

# Molecular features of new Peach Latent Mosaic Viroid variants suggest that recombination may have contributed to the evolution of this infectious RNA

I. Fekih Hassen<sup>a,b</sup>, S. Massart<sup>a</sup>, J. Motard<sup>c</sup>, S. Roussel<sup>a</sup>, O. Parisi<sup>a</sup>, J. Kummert<sup>a</sup>, H. Fakhfakh<sup>b</sup>,  
M. Marrakchi<sup>b</sup>, J.P. Perreault<sup>c</sup>, M.H. Jijakli<sup>a,\*</sup>

<sup>a</sup> Plant Pathology Unit, Faculté Universitaire des Sciences Agronomiques, Passage des Déportés, 2, B-5030 Gembloux, Belgium

<sup>b</sup> Laboratory of Molecular Genetic, Immunology and Biotechnology, Faculty of Sciences of Tunis, 2092 Elmanar Tunis, Tunisia

<sup>c</sup> RNA group/Groupe ARN, Département de Biochimie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Québec, Canada J1H 5N4

Received 2 July 2006; returned to author for revision 14 August 2006; accepted 12 October 2006

Available online 20 November 2006

## Abstract

Nucleotide sequences of a broad range of Peach Latent Mosaic Viroid (PLMVd) variants were determined. The variants were isolated from peach, pear, and almond tree samples collected in Tunisia. Sequence analysis confirmed the high variability of PLMVd, as no less than 119 new variants were identified. Variations included new polymorphic positions, insertions of 11 to 14 nucleotides, and new mutations within the hammerhead self-cleavage motifs. We provide the first covariation-based evidence for certain stems within the proposed secondary structure. Our covariation analysis also strengthens the view that a pseudoknot closes the replication domain. On the basis of phylogenetic tree studies and informative positions, PLMVd variants are proposed to cluster into groups and subgroups likely to have resulted from recombination events. PLMVd thus emerges as a suitable viroid for retracing the evolution of an RNA genome.

© 2006 Elsevier Inc. All rights reserved.

**Keywords:** Viroid; Sequence variation; Hammerhead self-cleavage; Recombination; Secondary structure

## Introduction

Viroids are small [246–401 nucleotides (nt)], circular, single-stranded RNA molecules infecting various plants and thereby causing important losses in agriculture (Daros et al., 2006). As there is no evidence for RNA translation, their pathogenic effects must result from direct interaction with host components. Viroids thus offer a unique model for exploring host–pathogen interactions that depend strictly on the pathogen's RNA sequence and structure. The 30 viroid species characterized so far have been classified into two major families (Daros et al., 2006). The *Pospiviroidae* family, whose type species is Potato Spindle Tuber Viroid (PSTVd), comprises 26 viroid species characterized by a rod-like secondary structure with five structural/functional domains including a central conserved region. These species replicate and accumulate in the nucleus via an asymmetric rolling-circle mechanism that

involves host enzymes for all RNA processing reactions. The *Avsunviroidae* family, whose type species is Avocado Sunblotch Viroid (ASBVd), comprises four species. Within this family, ASBVd and the Eggplant Latent Viroid (ELVd) adopt a quasi-rod-like secondary structure, whereas Peach Latent Mosaic Viroid (PLMVd) and Chrysanthemum Chlorotic Mottle Viroid (CChMVd) fold into a branched secondary structure. These four viroids replicate and accumulate in chloroplasts through a symmetric rolling-circle mechanism involving the hammerhead self-cleavage mechanism and, most likely, RNA self-ligation.

PLMVd is the causal agent of peach latent mosaic disease (Desvignes, 1986). Under field conditions, PLMVd symptoms became visible 2 years after planting infected material. The viroid can induce delays in foliation, flowering, and ripening, fruit deformations, bud necrosis, and rapid aging of the tree (Desvignes, 1986; Llacer, 1998). In greenhouse, PLMVd isolates have been classified as latent or severe and have been found to induce a broad variety of symptoms. An example is peach calico, caused by a 12- to 14-nt insertion folding itself into a hairpin structure (Malfitano et al., 2003; Rodio et al., 2006).

\* Corresponding author. Fax: +32 81 61 01 26.

E-mail address: [jijakli.h@fsagx.ac.be](mailto:jijakli.h@fsagx.ac.be) (M.H. Jijakli).

PLMVd is known to infect peach (*Prunus persica*) in Europe, Asia, Africa, and both North and South America (Hadidi et al., 1997; Pelchat et al., 2000; Fekih Hassen et al., 2004). It has also been detected occasionally in apricot (*Prunus armeniaca*), plum (*Prunus domestica*), sweet cherry (*Prunus avium*), cultivated pear (*Pyrus communis*), wild pear (*Pyrus amygdaliformis*), mume (*Prunus mume*), and almond (*Prunus amygdalus*) (Faggioli et al., 1997; Hadidi et al., 1997; Kyriakopoulou et al., 2001; Fekih Hassen et al., 2004, 2005). Molecular variants isolated from European and North American sources have 335 to 351 nt and showed high naturally occurring polymorphism (Hernandez and Flores, 1992; Shamloul et al., 1995; Ambros et al., 1998; Pelchat et al., 2000; Malfitano et al., 2003). Analyses of the progenies of individual PLMVd cDNA clones have revealed the extremely heterogeneous character of this viroid, due most probably to the ability of PLMVd RNA to tolerate changes rather than to repeated infection of the same individual tree (Ambros et al., 1999).

Previous investigations of PLMVd epidemiology in Tunisia have revealed a high incidence of this pathogen in peach trees (unpublished data, I. Fekih Hassen and M.H. Jijakli). To learn more about the PLMVd variants infecting trees on the African continent, we have collected isolates from trees in Tunisia and have performed the widest-ever PLMVd sequence alignment. Our results highlight interesting sequence features and provide insights into PLMVd secondary structure and population evolution.

