Variability and genetic structure of the population of watermelon mosaic virus infecting melon in Spain

I.M. Moreno,[a] J.M. Malpica,[b] J.A. Díaz-Pendón,[c] E. Moriones,[c] A. Fraile,[a] and F. García-Arenal[a],*

[a] *Departamento de Biotecnología, E.T.S.I. Agrónomos, Universidad Politécnica de Madrid, 28040 Madrid, Spain*
[b] *Departamento de Biotecnología, Instituto Nacional de Tecnología Agraria y Alimentaria, 28040 Madrid, Spain*
[c] *Laboratorio de Virología Vegetal, Estación Experimental "La Mayora"-CSIC, 29750 Algarrobo-Costa, Málaga, Spain*

Abstract

The genetic structure of the population of *Watermelon mosaic virus* (WMV) in Spain was analysed by the biological and molecular characterisation of isolates sampled from its main host plant, melon. The population was a highly homogeneous one, built of a single pathotype, and comprising isolates closely related genetically. There was indication of temporal replacement of genotypes, but not of spatial structure of the population. Analyses of nucleotide sequences in three genomic regions, that is, in the cistrons for the P1, cylindrical inclusion (CI) and capsid (CP) proteins, showed lower similar values of nucleotide diversity for the P1 than for the CI or CP cistrons. The CI protein and the CP were under tighter evolutionary constraints than the P1 protein. Also, for the CI and CP cistrons, but not for the P1 cistron, two groups of sequences, defining two genetic strains, were apparent. Thus, different genomic regions of WMV show different evolutionary dynamics. Interestingly, for the CI and CP cistrons, sequences were clustered into two regions of the sequence space, defining the two strains above, and no intermediary sequences were identified. Recombinant isolates were found, accounting for at least 7% of the population. These recombinants presented two interesting features: (i) crossover points were detected between the analysed regions in the CI and CP cistrons, but not between those in the P1 and CI cistrons, (ii) crossover points were not observed within the analysed coding regions for the P1, CI or CP proteins. This indicates strong selection against isolates with recombinant proteins, even when originated from closely related strains. Hence, data indicate that genotypes of WMV, generated by mutation or recombination, outside of acceptable, discrete, regions in the evolutionary space, are eliminated from the virus population by negative selection.
© 2003 Elsevier Inc. All rights reserved.

Keywords: Potyvirus; Recombination; Sequence space; Genetic variation; Biological variation

Introduction

Understanding the processes that determine the evolution of pathogens is an important part of evolutionary biology with consequences for the control of pathogen-caused diseases. During the last 15 years, plant virologists have become increasingly aware of the interest of studying virus evolution, as evidenced by the abundance of publications on this subject (reviewed in García-Arenal et al., 2001). Mutation and genetic exchange (recombination and reassortment of genomic segments) are processes that fuel evolution by the generation of genetic variation, but the genetic composition of the virus population will depend on evolutionary forces as selection and random genetic drift. RNA viruses, which are the largest fraction of plant viruses, have very high mutation rates (Drake and Holland, 1999; Malpica et al., 2002) and, most probably, also high rates of recombination (Aaziz and Tepfer, 1999; Bruyere et al., 2000). However, genetic stability is the rule for plant viruses, and the accumulation of mutations seems not to be favoured (García-Arenal et al., 2001). Recombination has been shown to have an important role in the speciation of plant viruses, particularly in some taxa (Gibbs, 1995; Guyader and Ducray, 2001; MacFarlane, 1997; Padidam et al., 1999), but its role in the evolution of particular viral species has been evaluated less often. The analysis of the genetic structure of plant virus populations may provide information to understand the genetic stability of most plant RNA

* Corresponding author. Fax: +34-913365757.
E-mail address: fga@bit.etsia.upm.es (F. García-Arenal).

viruses in spite of their high potential to vary. Comparison of nucleotide sequence data from plant virus isolates of different geographic origins has allowed the analysis of the genetic structure for some plant virus species over most of their geographic range (e.g., Abubakar et al., 2003; Bateson et al., 2002; Bousalem et al., 2000; Fraile et al., 1996; Higgins et al. 1999; Krause-Sakate et al., 2002; Ohshima et al., 2002). Analyses of the genetic structure of viral populations at the local or regional scale have also been reported (e.g., Azzam et al., 2000; Desbiez et al., 2002; Fraile et al., 1997; McNeil et al., 1996; Rodríguez-Cerezo et al., 1991; Skotnicki et al., 1993) and provide data of another level of detail, which should be relevant to interpret those from large-scale analyses.

The genus *Potyvirus* is the largest taxon of plant viruses and includes many economically important species (Shukla et al., 1994). Reports on the large-scale structure of populations of potyviruses show that they are often built of different strains. Incongruent evolutionary relationships among isolates according to the genomic region considered have been reported for many potyvirus species, and recombination events have been invoked to explain their evolution (Alegria et al., 2003; Bousalem et al., 2000; Higgins et al., 1998; Moury et al., 2002; Ohshima et al., 2002; Revers et al.,1996). It is of interest to evaluate the impact of recombination in shaping potyvirus populations. However, the structure of potyvirus populations has rarely been analysed at the local scale, so that evolutionary processes that may determine the population structure of the species are often unknown. We present here such an analysis for *Watermelon mosaic virus* (WMV) infecting melon in Spain. Results showed that distinct genotypes existed and recombinants among them occurred. Evolutionary dynamics varied for different genomic regions. Interestingly, genetic variation was restricted for all genomic regions, so that only discrete areas of the sequence space were occupied.

