Variability and Evolution of the Plant RNA Virus Pepper Mild Mottle Virus

EMILIO RODRÍGUEZ-CEREZO,¹ ANDRÉS MOYA,² AND FERNANDO GARCÍA-ARENAL^{1*}

Departamento de Patología Vegetal, E.T.S.I. Agrónomos, 28040 Madrid,¹ and Laboratorio de Genética, Facultad de Biología, C/Dr. Moliner 50, 46100 Burjassot, Valencia,² Spain

Received 24 October 1988/Accepted 10 January 1989

The RNA genomes of 26 isolates of pepper mild mottle virus were compared by their RNase T_1 fingerprints. Twenty-three isolates came from epidemic outbreaks in greenhouse-grown peppers in Almería (southeastern Spain) from 1983 to 1987; three other isolates, from 1980, came from Sicily (Italy) and Zaragoza (central Spain). The 26 fingerprints can be classified into 10 different types; nucleotide substitution rates show them to be very similar. Cluster and cladistic analyses group types corresponding to the Almería isolates separate from those of 1980. Intraannual and interannual nucleotide differences were estimated. An evolutionary model for pepper mild mottle virus built on these data indicates a highly stable population, maintaining its diversity through time, with a main prevailing haplotype from which closely related variants arise that do not replace it. This high stability could be due to strong functional constraints on variation, as suggested by the high proportion of invariant versus polymorphic sites in fingerprints.

Current methods of genome comparison are particularly adequate to the quantitative study of genetic variability of viruses, because they permit the random representation of a large fraction of viral genomes. These methods have been extensively applied during the last 15 years to animal and bacterial RNA viruses, and the accumulated data have provided estimates of the mutability of RNA genomes and of the heterogeneity and variability of RNA populations. Models on RNA population structure and evolution have been developed (8, 12, 13, 21-23), giving a general picture of the high variation of RNA viruses. In the case of plant RNA viruses, data on the appearance and reversion of mutants (10, 11, 14, 28), on changes associated with host shift experiments (1, 9, 16), and on virulence shifts in epidemics in crop plants (26) suggest a high potential to vary, in accordance with this view. On the other hand, data on field strains of viruses naturally infecting wild plants suggest a high stability of the genomes of plant RNA viruses (2, 20; A. Gibbs, Abstr. Sixth. Int. Congr. Virol. 1984, p. 12). These data, nevertheless, do not come from quantitative studies with enough precision to be comparable to those done on bacterial or animal RNA viruses.

We have studied the variation of a plant RNA virus, the tobamovirus pepper mild mottle virus (PMMV). PMMV is a plus-sense RNA virus with a 6,400-base-long monopartite genome encapsidated in rod-shaped particles (29). PMMV is transmitted only through the seed and through direct contact between plants. It induces mottling and mild mosaic symptoms in peppers (*Capsicum annuum* L.), and epidemics are severe in cultivars carrying the L1 gene of resistance to other tobamoviruses (3, 29).

We present here for the first time quantitative data on the genetic variability of a plant RNA virus, as derived from randomly chosen genetic markers. The comparison of PMMV isolates representing five successive epidemics in greenhouse-grown peppers in southeastern Spain shows a high stability of the genome of this virus, thus supporting the view of a high genetic stability for field populations of RNA plant viruses.

MATERIALS AND METHODS

Virus isolates. Twenty-six isolates of PMMV were compared (Table 1). Twenty-three isolates collected from greenhouse-grown peppers in Almería (southeastern Spain) represent five epidemics (growing seasons 1983–1984 to 1987– 1988; here from 1983 to 1987). Two isolates collected in