## Results and discussion

### Analysis of PLMVd primary structure

RNA samples were isolated from leaves of 31 peach trees (belonging to 14 different cultivars), 2 almond trees, and 1 pear tree. The samples were taken from trees in different regions of Tunisia: Ben Arous, Bizerte, Manouba, Nabeul, and Zaghouane in the north, Kairouane in the center, and Mahdia in the Sahel. Initially, PLMVd was detected by RT-PCR in all samples (Fekih Hassen et al., 2006). The selection of the primers was achieved according to a previous sequencing study (Ambros et al., 1998). As these authors observed little sequence variation in the P3 and P4 stems, this region was suitable for designing primers (positions 92 to 141 in Fig. 1). In order to improve the PCR fidelity, the RT-PCR reactions were performed using the Titan RT-PCR kit (Roche Applied Biosciences) containing the proof-reading enzyme Tgo polymerase, together with the Taq polymerases. The fidelity of our protocol was checked. A PLMVd cDNA clone was transcribed *in vitro* and the RNA was submitted to RT-PCR. The PCR fragment was cloned and 18 clones were sequenced. The new sequences revealed 13 identical clones and five clones with only one mutation each. Three of these clones contained a substitution of one nucleotide while the two other contained a single deletion that occurred in the antisense primer and which could originate from imperfect chemical synthesis of the primer. Regardless of whether any

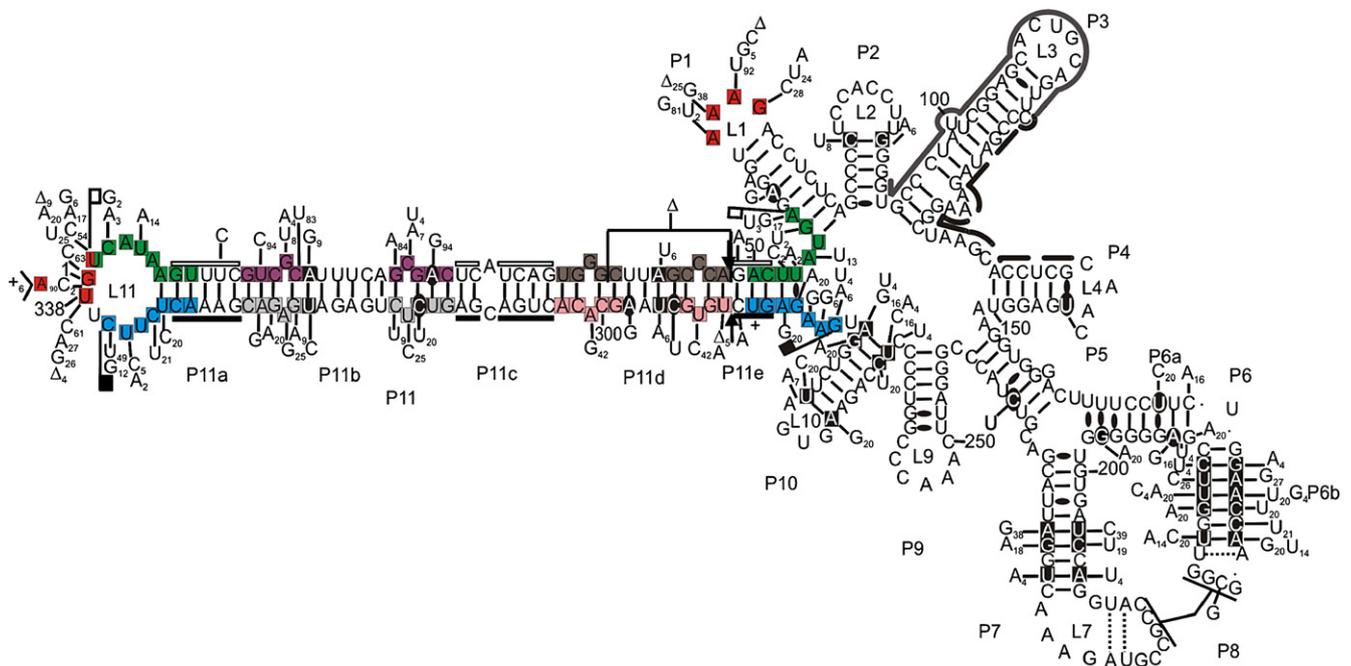


Fig. 1. Sequences and secondary structure of the PLMVd variants. Most of the mutations found in the 119 sequences are indicated along the secondary structure of the PLMVd reference sequence Ar1 which include a duplication of G at position 258 in this molecule (Beaudry et al., 1995). The frequency of mutations found more than once is indicated in subscript. Base-pairs supported by covariation are squared while those supported by one mutation are in ovals. Nucleotides forming the pseudoknot between L1 and L11 are in red. The symbol + on the left of the RNA molecule indicates the position of 11- to 14-nt insertions found in 6 variants. Regions involved in forming plus and minus polarities in hammerhead structures are flanked by closed and open flags, respectively. The hammerhead consensus sequences of plus and minus polarities are indicated by closed and open bars, respectively. Stems I, II, and III of the plus-polarity hammerhead are in pink, grey, and blue, while those of the minus-polarity hammerhead are in brown, purple, and green. Arrows indicate self-cleavage sites. Continuous and discontinuous lines along the P3 and P4 stems correspond to the positions of the antisense and sense primers, respectively.

potential bias was introduced by the *in vitro* transcription, the RT-PCR amplification followed by cloning and sequencing of the PCR fragment appeared to have good fidelity. Thus, we can assume that most of the observed sequence heterogeneity resulted from PLMVd heterogeneity in the sampled leaves.

So, the RT-PCR products obtained from our samples were cloned. Two to four clones were sequenced for each isolate. A total of 123 clones were sequenced, in both directions to avoid any sequencing error. This yielded 119 different PLMVd variants (reported in the GenBank Nucleotide Sequence Database accession number DQ680688 to DQ680806; see also Supplementary Table 1). On the basis of their sequence lengths the variants form two clusters: a cluster of 113 variants ranging in size from 336 to 341 nt, and a cluster of 6 variants with lengths between 341 and 351 nt (see Supplementary Table 1 for details). The latter variants show an insert of 11 to 14 nt in the left-hand loop L11, between nucleotides 1 and 338 of the reference variant Ar1 (GenBank accession no. M83545). One of the six variants (variant 235.2) displays a nine-nucleotide deletion (positions 41 to 49) leading to a sequence length of 341 nt. This shows that length conservation is not an essential feature for PLMVd.

