Effect of citrus hosts on the generation, maintenance and evolutionary fate of genetic variability of citrus exocortis viroid

Lucía Bernad,1 Núria Duran-Vila1 and Santiago F. Elena2,3

1Centro de Protección Vegetal y Biotecnología, Instituto Valenciano de Investigaciones Agrarias, Moncada, 46113 València, Spain
2Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-UPV, Campus UPV CPI 8E, 46022 València, Spain
3The Santa Fe Institute, Santa Fe, NM 87501, USA

Citrus exocortis viroid (CEVd) populations are composed of closely related haplotypes whose frequencies in the population result from the equilibrium between mutation, selection and genetic drift. The genetic diversity of CEVd populations infecting different citrus hosts was studied by comparing populations recovered from infected trifoliate orange and sour orange seedling trees after 10 years of evolution, with the ancestral population maintained for the same period in the original host, Etrog citron. Furthermore, populations isolated from these trifoliate orange and sour orange trees were transmitted back to Etrog citron plants and the evolution of their mutant spectra was studied. The results indicate that (i) the amount and composition of the within-plant genetic diversity generated varies between these two hosts and is markedly different from that which is characteristic of the original Etrog citron host and (ii) the genetic diversity found after transmitting back to Etrog citron is indistinguishable from that which is characteristic of the ancestral Etrog citron population, regardless of the citrus plant from which the evolved populations were isolated. The relationship between the CEVd populations from Etrog citron and trifoliate orange, both sensitive hosts, and those from sour orange, which is a tolerant host, is discussed.

INTRODUCTION

Viroids are small plant pathogens consisting of a naked single-stranded circular RNA molecule of 246–475 nt, that do not encode any proteins but are endowed with autonomous replication in their host plants. It is very likely that viroids have been co-evolving with their hosts since their origin, although many of them show a wide host range. The identification of viroids in wild and cultivated plants, as symptomless carriers, suggests that certain hosts may act as natural viroid reservoirs (Diener, 1995).

Viroids are taxonomically classified into two families, Pospiviroidae and Avsunviroidae (Elena et al., 1991). These families mainly differ in three characteristics. First, all pospiviroidae present a central conserved region (CCR) in their rod-like secondary structure, whereas the avsunviroidae lack such a region. Second, the pospiviroidae rely on cellular factors to process their multimeric replication intermediates into unit-length molecules, while the avsunviroidae present hammerhead ribozyme structures that are able to self-cleave the multimeric forms. Third, the pospiviroidae replicate in the nucleus whereas the avsunviroidae replicate in the chloroplast (reviewed by Flores et al., 2005).

Like most RNA and some DNA viruses, viroids replicate within their hosts as polymorphic populations composed of closely related sequence variants generally distributed around a predominant one. This polymorphic population structure arises as a result of (i) the high mutation rates inherent to the cellular DNA-dependent RNA polymerases involved in viroid replication subverted to replicate an RNA template and (ii) the diverse and fluctuating selective pressures imposed by the different host species. Surveys of diversity have been performed for different viroid species, including pospiviroidae such as citrus exocortis viroid (CEVd) (Visvader & Symons, 1985; Gandía et al., 2005, 2007), citrus dwarfing viroid (formerly citrus viroid III) (Owens et al., 2000), citrus bent leaf viroid (Foissac & Duran-Vila, 2000; Gandía & Duran-Vila, 2004), potato spindle tuber viroid (PSTVd) (Góra et al., 1994; Gruner et al., 1995; Góra-Sochacka et al., 1997), hop stunt viroid (HSVd) (Kofalvi et al., 1997; Palacio-Bielsa et al., 2004) and grapevine yellow speckle viroid 1 (Rigden & Rezaian, 1993; Polivka et al., 1996), as well as avsunviroidae such as chrysanthemum chlorotic mottle viroid (Navarro & Flores, 1997; Codoñer et al., 2006), peach latent mosaic viroid (Hernández & Flores, 1992; Ambroès et al., 1998, 1999) and avocado sunblotch viroid (Rakowski & Symons, 1989).
CEVd is a member of the genus *Pospiviroid* within the family *Pospiviroidae*. Like other members of this family, the highly base-paired rod-like secondary structure conforms to the model of five structural domains proposed by Keese & Symons (1985); these domains are terminal left (TL), pathogenicity (P), central (C), variable (V) and terminal right (TR). The genus *Pospiviroid*, in addition to its characteristic CCR, presents another conserved region in the TL domain, the terminal conserved region (TCR). CEVd is the causal agent of the exocortis disease characterized by a bark shelling or scaling disorder of trifoliate orange (*Poncirus trifoliata* L. Raf.) used as rootstock (Fawcett & Klotz, 1948). The citrange hybrids (*Citrus sinensis* L. × *P. trifoliata*) and Rangpur lime (*Citrus limonia* Osb.), both used as rootstocks, are also sensitive and develop bark scaling symptoms and stunting. In the Etrog citron indicator (*Citrus medica* L.), CEVd induces severe stunting, leaf epinasty and vein necrosis. The aim of the present work was to study the evolution of CEVd populations infecting different citrus hosts. More precisely, we have compared the genetic diversity of the CEVd populations recovered from infected trifoliate orange and sour orange seedling trees with that of the ancestral population maintained in Etrog citron that was used as the inoculum source. Our results indicate that the amount and composition of the genetic diversity generated after 10 years of evolution varied between these two hosts and was markedly different from the characteristics of the original Etrog citron population. In a second experiment, we transferred the populations isolated from trifoliate orange and sour orange trees back to Etrog citrons. The results of this second short-term evolution experiment show that the genetic diversity evolved was basically indistinguishable from the diversity characteristic of the original Etrog citron. Altogether, these results support the notion that the composition and structure of viroid populations is determined by the host in which they replicate.