Results

Biological characterisation of the WMV population

The possibility that the WMV population was built of different pathotypes was analysed. For this, 19 isolates were assayed on 13 different hosts from eight species in four dicotyledonous families (see Material and methods). Biological typing included analyses of virus presence in inoculated and upper leaves, and differences in symptom induction. Results were very similar for all analysed virus isolates in each host combination studied: all assayed isolates infected systemically the seven varieties and accessions of melon and squash, causing systemic mosaics and leaf lamina deformation in melon, and systemic mosaic, leaf lamina deformation and vein clearing in squash, except isolates MAD95.7 and MAD99.2 that caused a systemic necrotic mottle in melon PI414723, a genotype used as a

source of resistance to WMV (Moyer et al., 1985). Cucumber Kyoto three feet, with a resistance factor to WMV (Cohen et al., 1971), was systemically infected by all 19 isolates without inducing symptoms. All isolates infected systemically *Nicotiana benthamiana* plants, causing very mild mosaics. *Chenopodium amaranticolor* reacted to all isolates by chlorotic local lesions in the inoculated leaves, and *Chenopodium quinoa* exhibited chlorotic local lesions in the inoculated leaf and systemic mottle, except for isolates BAD95.3, SEG99.2 and ZAR99.2, which did not become systemic. Last, all isolates systemically infected pea cv Resal, without causing symptoms, and none infected pea cv. Bolero or bean cv. Great Northern 1140. This is noteworthy as Bolero pea has the *mo* resistance gene to WMV, expressed as immunity (Schroender and Provvidenti, 1971), and Great Northern 1140 bean has the *Wmv* resistance gene to WMV (Kyle and Provvidenti, 1987). Hence, in spite of minor differences in pathogenicity for some of the assayed hosts, all isolates belonged to the same pathotype, as judged by their reaction on the well-characterised resistances of cucumber, bean and pea.

Molecular characterisation of isolate VAL95.1

When this work was undertaken, databases only reported sequences for the capsid (CP) and part of the NIb cistrons of WMV. Thus, the nucleotide sequence of part of the approximately 10,000-nt-long genomic RNA of WMV was determined from a cDNA library constructed using as template RNA of isolate VAL95.1. This information was used to delimit regions of the WMV genome to be compared with the 44 isolates, and to design primers for RT-PCR. Three regions were delimited in three segments spanning the whole genomic RNA. The first segment was 708 nt long and BLAST searches in nucleotide sequence databases showed that it corresponded to the N-terminal half of the potyviral P1 protein cistron. For VAL95.1, similarity was highest for Bean common mosaic virus (BCMV) and *Peanut stripe virus* (PStV) sequences (GenBank Accession No. AY112735, AJ312438, U34972), but colinearity was not complete. Segment two was 702 nt long and roughly corresponded to the C-terminal half of the potyviral cylindrical inclusion (CI)-protein cistron. For VAL95.1, this segment was most similar to *Soybean mosaic virus* (SMV) (Accession No AY169284, AJ310200), but again, colinearity was not complete. The third segment was 873 nt long and included the complete coding sequence for the WMV CP (843 nt) and 30 nt in the cistron for the NIb protein. For VAL95.1, the CP cistron was >92% identical to that of WMV isolates in databases (Accession No. D00535, L22907, AB001994), and colinearity was complete.

Genetic diversity of WMV in Spain

The genetic diversity of the Spanish population of WMV was analysed by comparing the nucleotide sequences of

three selected regions in the 5′, central and 3′ regions of WMV genomic RNA, that is, the P1, CI and CP regions defined in the previous paragraph. A total of 132 RT-PCR-amplified fragments for these three genomic regions in 44 isolates were cloned and sequenced. Sequence data are available under accession numbers AJ577521–AJ577564 for P1, AJ577477–AJ577520 for CI and AJ579481–AJ579524 for CP fragments.

Genetic distances for each pair of isolates were estimated for each genomic region analysed by Kimura's two-parameter method. Genetic distance, averaged over all isolates, was 0.0258 when the concatenated sequences of the P1, CI and CP regions were considered. Nucleotide diversity was 0.0184 for the P1 region, 0.0308 for the CI region and 0.0232 for the CP region (Table 1). Isolate BAD95.3 was excluded from these analyses because of a double frameshift in the P1 region. However, when CI and CP regions were analysed separately including isolate BAD95.3, data did not vary. Nucleotide diversity values in the CI and CP regions did not differ significantly ($P = 0.698$ in a Wilcoxon nonparametric test), but the diversity of the P1 region was significantly less than for both the CI and CP regions ($P < 0.0000$). Nucleotide diversities within and among subpopulations were estimated (Nei, 1987) for all three genomic regions considering the WMV population divided into either two geographic (Mediterranean Coast and Central Spain) or two temporal subpopulations (1995 and 1999). Results did not provide evidence for population subdivision (data not shown).

Pairwise genetic differences at synonymous (d_S) and nonsynonymous (d_N) nucleotide positions were estimated according to the Pamilo, Bianchi and Li (PBL) method. The d_N/d_S ratio (Table 1) was significantly higher for the P1 region (0.620) than for the CI and CP regions (0.076 and 0.078, respectively), indicating that the CI and CP proteins are under tighter functional constraints than the P1 protein.

Phylogenetic relationships among the 44 analysed WMV isolates were inferred for the three analysed genomic regions using the NJ method (Fig. 1). Results differed according to the region considered: for the CP and CI regions, two major clusters were apparent, providing evidence for two genetic strains (strain I and strain II, Fig. 1), while for the P1 region, isolates did not significantly split into different clusters (Fig. 1A). For both the CI and the CP regions, the isolates clustering into the most frequent strain II can be divided into two subgroups. Isolates from 1995 and 1999 were present in both subgroups (Figs. 1B and C), but their frequencies were significantly different ($P < 0.01$ in a contingency test). This showed that there was a replacement of isolates with time, although between population diversity analyses did not provide evidence for population subdivision, probably because of close genetic distances among all isolates. The frequency of isolates from the two geographic subpopulations did not differ significantly between these two subgroups.

When pairwise genetic distances for the CP and CI region were plotted against each other, four discrete clusters of values were found (Fig. 2), with no points (i.e., pairs of isolates) among them. This indicates that the nucleotide sequences of WMV isolates occupy two discrete regions of the sequence space, and no sequences are found in intermediary positions between those two regions.

