distributions are folded in that for i shared substitutions from a sample size of n , it is the smaller of i or $n - i$ that is plotted. For comparison, expected values for three theoretical sampling distributions are presented in Fig. 6: (i) a Poisson distribution, (ii) the neutral infinite-sites model (63, 66), and (iii) the Luria-Delbrück sampling distribution (29, 31, 47) applied to nucleotide variation (54). Only LDSI-S10 was in concordance with a Poisson distribution. All of the passage 9 samples had ranges of allele size classes inconsistent with the Poisson distribution yet had singleton (allele size class of 1) frequencies in excess of those

TABLE 7. Nucleotide polymorphism^a in the CP cistron before and after nine passages in wheat, barley, or corn

PCR product ^b	No. of clones sequenced	No. of polymorphic sites/no. of nt ^c	Mutation frequency/nt		
			Uncorrected total	Corrected ^d total	Corrected ^d singleton
Reclone	10	3/13,710	2.2×10^{-4}	0 ^e	0 ^e
LDSI-S10	17	16/23,307	6.9×10^{-4}	4.3×10^{-4}	4.3×10^{-4}
9BA	13	19/17,823	10.7×10^{-4}	8.1×10^{-4}	6.4×10^{-4}
9WB	20	23/27,420	8.4×10^{-4}	5.8×10^{-4}	4.0×10^{-4}
9CB	15	20/20,565	9.7×10^{-4}	7.1×10^{-4}	6.6×10^{-4}

^a Determined by sequencing of clones derived from PCR products amplified by EHFS using a DNA template of known sequence.
^b LDSI-S10 shows polymorphism before passages; 9WB, 9BA, and 9CB show polymorphism after nine passages in wheat, barley, and corn, respectively.
^c All identical nonsingletons treated as a single mutation event at one site.
^d Sporadic changes attributed to combined error (2.6×10^{-4} /nt) due to RT (0.36×10^{-4} /nt) and EHFS-PCR error (2.2×10^{-4} /nt) were subtracted to yield corrected mutation frequencies/nucleotide.
^e All polymorphism due to in vitro polymerase error during PCR.

lowed by RT and PCR (15 cycles with *Pfu* polymerase). If only 20% of the composite mutation frequency was due to in vitro transcription, the RNA polymerase error rate equals $\sim 1 \times 10^{-5}$ /nt. In vitro transcription using an RNA polymerase with an error rate of $\sim 1 \times 10^{-5}$ /nt would result in many single-substitution mutants (for WSMV, a total of 28,152 are possible) represented in the initial inoculum, with an average of ~ 0.1 nucleotide substitution per viral genome transcribed. Although the frequency of each individual mutant is low, collectively the sum of all mutants generated would average ~ 1 mutation per 10 complete genomes transcribed. Note that the use of a 35S promoter construct to deliver the cloned virus genome as a DNA template for in vivo transcription (as for tobacco mosaic virus [TMV] in reference 52) does not eliminate variability (which in this case is equal to the error rate of the host RNA polymerase) but, rather, only delays the time when variation is generated. Although selection for infectious genomes occurs on inoculation and is expected to reduce primary variation, this would occur with both cloned and uncloned inocula.

Use of an infectious clone does not preclude rapid establishment of heterogeneity (52) after multiple rounds of virus replication and systemic movement in planta. Note that all LDSI-S10 mutants were singletons (Fig. 5) and that the corrected singleton mutation frequency per nucleotide obtained for LDSI-S10 (4.3×10^{-4}) (Table 7) was similar to the mutation frequency per nucleotide (albeit uncorrected for RT-PCR error) of cucumber mosaic virus (CMV) (4.6×10^{-4}) or TMV (5.2×10^{-4}) on systemic infection established with transcripts derived from cloned DNA (52). While neither method guarantees exclusion of all variation (an absolute bottleneck), both limit genetic diversity at inoculation and yield similar levels of variation after systemic infection. Finally, because the purpose of the LDSI and passage experiments was to monitor changes in the consensus sequence at specific points in time, it did not matter whether all or only some of the variation arose after a bottleneck event.

Variant haplotypes among LDSIs probably reflect preexisting variations in Type and Sidney 81. Differences in haplotype patterns among LDSIs are most easily explained as preexisting variations resident within Type and Sidney 81 that were biologically separated by limiting-dilution inoculation. This explanation is particularly relevant to the four variant haplotypes identified among Sidney 81 LDSIs, each of which was observed in at least two LDSIs (Table 2). Over half of the genomes

resident within the Sidney 81 culture bore at least one substitution, relative to the consensus sequence corresponding to the predominant haplotype pattern. However, since variation among Sidney 81 LDSIs was determined for only four cistrons,