**METHODS**

**Plant material and viroid sources.** A CEVd isolate (CEVd-117) maintained in the sensitive selection 861-S1 of Etrog citron grafted onto rough lemon (*Citrus jambhiri* Lush.) rootstock was grafted-transmitted to two trifoliate orange and two sour orange seedlings in 1992 and the inoculated plants were transplanted to an experimental field at the Instituto Valenciano de Investigaciones Agrarias the following year. Ten years later, the trifoliate orange seedlings that were severely stunted and presented the characteristic bark scaling symptoms were used as a source of inoculum for graft-transmission to two new Etrog citron plants. After the same period, the sour orange seedlings that remained symptomless and were indistinguishable from the non-inoculated controls were similarly used as a source of inoculum for graft-transmission to two new Etrog citron plants (Fig. 1). The inoculated citrons were maintained in a greenhouse at 28–32 ℃ for at least 6 months and the field-grown trifoliate and sour orange seedlings were used as source tissues for the characterization of their CEVd populations.

**Nucleic acid extraction, cDNA synthesis and cloning.** Samples (0.5 g bark and leaves) of citron, trifoliate orange and sour orange were homogenized inside sealed plastic bags containing 5 ml extraction buffer (0.1 M Tris/HCl pH 8.5, 50 mM EDTA, 0.5 M NaCl, 10 mM β-mercaptoethanol). The homogenates were subjected to alkaline denaturation with SDS (65 ℃ for 20 min) and potassium acetate (on ice for 20 min). The solubilized fraction was concentrated by ethanol precipitation and resuspended in 40 μl sterile water.

First-strand viroid cDNA was synthesized with 15 U ThermoScript RNase H- reverse transcriptase (ThermoScript-RT; Invitrogen) using the reverse specific primer complementary to the upper CCR strand, CEV-RT (5′-CTTCCTCAGGTGTTCCTCGGGAATCC-3′) (0.75 μM) and dNTPs (1 mM each). The ThermoScript-RT reaction mixture contained 50 mM Tris/acetate (pH 8.4), 75 mM potassium acetate, 8 mM magnesium acetate and 40 U RNase Out (Invitrogen). The reaction mixture (20 μl final volume) was incubated at 60 ℃ for 1 h. Second-strand DNA synthesis and PCR amplification (50 μl final volume) were performed using 4 μl of the first-strand mixture, 3.5 U Expand high fidelity PCR system (Roche), the reverse and forward primers CEV-R1 (5′-CCGGGGATCCCTGAAGGA-3′) and CEVd-F1 (5′-GGGAAAACCTGGAAGAATCG-3′) (0.5 μM each) and dNTPs (0.12 mM each) in a buffer containing 150 mM MgCl₂. PCR parameters consisted of a denaturation step at 94 ℃ for 5 min, followed by 35 cycles of 94 ℃ for 30 s, 60 ℃ for 30 s and 72 ℃ for 1 min (Bernard & Duran-Vila, 2006).

Electrophoretic analysis in 2 % agarose gels confirmed the synthesis of a DNA product of the expected size. The RT-PCR products were purified (using an Amersham kit) and ligated into the pGEM-T vector and used to transform *Escherichia coli* DH5α competent cells. Transformants were grown for approximately 20 h at 37 ℃ on ampicillin-containing plates and 30 colonies were randomly selected for sequencing.

