Effect of citrus hosts on the generation, maintenance and evolutionary fate of genetic variability of *Citrus exocortis viroid* (CEVd)

Lucía Bernad[1], Núria Duran-Vila[1*], Santiago F. Elena[2,3]

1. Centro de Protección Vegetal y Biotecnología. Instituto Valenciano de Investigaciones Agrarias, Moncada, 46113 València, Spain.

2. Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-UPV, Campus UPV CPI 8E, 46022 València, Spain.

3. The Santa Fe Institute, Santa Fe, NM 87501, USA.

*To whom correspondence should be addressed. E-mail: nduran@ivia.es

Keywords: experimental evolution; genetic diversity; host-driven evolution; plant diseases; population genetics; selective pressures; viroid evolution

Running title: Host-driven evolution of CEVd populations

No. of words in the Summary: 192

No. of words in the main text: 5336

No. of figures and tables: 4

SUMMARY

Citrus exocortis viroid (CEVd) populations are composed by 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 and 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 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 characteristic of the original Etrog citron host and (*ii*) the genetic diversity found after transmitting back to Etrog citron is undistinguishable from that characteristic of the ancestral Etrog citron population regardless the citrus plant from where 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 for any protein but endowed with autonomous replication in their host plants. Since their origin, viroids very likely have been co-evolving with their hosts 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 region. Second, the *Pospiviroidae* rely on cell factors to process their multimeric intermediates of replication into unit-length molecules while the *Avsunviroidae* present hammerhead ribozyme structures able to self-cleave the multimeric forms. Third, the *Pospiviroidae* replicate in the nucleus whereas the *Avsunviroidae* replicate in the chloroplast (reviewed in Flores *et al*., 2005).

Like most RNA and some DNA viruses, viroids replicate within their hosts as polymorphic populations composed by closely related sequence variants generally distributed around a predominant one. This polymorphic population structure arises as the 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

65 selective pressures imposed by the different host species. Surveys of diversity
66 have been performed for different viroid species, including pospiviroids such
67 as *Citrus exocortis viroid* (CEVd) (Visvader & Symons, 1985; Gandía *et al.*, 2005;
68 Gandía *et al.*, 2007), *Citrus dwarfing viroid* (formerly *Citrus viroid III*) (Owens *et al.*,
69 2000), *Citrus bent leaf viroid* (Foissac & Duran-Vila, 2000; Gandía & Duran-Vila,
70 2004), *Potato spindle tuber viroid* (PSTVd) (Góra *et al.*, 1994; Gruner *et al.*, 1995;
71 Góra-Sochacka *et al.*, 1997), *Hop stunt viroid* (HSVd) (Kofalvi *et al.*, 1997; Palacio-
72 Bielsa *et al.*, 2004), and *Grapevine yellow speckle viroid 1* (Rigden & Rezaian, 1993;
73 Polivka *et al.*, 1996), as well as avsunviroids like *Chrysanthemum chlorotic mottle*
74 *viroid* (Navarro & Flores, 1997; Codoñer *et al.*, 2006), *Peach latent mosaic viroid*
75 ((Hernández & Flores, 1992; Ambrós *et al.*, 1998; Ambrós *et al.*, 1999), and
76 *Avocado sunblotch viroid* (Rakowski & Symons, 1989).

77 CEVd is a member of the genus *Pospiviroid* within the family *Pospiviroidae*. Like
78 other members of this family, the highly base paired rod-like secondary
79 structure conforms to the model of five structural domains proposed by Keese
80 & Symons (1985): terminal left (T_L), pathogenicity (P), central (C), variable (V),
81 and terminal right (T_R). The genus *Pospiviroid*, in addition to its characteristic
82 CCR, presents another conserved region in the T_L domain named terminal
83 conserved region (TCR). CEVd is the causal agent of the exocortis disease
84 characterized by a bark shelling or scaling disorder of trifoliate orange (*Poncirus*
85 *trifoliata* (L.) Raf.) used as rootstock (Fawcett & Klotz, 1948). The citrange
86 hybrids (*Citrus sinensis* (L.) × *P. trifoliata*) and Rangpur lime (*Citrus limonia* Osb.),
87 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 that characteristic of the original Etrog citron. 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 undistinguishable from that characteristic of the original Etrog citron. All together, these results support the notion that the composition and structure of viroid populations is determined by the host where they replicate.

METHODS

Plant material and viroids sources. A CEVd isolate (CEV-117) maintained in the sensitive selection 861-S1 of Etrog citron grafted onto rough lemon (*Citrus jambhiri* Lush.) rootstock was graft-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 on inoculum for graft-transmission to two new Etrog citron plants (Fig. 1). The inoculated citrons that were maintained in a greenhouse at 28 - 32 ^{o}C 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 of bark and leaves) of citron, trifoliate orange and sour orange were homogenized inside sealed plastic bags containing 5 mL of 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 ^{o}C for 20 min) and potassium acetate (on ice for 20 min). The soluble fraction was concentrated by ethanol precipitation and resuspended in 40 μL of sterile water.

First–strand viroid cDNA was synthesized with 15 U ThermoScript^{TM} Rnase H^{-} reverse transcriptase (ThermoScript-RT, Invitrogen) using the reverse specific primer complementary to the upper CCR strand, CEV-RT (5'-CTTCCTCCAGGTTTCCCCGGGGATCCC-3') (0.75 μM) and dNTPs (1 mM each). The ThermoScript-RT reaction buffer contained 50 mM Tris-acetate (pH 8.4), 75 mM potassium acetate, 8 mM magnesium acetate and 40 U of RNase Out (Invitrogen). The reaction mixture (20 μL final volume) was incubated at 60

134 °C for 1 h. Second-strand DNA synthesis and PCR amplification (50 μL final
135 volume) were performed using 4 μL of the first-strand mixture, 3.5 U Expand
136 High Fidelity PCR System (Roche), the reverse and forward primers CEV-R1
137 (5′-CCGGGGATCCCTGAAGGA-3′) and CEVd-F1 (5′-
138 GGAAACCTGGAGGAAGTCG-3′) (0.5 μM each) and dNTPs (0.12 mM each) in
139 a buffer containing 150 mM $MgCl_2$. PCR parameters consisted of a denaturation
140 step at 94 °C for 5 min, followed by 35 cycles (94 °C for 30 s, 60 °C for 30 s and 72
141 °C for 1 min) (Bernad & Duran-Vila, 2006).