Recombination between WMV strains

Phylogenetic analyses in Figs. 1B and C showed that six isolates clustered into strain I or strain II depending on the genomic region, CI or CP, analysed. This was the case of isolates BAD95.3, MAL99.1, MAL99.2, MAL99.3, SG99.1 and SG.99.2. Phylogenetic incompatibility of this type is usually interpreted as evidence of recombination, and suggested that the above six isolates could be recombinants between strains I and II. This possibility was further analysed using two methods for detecting recombination based on different theoretical assumptions. First, Sawyer's test for gene conversion was used with 10,000 simulations. This test analyses the distribution over the pairs of sequences of the length of the fragments defined by the discordances (i.e., polymorphisms) versus the null hypothesis that the distribution of that length is random. When the concatenated sequences of the CI and CP regions of the 44 isolates were so analysed, the probability of the null hypothesis was smaller than 10^{-4}. The probability of the null hypothesis was smaller than 0.015 for all pairwise sequence comparisons involving each of the six putative recombinant isolates above. Hence, Sawyer's test analyses concluded that the distribution of polymorphic sites among the CI–CP sequences of the 44 WMV isolates analysed was most probably due to recombination, and identified six putative recombinant isolates. When the P1–CI concatenated sequences

Table 1
Origin of WMV isolates analysed in this work[a]

Region	Location	1995	1999	Total
Mediterranean Coast	Barcelona	BAR95.1, BAR95.2	BAR99.1– BAR99.5	7
	Valencia	VAL95.1, VAL95.2	–	2
	Murcia	MUR95.1– MUR95.3	–	3
	Malaga	–	MAL99.1– MAL99.5	5
	Subtotal	7	10	17
Central Spain	Segovia	–	SG99.1– SG99.4	4
	Madrid	MAD95.1– MAD95.7	MAD99.1– MAD99.4	11
	Badajoz	BAD95.1– BAD95.4	–	4
	Zaragoza	ZAR95.1– ZAR95.3	ZAR99.1– ZAR99.5	8
	Subtotal	14	13	27
Total		21	23	44

[a] Each isolate was identified by location, year of sampling and series number.
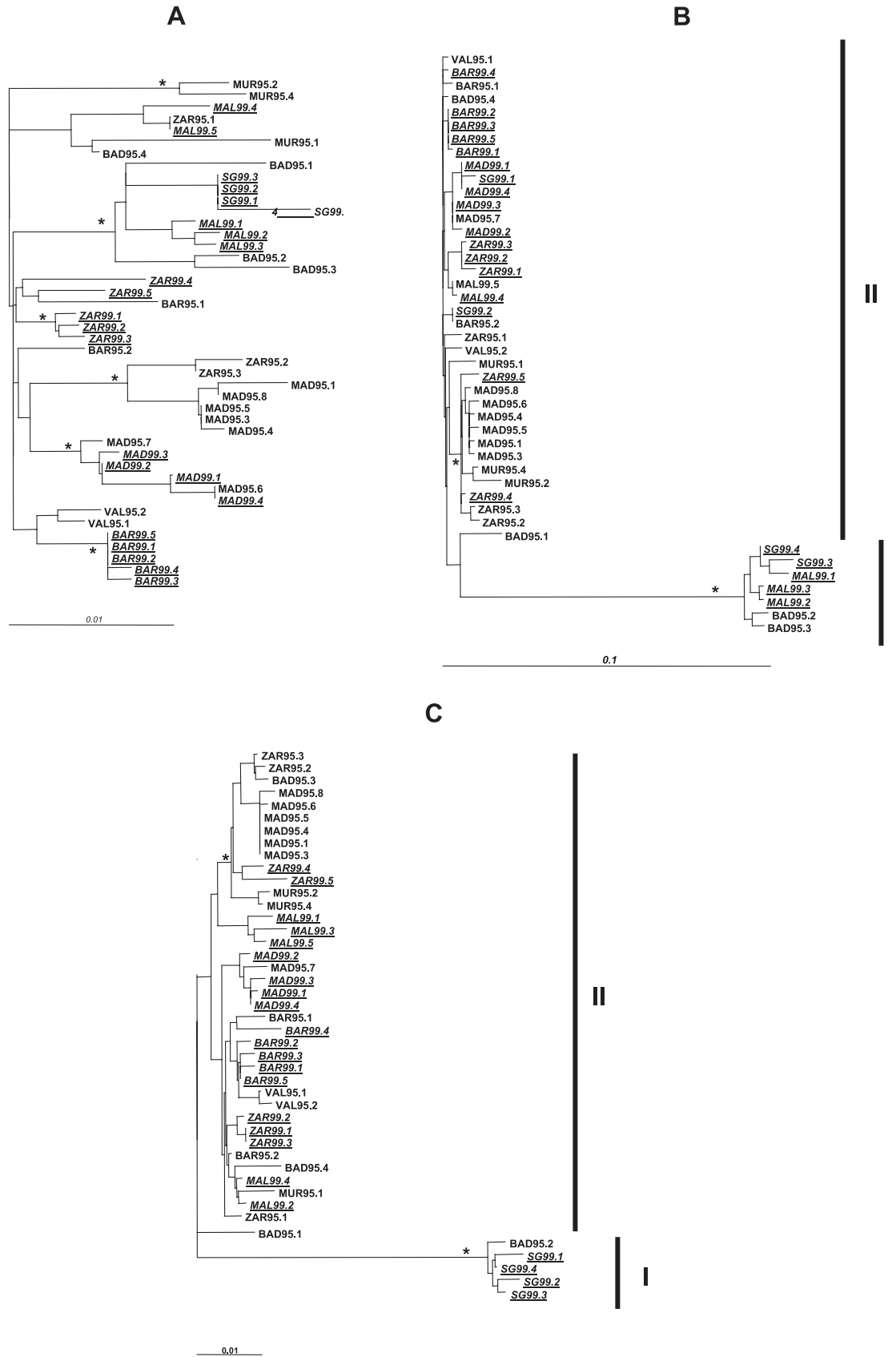
Fig. 1. Neighbor-joining unrooted phylogenetic trees for WMV isolates based on sequences in the P1 protein (panel A), cylindrical inclusion protein (panel B) and capsid protein (panel C) cistrons. The significance of nodes in a bootstrap analysis with 1000 replicates higher than 75% is shown with asterisks higher than 75. Isolates from 1999 are shown in italics and underlined.
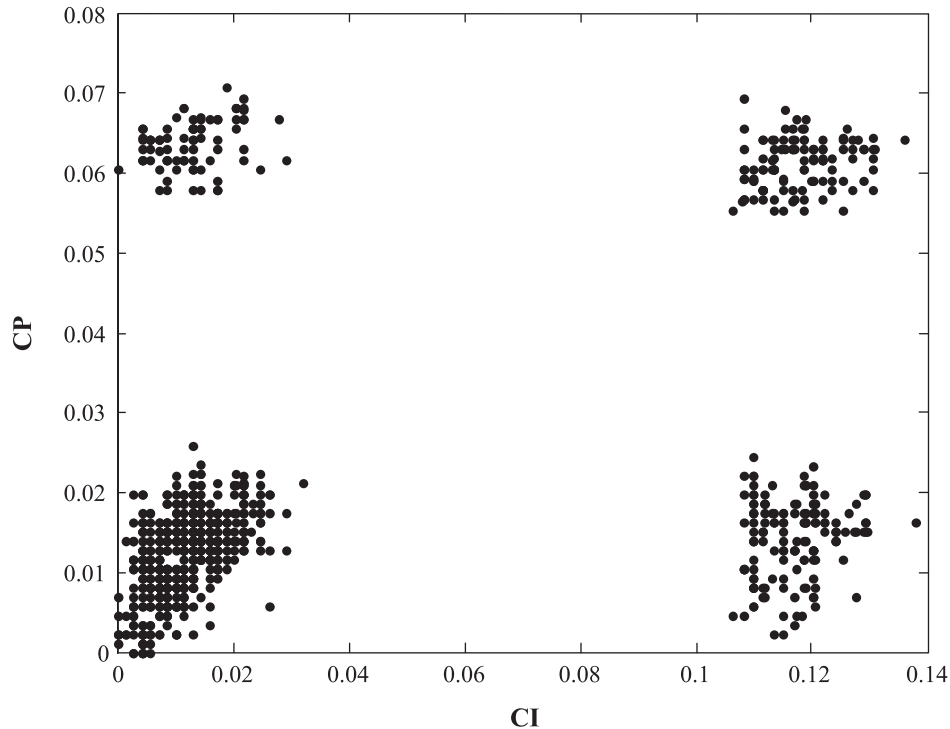
Fig. 2. Plot of pairwise genetic distances (Kimura two-parameter method) between sequences in the capsid protein cistron (CP) and sequences in the cylindrical inclusion cistron (CI).

