


Conserved Motifs and Domains in Members of *Pospiviroidae*

Kevin-Phil Wüsthoff [†] and Gerhard Steger ^{*,†} 

Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany; Kevin.Wuesthoff@hhu.de

* Correspondence: steger@biophys.uni-duesseldorf.de

† These authors contributed equally to this work.

Abstract: In 1985, Keese and Symons proposed a hypothesis on the sequence and secondary structure of viroids from the family *Pospiviroidae*: their secondary structure can be subdivided into five structural and functional domains and “viroids have evolved by rearrangement of domains between different viroids infecting the same cell and subsequent mutations within each domain”; this article is one of the most cited in the field of viroids. Employing the pairwise alignment method used by Keese and Symons and in addition to more recent methods, we tried to reproduce the original results and extend them to further members of *Pospiviroidae* which were unknown in 1985. Indeed, individual members of *Pospiviroidae* consist of a patchwork of sequence fragments from the family but the lengths of fragments do not point to consistent points of rearrangement, which is in conflict with the original hypothesis of fixed domain borders.

Keywords: pairwise sequence identity; hairpin I (HPI); hairpin II (HPII); loop E; RY motif; terminal conserved hairpin (TCH); terminal conserved region (TCR); NUCALN; JALI; VMATCH



Citation: Wüsthoff, K.-P.; Steger, G. Conserved Motifs and Domains in Members of *Pospiviroidae*. *Cells* **2022**, *11*, 230. <https://doi.org/10.3390/cells11020230>

Academic Editors: Ahmed Hadidi, Henryk Hanokh Czosnek, Kriton Kalantidis and Robert A. Owens

Received: 6 December 2021

Accepted: 7 January 2022

Published: 11 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Viroids are the smallest known plant pathogens with a genome length ranging from 246 to 401 bases in size depending on the viroid species (Table S1) and variant. They solely consist of a circular, single-stranded RNA [1], which does not code for any protein [2]. Thus, for all of their biological functions—such as replication, processing, and transport [3]—they have to present sequence or structural features to exploit host proteins. For example, viroids from the family *Pospiviroidae* are transcribed by DNA-dependent RNA polymerase II (pol II) using the viroid RNA as a template in an asymmetric rolling circle mechanism in the nuclei of infected cells [4–7]. The thermodynamically optimal, stable structure of most members of the family *Pospiviroidae* consists of an unbranched series of short helices and small loops (Figure 1); i.e., their sequence is an imperfect inverted repeat [8–10].

In 1985, Keese and Symons [11] proposed a model based on the secondary structure of viroids from the family *Pospiviroidae*: accordingly, their secondary structure can be subdivided into five structural and functional domains and “viroids have evolved by rearrangement of domains between different viroids infecting the same cell and subsequent mutations within each domain” [12]. These five domains, with precise borders [11–13], are as follows (Figure 1):

- The terminal left (TL) domain plays an important role in the replication of *Pospiviroidae* as it contains the starting point for pol II [14–17]. Furthermore, two subgroups of *Pospiviroidae* differ by presence and absence, respectively, of two sequence motifs in the TL domain [18–20]: members of genus *Pospiviroid* and some members of *Coleviroid* contain a terminal conserved region (TCR; see Supplemental Figure S20), while members of genera *Cocad-* and *Hostuviroid* contain a terminal conserved hairpin (TCH; Figure S21).
- The pathogenicity (P) domain is associated with symptom severity [21–23], but clearly not the only viroid part responsible for virulence (examples can be found in [22,24–27]).

The P domain includes an oligopurine stretch in the upper part and a partly complementary oligopyrimidine stretch in the lower part of the secondary structure. Thus, the region shows alternative structures of low thermodynamic stability, giving rise to the name “premelting region” [28].

- The central (C) domain is the most conserved part between different viroids. It contains a loop E motif (Figure S22b), which shows similarities to loop E of eukaryotic 5S rRNA, sarcin/ricin loop in 28S rRNA, and loop B of hairpin ribozymes [29,30]: it consists of five non-Watson–Crick basepairs and a bulged nucleotide involved in a triple pair; this unusual conformation allows for the formation of a UV-induced crosslink between two nucleotides of loop E [31–33]. The upper part of the C domain can be rearranged into a hairpin (HPI; Figure S22a), which is only present in thermodynamic metastable structures during replication or at biologically non-relevant temperatures [34]. Loop E and HPI are both involved in the processing of linear replication intermediates to mature circles [35–39].
- The variable (V) domain has the lowest sequence similarity even between closely related species [11] but contains one part of hairpin II (HPII; Figure S25), which is a metastable structural element critical for the transcription of (–)-stranded replication intermediates in pospiviroids [40,41]. The 3′ part of HPII is located in the lower part of the TL domain.
- The terminal right (TR) domain of genus *Pospiviroid* has been proposed to be involved in transport. For example, the RY motifs in the TR domain are the binding sites for viroid RNA-binding protein 1 (VirP1) [42–44], which is indispensable for replication and cell-to-cell transport [45,46].