TABLE 1. Details of PMMV isolates

Yr of Pepper isolation cultivar		Origin	Isolate name	Haplotype
1980	Lamuyo	Sicily (Italy) ^a	SR/80	I
1980	Lamuyo	Sicily (Italy) ^b	SW/80	II
1980	p68	Zaragoza (Spain) ^c	M03/80	11
1983	Skipper	Almería (Spain) ^d	P61/83	VI
1983	Skipper	Almería (Spain) ^d	P66/83	v
1983	p808	Almería (Spain) ^c	M09/83	IV
1984	Latino	Almería (Spain) ^c	M06/84	III
1984	Skipper	Almería (Spain) ^d	P77/84	VI
1984	Bellamy	Almería (Spain) ^d	P78/84	II
1985	Clovis	Almería (Spain) ^d	P132/85	v
1985	Latino	Almería (Spain) ^d	P138/85	v
1985	p808	Almería (Spain) ^c	P24/85	VII
1985	Latino	Almería (Spain) ^c	P37/85	VI
1985	Clovis	Almería (Spain) ^e	B23/85	VI
1985	Latino	Almería (Spain) ^e	B25/85	VI
1986	Latino	Almería (Spain) ^e	D15/86	VIII
1986	Red	Almería (Spain) ^e	D20/86	VIII
1986	Red	Almería (Spain) ^e	D22/86	VIII
1987	Latino	Almería (Spain) ^e	F01/87	IX
1987	Sonar	Almería (Spain) ^e	F14/87	VI
1987	Sonar	Almería (Spain) ^e	F18/87	VI
1987	Clovis	Almería (Spain) ^e	F22/87	IX
1987	Lamuyo	Almería (Spain)	4A1/87	Х
1987	Lamuyo	Almería (Spain) ^f	4A2/87	Х
1987	Lamuyo	Almería (Spain)	4A13/87	VI
1987	Orion	Almería (Spain) ^f	1C1/87	IX

" Provided by M. H. V. Van Regenmortel, Strasbourg, France.

^b Provided by C. Wetter, Saarbrücken, Federal Republic of Germany.

^c Provided by M. Luis, Zaragoza, Spain.

" Provided by J. R. Diaz-Ruiz, Madrid, Spain.

^e Provided by E. Rodríguez-Cerezo and F. García-Arenal, Madrid, Spain. ^f Provided by Sluis & Groot S.A., Almería, Spain.

^{*} Corresponding author.



FIG. 1. (I to X) Fingerprints of PMMV haplotypes I to X. Additional (\blacktriangleright) and missing (\bigcirc) oligonucleotides with respect to haplotype VI are indicated. (Bottom right panel) Universal fingerprint with all found invariant (\bigcirc) or polymorphic (\bullet) sites. (x) and (b), Positions of xylencyanol and bromophenol blue, respectively. Arrows indicate direction of first and second electrophoresis.

 TABLE 2. Annual distribution of isolates belonging to each haplotype

Year	Haplotype								Tota		
	I	II	III	IV	v	VI	VII	VIII	IX	x	no.
1980	1	2									3
1983				1	1	1					3
1984		1	1			1					3
1985					2	3	1				6
1986								3			3
1987						3			3	2	8

Sicily (Italy) and one collected in Zaragoza (central Spain) in 1980 were also included. Field isolates from individual pepper plants were cloned by single-lesion passaging in *Nicotiana glutinosa* L. and multiplied in *Physalis floridana* Rydb. Virions were purified from leaves, and the viral RNA was extracted by sodium dodecyl sulfate-phenol purification.

Genomic comparisons. Genomic RNAs (G-RNAs) were compared by two-dimensional fingerprinting of oligonucleotides generated by complete digestion with RNase T_1 (T_1 fingerprints) as described previously (10), except that 25% polyacrylamide gels were used for the second dimensions.

Similarities (shared oligonucleotides) and nucleotide substitutions among fingerprint types (haplotypes) were estimated by the method of Nei (17). The fingerprint type of each haplotype (a set of presences or absences of a given oligonucleotide pattern) were also used as input data to perform phylogenetic studies. We have followed Wagner criteria to develop a possible parsimonic unrooted tree, using a modified version of the Mix program included in the PHYLIP package (version 3.0) of J. Felsenstein, Seattle, Wash.

RESULTS

A typical fingerprint resolved about 45 oligonucleotides in the size range of 10 to 25 bases, representing about 9% of the G-RNAs (Fig. 1). Assuming that the resolved oligonucleotides are a random sample of the genome sequence and that any observed oligonucleotide change is due to a point mutation (7), the maximum divergence found among haplotypes (23 spots differing between haplotypes II and X) is equivalent to a sequence divergence in 4.1% of the G-RNAs.