142 Electrophoretic analysis in 2% agarose gels confirmed the synthesis of a DNA
143 product of the expected size. The RT-PCR products were purified (Amersham
144 Kit) and ligated into the pGEM-T vector and used to transform *Escherichia coli*
145 DH5 competent cells. Transformants were grown for approximately 20 h at 37
146 °C on ampicillin-containing plates, and 30 colonies were randomly selected for
147 sequencing.

148 To verify the sequence of the region of the upper strand corresponding to the
149 primers used for RT-PCR, a second RT-PCR reaction was performed under the
150 same conditions described above but using a reverse specific primer
151 complementary to the lower CCR strand, CEV-RT2 (5′–
152 CCGGGTAGTATCCAGAGAGAAGCTCCG-3′) and the reverse and forward
153 primers CEV-R2 (5′-GGGTAGTCTCCAGAGAGAAG-3′) and CEV-F2 (5′-
154 GGTGGAAACAACTGAAGCTT-3′).

155 **Sequence analysis.** Cloned full-length viroid cDNAs were sequenced with an
156 ABI PRISM DNA analyzer 377 (Perkin-Elmer). Chromatograms were edited

with Chromas v.1.43. Multiple-sequences alignments were generated with the ClustalW program (Higgins & Sharp, 1994). 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 www.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi server, choosing the option of circular RNA molecules and were drawn with the RNAviz 2.0 (De Rijk *et al.*, 2003) software.

Nucleotide diversity calculation and assessment of differences among host trees. As a measure of CEVd genetic diversity present on each analyzed 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 $H_j = -\sum_{i=1}^{M} P_i log_2 P_i$, 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 can then be estimated as the sum of entropies from all sites in the genome $H = \sum_{j=1}^{L} H_j$, where L is the length of the sequence alignment.

RESULTS

178 **Molecular characterization and genetic diversity of CEVd populations**

179 The CEVd isolate (CEV-117) maintained in citron (Ci) was used as source of
180 inoculum for transmissions to two trifoliate orange seedlings (T1, T2) and two
181 sour orange seedlings (S1, S2) trees (Fig. 1a). The consensus sequence was
182 determined by directly sequencing two full-length RT-PCR amplicons of CEVd
183 obtained using two different sets of primers. In order to minimize the
184 introduction of artifactual changes, PCR amplification was performed using a
185 DNA polymerase with proofreading activity. The consensus sequence showed
186 98% nucleotide identity with the reference sequence CEVd (class A) (Visvader
187 *et al.*, 1982) and the predicted MFESS was a highly base-paired rod-like
188 secondary structure characteristic of viroids of the family *Pospiviroidae* (Fig. 2).

189 Ten years after inoculation, the infected trifoliate orange seedling trees (T1, T2)
190 were severely stunted and presented the bark scaling symptoms characteristic
191 of the exocortis disease, whereas the infected sour orange seedling trees (S1, S2)
192 remained symptomless. The consensus CEVd sequences recovered from T1 and
193 T2 were found to differ by only one change (T1: A130Δ, T2: A313G) relative to
194 the consensus CEVd sequence of the inoculum source. By contrast, the
195 consensus CEVd sequences recovered from S1 and S2 presented multiple peaks
196 at different chromatogram positions, suggesting that the population of
197 sequence variants was genetically heterogeneous. These results, although
198 preliminary, indicated that the CEVd populations in these hosts were not
199 identical nor even similar. However, transmissions from each of the two
200 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 which 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 on 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 9 populations are summarized in Table 1.

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

First, the ancestral CEVd population derived from the initial citron source of inoculum (Ci) was constituted by 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

223 was constituted by haplotypes with frequencies between 3.3 and 6.6% (Fig. 1b).
224 The Shannon entropy of the ancestral population was 0.0051.

225 Second, the CEVd populations isolated from the trifoliate oranges (T1 and T2),
226 were both characterized by the presence of a numerically dominant haplotype
227 (T1-1 and T2-1, respectively), representing 63.3% and 60.0% of each population,
228 respectively. However, these two haplotypes were not identical among them
229 nor to the dominant ancestral Ci-1 haplotype. Interestingly, all haplotypes from
230 T1 shared a deletion, relative to the sequence of Ci-1, in the upper strand of the
231 V domain (–A in position 130). Similarly, most haplotypes from plant T2 shared
232 a nucleotide substitution (A313G) in the lower strand of the P domain. None of
233 these two characteristic mutations was detected in the ancestral CEVd
234 population used as inoculum. It is worth noting that the second most abundant
235 ancestral haplotype, Ci-2 has been also found in population T2, although its
236 frequency was reduced to 3.3%. The Shannon entropy for these two populations
237 was very similar (Table 1) and also in the same range that that obtained for the
238 ancestral Etrog citron population.

239 Third, CEVd populations replicating within sour oranges (S1 and S2) were
240 constituted by a heterogeneous assemble of haplotypes, with no clear dominant
241 one. Hence, the most abundant haplotypes in plant S1 were S1-4 and S1-12,
242 both at frequency 20%; whereas the most frequent haplotype in plant S2, S2-11,
243 was only at 16.6%. Remarkably, S1-12 and S2-11 haplotypes were identical to
244 the minority haplotype Ci-8 present at 3.3% in the ancestral Ci population.
245 Similarly, haplotype S1-2 was identical to haplotype Ci-5 present in the

246 ancestral Ci population at half frequency (3.3% in Ci versus 6.6% in S1). The
247 numerically dominant haplotypes in the ancestral Ci populations, Ci-1 and Ci-2,
248 are also present in populations S1 and S2, although their respective frequencies
249 were lower than in the ancestral population (Fig. 1b).

250 To quantitatively explore the above data obtained for the new two hosts, a
251 nested ANOVA model was fitted to the per site Shannon entropy values. In this
252 model, host species and replicate plants were treated as random factors and
253 plant replicates were nested within host species. Differences among sites in the
254 alignment were used to evaluate the within-plant variability. The test found no
255 significant differences among replicate plants within hosts ($F_{2,1500}$ = 0.345, P =
256 0.709), but the diversity levels in sour orange were 139% larger than those in
257 trifoliate orange, a difference which was statistically significant ($F_{2,2}$ = 65.178, P
258 = 0.015) (Table 1).

