The molecular characterization of 16 new sequence variants of *Hop stunt viroid* reveals the existence of invariable regions and a conserved hammerhead-like structure on the viroid molecule

Khalid Amari,1 Gustavo Gomez,1† Arben Myrta,2 Biagio Di Terlizzi3 and Vicente Pallás1†

1 Departamento de Mejora y Patología Vegetal, CEBAS-CSIC, Campus Universitario de Espinardo, PO Box 4195, 30071 Murcia, Spain
2 Ministry of Agriculture and Food, Tirana, Albania
3 Istituto Agronomico Mediterraneo, Valenzano, Bari, Italy

At present isolates of *Hop stunt viroid* (HSVd) are divided into five groups: three major groups (plum-type, hop-type and citrus-type) each containing isolates from only a limited number of isolation hosts and two minor groups that were presumed to derive from recombination events between members of the main groups. In this work we present the characterization of 16 new sequence variants of HSVd obtained from four Mediterranean countries (Cyprus, Greece, Morocco and Turkey) where this viroid had not previously been described. Molecular variability comparisons considering the totality of the sequence variants characterized so far revealed that most of the variability is found in the pathogenic and variable domains of the viroid molecule whereas both the terminal right (TR) and left (TL) domains are regions of low or no variability, respectively, suggesting the existence of constraints limiting the heterogeneity of the sequence variants. Phylogenetic analyses revealed that sequence variants belonging to the two minor recombinant subgroups are more frequent than previously thought. When the cruciform structure alternative to the typical rod-like conformation was considered it was observed that the upper part of this structure (hairpin I) was strictly conserved whereas in the lower part a reduced variability was found. The existence of a covariation in this lower part was notable. Interestingly, a hammerhead-like sequence was found within the TR domain of HSVd and it was strictly conserved in all the sequence variants. The evolutionary implications of the presence of this motif on the HSVd are discussed.

Introduction

Viroids are small circular single-stranded RNA molecules with sizes ranging between 246 and 399 nucleotides (Semancik, 1987; Diener, 1991; Symons, 1997; Flores et al., 1997). They are the smallest known pathogens and cause several economically significant crop diseases. Viroids are not known to code for any protein and therefore they are host-dependent in their life-cycle. The genome of most viroids characterized so far contains five structural domains (Keese & Symons, 1985): two terminal regions, left (TL) and right (TR), pathogenic (P), variable (V) and central domain (C) with a central conserved region (CCR). Those viroids lacking a CCR present a hammerhead self-cleavage ability, which introduces substantial differences in some steps of their replication cycle. These two features, the CCR domain or a hammerhead structure, have been used as the major criteria to classify these pathogens (Koltunow & Rezaian, 1988; Flores et al., 1998) and two main viroid families can be considered: members of the *Pospiviroidae* family are characterized by having a CCR domain and lacking hammerhead self-cleavage whereas those of the *Assunviroidae* family include viroids lacking the CCR domain and having hammerhead self-cleavage. Molecular variability studies have shown that these two features, the CCR domain or a

**Author for correspondence:** Vicente Pallás.
Fax +34 96 3877859. e-mail vpallas@ibmcp.upv.es

† Present address: Instituto de Biología Molecular y Celular de Plantas (UPV-CSIC), Universidad Politécnica de Valencia, Avenida de los Naranjos s/n, 46022 Valencia, Spain

The sequences described in this work have been deposited in the EMBL database and received accession numbers AJ297825 to AJ297840.
Hammerhead structure, are extremely conserved among different isolates of a viroid (e.g. Kofalvi et al., 1997; Ambrós et al., 1998; Pelchat et al., 2000). These kinds of studies contribute substantially to the determination of how important other structural motifs or stretches of sequences are in the viroid life-cycle.

