Evolutionarily Related Sindbis-Like Plant Viruses Maintain Different Levels of Population Diversity in a Common Host

WILLIAM L. SCHNEIDER AND MARILYN J. ROOSSINCK*

Plant Biology Division, Samuel Roberts Noble Foundation, Ardmore, Oklahoma 73402

Received 18 October 1999/Accepted 2 January 2000

The levels of population diversity of three related Sindbis-like plant viruses, Tobacco mosaic virus (TMV), Cucumber mosaic virus (CMV), and Cowpea chlorotic mottle virus (CCMV), in infections of a common host, Nicotiana benthamiana, established from genetically identical viral RNA were examined. Despite probably having a common evolutionary ancestor, the three viruses maintained different levels of population diversity. CMV had the highest levels of diversity, TMV had an intermediate level of diversity, and CCMV had no measurable level of diversity in N. benthamiana. Interestingly, the levels of diversity were correlated to the relative host range sizes of the three viruses. The levels of diversity also remained relatively constant over the course of serial passage. Closer examination of the CMV and TMV populations revealed biases for particular types of substitutions and regions of the genome that may tolerate fewer mutations.

The error-prone replication, large populations, and rapid replication times associated with RNA viruses result in the potential for genetically diverse populations, termed quasispecies, arising within a single host (16). Developed as a model for early forms of life (14), the theoretical quasispecies describes a steady-state collection of genetic mutants that vary around a consensus sequence. Viral quasispecies are complex and dynamic distributions where the level of population variation (quasispecies cloud size) reacts to changes in selection pressures (8). There are a number of biological and evolutionary implications associated with highly diverse populations (7). Maintaining a large quasispecies cloud size would allow a virus ready access to a pool of mutants which could become the selectively advantaged dominant RNA species in a shift to a new environment. Conversely, highly diverse populations that are subjected to repeated bottlenecks have been shown to lose fitness through a process known as Muller’s ratchet (4, 9, 10, 25, 26).

Both DNA (15, 39) and RNA plant viruses (6, 23, 30, 32) can maintain highly diverse populations in collections of field isolates, but the quasispecies variation of single plant virus isolates has not been examined. Plant viruses can use several mechanisms to generate quasispecies diversity, including replication error, recombination, and, for multipartite viruses, reassortment (for a review see reference 33). However, the extent of population variation is limited by selection pressure for variants that interact successfully with both host and viral proteins necessary for completion of the infection cycle. The Sindbis-like virus group includes a number of plant and animal viruses with similarities in their genome organizations and nonstructural proteins. The three virus species chosen for this study were Tobacco mosaic virus (TMV; genus Tobamovirus), Cucumber mosaic virus (CMV; genus Cucumovirus, family Bromoviridae), and Cowpea chlorotic mottle virus (CCMV; genus Bromovirus, family Bromoviridae). TMV is a monopartite rosspend virus, and CMV and CCMV are tripartite isocahedral viruses. Although the genomic organization of regions producing similar functional domains for nonstructural proteins (Fig. 1A) indicates that TMV, CMV, and CCMV evolved from a common ancestor, these three viruses differ greatly in terms of the sizes of their host ranges. CMV infects about 1,000 plant species (12), TMV infects 80 to 100 plant species (17), and CCMV infects only a few plant species (22). A correlation between diversity in the viral quasispecies and the sizes of viral host ranges seems possible (33). However, there are few experimental examinations of plant virus population diversity, and no controlled comparisons between different viruses, related or otherwise, have been reported. The well-characterized cDNA clones and the availability of a common host (Nicotiana benthamiana) for TMV, CMV, and CCMV make it possible to do a controlled experimental comparison of quasispecies diversity for these related viruses. Here we examine CMV, TMV, and CCMV quasispecies, both in initial infections and following consecutive passages in N. benthamiana.

MATERIALS AND METHODS

Plants and viral inoculations. N. benthamiana plants were grown under greenhouse conditions (16-h days at 24°C, 8-h nights at 20°C; supplemental light intensity, 500 μmol of photons m⁻² s⁻¹) and inoculated at the three- or four-leaf stage. Plants were maintained in the greenhouse after inoculation. Viral RNAs were generated from cDNA clones of Fny-CMV (29, 34), CCMV (1, 11), and U1 (36). For Fny-CMV and CCMV, transcripts were generated in vitro using SP6 or T7 RNA polymerase (Ambion). The TMV construct was inoculated as plasmid DNA, with viral RNA generated in planta from the Cauliflower mosaic virus 35S promoter (36). Transcript- or DNA-inoculated plants constituted passage zero. Subsequent passages (1 through 10) were at 14-day intervals, using sap extracts from the inoculated plant material. Two plants were inoculated at the passage zero level. One or two systemically infected leaves were removed from both plants, combined, and ground in 1 mM NaH₂PO₄ buffer (pH 7.0). The combined sap was used to inoculate both plants in the next passage.