259 It is worth noting that the mutation frequencies and entropy estimations may be
260 somehow inflated by the intrinsic error associated to the RT-PCR amplification
261 reaction, as the fidelity of PCR varies depending on reaction conditions and the
262 nature of target sequences (Cha & Thilly, 1993). However, Theycheney *et al*.
263 (2005) showed that, in general, RT-PCR produced less than 0.5% errors. In
264 addition, a previous work addressed to characterize CEVd and HSVd diversity
265 indicated that the observed mutations were unlikely due to PCR errors (Palacio
266 & Duran-Vila, 1999; Palacio-Bilesa *et al*., 2004). In fact, some of the
267 polymorphisms here discussed, such as the most frequent haplotypes (i.e., Ci-1
268 and Ci-2), the minority ones (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 pervasively occurring in independent PCR reactions. Furthermore, since we are interested in the relative value of diversity measures across hosts, and all our samples were treated identically, any bias induced 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 Etrog citron host

In the second transmission event, back to the ancestral Etrog citron host, population composition dramatically changed, basically reverting to the same situation described for the ancestral Ci tree. In the new populations 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 they were only beneficial in the trifoliate host but deleterious in the Etrog one. However, the A313G substitution characteristic of the T2 population was identified not only in three minority haplotypes in the CT2 population but also in 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, 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 Etrog

292 citron-characteristic haplotypes Ci-1 and Ci-2 were both found at high
293 frequency in the new populations.

294 **The P domain appears as the most polymorphic region of the CEVd molecule**

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

305 **Conservation of structural motifs**

306 Conserved sequence motives and functional conformations typical of viroids
307 belonging to the family *Pospiviroidae* were also examined. These motifs are the
308 TCR present only in some genera of this family and whose function remains
309 still unknown (Flores *et al.*, 1997), hairpins I and II (HP I and HP II) probably
310 involved in replication processes (Ding & Itaya, 2007), and the recently
311 identified "RY" motif located in the T_R domain which is thought to be involved
312 in long-distance viroid transport within plant (Gozmanova *et al.*, 2003).

313 The HP I, which comprises the segment of the viroid molecule used to design
314 the primers-set used for CEVd amplification, was not analyzed. Only one
315 sequence out of the 270 genomes sequenced contained a mutation (C240U) in
316 the palindrome sequences forming the HP II (Fig. 2). This mutation would
317 prevent the formation of the canonical pair C:G but still would allow forming a
318 non-canonical U:G pair, likely maintaining the right folded conformation of
319 hairpin HP II. In the TCR motif (Fig. 2), two changes, +U19 and G26A, were
320 observed in two of the less frequent haplotypes, respectively. However, our
321 data do not allow us to conclude whether these two changes are selectively
322 important or not. Finally, the sequences forming the "RY" motif (Fig. 2)
323 remained conserved as well as its predicted minimum free energy secondary
324 structure in all CEVd populations (data not shown).

325

326 **DISCUSSION**

327 Here we present the characterization of the genetic structure and evolution of
328 CEVd populations in both natural and indicator citrus hosts. CEVd populations,
329 as other viroids and most RNA viruses, are composed by closely related
330 haplotypes whose frequencies in the population result from the equilibrium
331 between mutation, selection and genetic drift. This population heterogeneity is
332 an intrinsic property of RNA replicons that has been broadly studied for many
333 RNA plant viruses (García-Arenal *et al.*, 2003). The characterization of CEVd
334 populations replicating for over ten years under field conditions in two natural
335 citrus hosts of economical importance has allowed us to study with

unprecedented detail the mechanisms of CEVd diversification as it adapts to different hosts.

On the 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, here measured as Shannon entropy, generated and maintained in this sensitive host was undistinguishable 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 which, indeed, were also different from that 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 IIa 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 with what was observed in trifoliate orange trees, these populations were much

359 more variable than the ancestral population isolated from citron. Furthermore,
360 these populations were not characterized by one or two majority haplotypes, as
361 it was the case for the citron and trifoliate orange trees, but by a much more
362 complex genetic mixture, although some haplotypes (e.g., Ci-8) were shared by
363 the two trees and were also present in the ancestral citron tree. The fact that Ci-8
364 was present in the ancestral population (Ci) as a minority haplotype would
365 provide to CEVd populations a host adaptation advantage. Ci-8 differed from
366 the dominant haplotype Ci-1 (Ci) in two insertions +A62 and +G74, both
367 located in the upper strand of the P domain, leading to a more relaxed
368 minimum free energy secondary structure (data not shown) which could
369 modify (facilitate or hinder) the interaction with unknown host-factors.

370 In both cases, after inoculating the various CEVd populations back into the
371 original Etrog citron host, both the amount of genetic diversity as well as the
372 haplotypes constitution (including the dominant one Ci-1) reversed to a
373 configuration that clearly reflected the ancestral population. These results
374 further support the conclusion of host-driven adaptation and, furthermore,
375 suggest that the evolvability of populations was not constrained by the actual
376 host. Cases of host-driven adaptation have been widely reported for both plant
377 RNA viruses and viroids (García-Arenal *et al.*, 2003). However, cases have also
378 been described in which certain citrus viruses such as *Citrus leaf blotch virus* and
379 *Citrus tristeza virus* did not respond in a specific way to different citrus host
380 species (Vives *et al.*, 2002). Viroid populations undergo bottlenecks upon
381 transmission to different hosts. In a previous study heterogeneous CEVd
382 population infecting a symptomless broad bean plant evolved to a more

383 homogeneous CEVd population after being inoculated through tomato
384 (Fagoaga *et al.*, 1995; Gandía *et al.*, 2007). In the host-viroid systems studied in
385 this work, differences in the genetic diversity were observed not only between
386 host species but also between two different trees of the same species.
387 Furthermore, an association between population diversity and host response to
388 infection has been found.