were analysed, the probability that gene conversion had occurred was again significant ($P < 10^{-4}$ for the null hypothesis of random distribution of fragment length), and in pair wise comparisons, it was again significant ($P < 0.05$) for some, but not all, comparisons including isolates SG99.1, SG99.2, MAL99.1, MAL99.2 and MAL99.3. When each of the P1, CI and CP regions were analysed separately, no evidence for gene conversion was found.

Second, phylogenetic incongruence was tested by the method described by Jakobsen and Eastel (1996). This method is based in the analysis of phylogenetic incompatibility between pairs of informative positions within a set of aligned nucleotide sequences, scored against a randomised set. The global score from the analysis of 1000 random matrices was significant ($P < 0.02$) when all three regions (the concatenated sequences P1–CI–CP) were analysed, and when the concatenated CI–CP regions were analysed, but not when the concatenated P1–CI regions were analysed or when each region was analysed separately. Hence, this test detected recombination between the CI and CP regions, but not between the P1 and CI regions, or within the P1, CI and CP regions. The distribution of the probability of recombination at each site is shown in Fig. 3 for the whole P1–CI–CP dataset. Results showed that recombination between the CI and CP regions was not detected within the sequenced regions.

Sawyer's and Jacobsen's methods have been reported as most efficient for detecting recombination (Posada and Crandall, 2001). Thus, joint analyses using these methods provide strong evidence of recombination at some part of the

genome between the analysed CI and CP regions, but not within any of the sequenced regions, which evolve as a unit.

Evidence for recombination between the CI and the CP regions was also provided by the data in Fig. 2: the four well-defined groups observed when the pairwise distances using the CP sequences were plotted against the pairwise distances calculated with the CI sequences corresponded to
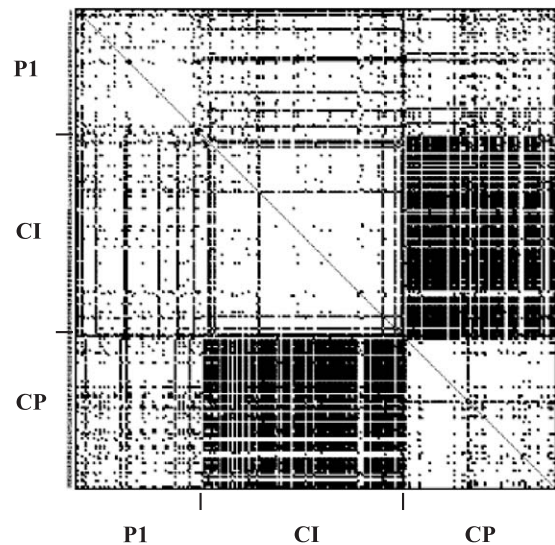

Fig. 3. Distribution of phylogenetically incompatible informative positions along the concatenated sequences of the P1–CI–CP cistrons. Phylogenetically incompatible positions according to Jakobsen and Eastel (1996) are represented by black dots. Limits among the P1, CI and CP sequences are indicated.

comparisons between isolates within strains I and II (low values for both axes), between isolates belonging to each of the two strains (high values for both axes), and between putative recombinant isolates and isolates in any of the two strains (high value for one axis and low value for the other).

Discussion

Understanding the factors that determine the evolution of pathogens is a central aspect of evolutionary biology with potentially important consequences for the control of pathogen-caused diseases. The characterisation of the genetic structure of pathogen populations is a powerful tool for unravelling those evolutionary factors, as their action determines the genetic structure. Analyses of population structure at the local scale, less frequently reported than those at the regional or continental scale, allow a higher level of detail that may be particularly informative. We report here such an analysis for the population of WMV infecting melon in Spain. WMV is the most prevalent *Potyvirus* causing mosaics in melon worldwide (Grafton-Cardwell et al., 1996), Spain being no exception (Luis-Arteaga et al., 1998). WMV has a relatively wide host range infecting plants in 23 families of dicotiledonous species, and occurs naturally in cucurbit and legume crops (Purcifull et al., 1984). Comparison of the sequences of the P1, CI and CP cistrons with those homologous from other potyvirus species indicates that WMV is most similar to legume-infecting potyviruses. This agrees with the broad host range of WMV within the legumes (Purcifull et al., 1984) and with the frequency of resistance to this virus in legume species (Kyle and Provvidenti, 1987; Schroeder and Provvidenti, 1971). In Spain, WMV is very infrequently found infecting noncucurbit hosts (Sacristán et al., 2001 and our unpublished results) and melon is by far the most important cucurbit crop. Hence, the WMV population should be well represented by isolates sampled from melon, and in this work, care was taken that all the geographic areas and agroecological conditions where melon is grown were represented.