To verify the sequence of the region of the upper strand corresponding to the primers used for RT-PCR, a second RT-PCR was performed under the same conditions described above but using a specific reverse primer complementary to the lower CCR strand, CEV-RT2 (5′-CCGGGTATGTATCCAGAGAGAAGCTCCG-3′) and the reverse and forward primers CEV-R2 (5′-GGGGATCTCCAGAGAGAAG-3′) and CEV-F2 (5′-GGGGTAAACCTGGAAGAATCG-3′).

**Sequence analysis.** Cloned full-length viroid cDNAs were sequenced with an ABI PRISM DNA analyser 377 (Perkin-Elmer). Chromatograms were edited with Chromas v.1.43. Multiple sequence alignments were generated with the CLUSTAL_W program ( Higgins & Sharp, 1989). Minor adjustments were introduced manually in the final alignment to maximize the sequence identity [all these programs are integrated in MEGA version 3.1 (Kumar et al., 2004)].

Minimum free energy secondary structures (MFESS) of viroids were predicted using the Mfold algorithm (Zuker, 2003) as implemented in the Mfold server (www.bioinfo.rpi.edu/applications/mfold/cgi-bin/ RNAform1.cgi), choosing the option of circular RNA molecules. These were drawn with the RNAviz 2.0 software (De Rijk et al., 2003).

**Nucleotide diversity calculation and assessment of differences among host trees.** As a measure of CEVd genetic diversity present on each analysed tree, the Shannon entropy (Shannon, 1948; Korber et al., 1994) was calculated from the alignment of the sequences recovered from each tree. For a multiple sequence alignment, the Shannon entropy (H) for every position (j) was calculated as $H_j = -\sum_i^n P_i \log P_i$, where $P_i$ is the fraction of residues of nucleotide $i$ and $M$ is the number of different characters in the sequence alphabet (five, including the four nucleotides and deletions). The total entropy of a population was then estimated as the sum of entropies from all sites in the genome using $H = \sum_{j=1}^L H_j$, where $L$ is the length of the sequence alignment.
RESULTS

Molecular characterization and genetic diversity of CEVd populations

The CEVd isolate (CEVd-117) maintained in citron (plant Ci) was used as a source of inoculum for transmissions to two trifoliate orange (T1 and T2) and two sour orange (S1 and S2) seedling trees (Fig. 1a). The consensus sequence was determined by directly sequencing two full-length RT-PCR amplicons of CEVd obtained using two different sets of primers. In order to minimize the introduction of artefactual changes, PCR amplification was performed using a DNA polymerase with proofreading activity. The consensus sequence showed 98% nucleotide identity with the reference sequence CEVd (class A) (Visvader et al., 1982) and the predicted MFESS was a highly base-paired rod-like secondary structure characteristic of viroids of the family Pospiviroidae (Fig. 2).

![Diagram](a) Inoculated plants

![Diagram](b) CEVd populations

---

Fig. 1. (a) CEVd transmission scheme. Etrog citron (Ci) infected with a CEVd isolate (CEVd-117) was used as the inoculum source for a first graft-transmission to two trifoliate orange seedlings (T1 and T2) and two sour orange seedlings (S1 and S2). After 10 years, the CEVd populations from the trifoliate orange and sour orange trees (maintained under field conditions) were graft-transmitted back to new Etrog citron plants (CT1, CT2, CS1 and CS2). (b) Frequency diagrams showing the genetic structures (haplotypes and frequencies) of CEVd populations retrieved from each infected tree. Shared haplotypes among populations are indicated by the following Greek letters: \( \Delta \), Ci-1; \( \sigma \), Ci-2; \( \varphi \), Ci-5; and \( \epsilon \), Ci-8.
Ten years after inoculation, the infected trifoliate orange seedling trees (T1 and T2) were severely stunted and presented the bark scaling symptoms characteristic of the exocortis disease, whereas the infected sour orange seedling trees (S1 and S2) remained symptomless. The consensus CEVd sequences recovered from T1 and T2 were found to differ by only one change (T1, A130Δ; T2, A313G) relative to the consensus CEVd sequence of the inoculum source. By contrast, the consensus CEVd sequences recovered from S1 and S2 presented multiple peaks at different chromatogram positions, suggesting that the population of sequence variants was genetically heterogeneous. These results, although preliminary, indicated that the CEVd populations in these hosts were not identical nor even similar. However, transmissions from each of the two trifoliate orange trees and from each of the two sour orange trees to new Etrog citron plants (CT1, CT2, CS1 and CS2) (Fig. 1a) resulted in the recovery of CEVd populations with consensus sequences that were all identical to that of the original Etrog citron used as the initial inoculum source (Ci), strongly suggesting that the fittest CEVd variant was different on each host tree.