Hop stunt viroid (HSVd) belongs to the Pospiviroidae family. It has been found in a wide range of hosts including hop, cucumber, grapevine, citrus, plum, peach, pear (Shikata, 1990) and, recently, apricot and almond (Astruc et al., 1996; Cañizares et al., 1999). The infection seems to be latent in some hosts such as grapevine (Shikata, 1990; Polivka et al., 1996) and apricot (Astruc et al., 1996). In other cases, specific disorders such as hop stunt (Shikata, 1990), dapple fruit disease of plum and peach (Sano et al., 1989) and citrus cachexia (Diener et al., 1988; Semancik et al., 1988) have been associated with HSVd infection.

Historically, HSVd sequence isolates have been divided into three groups (i.e. plum-type, hop-type and citrus-type) on the basis of overall homology (Shikata, 1990). The characterization of ten new sequence variants from three different Prunus species and the subsequent phylogenetic analysis revealed the appearance of two new groups that very probably derived from recombination events (Kofalvi et al., 1997). In addition, it

Table 1. HSVd sources used and HSVd sequence variants analysed in this work

<table>
<thead>
<tr>
<th>Origin</th>
<th>Apricot source</th>
<th>Sequenced clones</th>
<th>HSVd sequence</th>
<th>Size (nt)</th>
<th>Closest HSVd sequence</th>
<th>Nucleotide differences with closest sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morocco</td>
<td>Canino</td>
<td>Mo 1.4</td>
<td>apr9</td>
<td>298</td>
<td>HSVd.apr4</td>
<td>C192-, G205T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 1.9</td>
<td>apr10</td>
<td>298</td>
<td>HSVd.apr4</td>
<td>C192-, G205T, G206T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 1.10</td>
<td>apr9</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 2.1</td>
<td>apr11</td>
<td>298</td>
<td>HSVd.apr4</td>
<td>C192-, G205T, G206T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 2.3</td>
<td>apr11</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 2.6</td>
<td>apr12</td>
<td>298</td>
<td>HSVd.apr4</td>
<td>T60G, C192-, G205T, G263C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 44.2</td>
<td>apr13</td>
<td>297</td>
<td>HSVd.g3</td>
<td>C25T, A55G, C107G, -107A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 44.5</td>
<td>apr13</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mo 44.4</td>
<td>apr13</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tr 19.4</td>
<td>apr14</td>
<td>295</td>
<td>HSVd.h1</td>
<td>G54A, -106G, -107A, T204C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tr 19.5</td>
<td>apr14</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tr 19.9</td>
<td>apr14</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 10.1</td>
<td>apr2</td>
<td>299</td>
<td>HSVd.apr2</td>
<td>C240T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 10.2</td>
<td>apr15</td>
<td>299</td>
<td>HSVd.apr2</td>
<td>C240T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 14.1</td>
<td>apr2</td>
<td>299</td>
<td>HSVd.apr2</td>
<td>C240T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 14.2</td>
<td>apr16</td>
<td>299</td>
<td>HSVd.apr2</td>
<td>C240T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 15.1</td>
<td>apr17</td>
<td>300</td>
<td>HSVd.apr2</td>
<td>A103-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 15.2</td>
<td>apr2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 16.1</td>
<td>apr2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 16.2</td>
<td>apr18</td>
<td>297</td>
<td>HSVd.g3</td>
<td>C25T, A48T, G49C, C107G, -107A, A265G</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 17.1</td>
<td>apr2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 17.2</td>
<td>apr19</td>
<td>299</td>
<td>HSVd.apr2</td>
<td>T49G</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 18.1</td>
<td>apr20</td>
<td>299</td>
<td>HSVd.apr2</td>
<td>T127C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 18.2</td>
<td>apr2</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 25.1</td>
<td>apr21</td>
<td>296</td>
<td>HSVd.apr1</td>
<td>A25T, A26C, T32A, C152-, T158C, C159T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cy 25.2</td>
<td>aprCy7</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Greece</td>
<td>Kolioponlou</td>
<td>Gr 6.1</td>
<td>apr22</td>
<td>297</td>
<td>HSVd.g3</td>
<td>C18T, C25T, A48T, G49C, A55G, C107G, -107A</td>
</tr>
<tr>
<td></td>
<td>Pr. Porou</td>
<td>Gr 6.2</td>
<td>apr22</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gr 7.1</td>
<td>apr9</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gr 7.2</td>
<td>apr9</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gr 12.1</td>
<td>apr23</td>
<td>297</td>
<td>HSVd.g3</td>
<td>C25T, A48T, G49C, A55G, C107G, -107A</td>
</tr>
<tr>
<td></td>
<td>Bebecou</td>
<td>Gr 12.2</td>
<td>apr24</td>
<td>297</td>
<td>HSVd.g3</td>
<td>C25T, A48T, G49C, A55G, C107G, -107A</td>
</tr>
</tbody>
</table>
was shown that the previous hop-type group itself is likely to be the result of a recombination between members of the plum-type and citrus-type groups.