The domain model was established with seven sequences from different viroid species of family *Pospiviroidae* [11]. The borders between these domains “were defined by sharp changes in sequence homology” [12] between the aligned subgroups of the seven viroids. However, in the original publication, a description of the used procedure was not given [11]; a further publication [12] states the use of the alignment program NUCALN [47]. Here, we describe an attempt to redo the determination of the domain boundaries for the original seven sequences as well as extend it for more recently added *Pospiviroidae* species—today counting 30—by the use of three different programs:

1. NUCALN is no longer available, but its algorithm is still part of CLUSTALΩ [48] for the fast production of pairwise alignments. Thus, we modified CLUSTALΩ to output not only scores but also aligned positions from the fast alignment step (for details see Supplement and Figure S1). In the following, this modified CLUSTALΩ is called NUCALN.
2. JALI, short for “Jumping Alignments” [49,50], determines the positions of recombination in a candidate sequence when compared to a seed alignment of related sequences.
3. VMATCH [51] determines maximal substring matches between two sequences.

NUCALN as well as JALI were only able to reproduce a few of the domain borders for certain pairwise combinations of viroid sequences; however, in general, exact positions were not deducible neither for the original viroid sequences nor for more recent ones. We conclude that Keese and Symons used some undisclosed method to determine the domain borders. Such a process—most probably guided by human intuition—is difficult to implement in a computer program. Regardless, results from JALI and VMATCH hint to past recombination events between members of *Pospiviroidae*. In these events, sequence fragments were exchanged between viroids, supporting the observations of [11] about the possible existence of a sequence toolbox, which allows viroids to easily evolve by recombination between different viroids infecting the same plant.