Since RNA viruses and viroids replicate as complex populations of sequence variants, consensus nucleotide sequences provide rather limited information about the genetic diversity of the viroid populations and their evolution after transmission to different hosts (Domingo et al., 2006). Therefore, analysis of individual genomic sequences within the mutant spectra is necessary to gain deeper insights into the evolutionary processes taking place in each host. The RT-PCR amplicons obtained from each infected host (Ci, T1, T2, S1, S2, CT1, CT2, CS1 and CS2) were ligated into a cloning vector and 30 clones per sample were randomly selected for sequencing. Hence, each CEVd population was represented by 30 clones ranging from 369 to 374 nt in size, most of them being 371 nt. The characteristics of these nine populations are summarized in Table 1.

Sensitive and tolerant citrus trees determine the level of CEVd genetic diversity

There were three major findings. First, the ancestral CEVd population derived from the initial citron source of inoculum (Ci) was made up of a dominant haplotype (Ci-1), which represented 50 % of the population. A second haplotype (Ci-2), representing 23.3 % of the genetic variability, only differed from Ci-1 in one insertion at the upper strand of the P domain (+G at position 74). The rest of the population was made up of haplotypes with frequencies between 3.3 and 6.6 % (Fig. 1b). The Shannon entropy of the ancestral population was 0.0051.

Second, the CEVd populations isolated from the trifoliate oranges (T1 and T2), were both characterized by the presence of a numerically dominant haplotype (T1-1 and T2-1, respectively), representing 63.3 and 60.0 % of each population, respectively. However, these two haplotypes...
Table 1. Descriptive parameters of heterogeneous CEVd populations

<table>
<thead>
<tr>
<th>CEVd population</th>
<th>No. haplotypes/total no. sequenced clones</th>
<th>Most frequent haplotype in CEVd population (%)</th>
<th>Min.&lt; consensus&lt; max. haplotype length (nt)</th>
<th>No. polymorphic sites in the genome (%)†</th>
<th>Average no. observed mutations per haplotype‡</th>
<th>Mutation frequency (×10⁻⁵)§</th>
<th>CEVd population diversity (H)ǁ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ci</td>
<td>9/30</td>
<td>Ci-1 (50)</td>
<td>370&lt;371&lt;373</td>
<td>8 (2.16)</td>
<td>0.6331</td>
<td>7.18</td>
<td>0.0051</td>
</tr>
<tr>
<td>T1</td>
<td>9/30</td>
<td>T1-1 (63.3)</td>
<td>369&lt;370&lt;373</td>
<td>11 (2.97)</td>
<td>0.5664</td>
<td>9.99</td>
<td>0.0056</td>
</tr>
<tr>
<td>T2</td>
<td>9/30</td>
<td>T2-1 (60)</td>
<td>369&lt;371&lt;372</td>
<td>11 (2.96)</td>
<td>0.5329</td>
<td>10.80</td>
<td>0.0054</td>
</tr>
<tr>
<td>S1</td>
<td>15/30</td>
<td>S1-12 (20)</td>
<td>369&lt;373&lt;373</td>
<td>19 (5.09)</td>
<td>1.7660</td>
<td>19.80</td>
<td>0.0149</td>
</tr>
<tr>
<td>S2</td>
<td>17/30</td>
<td>S2-11 (16.6)</td>
<td>369&lt;373&lt;374</td>
<td>16 (4.29)</td>
<td>0.7595</td>
<td>14.00</td>
<td>0.0114</td>
</tr>
<tr>
<td>CT1</td>
<td>7/30</td>
<td>CT1-10 (60)</td>
<td>371&lt;373&lt;373</td>
<td>6 (1.62)</td>
<td>0.4997</td>
<td>5.39</td>
<td>0.0040</td>
</tr>
<tr>
<td>CT2</td>
<td>10/30</td>
<td>CT2-6 (50)</td>
<td>371&lt;373&lt;372</td>
<td>11 (2.96)</td>
<td>0.7328</td>
<td>8.98</td>
<td>0.0070</td>
</tr>
<tr>
<td>CS1</td>
<td>10/30</td>
<td>CS1-12 (50)</td>
<td>371&lt;373&lt;372</td>
<td>9 (2.43)</td>
<td>0.6997</td>
<td>8.08</td>
<td>0.0062</td>
</tr>
<tr>
<td>CS2</td>
<td>12/30</td>
<td>CS2-12 (36.6)</td>
<td>370&lt;373&lt;373</td>
<td>12 (3.23)</td>
<td>0.8993</td>
<td>10.80</td>
<td>0.0075</td>
</tr>
</tbody>
</table>