Until now, characterization of the primary structure of HSVd isolates has been carried out using isolates from Spain, France, Italy, USA and Japan. In this work we present the characterization of 16 new sequence variants of HSVd obtained from four Mediterranean countries (Cyprus, Greece, Morocco and Turkey) from where no sequence data were available before this work. Molecular variability comparisons and phylogenetic analyses revealed that sequence variants belonging to the two minor recombinant subgroups are more frequent than previously thought and that there are stretches of sequences on the viroid molecule that are highly conserved, suggesting key roles in the viroid life-cycle. In addition we identified a hammerhead-like structure within the T_R domain that is strictly conserved in all the sequence variants characterized so far and that can be considered as an evolutionary link between typical viroids and those having the ability to undergo self-cleavage via hammerhead ribozymes.

**Methods**

**Plant and viroid sources.** Fourteen HSVd-infected sources were used in this study (Table 1). All of them were from apricot (*Prunus armeniaca* L.) from four Mediterranean countries: Cyprus, Greece, Morocco and Turkey (Amari et al., 2000). A total of ten different cultivars were covered: Monaco bello, Cafiola, Boccia spina, Palamella and Canino from Cyprus, Kolioponlou, Pr. Porou and Bebecou from Greece, Canino and Maoui 1.2 from Morocco and Septik from Turkey.

**Preparation of low molecular mass RNAs.** Processing of frozen leaves was based on a protocol previously described that avoids the use of organic solvents (Pallás et al., 1987; Astruc et al., 1996). Briefly, 0.5 g of leaf tissue was homogenized inside sealed plastic bags in the presence of 5 ml of extraction buffer (0.1 M Tris–HCl pH 8.1, 50 mM EDTA, 0.5 M NaCl, 10 mM 2-mercaptoethanol) using a hand-homogenizer. An aliquot (1 ml) of the homogenate was transferred to an Eppendorf tube of organic solvents (Palla et al., 2005). Preparation of low molecular mass RNAs. Canino and Maoui 1.2 from Morocco and Septik from Turkey.

**RT–PCR amplification, cloning and sequencing of viroid isolates.** RT–PCR was performed as described (Astruc et al., 1996) by using avian myeloblastosis virus reverse transcriptase (Promega) for the RT and *Pfu* DNA polymerase (Stratagene) for PCR amplification. The oligonucleotides used were the antisense 26-mer VP-19 (5′ dGCCGCCGGCTCTTCTCAGAGTAAAG 3′, complementary to HSVd residues 60–85) and the sense 27-mer VP-20 (5′ dGCCGCCGGCCAGCTCTTCTCAGAATCC 3′, residues 78–102). Both primers lie in the strictly conserved central region of HSVd and contain the unique endonuclease restriction site *Smal* (underlined). Following RT–PCR, electrophoretic analysis confirmed the presence of a monomeric PCR product of the expected size. The PCR products were phenol-extracted, ethanol-precipitated and digested with the *Smal* endonuclease. The resulting DNA fragments were cloned in the *Smal* site of dephosphorylated pUC18 plasmid. Since this pair of primers covers almost the totality of the CCR, another pair of primers (VP-98 and VP-99) was designed to study the molecular variability of this part of the molecule. These primers (VP-98 5′, dTCGCCAGACCCGGGCCCCCTC 3′, complementary to residues 120–140; and VP-99, 5′ dTCGGG-AAATTCTCGAGTTGCCGC 3′, HSVd residues 1–23) flank the CCR of HSVd and contain EcoRI and SacI restriction sites, respectively (underlined). The PCR products were phenol-extracted, ethanol-pre- cipitated and digested with EcoRI and SacI. The resulting DNA fragments were cloned in a previously digested Bluescript II KS + plasmid.