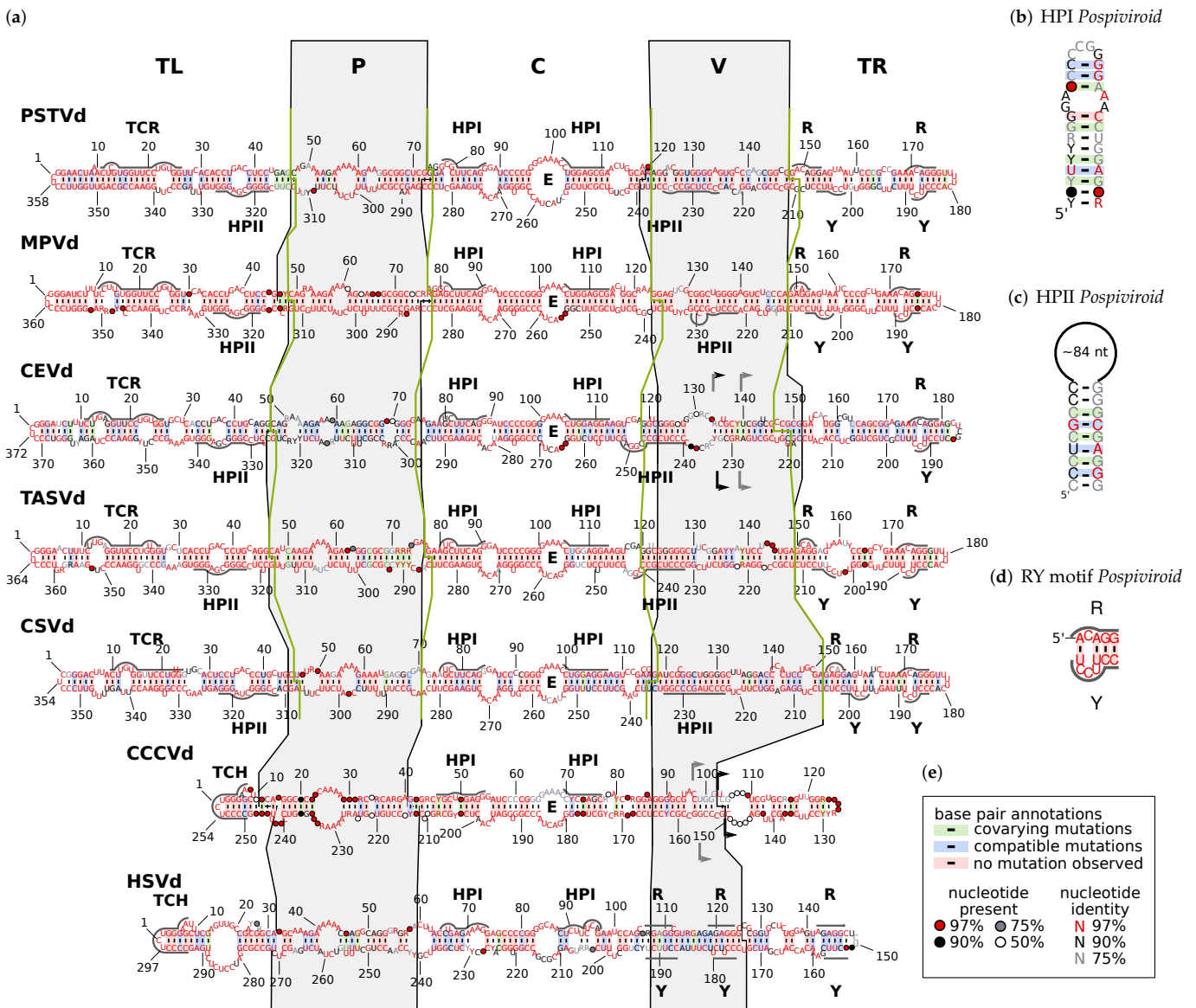


Figure 1. Consensus sequences and secondary structures of viroid species used by Keese and Symons [11,52] to determine the domain boundaries. Sequence alignments were produced by MAFFT X-INS-1 [53]. Consensus structures were generated by CONSTRUCT [54]. Consensus sequences and drawings were produced by R2R [55]. The color code used by R2R is explained in box (e); a co-varying position has Watson–Crick pairs that differ at both nucleotide positions among sequences; if only one position differs, the occurrence is classified as a compatible mutation [56]. Note that no basepair annotated as “covarying” or “compatible” is statistically significant according to the analysis by R-SCAPE [57]. (a) Domain names are given at top: TR, terminal right; P, pathogenicity; C, central; V, variable; TL, terminal left. The P and V domains given by [52] are marked by black lines and gray background; the “consistent” domains are marked by green lines (see Section 3.2). For viroid names, see Table S1. E, loop E (Figure S22b); HPI, the two sequence regions form hairpin I in metastable structures of *Pospiviroidae* sequences ((b) for further details see Figures S22a and S23; HPII, the two sequence regions form hairpin II in metastable structures of *Pospiviroid* sequences ((c) and Figure S25); RY, binding motifs for VirP1 of *Pospiviroid* sequences ((d) and Figure S26); TCR, terminal conserved region (Figure S20); TCH, terminal conserved hairpin (Figure S21). Flag pairs in CEVd and CCCVd mark examples of sequence stretches that are duplicated in natural variants [58–62].

2. Materials and Methods

Software tools used to produce the results are listed in Table 1. GLE was used to visualize the results of JALI and VMATCH.

Table 1. Programs and Web tools. All web pages were accessed on 1 November 2021.

Program	Version	URL	Ref.
BLASTN	2.9.0	https://blast.ncbi.nlm.nih.gov/Blast.cgi	[63]
CLUSTALΩ	1.2.4	http://www.clustal.org/omega/#Download	[48]
CONSTRUCT	3.2.2	http://www.biophys.uni-duesseldorf.de/html/local/construct3/	[54]
GLE	4.2.5	https://sourceforge.net/projects/glx/	[64]
JALI	1.3	https://bibiserv.cebitec.uni-bielefeld.de/jali	[49,50]
MAFFT X-INS-I, L-INS-I	7.402	https://mafft.cbrc.jp/alignment/software/	[53,65,66]
R2R	1.0.5	https://sourceforge.net/projects/weinberg-r2r/	[55]
RNAFOLD	2.4.6	https://www.tbi.univie.ac.at/RNA/	[67]
R-SCAPE	1.2.3	http://eddylab.org/R-scape/	[57]
VMATCH	2.3.0	http://www.vmatch.de/	[51]

2.1. Consensus Sequence Compilation

Viroid reference sequences from *Pospiviroidae* were downloaded from the Viral Genomes database [68,69]; further members of each viroid were found via BLASTN in NCBI's "standard databases". A preliminary alignment of sequences from a single viroid species was produced via MAFFT L-INS-I. Based on this alignment, incomplete sequences were removed; if necessary, sequences were reverse complemented to (+)-strands, the start and end of sequences were adjusted to standard positions, and redundant sequences were removed. Then, a final alignment was produced via MAFFT X-INS-I with options `--maxiterate 1000` and `--retree 100`. MAFFT X-INS-I takes into account structural information for the alignment; the iterative refinement and the guide tree prediction are repeated until no more improvement in the final score is made or the number of cycles reaches 1000 and 100, respectively. For each of the final alignments, a consensus sequence and structure were calculated via CONSTRUCT using RNAFOLD with options `-p -c` (partition folding, circular RNA). Drawings of the consensus structures were performed via R2R.

2.2. NUCALN

The algorithm, implemented in NUCALN [47], is still part of CLUSTALΩ for the fast production of pairwise alignments. Thus, we modified CLUSTALΩ to output not only the scores but also the k -tuple matches from the fast alignment step. For details, see Supplementary Materials.

NUCALN searches for sequence stretches of at least k nucleotides in length that are identical between sequences x and y ; these stretches are called k -tuple matches. In a dotplot of sequences x and y , the k -tuple matches are diagonals of at least k dots (Figure S1). That is, each k -tuple match lies on a diagonal d_m containing all nucleotide pairs x_i and y_j with $m = i - j$. A diagonal is significant if its number of k -tuple matches is above the mean number of k -tuple matches for all diagonals with matches. The k -tuple matches that occur in a window of length w around a significant diagonal lie in window space. Finally, from the k -tuple matches within window space, an alignment is produced via standard dynamic programming with scores $+1$ for each matching nucleotide pair of the k -tuple matches and $-g$ for gaps in between the matches independent of the gap length. We slightly varied the parameters $2 \leq k \leq 6$ and $1 \leq g \leq 4$ around those used in [12] ($k = 4, g = 4$); these variations did not alter the results. The window length $w = 100$ used by [12] is clearly too large, because with members of *Pospiviroidae*, most k -tuple matches are quite close to the main diagonal; thus, we chose $w = 25$ to reduce the runtime of NUCALN and reduce the number of uninformative matches.