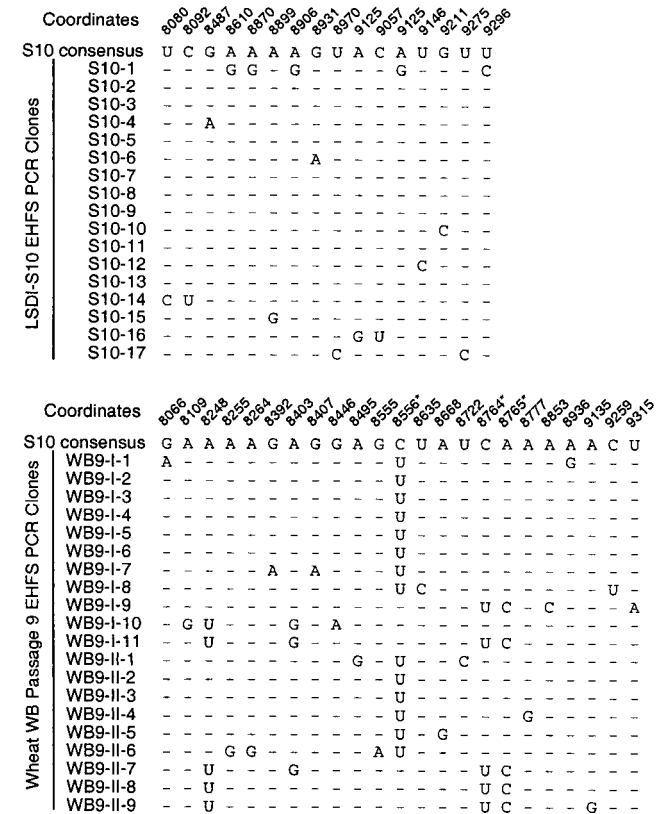


FIG. 5. Nucleotide sequence variation in the WSMV coat protein cistron and 3'-untranslated region before (LDSI-S10) and after (wheat WB) nine passages in wheat. Sequences were determined from multiple clones derived from RT-PCR products amplified with the EHFS employing a proofreading polymerase. Only positions at which polymorphisms were observed are shown. WSMV genomic coordinates corresponding to each polymorphic site are listed above each sequence data set. Substitutions present within clones are identified by specific nucleotide changes from the LDSI-S10 (S10) consensus sequence. Dashes indicate positions within clones that are identical to the LDSI-S10 consensus sequence. Note that all nonsingleton substitutions in the wheat WB passage 9 data set occur in clones derived from two independent PCR products (I and II).

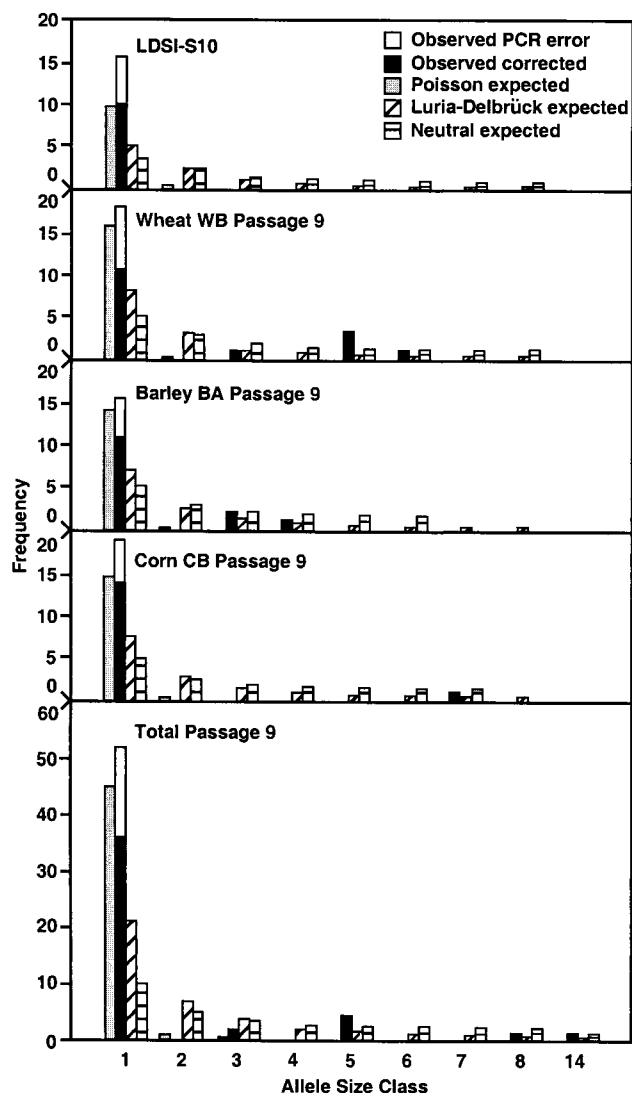


FIG. 6. Distribution of alleles size classes of variants detected among EHFS PCR clones of the coat protein cistron and 3'-untranslated region of WSMV before (LDSI-S10) or after nine passages in wheat, barley, or corn. Black bars indicate the frequency of allele size classes corrected for the proportion of singletons estimated to be derived from in vitro RT-PCR error (open bars). Shaded or hatched bars denote allele size class frequencies expected for each of three theoretical distributions. The expected frequencies for the Luria-Delbrück distribution were calculated using equations 17 and 16 in reference 47, and those for the neutral, infinite alleles distribution were calculated using equation 51 in reference 62. Note that the LDSI-S10 variation is composed only of singletons whereas the passage 9 data sets contain both singletons and nonsingletons. Two novel size classes (classes 8 and 14) appear in the total passage 9 data set and represent cases where the majority allele within a single data set (wheat WB allele size class 5 and corn CB allele size class 7) become the minority allele in the total data set.