Extraction of total RNA and cDNA synthesis. Fourteen to 20 days postinoculation total RNA was extracted from systemically infected leaves of passage 0, passage 1, and passage 10 plants as previously described (34). One-fifth of the total RNA extraction was used as a template for reverse transcription (RT) with Superscript reverse transcriptase as prescribed by the manufacturer (Gibco). Two RT reactions were done for each infected plant. RT reactions were primed with primers 4144 (AACCTACTGCAGGATTAACCAATGCGGAAAGAATCCGACAC) for TMV, 4140 (AACCTACTGCAGGATTAACCAATGCGGAAAGAATCCGACAC) for CMV, and 4145 (TAATCGGAATTCAGGAAACAGCTATGACCCGCGCGATCCAGACAC) for CCMV. The cDNAs were used as templates for thermal cycling reactions using the following primers: 4139 (AACCTACTGCAGGATTAACCAATGCGGAAAGAATCCGACAC) and 4144 for CMV, 4140 and 4146 (AACCTACTGCAGGATAAAATGCCGAGGATTACAC) for CCMV, and 4143 (AACCTACTGCAGGATAAAATGCCGAGGATTACAC) and 4145 for TMV (Fig. 1B). Thermal cycling reactions were carried out for 15 cycles (94°C denaturation for 30 s, 50°C annealing for 1 min, 72°C extension for 1 min), and included a polymerase with proofreading capability (β′; Stratagenec). Between 82 and 123 nucleotides at the 3′ termini of the viral RNAs were
A

Tobamovirus Genus (TMV)

Methyltransferase Domain

Helicase Domain

RdRP Domain

Movement protein

Coat protein

Bromoviridae Family (CMV, CCMV)

RNA 1

RNA 2

RNA 3

FIG. 1. Genomic organization of the three viruses used in this study and locations of analyzed sequence regions. (A) Genomic maps of TMV, CMV, and CCMV with the common regions labeled. (B) Portions of viral genomes encoding the coat protein and flanking regions that were amplified and cloned for sequence analysis. Arrows, primer sites; numbers, nucleotide positions.

not included in the cloned and sequenced segments. In addition to reactions using the N. benthamiana viral populations, RT and thermal cycling reactions were done using in vitro transcripts of all three viruses as templates, in order to establish the level of error introduced by the experimental method.

Analysis of viral clones. The amplified products generated from the viral RNAs were cloned into the vector pBSKS+ (Stratagene). The viral clones were sequenced using dideoxy sequencing and the T7 DyeDeoxy terminator cycle sequencing kit (Applied Biosystems). Sequencing gels were run on an ABI 377 capillary sequencer as described by the manufacturer (Perkin-Elmer). Eleven to 16 clones (at least 10,000 nucleotides) were sequenced for each virus population. The sequences of all viral clones were compared to the sequences of the source infectious clones, and the percentage of clones with mutations and the mutation frequency were calculated for each population.

A number of steps were taken to minimize the level of error introduced by the experimental procedure. Eighty-two to 123 bases from the 3' terminus of the viral RNAs were excluded from the cloned region, because RNA polymerases used for in vitro transcription have a tendency to make errors at the ends of transcripts (24). Errors were reduced during cDNA synthesis by using a high-fidelity reverse transcriptase and by thermal cycling using a polymerase with proofreading capabilities. Only 15 cycles were completed for each thermal cycling reaction in order to prevent mutations introduced by fluctuations in nucleotide levels (38). In addition, control reactions were done using in vitro transcripts as the template RNA to estimate the background level of experimental mutation frequency at 4.5 ± 10^{-5} mutations per nucleotide.