Sequence alignment of the Tunisian PLMVd variants with all the PLMVd variants existing in databases revealed that the 119 sequences are novel variants. This set of sequences showed 127 polymorphic positions corresponding to 36.8% of the 345 positions in the alignment (see Supplementary Fig. 1). Among them, 46 mutations have never been described previously. The observed substitutions, insertions, and deletions are unevenly distributed along the PLMVd molecule. Mutational analysis showed that most of the variations are located in the regions spanning nucleotides 1 to 70 and 170 to 345 of the alignment (see Supplementary Fig. 1). A tiny amount of mutations, mainly deletions, were observed at the level of the primer sequences. These mutations might originate from imperfect chemical synthesis of the oligonucleotides. However, the possibility that they originated from PLMVd evolution cannot be completely excluded, and consequently, the sequences reported in Genebank included these mutations.

Different sequences were obtained for almost all isolates, thus confirming the quasi-species nature of PLMVd. We observed no sequence variant–cultivar correlation and no regional specificity. For example, an identical sequence was obtained from two different cultivars: Early May Crest and Royal Glory (variants 244.1 and 253.2). We likewise found no sequence basis for distinguishing PLMVd variants isolated from peach, pear, and almond trees. This suggests that there is no host-related sequence specificity.

### Secondary structure of PLMVd

In regards to the new collection of variants, the observed sequence variations have been illustrated using the PLMVd reference variant and the consensus secondary structure that has been established previously following sequence analysis as well as nuclease probing (Bussi ere et al., 2000; Pelchat et al., 2000). Variations appear predominantly in the regions including the P1,

P6, P7, P9, P10, and P11 stems, as opposed to those forming the P2, P5, and P8 stems (see Fig. 1 and Supplementary Table 2).

PLMVd secondary structure was analyzed with two different approaches. Firstly, the Mfold package (Zuker, 1989) was used to identify secondary structures showing the highest thermodynamic stability (i.e. the lowest Gibbs free energy,  $\Delta G$ ). In all variants the P7, P9, P10, and P11 stems appeared in the most stable structures (see Fig. 1), and in most cases so did the P1 and P6 stems. The P8 stem cannot be predicted using Mfold since pseudoknot prediction is not a functionality of this software. The existence of this pseudoknot received physical support from nuclease probing data (Bussi ere et al., 2000). Only one mutation was observed in all our sequence collection suggesting that its GC rich composition is probably important to preserve its formation. Finally, the variants having an insert sequence in the L11 loop (i.e. at the extremity of the P11 stem) displayed various potential structures for this region (discussed below).

Secondly, we performed nucleotide covariation analyses. Either, covariation of base-paired residues or nucleotide mutations was detected within the P6a, P6b, P7, P10, and P11 stems (see Fig. 1). Most importantly, covariation was found for 5 base-pairs (bp) of the P6b stem and 3 bp of both the P7 and P10 stems, thus supporting the existence of these helical regions. This contrasts with previous analyses that failed to detect covariation in the P6b and P7 stems (Pelchat et al., 2000). Conversely, we observed only one covariation in the P2 stem, and none in the P5 stem and the P8 pseudoknot. Since mapping experiments have proven the presence of paired nucleotides in the P2, P5, and P8 stems and since these sequences are highly conserved, it could be that selective pressure has preserved them intact. Thus, only the P3 and P4 stems might have the potential to fold into alternative structures. They provide no covariation-based support for a specific structure. Yet as the oligonucleotides used to amplify the entire PLMVd molecule corresponded to the sequences forming these stems, this may explain why only few variations were observed yet in both the P3 and P4 stem (data not shown).

Despite the high variability observed within the L1 and L11 loops, frequent nucleotide covariations were observed for the formation of a pseudoknot by the nucleotides of the positions 338 to 2 with those of the positions 64 to 67 of these regions (Fig. 2A). Of the 119 variants, 86 have the potential to adopt a 4-bp pseudoknot, while six others might establish a shorter 3-bp version. A similar interaction emerged from analyzing the minus-polarity PLMVd strand (data not shown). It should be mentioned that the presence of this pseudoknot has been proposed previously, although it was suggested to involve 4 to 7 bp and to tolerate a mismatch of 1 or 2 bp (Ambros et al., 1998, 1999). Importantly, the formation of this pseudoknot would have the effect of closing the replication domain forming the P11 stem. This proposed pseudoknot structure could increase the stability of the PLMVd molecule and confer a compact form to the PLMVd RNA. The inability of an *in vitro* assay to reveal the L11–L1 pseudoknot motif (Bussi ere et al., 2000) might be due to the absence of host proteins that could intervene *in vivo* to stabilize the pseudoknot structure. Furthermore, our proposed structures are not necessarily the