representing ~45% of the genome, this represents a minimum estimate and does not account for singleton variation, which could not be distinguished from PCR artifacts in the LDSI experiment.

Fewer variant haplotypes were separated as LDSIs by limiting-dilution inoculation of the Type culture (Table 2). This

difference could have resulted from random probability of sampling and/or arbitrary choice of genomic regions examined in the LDSI experiment. Differences in within-isolate genetic diversity also were observed for citrus tristeza virus (25). It is unlikely that the Type genome encodes a viral polymerase with greater fidelity or is subject to greater negative selective pressure than Sidney 81; however, these hypotheses cannot be excluded. Given the results of the passage experiment, in which nonsingleton variation increased over time, the difference in variant haplotype pattern frequencies among LDSIs of Sidney 81 and Type may reflect differences in the population growth histories of the two cultures. Although the laboratory propagation histories of the two WSMV strains are not completely known, Type was subjected to a bottleneck event by mite transmission (4) prior to storage as dried infected tissue and was passaged in the laboratory at high MOI fewer times than was Sidney 81 prior to the beginning of our experiments. Thus, sampling at different stages of population expansion following a genetic bottleneck could readily explain differences in the polymorphism frequency between Type and Sidney 81.

In the LDSI experiment, a deliberate bottleneck was introduced. The outcome measured was essentially that of joint probability; i.e., given low MOI, how often were differences in the consensus sequence of WSMV detected? Type and Sidney 81 vary by about 0.02/nt (~2%), confirming that WSMV can tolerate multiple substitutions, and the preponderance of synonymous substitutions within and between these isolates are compatible with the idea that most of the observed polymorphisms among consensus sequences are essentially neutral (9). Nonsingleton variation uncovered within the two LDSI populations was limited, with mean pairwise differences (π) ranging from 0 to 7.7×10^{-4} /nt among the four cistrons examined. The average π across all four cistrons was 2.9×10^{-4} /nt (Sidney 81) or 0.4×10^{-4} /nt (Type). These are 2 orders of magnitude smaller than the nucleotide diversity ($\pi \approx 2 \times 10^{-2}$) estimated for WSMV field populations (8, 33).

Selection explains one shared substitution during passage.

The near-global reversion ($G_{6016} \rightarrow A$) represented the sole example of a change most consistent with selection. Since both alleles at this polymorphic site yield synonymous codons (Fig. 3), a selective advantage of A over G at this position could reflect RNA structural features or codon bias during translation. The eight lineages with A_{6016} at passage 9 have a common node in a phylogenetic analysis (Fig. 4). Since G_{6016} was present in all 22 LDSI-S10 NIa clones sequenced, A_{6016} was either absent from LDSI-S10 or present at low frequency (<5%). If A_{6016} was absent from LDSI-S10, then the substitution $G_{6016} \rightarrow A$ occurred de novo in each lineage, such that the common node shared by all lineages (except corn CC) represents a point of convergence rather than divergence. Regardless of whether A_{6016} occurred at a low level in LDSI-S10 or was generated de novo in each lineage, a selective advantage of A over G at position 6016 remains the simplest explanation for reversion to A_{6016} in eight of nine lineages by passage 6.

LDSI-S10 was known to be a mixture at nucleotide position 961 in P1, with both U and C present. Throughout the passage experiment, the haplotype associated with C_{961} was detected by SSCP only once in corn CC at passage 6 (Table 5). Sequences of two clones from each passage 6 P1 PCR product for all nine lineages indicated that only the two clones derived

from corn CC had C₉₆₁. Thus, unlike A₆₀₁₆, there did not appear to be selective pressure for reversion to C₉₆₁, the nucleotide most common at this position among Sidney 81 LDSIs.

It was anticipated that changes in consensus sequence driven by host-specific selection might be revealed as identical changes in all lineages passaged in the same host species. Although the adjacent substitutions C₈₇₆₄→U and A₈₇₆₅→C occurred in two wheat lineages, it is unlikely that this represents an example of host-specific selection, since U₈₇₆₄ or C₈₇₆₅ did not appear in all three wheat lineages and did not successfully displace the original alleles in any lineage at passage 9 (Tables 5 and 6). Thus, no clear example of host-specific selection was observed, a result similar to those reported with TMV virus passaged in different host species (23). While a signature of negative selection (fewer substitutions in a certain region of the genome) was detected among divergent strains of WSMV (9), too few substitutions accrued during passage to perform similar analyses. Although the reversion G₆₀₁₆→A was probably due to selection, occurrence in all three host species suggests that selective forces operating here did not vary among the three host species tested.