*CEVd populations from the following hosts were used: Ci, original citron; T1 and T2, trifoliate orange; S1 and S2, sour orange; and CT1, CT2, CS1 and CS2, citron plants graft-inoculated with T1, T2, S1 and S2 CEVd sources.
†Number of sites where one or more mutations were observed in each individual CEVd population. Percentages were obtained by dividing this by the number of nucleotides of the consensus sequence of the population.
‡Total number of mutations observed in each CEVd population/total number of sequenced clones (n=30).
§Total number of non-repeated mutations observed in each CEVd population/(number of nucleotides of the consensus sequence of the population × total number of sequenced clones).
ǁH, Shannon entropy.

were not identical to each other or to the dominant ancestral Ci-1 haplotype. Interestingly, all haplotypes from T1 shared a deletion in the upper strand of the V domain relative to the sequence of Ci-1 (−A in position 130). Similarly, most haplotypes from plant T2 shared a nucleotide substitution (A313G) in the lower strand of the P domain. None of these two characteristic mutations was detected in the ancestral CEVd population used as inoculum. It is worth noting that the second most abundant ancestral haplotype, Ci-2 has also been found in population T2, although its frequency was reduced to 3.3%. The Shannon entropy for these two populations was very similar (Table 1) and also in the same range as that obtained for the ancestral Etrog citron population.

Third, CEVd populations replicating within sour oranges (S1 and S2) were made up of a heterogeneous assembly of haplotypes, with no clearly dominant one. Hence, the most abundant haplotypes in plant S1 were S1-4 and S1-12, both at a frequency of 20%; whereas S2-11, the most frequent haplotype in plant S2, was only found at 16.6%. Remarkably, S1-12 and S2-11 haplotypes were identical to the minority haplotype Ci-8 present at 3.3% in the ancestral Ci population. Similarly, haplotype S1-2 was identical to haplotype Ci-5 present in the ancestral Ci population, but Ci-5 was found at half the frequency (3.3% in Ci versus 6.6% in S1). The numerically dominant haplotypes in the ancestral Ci populations, Ci-1 and Ci-2, are also present in populations S1 and S2, although their frequencies were lower than in the ancestral population (Fig. 1b).

To quantitatively explore the above data obtained for the new two hosts, a nested ANOVA model was fitted to the per site Shannon entropy values. In this model, host species and replicate plants were treated as random factors and plant replicates were nested within host species. Differences among sites in the alignment were used to evaluate the within-plant variability. The test found no significant differences among replicate plants within hosts (F_{2,150} = 0.345, P = 0.709), but the diversity levels in sour orange were 139% larger than those in trifoliate orange, a difference which was statistically significant (F_{2,2} = 65.178, P = 0.015) (Table 1).

It is worth noting that the mutation frequencies and entropy estimations may be somehow inflated by the intrinsic error associated with the RT-PCR amplification reaction, as the fidelity of PCR varies depending on the reaction conditions and the nature of target sequences (Cha & Thilly, 1993). However, Teycheney et al. (2005) showed that, in general, RT-PCR produced less than 0.5% errors. In addition, a previous work characterizing CEVd and HSVd diversity indicated that the observed mutations were unlikely to be due to PCR errors (Palacio & Duran-Vila, 1999; Palacio-Bielsa et al., 2004). In fact, some of the polymorphisms discussed here, such as the most frequent haplotypes (i.e. Ci-1 and Ci-2), the minority haplotypes (i.e. Ci-5 and Ci-8) and some of the nucleotides apparently more prone to mutation (i.e. A313G), were detected several times in independent hosts, thus indicating that they are real polymorphisms rather than the same PCR errors persistently occurring in independent PCRs. Furthermore,
since we are interested in the relative value of diversity measures across hosts, and all our samples were treated identically, any bias introduced by the RT-PCR treatment would be consistent, if not identical, across all samples and, therefore, would not invalidate our conclusions.

