Genetic Variation of *Citrus Tristeza Virus* Isolates from California and Spain: Evidence for Mixed Infections and Recombination

LUIS RUBIO,† MARÍA ANGELES AYLLÓN,‡ PING KONG, ANDRES FERNÁNDEZ, MARYLOU POLEK, JOSÉ GUERRI, PEDRO MORENO, AND BRYCE W. FALK†

Plant Pathology Department, University of California, Davis, California 95616; Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain; and Central California Tristeza Eradication Agency, Tulare, California 93224

Received 9 March 2001/Accepted 18 May 2001

We examined the population structure and genetic variation of four genomic regions within and between 30 *Citrus tristeza virus* (CTV) isolates from Spain and California. Our analyses showed that most isolates contained a population of sequence variants, with one being predominant. Four isolates showed two major sequence variants in some genomic regions. The two major variants of three of these isolates showed very low nucleotide identity to each other but were very similar to those of other isolates, suggesting the possibility of mixed infections with two divergent isolates. Incongruencies of phylogenetic relationships in the different genomic regions and statistical analyses suggested that the genomes of some CTV sequence variants originated by recombination events between diverged sequence variants. No correlation was observed between geographic origin and nucleotide distance, and thus from a genetic view, the Spanish and Californian isolates analyzed here could be considered members of the same population.

*Citrus tristeza virus* (CTV) is distributed worldwide and is the causal agent of one of the most economically important diseases of citrus. CTV, a member of the genus *Closterovirus*, is phloem limited and is transmitted by aphids in a semipersistent manner. CTV virions are filamentous flexuous particles about 2,000 nm long, with two coat proteins (CP and CPm) covering 95 and 5% of the particle length, respectively (8). The CTV genome is a single-stranded, positive-sense RNA of 19,226 to 19,296 nucleotides (nt) (18, 27, 48, 51) organized in 12 open reading frames encoding at least 19 proteins. These include two papain-like proteases, replication-associated proteins (RNA polymerase, helicase, and methyltransferase), a homologue of the HSP70 protein, two coat proteins (CP and CPm), RNA-binding protein p23 (23), a p20 protein that accumulates in the amorphous inclusion bodies (14), and other proteins of so far unknown function (p61, p13, and p18) (Fig. 1). CTV-infected plants contain, in addition to the genomic RNA, 3′-coterminal subgenomic RNAs (15) and defective RNAs (D RNAs), the latter resulting from extensive internal deletions of the genomic RNA (2, 26, 28, 50).

CTV isolates differing in the type and intensity of symptoms induced in different citrus species and cultivars and in their aphid transmissibility have been reported worldwide (38). In the last two decades, efforts have been taken to develop molecular techniques for rapid differentiation of CTV isolates and identification of molecular markers related to CTV-induced symptoms. Variation in serological reactivity, peptide maps of the coat protein, double-stranded RNA (dsRNA) patterns, hybridization with cDNA probes, restriction fragment length polymorphism, and single-strand conformation polymorphism (SSCP) have been described in attempts to differentiate CTV isolates (29).

Nucleotide sequence analysis is the most accurate procedure for CTV differentiation and estimation of molecular or genetic variation. To date, the complete genome nucleotide sequences of the five CTV isolates T36 and T30 from Florida (1, 18, 34), VT from Israel (27), SY568 from California (51), and T385 from Spain (48) have been reported. Also the partial nucleotide sequences of several CTV isolates have been reported (1, 17, 22, 25, 35, 36). Recently, it has been shown that individual CTV isolates are composed of a population of sequence variants (3, 19, 22). These reports showed the genetic differences between CTV isolates but did not estimate the genetic diversity of natural populations of CTV. Only Moya and García-Arenal (31) estimated the genetic diversity of CTV in Spain based on the number and position of D RNAs associated with Spanish CTV isolates (14). However, for CTV, nucleotide identity seems not to be correlated with the similarity of D RNA patterns (2).

In this study, we assessed the structure and genetic diversity of four genomic regions from two natural CTV populations located in two important citrus-growing areas, Spain and California. By SSCP analysis we estimated the population structure of sequence variants within individual isolates. Nucleotide analysis was used to estimate the genetic distance between sequence variants. These analyses open new insights about mixed infections, recombination, and spatial population structure in the effort to understand CTV complexity.

