High Genetic Stability of the Begomovirus *Tomato yellow leaf curl Sardinia virus* in Southern Spain Over an 8-Year Period

S. Sánchez-Campos, J. A. Díaz, F. Monci,
E. R. Bejarano, J. Reina, J. Navas-Castillo, M. A. Aranda, and E. Moriones

First, second, third, sixth, seventh, and eighth authors: Estación Experimental "La Mayora," Consejo Superior de Investigaciones Científicas, 29750 Algarrobo-Costa, Málaga, Spain; and fourth and fifth authors: Departamento de Genética, Facultad de Ciencias, Universidad de Málaga, Spain.

Current address of S. Sánchez-Campos: Rijk Zwaan Ibérica, La Cañada, Almería, Spain.

Accepted for publication 29 March 2002.

ABSTRACT

Sánchez-Campos, S., Díaz, J. A., Monci, F., Bejarano, E. R., Reina, J., Navas-Castillo, J., Aranda, M. A., and Moriones, E. 2002. High genetic stability of the begomovirus *Tomato yellow leaf curl Sardinia virus* in southern Spain over an 8-year period. *Phytopathology* 92:842-849.

The evolution of the plant single-stranded DNA virus *Tomato yellow leaf curl Sardinia virus* (TYLCSV) (genus *Begomovirus*, family *Geminiviridae*) has been monitored for 8 years after its appearance in southern Spain. Variation within three genomic regions of 166 TYLCSV isolates collected from three locations was assessed by single-strand conformation polymorphism (SSCP) analysis. According to SSCP, the intergenic region (IR) was the most variable. Low genetic diversity was found within the population and geographical or temporal differences were not evident. Nucleotide sequences of specific genomic regions of haplotypes

identified by SSCP indicated close relationships among them. Therefore, the Spanish TYLCSV population appears to represent a single, undifferentiated population. The analysis of IR sequences for a subsample of 76 randomly chosen isolates confirmed the limited genetic diversity revealed by the SSCP analysis. A tendency to a lineal increase in diversity over time was observed in Málaga and Almería subpopulations; however, no accumulation of mutations in single isolates was evident. Negative selection to variation seems to operate to conserve certain regions of the genome. Thus, the low genetic diversity found in the studied TYLCSV population might be the result of a founder effect with subsequent selection against less fit variants arising by mutation.

Additional keywords: *Lycopersicon esculentum*, tomato, viral disease.

The ability of viruses to colonize new ecological niches is well documented and has been frequently associated with the development of important epidemic outbreaks (17,33,36). During the last decades, virologists have devoted considerable effort to understand the evolutionary changes accompanying the expansion of virus populations (14,16,30). As a result, we have an increased knowledge of evolutionary mechanisms used by viruses, which may facilitate epidemiological predictions for the control of viral diseases. Nevertheless, the mechanisms underlying virus evolution are still incompletely understood (6,13,25,35,48). In addition, most of the detailed genetic analyses of virus populations reported so far refer to single-stranded (ss) RNA viruses, whereas studies on populations of DNA viruses are scarce (1,4,9,11,12,14,19,40,47,51,55). Because the intrinsic properties and infection cycle of RNA and DNA viruses might differ, they may be affected differently by forces driving evolution.

The present report considers field populations of *Tomato yellow leaf curl Sardinia virus* (TYLCSV), a plant DNA virus. TYLCSV is a species of the genus *Begomovirus* in the family *Geminiviridae* (8,49), one of only two families of plant viruses with recognized DNA genomes. TYLCSV is one of a number of closely related viruses associated with tomato yellow leaf curl disease (TYLCD), which represents a severe threat to tomato (*Lycopersicon esculentum* Mill.) production in many warm and temperate regions world-

wide (34). In nature, these viruses are transmitted by the whitefly *Bemisia tabaci* Gen. (Hemiptera: Aleyrodidae). The TYLCSV genome consists of one circular ssDNA molecule of approximately 2,800 nucleotides (nts) with two (V2 and coat protein [CP]) open reading frames (ORFs) in the virion sense and four ORFs (replication-associated protein [Rep], transcriptional activator protein [TrAP], replication enhancer protein [REn], and C4) in the complementary sense, separated by an intergenic region (IR) of approximately 300 nts.

Geminiviruses replicate in the nuclei of infected plant cells and rely on host DNA polymerases (18,20). The source of the variation in the geminiviruses is not clear. Some authors suggested that mutation repair mechanisms of the host probably do not function in the geminivirus replication cycle (23). Nevertheless, there are a number of reports that indicate that in the absence of recombination with other strains or species, the level of intra-strain variability in the geminiviruses is limited (7,12,14,55). Moreover, in recent years, it is becoming clear that recombination is an efficient way for geminiviruses to evolve (22,31,45,51–53). However, relatively few studies have been conducted to address temporal and spatial aspects of geminivirus population diversity.

Whitefly-transmitted geminiviruses are responsible for emergent diseases worldwide. One such virus, TYLCSV, appeared in southern Spain and has caused severe damage to tomato crops since 1992 (34,39). We have conducted an extensive survey of the TYLCSV population during eight consecutive epidemic outbreaks of TYLCD within the main tomato growing areas of southern Spain between 1992 and 1999. This provided a unique opportunity to analyze the evolution of a plant DNA virus species in both

Corresponding author: E. Moriones; E-mail address: moriones@eelm.csic.es

Publication no. P-2002-0606-03R

© 2002 The American Phytopathological Society

space and time. In this study, we report the genetic diversity and structure of the TYLCSV population present in southern Spain, and its temporal change over an 8-year period.