Evolutionary reversal to the ancestral Etorq citron host

In the second transmission event, back to the ancestral Etorq citron host, the population composition dramatically changed, basically reverting to the same situation described for the ancestral Ci tree. In the new populations of CT1 and CT2 haplotypes, Ci-1 and Ci-2 become dominant again, with frequencies close to the values described in the ancestral plant (Fig. 1b). Moreover, none of the haplotypes characteristic of population T1 that carried the A130Δ deletion was observed in the CT1 population, strongly suggesting that they were only beneficial in the trifoliate host but deleterious in the Etorq one. However, the A313G substitution characteristic of the T2 population was identified in not only three minority haplotypes in the CT2 population but also some haplotypes of the other CEVd populations except for Ci and S1. In the new CS1 and CS2 populations, the two most frequent haplotypes, S1-12 and S2-11 (identical to the minority variant Ci-8), were not found. However, the CS1 population was characterized by a dominant haplotype identical to the Ci-1 dominant in the Ci population, whereas the secondary haplotype Ci-2 was not detected. The Etorq citron-characteristic haplotypes Ci-1 and Ci-2 were both found at high frequencies in the new populations.

The P domain appears to be the most polymorphic region of the CEVd molecule

The division of the viroid rod-like secondary structure into five structural domains termed TL, P, C, V and TR (Keese & Symons, 1985) was proposed as an attempt to associate biological functions to the different regions of the viroid molecule. This model is still widely used to describe the molecular characteristics of the viroids of the family Pospiviroidae (Fig. 2). Next, we sought to explore whether the above nucleotide diversity was distributed equally among the five domains or whether variation hotspots existed. The distribution of the entropies along the CEVd domains was not uniform (one-way ANOVA: $F_{4,375} = 6.027, P < 0.001$) (Fig. 3), although differences were entirely due to the larger variability found in the P domain (Tukey’s post-hoc test, $P > 0.942$).

Conservation of structural motifs

Conserved sequence motifs and functional conformations typical of viroids belonging to the family Pospiviroidae were also examined. These motifs are the TCRs present only in some genera of this family and whose functions remain unknown (Flores et al., 1997), hairpins (HP) I and II probably involved in replication processes (Ding & Itaya, 2007) and the recently identified ‘RY’ motif located in the TR domain which is thought to be involved in long-distance viroid transport within plants (Gozmanova et al., 2003).

HP I, which comprises the segment of the viroid molecule used to design the primer set used for CEVd amplification, was not analysed. Only one sequence of the 270 genomes sequenced contained a mutation (C240U) in the palindrome sequences forming HP II (Fig. 2). This mutation would prevent the formation of the canonical pair C:G but would still allow formation of a non-canonical U:G pair, likely maintaining the correct folded conformation of HP II. In the TCR motif (Fig. 2), two changes, +U19 and G26A, were observed in two of the less-frequent haplotypes. However, our data do not allow us to conclude whether these two changes are selectively important or not. Finally, the sequences forming the ‘RY’ motif (Fig. 2) as well as its predicted MFESS in all CEVd populations remained conserved (data not shown).
DISCUSSION

Here we present the characterization of the genetic structure and evolution of CEVd populations in both natural and indicator citrus hosts. CEVd populations, as in other viroids and most RNA viruses, are composed of closely related haplotypes whose frequencies in the population result from the equilibrium between mutation, selection and genetic drift. This population heterogeneity is an intrinsic property of RNA replicons that has been broadly studied for many RNA plant viruses (García-Arenal et al., 2003). The characterization of CEVd populations replicating for over 10 years under field conditions in two economically important natural citrus hosts has allowed us to study the mechanisms of CEVd diversification as it adapts to different hosts in unprecedented detail.

On one hand, in trifoliate orange, a sensitive CEVd host, populations reach a remarkably constant level of genetic diversity among replicate plants. Furthermore, the amount of genetic diversity, measured here as Shannon entropy, generated and maintained in this sensitive host was indistinguishable from the variability characteristic of the ancestral Etrog citron. However, despite this similarity in diversity parameters, differences were found in the details of the mutant spectra. Different trifoliate orange plants were characterized by different master sequences and different dominant haplotypes that were also different from those characteristic of the ancestral citron population. In all cases, differences among dominant haplotypes were given by a single mutation located in the P or V domain. The evolution of these populations could be explained by genetic drift processes that occurred during the course of successive bottlenecks in both transmissions and systemic infections, and by selection of the fittest variants after colonization. A similar phenomenon had been reported previously in HSVd strain Ia transmission carried out from Etrog citron to cucumber (sensitive host) (Palacio-Bielsa et al., 2004).