For all isolates, cDNA clones were identified by restriction analysis. Selected clones were sequenced in both orientations by using universal primers with an automated DNA sequence (ABI PRISM 337; Perkin-Elmer). The new sequence variants were named following the rules described previously (Kofalvi et al., 1997).

**Computer analysis of the sequences.** Multiple alignments of HSVd sequences were obtained using ClustalW (Thompson et al., 1994). The alignment was corrected manually to maximize sequence homology. Phylogenetic analyses were performed using the following programs of the PHYLIP 3.5c package (Felsenstein, 1993). DNADIST was used to calculate genetic distances, NEIGHBOUR (UPGMA or neighbour-joining methods) to cluster the variants from the distance data, DRAWTREE to draw the resulting phylogenetic tree and SEQBOOT (100 repetitions) and CONSENSE to perform bootstrap analysis.

**Results and Discussion**

**Characterization of new HSVd sequence variants**

HSVd was isolated from ten different apricot cultivar sources from which 14 isolates were characterized (Table 1 and below): one from Turkey, three from Greece, three from Morocco and seven from Cyprus. For each isolate two or three full-length cDNA clones and three partial clones were sequenced (Table 1). After the comparison between the sequences obtained from these isolates and the previously characterized ones, 16 new HSVd variants were found, one from Turkey, three from Greece, five from Morocco and seven from Cyprus, giving the widest analysis of HSVd diversity reported to date. Most of the new sequence variants had a length varying between 297 and 299 nucleotides, except apr14 and apr21, which had 295 and 296 nt, respectively, and apr17 with 300 nt. Six out of the 36 sequence variants characterized (all of them from Cyprus) were found to be identical to the previously described HSVd.apr2, which was isolated from apricot cultivars Roma 1 from Italy, Bulida d’Arques from Spain and from Japanese plum cv. Taito from Japan (Kofalvi et al., 1997).

Viroid apricot sources found to be homogeneous within the limits of the sparse sample group used include Maou 1.2 from Morocco, Septik from Turkey, Canino from Cyprus and Kolioponlou and Pr. Porou from Greece. However the other sources were heterogeneous, with two or three different sequences for two or three cDNA clones sequenced. It is worth noting that the two cDNA clones sequenced from the source...
Phylogenetic tree of HSVd sequence variants. Phylogenetic analysis was based on genetic distances (PHYLIP 3.5c package) calculated between all available HSVd sequence variants. Variants sequenced in this work are indicated with bold type. The position and branching of PSTVd, included in the analysis as an out-group, are indicated by an arrow. The five phylogenetic groups identified are delineated. Dots near nodes indicate the statistical value of the node as determined by bootstrap analysis (100 replicates). ●●●, Node detected in 100% of replicates; ●●, node detected in 80–100% of replicates; ●, node detected in > 50% of replicates.

Fig. 1. Phylogenetic tree of HSVd sequence variants. Phylogenetic analysis was based on genetic distances (PHYLIP 3.5c package) calculated between all available HSVd sequence variants. Variants sequenced in this work are indicated with bold type. The position and branching of PSTVd, included in the analysis as an out-group, are indicated by an arrow. The five phylogenetic groups identified are delineated. Dots near nodes indicate the statistical value of the node as determined by bootstrap analysis (100 replicates). ●●●, Node detected in 100% of replicates; ●●, node detected in 80–100% of replicates; ●, node detected in > 50% of replicates.