MATERIALS AND METHODS

Field surveys and virus isolates. A field survey for TYLCSV-infected samples was conducted in the main tomato growing areas of southern Spain during eight consecutive epidemic outbreaks of TYLCD between 1992 and 1999. Sampling regions were 150 to 400 km apart, and are in the coastal areas of the provinces of Almería, Málaga, and Murcia (Fig. 1). Tomato plants showing yellow leaf curl symptoms were sampled in all three provinces and during each growing season. Approximately three to five fields were visited per province each growing season, and three to five samples were collected per field. In Murcia in 1992 and during 1996 and 1997, 5 to 10 fields were visited per province and 5 to 10 samples were collected per field. Samples were stored frozen at -70°C and lyophilized and, from 1997 onward, fresh prints of petiole cross sections on Whatman 3MM paper were stored at room temperature (37). Each sample was tested for infection by TYLCSV and *Tomato yellow leaf curl virus* (TYLCV), because the latter begomovirus species recently was reported to be associated with TYLCD-affected plants in Spain (50). Samples were tested by specific molecular hybridization as described by Sánchez-Campos et al. (50) on dot blots of total nucleic acids (TNA) extracted as reported by Noris et al. (39), or on tissue prints of fresh petiole cross sections performed on positively charged nylon membranes (Boehringer-Mannheim, Mannheim, Germany). For the genetic

study, isolates from samples positive only for TYLCSV were randomly selected from different regions and years.

Genetic characterization of virus isolates. Variation was studied in three regions of the TYLCSV genome (C2, V2, and IR). The C2 region corresponds to most (85%) of the TrAP ORF, the V2 region covers most (83%) of the V2 ORF plus the 5'-terminal part of the CP ORF, and the IR includes the entire noncoding region. DNA fragments from these regions were amplified by polymerase chain reaction (PCR) using primers designed based on the sequence of a TYLCSV isolate from Murcia (39). Primer MA11 (5'-TCAGGGCTTCTGTACATTC-3') is complementary to nucleotide 496-478, MA13 (5'-AATGCAATCTTCGTCACC-3') is complementary to nucleotide 1632-1615, MA14 (5'-TGCATTATTGAAAACG-3') is identical to nucleotide 2587-2604, MA15 (5'-AAAGGATCCACATATTG-3') is complementary to nucleotide 162-145, MA16 (5'-GCAATAAAATATTTGCAG-3') is identical to nucleotide 209-226, and MA17 (5'-GAAAACATTTGTGATCC-3') is identical to nucleotide 1285-1302. The nucleotide positions refer to the sequence of the Murcia isolate (GenBank Accession No. Z25751). PCR reactions were carried out in 50-µl final volume of PCR reaction mix containing either 0.5 µl of TNA or a piece of Whatman 3MM paper with one print from a fresh petiole cross section (37), 200 µM each dNTP, 100 ng of each primer, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1 mM MgCl₂, and 2 units of AmpliTaq DNA polymerase (Perkin-Elmer, Foster City, CA). Thermal cycling was performed in a GeneAmp PCR System 9700 (Perkin-Elmer) according to the following conditions: 1 min at 94°C followed by 30 cycles of 1 min at 94°C, 1 min at 50°C, and 1 min at 72°C, and a final step of 5 min at 72°C.

Fig. 1. Location of the geographical regions (shadowed) in which the field surveys for *Tomato yellow leaf curl Sardinia virus* isolates were conducted between 1992 and 1999.

TABLE 1. Distribution of the *Tomato yellow leaf curl Sardinia virus* (TYLCSV) population analyzed in this study^a

Province	Number of TYLCSV isolates collected/analyzed								Total
	1992	1993	1994	1995	1996	1997	1998	1999	
Almería	5/5 (4)	13/13 (9)	7/7 (7)	...	43/16 (4)	98/15 (7)	...	...	166/56 (31)
Málaga	...	1/0	...	7/7 (7)	49/17 (4)	117/16 (5)	2/0	6/0	182/40 (16)
Murcia	38/15 (8)	5/5 (3)	3/3 (3)	1/0	20/15 (4)	25/8 (3)	28/14 (4)	27/10 (4)	147/70 (29)
Total	43/20 (12)	19/18 (12)	10/10 (10)	8/7 (7)	112/48 (12)	240/39 (15)	30/14 (4)	33/10 (4)	495/166 (76)

^a The number of samples for which the intergenic region sequence has been obtained is shown in parentheses.

PCR products were submitted to single-strand conformation polymorphism (SSCP) analysis. This analysis was performed by running 1 µl of the PCR product in TBE buffer (89 mM Tris-borate, 2 mM EDTA, pH 8.0) nondenaturing 14% polyacrylamide gel at 200 V and 4°C for 5 h for C2 and V2 fragments, and 12 h for the IR fragment. Data from SSCP analysis were used to estimate population diversity by calculating its heterozygosity as described by Nei (38), adapted for haploid organisms.

Nucleotide sequences were obtained by direct sequencing of PCR DNA products in both directions in an automatic sequencer (ABI PRISM 310 Genetic Analyzer, Perkin-Elmer) using the same primers used for DNA amplification. Sequences obtained for C2, V2, and IR were aligned with available sequences of other begomovirus species associated with TYLCD (34) using Clustal W (56). Nucleotide sequences of the IR were aligned using the PileUp program of GCG (Genetics Computer Group, Madison, WI) and analyzed with the PHYLIP 3.5 software package of J. Felsenstein (University of Washington, Seattle). This software also was used to estimate genetic distances by the Kimura two-parameter method (27). Genetic distances were used to estimate population nucleotide diversities as described by Nei (38). The increase of genetic diversity with time in each location was tested by lineal regression analysis (SPSS, Chicago) of intrapopulation nucleotide diversity values over time.

Nucleotide sequence accession numbers. The nucleotide sequences obtained for this study were deposited as GenBank

Accession Nos. AF376002 to AF376007 for the C2 region, AF376009 to AF376015 for the V2 region, and AF373616 to AF373691 for the IR.