389 The fact that the highest viroid nucleotide diversity was observed in the
390 tolerant citrus host (sour orange) could be explained by three non-mutually
391 exclusive reasons: (*i*) differences in the strength of defense responses by the host,
392 (*ii*) differences in systemic movement and accumulation and (*iii*) host-mediated
393 differences in mutation rate. In the following paragraphs we will comment on
394 each of these putative mechanisms. Differences in host defense mechanisms
395 could impose different strengths of diversifying selection that may result in
396 higher or lower accumulation of mutant haplotypes. In sensitive hosts (citrons
397 and trifoliate oranges) defense mechanisms may allow the viroid to induce
398 symptoms as a population characterized by a dominant haplotype. For instance,
399 concerning the plant RNA silencing defense strategy, viroids are thought to
400 resist it by adopting highly packed secondary structures (Wang *et al.*, 2004;
401 Gómez & Pallás, 2007). Indeed, it has been determined that the viroid RNA
402 itself serves as a substrate for DICER-like cleavage (Itaya *et al.*, 2007; Martín *et*
403 *al.*, 2007; Gómez & Pallás, 2007; Carbonell *et al.*, 2008). The heterogeneity of
404 these CEVd populations, composed by different haplotypes adopting slightly
405 different secondary structures, could aim to confer certain level of resistance
406 towards the plant RNA silencing machinery. In fact, it has been recently

407 demonstrated for *Turnip mosaic virus* that substitutions in the target of artificial
408 miRNAs allow the virus to escape from RNA silencing (Lin *et al*., 2009).
409 Furthermore, taking into consideration that in the sour orange populations, the
410 P domain appears to be the most variable one, suggests its implication in the
411 viroid defense process. Supporting this possibility, when small RNAs (sRNAs)
412 from both PSTVd (from the same genus as CEVd) (Itaya *et al*., 2007) and CEVd
413 (Martín *et al*., 2007) were mapped into the viroid molecule, it was found that
414 those derived from the P domain were the less abundant ones, thus probably
415 more resistant to the DICER-like cleavage. If one assumes that sRNAs derived
416 from the P domain are involved in symptom expression, and that a positive
417 correlation between the accumulation levels of sRNAs and symptom severity
418 exists (Itaya *et al*., 2001), then the resistance of CEVd sour orange populations in
419 this domain towards DICER machinery might contribute to the absence of
420 symptoms in this host. At present, viroid silencing mechanisms including its
421 direct implication in viroid pathogenicity are not fully understood (Ding &
422 Itaya, 2007). Additional experiments analyzing the effect of the host in the
423 biological properties of specific sRNAs could provide insights in this sense.

424 A second plausible mechanism for the larger genetic variability observed in
425 sour orange trees is possible differences in systemic movement and, henceforth,
426 in viroid accumulation. Sour orange was the only host in which the T_R domain
427 of CEVd populations remained fully conserved. This T_R domain is believed to
428 be involved in long-distance viroid transport (Hammond, 1994; Maniataki *et al*.,
429 2003) mediated by the interaction of an "RY" motif formed in this domain with
430 a phloematic host protein (Gozmanova *et al*., 2003). When CEVd concentration

431 was determined in the trees described in this experiment, a higher
432 accumulation was observed in the trifoliate orange leaves than in the sour
433 orange ones. However, the concentration was similar in the stems of both plant
434 types (data not shown). It has been previously suggested that PSTVd could
435 have a motif mediating trafficking (Ding *et al.*, 1997) and that the phloem may
436 have a factor able of recognizing and traffic PSTVd into selective sink organs.
437 Furthermore, phloem entry and exit appears to be differently regulated (Zhu *et*
438 *al.*, 2001; Zhu *et al.*, 2002). Further work with PSTVd mutants defined a bipartite
439 trafficking motif, one part formed in the T_R domain (U201) and the other in the
440 canonical P domain (U309 and U47/A313). This motif would mediate
441 unidirectionally the exit from bundle sheath (phloem) to mesophyll in young
442 tobacco leaves (Qi *et al.*, 2004). Hence, the lack of plasticity of the T_R domain
443 together with the high variability of the P domain in sour orange populations
444 could be related to the difficulty of CEVd transport from stems to leaves. And
445 in the other sense, dominant CEVd variants in sensitive hosts could be more
446 fitted to get into mesophyll cells and to promote the spreading of the infection.

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

454 Concerning the mutations observed in CEVd populations, the five nucleotide
455 positions particularly prone to changes were all located within the P domain
456 except one which was situated in the upper strand of the V domain (Fig. 2 and
457 Fig.3). The mutations +A62, +G74 and A130□ were basically identified in the
458 principal haplotypes of the S1, S2 and T2 populations. Additionally, mutations
459 +G74 and A313G were both detected in many haplotypes from almost all CEVd
460 populations. The A300U/C/□ changes were observed only in some haplotypes
461 of the S1 and T1 populations. Similarly, a previous study concerning naturally
462 occurring PSTVd isolates indicated that the P domain accumulated the majority
463 of neutral mutations (Owens *et al.*, 2003). Many works have showed that
464 sequence variability occurred at specific positions of the viroid molecule (Keese
465 & Symons, 1985; Góra-Sochacka *et al.*, 2001; Ambrós *et* al., 1998; Owens *et al.*,
466 2003). However, in spite of this flexibility which allows the pathogen to adapt to
467 different hosts and environmental changes, a sufficient degree of conservation
468 is also maintained (Tabler & Tsagris, 2004). Thus, it is not surprisingly that the
469 mutations observed in this study were not randomly distributed throughout the
470 CEVd genome (Fig. 3).

471 We evaluated the contribution of each structural domain to CEVd population
472 differentiation through hosts and we found that the T_L, V and P domains were
473 involved, although it is important to highlight that the P domain was the only
474 domain always implicated in cases of reversion to the initial Ci population
475 occurred in the second transmission event. The P domain appeared as the most
476 polymorphic one. This observation was in partial concordance to a previous
477 observation that defined the V, first, and the P, second, as the most variables

478 regions of the viroid molecule (Keese & Symons, 1985). This domain linked to
479 viroid pathogenicity (Visvader *et al.*, 1982; Visvader & Symons, 1985; Góra *et al.*,
480 1996; Skoric *et al.*, 2001) could also be involved in CEVd host-driven adaptation.
481 In addition, since the P domain includes conserved sites among most of the
482 members of the family *Pospiviroidae*, such as an A-rich and a U-rich regions of
483 the rod-like structure (Elena *et al.*, 2001), it cannot be ruled out its possible
484 implication in essential biological processes. In fact, it has been reported that
485 two PSTVd mutants in the P domain, NBU47A and NBA313U, tested in tobacco
486 BY2 protoplasts resulted defective in replication (Qi *et al.*, 2004). Furthermore,
487 one of the critical loops for PSTVd trafficking has also been mapped in this
488 domain (Zhong *et al.*, 2008). Unlike the T_L and T_R domains, the P domain has
489 not been associated to the contribution of the origin of new viroids (divergence)
490 by molecular recombinant or rearrangement processes (Haseloff *et al.*, 1982;
491 Diener, 1983; Keese & Symons, 1985; Hammond *et al.*, 1989; Szychowski *et al.*,
492 2005; Daròs *et al.*, 2006). Therefore, the functional role of the P domain in
493 evolution could be restricted to intra-host differentiation.