Most newly generated mutations are sequestered in virions and have yet to be tested for fitness. Variation within a plant virus isolate may be separated into two classes: (i) new mutants that have yet to serve as replication templates and (ii) older mutants that have been replicated. The RNA genome contained in each virion was transcribed independently from a minus-strand template by the viral polymerase with some probability of nucleotide misincorporation. By definition, any “new” mutation appearing in a sample will be a singleton. The sample frequency of all singletons representing new mutations is dependent solely on the viral polymerase error rate. Because most positive strands become encapsidated and thereby are removed from the pool of replicating templates (14), the bulk of viral RNA sampled for variation represents end products of replication. If this is so, most variation is sequestered and is not part of the effective population (that fraction of genomes giving rise to progeny). Thus, a substantial fraction of variation generated *de novo* in a plant cell has not been subject to selection, the essential *cis*-acting requirements of encapsidation notwithstanding. Therefore, the virion population could include deleterious or lethal mutations that have yet to be tested for fitness. The observation that 3 of 16 satellite tobacco mosaic virus (STMV) cDNA clones bearing nucleotide substitutions yielded noninfectious transcripts (27) supports the “sequestered and yet to be tested” hypothesis. A corollary to this hypothesis is that the error rate of the viral polymerase will be no greater than the per-nucleotide frequency of singletons observed in a population.

Although most new mutants are sequestered by encapsidation and are not immediate objects of selection, there remains an evolutionary advantage for high viral RNA polymerase error rates. Constant production of mutants ensures that alternative genotypes are poised to take advantage of new environments or selective conditions. A fraction of new mutants must serve as replication templates, otherwise populations could not evolve. Furthermore, encapsidated genomes are available to initiate new infections upon transmission, such that the pool of sequestered variation is not an evolutionary dead end.

Replicated mutants have been subjected to evolutionary processes. Only mutants that grow (serve as replication templates and move) are subjected to selection and drift within a plant, and those that survive lead to a distribution of shared “older” polymorphisms in the sample. Examples of older mutants identified in this study include the variant haplotypes observed among Sidney 81 LDSIs and nonsingleton variants detected after passage. In contradistinction to most other organisms with DNA genomes, RNA virus growth is not semi-conservative and has a significant linear component (14) that is not accounted for by population growth models based on binary (exponential) growth. The linear aspect of virus growth should have a profound quantitative effect on population divergence and polymorphism, which warrants future theoretical examination.

The distribution of allele size classes (Fig. 6) expected would be a composite of the Poisson distribution (singletons) and another distribution reflective of underlying parameters of mutant growth, selection, and drift. Differences in allele size class distribution before and after passage were due to the accumulation of nonsingleton mutants. All variants in LDSI-S10 were singletons, consistent with a Poisson distribution. This is not unexpected, since bottlenecks by limiting-dilution inoculation should have eliminated most preexisting variation. On passage at high MOI, the accumulation of nonsingleton variants shifted the population away from a Poisson distribution, although the frequency of singletons remained higher than that predicted from population models for which probability distributions may be calculated: the constant-sized, infinite-sites model (64), or a model incorporating population growth (28, 30, 51). The latter predicts more singleton sites than does the infinite-site model but clearly not as many as were actually observed (Fig. 6). Moreover, singleton sites were randomly distributed among WB passage 9 clones, which also contained sites with moderate allele class sizes (Fig. 5). It is difficult to imagine a population model that retains some sites with moderate allele class sizes while predicting many more singleton sites unless two different distributions are superimposed. The excess of singletons is consistent with the hypothesis that most genomes in this allele size class represent end products of replication.

The two classes of mutations (singleton and nonsingleton) observed in our data set are not unique to WSMV. Heterogeneity within STMV and CMV satellite RNA also partitioned into these two classes (26–28). The lack of nonsingleton variation observed (52) on passage of TMV, CMV, or cowpea chlorotic mottle virus may have resulted from examination of only a small portion (12 to 16%) of each genome and/or characterization of only a single lineage for each of the three viruses studied. Nonuniform distribution of nucleotide substitutions accumulating in WSMV (Table 6) or vesicular stomatitis virus (34) after multiple passages clearly indicates that the amount of variation detected may differ depending on which region and proportion of the genome is examined. Alternatively, the relative titer may affect the ratio of singletons to nonsingletons, with high-titer viruses such as TMV, CMV, and cowpea chlorotic mottle virus having higher ratios of singletons to nonsingletons relative to a lower-titer virus such as WSMV. For the relative-titer explanation to hold, the number of genomes participating in cell-to-cell movement would be roughly

equivalent regardless of the titer but the total number of genomes produced within individual cells would differ significantly. In this scenario, higher-titer viruses experience more stringent bottlenecks, not because fewer viral genomes move but simply because a greater percentage do not. Note that this explanation would not affect the singleton frequency per nucleotide, which does not correspond to relative virus titer as long as the majority of singletons do not serve as replication templates. Nonetheless, the consensus sequence of TMV can change on passage (23), demonstrating that fundamental properties of population growth are similar among divergent plant virus species.