Examination of the presence or absence of oligonucleotide spots compared with a universal pattern representing all observed spots (Fig. 1, bottom right panel) allowed us to



SIMILARITY COEFFICIENT

FIG. 2. UPGMA cluster analysis of PMMV haplotypes according to the similarity values in Table 3.

classify the fingerprints of the 26 isolates into 10 different haplotypes. Of a total of 62 oligonucleotides, 33 were polymorphic, the rest were invariant. The yearly distribution of haplotypes is shown in Table 2. Table 1 indicates the haplotype to which each isolate belongs.

The estimates of shared sites (oligonucleotides) and nucleotide substitution values (17) presented in Table 3 show a high homogeneity of types. The average (mean value of 45 comparisons) fraction of shared sites is 0.888 ± 0.011 , 0.931 ± 0.007 if types I and II are not considered (mean value of 28 comparisons). For all possible pairs, the mean value of nucleotide substitution is 0.132 ± 0.014 , 0.073 ± 0.008 when I and II haplotypes are excluded. UPGMA cluster analysis of the 10 haplotypes on the basis of the similarity coefficient values (Table 3) among pairs of haplotypes (Fig. 2) clearly shows the clustering of haplotypes I and II and III to X into two subgroups.

Figure 3 shows a possible Wagner unrooted tree that yields 39 mutational steps among the different haplotypes. As in the cluster analysis, haplotypes I and II are in a distant position relative to the rest. Haplotype VI appears to be in a central position; it is the more frequently found type (Table 2) and was present in each of the recorded epidemics. Its absence in 1986 can be explained by the limited sample size. All other types in groups III to X, which appeared in later years, can be derived from it.

Estimates for the intraannual and interannual nucleotide

TABLE 3. Similarity (shared sites [s] right upper half)" and nucleotide substitution values ([d] left lower half)^b among 10 PMMV haplotypes

Haplotype	I	II	111	ĪV	v	VI	VII	VIII	IX	x
I		0.952	0.828	0.828	0.805	0.814	0.835	0.800	0.837	0.761
11	0.040		0.771	0.771	0.771	0.780	0.782	0.744	0.805	0.727
III	0.192	0.266		0.953	0.930	0.941	0.867	0.921	0.918	0.879
IV	0.192	0.266	0.048	_	0.977	0.988	0.911	0.966	0.965	0.923
v	0.222	0.266	0.073	0.024		0.988	0.889	0.944	0.965	0.945
VI	0.210	0.253	0.061	0.012	0.012	_	0.899	0.955	0.976	0.933
VII	0.183	0.252	0.145	0.094	0.119	0.108		0.925	0.921	0.842
VIII	0.228	0.303	0.082	0.034	0.058	0.047	0.079	_	0.932	0.894
IX	0.180	0.221	0.087	0.036	0.036	0.024	0.082	0.071	—	0.911
х	0.280	0.328	0.130	0.080	0.057	0.069	0.174	0.114	0.094	

a = 2nxy/(nx + ny), where nxy = sites shared between haplotypes x and y; nx, ny = number of sites for x and y, respectively.

 $b d = -\operatorname{Ln}(s).$

TABLE 4. Intraannual nucleotide differences^a

Yr	Avg
1980	
1983	0.0020
1984	0.0305
1985	0.0404
1986	0.0000
1987	
Avg	0.0184
Avg (excluding 1980)	0.0186

^{*a*} Intraannual nucleotide difference for year x, δx being:

$$\delta x = \left(\frac{nx}{nx-1}\right)\sum_{ij} x_i x_j dij$$

where nx = number of haplotypes sampled at year x; xi, xj = frequency of haplotypes i and j in year x; dij = nucleotide substitution value between haplotypes i and j.

differences have been calculated from nucleotide substitutions and haplotype frequencies (17). Table 4 shows the results for intraannual nucleotide differences. In the first documented epidemic, this value is 0.0020, increasing in later years. This increase, though, is not monotonous with time but appears to have an upper limit. Interannual divergences (Table 5) again do not increase with time. For the Almería epidemics, these values are only about twice as large as those for annual divergences.