Pr. Porou from Greece were identical to HSVd.apr9 obtained from the Canino cultivar from Morocco (Table 1). The closest HSVd sequence for most of the sequence variants from Morocco, with the exception of apr13, was HSVd.apr4, a sequence variant previously identified in Bulida apricots from Spain (Kofalvi et al., 1997). The variant from Turkey was found to be very similar to HSVd.h1, a sequence variant originating from Japan. All variants from Greece, except the one obtained from Gr7.1 and Gr7.2 cDNA clones, were found to be very similar to HSVd.g3, which came from a German grapevine. Finally, all the variants from Cyprus, except apr18 and apr21, which are very similar to HSVd.g3 and HSVd.apr1 respectively, were found to be very similar to HSVd.apr2, a sequence variant that, as stated above, was previously detected in apricot cultivars originating from Spain and Italy (Kofalvi et al., 1997) and in plum from Japan.

The above results could be explained by the frequent plant exchange between different countries or, alternatively, by a parallel evolution of the viroid molecule, whose variability is restricted to certain polymorphic positions.
Several lines of evidence allowed us to conclude that we obtained a high degree of fidelity in the characterization of the new sequences variants. (i) Two different RT–PCR reactions were carried out for each isolate. The first one was done by using VP-19 and VP-20 primers that were designed to the central part of the molecule (see Methods). With this pair of primers the sequence of 251 out of 300 nt of the viroid molecule was determined for each isolate. To obtain clones representing the central part and surrounding areas VP-98 and VP-99 were designed and a different RT–PCR was carried out. The sequence of three clones for each isolate revealed that all of the mutations observed in these partial clones were coincident in all isolates with the mutations observed with the almost full-length clones obtained in the other RT–PCR. (ii) We used a thermostable DNA polymerase endowed with proofreading activity to minimize the introduction of substitutions during PCR amplification. It has been determined that the error rate of Pfu DNA polymerase (< 3 x 10^-8 errors per bp per cycle) is ten times lower than the error rate of Taq polymerase (Bracho et al., 1998). (iii) All the polymorphic positions found are present in at least five out of the 16 new sequence variants characterized in this work. In addition they are identical to other polymorphic positions previously described for other Prunus hosts (Kofalvi et al., 1997) and even for different non-Prunus hosts.

Phylogenetic analysis of the new HSVd variants

Alignment and phylogenetic analyses of the HSVd sequence variants characterized in this work were carried out together with the 38 HSVd sequences previously reported, giving a total of 54 HSVd sequences. Previously, HSVd variants were divided, according to overall sequence homologies (Sano et al., 1989) and phylogenetic analysis (Hsu et al., 1994), into three major groups: citrus-, hop- and plum-type. After characterizing ten new sequence variants, Kofalvi et al. (1997) redefined this classification into five groups, including the three previous ones and two new groups, having only two or three members, respectively. These two new groups could be considered as the results of recombination events between members of the plum- and citrus-type (now named P-C group) or between members of the plum- and hop-type or cit3 sequence variant (now named P-H/cit3 group) (Kofalvi et al., 1997).

Phylogenetic analysis of all sequences, including the sequence variants characterized in this work, showed that five out of the seven variants from Cyprus are included in the recombinant group P-C, which previously contained the variants HSVd.apr2 and HSVd.apr5 (Fig. 1). In the other recombinant group (P-H/cit3) are the three new variants from Greece, HSVd.apr18 from Cyprus and HSV.d.apr13 from Morocco (Fig. 1). These results reflect the facts that recombination events are more frequent than previously expected on HSVd and that intraspecific recombination could be a general mechanism in the evolution of viroids (Candresse et al., 1997). Finally, four out of the five variants from Morocco clustered into the plum-type group, confirming the homogeneous origin of the Moroccan isolates and showing their close relationship to the Spanish sequence variant isolated from the Bulida apricot, the most extended cultivar growth in the southeast of Spain having 81% HSVd infection (Cañizares et al., 1998).