494

495 **ACKNOWLEDGEMENTS**

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

501

502 **REFERENCES**

503 **Ambrós, S., Hernández, C., Desvignes, J.C. & Flores, R. (1998).** Genomic
504 structure of three phenotypically different isolates of Peach latent mosaic viroid:
505 implications of the existence of constraints limiting the heterogeneity of viroid
506 quasispecies. *J Virol* **72**, 7397-7406.

507 **Ambrós, S., Hernández, C. & Flores, R. (1999).** Rapid generation of genetic
508 heterogeneity in progenies from individual cDNA clones of Peach latent mosaic
509 viroid in its natural host. *J Gen Virol* **80**, 2239-2252.

510 **Bernad, L. & Duran-Vila, N. (2006).** A novel RT-PCR approach for detection
511 and characterization of citrus viroids. *Mol Cell Probes* **20**, 105-113.

512 **Carbonell, A., Martínez de Alba, A.E., Flores, R. & Gago, S. (2008).** Double-
513 stranded RNA interferes in a sequence-specific manner with the infection of
514 representative members of the two viroid families. *Virology* **371**, 44-53.

515 **Cha, R.S. & Thilly, W.G. (1993).** Specificity, efficiency, and fidelity of PCR.
516 *Genome Res* **3**, 18-29.

517 **Codoñer, F.M., Daròs, J.A., Solé, R.V. & Elena, S.F. (2006).** The fittest versus
518 the flattest: experimental confirmation of the quasispecies effect with subviral
519 pathogens. *PLoS Pathog* **2**, e136.

520 **Daròs, J.A., Elena, S.F. & Flores, R. (2006).** Viroids: an Ariadne's thread into the
521 RNA labyrinth. *EMBO Rep* **7**, 593-598.

522 **De Rijk, P., Wuyts, J. & De Wachter, R. (2003).** RnaViz 2: an improved
523 representation of RNA secondary structure. *Bioinformatics* **19**, 299-300.

Diener, T.O. (1983). The viroid-a subviral pathogen. *Am Sci* **71**, 481-489.

Diener, T.O. (1995). Origin and evolution of viroids and viroid-like satellite RNAs. *Virus Genes* **11**, 119-131.

Ding, B., Kwon, M.O., Hammond, R. & Owens, R. (1997). Cell-to-cell movement of potato spindle tuber viroid. *Plant J* **12**, 931-936.

Ding, B. & Itaya, A. (2007). Viroid: a useful model for studying the basic principles of infection and RNA biology. *Mol Plant Microbe Interact* **20**, 7-20.

Domingo, E., Martin, V., Perales, C., Grande-Pérez, A., García-Arriaza, J. & **Arias, A. (2006).** Viruses as quasispecies: biological implications. *Curr Top Microbiol Immunol* **299**, 51-82.

Elena, S.F., Dopazo, J., Flores, R., Diener, T.O. & Moya, A. (1991). Phylogeny of viroids, viroidlike satellite RNAs, and the viroidlike domain of hepatitis δ virus. *Proc Natl Acad Sci U S A* **88**, 5631-5634.

Elena, S.F., Dopazo, J., De la Peña, M., Flores, R., Diener, T.O. & Moya, A. (2001). Phylogenetic analysis of viroid and viroid-like satellite RNAs from plants: a reassessment. *J Mol Evol* **53**, 155-159.

Fagoaga, C., Semancik, J.S. & Duran-Vila, N. (1995). A Citrus exocortis viroid variant from broad bean (*Vicia faba* L.): infectivity and pathogenesis. *J Gen Virol* **76**, 2271-2277.

Fawcett, H.S. & Klotz, L.J. (1948). Exocortis on trifoliate orange. *Citrus Leaves* **28**, 8.

Flores, R., Di Serio, F. & Hernández, C. (1997). Viroids: The noncoding genomes. *Semin Virol* **8**, 65-73.

Flores, R., Hernández, C., Martínez de Alba, A.E., Daròs, J.A. & Di Serio, F. (2005). Viroids and viroid-host interactions. *Annu Rev Phytopathol* **43**, 117-139.

Foissac, X. & Duran-Vila, N. (2004). Characterization of two citrus apscaviroids isolated in Spain. *Arch Virol* **145**, 1075-1983.

Gandía, M. & Duran-Vila, N. (2004). Variability of the progeny of a sequence variant Citrus bent leaf viroid (CBLVd). *Arch Virol* **149**, 407-416.

Gandía, M., Rubio, L., Palacio, A. & Duran-Vila, N. (2005). Genetic variation and population structure of an isolate of Citrus exocortis viroid (CEVd) and of the progenies of two infectious sequence variants. *Arch Virol* **150**, 1945-1957.

Gandía, M., Bernad, L., Rubio, L. & Duran-Vila, N. (2007). Host effect on the molecular and biological properties of a Citrus exocortis viroid isolate from *Vicia faba*. *Phytopathology* **97**, 1004-1010.

García-Arenal, F., Fraile, A. & Malpica, J.M. (2003). Variation and evolution of plant virus populations. *Int Microbiol* **6**, 225-232.

Gómez, G. & Pallás, V. (2007). Mature monomeric forms of Hop stunt viroid resist RNA silencing in transgenic plants. *Plant J* **51**, 1041-1049.

Góra, A., Candresse, T. & Zagorski, W. (1994). Analysis of the population-structure of 3 phenotypically different PSTVd isolates. *Arch Virol* **138**, 233-245.

Góra, A., Candresse, T. & Zagorski, W. (1996). Use of intramolecular chimeras to map molecular determinants of symptom severity of potato spindle tuber viroid (PSTVd). *Arch Virol* **141**, 2045-2055.

Góra-Sochacka, A., Kierzek, A., Candresse, T. & Zagorski, W. (1997). The genetic stability of potato spindle tuber viroid (PSTVd) molecular variants. *RNA* **3**, 68-74.

Góra-Sochacka, A., Candresse, T. & Zagorski, W. (2001). Genetic variability of potato spindle tuber viroid RNA replicon. *Acta Biochim Pol* **48**, 467-476.

Gozmanova, M., Denti, M.A., Minkov, I.N., Tsagris, M. & Tabler, M. (2003). Characterization of the RNA motif responsible for the specific interaction of potato spindle tuber viroid RNA (PSTVd) and the tomato protein Virp1. *Nucleic Acids Res* **31**, 5534-5543.