RESULTS

Genetic diversity of TYLCSV isolates. As a result of surveys conducted during 1992 to 1999 in the main tomato growing areas of southern Spain, a total of 495 TYLCSV isolates were obtained (Table 1). From these, 166 isolates were randomly selected for the genetic analysis; in those regions and years in which less than 15 TYLCSV isolates were collected, all of them were analyzed (Table 1). The diversity of the TYLCSV isolates was analyzed for three genomic regions (C2, V2, and IR). The genomic locations of these regions are shown in Figure 2. IR is highly variable among geminiviruses (43), whereas C2 and V2 are more conserved. DNA fragments from these genomic regions were amplified by PCR from each isolate using primers MA17 and MA13, which direct the amplification of a 349-bp fragment (C2 region); MA16 and MA11, which direct the amplification of a 288-bp fragment (V2 region); and MA14 and MA15, which direct the amplification of a 353-bp fragment (IR). Thus, a total of ≈36% of the genome was analyzed.

SSCP is a powerful technique for scanning mutations in small DNA fragments, and has been used successfully to characterize variation of nuclear genes and virus populations (2,21,28,41). This technique is a simple method for reliably detecting single-base polymorphisms (24,46). As discussed elsewhere (19), in SSCP, PCR products rather than cloned fragments are tested; therefore, the study is not affected by nucleotide variation due to misincorporation by *Taq* DNA polymerase. To evaluate the variation of the 166 TYLCSV isolates included in the analysis, SSCP patterns for each of the three genomic regions were assessed. A reproducible pattern for each of the regions was assigned to every virus isolate, and isolates with identical SSCP patterns were grouped into classes, each of which was considered to correspond to a single genetic type. Some examples of SSCP patterns are shown in Figure 3. The number of genetic types for C2 and V2 was only six and seven, respectively, whereas the number of types for the IR was 30. The ratio between the resulting number of different genetic types deduced by SSCP analysis with respect to number of TYLCSV isolates analyzed was more than four times smaller for C2 and V2 regions (0.036 and 0.042, respectively) compared with IR (0.181). For C2 and V2, one predominant type (type A) represented to more than 90% of the isolates analyzed, whereas for the IR, there were several frequent types, with the most abundant representing only about 20% of the isolates (Tables 2 and 3). Accordingly, the estimated probability that two individuals taken at random from the population did not share the same genetic type (heterozygosity) (38) was substantially smaller for C2 and V2 than for the IR (Table 2).

The nucleotide sequences of the considered genomic regions from at least one isolate corresponding to each SSCP type were obtained. As before, to reduce artifactual nucleotide variation

Fig. 2. Schematic representation of the genome organization of *Tomato yellow leaf curl Sardinia virus*. Arrows show the open reading frames, the black box in the circle shows the intergenic region containing the stem-loop structure, and boxes within the circle represent the genomic regions examined in this study.

Fig. 3. Electrophoretic mobility patterns obtained for the polymerase chain reaction-amplified DNA fragments from tomato leaf samples infected with *Tomato yellow leaf curl Sardinia virus* after single-strand conformation polymorphism analysis. A selection of the patterns obtained for the DNA fragments amplified from regions V2, C2, and IR is shown.

results due to misincorporation by *Taq* DNA polymerase, direct sequencing of PCR DNA products was performed. These sequences showed little genetic divergence among SSCP types, being 98 to 99% identical for C2 and V2, and 94 to 98% identical for the IR. Identities of these sequences with those of the Murcia and Almería TYLCSV isolates collected in Spain during the first TYLCD epidemic outbreak in 1992 (GenBank Accession Nos. Z25751 and L27708, respectively) were 98 to 99%, 98 to 100%, and 94 to 99% for C2, V2, and the IR, respectively. However, identity with the corresponding genomic regions of TYLCSV isolates reported from other locations, e.g., Sardinia (GenBank Accession No. X61153) and Sicily (GenBank Accession No. Z28390), revealed a more distant relationship ($\leq 92\%$ for C2, 96% for V2, and $\leq 83\%$ for IR). Therefore, all the isolates analyzed in the present study seem to belong to a single TYLCSV strain, and are more distantly related to TYLCSV strains reported from Sardinia or Sicily.

Genetic structure of TYLCSV populations. The combination of the SSCP types for each of the C2, V2, and the IR defined the haplotype of the isolate (e.g., an isolate with an SSCP pattern of type A in C2, B in V2, and P in IR corresponded to the haplotype ABP). The 166 TYLCSV isolates analyzed were classified into 48 haplotypes (Table 2). The occurrence of the different haplotypes in each geographical region and year is shown in Table 3. The genetic structure of the TYLCSV population was first analyzed by comparing the distribution of the estimated frequencies of different haplotypes among the three geographical regions. The most frequent haplotype differed in totals from each of the regions (Table 3). Thus, in Almería, AAB accounted for 28.6% of the isolates, in Málaga, AAC accounted for 40.0%, and in Murcia, AAA accounted for 22.9%. In accordance, a *G* test of independence (54) found significant differences ($P < 0.005$) when comparing the distributions of haplotype frequencies for the totals corresponding to each geographical region. Therefore, some degree of geographical differentiation could exist. The real significance of this presumed genetic differentiation was estimated between and within geographical subpopulations by means of the Nei's G_{st} test (38). For this analysis, a subpopulation was considered to be constituted by all the isolates coming from a single geographical region, and the three genomic regions under study were included in the analysis as three independent loci (38). The value of the G_{st} estimator for the three subpopulations was 0.0797. Therefore, although the *G* test of independence identified some degree of geographical differentiation, the existence of geographical structure is not sustained by the G_{st} test because only about 8% of the total diversity was distributed among subpopulations. G_{st} values calculated for each pair of subpopulations showed that those of Almería and Murcia were more similar to each other ($G_{st} = 0.0425$) than to the Málaga subpopulation (Almería versus Málaga, $G_{st} = 0.0658$; Murcia versus Málaga, $G_{st} = 0.0750$).