DISCUSSION

We present here for the first time quantitative data on the genetic variation of a plant RNA virus as derived from randomly chosen genetic markers (T_1 oligonucleotides). These data come from the comparison of 26 PMMV isolates, 23 of them representing five epidemics in peppers grown in southeastern Spain. According to their fingerprints, the 26 isolates may be classed into 10 haplotypes. Maximum divergence among haplotypes was equivalent to differences in 4.1% of the G-RNA sequences. This value is comparable to those reported for other RNA viruses, such as foot-and-mouth disease virus (7) or enterovirus EV70 (25). The significance of this datum becomes clear from an analysis of the variability and evolution of the population under study.

The nucleotide substitution (17) of the 10 haplotypes (Table 3) shows them to be very similar, although two subgroups, groups I and II and III to X, can be established. Cluster and cladistic analyses (Fig. 2 and 3) also confirm this grouping. This finding correlates with the origin of the isolates. Group III to X is formed from isolates sampled in

TABLE 5. Interannual corrected nucleotide difference"

Yr	1980	1983	1984	1985	1986
1983	0.2348				
1984	0.2168	0.0181			
1985	0.2732	0.0351	0.0369		
1986	0.2702	0.0453	0.0493	0.0493	
1987	0.2259	0.0202	0.0309	0.0338	0.0618

" Interannual corrected nucleotide difference, δ being

$$\delta = \delta xy - \left(\frac{\delta x + \delta y}{2}\right), \ \delta xy = \sum_{ij} x_i \ y_j \ dij$$

where xi and yj = frequencies of haplotypes i and j in years x and y, respectively; dij = nucleotide substitution value between haplotypes i and j (Table 3); x and y = intraannual nucleotide differences for years x and y (Table 4).



FIG. 3. Most parsimonic Wagner unrooted tree of 10 PMMV haplotypes. The year(s) in which a certain haplotype appears is indicated in parentheses.

pepper epidemics in Almería between 1983 and 1987; subgroup I and II is formed from isolates sampled in 1980 at such distant places as Sicily and Zaragoza, with isolate P78/84 from Almería (in group II) being the only exception. The phylogenetic analysis of the 10 haplotypes (Fig. 3) shows that haplotype VI is in a central position in group III to X (Almería isolates); haplotype VI, the prevalent type among Almería isolates, is found from 1983 until 1987, and all other types may be derived from it since they appeared in later years (Fig. 3, Table 2). Another unrooted tree, which is a bit less parsimonic (40 steps), places haplotype IV as the origin of the other types. Type IV corresponds to an isolate found in a seed batch of the cultivar associated by the growers with the origin of the 1983-1984 epidemic. Since haplotypes IV and VI are extremely similar, both trees are compatible with an epidemiological explanation that would consider the origin of the group III to X haplotypes and of the 1983 epidemic in infected seed. This conclusion is also compatible with the small variability observed among 1983 isolates (Table 4). The presence of type II in Almería in 1984 could be due to its introduction from a second source. PMMV was first detected in Almería in 1980 (15) even though the first serious epidemic did not occur until 1983. Thus, a different interpretation could consider haplotypes in group I and II as the prevalent ones for the 1980 to 1983 period, when they would have been quickly replaced by the newly introduced, more fit, type III to X haplotypes. Isolate P78/84 in type II would be a relict of the previously prevailing haplotypes. Since there are no extant isolates from 1980 to 1983 nor is there data on the relative fitness of haplotypes I and II versus III to X, this hypothesis cannot at present be tested.

According to our data, the PMMV population in Almería seems to be highly homogeneous and stable, as is also shown by the low intraannual and interannual variability rates found (Tables 4 and 5). The first recorded epidemic (1983) shows a variability value that is consistent with the circulating population coming from a single inoculum source; afterwards, the intraannual variability increases but the increase seems to be checked out at an upper limit. This situation clearly differs from that described for some animal RNA viruses, i.e., enterovirus EV70 (25). Also, interannual nucleotide differences suggest a certain but slow replacement of haplotypes with time, although haplotype VI prevails throughout.