As described previously (Hsu et al., 1994; Kofalvi et al., 1997), only 30 of the so-called ‘informative’ changes are required to discriminate phylogenetic clusters of HSVd variants. The positions and numbering as well as the frequency of these 30 ‘informative’ changes within the rod-like viroid structure are shown in Fig. 2(a). In addition to the 30 ‘informative’ changes, 15 extra changes were found in the 16 new HSVd sequences analysed in this work. The precise positions are depicted in Fig. 2 by the letters ‘a’ to ‘o’. Fig. 2(b, c) gives a condensed alignment of the 16 new sequence variants together with other representative sequence variants characterized previously, showing only those 30 ‘informative’ positions (Fig. 2b) which account for the discrimination of the five phylogenetic groups defined in Fig. 1 or the exclusive changes of the new 16 sequence variants (Fig. 2c).

Analysis of key structural elements on the viroid molecule

As stated above it has been proposed that the genome of typical viroids contains five structural domains (Keese & Symons, 1985). Fig. 2(a) shows the predicted rod-like conformation of HSVd in which structural domains are indicated. As can be seen, when all the 54 known sequence variants are compared, most of the variability is located on the pathogenic (P) and variable (V) domains and to a lesser extent in the lower strand of the central (C) domain. Both T_R and T_L domains are regions of very low or no variability, respectively, which suggests that they have a key role in the viroid life-cycle. The involvement of the T_R region on viroid replication has recently been suggested (Sano & Ishiguro, 1998). In addition, three nucleotide substitutions in the left terminal loop of the T_L domain of Potato spindle tuber viroid (PSTVd) resulted in the absence of systemic infection (Hu et al., 1997). Important here is that the left terminal conserved hairpin within the T_L region present in HSVd and in viroids belonging to the Cocavirus genus (represented by Coconut cadang-cadang viroid; CCCVd) is extremely conserved (Fig. 2), reinforcing the hypothesis that it must have an important role in the viroid molecule (Flores et al., 1997). Keese et al. (1988) showed that all the PSTVd-like viroids share a CCUC sequence (positions 295–1 in sequence variant h1) in the T_L domain and a CCUCUC sequence (position 152–156) in the T_R domain. Fig. 2 shows that such a sequence at the T_L domain is strictly conserved in all the sequence variants analysed, whereas sequence CCUC in the T_R domain changed to CUUC, losing the unpaired C (informative change...
Fig. 2. (a) Predicted secondary folding of HSVd showing the phylogenetically informative changes (numbers 1–30) as described by Hsu et al. (1994) and Kofalvi et al. (1997) and the polymorphic positions described in this work (letters ‘a’ to ‘o’). The variation frequency is indicated with colours (red, > 50%; green, 25–49%; and blue, < 25%). The most stable predicted structure for HSVd-1 is shown with the boundaries of the terminal left (TL), pathogenicity (P), central conserved (C), variable (V) and terminal right (TR) domains, indicated at the top of the figure. Regions with low and no variability are shaded with yellow and purple colours, respectively. Within the invariable region, the terminal conserved hairpin (TCH) described for HSVd and CCCVd is shown in yellow. Within the C domain, sequences involved in the formation of an alternative cruciform structure (see Fig. 3) are shown in blue. The sequence painted in magenta shows the location (in the TR domain) of the central region of the hammerhead-like structure described in Fig. 4 and described in the text. (b) Condensed alignment of the 30 phylogenetically informative changes for all HSVd variants. The numbering of the positions is shown at the top of the alignment. The sequence variants characterized in this work are indicated in bold type. (c) Condensed alignment of the additional 15 polymorphic positions of the sequence variants characterized in this work. The reference isolate for the two alignments was selected as HSVd.h1 (Ohno et al., 1983) and is shown at the top with the corresponding nucleotide numbering (where the alignment shows a gap, the numbering is that of the preceding residue). Residues (or gaps) identical to the reference isolate are not displayed in the alignments.
Fig. 3. Possible alternative cruciform structure of the central conserved domain in HSVd. In the upper part of this conformation (hairpin I) no variability was found, whereas in the lower hairpin the presence of a G residue (marked with an arrowhead) is always accompanied with the deletion of U207 and A225 (marked with asterisks). This covariation is present in 100% of the HSVd sequences analysed. Domains at the ends of both lower and upper parts of the cruciform structure having a high degree of complementarity are shaded. Nucleotide residues that vary among cachexia-inducing variants and non-pathogenic variants are marked with a large arrow (see text for details).