2.3. JALI

The parameters of JALI were optimized to closely reproduce the domain boundaries as given by Keese and Symons [52]. First, a sequence file with all seven sequences (PSTVd, GENBANK AC NC_002030 or PTVA; MPVd, TPM; CEV-A, CEV; TASVd, TASCg; CSVd, V01107; CCCVd, CCC; HSVd, X00009) given in [11,52] was assembled. A PERL script then selected one of the sequences in a loop as the candidate and then repeatedly called JALI to align the candidate sequence to the seed alignment using a range of scoring parameters (gap open $-70 \leq o \leq -40$, gap extension $-6 \leq e \leq -2$, jump $-125 \leq j \leq -45$); for match scoring, we used the RIBOSUM weights w [70] multiplied by 10, which gives a range $w(A|C) = 6$ to $w(A|A) = 47$. The positions of jumps were recorded and compared to the domain borders as given in [52] to find optimal parameter combinations. We finally used $o = -60$, $e = -2$, and $j = -95$. A toy example with JALI is shown in Figure S2; examples of JALI's output with different parameters are shown in Figure S3.

2.4. VMATCH

VMATCH is a tool that finds, for example, long matches between two sequences or repeats in a single sequence. We used the following options for a comparison of the two sequences:

- d: report (only) direct matches;
- l 50: report matches of length ≥ 50 ;
- e 20: edit distance—only report matches that contain at most 20 mismatches, insertions or deletions;
- leastscore 50: only report matches if their score is ≥ 50 ;
- seedlength 4: length of exact seeds.

2.5. Pairwise Sequence Identity

We calculated pairwise sequence identity (PSI; in percent) for two aligned sequences by

$$\text{PSI} = \frac{\text{number of nucleotide matches}}{\min(\text{length of non-aligned sequences})} \cdot 100;$$

that is, the number of nucleotides was normalized to the length of the shorter of both sequences without gaps. There are other methods to calculate PSI, that all have some disadvantages. For example, normalization to the average length of both sequences, as used in [52], assigns artifactually low PSI values to an alignment of a short and a long sequence. This is of relevance because members of *Pospiviroidae* have sequence lengths in a wide range from 250 to 380 nt.

To determine the PSI of a domain X with sequence ranges $i \dots j$ and $k \dots l$ for viroid v , we extracted from a pairwise alignment of viroids v and w the sequence blocks $i \dots j$ and $k \dots l$, calculated PSI for the individual blocks, and then averaged both values. If the domain ranges of both viroids differ from each other in the alignment, $\text{PSI}_{X,v,w}$ will differ from $\text{PSI}_{X,w,v}$.

Average pairwise sequence identity (APSI; in percent) is calculated for N aligned sequences by

$$\text{APSI} = \frac{2}{N(N-1)} \cdot \sum_{i=1}^{N-1} \sum_{j=i+1}^N \text{PSI}_{ij}.$$

3. Results and Discussion

3.1. Can We Retrace the Results of Keese and Symons?

Keese and Symons [11,52] based their domain hypothesis on pairwise alignments of single sequences from seven viroid species, namely PSTVd, MPVd, CEVd, TASVd, CSVd (all belonging to the genus *Pospiviroid*), CCCVd (genus *Cocadviroid*), and HSVd (genus

Hostuviroid; for full viroid names and genus affiliation see Table S1). Hence, we started with these sequences and aligned them using NUCALN with parameters as mentioned in [12].

In principle, one should compare all 21 pairwise alignments of the seven viroid sequences. As examples, pairwise alignments of PSTVd to the other six viroids are shown in Supplemental Figure S4. Each alignment can also be visualized as a dotplot; an example is shown in Figure 2a, and the further dotplots of PSTVd with the other six viroid sequences are shown in Figure S5. To ease this comparison, we show in Figure 2b an overlay of all optimal *k*mers from the 21 pairwise alignments produced with NUCALN on the basis of a MAFFT X-INS-I alignment. Figure 2d differs from Figure 2b by the use of consensus sequences; the latter should avoid any misjudgment due to a pathological choice of sequences. The dotplots of Figure 2b,d, however, are very similar to each other and their small differences are not of importance. The following results and discussions are based on pairwise sequence identities of these viroids; for an overview, we already refer to Table 2.

Table 2. Sequence homology between the domains of different viroids. Values in columns labeled “KS” are copied from Keese and Symons (1987) [52]. The table in Keese and Symons (1985) [11] only differs by the missing values for the V and TR domains of HSVd. Pairwise sequence identity (PSI) values in columns labeled “WS” are calculated on basis of a MAFFT X-INS-I alignment with options `--maxiterate 1000` and `--retree 100`. Bold values in columns V and TR point to viroid pairs with large differences in the PSI values of these domains.

Viroids Used for Pairwise Comparisons		% Sequence Homology											
		Overall		Domains									
				TL		P		C		V		TR	
1	2	KS	WS	KS	WS	KS	WS	KS	WS	KS	WS	KS	WS
TASVd	CEVd	73	79.4	91	93.4	54	63.3	99	97.9	49	59.2	46	61.9
	PSTVd	64	69.9	67	68.1	59	69.1	65	74.2	30	44.9	90	92.1
MPVd	PSTVd	76	80.2	67	76.6	73	82.1	94	96.8	42	42.3	95	95.2
	CEVd	60	66.1	80	76.1	70	62.5	69	78.3	29	42.3	37	61.9
CCCVd	PSTVd	38	59.3	25	52.2	14	47.5	70	71.6	37	57.1	27	53.7
	HSVd	39	56.5	52	82.6	33	39.0	42	63.2	31	46.4	50	58.5
HSVd	PSTVd	35	53.9	23	49.1	58	71.2	35	52.6	37	43.2	28	51.1
CSVd	CEVd	59	67.4	77	80.9	42	57.4	82	91.2	28	36.7	38	60.4
	PSTVd	61	69.7	69	71.9	49	64.8	71	78.9	31	46.6	81	92.5
CEVd	PSTVd	55	64.6	62	70.0	71	67.9	65	74.2	31	43.4	38	60.0