On the other hand, in sour orange, a tolerant CEVd host, populations derived from two different trees were also remarkably similar in the amount of genetic diversity contained even after 10 years of infection. However, in sharp contrast to what was observed in trifoliate orange trees, these populations were much more variable than the ancestral population isolated from citron. Furthermore, these populations were not characterized by one or two majority haplotypes, as was the case for the citron and trifoliate orange trees, but by a much more complex genetic mixture, although some haplotypes (e.g. Ci-8) were shared by both trees and were also present in the ancestral citron tree. The fact that Ci-8 was present in the ancestral population (Ci) as a minority haplotype would provide a host adaptation advantage to CEVd populations. Ci-8 differed from the dominant haplotype Ci-1 (Ci) in two insertions +A62 and +G74, both located in the upper strand of the P domain, leading to a more relaxed MFESS (data not shown), which could modify (facilitate or hinder) the interaction with unknown host factors.

In both cases, after inoculating the various CEVd populations back into the original Etrog citron host, both the amount of genetic diversity, as well as the haplotype constitution (including the dominant Ci-1), reversed to a configuration that clearly reflected the ancestral population. These results further support the conclusion of host-driven adaptation and, furthermore, suggest that the evolvability of populations was not constrained by the actual host. Cases of host-driven adaptation have been widely reported for both plant RNA viruses and viroids (García-Arenal et al., 2003). However, cases have also been described in which certain citrus viruses such as citrus leaf blotch virus and citrus tristeza virus did not respond to a specific way to different citrus host species (Vives et al., 2002). Viroid populations undergo bottlenecks upon transmission to different hosts. In a previous study, a heterogeneous CEVd population infecting a symptomless broad bean plant evolved to a more homogeneous CEVd population after being inoculated through tomato (Fagoaga et al., 1995; Gandía et al., 2007). In the host-viroid systems studied in this work, differences in the genetic diversity were observed between not only host species but also two different trees of the same species. Furthermore, an association between population diversity and host response to infection has been found.

The fact that the highest viroid nucleotide diversity was observed in the tolerant citrus host (sour orange) could be explained by three non-mutually exclusive reasons: (i) differences in the strength of defence responses by the host, (ii) differences in systemic movement and accumulation and (iii) host-mediated differences in mutation rate. In the following paragraphs we will comment on each of these putative mechanisms. Differences in host defence mechanisms could impose different strengths of diversifying selection that may result in higher or lower accumulation of mutant haplotypes. In sensitive hosts (citrons and trifoliate oranges) defence mechanisms may allow the viroid to induce symptoms as a population characterized by a dominant haplotype. For instance, viroids are thought to resist the plant RNA silencing defence strategy by adopting highly packed secondary structures (Wang et al., 2004; Gómez & Pallàs, 2007). Indeed, it has been determined that the viroid RNA itself serves as a substrate for DICER-like cleavage (Itaya et al., 2007; Martín et al., 2007; Gómez & Pallàs, 2007; Carbonell et al., 2008). The heterogeneity of these CEVd populations, composed of different haplotypes adopting slightly different secondary structures, could aim to confer a certain level of resistance towards the plant RNA silencing machinery. In fact, it has been recently demonstrated for turnip mosaic virus that substitutions in the target of artificial micro RNAs allow the virus to escape from RNA silencing (Lin et al., 2009). Therefore, the fact that the P domain appears to be the most variable one in the sour orange populations suggests that it is involved in the viroid defence process. Supporting
this possibility, when small RNAs (sRNAs) from both PSTVd (from the same genus as CEVd) (Itaya et al., 2007) and CEVd (Martin et al., 2007) were mapped into the viroid molecule, it was found that those derived from the P domain were the less abundant ones, and thus probably more resistant to the DICER-like cleavage. If one assumes that sRNAs derived from the P domain are involved in symptom expression, and that a positive correlation between the accumulation levels of sRNAs and symptom severity exists (Itaya et al., 2001), then the resistance of CEVd sour orange populations in this domain towards DICER machinery might contribute to the absence of symptoms in this host. At present, viroid silencing mechanisms, including their direct implication in viroid pathogenicity, are not fully understood (Ding & Itaya, 2007). Additional experiments analysing the effect of the host on the biological properties of specific sRNAs could provide insights into this matter.