Host species effects on mutant accumulation. The singleton frequency per nucleotide remained constant after passage in wheat (Table 7), a result in congruence with that observed (52) for TMV and CMV passaged in *Nicotiana benthamiana*. Because viral polymerases have relatively high error rates and lack proofreading capabilities (13, 14), it is not surprising that singletons were the most common allele size class observed. That the singleton frequency per nucleotide did not significantly differ on passage suggests that given a constant rate of singleton production (the viral polymerase error rate), repeated bottlenecks at each movement event dictate that the bulk of the "quasispecies cloud" must be regenerated within each cell. Mutant frequencies at passage 9 for lineages passaged in barley or corn were not significantly higher than for lineages passaged in wheat (Table 7), implying that the three cereal host species do not provide differential environments affecting the accumulation of mutants. However, passage of TMV and CMV in different solanaceous hosts did result in altered mutant frequencies (53), indicating that certain host-virus combinations may differentially affect mutant accumulation.

Changes on passage are consistent with genetic drift. By examining nearly the entire WSMV genome, we measured changes in consensus sequence during serial passage at high MOI and under constant environmental conditions. From these results, it was clear that WSMV populations are dynamic, with genotypes arising, gaining predominance, losing dominance to other sequences, or serving as the foundation for the accumulation of additional changes.

Variation within the LDSI-S10 population was dominated by singletons. Upon passage, changes in SSCP patterns facilitated the identification of changes in consensus sequence due to point mutations. At passage 9, changes in consensus sequence were reflected by a majority of clones bearing the point mutation associated with the altered SSCP pattern (Fig. 5 and data not shown). However, not all clones examined in a lineage at passage 9 contained the point mutation which rose to prominence. Thus, mutations do not necessarily need to be fixed to give rise to SSCP differences between populations, but they need only achieve a threshold frequency.

Stochastic processes are now recognized as contributing to viral evolution (9, 17, 20, 30, 34, 38, 48). Most changes in consensus sequence during passage were unique to individual lineages and appear nondeterministic. That we observed consensus sequence shifts in populations of $\sim 10^{12}$ viral particles per plant implies either strong selection, transmission bottlenecks, or population subdivision within the host plant. The different populations were maintained in a constant environ-

ment, and only one consensus sequence change ($G_{6016} \rightarrow A$) appears to have been driven by selection.

High-MOI inoculations might be expected to preclude plant-to-plant transmission bottlenecks. However, the signature of bottlenecks was clearly evident in the distribution and composition of variation during passage at high MOI. Genotypic complexity among nonsingletons within each lineage at passage 9 was limited. For example, within wheat WB passage 9 (Fig. 5), clones appear to be from two groups or sublineages, suggesting that genetic bottlenecks must have occurred despite a high-MOI inoculation regime. Several pairs of polymorphic sites (e.g., $A_{8248} \rightarrow U$ and $A_{8403} \rightarrow G$) were linked in the majority of clones in which they were present (Fig. 5). This is consistent with a stepwise accumulation of mutations in a sublineage that increases in frequency by fortuitously passing through bottlenecks.

A simple and plausible explanation for bottlenecks during transmission at high MOI is subdivision of viral populations within a plant. Population subdivision within mixed infections of WSMV strains has been documented (17). Although that study examined the distribution of Type and Sidney 81 in plants simultaneously inoculated with a mixture of the two strains, variants of a single strain should subdivide in a similar fashion. Subdivision probably results from a small number of genomes participating in movement, followed by exclusion via cross-protection at the cellular level (17). One consequence of subdivision is that different regions of a systemically infected plant may harbor different proportions of variants such that an inoculum harvested from a given region of systemically infected tissue may be randomly enriched for a specific genotype. Subdivision will enhance the probability of fixation (i.e., genetic drift) of a mutant upon passage, even at high MOI. Over multiple passages, this process is repeated, resulting in accrual of mutations in a stepwise manner. Because the process of drift is essentially random, the outcome within any single lineage is not predetermined. The stochastic processes of bottlenecks, subdivision, and drift actually reduce variation within any given lineage, yet the random nature of these processes facilitates divergence among genetically isolated lineages within the constraints of selection.

ACKNOWLEDGMENTS

We acknowledge the excellent technical assistance of Melissa Morris and the support and advice of Casey Carmelea, Kempton Horken, and Il-Ryong Choi.