Most of the conclusions presented in [11,12,52] are comprehensible from these dotplot overlays (Figure 2b,d). These conclusions are given verbatim below; changes due to more recent nomenclature are given in square brackets:

- The V domain shows the “greatest variability between closely related viroids” [12] and “is the most variable region . . . between otherwise closely related viroids, such as between TASV[d] and CEV[d] or TPMV [MPVd] and PSTV[d]” [11].
Indeed, only a few diagonals show up in the dotplots (Figure 2) in the V domain; minor similarities in the lower part of the V domain are due to the 3′ part of HP11. The average pairwise sequence identity (APSI) in a MAFFT X-INS-I alignment of consensus sequences of species used in [11] is 64%, while the APSI of the V domain in the same alignment is only around 49% (Table S4b).
- “In pairwise sequence comparisons of viroids containing highly homologous C domains . . . there is significantly less sequence homology in the P and V regions, occurring about 5–9 residues 5′ and about 7–15 residues 3′ of the inverted repeat” [11]. The number of *k*mer matches in the C domain regions is higher than in P and V regions; compare the dot colors of these domains in the overlay dotplots (Figure 2b,d).

The APSI of full-length sequences PSTVd, MPVd, CEVd, TASVd, CSVd, CCCVd, aligned with MAFFT X-INS-I, is 64% (Table S4a), their C, P, and V domains have $71 < \text{APSI} < 77\%$, $62 < \text{APSI} < 67\%$, and $48 < \text{APSI} < 50\%$, respectively (Table S4b). That is, for each of these viroids, the order $\text{APSI}(C) > \text{APSI}(P) > \text{APSI}(V)$ is valid.

- "...in comparisons of PSTV[d], TPMV [MPVd], TASV[d], and CSV[d], a change from low homology in the V domain to high homology in the T2 [TR] domain defines the boundary for these two domains" [11].

The overlay dotplots clearly show a higher number of *k*mer matches in the TR domain regions than in the V domain regions (Figure 2b,d). The TR domains have $\text{APSI} > 50\%$, while the V domains have $\text{APSI} < 50\%$ (Table S4b). Only the mentioned viroids have much larger differences in their individual PSI values with $\text{PSI}(V) \approx 45\%$ and $\text{PSI}(\text{TR}) > 92\%$ (see values in bold font in columns V and TR of Table 2). Other viroid pairs have more similar PSI values of their V and TR domains, which is not sufficient to define their V/TR borders.

- "The P domain, with a conserved oligo(A) sequence flanked by regions with greater variability, has its borders based on homologies between the P region of HSV[d] and other viroids such as PSTV[d] and by certain pairwise comparisons such as CEV[d]-A and TASV[d] in which there is significant change from relatively low sequence homology in the P region to higher homology in the adjacent T1 [TL] and C domains" [11]. Pairwise dotplots produced by NUCALN of HSVd with PSTVd and CEVd with TASVd, mentioned in the above sentence, are shown in Figure S5f,h. Here, as well as in further dotplots (f. e. Figure 2), the oligo-purine sequence in the upper and the partially complementary oligo-pyrimidine sequence in the lower P domain regions give rise to long *k*mer diagonals. In alignments (f. e. Figures S4 and S7), these sequences are also easily detectable.
- "In addition, sequence data show that three viroids (TASV[d], TPMV [MPVd], and CCCV[d]) exhibit unusual relationships with respect to their terminal sequences. For example, TASV[d] shares 73% overall sequence homology with CEV[d]-A but the T2 [TR] domains are only 46% homologous ... In contrast, TASV[d] shares less overall sequence homology with PSTV[d] (64%) but the T2 [TR] domains are highly homologous (90%). Therefore, TASV[d] appears to be a recombinant between the T2 [TR] domain of a PSTV[d]-like viroid and all but the T2 [TR] domain of a CEV[d]-like viroid" [11].

Clearly, several viroid sequences show PSI values of individual domains largely deviating (>20%) from the PSI values of their full sequences (Tables S2 and S3). One of the examples is the relation between TASVd, CEVd, and PSTVd (or MPVd), mentioned above: the full-sequence PSI values of pairs TASVd/CEVd, TASVd/PSTVd, and TASVd/MPVd are around 80%, 71%, and 76%, respectively; in contrast, the TR PSI values are around 57%, 91%, and 97%, respectively. That is, despite the highest full-sequences PSI of TASVd/CEVd, their TR PSI is much lower, and the opposite is true for the TASVd/PSTVd and TASVd/MPVd pairs, which led Keese and Symons to suggest that TASVd is recombinant of a CEVd-like sequence and a TR domain of a PSTVd- (or MPVd)-like sequence.