Gruner, R., Fels, A., Qu, F., Zimmat, R., Steger, G. & Riesner, D. (1995). Interdependence of pathogenicity and replicability with potato spindle tuber viroid. *Virology* **209**, 60-69.

Hammond, R., Smith, D.R. & Diener, T.O. (1989). Nucleotide sequence and proposed secondary structure of Columnea latent viroid: a natural mosaic of viroid sequences. *Nucleic Acids Res* **17**, 10083-10094.

Hammond, R.W. (1994). Agrobacterium-mediated inoculation of PSTVd cDNAs onto tomato reveals the biological effect of apparently lethal mutations. *Virology* **201**, 36-45.

Haseloff, J., Mohamed, N.A. & Symons, R.H. (1982). Viroid RNAs of cadang-cadang disease of coconuts. *Nature* **299**, 316-321.

Hernández, C. & Flores, R. (1992). Plus and minus RNAs of peach latent mosaic viroid self-cleave in vitro via hammerhead structures. *Proc Natl Acad Sci U S A* **89**, 3711-3715.

Higgins, P.G. & Sharp, P.M. (1994). Fast and sensitive multiple sequence alignment on a microcomputer. *Comput Appl Biosci* **5**, 151-153.

Itaya, A., Folimonov, A., Matsuda, Y., Nelson, R.S. & Ding, B. (2001). Potato spindle tuber as inducer of RNA silencing in infected tomato. *Mol Plant-Microbe Interact* **14**, 1332-1334.

Itaya, A., Zhong, X., Bundschuh, R., Qi, Y., Wang, Y., Takeda, R., Harris, A.R., Molina, C., Nelson, R.S. & Ding, B. (2007). A structured viroid RNA serves as a substrate for dicer-like cleavage to produce biologically active small RNAs but is resistant to RNA-induced silencing complex-mediated degradation. *J Virol* **81**, 2980-2994.

Keese, P. & Symons, R.H. (1985). Domains in viroids: evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. *Proc Natl Acad Sci U S A* **82**, 4582-4586.

Kofalvi, S.A., Marcos, J.F., Cañizares, M.C., Pallás, V. & Candresse, T. (1997). Hop stunt viroid (HSVd) sequence variants from Prunus species: evidence for recombination between HSVd isolates. *J Gen Virol* **78**, 3177-3186.

Korber, B.T., Kunstman, K.J., Patterson, B.K., Furtado, M., McEvilly, M.M., Levy, R. & Wolinsky S.M. (1994). Genetic differences between blood and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. *J Virol* **68**, 7467-7481.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinformatics* **5**, 150-163.

615 **Lin, S.S., Wu, H.W., Elena S.F., Chen K.C., Niu, Q.W., Yeh, S.D., Chen, C.C. &**
616 **Chua, N.H. (2009).** Molecular evolution of a viral non-coding sequence under
617 the selective pressure of amiRNA-mediated silencing. *PLoS Pathog* **5**, e1000312.
618 **Maniataki, E., Martínez de Alba, A.E., Sagesser, R., Tabler, M. & Tsagris, M.**
619 **(2003).** Viroid RNA systemic spread may depend on the interaction of a 71-
620 nucleotide bulged hairpin with the host protein VirP1. *RNA* **9**, 346-354.
621 **Martín, R., Arenas, C., Daròs, J.A., Covarrubias, A., Reyes, J.L. & Chua, N.H.**
622 **(2007).** Characterization of small RNAs derived from Citrus exocortis viroid
623 (CEVd) in infected tomato plants. *Virology* **367**, 135-146.
624 **Navarro, B. & Flores, R. (1997).** Chrysanthemum chlorotic mottle viroid:
625 unusual structural properties of a subgroup of self-cleaving viroids with
626 hammerhead ribozymes. *Proc Natl Acad Sci U S A* **94**, 11262-11267.
627 **Owens, R.A., Yang, G., Gundersen-Rindal, D., Hammond, R.W., Candresse, T.**
628 **& Bar-Joseph, M. (2000).** Both point mutation and RNA recombination
629 contribute to the sequence diversity of citrus viroid III. *Virus Genes* **20**, 243-252.
630 **Owens, R., Thompson, S.M. & Kramer, M. (2003).** Identification of neutral
631 mutants surrounding two naturally occurring variants of Potato spindle tube
632 viroid. *J Gen Virol* **84**, 751-756.
633 **Palacio, A. & Duran-Vila, N. (1999)**. Single-strand conformation polymorphism
634 (SSCP) analysis as a tool for viroid characterisation. *J Virol Methods* **77,** 27-36.
635 **Palacio-Bielsa, A., Romero-Durbán, J. & Duran-Vila, N. (2004).**
636 Characterization of citrus HSVd isolates. *Arch Virol* **149**, 537-552.

637 **Pita J.S., de Miranda, J.R., Schneider, W.L. & Roossinck, M.J. (2007).**
638 Environment determines fidelity for an RNA virus replicase. *J Virol* **81**, 9072-
639 9077.
640 **Polivka, H., Staub, U. & Gross, H.J. (1996).** Variation of viroid profiles in
641 individual grapevine plants: novel grapevine yellow speckle viroid 1 mutants
642 show alterations of hairpin I. *J Gen Virol* **77**, 155-161.
643 **Qi, Y., Pelissier, T., Itaya, A., Hunt, E., Wassenegger, M. & Ding, B. (2004).**
644 Direct role of a viroid RNA motif in mediating directional RNA trafficking
645 across a specific cellular boundary. *Plant Cell* **16**, 1741-1752.
646 **Rakowski, A.G. & Symons, R.H. (1989).** Comparative sequence studies of
647 variants of avocado sunblotch viroid. *Virology* **173**, 352-356.
648 **Rigden, J.E. & Rezaian, M.A. (1993).** Analysis of sequence variation in
649 grapevine yellow speckle viroid 1 reveals two distinct alternative structures for
650 the pathogenic domain. *Virology* **193**, 474-477.
651 **Shannon, C.E. (1948).** A mathematical theory of communication. *Bell System*
652 *Tech J* **27**, 379-423 & 623-656.
653 **Skoric, D., Conerly, M., Szychowski, J.A. & Semancik, J.S. (2001).** CEVd-
654 induced symptom modification as a response to a host-specific temperature-
655 sensitive reaction. *Virology* **280**, 115-123.
656 **Szychowski, J.A., Vidalakis, G. & Semancik, J.S. (2005).** Host-directed
657 processing of Citrus exocortis viroid. *J Gen Virol* **86**, 473-477.
658 **Tabler, M. & Tsagris M. (2004).** Viroids: petite RNA pathogens with
659 distinguished talents. *Trends Plant Sci* **9**, 339-348.