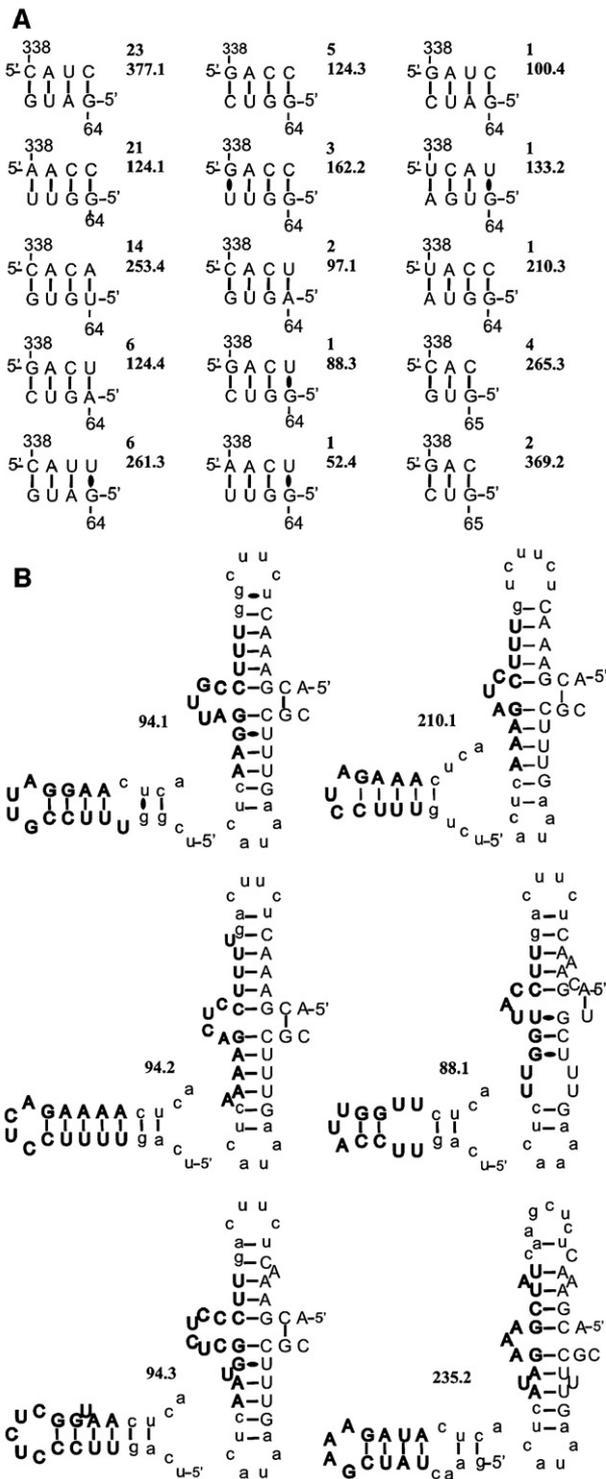


Fig. 2. Sequences of specific secondary structure motifs. (A) Potential pseudoknot interactions between nucleotides of the L11 and L1 loops. The first number indicates the number of PLMVd variants presenting this pseudoknot, while the second number is an example of variant. (B) Potential structures adopted by the 11- to 14 nt-insertion in the L11 loop. Nucleotides corresponding to the insertion are indicated in extra-bold uppercase letters, while those of the P11a stem and L11 loop are indicated in uppercase and lowercase letters, respectively.

biologically active ones, as illustrated by the absence, in the most stable secondary structure, of the hammerhead structures known to be crucial to PLMVd replication. Clearly, additional experiments are required to confirm the formation of this pseudoknot *in vivo* and to identify its function.

The PLMVd sequences having an insertion within the L11 loop, reminiscent of the peach calico insertion (Malfitano et al., 2003), lack the potential to form the L1–L11 pseudoknot. The sequence of the 11- to 14-nt insertion was found to differ among variants and with respect to all other insertions reported previously. These insertions can fold itself or with the sequence of the loop L11 and the stem P11a into alternative structures (Fig. 2B). In several cases, the resulting hairpins were found to include a 4- to 7-bp stem capped by a 3- to 5-nt loop. Alternatively, two contiguous stems or palindromic structures can be formed. Clearly, biochemical experiments are required to establish the structure of this domain. In none of these variants did the overall PLMVd secondary structure or the hammerhead structures appear to be altered.

Our results constitute additional evidence in favor of a branched secondary structure for PLMVd. This compact structure, also proposed for CChMVd (Bussière et al., 2000), appears as a characteristic of these viroids. It might explain their unique insolubility in 2 M lithium chloride as compared to other viroid species (Daros et al., 2006).

Regarding sequence variations and secondary structure in PLMVd, our data support the notion that this viroid comprises two domains: a left domain (proposed to be essential to replication) having a conserved secondary structure despite sequence variations and a right domain (suggested to be important for a non-crucial function such as pathogenesis) showing a relatively conserved sequence but possible structural variations. Yet our results suggest a somewhat different definition of the two domains than proposed by Pelchat et al. (2000) for 34 North American PLMVd variants. These authors describe a left domain comprising stems P1, P2, P10, and P11 and a right domain formed by stems P3 to P9. The main difference between previously studied variants (Ambros et al., 1998; Pelchat et al., 2000) and our new Tunisian variants concerns stems P6 and P7: in the former these stems show low variability, but in the latter they show a high level of nucleotide covariation. Moreover, stem P2 showed less variability, as opposed to the previously published PLMVd variants. Our data thus suggest that the left domain is larger than initially believed and that the right domain is limited to stems P2 to P5.

*Hammerhead motifs and self-cleavage activity*

The P11 stem contains the hammerhead self-cleaving RNA motifs of both plus and minus polarity (i.e. the lower and the upper strand respectively; see Fig. 1). A relatively large number of mutations were observed in these sequences, but only the A-to-G mutation at position 11 of variant 261.3 affects one of the 13 conserved nucleotides of the catalytic core (Fig. 3B). All three essential helices of a hammerhead motif (stems I, II, and III) can be formed in each variant, and nucleotide covariation was detected for almost all base-pairs (Figs. 3A and B). Thus,