- "TPMV [MPVd] shares 76% overall sequence homology with PSTV[d] but the T1 [TL] domains are less homologous (67%). In contrast, TPMV [MPVd] shares less overall sequence homology with CEV[d]-A (60%) but the T1 [TL] domains are more homologous (80%). Thus, TPMV [MPVd] appears to be a recombinant between the T1 [TL] domain of a CEV[d]-like viroid and all but the T1 [TL] domain of a PSTV[d]-like viroid" [11].

According to alignments by NUCALN with $k = 4$, $g = 4$ (Table S2) or $k = 3$, $g = 3$ (Table S3), and by MAFFT X-INS-I (Table 2), the PSI values of full-sequence alignments for MPVd and PSTVd are around 77%, 82%, and 80%, respectively, and for MPVd and CEVd, are around 66%, 67%, and 66%, respectively; the PSI values of TL domains are lower with around 68%, 75%, and 77% for MPVd and PSTVd, and higher with

around 75%, 77%, and 76% for MPVd and CEVd. Despite the difference between our values and the values from [11] (see below), all values support the given conclusion by Keese and Symons.

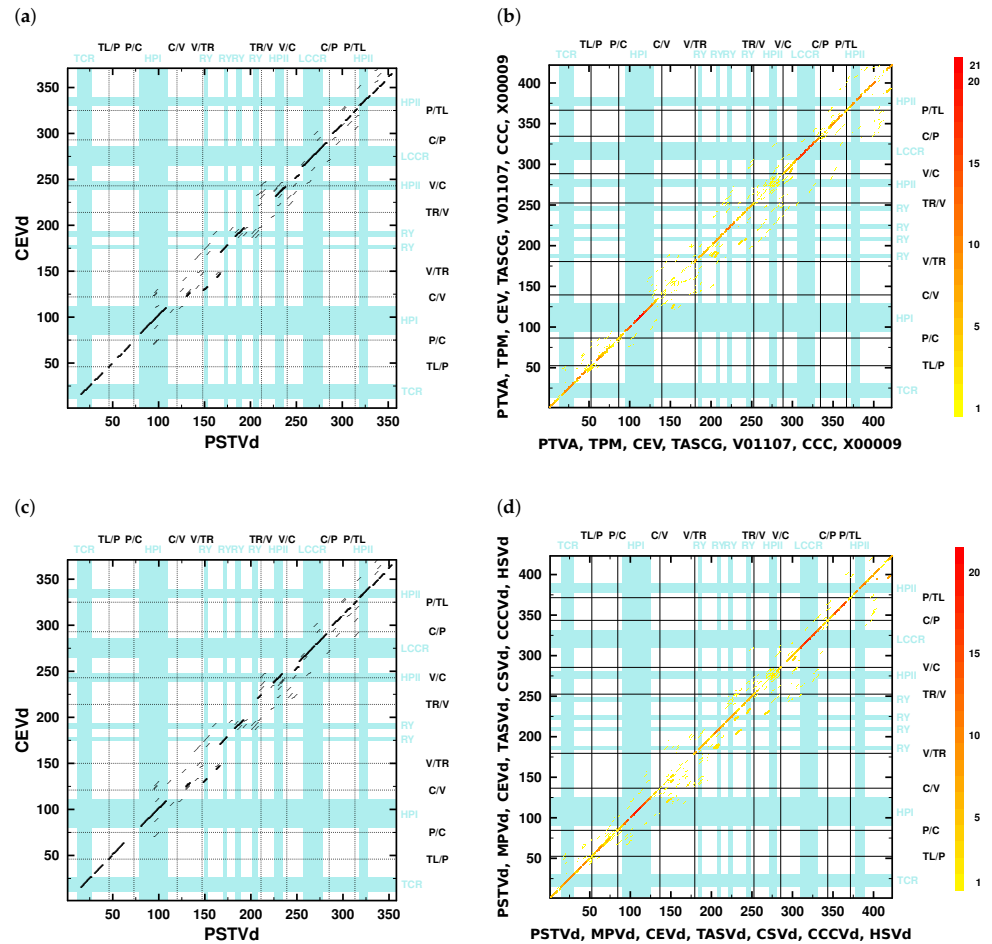


Figure 2. Dotplots of single sequences (left) and dotplot overlay of the seven sequences used by Keese and Symons [11,52] to establish the domain boundaries. (a,c): gray and black diagonal lines show regions with at least four consecutive nucleotides that match between the two sequences, as black lines show the best overall matches according to NUCALN with parameters $k = 4$, $g = 4$ and $w = 25$. In (a), the original sequences were used (PSTVd, GenBank Locus PTVA, NC_002030; CEVd, CEV); in (c), consensus sequences for PSTVd (274 sequences) and CEVd (217 sequences) were used. (b,d): Overlay of NUCALN best matches ($k = 4$, $g = 4$, $w = 25$) on a MAFFT X-INS-I alignment of the original seven sequences ((b); PSTVd, PTVA, NC_002030; MPVd, TPM; CEVd, CEV; TASVd, TASCG; CSVd, V01107; CCCVd, CCC; HSVd, X00009) and the corresponding seven consensus sequences ((d); 274 PSTVd sequences, 15 MPVd, 217 CEVd, 24 TASVd, 215 CSVd, 13 CCCVd, 433 HSVd), respectively. That is, plot (b,d) show an overlay of the dotplots of Figure S5 and of Figure S6, respectively. The number of coinciding matches is color-coded from red ($7 \times 6/2 = 21$) to yellow (1). Domain boundaries according to Keese and Symons [52] are marked by black lines (TL, terminal left; P, pathogenicity; C, central; V, variable; TR, terminal right); in (c,d), boundaries for PSTVd are shown. Blue bars and labels mark TCR (Figure S20), HPI (Figures 1b and S22a), LCCR (Figure S22b), HPII (Figures 1c and S25), and RY motifs (Figures 1d and S26).