RESULTS AND DISCUSSION

Generation and analysis of quasispecies. Infections were initiated for all three viruses using transcripts derived from cDNA clones. This ensured that the infections of each virus began with genetically identical sources of RNA. All three sets of viral clones used for in vitro transcription were previously thoroughly characterized, and transcript infections displayed no detectable phenotypic differences from wild-type viral infections. A minimum of two plants were used for each mechanical inoculation of N. benthamiana plants, and 14 to 20 days postinoculation total RNA was extracted from systemically infected leaves. A segment of viral RNA was cloned from total RNA pools of all three viruses. The cloned region represented the coat protein and flanking regions (Fig. 1), including both translated and nontranslated sequences (roughly a 1-kb segment for all three viruses). An excess of viral RNA was used in the cDNA reactions to prevent amplification of only a small subset of the viral population. Eleven to 16 clones (at least 10,000 nucleotides) were sequenced for each virus population. The sequences of all viral clones were compared to the sequences of the source infectious clones, and the percentage of clones with mutations and the mutation frequency were calculated for each population.

The single mutant derived from the CCMV passage 10 plants was a recombinant which had incorporated a stretch of 24 bases in the nontranslated intercistronic region.

TABLE 1. Quasispecies variation in CMV, TMV, and CCMV populations after serial passage in N. benthamiana

<table>
<thead>
<tr>
<th>RNA</th>
<th>Passage</th>
<th>% Mutated clones (no. mutated/total no.)</th>
<th>No. of mutations/no. of bases sequenced</th>
<th>Mutation frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT-PCR controls</td>
<td>5 (1/22)</td>
<td>1/23,488</td>
<td>4.5 × 10^{-5}</td>
<td></td>
</tr>
<tr>
<td>CMV</td>
<td>0</td>
<td>43 (6/14)</td>
<td>8/16,454</td>
<td>4.6 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>60 (9/15)</td>
<td>15/18,496</td>
<td>8.1 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>69 (9/13)</td>
<td>11/15,827</td>
<td>7.0 × 10^{-4}</td>
</tr>
<tr>
<td>Total</td>
<td>57 (24/42)</td>
<td>34/50,777</td>
<td>6.6 × 10^{-4}</td>
<td></td>
</tr>
<tr>
<td>TMV</td>
<td>0</td>
<td>29 (4/14)</td>
<td>7/13,227</td>
<td>5.2 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>33 (4/12)</td>
<td>5/11,350</td>
<td>4.4 × 10^{-4}</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>27 (3/11)</td>
<td>3/10,234</td>
<td>2.9 × 10^{-4}</td>
</tr>
<tr>
<td>Total</td>
<td>50 (11/37)</td>
<td>15/34,811</td>
<td>4.3 × 10^{-4}</td>
<td></td>
</tr>
<tr>
<td>CCMV</td>
<td>0</td>
<td>0 (0/14)</td>
<td>0/13,454</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>7 (1/14)</td>
<td>1/13,969</td>
<td>7.1 × 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6 (1/16)</td>
<td>24/15,847</td>
<td>1.5 × 10^{-3}</td>
</tr>
<tr>
<td>Total</td>
<td>16 (2/44)</td>
<td>25/43,270</td>
<td>5.7 × 10^{-4}</td>
<td></td>
</tr>
</tbody>
</table>

*The single mutant derived from the CCMV passage 10 plants was a recombinant which had incorporated a stretch of 24 bases in the nontranslated intercistronic region.
of the passage 10 CCMV populations (Table 1), where one cloned RNA had undergone a recombination event incorporating 24 nucleotides into the intercistronic region. This single mutational event artificially raised the mutation frequency, but if the percentage of mutated clones is considered or if the recombination event is considered as a single mutation, the levels of population diversity in CCMV populations remain statistically unchanged over the course of serial passage. The CMV populations showed a slightly significant increase in mutation frequency between passage 0 and passage 1 but then stabilized through passage 10.

There are a number of possible scenarios that could affect diversity levels of RNA viruses in serial passage experiments. The levels of diversity could rise as the virus population expands exponentially and each new infection is started with an increasing spectrum of mutants. Alternatively, the levels of diversity could fall as selection pressures select for mutants that are better suited to infecting *N. benthamiana* and these mutants begin to dominate the population. However, these experiments seem to indicate that the viral populations rapidly reach a level of diversity in the initially inoculated plant which is maintained relatively constant over the course of passaging. This would imply that the sap passaging technique used is not a severe bottleneck, since the accumulation of deleterious mutants associated with Muller’s ratchet (7) did not occur in the higher-diversity CMV and TMV populations. In addition, this raises the possibility of a threshold limit to population diversity for plant viruses. Quasispecies clouds for these three viruses appear to follow Eigen’s prediction (13) that there is selection for a level of variation. Alternatively, these viruses might establish each subsequent infection in the passage with viral RNAs having sequences that are identical to the consensus sequence, in effect rendering infections in later passages the same as the initial-transcript infections. No changes to the consensus sequence were observed in any of the passaged populations; mutations arose and then disappeared without becoming fixed. However, it should be noted that for most plant viruses, especially those with divided genomes, there is nothing equivalent to a plaque assay. Hence, there is no way of accurately quantifying the amount of virus used in the inoculum or quantifying the number of viruses that initiate the infection in the new host.