Population structure was analysed on isolates biologically cloned through single-lesion passage. This was necessary because of the high frequency of mixed infections in the field with other viruses and, possibly, with different strains of WMV (Luis-Arteaga et al., 1998). Biological cloning was then necessary to interpret sequences from three noncontiguous genomic regions (i.e., from the P1, CI and CP cistrons) as part of the same genome. Although changes in within isolate sequence composition can occur during single lesion passage, this should not affect our results: we analyse the WMV population from Spain from the characterisation of 44 isolates randomly sampled. The biologically cloned isolate or, for that matter, the molecularly cloned genomic fragment used for sequence determination, should not affect

the quality of the sample, as cloning would pick an individual from the within isolate population.

The results of the molecular characterisation of 44 isolates and of the biological characterisation of a subset of 19 isolates randomly sampled from those 44 show a highly homogeneous population. All isolates belonged to the same pathotype, as judged by their ability to infect those hosts (melon, cucumber, bean and pea) with well-characterised resistance factors (Cohen et al., 1971; Kyle and Provvidenti, 1987; Moyer et al., 1985; Schroender and Provvidenti, 1971). Also, nucleotide diversity was very low for the analysed population. No evidence of spatial differentiation of the population was found. Data suggest that there may be a temporal differentiation. Data show that the population of WMV in Spain has no metapopulation structure with local extinction and recolonisation, as is the case for *Cucumber mosaic virus* (CMV) (Fraile et al. 1997, García-Arenal et al., 2000), another aphid-transmitted virus frequently found infecting melon (Luis-Arteaga et al., 1998).

In spite of the low genetic diversity of the WMV population, two genetic strains, strain I and strain II, were apparent. Interestingly, these two strains were detectable when sequences from the 3′-half of the genome, that is, from the CI and CP cistrons, were analysed, but not when the P1 cistron, most 5′ wards, was analysed. Hence, data show different evolutionary dynamics for different regions in the genome of WMV. It should be pointed that for *Turnip mosaic virus* (TuMV), *Potato virus Y* (PVY) and *Potato virus V,* for which sequence data on different genomic regions have been reported (Moury et al., 2002; Ohshima et al. 2002; Oruetxebarria et al. 2000), the P1 cistron was neither less variable nor was it under lower evolutionary constraint than the CP cistron. However, these data should be compared with ours with caution due to differences in the analysed samples. Different evolutionary dynamics for different genomic regions have been described for other plant RNA viruses, a phenomenon that has been analysed in detail for *Citrus tristeza virus* (e.g., D'Urso et al., 2003; López et al., 1998) and has been associated to stronger functional constrains on the less variable regions. However, the fact that differentiation in two strains does not occur for the P1

Table 2

Nucleotide diversity for three genomic regions of WMV isolates in Spain

	Genomic region[a]			
	P1	CI	CP	Concatenated
d[b]	0.0184	0.0380	0.0232	0.0258
d_N[b]	0.0151	0.0096	0.0054	0.0097
d_S[b]	0.0267	0.1173	0.0670	0.0649
d_N/d_S[b]	0.566	0.082	0.081	0.149

[a] The analysed regions are 709 nt in the P1 region, 702 nt in the CI region, 873 nt in the CP region and the three sequences together (concatenated).
[b] d = nucleotide diversity according to Kimura's two-parameter method. d_N and d_S = nucleotide diversity at nonsynonymous and at synonymous position, respectively, according to the PBL method.

protein cistron is not due to this genomic region being under a stronger purifying selection than the CI or CP cistrons, as the data in Table 2 clearly shows. Note that a higher diversity with stronger selection is possible if selection results in several genetically distinct groups. In our case, this occurs for the CI and CP regions, and intragroup diversity for these regions is even smaller than the diversity for the P1 region (e.g., 0.012 and 0.013 for the CI and CP, respectively, for strain II).

A noteworthy phenomenon is evident from the comparison of the CI and CP cistron sequences of the analysed isolates: sequences cluster into two discrete areas of the sequence space, and no intermediary states were found (Fig. 2). Because mutation rates for RNA viruses are high, and mutations would be random (Drake and Holland, 1999; Malpica et al., 2002), it could have been expected that the sequences of the analysed isolates will be continuously distributed over the sequence space. Thus, in this case, selection seems to prevent mutational "smear", as it would have been predicted under the quasi-species model of virus evolution. Again, this was not the case for the P1 cistron. Because the P1 protein is less constrained evolutionarily than the CI or CP (Table 2), it is possible that a higher number of sequence states are accepted and, hence, two fitness regions, corresponding to two strains, do not occur.

Recombination may be an important source of genetic variation for WMV, with at least 7% of the analysed isolates being recombinants between the two detected genetic strains I and II. The observed recombinants presented two interesting features: (i) crossover points were not detected over the whole genome: they occurred between the analysed regions in the CI and CP cistrons, but not between the P1 and CI cistrons, (ii) crossover points were not observed within the analysed regions coding for the P1, CI or CP proteins (Fig. 3). This is at odds with most reported analyses for potyviral species, such as BCMV, PStV, PVY, TuMV or *Yam mosaic virus*, for which recombinants within the CP were frequently found using procedures comparable to those employed here (Bousalem et al, 2000; Moury et al., 2002; Ohshima et al., 2002; Revers et al. 1996). However, recombinants within the CP cistron are infrequent, or have not been reported for other species in the genus, as is the case of *Papaya ringspot virus*, *Lettuce mosaic virus* or Sugarcane mosaic virus (Alegria et al., 2003; Bateson et al., 2002; Krause-Sakate et al., 2002). Recombination affecting other potyviral proteins, including the P1 protein, has also been reported for PVY and TuMV (Glais et al., 2002; Moury et al., 2002; Ohshima et al., 2002). These data from different potyviral species show that particular recombinant genotypes can occur having a relative fitness enough for being maintained in the virus population, but that the probability of their occurrence is not the same for all virus species, even those from a single genus. For WMV, data show that each of the analysed cistrons behaves as an evolutionary

unit, and that recombinants that exchange parts of the same protein among different strains are selected against the virus population so that they are not detected in a small-scale analysis. It is to be noted that selection against intracistronic recombinants operates even when parentals are so closely related as strains I and II in this work. It would be interesting to explore how general is this phenomenon: one major concern for the widespread use of transgenic plants with pathogen-derived resistance to viruses is the ecological risk of transgene escape though recombination with infecting viruses (Tepfer, 2002). Selection against part of the possible recombinants, as shown here for WMV, should be considered when designing transgenes and when evaluating ecological risks of the transgenic plants.