**MATERIALS AND METHODS**

**Virus isolates.** Nineteen Californian and 11 Spanish CTV isolates were obtained from field trees and maintained in sweet orange (*Citrus sinensis*) plants in insect-proof greenhouses. Symptom evaluation was performed for these isolates in sweet orange plants and in graft-inoculated Mexican lime (*Citrus aurantifolia*) and grapefruit (*Citrus paradisi*) plants. Geographic origin and symptoms induced by these CTV isolates are summarized in Table 1. More detailed information about the biological characteristics of CTV Spanish isolates can be found in Ballester-Olmos et al. (4).
FIG. 1. Unrooted maximum-likelihood phylogenetic trees of genomic regions A, F, C, and P (see also Table 2) of 34 CTV isolates (Table 1), constructed using the PHYLIP program DNAML. Bootstrap values of between 600 and 800 for 1,000 replicates are indicated by #, and values greater than 800 are indicated by *.

Branch lengths are proportional to the genetic distances. Boxes include sequences with a nucleotide identity of greater than 99%. Above is a layout of the CTV genome, with the regions analyzed in black boxes. When an isolate contained more than one major sequence, these are indicated by a letter (corresponding to the genomic region) and a number (corresponding to the sequence variant), e.g., 65A1 and 65A2 occurred in genomic region A of isolate 65.
**TABLE 1.** Origin and biological characterization of CTV isolates

<table>
<thead>
<tr>
<th>Origin</th>
<th>Isolate no.</th>
<th>Specific origin (city, county or province)</th>
<th>Symptoms induced(^a)</th>
<th>Mexican lime</th>
<th>Sweet orange</th>
<th>SY</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>5</td>
<td>Lindsay, Tulare</td>
<td>2 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Exeter, Tulare</td>
<td>3 0 –</td>
<td>2 0 +</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>Edison, Kern</td>
<td>2 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>Exeter, Tulare</td>
<td>0 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>G103</td>
<td>Ventura</td>
<td>2 2 +</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>North Exeter, Tulare</td>
<td>2 2 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>122</td>
<td>Exeter, Tulare</td>
<td>2 1 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>143</td>
<td>Mc Farland, Kern</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>161</td>
<td>Fresno</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>162</td>
<td>McFarland, Kern</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>173</td>
<td>Delano, Kern</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>McFarland, Kern</td>
<td>1 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>364</td>
<td>Delano, Kern</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>379</td>
<td>Lindsay, Tulare</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>381</td>
<td>Lindsay, Tulare</td>
<td>0 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>384</td>
<td>Hemet, Riverside</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>386</td>
<td>Riverside City, Riverside</td>
<td>0 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>416</td>
<td>Fresno</td>
<td>N N N</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>519</td>
<td>Riverside</td>
<td>0 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td>Spain</td>
<td>T300</td>
<td>Puebla Larga, Valencia</td>
<td>1 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T308</td>
<td>Burjasot, Valencia</td>
<td>3 0 +</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T303</td>
<td>Almenara, Castellon</td>
<td>3 0 +</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T311</td>
<td>Xeresa, Valencia</td>
<td>1 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T315</td>
<td>Huercal, Almeria</td>
<td>1 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T340</td>
<td>Godella, Valencia</td>
<td>2 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T346</td>
<td>Bences, Sevilla</td>
<td>2 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T362</td>
<td>Tortosa, Tarragona</td>
<td>3 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T373</td>
<td>Alahma, Murcia</td>
<td>2 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T398</td>
<td>Alcira, Valencia</td>
<td>2 2 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
<tr>
<td></td>
<td>T405</td>
<td>Alcira, Valencia</td>
<td>2 0 –</td>
<td></td>
<td></td>
<td>0  -</td>
</tr>
</tbody>
</table>

\(^a\) Symptoms induced by CTV isolates in different hosts: Mexican lime (vein clearing and stem pitting) and sweet orange (stunting and stem pitting). Symptom intensity is scored from 1 to 3, with 1 being mild and 3 very severe. SY, seedling yellows reaction in sour orange and grapefruit; and +, absence and presence of symptoms, respectively, N, characterization not done.