**Comparing quasispecies diversity.** Comparing the sequence data from the clones derived from the *N. benthamiana* viral populations demonstrated differences in the level of quasispecies variation between the three viruses (Table 1). Because the populations changed very little over the course of serial passage, the data from all three sampled passages were combined for purposes of comparison. CMV populations showed the highest level of variation, significantly higher than that of the controls (*P* < 0.05). Similarly, TMV populations were also significantly more diverse than the controls, but somewhat less diverse than CMV populations (*P* < 0.05). In contrast, CCMV populations had no more variation than the controls. Additionally, a comparison of the percentages of mutated clones shows that this trend is consistent within each individual passage. Interestingly, these results correlated with the relative host range sizes of the three viruses. No two clones had the same mutations, indicating that each mutant clone represented a unique viral RNA. Thus, we can infer that each clone with a wild-type sequence also represents a cDNA generated from a unique RNA template and that the experimental procedure provides a representative sampling of the population.

There are a number of possible explanations for the observed differences in the levels of genetic diversity of viral populations. Certainly the effects of host selection play a role in limiting error accumulation in the population, but here all three viruses were subjected to the same environment. The differences in diversity are likely not due to differences in population size, since all three viruses replicate to high levels in *N. benthamiana*. In fact, CMV, with the highest level of diversity, has the lowest viral titer (0.1 to 0.3 mg/gm) (28). CCMV accumulates to 0.3 to 0.5 mg/gm (22), and TMV accumulates to 1 mg/gm (41). The relative viral titers do not correlate with the levels of population diversity.

Another potential source for the different levels of diversity may be differences in the fidelity of replication. The differences in population variation may simply be due to differences in the error rates and recombination rates of the viral replicates. For example, the high mutation frequencies in the CMV populations may be due to higher error rates in replication. Certainly the replication fidelity of these viruses could be affected by the intracellular site of replication due to nucleotide levels, but there are no data on this for any of these viruses. Recombination, which can act as a purifying mechanism that reduces diversity (5), has been well documented in members of the Bromovirus genus such as CCMV (3, 37). Recombination may reduce diversity in the CCMV population. However, these viruses have similar origins and similar replication proteins. It seems likely that in a common host these three viruses utilize the same host factors and have replicating units with similar characteristics. While there are no studies of replicate error rates or recombination rates for any plant viruses to date, it is hard to imagine that these viruses vary greatly in their abilities to generate diversity through replication. Alternatively, the differences could be explained by a factor unique to the infection cycle of each particular virus. Perhaps the low-diversity CCMV populations are subjected to frequent bottlenecks, which would prevent the population from developing high diversity without losses of fitness. CMV and TMV quasispecies could also be limited by bottlenecks of various sizes.

**Distribution and types of mutations.** None of the mutations in either the CMV or TMV populations became fixed during the passage experiments. All of the mutations observed in the passage 0 clones were not evident in the passage 1 clones, and the mutations observed in the passage 1 clones were subsequently lost by passage 10 (data not shown). There was no change in the consensus sequence, indicating that none of the mutations conferred any selective advantage to the virus and emphasizing the rare nature of adaptive mutations in viral populations. Occasionally multiple mutations were observed in close proximity to each other in the same viral clone (data not shown). These mutations may have arisen as a result of the same replication event, although this is difficult to test. When mutation frequencies were calculated, each altered nucleotide was considered an individual mutation, regardless of its proximity to other mutations in the same viral clone. If mutations in close proximity were to be considered as single mutational events, the mutation frequencies observed would be lowered slightly but the comparative mutation frequencies for the three viruses as well as the relative levels of diversity in passaged populations would remain the same.