In summary, the similarities and relationships between the members of *Pospiviroidae* are comprehensible. However, we were not able to exactly reproduce the PSI and APSI values of Keese and Symons. This is not simply due to different equations (see Section 2.5) used here and in [12]: we tried different approaches to calculate PSI, as well as different

NUCALN parameters to create the basic alignments, all without reaching coincident results. Further possibilities, which we did not explore, include the alignments of domains instead of full-length sequences or the manual optimization of non-aligned regions in between the aligned *kmers* of NUCALN. Anyway, our PSI and APSI values are also in support of the above conclusions. More critical than the exact reproduction of PSI and APSI values, however, is the missing consistency of domain borders in our alignments of different pospiviroids and a missing idea of how to define these borders.

3.2. Consistent Domain Borders of Pospiviroid Members

Without having a proper definition for a domain border, we can still improve the position of borders given in [11,12] to be consistent between species of *Pospiviroids*. That is, the domain borders should align between species in an alignment of all *Pospiviroid* species (Figure S7). As a starting point, we used the original domain borders of PSTVd (Figure S7, top bars for PSTVd and background shading of all sequences). In the aligned sequences of most species, gaps are located close to the original borders. We positioned the new borders (Figure 1, green lines; Figure S7, top) into these gap positions, which was equivalent to a shift of original borders by less than six nucleotides. Table S5 summarizes the positions of the now consistent domain borders and of the sequence stretches around the borders in consensus sequences and in more than 90% of individual sequences, which should help identify domain borders in new sequences.

We selected the new domain borders only for *Pospiviroid* members. The inclusion of sequences from the other genera of *Pospiviroidae* led to alignments with a high number of gaps because *Hostu-* and *Cocadviroid* sequences are a lot shorter than *Pospiviroid* sequences and *Apscaviroid* sequences have a low similarity to *Pospiviroid* sequences. This is also obvious from dotplots; f. e., see the dotplots of PSTVd vs. CCCVd (Figure S5e) and PSTVd vs. HSVd (Figure S5f). The seven *Coleviroid* members consist of combinations of six different structural elements with a common CCR (Figure S15). The low number of sequences for the structural elements (two to three) does not allow us to conclude on domains.

3.3. Can We Extend the Domain Hypothesis to the Other Genera of Pospiviroidae?

With descriptions of further viroid species, beyond those used to establish the domain model [11], the model was extended to these new sequences. The additional species were CTiVd [71] from genus *Cocadviroid* and ASSVd [13], GYSVd-1 [13], and GYSVd-2 [72] from genus *Apscaviroid* (Table S1). Today, ten species belong to the genus *Apscaviroid* and a further ten have been preliminary assigned to *Apscaviroid*. An overlay of all optimal *kmers* from the $20 \times 19/2 = 190$ pairwise alignments produced with NUCALN on the basis of a MAFFT X-INS-I alignment is shown in Figure 3; alignment, consensus sequences, and structures of the ten officially accepted *apscaviroids* are shown in Figures S10 and S11, respectively.

In the dotplot (Figure 3), only two highly conserved motifs are clearly detectable. These are the TCR (Figure S10b) and the CCR including HPI (Figures S23 and S24). An oligo-purine and the partially complementary oligopyrimidine stretch in the P domain [13,18] are obvious in an alignment (Figure S10) but only marginally in the dotplot. The absence of further motifs, such as HPII and RY in *Pospiviroid* members, makes it difficult to deduce any domain borders.

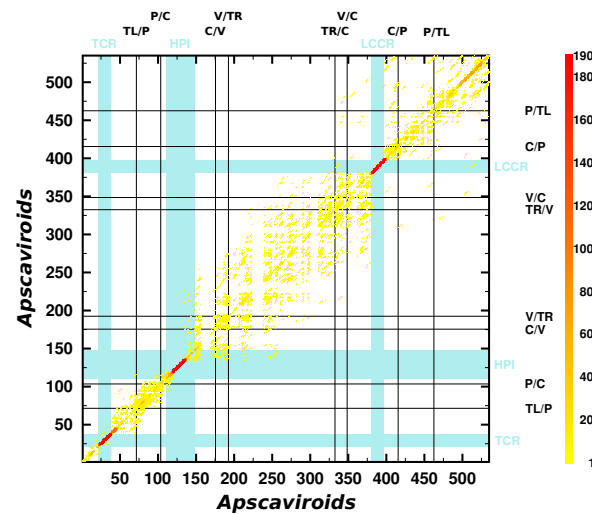


Figure 3. Overlay of NUCALN matches on a MAFFT X-INS-I alignment of all members of *Apscaviroid*. Parameters for NUCALN were $k = 4$, $g = 4$, and $w = 25$. In the alignments, consensus sequences of ACFSVd (1 sequence), ADFVd (28 sequences), AFCVd (75), AGVd (157), ASSVd (118), CBLVd (52), CDVd (130), CVd-V (19), CVd-VI (18), CVd-VII (4), DVd (17), GLVd (3), GYSVd-1 (109), GYSVd-2 (70), GYSVd-3 (23), LVd (1), PBCVd (61), PIVd-I (39), PVd-2 (2), and PVd (3) were used (Table S1). Domain borders of ASSVd [13] are marked by black lines. Blue bars and labels mark TCR (Figure S10b), HPI (Figure S23), and LCCR (Figure S24).