The results of the analysis of the genetic variability and structure of the WMV population in Spain add to our current view of high genetic conservation of plant RNA viruses (García-Arenal et al., 2001) and contribute to understand why this is so in spite of the high potential for genetic variation both through mutation and recombination of RNA viruses (Aaziz and Tepfer, 1999; Bruyere et al., 2000; Drake and Holland, 1999; Malpica et al., 2002). For WMV only sequences in limited, discrete regions of the evolutionary space are acceptable. Sequences outside those regions, generated either by mutation or recombination, would be eliminated from the virus population by selection.

Materials and methods

Virus isolates and biological cloning

WMV isolates were collected from melon plants (*Cucuumis melo* L.) in commercial fields. Melon samples were randomly selected among plants showing mosaic-like symptoms in every field and WMV positive samples were identified by double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) using polyclonal antisera from Loewe Biochemica GmbH (Otterfing, Germany) according to manufacturer's instructions. Samples were collected from all the major areas of melon production in Spain, in 1995 and 1999. The original locations where the samples were collected were grouped according to agro-ecological considerations in two regions: the Mediterranean Coast region (temperate with mild winters, melons grown in the spring) and Central Spain (Southern Central plateau and Ebro river valley, temperate with cold winters, melons grown in the summer). From this sample set, 44 specimens were randomly chosen, and WMVs infecting these 44 specimens were biologically cloned by two local lesion passages in *C. amaranticolor*, followed by a single local lesion passage in *C. quinoa*. Such biologically cloned viruses are hereafter called isolates, and after multiplication in squash plants they were used for biological and molecular

characterisation. The isolates were named and numbered according to their geographical origins and to the year of collection (Table 1). VAL95.1 was taken as the reference isolate.

Biological characterisation of the WMV population

Eighteen randomly sampled isolates, plus the reference isolate VAL95.1, were characterised biologically by reaction to inoculation of a variety of hosts. The 18 isolates were BAR95.1, BAR95.2, BAR99.3, BAR99.5, MUR95.2, MUR95.4, MAL99.1, MAL99.5, BAD95.1, BAD95.3, MAD95.1, MAD95.4, MAD95.7, MAD95.8, MAD99.2, SG99.2, ZAR95.1 and ZAR99.2. These isolates were inoculated mechanically into 13 cultivars or accessions in eight plant species of four botanical families: *C. melo* (melon, accession PI 414723 and cvs. Panal, and Bola de Oro), *C. sativus* (cucumber, cv. Kyoto three feet) *Cucurbita pepo* (squash, cvs. Diamante, Afrodita and Sofia), *Phaseolus vurlgaris* (bean, accession Great Northern 1140)*, Pisum sativum* (pea, cvs. Bolero and Resal), *N. benthamiana*, *C. amaranticolor* and *C. quinoa.* Inocula consisted of freshly ground, systemically infected squash leaves in 20 mM Na-phosphate, 0.03% DIECA and at least two plants were inoculated for every isolate–host combination. After inoculation, the type of symptoms and the timing of symptom appearance were recorded. Accumulation of viral RNA in the inoculated and upper leaves was also analysed by dot-blot hybridisation of nucleic acid extracts from inoculated and upper leaves (Sambrook and Russel, 2001) with a ^{32}P-labelled RNA probe complementary to the CP cistron of isolate VAL95.1. Each extract was scored as positive or negative for WMV RNA by comparison with proper controls.

Cloning and partial molecular characterisation of isolate VAL95.1

The purified RNA of the reference isolate was used as template for the construction of a cDNA library. WMV particles were extracted from 50 g (fresh weight) of infected squash tissue by grinding in 300 ml of 500 mM potassium phosphate buffer pH 7.5, 5 mM EDTA, 10 mM diethyl-dithyocarbamate, 20 mM Na_2SO_3 and gently stirring for 1 h at 4 °C. Then the slurry was strained through four layers of cheesecloth, clarified by low speed centrifugation (10,000 × g, 10 min) and the supernatant stirred for 1 h at 4 °C in the presence of 1% Triton X-100. The cleared homogenate was layered over a 30% sucrose cushion and ultracentrifuged (1 h, 185,000 × g). The pellet was resuspended in 4 ml of 50 mM potassium phosphate buffer pH 8.2, 10 mM EDTA by gently stirring for 1 h at 4 °C. After low speed centrifugation, the supernatant was mixed with 0.26 g/ml of Cs_2SO_4, a density gradient was formed by ultracentrifugation for 16 h at 75,000 × g, in which WMV particles concentrated in an opaque band that was collected using a syringe. The

preparation was 3-fold diluted in 50 mM potassium phosphate buffer pH 8.2 and ultracentrifuged (1 h, 185,000 × g). The pellet contained purified WMV particles and was resuspended in 200 μl of TE pH 8.0. WMV RNA was purified from viral particle preparation by incubation for 1 h at 37 °C in 10 mM Tris–ClH pH 7.5, 0.1 M NaCl, 0.5% SDS, 50 μg/ml *Pronase* E, followed by one phenol, one phenol/chlorophorm and one chlorophorm extractions. WMV RNA was concentrated by ethanol precipitation and quantified by UV-absortion spectrophotometry.

Purified RNA was used for the synthesis of double-stranded cDNA using the Time Saver cDNA synthesis kit (Amersham Biosciences, UK). Unfractionated cDNA was ligated to 100 ng of *Eco*RV digested PSK(−) bluescript DNA, and the reaction was used to transform electrocompetent *E. coli* JM109 cells. Recombinant clones were selected and five clones containing inserts ranging 1200–6000 pb were partially sequenced. Identity search in sequence databases using *Fasta 3.4t11* (Pearson and Lipman, 1988) showed that all sequences were homologous to known WMV sequences or to related *Potyvirus* sequences (see Results).