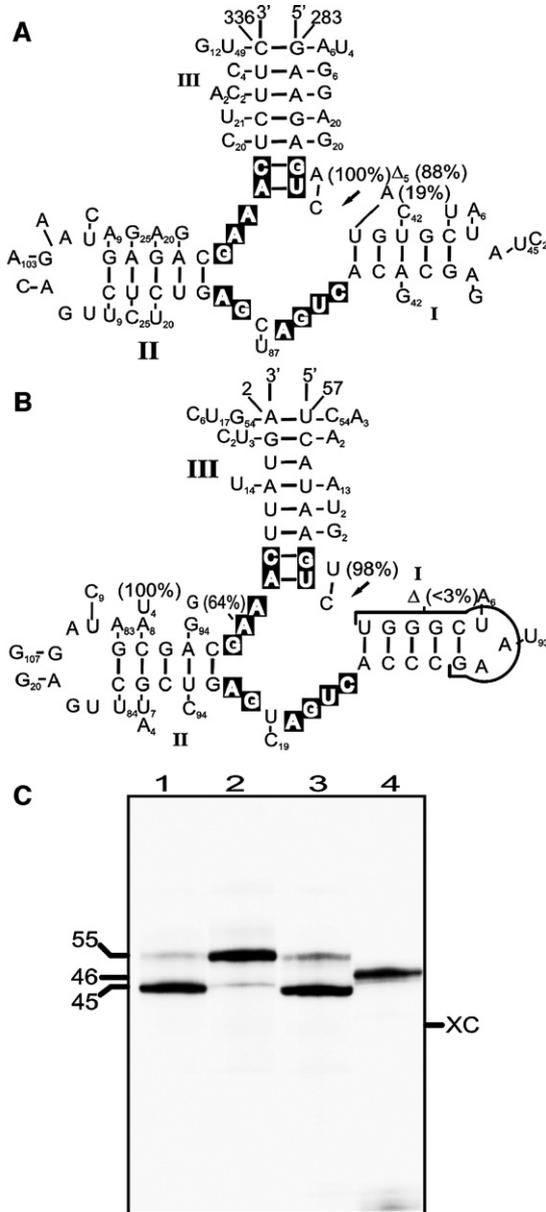


Fig. 3. Hammerhead structures and self-cleavage assays. (A, B) Nucleotide variations in the plus- and minus-polarity hammerheads, respectively. The nucleotides of the catalytic core are in black boxes and the stems are numerated (I, II and III). For mutations found more than once, the frequency of occurrence is indicated in subscript. For the tested variants, the relative percentage of self-cleavage is also indicated. Arrows show the self-cleavage sites. (C) Autoradiogram of a polyacrylamide gel run to test the hammerhead variants for self-cleavage. Lanes 1 to 4 correspond respectively to variants Ar1 (a positive control, to which a value of 100% is arbitrarily attributed), 52.4 (mutation U291A), 59.3 (deletion of U291), and 235.2 (deletion of 9 nt in stem I), respectively. Positions of molecular size markers in nucleotides and xylene cyanol (XC) are indicated.

the hammerhead sequences of both polarities in each variant can fold into their active secondary structure.

Since many of the detected mutations might influence the self-cleaving capability, several mutated minimal transcripts were synthesized *in vitro* from a T3 RNA polymerase promoter and their catalytic activities were assessed. As the transcription buffer contained 24 mM MgCl<sub>2</sub>, the RNA products self-cleaved

as they were synthesized. Fig. 3C illustrates a typical autoradiogram of a denaturing polyacrylamide gel used to visualize the self-cleavage activity of four transcripts (the activity is set arbitrarily at 100% for the positive control Ar1). When the uridine residue (U291) adjacent to the cleavage site was mutated to an adenosine (as observed in variant 52.4), the self-cleavage activity was drastically decreased to 19% (Fig. 3C, compare lanes 2 and 1). Yet surprisingly when U291 was absent (as in variant 59.3 and four others), self-cleavage was restored to 88% (Fig. 3C, lane 3). This contrasts with a previous study showing inefficient self-cleavage of PLMVd RNA transcripts displaying this deletion (Ambros et al., 1998). This discrepancy might be due to the fact that the hammerhead sequences differ at surrounding positions: selective pressure in favor of efficient self-cleavage should not preclude deletion of U291 if the surrounding nucleotides can compensate for the mutation. A mutant including a major deletion in stem I was found to lack self-cleaving activity (Fig. 3C, lane 4). This deletion could originate from the aberrant PLMVd replication or, alternatively, the reverse transcription step. In both cases, it resulted from an aberrant reaction in which the polymerase created the deletion. The presence of an atypical 2'-phosphate group or a 2'-5'-phosphodiester backbone at the ligation site (Côté et al., 2001), which corresponds to the hammerhead self-cleavage site, is probably involved in the formation of this deletion. Regardless how it is created, this PLMVd variant would have lost its potential to act as a template for further replication. Four other mutations were also tested and the results are compiled in panels A and B of Fig. 3. Briefly, all other mutants exhibited efficient self-cleavage (i.e. varying between 98% and 100%), with the exception of the variant having the A-to-G mutation at position 11 (i.e. within the GAAAC box of variant 261.3). In this case, only 64% self-cleavage was observed. Further investigations using various biophysical approaches should reveal whether or not some of these mutations contributed to modify the tertiary structure of some of the PLMVd hammerhead self-cleavage motif. Importantly, together these results highlight the strong selective pressure that must exist in favor of self-cleavage activity.

#### PLMVd variants belong to two groups

Using the sequence alignment of the 119 variants, we computed phylogenetic trees. Virtually identical trees were obtained with both softwares used (raw data not shown). A condensed version is presented in Fig. 4. Clustering of the variants yielded two groups with bootstrap values of 100%. More importantly, the results are supported by sequence and structural features. Briefly, between the two groups there is 85% homology. Group I comprises 20 sequences sharing at least 92% homology. Within it two subgroups, IA (16 variants) and IB (4 variants), can be distinguished on the basis of 18 informative positions (Fig. 4 and Supplementary Fig. 2). Group II includes 99 variants showing at least 93% homology. On the basis of 7 informative polymorphisms, this group can be split into three subgroups: IIA (10 variants), IIB (74 variants), and IIC (15 variants) (Fig. 4). For example, variants of subgroups