Thus, an evolutionary model for PMMV would point to a highly stable population, maintaining its diversity through time, with a main prevailing haplotype from which variants arise that do not replace it. This model is clearly different from those of fast viral evolution, such as the one proposed for influenza A virus (5). It may be more similar to models that describe situations of genetic stability for RNA viruses, such as those reported for influenza B virus (30), influenza C virus (4), foot-and-mouth disease virus (7, 19, 24), or vesicular stomatitis virus (18); still, the interannual variations seem low for PMMV compared with those cases. It must be pointed out that due to differences in the studied populations and in the analytical methodology used the comparison of the variation rates reported for different systems is at least problematic. However, one major difference is the prevalence of an haplotype during the five PMMV epidemics, as opposed to a situation of types cocirculating with similar frequencies for the animal viruses cited.

The high frequency of variant haplotypes derived from a prevailing one is in agreement with the high potential to vary described for plant, bacterial, and animal viruses (12, 13, 21, 22, 27). On the other hand, the absence of haplotype replacement during the studied period agrees with views of the high stability of plant RNA viruses under field conditions (2; Gibbs, Abstr. Sixth. Int. Congr. Virol. 1984). These views have arisen from the study of viruses naturally infecting wild plants, and it is indeed remarkable that a similar picture derives from the study of epidemic populations in which variation is widely assumed to be higher due to selection pressures from the host (6). A clue to the observed stability of the PMMV genome may be given by the high proportion (two of three) of invariant sites in fingerprints; there seem to be strong functional restrictions that keep the ability of PMMV to vary inside tight limits. Variation would be restricted to parts of the genome that would not affect relevant functions.

ACKNOWLEDGMENTS

We thank J. R. Díaz-Ruíz, M. Luis-Arteaga, M. H. V. Van Regenmortel, and C. Wetter for providing isolates; Isabel Cuadrado and Julio Gómez (CIDA, Almería, Spain) for helping gather our field data; and J. Felsenstein, University of Washington, Seattle, for sending his PHYLIP package.

This work was in part supported by grants PR83-2032 from the Comisión Asesora para la Investigación Científica y Técnica and PB86-0517 from Dirección General de Investigación Científica y Técnica, Spain. E.R.-C. was in receipt of a Formación del Personal Investigador fellowship, MEC, Spain.

LITERATURE CITED

- 1. Bawden, F. C. 1956. Reversible, host-induced changes in a strain of tobacco mosaic virus. Nature (London) 177:302-304.
- 2. Blok, J., A. MacKenzie, P. Guy, and A. Gibbs. 1987. Nucleotide sequence comparisons of turnip yellow mosaic virus isolates from Australia and Europe. Arch. Virol. 97:283–295.
- Boukema, L. W. 1980. Allelism of genes controlling resistance to TMV in *Capsicum* L. Euphytica. 29:433–439.
- 4. Buonagurio, D. A., S. Nakada, W. M. Fitch, and P. Palese. 1986. Epidemiology of influenza C virus in man: multiple evolutionary lineages and low rate of change. Virology 153:12–22.
- Buonagurio, D. A., S. Nakada, J. D. Parvin, M. Krystal, P. Palese, and W. M. Fitch. 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in the NS gene. Science 232:980–982.
- Clarke, D. D., J. R. Bevan, and I. R. Crute. 1987. Genetic interactions between wild plants and their parasites, p. 195–207. *In* P. R. Day and G. J. Jellis (ed.), Genetics and plant patho-

genesis. Blackwell Scientific Publications, Ltd., Oxford.