The structure of the TYLCSV population also was analyzed by comparing the distribution of the estimated frequencies of haplotypes between yearly subpopulations. The data in Table 3 show that the frequencies of haplotypes within each geographical region changed with time and that haplotypes different from those initially prevalent can emerge and become predominant. For example, the haplotype AAP in Málaga was not detected during 1995 and 1996, but represented 43.7% of the isolates analyzed in 1997. Similarly, in Murcia, none of the haplotypes detected from 1992 to 1996 was represented in 1998 and 1999. These data may indicate a tendency toward haplotype replacement with time. However, such a tendency is not sustained statistically, because a *G* test of independence (54) did not find significant differences in the distribution of haplotype frequencies for the yearly subpopulations within each geographical region.

Increase in genetic diversity of TYLCSV populations over time. The data in Table 3 also suggest that there is a tendency toward an increase in genetic diversity over time. For example, for

Almería in 1992, the ratio of haplotypes to isolates was 0.20, whereas in 1997 it was 0.53. Correspondingly, the estimated average genetic diversity (average heterozygosity) (38) for Almería in 1992 was 0.00, whereas in 1997 it was 0.37. Nucleotide diversity cannot be estimated from SSCP data. Therefore, in order to get a deeper insight into this aspect, a subsample of the whole population (76 isolates) was taken at random, and sequences were obtained for the IR, the most variable region according to the SSCP results. Sequence information was used to estimate the genetic distances between any two isolates by the Kimura two-parameter method (27) and from these, nucleotide diversities were calculated according to Nei (38). The average nucleotide diversity was estimated to be 0.0087 for the entire population, which shows the limited genetic divergence among the TYLCSV isolates that constitute the population studied. However, as was evident from SSCP data in Table 3, a tendency toward a lineal increase in intrapopulation nucleotide diversity over time could be perceived for the Almería ($R^2 = 0.9804$) and Málaga subpopulations ($R^2 = 0.9973$) (Table 4). No such tendency was detected in the Murcia subpopulation, although the diversity for the final date was substantially higher than for the initial date but values fluctuated randomly ($R^2 = 0.4789$) (Table 4). When values of genetic distances for every pair of isolates were plotted against the time of collection, no significant increase of genetic distance with time was observed either for the subpopulations or for the entire population (data not shown). Therefore, no accumulation of mutations in single isolates seemed to occur. Hence, the increase in the intrapopulation genetic diversity over time observed in certain locations seemed to result from an increased number of genetic types.

Distribution of mutations in C2, V2, and IR of TYLCSV isolates. The C2 and V2 regions encompass parts of five overlapping ORFs of TYLCSV (Fig. 2). The number and nature of nucleotide substitutions in these regions were analyzed by aligning the nucleotide sequences obtained for isolates corresponding to each SSCP type with a putative ancestral sequence deduced by maximum parsimony (15). In the V2 ORF, seven nucleotide substitutions were found, six of which were silent and one resulted in a conservative amino acid change in the V2 protein. In the CP ORF within the V2 region, which encodes the most variable part of the CP (43), three nucleotide substitutions were found; one resulted in a conservative change and two in nonconservative amino acid changes in the CP sequence. None of these changes affected the conserved nuclear localization motif present in this part of the CP (29). In the REN ORF within the C2 region, five nucleotide substitutions were found, two silent, two resulted in conservative changes, and one in a nonconservative amino acid change in the REN protein. In the TrAP ORF, seven nucleotide substitutions were found, one silent change, two resulted in conservative changes, and four in nonconservative amino acid changes in the TrAP protein. One substitution was found in Rep ORF fragment within the C2 region and it was silent in the Rep protein. These results suggest that relatively few nucleotide substitutions occurred

TABLE 2. Number of genetic types per isolate, frequency of the more abundant genetic type, and heterozygosity estimated from the data of the single-strand conformation polymorphism analysis performed on the C2, V2, and intergenic region (IR) of the *Tomato yellow leaf curl Sardinia virus* genome

Region	Genetic types/isolates	Frequency of the more abundant genetic type	Heterozygosity ^a
C2	6/166	158/166 (95.2%)	0.0946 $\pm$ 0.0157
V2	7/166	153/166 (92.2%)	0.1502 $\pm$ 0.0203
IR	30/166	34/166 (20.5%)	0.9526 $\pm$ 0.0177
Total	48 ^b /166	26/166 (15.7%)	0.9290 $\pm$ 0.0148

^a As described by Nei (38).

^b Number of different genetic types based on the combination of single-strand conformation polymorphism pattern in C2, V2, and IR (Table 3).

in the C2 and V2 regions of the analyzed population, and that a limited number of these resulted in amino acid replacements.

Several functional domains have been defined in the IR of geminiviruses, some of them associated with predicted structural features. The conserved stem-loop structure containing the cleavage site for initiation of positive-strand DNA synthesis, the TATA box elements for transcription promoters of the mRNAs for virion-sense and complementary-sense genes, the Rep binding site for the initiation of replication and down-regulation of transcription, and the sequence element for TrAP-mediated transactivation of virion-sense gene expression are located within this region (18,20). The position of the 78 substitutions identified in the IR sequences (between 1 and 14 substitutions per sequence) obtained in this study were compared with an ancestral sequence deduced by maximum parsimony (15). This analysis showed that (i) none of the substitutions altered either the TATA box or Rep binding

motifs, (ii) three substitutions (only one per sequence) were found in the stem-loop region and these were in the 5' part of the stem, a region more permissive to mutations, as demonstrated for *Tomato golden mosaic virus* (42), and (iii) two substitutions were found in the TrAP recognition motif, one of which is also present in related viruses associated with TYLCD (data not shown). Therefore, nucleotide substitutions in the IR tend to accumulate at positions where no structural and/or functional motifs are found.