|     | Number of variants | Example | Specific informative alignment positions   | Structural features  |
|-----|--------------------|---------|--|--|
| IA  | 16                 | 143.4   | C1, U5 <sub>15</sub> , A54 <sub>15</sub> , A171, Δ194, G195,   | Ultrastable P10-L10 stem-loop<br>Pseudoknot between L11-L1                   |
|     |                    |         | A210, U228, A243, C244, G248, C258 <sub>15</sub> ,<br>U262, C266, C267, G284, U296, A304   |  |
| IB  | 4                  | 377.1   | A1, A5, U54, C171, U194, A195,<br>U210, A228, U243, A244, A248, U258,<br>C262, U266, A267, U284, C296, G304  | Ultrastable P10-L10 stem-loop<br>Pseudoknot between L11-L1                   |
| IIA | 10                 | 124.3   | U17, A24, A42, U178, A185, U299,<br>A321   | Pseudoknot between L11-L1  |
| IIB | 74                 | 66.3    | U17 <sub>73</sub> , A24, A42 <sub>19</sub> or U42 <sub>55</sub> , A178 <sub>70</sub> ,<br>U185 <sub>70</sub> , U299 <sub>18</sub> or A299 <sub>55</sub> , A321 <sub>71</sub> | (Pseudoknot between L11-L1) <sub>82</sub><br>(Insertion in L11) <sub>1</sub> |
| IIC | 15                 | 155.1   | C17, G24, U42 <sub>111</sub> , A178, U185, A299 <sub>13</sub> ,<br>G321  | (Insertion in L11) <sub>5</sub>  |

Fig. 4. Schematic phylogenetic tree clustering the variants into 2 groups and 5 subgroups. Numbers in subscript indicate the number of variants possessing a characteristic when not all of them possessed it.

IIA and IIB include an adenosine at position 24, where the members of subgroup IIC have a guanosine. Likewise, the members of subgroup IIA have a uridine in position 178 whilst those of subgroups IIB and IIC possess an adenosine at that position of the alignment (Supplementary Fig. 3). Five of the six variants displaying the 11- to 14-nt insertion in the L11 loop belong to subgroup IIC and the sixth one belongs to subgroup IIB.

The different groups also show specific structural features. On the one hand, all group-I variants possess a G<sub>270</sub>–C<sub>275</sub> base-pair at the top of the P10 stem, which is capped by a GUGA tetraloop. This combination yields a GNRA tetraloop that is proposed to be ultrastable (Varini, 1995). On the other hand, almost all group-II variants include an A<sub>270</sub>–U<sub>275</sub> base-pair at this position (except for 7 variants showing an insertion of A at position 275). Both conformations should provide a stable hairpin. The implications of different P10–L10 stem-loop structures remain elusive.

More generally, the subdivision proposed is applicable to all PLMVd sequence variants that can be retrieved from public sequences databases (Rocheleau and Pelchat, 2006). They all belong to group II except the Hd6 and Hd8 variants (GenBank accession nos. AF170501 and AF170503). As 90% homology is observed among all group-II variants described to date (here or elsewhere), they meet the criterion for being considered variants of a single species (Flores et al., 1998). However, only 85% homology is observed among the group-I variants. This is below the 90%-homology criterion proposed for distinguishing variants of a same species from different species. Yet all other physical and biological features of these variants indicate that they do belong to the PLMVd species. This suggests that the 90%-homology criterion is not adequate for distinguishing species. Structural similarity and biological properties, such as the ecological niche used for the virus classification (Mayo and Horzinek, 1998), might provide a better criterion for classification.

#### Evolution of PLMVd populations

The analysis of the informative positions for classifying variants into groups and subgroups showed that the subgroup-

IIB variants might have resulted from recombination events between subgroup-IIA and subgroup-IIC variants. Subgroup-IIB variants are nearly identical to subgroup IIA variants in a region comprising almost the entire P11 stem, whereas they share high homology with subgroup-IIC variants in the region beginning near the middle of the P11d stem (Supplementary Fig. 3 and Fig. 5A). The informative positions 42 and 299 in the alignment form a UA base-pair at the middle of stem P11d, which could be a potential recombination site. Interestingly, this proposed site is located near the self-cleavage sites and near one of the presumed initiation sites on the strands of both polarities (Delgado et al., 2005). This might favor the occurrence of recombination events. The recombination hypothesis is supported by the simultaneous detection of both subgroup-IIA and

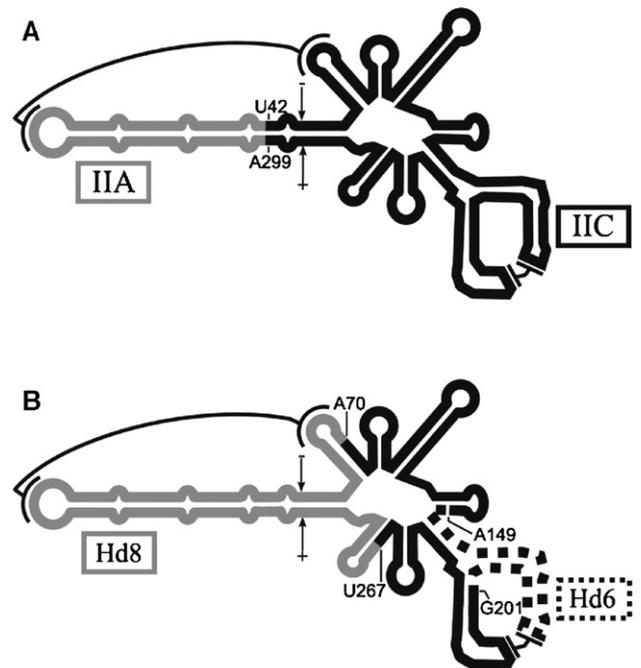


Fig. 5. Schematic representation of the hypothesized recombinations. (A) Between subgroup-IIA and -IIC variants, yielding subgroup IIB. (B) Between the Hd6 and Hd8 variants, yielding group I. Potential recombination sites are shown in black and correspond to the alignment nucleotides U42 and A299 in panel A and A70, A149, G201, and U267 in panel B.

subgroup-IIC variants (e.g. variants 88.2 and 88.1) in the same infected trees.