REFERENCES

1. Albiach-Martí, M. R., M. Mawassi, S. Gowda, T. Satyanarayana, M. E. Hilf, S. Shanker, E. C. Almira, M. C. Vives, C. López, J. Guerri, R. Flores, P. Moreno, S. M. Garnsey, and W. O. Dawson. 2000. Sequences of citrus tristeza virus separated in time and space are essentially identical. *J. Virol.* **74**:6856–6865.
2. Aleman-Verdaguier, M.-E., C. Goudou-Urbino, J. Dubern, R. N. Beachy, and C. Fauquet. 1997. Analysis of the sequence diversity of the P1, HC, P3, N1b, and CP genomic regions of several yam mosaic potyvirus isolates: implications for the intraspecies molecular diversity of potyviruses. *J. Gen. Virol.* **78**:1253–1264.
3. Bracho, M. A., A., Moya, and E. Barrio. 1998. Contribution of *Taq* polymerase-induced errors to the estimation of RNA virus diversity. *J. Gen. Virol.* **79**:2921–2928.
4. Brakke, M. K. 1958. Properties, assay, and purification of wheat streak mosaic virus. *Phytopathology* **48**:439–445.
5. Brakke, M. K. 1970. Systemic infections for the assay of plant viruses. *Annu. Rev. Phytopathol.* **8**:61–84.
6. Brakke, M. K., R. N. Skopp, and L. C. Lane. 1990. Degradation of wheat