DISCUSSION

TYLCSV appeared in southern Spain and has caused severe damage to tomato crops since 1992 (39). Similar TYLCSV variants were presumed to be involved in the epidemics because almost identical nucleotide sequences (99.5% identity) were obtained for complete genomes of two TYLCSV isolates collected in

TABLE 3. Genetic structure of the population of *Tomato yellow leaf curl Sardinia virus* as revealed by single-strand conformation polymorphism (SSCP) analysis

Haplotype ^a			Number of isolates per year - frequency (f)																				Total		
			Almería							Málaga					Murcia										
C2	V2	IR	92	93	94	96	97	Total	f	95	96	97	Total	f	92	93	94	96	97	98	99	Total	f	Total	f
A	A	A	...	1	...	...	...	1	0.018	...	...	...	...	...	11	2	1	2	...	...	...	16	0.229	17	0.102
A	A	B	5	7	2	2	...	16	0.286	...	...	...	...	...	1	1	1	1	...	...	...	4	0.057	20	0.120
A	A	C	...	...	...	...	...	...	...	5	9	2	16	0.400	...	...	...	4	1	...	...	5	0.071	21	0.126
A	A	D	...	...	4	1	2	7	0.125	...	3	3	6	0.015	...	...	...	...	5	7	1	13	0.186	26	0.157
A	A	E	...	...	...	6	2	8	0.143	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	9	0.054
A	A	F	...	...	...	...	1	1	0.018	...	...	...	...	...	...	...	...	3	...	...	...	3	0.042	4	0.024
A	A	G	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	...	...	1	0.014	1	0.006
A	A	H	...	...	...	2	...	2	0.036	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	3	0.018
A	A	I	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	...	1	0.014	1	0.006
A	A	J	...	...	...	1	...	1	0.018	...	...	...	...	...	...	...	...	1	...	...	...	1	0.014	2	0.012
A	A	K	...	1	...	...	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
A	A	L	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	2	0.028	2	0.012
A	A	M	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	...	...	...	1	0.014	1	0.006
A	A	N	...	...	1	...	...	1	0.018	1	...	...	1	0.025	...	...	...	...	...	...	...	...	...	2	0.012
A	A	Ñ	...	...	...	...	...	...	...	...	1	...	...	0.025	...	...	...	...	...	...	...	...	...	1	0.006
A	A	O	...	...	...	1	...	1	0.018	...	...	1	1	0.025	...	...	...	...	1	3	...	4	0.057	6	0.036
A	A	P	...	...	...	...	...	...	...	...	...	7	7	0.175	...	...	...	...	...	...	...	...	...	7	0.042
A	A	Q	...	1	...	...	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
A	A	R	...	1	...	...	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
A	A	S	...	1	...	...	4	5	0.089	...	...	...	...	...	...	...	...	...	...	...	...	...	...	5	0.030
A	A	T	...	1	...	...	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
A	A	V	...	...	...	1	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
A	A	W	...	...	...	...	1	1	0.018	...	...	...	...	...	...	...	...	...	2	...	...	2	0.028	3	0.018
A	A	X	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	1	0.014	1	0.006
A	A	Z	...	...	...	...	3	3	0.054	...	...	...	...	...	...	...	...	...	...	1	2	3	0.042	6	0.036
A	A	α	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	1	0.006	
A	A	β	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	1	0.006	
A	A	γ	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	1	0.006	
A	B	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	1	0.014	1	0.006
A	B	D	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	1	0.014	1	0.006
A	C	A	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	1	0.014	1	0.006
A	C	B	...	...	...	1	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
A	D	B	...	...	...	...	...	...	...	...	1	...	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
A	D	D	...	...	...	...	...	...	...	...	...	1	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
A	D	F	...	...	...	...	...	...	...	...	...	1	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
A	D	H	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	1	0.006	
A	E	C	...	...	...	...	...	...	...	...	1	...	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
A	F	D	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	...	1	0.014	1	0.006
A	F	Y	...	...	...	...	...	...	...	...	...	1	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
A	G	H	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	1	0.014	1	0.006	
B	A	Z	...	...	...	...	1	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
B	E	U	...	...	...	1	...	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
C	A	D	...	...	...	...	...	...	...	...	1	...	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
D	A	C	...	...	...	...	...	...	...	...	1	...	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
E	A	C	...	...	...	...	...	...	...	...	1	...	1	0.025	...	...	...	...	...	...	...	...	...	1	0.006
E	A	F	...	...	...	...	1	1	0.018	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	0.006
E	A	P	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	1	0.014	1	0.006
F	A	A	...	...	...	...	...	...	...	...	...	...	...	...	...	1	...	...	...	...	...	1	0.014	1	0.006
Total			5	13	7	16	15	56	1	7	17	16	40	1	15	5	3	15	8	14	10	70	1	166	1

^a Defined by the combination of the SSCP patterns obtained for the polymerase chain reaction DNA fragments amplified from the regions C2, V2, and IR.

1992 from the Murcia and Almería provinces, which are separated by more than 200 km (GenBank Accession Nos. Z25751 and L27708, respectively). In this study, the examination of a large number of TYLCSV isolates confirmed that closely related genetic variants of a single TYLCSV strain were present in southern Spain. These TYLCSV isolates exhibited a more distant relationship to other TYLCSV strains reported so far (5,26). The SSCP analysis of three genomic regions (C2, V2, and IR) of the TYLCSV isolates showed that the population analyzed is quite uniform genetically. It was composed of few common genetic types (haplotypes AAA, AAB, AAC, and AAD accounted for 50% of the isolates) plus several infrequent ones. This gamma distribution of the frequency of genetic types (13) was also observed in the populations of the geminivirus *Cotton leaf curl virus* (CLCuV) in Pakistan (51), whereas in other cases, as for the population of the geminivirus *Beet curly top virus* in western United States, such a distribution was not evident (55).