**RESULTS AND DISCUSSION**

Gene expression and disease development. CTV isolates induce a range of symptoms in different Citrus species and cultivars, which are often used as indication of disease presence. For example, isolates from California and Spain are known to induce severe seedling yellows in sour orange and grapefruit, while isolates from Japan are associated with mild symptoms in these hosts (19, 40). To confirm the presence of CTV in field samples, RT-PCR was performed using specific primer pairs (Table 2) and a 20-μl reaction mixture containing 1× PCR buffer, 1.5 mM MgCl\(_2\), 200 μM each dNTPs, 20 ng of each primer, and Taq DNA polymerase (Promega). The following PCR conditions were used: 94°C for 2 min; 30 cycles each of 94°C for 30 s, 50°C for 30 s, and 72°C 40 s; and 72°C for 5 min. The resulting RT-PCR products were separated by electrophoresis in a 2% agarose gel and detected by ethidium bromide staining.

**TABLE 2.** Primers designed for PCR amplification of four regions of CTV genomic RNA

<table>
<thead>
<tr>
<th>Primer(^a)</th>
<th>Nucleotide sequence</th>
<th>Genomic region</th>
<th>Positions(^b)</th>
<th>Size(^c) (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-forward</td>
<td>5′-ACGTTGTCGGAAACGCGG-3′</td>
<td>A</td>
<td>2021–2039</td>
<td>528</td>
</tr>
<tr>
<td>A-reverse</td>
<td>5′-GCGTTCGCGGAAAGCGG-3′</td>
<td></td>
<td>2527–2548</td>
<td>231</td>
</tr>
<tr>
<td>F-forward</td>
<td>5′-GCGTTCGCGGAAAGCGG-3′</td>
<td>F</td>
<td>3561–3582</td>
<td>21</td>
</tr>
<tr>
<td>F-reverse</td>
<td>5′-GCGTTCGCGGAAAGCGG-3′</td>
<td></td>
<td>3978–3998</td>
<td>19</td>
</tr>
<tr>
<td>CP-forward</td>
<td>5′-ACGCGCCTTGGTTGGGTA-3′</td>
<td>C</td>
<td>16515–16539</td>
<td>273</td>
</tr>
<tr>
<td>CP-reverse</td>
<td>5′-ACGCGCCTTGGTTGGGTA-3′</td>
<td>P</td>
<td>16763–16787</td>
<td>240</td>
</tr>
<tr>
<td>P20-forward</td>
<td>5′-ACCGGCTGCGCCAGCCTTA-3′</td>
<td></td>
<td>17720–17740</td>
<td>283</td>
</tr>
<tr>
<td>P20-reverse</td>
<td>5′-ACCGGCTGCGCCAGCCTTA-3′</td>
<td></td>
<td>18262–18280</td>
<td>258</td>
</tr>
</tbody>
</table>

\(^a\) Oligonucleotide primers used. Forward indicates primers complementary to the negative strand, and reverse indicates those complementary to the positive strand.

\(^b\) Nucleotide positions of the primers in the genomic RNA of CTV isolate T385 (48).

\(^c\) Expected size of the RT-PCR products.
isolate of RNA virus is expected to consist of a population of genetically related variants, termed quasispecies (16). To estimate the within-isolate population structure, we performed RT-PCR of four genomic regions (Table 2) of 11 Spanish and 19 Californian CTV isolates and analyzed the SSCP patterns of 10 clones obtained from each genomic region and isolate. Previously we assessed the accuracy of these SSCP analyses (19). The average number of nucleotide differences per site between pairs of CTV sequence variant clones was only 0.001346 ± 0.0002126, supporting SSCP analysis as a precise tool for population studies. To minimize primer-directed selection of sequence variants within a isolate, the primers used here (Table 2) were designed from nucleotide sequences conserved for the five CTV isolates whose complete genome sequences are known (1, 18, 27, 34, 48, 51). Previously, we also found that the intensities of the DNA bands in SSCP profiles reflected the relative proportion of CTV RNA variants within an isolate, suggesting that these primers do not bind preferentially to some CTV sequence variants (42). Furthermore, to minimize the possibility that nucleotide incorporation errors in the initial phases of the RT-PCR give rise to a detectable subpopulation of false mutants (20), we used as a template a large amount of dsRNAs (~200 ng). Nonetheless, because we analyzed only the major sequence variants, RT-PCR-induced errors, if present, did not affect our analysis. Finally, to minimize possible variations due to irregular distribution of sequence variants in infected tissues, the dsRNAs were extracted from different branches of a large plant and pooled.