number 13; Fig. 2), in all citrus variants, in HSVd.apr4, HSVd.apr2 and HSVd.apr5, and in most of the new sequence variants from Morocco and Cyprus (Fig. 2b).

The presence of the inverted repeat sequences within the C domain of typical viroids allows the formation of a cruciform structure alternative to the rod-like conformation (Fig. 3), similar to the alternative structure described in the PSTVd group (Liu & Symons, 1998). Interestingly, the upper part of this cruciform structure (hairpin I) is strictly conserved in all the 54 sequence variants characterized, whereas in the lower part a reduced variability is allowed (Fig. 3). The presence of a G residue (marked with an arrowhead in Fig. 3) between U207 and C225 always led to the disappearance of U207 and A225 (marked with asterisks). This covariation might indicate the requirement of an unpaired residue in the proximal region of the lower stem of the cruciform structure that could be involved in a tertiary interaction with another residue on the loop located immediately downstream of the stem (U207–A225 or C212–extra G). Another important feature that is apparent from the cruciform representation is that the ends of both stem–loops are strictly conserved and complementary (see shaded area in Fig. 3), suggesting a tertiary structure that may have an important role in the viroid infection process.

Citrus viroids (CVds) have been classified into five groups of viroids variants (Durán-Vila et al., 1988). Group II contains HSVd-related variants including those inducing cachexia disease (CV-IIb, CV-IIc and CVd-903), no pathogenic symptoms (Cvd-IIa) and one inducing mild cachexia reactions (Cvd-909). Within the variable domain of the HSVd sequence variants infecting citrus (CVds-II) it has been proposed that a cluster of six nucleotide changes regulates the induction of cachexia disease (Reanwarakorn & Semancik, 1998). It is relevant that the cachexia-inducing sequence (CVd-IIb) is not present in any of the sequence variants obtained from any non-Citrus host. The most similar sequence variant to CVd-IIb is apr14, in which only two substitutions are needed to convert it to a cachexia-inducing sequence. For all the Moroccan and Greek sequence variants, three nucleotides would have to change to revert them to CV-IIb. In these two countries, Prunus and Citrus hosts are cultivated in close proximity, emphasizing
the need to control the sanitary status of apricot even though HSVd is considered to be latent in this crop. Curiously, a detailed analysis of the lower part of this cruciform structure revealed that in all citrus sequence variants except those inducing cachexia disease (CVd-IIb, CVd-Iic and Ca-909) the upper stem (marked with a shaded arrow in Fig. 3) was disturbed by the changes of C=G and G=A. This could explain why CVd-909, in spite of lacking all six nucleotide changes defined above in the cachexia-inducing sequence (in the V domain), incites mild cachexia reactions (Reanwarakorn & Semancik, 1999). Thus, a rule could be drawn for all citrus sequences characterized so far: those sequence variants with the six nucleotide changes within the V domain and with a very stable stem in the lower part of the cruciform structure would induce a severe cachexia (e.g. CVd-IIb, CVd-Iic); those sequence variants having one of the two features would induce mild cachexia (e.g. Ca 993 and Ca 903, see Table 1 in Reanwarakorn & Semancik, 1999) and those sequences having an unaltered stem (e.g. CVd-IIa and the rest of citrus sequences) would not induce cachexia disease.