The SSCP analysis may have resulted in an underestimation of the actual variability of the population due to limitations of the technique, because some isolates with indistinguishable SSCP patterns might differ in their nucleotide sequences. This does not seem to have been the case for C2 and V2 regions because identical nucleotide sequences were obtained for PCR products from 15 isolates corresponding to the most abundant SSCP type for each region (data not shown). In contrast, comparison of SSCP results and sequences of the IR PCR products obtained from the 76 randomly chosen isolates revealed that the SSCP analysis was able to detect only 60% of the actual sequence variation present in the population (data not shown). However, the analysis of the IR sequences indicated that a high genetic homogeneity was found in the population thus confirming the results obtained from the SSCP analysis. In fact, the average nucleotide diversity (38) in the IR for the entire population was 0.0087. This value is 10 times smaller than that obtained for a plant pathogenic noncoding RNA (the satellite RNA of *Cucumber mosaic virus*) (1), and four times smaller than the value obtained for the plant RNA virus, *Tobacco mild green mosaic virus*, which is relatively genetically stable (11, 47). Similar limited genetic divergence was also detected in other geminiviruses for which field isolates of a single virus strain have been examined (7,12,14,55). Therefore, this work and those studies suggest that the levels of intra-strain variability in field populations of geminiviruses seem to be limited. This is consistent with data available for plant virus populations that indicate that genetic stability rather than diversity seems to be the rule irrespective of whether their genome is RNA or DNA (13).

Although it is not known whether TYLCSV was introduced in southern Spain or it emerged from indigenous reservoirs, the low genetic diversity found in the population could be the result of a founder effect associated with a population bottleneck during the transmission to a new host or area (11,40). However, our analyses also suggest that negative selection against variation is operating in the population. Heterozygosity indices calculated from SSCP data showed that selection in protein-encoding genes is limiting genetic divergence as observed in other cases (13,51). Moreover, in the noncoding IR, mutations accumulate in presumed nonfunctional areas, therefore suggesting that selection is also operating for maintenance of structures that are important for the viability of

the virus. Thus, constraints to genetic divergence are operating in the TYLCSV population. These restrictions to nucleotide variation are common in plant viruses (13) and also were detected in the IR and in nonsynonymous positions of coding regions of the begomovirus CLCuV (51).

No marked geographical structure was detected in the TYLCSV population, suggesting it is a unique, undifferentiated population. Only about 8% of the total diversity was distributed among the geographical regions sampled based upon Nei's G_{st} estimator. This result, from SSCP typing, agreed with that derived from the analysis of the IR sequences, which showed that the interpopulation nucleotide diversity was always smaller than the intrapopulation nucleotide diversity (data not shown). This situation is frequent in plant virus populations (13,47) and has also been observed for the geminivirus CLCuV in Pakistan (51). The Nei's test also revealed that the TYLCSV subpopulations from Almería and Murcia were more closely related to each other than to the Málaga subpopulation, which is consistent with a more frequent movement of tomato transplants between these regions.

No significant temporal differences were detected in the TYLCSV populations present in Málaga, Almería, or Murcia. However, a tendency for genetic diversity to increase over time seemed to occur in Almería and Málaga, whereas a fluctuation was observed in Murcia between years. Fluctuation in Murcia might be explained by the occurrence of more severe population bottlenecks between epidemics because of less intensive overlapping of TYLCSV-susceptible crops and a more pronounced disruption of the cycle during winter for both host plants and the whitefly vector. Then, in years with severe bottlenecks, local extinctions and recolonization from near or distant virus reservoirs may result in a decrease in genetic diversity as documented for other viruses (9,40). The differences detected among the absolute values of the intrapopulation genetic diversity deduced for the TYLCSV populations present in the geographical regions studied (Table 4) may reflect differences in the onset of the epidemics. It is interesting to note that our data suggest that no accumulation of mutations over time is occurring in single isolates, and that no evident replacement of genetic types occurs with time. Therefore, the tendency to increase in genetic diversity observed in Málaga and Almería might be better explained by an increased number of closely related genetic variants cocirculating in the populations present there with time.

This work advances in the understanding of the epidemics caused by TYLCSV and provides a picture of the evolution of this virus after its appearance in southern Spain. The major finding is that there was not much genetic diversity detected over the time period analyzed. However, it should be noted that ecological circumstances have dramatically changed for TYLCSV epidemics in this geographical region in the last few years. In 1997, we reported the spread of a new begomovirus species, TYLCV, associated with the TYLCD and that this resulted in the progressive displacement of TYLCSV in the epidemics (50). This may have direct consequences for the evolution of the TYLCSV population because the decrease in the population size might result in a TYLCSV fitness loss by accumulation of mutations according to Muller's ratchet mechanism (3) as observed in other cases (10). Moreover, we have provided evidence that recombination between TYLCSV

TABLE 4. Nucleotide diversity within *Tomato yellow leaf curl Sardinia virus* populations inferred from intergenic region (IR) sequences for each location and their R^2 values of lineal regression over time^a

Province	1992	1993	1994	1995	1996	1997	1998	R^2
Almería	0.0000 (4)	0.0028 (9)	0.0052 (7)	...	0.0080 (4)	0.0098 (7)	...	0.9804 ^b
Málaga	...	...	...	0.0018 (7)	0.0058 (4)	0.0106 (5)	...	0.9973
Murcia	0.0066 (8)	0.0043 (3)	0.0085 (3)	...	0.0101 (4)	0.0085 (3)	0.0096 (4)	0.4789

^a Nucleotide diversities were estimated as proposed by Nei (38). Parentheses indicates the number of isolates randomly selected from the whole population, for which the sequence of the IR was obtained.

^b In each location, R^2 resulted from lineal regression analyses (SSCP, Chicago) of nucleotide diversity values over time.

and TYLCV genomes has occurred in southern Spain (32). This is especially important because it is becoming clear from the literature that major changes in geminivirus genomes occur via recombination (22,31,44,45,51–53). Therefore, a continued analysis of the epidemics caused by these viruses in southern Spain is needed to better understand their evolution and the dynamics of their populations.

ACKNOWLEDGMENTS

This work was financed by grants AGF98-0439-C05 (Comisión Interministerial de Ciencia y Tecnología, CICYT, Spain), 1FD97-0271, and 1FD97-0651 (CICYT, European Commission). S. Sánchez-Campos and J. Navas-Castillo were recipients of a fellowship and a contract, respectively, from Ministerio de Educación y Cultura (MEC), Spain. We thank F. García-Arenal (ETSIA, Madrid, Spain) for providing part of the samples from 1992, J. Stanley for critical reading of the manuscript, and M. Victoria Martín for technical assistance.