Teycheney, P.Y., Laboureau, N., Iskra-Caruana, M.L. & Candresse, T. (2005). High genetic variability and evidence for plant-to-plant transfer of *Banana mild mosaic virus*. *J Gen Virol* **86**, 3179-3187.

Visvader, J.E., Gould, A.R., Bruening, G.E. & Symons, R.H. (1982). Citrus exocortis viroid: nucleotide sequence and secondary structure of an Australian isolate. *FEBS Lett* **137**, 288-292.

Visvader, J.E. & Symons, R.H. (1985). 11 New sequence variants of Citrus exocortis viroid and the correlation of sequence with pathogenicity. *Nucleic Acids Res* **13**, 2907-2920.

Vives, M.C., Rubio, L., Galipienso, L., Navarro, L., Moreno, P. & Guerri, J. (2002). Low genetic variation between isolates of Citrus leaf blotch virus from different host species and of different geographical origins. *J Gen Virol* **83**, 2587-2591.

Wang, M.B., Bian, X.Y., Wu, L.M., Liu, L.X., Smith, N.A., Isenegger, D., Wu, R.M., Masuta, C., Vance, V.B. & other authors. (2004). On the role of RNA silencing in the pathogenicity and evolution of viroids and viral satellites. *Proc Natl Acad Sci U S A* **101**, 3275-3280.

Zhong, X., Archual, A.J., Amin, A.A. & Ding, B. (2008). A genomic map of viroid RNA motifs critical for replication and systemic trafficking. *Plant Cell* **20**, 35-47.

Zhu, Y., Green, L., Woo, Y.M., Owens, R. & Ding, B. (2001). Cellular basis of potato spindle tuber viroid systemic movement. *Virology* **279**, 69-77.

682 **Zhu, Y., Qi, Y., Xun, Y., Owens, R. & Ding, B. (2002).** Movement of potato
683 spindle tuber viroid reveals regulatory points of phloem-mediated RNA traffic.
684 *Plant Physiol* **130**, 138-146.
685 **Zuker, M. (2003).** Mfold web server for nucleic acid folding and hybridization
686 prediction. *Nucleic Acids Res* **31**, 3406-3415.
687

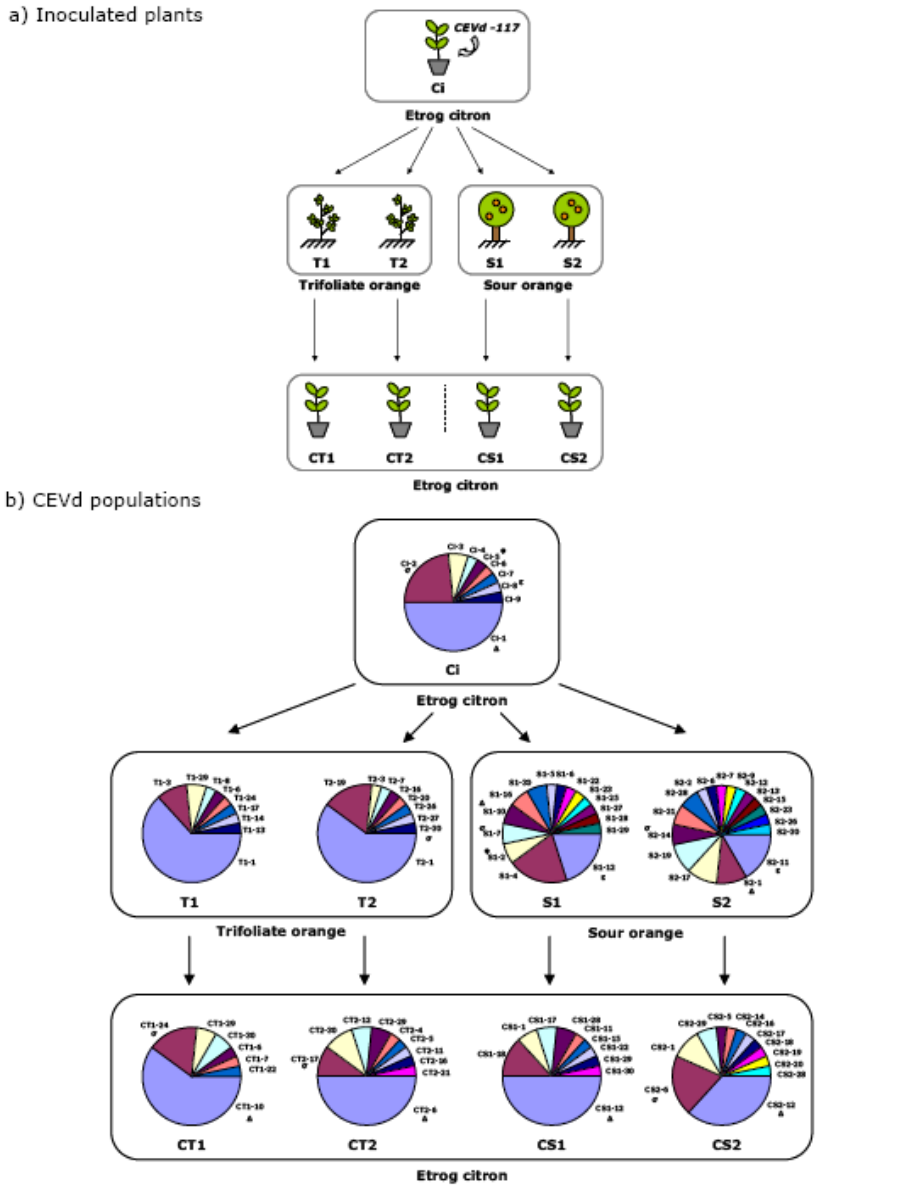
687 **FIGURE LEGENDS**

688

689 **Fig. 1.** a) CEVd transmissions scheme. The Etrog citron (Ci) infected with a
690 CEVd isolate (CEVd-117) was used as inoculum source for a first graft-
691 transmission to two trifoliate orange seedlings (T1 and T2) and two sour orange
692 seedlings (S1 and S2). After 10 years, the CEVd populations from the trifoliate
693 orange and sour orange trees (maintained under field conditions) were graft-
694 transmitted back to new Etrog citron plants (CT1, CT2, CS1 and CS2). b)
695 Diagrams of frequency showing the genetic structures (haplotypes and
696 frequencies) of CEVd populations retrieved from each infected tree. Shared
697 haplotypes among populations are indicated by the following letters: Δ (Ci-1), σ
698 (Ci-2), φ (Ci-5) and ϵ (Ci-8).