- streak mosaic virus capsid protein during leaf senescence. *Phytopathology* **80**:1401–1405.
7. **Cha, T., J. Zhao, E. Lane, M. A. Murray, and D. S. Stec.** 1997. Determination of the genome composition of influenza virus reassortants using multiplex reverse transcription-polymerase chain reaction followed by fluorescent single-strand conformation polymorphism analysis. *Anal. Biochem.* **252**:24–32.
 8. **Chenault, K. D., R. M. Hunger, and J. L. Sherwood.** 1996. Comparison of the nucleotide sequence of the coat protein open reading frame of nine isolates of wheat streak mosaic Rymovirus. *Virus Genes* **13**:187–188.
 9. **Choi, I.-R., J. S. Hall, M. Henry, L. Zhang, G. L. Hein, R. French, and D. C. Stenger.** 2001. Contributions of genetic drift and negative selection on the evolution of three strains of wheat streak mosaic tritrovirus. *Arch. Virol.* **146**:619–628.
 10. **Domingo, E. and J. J. Holland.** 1994. Mutation rates and rapid evolution of RNA viruses, p. 161–184. *In* S. S. Morse (ed.), *The evolutionary biology of viruses*. Raven Press, New York, N.Y.
 11. **Domingo, E., J. J. Holland, C. Biebricher, and M. Eigen.** 1995. Quasispecies: the concept and the word, p. 181–191. *In* A. J. Gibbs, C. H. Calisher, and F. García-Arenal (ed.), *Molecular basis of virus evolution*. Cambridge University Press, Cambridge, England.
 12. **Domingo, E., C. Escarmis, N. Sevilla, A. Moya, S. F. Elena, J. Quer, I. Novella, I. and J. J. Holland.** 1996. Basic concepts in RNA virus evolution. *FASEB J.* **10**:859–864.
 13. **Drake, J. W.** 1993. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. USA* **90**:4171–4175.
 14. **Drake, J. W., and J. J. Holland.** 1999. Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. USA* **96**:13910–13913.
 15. **Fraile, A., F. Escriu, M. A. Aranda, J. M. Malpica, A. J. Gibbs, and F. García-Arenal.** 1997. A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *J. Virol.* **71**:8316–8320.
 16. **Frost, S. D. W., M. Nijhuis, R. Schuurman, C. A. B. Boucher, and A. J. L. Brown.** 2000. Evolution of lamivudine resistance in human immunodeficiency virus type 1 infected individuals: the relative roles of drift and selection. *J. Virol.* **74**:6262–6268.
 17. **Hall, J. S., R. French, G. L. Hein, T. J. Morris, and D. C. Stenger.** 2001. Three distinct mechanisms facilitate genetic isolation of sympatric wheat streak mosaic virus lineages. *Virology* **282**:230–236.
 18. **Hey, J., and J. Wakely.** 1997. A coalescent estimator of the population recombination rate. *Genetics* **145**:833–846.
 19. **Hirsch, V. M., G. Dapolito, A. Hahn, J. Lifson, D. Montefiori, C. R. Brown, and R. Goeken.** 1998. Viral genetic evolution in macaques infected with molecularly cloned simian immunodeficiency virus correlates with the extent of persistent viremia. *J. Virol.* **72**:6482–6489.
 20. **Jenkins, G. M., M. Wroboey, C. H. Woelk, and E. C. Holmes.** 2001. Evidence for the non-quasispecies evolution of RNA viruses. *Mol. Biol. Evol.* **18**:987–994.
 21. **Ji, J. P., and L. A. Loeb.** 1992. Fidelity of HIV-1 reverse transcriptase copying RNA *in vitro*. *Biochemistry* **31**:954–958.
 22. **Kearney, C. M., J. Donson, G. E. Jones, and W. O. Dawson.** 1993. Low level of genetic drift in foreign sequences replicating in an RNA virus in plants. *Virology* **192**:11–17.
 23. **Kearney, C. M., M. J. Thomson, and K. E. Roland.** 1999. Genome evolution of tobacco mosaic virus populations during long-term passaging in a diverse range of hosts. *Arch. Virol.* **144**:1513–1526.
 24. **Keese, P., and A. Gibbs.** 1993. Plant viruses: master explorers of evolutionary space. *Curr. Opin. Genet. Dev.* **3**:873–877.
 25. **Kong, P., L. Rubio, M. Polek, and B. W. Falk.** 2000. Population structure and genetic diversity within California citrus tristeza virus (CTV) isolates. *Virus Genes* **21**:139–145.
 26. **Kurath, G., J. A. Heick, and J. A. Dodds.** 1993. RNase protection analyses show high genetic diversity among field isolates of satellite tobacco mosaic virus. *Virology* **194**:414–418.
 27. **Kurath, G., and P. Palukaitis.** 1990. Serial passage of infectious transcripts of a cucumber mosaic virus satellite RNA clone results in sequence heterogeneity. *Virology* **176**:8–15.
 28. **Kurath, G., C. M. E. Rey, and J. A. Dodds.** 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.
 29. **Lea, D. E., and C. A. Coulson.** 1949. The distribution of the numbers of mutants in bacterial populations. *Genetics* **49**:264–285.
 30. **Leigh-Brown, A. J.** 1997. Analysis of HIV-1 *env* gene sequences reveals evidence for a low effective number in the viral population. *Proc. Natl. Acad. Sci. USA* **94**:1862–1865.
 31. **Luria, S. E., and M. Delbrück.** 1943. Mutations of bacteria in virus sensitivity to virus resistance. *Genetics* **28**:491–511.
 32. **McKinney, H. H.** 1937. Mosaic diseases of wheat and related cereals. U.S. Department of Agriculture Circular 442. U.S. Department of Agriculture, Washington, D.C.
 33. **McNeil, J. E., R. French, G. L. Hein, P. S. Baenziger, and K. M. Eskridge.** 1996. Characterization of genetic variability among natural populations of wheat streak mosaic virus. *Phytopathology* **86**:1222–1227.
 34. **Moya, A., S. F. Elena, A. Bracho, R. Miralles, and E. Barrio.** 2000. The evolution of RNA viruses: a population genetics view. *Proc. Natl. Acad. Sci. USA* **97**:6967–6973.
 35. **Moya, A., and F. García-Arenal.** 1995. Population genetics of viruses: an introduction, p. 213–223. *In* A. J. Gibbs, C. H. Calisher, and F. García-Arenal (ed.), *Molecular basis of virus evolution*. Cambridge University Press, Cambridge, England.
 36. **Moya, A., E. Rodríguez-Cerezo, and F. García-Arenal.** 1993. Genetic structure of natural populations of the plant RNA virus tobacco mild green mosaic virus. *Mol. Biol. Evol.* **10**:449–456.
 37. **Neildez, O., R. Le Grand, P. Caufour, B. Vaslin, A. Cheret, F. Matheux, F. Theodoro, P. Roques, and D. Dormont.** 1998. Selective quasispecies transmission after systemic or mucosal exposure of macaques to simian immunodeficiency virus. *Virology* **243**:12–20.
 38. **Nijhuis, M., C. A. B. Boucher, P. Schipper, T. Leitner, R. Schuurman, and J. Albert.** 1998. Stochastic processes strongly influence HIV-1 evolution during suboptimal protease-inhibitor therapy. *Proc. Natl. Acad. Sci. USA* **95**:14441–14446.
 39. **Novella, I. S., E. A. Duarte, S. F. Elena, A. Moya, E. Domingo, and J. J. Holland.** 1995. Exponential increases of RNA virus fitness during large population transmissions. *Proc. Natl. Acad. Sci. USA* **92**:5841–5844.
 40. **Novella, I. S., S. F. Elena, A. Moya, E. Domingo, and J. J. Holland.** 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. *J. Virol.* **69**:2869–2872.
 41. **Novella, I. S., J. Quer, E. Domingo, and J. J. Holland.** 1999. Exponential fitness gains of RNA virus populations are limited by bottleneck effects. *J. Virol.* **73**:1668–1671.
 42. **Orita, M., H. Iwahana, H. Kanazawa, K. Hayashi, and T. Sekiya.** 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphism. *Proc. Nat. Acad. Sci. USA* **86**:2766–2770.
 43. **Pawlotsky, J.-M., G. Germanidis, A. U. Neumann, M. Pellerin, P. O. Frainais, and D. Dhumeaux.** 1998. Interferon resistance of hepatitis C virus genotype 1b: relationship to nonstructural 5A gene quasispecies mutations. *J. Virol.* **72**:2795–2805.
 44. **Rodríguez-Cerezo, E., S. F. Elena, A. Moya, and F. García-Arenal.** 1991. High genetic stability in natural populations of the plant RNA virus tobacco mild green mosaic virus. *J. Mol. Evol.* **32**:328–332.
 45. **Rodríguez-Cerezo, E., and F. García-Arenal.** 1989. Genetic heterogeneity of the RNA genome population of the plant virus U5-TMV. *Virology* **170**:418–423.
 46. **Roossinck, M.** 1997. Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol.* **35**:191–209.
 47. **Rosche, W. A., and P. L. Foster.** 2000. Determining mutation rates in bacterial populations. *Methods* **20**:4–17.
 48. **Rouzine, I. M., and J. M. Coffin.** 1999. Linkage disequilibrium test implies a large effective population number for HIV *in vivo*. *Proc. Natl. Acad. Sci. USA* **96**:10758–10763.
 49. **Rozas, J., and R. Rozas.** 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**:174–175.
 50. **Sala, M., and S. Wain-Hobson.** 2000. Are RNA viruses adapting or merely changing? *J. Mol. Biol.* **51**:12–20.
 51. **Sánchez-Sánchez, H., M. Henry, E. Cárdenas-Soriano, and H. F. Alviso-Villasana.** 2001. Identification of *Wheat streak mosaic virus* and its vector *Aceria tosichella* in Mexico. *Plant Dis.* **85**:13–17.
 52. **Schneider, W. L., and M. J. Roossinck.** 2000. Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host. *J. Virol.* **74**:3130–3134.
 53. **Schneider, W. L., and M. J. Roossinck.** 2001. Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *J. Virol.* **75**:6566–6571.
 54. **Shen, P. D., F. Wang, P. A. Underhill, C. Franco, W. H. Yang, A. Roxas, R. Sung, A. A. Lin, R. W. Hyman, D. Vollrath, R. W. Davis, L. L. Cavalli-Sforza, and P. J. Oefner.** 2000. Population genetic implications from sequence variation in four Y chromosome genes. *Proc. Natl. Acad. Sci. USA* **97**:7354–7359.
 55. **Skotnicki, M. L., S. W. Ding, A. M. Mackenzie, J. Q. Mo, and A. J. Gibbs.** 1993. RNA hybrid mismatch polymorphisms in Australian populations of turnip yellow mosaic tymovirus. *Arch. Virol.* **132**:83–99.
 56. **Skotnicki, M. L., A. M. Mackenzie, and A. J. Gibbs.** 1992. Turnip yellow mosaic virus variants produced from DNA clones encoding their genomes. *Arch. Virol.* **127**:25–35.
 57. **Skotnicki, M. L., A. M. Mackenzie, and A. J. Gibbs.** 1996. Genetic variation in populations of kennedy yellow mosaic tymovirus. *Arch. Virol.* **141**:99–110.
 58. **Smith, D. B., J. McAllister, C. Casino, and P. Simmonds.** 1997. Virus 'quasispecies': making a mountain out of a molehill? *J. Gen. Virol.* **78**:1511–1519.
 59. **Stavalone, L., D. Alioto, A. Ragozzino, and J.-F. Libalberté.** 1998. Variability among turnip mosaic potyvirus isolates. *Phytopathology* **88**:1200–1204.
 60. **Steinhauer, D. A., J. A. de la Torre, E. Meier, and J. J. Holland.** 1989.