- 7. Domingo, E., M. Davila, and J. Ortin. 1980. Nucleotide sequence heterogeneity of the RNA from a natural population of foot-and-mouth disease virus. Gene 11:333-346.
- Domingo, E., E. Martinez-Salas, F. Sobrino, J. C. de la Torre, A. Portela, J. Ortin, C. Lopez-Galindez, P. Perez-Breffa, N. Villanueva, R. Najera, S. VandePol, D. Steinhauer, N. DePolo, and J. J. Holland. 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. Gene 40:1–8.
- 9. Donis-Keller, H., K. S. Browning, and J. M. Clark. 1981. Sequence heterogeneity in satellite tobacco necrosis virus RNA. Virology 110:43-54.
- 10. García-Arenal, F., P. Palukaitis, and M. Zaitlin. 1984. Strains and mutants of tobacco mosaic virus are both found in virus derived from single-lession-passaged inoculum. Virology 132: 131-137.
- 11. Hennig, B., and H. G. Wittmann. 1972. Tobacco mosaic virus: mutants and strains, p. 546–587. In C. J. Kado and H. D. Agrawal (ed.), Principles and techniques in plant virology. Van Nostrand, Reinhold Co., Inc., New York.
- Holland, J. J. 1984. Continuum of change in RNA virus genomes, p. 137-143. In A. L. Notkins and M. B. A. Oldstone (ed.), Concepts in viral pathogenesis. Springer-Verlag, New York.
- Holland, J. J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol, and S. VandePol. 1982. Rapid evolution of RNA genomes. Science 215:1577-1585.
- 14. Kassanis, B., and R. D. Woods. 1969. Properties of some defective strains of tobacco mosaic virus and their behaviour as affected by inhibitors during storage in sap. Ann. Appl. Biol. 64:213-224.
- 15. Luis-Arteaga, M., and R. Gil. 1981. Primeras observaciones de cepas pimiento del TMV en pimiento. Información Técnica Economica Agraria 42:27-32.
- MacNeill, B. H., and M. Boxall. 1974. The evolution of a pathogenic strain of tobacco mosaic virus in tomato: a hostpassage phenomenon. Canad. J. Bot. 52:1305–1307.
- 17. Nei, M. 1987. Molecular evolutionary genetics, p. 97–105. Columbia University Press, New York.
- Nichol, S. T. 1987. Molecular epizootiology and evolution of vesicular stomatitis virus New Jersey. J. Virol. 61:1029–1036.
- Piccone, M. A., G. Kaplan, L. Giavedoni, E. Domingo, and E. L. Palma. 1988. VP1 of serotype C foot-and-mouth disease virus: long-term conservation of sequences. J. Virol. 62:1469–1473.
- Randles, J., P. Palukaitis, and C. Davies. 1981. Natural distribution, spread, and variation in the tobacco mosaic virus infecting *Nicotiana glauca* in Australia. Ann. Appl. Biol. 98:109–119.
- Reanney, D. C. 1982. The evolution of RNA viruses. Annu. Rev. Microbiol. 36:47-73.
- Reanney, D. C. 1984. The molecular evolution of viruses, p. 175–196. In B. W. J. Mahy and J. R. Pattison (ed.), The Microbe. Part I: Viruses. Cambridge University Press, Cambridge.
- 23. Smith, D. B., and S. C. Inglis. 1987. The mutation rate and variability of eukaryotic viruses: an analytical review. J. Gen. Virol. 68:2729-2740.
- 24. Sobrino, F., E. L. Palma, E. Beck, M. Davila, J. C. de la Torre, P. Negro, N. Villanueva, J. Ortin, and E. Domingo. 1986. Fixation of mutations in the viral genome during an outbreak of foot-and-mouth disease: heterogeneity and rate variations. Gene 50:149–159.
- Takeda, N., K. Miyamura, T. Ogino, S. Natori, S. Yamazaki, S. Sakurai, N. Nakazono, K. Ishii, and R. Kono. 1984. Evolution of enterovirus type 70: oligonucleotide mapping analysis of RNA genome. Virology 134:375–389.
- 26. Thresh, J. M. 1987. The population dynamics of plant virus diseases, p. 135–149. *In* M. S. Wolfe and C. E. Caten (ed.), Populations of plant pathogens. Their dynamics and genetics. Blackwell Scientific Publications, Ltd., Oxford.
- Van Vloten-doting, L., and J. F. Bol. 1988. Variability, mutant selection, and mutant stability in plant RNA viruses, p. 37-51. *In E. Domingo, J. J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. 3. CRC Press, Boca Raton, Fla.*

- Van Vloten-doting, L., J. F. Bol, and B. Cornelissen. 1985. Plant virus based vectors for gene transfer will be of limited use because of the high error frequency during viral RNA synthesis. Plant. Mol. Biol. 4:323–326.
- 29. Wetter, C., M. Conti, D. Altschuh, R. Tabillion, and M. H. V. Van Regenmortel. 1984. Pepper mild mottle virus, a tobamovi-

rus infecting pepper cultivars in Sicily. Phytopathology 74: 405-410.

30. Yamashita, M., M. Krystal, W. M. Fitch, and P. Palese. 1988. Influenza B virus evolution: co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. Virology 163:112–122.