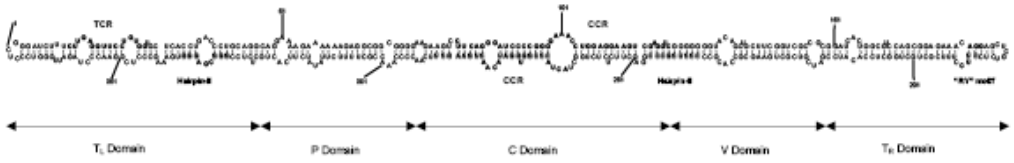
699

FIGURE 1


a) Inoculated plants
CEVd -117
Ci
Etrog citron
T1
T2
Trifoliate orange
S1
S2
Sour orange
CT1
CT2
CS1
CS2
Etrog citron
b) CEVd populations
Ci
Etrog citron
T1
T2
Trifoliate orange
S1
S2
Sour orange
CT1
CT2
CS1
CS2
Etrog citron

699 **Fig. 2.** Rod-like secondary structure of minimum free energy of the Ci-1
700 haplotype of CEVd. Arrows delimit the five structural domains characteristic
701 of members of the family *Pospiviroidae* (Keese and Symons, 1985): terminal left
702 (T_L), pathogenicity (P), central (C), variable (V) and terminal right (T_R). Bold
703 letters indicate the conserved structural motifs: terminal conserved region
704 (TCR), central conserved region (CCR), Hairpin-II and "RY" motif.
705

FIGURE 2



705 **Fig. 3.** Molecular diversity at each nucleotide site of the CEVd molecule
706 computed for the overall CEVd populations as Shannon entropy. The five
707 structural domains characteristic of members of the family *Pospiviroidae* (Keese
708 and Symons, 1985): terminal left (T_L), pathogenicity (P), central (C), variable (V)
709 and terminal right (T_R) are indicated in different colours.
710

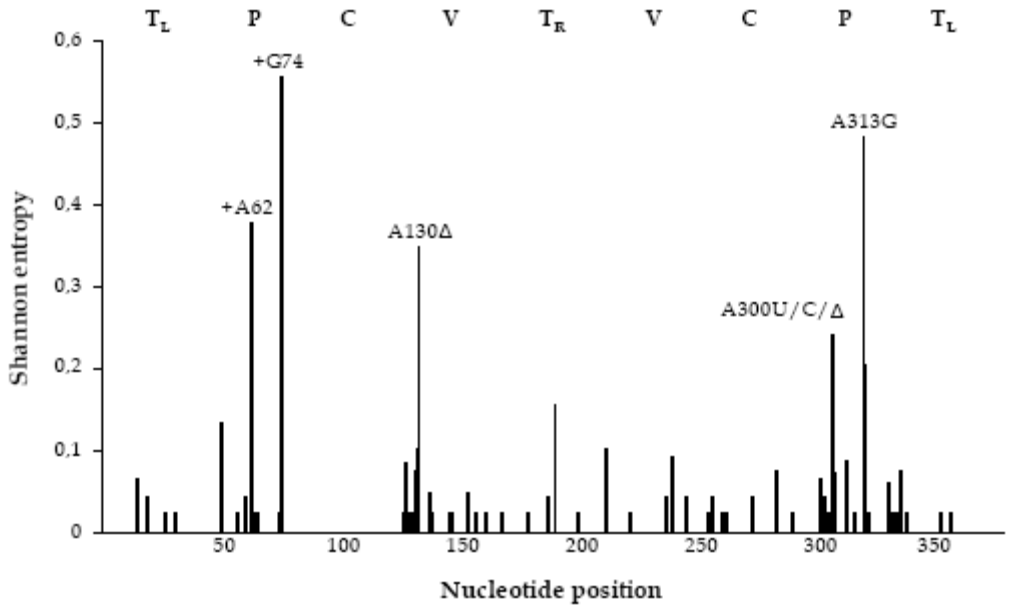


TABLE 1: Descriptive parameters of CEVd heterogeneous populations.

CEVd populations*	No. of haplotypes/ total no. of sequenced clones	Most frequent haplotype in CEVd population	Min. <Consensus<Max. haplotype length (nt)	No. and percentage of polymorphic sites in the genome†	Average no. of observed mutations per haplotype‡	Mutation frequency§	CEVd population diversity (*H*)□
Ci	9/30	Ci-1 (50%)	370<371<373	8 (2.16%)	0.6331	7.18×10^{-4}	0.0051
T1	9/30	T1-1 (63.3%)	369<370<373	11 (2.97%)	0.5664	9.99×10^{-4}	0.0056
T2	9/30	T2-1 (60%)	369<371<372	11 (2.96%)	0.5329	10.8×10^{-4}	0.0054
S1	15/30	S1-12 (20%) S1-4 (20%)	369<373-373	19 (5.09%)	1.7660	19.8×10^{-4}	0.0149
S2	17/30	S2-11 (16.6%)	369<373<374	16 (4.29%)	0.7595	14.0×10^{-4}	0.0114
CT1	7/30	CT1-10 (60%)	371-371<373	6 (1.62%)	0.4997	5.39×10^{-4}	0.0040
CT2	10/30	CT2-6 (50%)	371-371<372	11 (2.96%)	0.7328	8.98×10^{-4}	0.0070
CS1	10/30	CS1-12 (50%)	371-371<372	9 (2.43%)	0.6997	8.08×10^{-4}	0.0062
CS2	12/30	CS2-12 (36.6%)	370<371<373	12 (3.23%)	0.8993	10.8×10^{-4}	0.0075

*CEVd populations from the following hosts: 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).

†No. of sites where one or more mutations have been observed into each individual CEVd population. Percentages are obtained when this no. is divided by no. of nucleotides of the consensus sequence of the population

‡Total no. of mutations observed in each CEVd population/ Total no. of sequenced clones (30 clones)

§Total no. of non- repeated mutations observed in each CEVd population/ (No. of nucleotides of the consensus sequence of the population * total no. of sequenced clones)

□Shannon entropy (*H*)