An examination of the locations of the mutations observed suggests that mutations may not be completely random. The observed mutations in the CMV populations were distributed throughout the sequenced region, both in translated and nontranslated regions (Fig. 2), with a bias for nontranslated regions over translated regions. In addition, there were large areas where no mutations were observed, in particular the area between nucleotides 1577 and 1846 of the coat protein gene. This could represent a region where selection acts against mutation tolerance, but more mutations need to be mapped to

---

*SCHNEIDER AND ROOSSINCK J. VIROL.*
confirm that it is not occurring by random chance. The distribution of TMV mutations also covered the majority of the sequenced portions of the genome (Fig. 3), although in TMV 14 of the 15 mutations observed were in the coding regions. However, less of the TMV cloned segment represents non-translated regions. Both of the CCMV mutational events occurred in the nontranslated intercistronic region (data not shown). Any selection that might be in effect here must be working at the RNA level. All populations have more than enough functional copies of the genes in question to supply the appropriate proteins in trans. This is confirmed by examining the classes of mutations found in the coding regions, where there is no bias for synonymous mutations. Four of the 11 mutations in the CMV coding region were silent, and 5 of the 14 mutations in the TMV coding region were silent. Thus, any selection is likely to be for the capacities of the viral RNA in replication, movement, encapsidation, or stability.

The majority of observed mutations in the viral populations were substitutions: 31 of the 34 observed mutations in CMV populations were substitutions, all of the observed mutations in TMV populations were substitutions, and 1 of 2 mutations in the CCMV populations was a substitution. The three remaining mutations in CMV were single-base deletions. The CCMV recombination event was the only example of an addition mutation; no recombination events in the CMV and TMV populations were observed. Close examination of the specific changes indicates a bias for transitions, in particular G-to-A and C-to-U transitions. There was a strong bias for G-to-A transitions in the CMV populations (12 of 31 substitutions; Table 2). The TMV populations also demonstrated a slight preference for G-to-A and C-to-U transitions (7 of 15 substitutions; Table 3). This sort of transition bias, noted in other viral systems (21, 40), is higher than what is seen in DNA evolution (40). The bias may be due to the ability of guanosine to form a hydrogen bond with uridine in RNA base pairing. A phylogenetic study of CMV RNA 3 sequences has also indicated a strong bias for G-to-A and C-to-U transitions, although C-to-U transitions appear to be more common (35).

**Conclusions.** Different levels of population diversity in field isolates of plant viruses have been observed, but until now there have been no controlled systematic studies of plant viral quasispecies and how they related to other biological properties. This study is also the first controlled comparison of related viral populations in a common host, CMV, CCMV, and TMV represent three viruses with a common evolutionary link that have distinctly different host range sizes. Previous studies on the population diversity of these three viruses suggest that CMV replicated-RNA populations (especially CMV satellite RNA populations) can be highly diverse (2, 20, 27, 30). Reports indicate that other strains of TMV maintain a lower level of diversity for genomic RNAs (31), even though TMV replicated satellite virus RNAs have been shown to have high levels of population diversity (19). A recent study of TMV suggested a mutation frequency of $3.1 \times 10^{-4}$ per nucleotide in passaged TMV populations on a variety of hosts (18). CCMV populations have not been previously studied for diversity levels. A controlled study of these three viruses in a common host provides an opportunity to better understand the quasispecies nature of plant RNA viruses and how quasispecies may relate to host range.

Interestingly, the levels of quasispecies diversity for CMV, TMV, and CCMV in the common host *N. benthamiana* correlate directly with the relative sizes of the viral host ranges (Table 1). This has important evolutionary implications for the quasispecies structure of viral populations, since it suggests that highly diverse viruses such as CMV have a better chance of expanding into a new niche and thus pose a greater threat of emerging as new crop diseases. Several studies have examined the fitness effects of highly diverse viral populations in different circumstances. However, the overall effects of population diversity on the long-term evolutionary trajectories of viruses, particularly as they relate to the emergence of new diseases, are not well understood. The advantage of genetic diversity in

TABLE 3. Types of substitution mutations observed in TMV populations

<table>
<thead>
<tr>
<th>Original base</th>
<th>G</th>
<th>A</th>
<th>U</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>U</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
populations that encounter new environmental challenges is a consistent theme that runs through all levels of biology. Theoretically, the ability to maintain genetic diversity in viral populations should enhance chances for adaptation to new selective regimes. Alternatively, if high diversity in the viral population resulted in fitness losses, the forces of selection would rapidly eliminate viruses that surpass viable limits of population diversity. Although a correlation between viral population diversity and expansion into new hosts is intriguing, only one host species and three viruses were used in this study. It will be important to analyze more hosts and more viruses to confirm this observation.

ACKNOWLEDGMENTS

We thank Rick Nelson and Shelly Carter for supplying U1 TMV; Robert Gonzalez, Angela Scott, and Ann Harris for assistance in sequencing; Joachim de Miranda, Jim Bull, Holly Wichman, and Isabel Novella for helpful comments and discussion; and Greg May and Xin Shun Ding for critical reviews of the manuscript.

REFERENCES