A second plausible mechanism for the larger genetic variability observed in sour orange trees is the possible differences in systemic movement and, hence, in viroid accumulation. Sour orange was the only host in which the T_R domain of CEVd populations remained fully conserved. This T_R domain is believed to be involved in long-distance viroid transport (Hammond, 1994; Maniataki et al., 2003) mediated by the interaction of an ‘RY’ motif formed in this domain with a phloematic host protein (Gozmanova et al., 2003). When the CEVd concentration was determined in the trees described in this experiment, a higher accumulation was observed in the trifoliate orange leaves than in the sour orange ones. However, the concentration was similar in the stems of both plant types (data not shown). It has been suggested previously that PSTVd could have a motif mediating trafficking (Ding et al., 1997) and that the phloem may have a factor able to recognize and traffic PSTVd into selective sink organs. Furthermore, phloem entry and exit appear to be differently regulated (Zhu et al., 2001, 2002). Further work with PSTVd mutants defined a bipartite trafficking motif, one part formed in the T_R domain (U201) and the other in the canonical P domain (U309 and U47/A313). This motif would unidirectionally mediate the exit from bundle sheath (phloem) to mesophyll in young tobacco leaves (Qi et al., 2004). Hence, the lack of plasticity of the T_R domain together with the high variability of the P domain in sour orange populations could be related to the difficulty of CEVd transport from stems to leaves. Alternatively, dominant CEVd variants in sensitive hosts could be more fitted to get into mesophyll cells and to promote the spread of infection.

Finally, a third plausible reason for the larger genetic variability observed in the sour orange is that different hosts impose different mutation rates on their RNA pathogens. This host effect on mutation rate has recently been described for cucumber mosaic virus populations infecting pepper and tobacco, with mutation rates in the pepper being one order of magnitude larger than in tobacco (Pita et al., 2007). Whether a similar situation may exist for viroids is something that cannot be answered with the currently available data.

Concerning the mutations observed in CEVd populations, the five nucleotide positions particularly prone to changes were all located within the P domain, except one which was situated in the upper strand of the V domain (Figs 2 and 3). The mutations +A62, +G74 and A130Δ were basically identified in the principal haplotypes of the S1, S2 and T2 populations. Additionally, mutations +G74 and A313G were both detected in many haplotypes from almost all CEVd populations. The A300U/C/A changes were observed only in some haplotypes of the S1 and T1 populations. Similarly, a previous study concerning naturally occurring PSTVd isolates indicated that the P domain accumulated the majority of neutral mutations (Owens et al., 2003). Many studies have showed that sequence variability occurs at specific positions of the viroid molecule (Keese & Symons, 1985; Góra-Sochacka et al., 2001; Ambroś et al., 1998; Owens et al., 2003). However, in spite of this flexibility which allows the pathogen to adapt to different hosts and environmental changes, a sufficient degree of conservation is also maintained (Tabler & Tsagris, 2004). Thus, it is not surprising that the mutations observed in this study were not randomly distributed throughout the CEVd genome (Fig. 3).

We evaluated the contribution of each structural domain to CEVd population differentiation through hosts and we found that the T_L, V and P domains were involved, although it is important to highlight that the P domain was the only domain always implicated in cases of reversion to the initial Ci population that occurred in the second transmission event. The P domain appeared to be the most polymorphic one. This observation was in partial concordance with a previous observation that defined the V and the P regions as the first and second most variable regions of the viroid molecule, respectively (Keese & Symons, 1985). This domain linked to viroid pathogenicity (Visvader et al., 1982; Visvader & Symons, 1985; Góra et al., 1996; Skoric et al., 2001) could also be involved in CEVd host-driven adaptation. In addition, since the P domain includes conserved sites among most of the members of the family Pospiviroidae, such as an A-rich and a U-rich region of the rod-like structure (Elena et al., 2001), its possible implication in essential biological processes cannot be ruled out. In fact, it has been reported that two PSTVd mutations in the P domain, NBU47A and NBA313U, tested in tobacco BY2 protoplasts resulted in defective replication (Qi et al., 2004). Furthermore, one of the critical loops for PSTVd trafficking has also been mapped to this domain (Zhong et al., 2008). Unlike the T_L and T_R domains, the P domain has not been associated with the contribution of the origin of new viroids (divergence) by molecular recombinant or rearrangement processes (Haseloff et al., 1982; Diener, 1983; Keese & Symons, 1985; Hammond et al., 1989; Szychowski et al., 2005; Daróς et al., 2006). Therefore, the functional role of
the P domain in evolution could be restricted to intra-host differentiation.

ACKNOWLEDGEMENTS

L. B. and N.D.V. were supported by grants RTA01-119, AGL2005-01468 and AGL2008-01491 from the Spanish MICINN. L. B. received a fellowship of the Conselleria de Agricultura – IVIA. S. F. E. was supported by grant BFU2006-14819-C02-01/BMC from the Spanish MICINN. We thank Luis Rubio for helpful comments.

REFERENCES