Our results and the results reported previously (19) showed that 26 of 30 field CTV isolates, in the four genomic regions analyzed, had a within-isolate population structure consisting of one major variant (frequency greater than 0.7) and other sequence variants of lower frequencies (minor variants), a typical quasispecies structure (16). In contrast, genomic region A of isolate 65 (19), genomic region C of isolates T398 and 386, and genomic region P of isolates T398 and T405 had two major variants with frequencies of 0.4 or greater. SSCP frequencies and nucleotide distances for all variants from Californian isolates 65, 107, 122, 173, and 190 have been described in detail by Kong et al. (19). We determined the nucleotide sequences of all CTV major variants and estimated the nucleotide distances between them. The nucleotide distances between sequences are represented in Fig. 1 as branch lengths of phylogenetic trees (see next section). Interestingly, for 65 A, T398 and 386 C, and T398 P, the nucleotide distances between the two major variants within the same isolate were high (Fig. 1). For example, the nucleotide distance of genomic region P between the two major variants of isolate T398 (T398P1 and T398P2) was 0.0795. Furthermore, within these isolate regions, each major variant was significantly more similar to major variants of other isolates than to the other major variant from the same isolate (Fig. 1). For example, the major variant T398P2 and that of isolate T300 showed a nucleotide distance of 0.0019, whereas T398P1 and T346 had a nucleotide distance of 0.0039. These data suggest that 65 A, T398 and 386 C, and T398 P could have originated from mixed infections of two CTV isolates with diverged sequence variants. Also, the fact that some genomic regions in the same isolate have two diverged major variants whereas other genomic regions have only one suggests the possibility of recombination events between sequence variants of the original coinfecting CTV isolates.

Mixed infections are possible, as CTV hosts are long-lived perennial plants (some living 100 years or more), allowing the possibility of repeated inoculations of CTV by viruliferous aphids. Although we have evidence for mixed infections in a small proportion of CTV isolates here analyzed (3 of 30), it is probable that mixed CTV infections occur more frequently in nature. With the approaches used here, mixed infections could only be detected between two isolates with diverged sequence variants, and both sequences should be in relatively high proportion when they were sampled. The coexistence of two quasispecies in the same host is probably not common, as differences in fitness would likely cause the displacement of one quasispecies.

**Genetic variation in different genomic regions: evidence for natural recombination events.** The major variant sequences from the CTV isolates analyzed here and the sequences of CTV isolates T36, VT, T385, and SY568 (18, 27, 34, 48, 51) were used to estimate the nucleotide diversity (average number of nucleotide substitutions per site in each pair of sequence variants) for each genomic region. These analyses showed appreciable differences between the different genomic regions. Genomic region A showed the greatest diversity, twice that for F and P and three times that for the C genomic region (Table 3). To estimate the degree of selective constraints on each genomic region, nonsynonymous and synonymous substitutions were computed separately. The number of synonymous substitutions per synonymous site (dS) was similar over the four genomic regions analyzed (Table 3). However, the number of nonsynonymous substitutions per nonsynonymous site (dN) was smaller than dS and varied considerably between genomic regions (Table 3). This suggests a negative selective pressure for most amino acid changes (functional constraints) and that the degree of the functional constraints varies for the different genomic regions. Genomic regions F and P had a ratio dN/dS in the range of most DNA and plant RNA virus protein-coding sequences (12, 32). F corresponds to part of the methyltransferase domain, involved in virus replication, and P corresponds to gene p20, coding for a protein of unknown function that accumulates in amorphous inclusion bodies (13). Genomic region A, which so far is not known to be part of any known functional domain, showed the greatest dN/dS ratio. Finally, C, corresponding to a portion of the coat protein gene, showed the greatest functional constraints. The CTV coat protein, in addition to constraints related to virion structure and

---

**Table 3. Average number of nucleotide substitutions between CTV isolates in different genomic regions (A, F, CP, and P20)**

<table>
<thead>
<tr>
<th>Region</th>
<th>dD</th>
<th>dN</th>
<th>dS</th>
<th>dN/dS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.13656 (0.04209)</td>
<td>0.11125</td>
<td>0.20472</td>
<td>0.56323</td>
</tr>
<tr>
<td>F</td>
<td>0.05993 (0.01601)</td>
<td>0.03455</td>
<td>0.29849</td>
<td>0.12302</td>
</tr>
<tr>
<td>C</td>
<td>0.03792 (0.00626)</td>
<td>0.00511</td>
<td>0.25128</td>
<td>0.02739</td>
</tr>
<tr>
<td>P</td>
<td>0.06029 (0.00763)</td>
<td>0.03221</td>
<td>0.25093</td>
<td>0.13424</td>
</tr>
</tbody>
</table>