On the basis of phylogenetic analysis of all known viroid sequences, the viroid-like satellite RNAs and the viroid-like domain of the Hepatitis delta virus RNA, it has been suggested that viroids may have evolved from satellite RNAs while still free-living molecules, with both presumably acquiring a dependence on their host (viroids) or helper virus (satellite RNAs) after becoming intracellular entities. It is then reasonable to think that, in this scenario, typical viroids could have maintained relics of the self-cleaving structures. In Fig. 4, we show a partial hammerhead structure formed in the TR domain of the HSVd molecule in which nine out of 13 absolutely conserved nucleotides of a typical hammerhead ribozyme (Hertel et al., 1992) are present (boxed nucleotides in Fig. 4). As a reference, the ribozyme of Peach latent mosaic viroid (PLMVd) (Hernández & Flores, 1992) has been included in Fig. 4. It is worth noting that a similar hammerhead-like structure with identical levels of similarity to the one observed in the plus-strand was also found in the negative-strand of the HSVd molecule (not shown). In addition, the presence of a high number of nucleotides that are conserved in a hammerhead ribozyme in a region of the negative polarity of a self-cleaving circular RNA associated with Rice yellow mottle virus has been considered as a vestige of an ancestral functional hammerhead (Collins et al., 1998). By introducing three substitutions and one insertion between helix I and II in the HSVd pseudo-ribozyme, a canonical hammerhead ribozyme could be reverted. To the best of our knowledge this is the first time that protein catalysts, the present day RNA-catalysed reactions could then be considered as ancestral to similar reactions that have become protein-dependent. In this context, Diener (1996) has suggested that viroids may have evolved from satellite RNAs while still free-living molecules, with both presumably acquiring a dependence on their host (viroids) or helper virus (satellite RNAs) after becoming intracellular entities. It is then reasonable to think that, in this scenario, typical viroids could have maintained relics of the self-cleaving structures. In Fig. 4, we show a partial hammerhead structure formed in the TR domain of the HSVd molecule in which nine out of 13 absolutely conserved nucleotides of a typical hammerhead ribozyme (Hertel et al., 1992) are present (boxed nucleotides in Fig. 4). As a reference, the ribozyme of Peach latent mosaic viroid (PLMVd) (Hernández & Flores, 1992) has been included in Fig. 4. It is worth noting that a similar hammerhead-like structure with identical levels of similarity to the one observed in the plus-strand was also found in the negative-strand of the HSVd molecule (not shown). In addition, the presence of a high number of nucleotides that are conserved in a hammerhead ribozyme in a region of the negative polarity of a self-cleaving circular RNA associated with Rice yellow mottle virus has been considered as a vestige of an ancestral functional hammerhead (Collins et al., 1998). By introducing three substitutions and one insertion between helix I and II in the HSVd pseudo-ribozyme, a canonical hammerhead ribozyme could be reverted. To the best of our knowledge this is the first time that
a hammerhead-like structure has been described for a typical viroid and it could represent an evolutionary link between typical viroids and those having the ability to undergo self-cleavage via hammerhead ribozymes.

Interestingly, the motif covering the putative HSVd pseudo-hammerhead is strictly conserved among the 54 known sequence variants, indicating that this viroid region is not prone to sequence variability and suggesting a putative key role in the viroid life-cycle. It has been proposed that self-cleavage reactions will also be involved in the replication of the PSTVd group of viroids (Symons, 1997) and preliminary supporting evidence was recently provided for CCCvd (Liu & Symons, 1998). Experiments are in progress to determine the in vivo viability of an HSVd mutant bearing the four changes required to acquire ribozyme activity.

This work was partially supported by an AIR2 project of the European Community (#93-1567) and by the research co-operation activities of the Mediterranean Network on Virus Disease Assessment and Sanitation of Stone Fruit Trees (MNFT) promoted by CIHEAM-IAM.B/EC-CDG.1. We thank Dr. C. Hernández for her valuable comments and the critical reading of the manuscript and Dr. R. Flores for the facilities given during the realization of part of this work. G. Gomez was the recipient of a fellowship from the CONICET of Argentina.

References


K. Amari and others


Received 13 November 2000; Accepted 22 December 2000
Published ahead of print (24 January 2001) in JGV Direct as DOI 10.1099/vir.0.17540-0