- Extreme heterogeneity in populations of vesicular stomatitis virus. *J. Virol.* **63**:2072–2080.
61. **Stephan, W., L. Xing, D. A. Kirby, and J. M. Braverman.** 1998. A test of the background selection hypothesis based on nucleotide data from *Drosophila ananassae*. *Proc. Nat. Acad. Sci. USA* **95**:5649–5654.
62. **Stenger, D. C., J. S. Hall, I.-R. Choi, and R. French.** 1998. Phylogenetic relationships within the family *Potyviridae*: wheat streak mosaic virus and brome streak mosaic virus are not members of the genus *Rymovirus*. *Phytopathology* **88**:782–787.
63. **Tajima, F.** 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585–595.
64. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**:4876–4882.
65. **van Vloten-Doting, L., and J. F. Bol.** 1988. Variability, mutant selection, and mutant stability in plant RNA viruses, p. 37–45. *In* E. Domingo, J. J. Holland, and P. Ahlquist (ed.), *RNA genetics*, vol. III. CRC Press, Inc., Boca Raton, Fla.
66. **Watterson, G. A.** 1975. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* **7**:256–276.
67. **Weiss, G., and A. von Haeseler.** 1997. A coalescent approach to the polymerase chain reaction. *Nucleic Acid Res.* **25**:3082–3087.
68. **Welsh, J., K. Castren, and K. H. Vahakangas.** 1997. Single-strand conformation polymorphism analysis to detect p53 mutations: characterization and development of controls. *Clin. Chem.* **43**:2251–2255.