* dD, nucleotide diversity; average number of nucleotide substitutions per site between pairs of sequences. Standard errors are indicated in parentheses. dN, average number of nonsynonymous substitutions per nonsynonymous site. dS, average of synonymous substitutions per synonymous site. dN/dS, average of the ratio between nonsynonymous and synonymous substitutions for each pair of comparisons. dD, dN, and dN/dS were estimated by the PBL method (21, 33).
stability, might play a critical role in interactions with its aphid vector and/or plant host. The \(dN/dS\) ratio of CTV genomic region C was very low in comparison with that of other plant virus coat protein genes (12).

Phylogenetic relationships between the CTV isolates were inferred using neighbor-joining and maximum-likelihood methods (9). Both gave basically the same relative phylogenetic grouping and showed bootstrap values greater than 80% in the main nodes (Fig. 1). Most CTV isolate major variants showed the same grouping in the phylogenetic trees obtained from the four genomic regions analyzed here. However, some isolate major variants showed sharp differences in their genetic relationships with other CTV sequences in different genomic regions. For example, Spanish isolate T308 was genetically very close to Florida isolate T36 and very different from isolate SY568 RUBIO ET AL. J. VIROL. between nt 40 and 80 (Fig. 3). Another clear example would be genomic region C of the major variant T398C2. When the identity profiles of T398C2 with respect to T398C1 and T385 were displayed, it was observed that from positions 110 to 170, the three sequences show 100% nucleotide identity; T398a and T398b have 100% nucleotide identity in the segment between positions 110 and 273; and T398C2 and T385 have 100% identity between positions 1 and 170. All this suggests that the T398C2 sequence might have resulted from recombination between T398C1-like and T385-like sequences at some point between positions 110 and 170. It is possible that the complete identity between nucleotides 110 and 170 resulted from a previous recombination of the ancestor of T385 and T398C1 with another sequence variant (Fig. 3). In all cases, the recombination events inferred from the nucleotide identity profiles (Fig. 3) confirmed those obtained from the phylogenetic correlation profiles (Fig. 2). It seems unlikely that the possible recombination sites could have arisen in vitro during RT-PCR, as in our previous work no recombinant variants were observed when dsRNAs from two isolates, each with one major variant, were mixed and amplified by RT-PCR (42).

We also performed conversion analyses using the program GENECONV (45) on the complete genome nucleotide sequences of isolates T36, VT, T385, and SY568. This program is based on the analysis of whether some regions of a pair of sequences have more consecutive identical silent polymorphic sites in common than would be expected by chance. GENECONV finds and ranks fragments with the highest score (number of matches for pairs of sequences). To avoid the selection of high-scoring fragments by chance, the polymorphic sites are permuted randomly among themselves 10,000 times and scored each time. Statistical significance is evaluated by the parameter \(P\), the proportion of permuted alignments for which the maximal fragment score for that pair of sequences is greater than or equal to the original fragment score. Because mutations could accumulate after the recombination event, we used the program option to allow mismatches (penalty was set for a gscale of 2). Our GENECONV analyses suggested recombination between isolates T385 and SY568, which showed a nucleotide identity of greater than 99% between positions 9304 and 16107 \((P = 0.0000)\), whereas the rest was below 93%, which corroborates the results obtained by Vives et al. (48). Also, we found a possible double recombination between VT and SY568: SY568 and VT showed a nucleotide identity of about 96% in positions 1309 to 2781 \((P = 0.0027)\) and 6546 to 9302 \((P = 0.0000)\), whereas the rest was below 92%.