LITERATURE CITED

- Aranda, M. A., Fraile, A., and García-Arenal, F. 1993. Genetic variability and evolution of the satellite RNA of cucumber mosaic virus during natural epidemics. *J. Virol.* 67:5896-5901.
- Ayllón, M. A., Rubio, L., Moya, A., Guerri, J., and Moreno, P. 1999. The haplotype distribution of two genes of citrus tristeza virus is altered after host change or aphid transmission. *Virology* 255:32-39.
- Chao, L. 1990. Fitness of RNA virus decreased by Muller's ratchet. *Nature* 348:454-455.
- Clarke, D. K., Duarte, E. A., Moya, A., Elena, S. F., Domingo, E., and Holland, J. J. 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. *J. Virol.* 67:222-228.
- Crespi, S., Noris, E., Vaira, A. M., and Accotto, G. P. 1995. Molecular characterization of cloned DNA from a tomato yellow leaf curl virus isolate from Sicily. *Phytopathol. Mediterr.* 34:93-99.
- Domingo, E., and Holland, J. J. 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* 51:151-178.
- Faria, J. C., and Maxwell, D. P. 1999. Variability in geminivirus isolates associated with *Phaseolus* spp. in Brazil. *Phytopathology* 89:262-268.
- Fauquet, C. M., Maxwell, D. P., Gronenborn, B., and Stanley, J. 2000. Revised proposal for naming geminiviruses. *Arch. Virol.* 145:1743-1761.
- Fraile, A., Alonso-Prados, J. L., Aranda, M. A., Bernal, J. J., Malpica, J. M., and García-Arenal, F. 1997. Genetic exchange by recombination or reassortment is frequent in natural populations of a tripartite RNA plant virus. *J. Virol.* 71:934-940.
- Fraile, A., Escriu, F., Aranda, M. A., Malpica, J. M., Gibbs, A. J., and García-Arenal, F. 1997. A century of tobamovirus evolution in an Australian population of *Nicotiana glauca*. *J. Virol.* 71:8316-8320.
- Fraile, A., Malpica, J. M., Aranda, M. A., Rodríguez-Cerezo, E., and García-Arenal, F. 1996. Genetic diversity in tobacco mild green mosaic tobamovirus infecting *Nicotiana glauca*. *Virology* 223:148-155.
- Garrido-Ramírez, E. R., Sudarshana, M. R., and Gilbertson, R. L. 2000. *Bean golden yellow mosaic virus* from Chiapas, Mexico: Characterization, pseudorecombination with other bean-infecting geminiviruses and germ plasm screening. *Phytopathology* 90:1224-1232.
- Gibbs, A. J., Keese, P. L., Gibbs, M. J., and García-Arenal, F. 1999. Plant virus evolution: Past, present and future. Pages 263-285 in: *Origin and Evolution of Viruses*. E. Domingo, R. Webster, and J. Holland, eds. Academic Press, New York.
- Gilbertson, R. L., Rojas, M. R., Russell, D. R., and Maxwell, D. P. 1991. Use of asymmetric polymerase chain reaction and DNA sequencing to determine genetic variability of bean golden mosaic geminivirus in the Dominican Republic. *J. Gen. Virol.* 72:2843-2848.
- Gojobori, T., Li, W.-H., and Graur, D. 1982. Patterns of nucleotide substitutions in pseudogenes and functional genes. *J. Mol. Evol.* 18:360-369.
- Gojobori, T., Moriyama, E. N., Ina, Y., Ikeo, K., Miura, T., Tsujimoto, H., Hayami, M., and Yokoyama, S. 1990. Evolutionary origin of human and simian immunodeficiency viruses. *Proc. Natl. Acad. Sci. USA* 87:4108-4111.
- Goldbach, R., and Peters, D. 1994. Possible causes of the emergence of *Tospovirus* diseases. *Sem. Virol.* 5:113-120.
- Gutiérrez, C. 1999. Geminivirus DNA replication. *Cell. Mol. Life Sci.* 56:313-329.
- Hall, J. S., French, R., Morris, T. J., and Stenger, D. C. 2001. Structure and temporal dynamics of populations within Wheat streak mosaic virus isolates. *J. Virol.* 75:10231-10243.
- Hanley-Bowdoin, L., Settlege, S. B., Orozco, B. M., Nagar, S., and Robertson, D. 1999. Geminiviruses: Models for plant DNA replication, transcription, and cell cycle regulation. *Crit. Rev. Plant Sci.* 18:71-106.
- Hirsch, V. M., Dapolito, G., Hahn, A., Lifson, J., Montefiori, D., Brown, C. R., and Goeken, R. 1998. Viral genetic evolution in macaques infected with molecularly cloned simian immunodeficiency virus correlates with the extent of persistent viremia. *J. Virol.* 72:6482-6489.
- Hou, Y. M., and Gilbertson, R. L. 1996. Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. *J. Virol.* 70:5430-5436.
- Inamdar, N. M., Zhang, X.-Y., Brough, C. L., Gardiner, W. E., Bisaro, D. M., Ehrlich, M. 1992. Transfection of heteroduplexes containing uracil-guanine or thymine-guanine mispairs into cell plants. *Plant Mol. Biol.* 20:123-131.
- Jäckel, S., Epplen, J. T., and Epplen, C. 1999. PCR-SSCP: A method for reliably detecting single nucleotide polymorphisms in multifactorial diseases. Pages 167-176 in: *DNA Profiling and DNA Fingerprinting*. J. T. Epplen and T. Lubjuhn, eds. Birkhäuser Verlag, Basel, Switzerland.
- Keese, P., and Gibbs, A. K. 1993. Plant viruses: Master explorers of evolutionary space. *Curr. Opin. Genet. Dev.* 3:873-877.
- Kheyr-Pour, A., Bendahmane, M., Matzeit, N., Accotto, G. P., Crespi, S., and Gronenborn, B. 1991. Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted geminivirus. *Nucleic Acids Res.* 19:6763-6769.
- 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.
- Koenig, R., Lüdekke, P., and Haeberle, A. M. 1995. Detection of beet necrotic yellow vein virus strains, variants and mixed infections by examining single-strand conformation polymorphism of immunocapture RT-PCR products. *J. Gen. Virol.* 76:2051-2055.
- Kunik, T., Palanichelvam, K., Czosnek, H., Citovsky, V., and Gafni, Y. 1998. Nuclear import of the capsid protein of tomato yellow leaf curl virus (TYLCV) in plant and insect cells. *Plant J.* 13:393-399.
- Levin, S. A., and Pimentel, D. 1981. Selection for intermediate rates of increase in parasite-host system. *Am. Nat.* 117:308-315.
- Martin, D. P., Willment, J. A., Billharz, R., Velders, R., Odhiambo, B., Njuguna, J., James, D., and Rybicki, E. P. 2001. Sequence diversity and virulence in *Zea mays* of Maize streak virus isolates. *Virology* 288:247-255.
- Monci, F., Navas-Castillo, J., and Moriones, E. 2001. Evidence of a naturally occurring recombinant between *Tomato yellow leaf curl virus* and *Tomato yellow leaf curl Sardinia virus* in Spain. *Plant Dis.* 85:1289.
- Morales, F. J., and Anderson, P. K. 2001. The emergence and dissemination of whitefly-transmitted geminiviruses in Latin America—Brief review. *Arch. Virol.* 146:415-441.
- Moriones, E., and Navas-Castillo, J. 2000. Tomato yellow leaf curl virus, an emerging virus complex causing epidemics worldwide. *Virus Res.* 71:123-134.
- Moya, A., Elena, S. F., Bracho, A., Miralles, R., and Barrio, E. 2000. The evolution of RNA viruses: A population genetics view. *Proc. Natl. Acad. Sci. USA* 97:6967-6973.
- Myers, G., MacInnes, K. A., and Korber, B. 1992. The emergence of simian/human immunodeficiency viruses. *AIDS Res. Hum. Retrovir.* 8:373-386.
- Navas-Castillo, J., Díaz, J. A., Sánchez-Campos, S., and Moriones, E. 1998. Improvement of the print-capture polymerase chain reaction procedure for efficient amplification of DNA virus genomes from plants and insects vectors. *J. Virol. Methods* 75:195-198.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Noris, E., Hidalgo, E., Accotto, G. P., and Moriones, E. 1994. High similarity among the tomato yellow leaf curl virus isolates from the West Mediterranean Basin: The nucleotide sequence of an infectious clone from Spain. *Arch. Virol.* 135:165-170.
- Novella, I. S., Elena, S. F., Moya, A., Domingo, E., and Holland, J. J. 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. *J. Virol.* 69:2869-2872.
- Orita, M., Iwanhama, H., Hayashi, K., and Sekiya, T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA* 86:2766-2770.
- Orozco, B. M., and Hanley-Bowdoin, L. 1996. A DNA structure is required for geminivirus replication origin function. *J. Virol.* 70:148-158.
- Padidam, M., Beachy, R. N., and Fauquet, C. M. 1995. Classification and identification of geminiviruses using sequence comparisons. *J. Gen. Virol.* 76:249-263.
- Padidam, M., Beachy, R. N., and Fauquet, C. M. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* 265:218-225.
- Pita, J. S., Fondong, V. N., Sangare, A., Otim-Nape, G. W., Ogwal, S., and Fauquet, C. M. 2001. Recombination, pseudorecombination and

- synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *J. Gen. Virol.* 82:655-665.
46. Prosser, J. 1993. Detecting single-base mutations. *Trends Biotechnol.* 11:238-246.
 47. Rodríguez-Cerezo, E., Elena, S. F., Moya, A., and 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.
 48. Roossinck, M. J. 1997. Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol.* 35:191-209.
 49. Rybicki, E. P., Briddon, R. W., Brown, J. K., Fauquet, C. M., Maxwell, D. P., Harrison, B. D., Markham, P. G., Bisaro, D. M., Robinson, D., and Stanley, J. 2000. Family *Geminiviridae*. Pages 285-297 in: *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*. M. H. V. Van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner, eds. Academic Press, New York.
 50. Sánchez-Campos, S., Navas-Castillo, J., Camero, R., Soria, C., Díaz, J. A., and Moriones, E. 1999. Displacement of tomato yellow leaf curl virus (TYLCV)-Sr by TYLCV-Is in tomato epidemics in Spain. *Phytopathology* 89:1038-1043.
 51. Sanz, A. I., Fraile, A., Gallego, J. M., Malpica, J. M., and García-Arenal, F. 1999. Genetic variability of natural populations of cotton leaf curl geminivirus, a single-stranded DNA virus. *J. Mol. Evol.* 46:672-681.
 52. Sanz, A. I., Fraile, A., Garcia-Arenal, F., Zhou, X. P., Robinson, D. J., Khalid, S., Butt, T., and Harrison, B. D. 2000. Multiple infection, recombination and genome relationships among begomovirus isolates found in cotton and other plants in Pakistan. *J. Gen. Virol.* 81:1839-1849.
 53. Saunders, K., Salim, N., Mali, V. R., Malathi, V. G., Briddon, R., Markham, P. G., and Stanley, J. 2002. Characterisation of Sri Lankan cassava mosaic virus and Indian cassava mosaic virus: Evidence for acquisition of a DNA B component by a monopartite begomovirus. *Virology* 293:63-74.
 54. Sokal, R. R., and Rohlf, F. J. 1995. *Biometry*. 3rd ed. Freeman & Co., New York.
 55. Stenger, D. C., and McMahon, C. L. 1997. Genotypic diversity of beet curly top virus populations in the western United States. *Phytopathology* 87:737-744.
 56. Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.