The situation of group I appears slightly different, as only two subgroups are distinguished. The Hd6 and Hd8 variants showed the least homology among all members of this group being at the extremities of the sequence spectrum. All other group-I variants show similarity to Hd8 in stems P10, P11 and P1 (regions 1 and 5), to Hd6 in stems P5 and P6 (region 3), and to both Hd6 and Hd8 in the regions spanning positions 70–149 (region 2) and 201–267 (region 4) (Supplementary Fig. 2 and Fig. 5B). Therefore it is tempting to hypothesize that variants in this group result from recombination events between previously coexisting Hd6 and Hd8 variants. This hypothesis is consistent with the detection of both Hd6 and Hd8 in the same peach tree (Hardired cultivar) in North America (Pelchat et al., 2000), a factor likely to favor recombination. Transport of propagated material from the USA to Tunisia might explain in part the existence of this group-I PLMVd population in Tunisia.

Intra- and interspecies recombination has been suggested as a motor of viroid evolution (e.g. see Amari et al., 2001; Stasys et al., 1995). Regarding PLMVd evolution, it would be very exciting to confirm with biological experiments the above hypotheses. Furthermore, it should be interesting to explore, in this viroid, the interplay between variability-generating mechanisms (recombination and error-prone replication) and the selective pressures limiting sequence divergence (e.g. the requirement for hammerhead self-cleavage activity and a branched structure). PLMVd now appears clearly as a suitable viroid for retracing the evolution of a RNA genome.

## Materials and methods

### *Viroid sources*

PLMVd-infected leaves were collected from 31 peach trees belonging to 14 different cultivars, 2 almond trees, and 1 pear tree growing in different regions of Tunisia: the north (Ben Arous, Bizerte, Manouba, Nabeul and Zaghouane), the center (Kairouane), and the Sahel (Mahdia). RNA was extracted from the leaves as reported previously (Fekih Hassen et al., 2006).

### *RT-PCR amplification, cloning, and sequencing*

RT-PCR amplifications of PLMVd RNA were carried out on RNA preparations as described previously (Fekih Hassen et al., 2006). The polymerase used for the amplification is the High Fidelity enzyme of the Titan one tube RT-PCR system (Roche Applied Biosciences). A control reaction without RNA template was always amplified with the samples in order to check for contaminating DNA amplicons from previous PCR. Primers are shown in Fig. 1. For cloning and sequencing purposes, RT-PCR products were purified by agarose gel electrophoresis using the QIAEX II® Gel Extraction Kit (Qiagen, KJ Venlo, Netherlands) and directly cloned into the pCR® 2.1 cloning vector with the help of the TA Cloning Kit (Invitrogen, Groningen, the Netherlands), according to the manufacturer's recommendations. Two to four PLMVd clones

from each isolate were sequenced in both directions by the dideoxyribonucleotide chain termination method using the BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). The nucleotide sequences were reported to the Genbank Nucleotide Sequence Database under accession numbers DQ680688 to DQ680806. Characteristics of the PLMVd variants are reported in Supplementary Table 1.

### *Sequence analysis*

Multiple sequence alignments were generated with the Clustal W program, version 1.8 (Thompson et al., 1994). Minor adjustments were introduced manually in the final alignment to maximize the sequence homology. Phylogenetic trees were obtained with both the DNAMAN software, version 5.2.2 (Lynnon Biosoft) and the Genebee software (<http://www.genebee.msu.su>). Bootstrap analyses based on 100 replicates were performed. The most stable secondary structures (having the lowest free energy) were predicted with the Mfold program (Zuker, 1989).

### *RNA synthesis and self-cleavage assays*

Eight hammerhead RNAs were synthesized by run-off transcription from overlapping oligonucleotides corresponding to the T3 RNA polymerase promoter followed by the self-cleaving sequences (see Supplementary Table 3 for detailed sequences). After annealing of the oligonucleotides (2 µM each), double-stranded DNA was obtained in a final volume of 100 µl reaction mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 24 mM MgCl<sub>2</sub>, 0.05 mM of each dNTP, 2 µM of T3 primer and 2.5 U Taq DNA polymerase. The resulting products were precipitated with ethanol and used as templates for subsequent *in vitro* transcriptions as previously described (Nehdi and Perreault, 2006). Briefly, the DNA templates were incubated in a total volume of 50 µl containing 80 mM HEPES-KOH (pH 7.5), 24 mM MgCl<sub>2</sub>, 2 mM spermidine, 40 mM DTT, 5 mM of each NTP, and 10 µCi [ $\alpha$ -<sup>32</sup>P]UTP (3000 Ci/mmol; New England Nuclear), 37 U RNA Guard (Amersham Biosciences), 0.004 U/µl pyrophosphatase (Roche Diagnostics), and 10 µg purified T3 RNA polymerase for 2 h at 37 °C. Upon completion, the reaction mixtures were treated with DNase RQ1 (Amersham Biosciences) at 37 °C for 10 min. The RNA was then purified by phenol/chloroform extraction, ethanol precipitation, and fractionation by denaturing (7 M urea) 15% polyacrylamide gel electrophoresis (19:1 ratio of acrylamide to bisacrylamide) in a buffer containing 45 mM Tris-borate (pH 7.5), 7 M urea, and 1 mM EDTA. Reaction products were visualized with a radioanalytic scanner (PhosphorImager).

## Acknowledgments

This work was supported by grants from the “Commissariat Général aux Relations Internationales” (CGRI) of the Communauté française Wallonie-Bruxelles (Belgium), from the “Ministère Tunisien de l'Enseignement Supérieur de la Recherche

Scientifique et de la Technologie” and the “Groupement Obligatoire des Viticultures et Producteurs des Fruits” (GOVPF) (Tunisia), to M.H.J., and from the Natural Sciences and Engineering Research Council (NSERC) of Canada to J.P.P. The authors are grateful to J. Perreault, A. Nehdi, and F. Bolduc for analysis and technical assistance. J.P.P. holds the Canada Research Chair in Genomics and Catalytic RNA.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.virol.2006.10.021](https://doi.org/10.1016/j.virol.2006.10.021).