Recombination in natural populations has been reported for other plant viruses (5, 10). It has been proposed that recombination can be advantageous for RNA viruses. The high mutation rates of RNA viruses cause accumulation of deleterious mutations, limiting the RNA genome size. In large populations, fitness can be maintained or increased by natural selection, but in small populations, genetic drift will lead to progressive loss of replicative fitness. Recombination of viral genomes with deleterious mutations can regenerate functional genomes (6, 39, 46). For CTV, having the largest genome among the known single-stranded plus-sense plant RNA viruses, recombination could act as a compensatory mechanism to offset accumulation of deleterious mutations in bottleneck episodes, such as aphid transmission. Also, coinfection and recombination of different genomic regions between diverged
FIG. 2. Phylogenetic correlation profiles (graphic representation of the coherence of sequence relationships) of 34 CTV isolates (Table 1) for genomic regions A, F, C, and P (Table 2). Only the variable sites (x axis) are represented in the graph. Phylogenetic correlation (y axis) was obtained at each variable site from pairwise distance analysis of all aligned sequences by using the program PHYLPRO, with a fixed window of 40 bp. Numbers under low phylogenetic correlation areas (possible recombination signals) indicate nucleotide positions. Some CTV sequences are indicated near their individual correlation profiles.
virus genomes allow greater genome diversity and adaptability to new environments (39, 46). Thus, recombination might explain in part the great number of CTV isolates with different biological characteristics described worldwide (38). Recombination among CTV isolates also has important practical implications. For example, for application of disease control measures, such as cross-protection or transgenic plant resistance, caution must be taken to avoid the introduction of exotic CTV sequences that might recombine and give rise to CTV isolates with new biological properties.

Genetic variation with respect to geographic distribution: absence of correlation between genetic and geographic proximity. According to CTV genomic RNA 5′ untranslated region nucleotide identity, Lopez et al. (22) classified clones from 11 CTV isolates in three groups: I, represented by isolate T36; II, represented by VT; and III, represented by T385. According to our analyses, genomic regions A and F (both located in the 5′ half of the genome; Fig. 1) of all the Spanish and Californian isolates analyzed here can be included in these three groups. The major variants of Spanish isolates T346 and T373 are included in group I, three to five Californian isolates are in group II, and about 80% of the Spanish and California isolates belonged to group III (Fig. 1). For genomic regions C and P, the CTV isolates could not be readily assigned to these three discrete groups. Curiously, group III isolates showed a nucleotide identity of greater than 99% in the four genomic regions (Fig. 1). Group III isolates, with a very high nucleotide identity, has been also found in Florida, Taiwan, and Colombia (1).

For a better comparison of CTV populations, we estimated the nucleotide diversity (average nucleotide distance between two pairs of CTV major sequences chosen randomly) of Spanish and Californian CTV isolates (including T385 and SY568) for each genomic region (Table 4). We found that in each genomic region, the nucleotide diversity of CTV isolates from the same geographic population was remarkably higher than that assessed between the two geographic populations (Table 4). For example, the nucleotide diversity of CTV genomic region P within California and within Spain was about 0.050 for both, whereas the nucleotide diversity considering only nucleotide distances between Spanish and Californian isolates was less than 0.004 (Table 4). Application of the $D$ statistic (24) showed that the Spanish and Californian CTV populations were not significantly genetically different ($D = 0.0000$), and hence, from a genetic view, the Spanish and Californian CTV
isolates can be considered part of the same population. We wanted to know if this was also true for other CTV geographic populations. Unfortunately, not many CTV nucleotide sequences from the same geographic region are available. When genomic region C of six Portuguese isolates (GenBank accession nos. AF184113-AF84118) was included in the analysis, we found that the Portuguese CTV isolates were part of the same genetic population as the Spanish and Californian CTV isolates (Table 4). Analysis of the first 462 nt of the coat protein gene (C2) of 6 Portuguese and 13 Japanese CTV isolates (17) also showed higher diversity within a geographic population than between the two geographic populations (Table 4). Low genetic variation among geographically distant isolates has also been observed for *Cucurbit yellow stunting disorder virus* and *Beet pseudo-yellows virus* in the genus *Crinivirus*, the other genus of the family *Closteroviridae* (41, 43), and for *Tobacco mild green mosaic virus* and *Pepper mild mottle virus* in the genus *Tobamovirus* (11, 37).

The genetic structure observed here suggests migration of CTV isolates among geographically isolated CTV populations. It is well known that an intense traffic of CTV-infected propagative citrus material has occurred between distant regions in the world (38). Within each geographic region, local dispersion is then effected by aphids.

**ACKNOWLEDGMENTS**

Part of this work was supported by grants from the USDA special grants Citrus Tristeza Virus Research Program, the California Tristeza Research Coalition, and the University of California to B.W.F. Part of this work was supported by projects SC93–111 and SC97–098 (INIA). L.R. was supported in part by a postdoctoral fellowship from Ministerio de Educación y Ciencia, Spain.

We acknowledge the technical assistance of T. Olopona, M. Boil, E. Estela, and M. Martínez. We thank F. Garcia-Arenal for excellent critical review of the manuscript.

**REFERENCES**