Molecular characterisation of the WMV population

Total nucleic acids were extracted from squash plants infected with each biologically cloned isolate and used as templates for RT-PCR amplification of three WMV genome segments. Oligonucleotides WCPd 5′ GCACTAGT-GAYGGKTGCTGTGAATCAG 3′ (virus sense) and WCPr 5′ GCACTAGTCGACCCGAAATGCTAACTG 3′ (complementary sense) amplified a 930-bp fragment including the complete CP coding sequence (843 nt). Oligonucleotides WCId 5′ GGGTAACGTTGAAATTACCA 3′ (virus sense) and WCIr 5′ CCAAAACCTGCATCACACTT 3′ (complementary sense) amplified a 702-bp fragment representing approximately the C-terminal half of the putative WMV CI cistron. Oligonucleotides WP1d 5′ CATCAAGCTCGAA-CAAGAA 3′ (virus sense) and WP1r 5′ GCTGTTCATACT-TAACATTA 3′ (complementary sense) amplified a 708-bp fragment representing part of the putative WMV P1 cistron. Fragments of the expected sizes were specifically amplified using *Taq* polymerase for all isolates. RT-PCR fragments were purified from 1% agarose gels using Geneclean II (Bio101, California, USA) and cloned in pGEM-T (Promega, Wisconsin, USA) using the manufacturers materials and methods. Plasmid DNA from recombinant clones was purified using the high pure plasmid isolation kit (Roche, Germany) and used as template for automatic sequencing with ABI prism A310 (Applied Biosystems, Foster City, CA). Double sequencing with reverse and direct oligonucleotides was done to avoid sequence ambiguities.

Database homology searches were done using BLAST. Multiple sequence alignments were obtained with ClustalX (Thompson et al., 1997) with the default parameters and these alignments were used to calculate pairwise genetic

distances using Kimura's two-parameter method (Kimura, 1980). Pairwise synonymous and nonsynonymous substitutions were calculated according to the method of Pamilo, Bianchi and Li (Li, 1993; Pamilo and Bianchi,1993) implemented with MEGA2 (Kumar et al., 2001). Phylogenetic analyses were done by the neighbour-joining (NJ) method (Saitou and Nei, 1987) implemented with ClustalX and displayed by TREE VIEW (Page, 1996). Robustness of the inferred evolutionary relationships was assessed by 1000 bootstrap replicates. Putative recombination events were analysed by two methods with different assumptions: that of Jakobsen and Eastel (1996), based on phylogenetic incompatibility, and that of Sawyer (1989), based on the probability that polymorphic sites come from the same parent sequence.

Acknowledgments

References

Aaziz, R., Tepfer, M., 1999. Recombination between genomic RNAs of two cucumoviruses under conditions of minimal selection pressure. Virology 263, 282–289.

Abubakar, Z., Ali, F., Pinel, A., Traoré, O., ŃGuessan, P., Notteghem, J.-L., Konaté, G., Fargette, D., 2003. Phylogeography of *Rice yellow mottle virus* in Africa. J. Gen Virol. 84, 733–743.

Alegria, O.M., Royer, M., Bousalem, M., Chatenet, M., Peterschmitt, M., Girard, J.C., Rott, P., 2003. Genetic diversity in the coat protein coding region of eighty-six sugarcane mosaic virus isolates from eight countries, particularly from Cameroon and Congo. Arch. Virol. 148, 357–372.

Azzam, O., Yambao, M.L.M., Muhsin, M., McNally, K.L., Umadhay, K.L.M., 2000. Genetic diversity of rice tungro spherical virus in tungro-endemic provinces of the Philippines and Indonesia. Arch. Virol. 145, 1183–1197.

Bateson, M.F., Lines, R.E., Revill, P., Chaleeprom, W., Ha, C.V., Gibbs, A.J., Dale, J.L., 2002. On the evolution and molecular epidemiology of the potyvirus Papaya ringspot virus. J. Gen. Virol. 83, 2575–2585.

Bousalem, M., Douzery, E.J.P., Fargette, D., 2000. High genetic diversity, distant phylogenetic relationships and intraspecies recombination events among natural populations of Yam mosaic virus: a contribution to understanding potyvirus evolution. J. Gen. Virol. 81, 243–255.

Bruyere, A., Wantroba, M., Flasinsky, S., Dzianott, A., Bujarski, J.J., 2000. Frequent homologous recombination events between molecules of one RNA component in a multipartite RNA virus. J. Virol. 74, 4214–4219.

Cohen, S., Gertman, E., Kedar, N., 1971. Inheritance of resistance to melon mosaic virus in cucumbers. Phytopathology 61, 253–255.

Desbiez, C., Wipfscheibel, C., Lecoq, H., 2002. Biological and serological variability, evolution and molecular epidemiology of Zucchini *yellow mosaic virus* (ZYMV, *Potyvirus*) with special reference to Caribbean islands. Virus Res. 85, 5–16.

Drake, J.W., Holland, J.J., 1999. Mutation rates among lytic RNA viruses. Proc. Natl. Acad. Sci. U.S.A. 96, 13910–13913.

D'Urso, F., Sambade, A., Moya, A., Guerri, J., Moreno, P., 2003. Variation of haplotype distribution of two genomic regions of *Citrus tristeza virus* populations from Eastern Spain. Mol. Ecol. 12, 517–526.

Fraile, A., Malpica, J.M., Aranda, M.A., Rodríguez-Cerezo, E., García-Arenal, F., 1996. Genetic diversity in tobacco mild green mosaic tobamovirus infecting the wild plant *Nicotiana glauca*. Virology 223, 148–155.

Fraile, A., Alonso-Prados, J.L., Aranda, M.A., Bernal, J.J., Malpica, J.M., García-Arenal, F., 1997. Genetic exchange by recombination or reassortment is infrequent in natural populations of a tripartite RNA plant virus. J. Virol. 71, 934–940.

García-Arenal, F., Escriu, F., Aranda, M.A., Alonso-Prados, J.L., Malpica, J.M., Fraile, A., 2000. Molecular epidemiology of *Cucumber mosaic virus* and its satellite RNA. Virus Res. 71, 1–8.

García-Arenal, F., Fraile, A., Malpica, J.M., 2001. Variability and genetic structure of plant virus populations. Annu. Rev. Phytopathol. 39, 157–186.

Gibbs, M., 1995. The luteovirus supergroup: rampant recombination and persistent partnerships. In: Gibbs, A.J., Calisher, C.H., García-Arenal, F. (Eds.), Molecular Basis of Virus Evolution. Cambridge Univ. Press, Cambridge, pp. 351–368.

Glais, L., Tribodet, M., Kerlan, C., 2002. Genome variability in Potato potyvirus Y (PVY): evidence that PVY(N)W and PVY(NTN) variants are single to multiple recombinants between PVY(O) and PVY(N) isolates. Arch. Virol. 147, 363–378.

Grafton-Cardwell, E.E., Perring, T.M., Smith, R.F., Valencia, J., Farrar, C.A., 1996. Occurrence of mosaic viruses in melons in Central Valley of California. Plant Dis. 80, 1092–1097.