## References

- Amari, K., Gomez, G., Myrta, A., Di Terlizzi, B., Pallas, V., 2001. 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. *J. Gen. Virol.* 82, 953–962.
- Ambros, S., Hernandez, C., Desvignes, J.C., Flores, R., 1998. Genomic structure of three phenotypically different isolates of peach latent mosaic viroid: implications of the existence of constraints limiting the heterogeneity of viroid quasi-species. *J. Virol.* 72, 7397–7406.
- Ambros, S., Hernandez, C., Flores, R., 1999. Rapid generation of genetic heterogeneity in progenies from individual cDNA clones of peach latent mosaic viroid in its natural host. *J. Gen. Virol.* 80, 2239–2252.
- Beaudry, D., Busiere, F., Lareau, F., Lessard, C., Perreault, J.P., 1995. The RNA of both polarities of the peach latent mosaic viroid self-cleaves in vitro solely by single hammerhead structures. *Nucleic Acids Res.* 23, 745–752.
- Bussière, F., Ouellet, J., Côté, F., Lévesque, D., Perreault, J.P., 2000. Mapping in solution shows the peach latent mosaic viroid to possess a new pseudoknot in a complex, branched secondary structure. *J. Virol.* 74, 2647–2654.
- Côté, F., Lévesque, D., Perreault, J.P., 2001. Natural 2',5'-phosphodiester bonds found at the ligation sites of peach latent mosaic viroid. *J. Virol.* 75, 19–25.
- Daros, J.A., Elena, S., Flores, R., 2006. Viroids: an Ariadne's thread into the RNA labyrinth. *EMBO Rep.* 7, 593–598.
- Delgado, S., Martinez de Alba, A., Hernandez, C., Flores, R., 2005. A short double stranded RNA motif of Peach latent mosaic viroid contains the initiation and the self-cleavage sites of both polarity strands. *J. Virol.* 79, 12934–12943.
- Desvignes, J.C., 1986. Peach latent mosaic and its relation to peach mosaic and peach yellow virus diseases. *Acta Hort.* 193, 51–57.
- Faggioli, F., Loreti, S., Barba, M., 1997. Occurrence of peach latent mosaic viroid (PLMVd) on plum in Italy. *Plant Dis.* 81, 423.
- Fekih Hassen, I., Kummert, J., Marbot, S., Fakhfakh, H., Marrakchi, M., Jijakli, M.H., 2004. First report of Pear blister canker viroid, Peach latent mosaic viroid and Hop stunt viroid infecting fruit trees in Tunisia. *Plant Dis.* 88, 1164.
- Fekih Hassen, I., Marbot, S., Kummert, J., Fakhfakh, H., Marrakchi, M., Jijakli, M.H., 2005. Peach latent mosaic viroid detected for the first time on almond trees in Tunisia. *Plant Dis.* 89, 1244.
- Fekih Hassen, I., Roussel, S., Kummert, J., Fakhfakh, H., Marrakchi, M., Jijakli, M.H., 2006. Development of a rapid RT-PCR test for the detection of Peach latent mosaic viroid, Pear blister canker viroid, Hop stunt viroid and Apple scar skin viroid in fruit trees from Tunisia. *J. Phytopathol.* 153, 217–223.
- Flores, R., Randles, J.W., Bar-Joseph, M., Diener, T.O., 1998. A proposed scheme for viroid classification and nomenclature. *Arch. Virol.* 143, 623–629.
- Hadidi, A., Giunchedi, L., Shamloul, A.M., Poggi-Pollini, C., Amer, M.A., 1997. Occurrence of peach latent mosaic viroid in stone fruits and its transmission with contaminated blades. *Plant Dis.* 81, 154–158.
- Hernandez, 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.
- Kyriakopoulou, P.E., Giunchedi, L., Hadidi, A., 2001. Peach latent mosaic and pome fruit viroids in naturally infected cultivated pear (*Pyrus Communis*) and wild pear *P. Amygdaliformis*: implications on possible origin of these viroids in the mediterranean region. *J. Plant Pathol.* 83, 51–62.
- Llacer, G., 1998. General aspects of peach latent mosaic disease. *Acta Hort.* 472, 561–564.
- Malfitano, M., Di Serio, F., Covelli, L., Ragazzino, A., Hernandez, C., Flores, R., 2003. Peach latent mosaic viroid variants inducing peach calico (extreme chlorosis) contain a characteristic insertion that is responsible for this symptomatology. *Virology* 313, 492–501.
- Mayo, M.A., Horzinek, M.C., 1998. A revised version of the international code of virus classification and nomenclature. *Arch. Virol.* 143 (8), 1645–1654.
- Nehdi, A., Perreault, J.P., 2006. Unbiased in vitro selection reveals the unique character of the self-cleaving antigenomic HDV RNA sequence. *Nucleic Acids Res.* 34, 584–592.
- Pelchat, M., Lévesque, D., Ouellet, J., Laurendeau, S., Lévesque, S., Lehoux, J., Thompson, A., Eastwell, K.C., Skrzeczkowski, J., Perreault, J.P., 2000. Sequencing of Peach Latent Mosaic Viroid variants from nine North American peach cultivars shows that this RNA folds into a complex secondary structure. *Virology* 271, 37–45.
- Rocheleau, L., Pelchat, M., 2006. The Subviral RNA Database: a toolbox for viroids, the hepatitis delta virus and satellite RNAs research. *BMC Microbiol.* 6, 24.
- Rodio, M.E., Delgado, S., Flores, R., Di Serio, F., 2006. Variants of peach latent mosaic viroid inducing peach calico: uneven distribution in infected plants and requirements of the insertion containing the pathogenicity determinant. *J. Gen. Virol.* 87, 231–240.
- Shamloul, A.M., Minafra, A., Hadidi, A., Giunchedi, L., Waterworth, H.E., Allam, E.K., 1995. Peach latent mosaic viroid: Nucleotide sequence of an Italian isolate, sensitive detection using RT-PCR and geographic distribution. *Acta Hort.* 386, 522–530.
- Stasys, R.A., Dry, I.B., Rezaian, M.A., 1995. The termini of a new citrus viroid contain duplications of the central conserved regions from two viroid groups. *FEBS Lett.* 358, 182–184.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Varini, G., 1995. Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct.* 24, 379–404.
- Zuker, M., 1989. On finding all suboptimal foldings of an RNA molecule. *Science* 244, 48–52.