Guyader, S., Ducray, D.G., 2001. Sequence analysis of Potato leafroll virus isolates reveals genetic stability, major evolutionary events and differential selection pressure between overlapping reading frame products. J. Gen. Virol. 83, 1799–1807.

Higgins, C.M., Cassidy, B.G., Teycheney, P.Y., Wongkaew, S., Dietzgen, R.G., 1998. Sequences of the coat protein gene of five peanut stripe virus (PStV) strains from Thailand and their evolutionary relationship with other bean common mosaic virus sequences. Arch. Virol. 143, 1655–1667.

Higgins, C.M., Dietzgen, R.G., Mat Akin, H., Sudarsono, Chen, K., Xu, Z., 1999. Biological and molecular variability of peanut stripe potyvirus. Curr. Top. Virol. 1, 1–26.

Jakobsen, I.B., Eastel, S.A., 1996. A program for calculating and displaying compatibility matrices as an aid in determining reticulate evolution in molecular sequences. CAABIOS 12, 291–295.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Krause-Sakate, R., Le Gall, O., Fakhfakh, H., Peypelut, M., Marrakchi, M., Varveri, C., Pavan, M.A., Souche, S., Lot, H., Zerbini, M., Candresse, T., 2002. Molecular and biological characterization of *Lettuce mosaic virus* (LMV) isolates reveals a distinct and widespread type of resistance-breaking isolate: LMV-Most. Phytopathology 92, 563–572.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular Evolutionary Genetics Analysis software. Bioinformatics 17, 1244–1245.

Kyle, M.M., Provvidenti, R., 1987. Inheritance of resistance to potato Y viruses in *Phaseolus vulgaris* L.: two independent genes for resistance to Watermelon mosaic virus-2. Theor. Appl. Genet. 74, 595–600.

Li, W.-H., 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitutions. J. Mol. Evol. 44, 583–590.

López, C., Ayllón, M.A., Navas-Castillo, J., Guerri, J., Moreno, P., Flores, R., 1998. Molecular variability of the 5ʹ and 3ʹterminal regions of Citrus tristeza virus RNA. Phytopathology 88, 685–691.

Luis-Arteaga, M., Alvarez, J.M., Alonso-Prados, J.L., Bernal, J.J., García-Arenal, F., Laviña, A., Batlle, A., Moriones, E., 1998. Occurrence, distribution, and relative incidence of mosaic viruses infecting field-grown melon in Spain. Plant Dis. 82, 979–982.

MacFarlane, S., 1997. Natural recombination among plant virus genomes: evidence from tobraviruses. Semin. Virol. 8, 25–31.

Malpica, J.M., Fraile, A., Moreno, I., Obies, C.I., Drake, J.W., García-

Arenal, F., 2002. The rate and character of spontaneous mutation in an RNA virus. Genetics 162, 1505–1511.

McNeil, J.E., French, R., Hein, G.L., Baezinger, P.S., Eskridge, K.M., 1996. Characterization of genetic variability among natural populations of wheat streak mosaic virus. Phytopathology 86, 1222–1227.

Moury, B., Morel, C., Johansen, E., Jacquemond, M., 2002. Evidence for diversifying selection in *Potato virus Y* and in the coat protein of other potyviruses. J. Gen. Virol. 83, 2563–2573.

Moyer, J.W., Kennedy, G.G., Romanow, L.R., 1985. Resistance to watermelon mosaic virus-II multiplication in *Cucumis melo*. Phytopathology 75, 201–205.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia Univ. Press, New York.

Ohshima, K., Yamaguchi, Y., Hirota, R., Hamamoto, T., Tomimura, K., Tan, Z., Sano, T., Azuhata, F., Walsh, J.A., Fletcher, J., Chen, J., Gera, A., Gibbs, A., 2002. Molecular evolution of *Turnip mosaic virus*: evidence of host adaptation, genetic recombination and geographical spread. J. Gen. Virol. 83, 1511–1521.

Oruetxebarria, I., Kekarainen, T., Spetz, C., Valkonene, J.P.T., 2000. Molecular characterization of Potato virus V genomes from Europe indicates limited spatiotemporal strain differentiation. Phytopathology 90, 437–444.

Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265, 218–225.

Page, R.D.M., 1996. Treeview-an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 124, 357–358.

Pamilo, P., Bianchi, N.O., 1993. Evolution of the Zfx and Zfy genes: rates and interdependence between the genes. Mol. Biol. Evol. 10, 271–281.

Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. U.S.A. 85, 2444–2448.

Posada, D., Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. U.S.A. 98, 13757–13762.

Purcifull, D., Hiebert, E., Edwardson, J., 1984. Watermelon mosaic virus 2. CMI/AAB Descr. Plant Viruses 293.

Revers, F., Le Gall, O., Candresse, T., Le Romancer, M., Dunez, J., 1996. Frequent occurrence of recombinant potyvirus isolates. J. Gen. Virol. 77, 1953–1965.

Rodríguez-Cerezo, E., Elena, S.F., Moya, A., García-Arenal, F., 1991. High genetic stability in natural populations of the plant RNA virus tobacco mild green mosaic virus. J. Mol. Evol. 32, 328–332.

Sacristán, S., del Monte, J.P., Fraile, A., García-Arenal, F., 2001. Papel de la flora arvense en las infecciones virales de plantas hortícolas en la Vega del Tajo, Comunidad de Madrid. Proc. Congr. SEMh, 47–50.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sambrook, J., Russell, D.W., 2001. Molecular cloning—A laboratory manual, third edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sawyer, S., 1989. Statistical test for detecting gene conversion. Mol. Biol. Evol. 6, 526–538.

Schroender, W.T., Provvidenti, R., 1971. A common gene for resistance to bean yellow mosaic virus and watermelon mosaic virus 2 in *Pisum sativum*. Phytopathology 61, 846–848.

Shukla, D.D., Ward, W.W., Brunt, A.A., 1994. The Potyviridae. CAB International, Wallingford.

Skotnicki, M.L., Mackenzie, A.M., Ding, S.W., Mo, J.Q., Gibbs, A.J., 1993. RNA hybrid mismatch polymorphisms in Australian populations of turnip yellow mosaic tymovirus. Arch. Virol. 132, 83–99.

Tepfer, M., 2002. Risk assessment of virus-resistant transgenic plants. Annu. Rev. Phytopathol. 40, 467–491.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.