Anyway, in most if not all publications with sequence descriptions of apscaviroids, it was mentioned that *Apscaviroid* sequences are composed of sequence stretches with high similarity to other members of *Pospiviroidae*. Given the correctness of the domain hypothesis, the borders of such a sequence stretches should coincide with the domain borders. To confirm this, we used the programs JALI and VMATCH. For example, in the publication of the first sequence of CDVd (formerly called citrus viroid IIIA, CVdIIIA) [73], the authors state: “the extended upper CCR [of CDVd] is most related to ASSVd ... In contrast, the sequence of the extended lower CCR is most related to PBCVd ... Thus, in the case of CVdIIIA, it appears that the CCR sequences were exchanged as two individual RNA strands from two different viroids.” Indeed, JALI aligns with the TL, upper P, and upper C domain of CDVd to ASSVd and the TR and lower C domain to PBCVd (see black line representing aligned segments of CDVd in Figure 4a). In addition, VMATCH finds the longest and most significant matches in the TL and upper C domain between ASSVd and CDVd and in the lower C domain between PBCVd and CDVd (see red and blue lines in Figure 4b,c). Note, however, that none of the recombination points predicted by JALI exactly coincide with the published domain borders of ASSVd [13] and none of the sequence matches found by VMATCH coincide with domains of ASSVd.

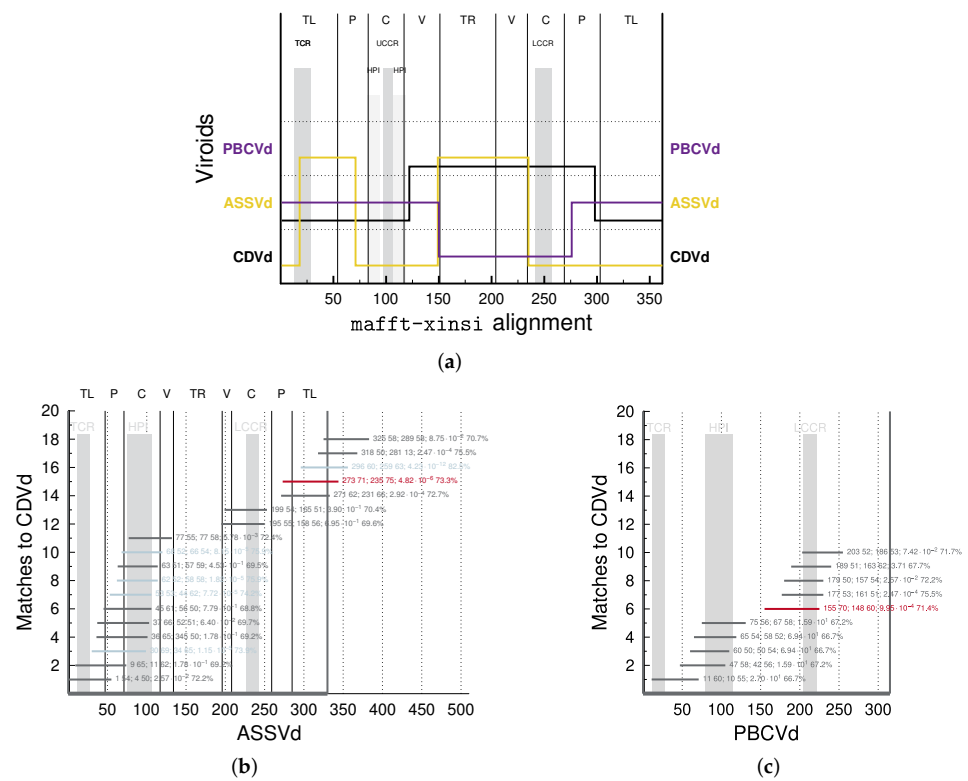


Figure 4. Similarity of CDVd with ASSVd and PBCVd. (a): JALI output with a seed alignment of CDVd, ASSVd, and PBCVd. The seed alignment of the three consensus sequences was produced with MAFFT X-INS-I. Horizontal lines in the graph show which parts of the candidate sequence were aligned by JALI to which sequence of the MAFFT alignment; vertical lines show jumps between two sequences. Line color coincides with color of viroid names. Domains of ASSVd are given at the top; domain borders are marked by black vertical lines. Motif regions are marked by vertical gray bars. For parameters used by JALI, see Material & Methods. (b,c): Graphical vMATCH output. Horizontal lines mark sequence stretches of high similarity between query—PBCVd (c) and ASSVd (b), respectively—and CDVd. Red lines denote stretches of length ≥ 70 nts; blue lines denote stretches with expect value $E < 10^{-3}$. The numbers right of match lines are position and length of subject given on x axis, position and length of query given on y axis, expect value, and PSI. In (b), the ASSVd sequence was doubled to allow for matches across the unit-length circle. For parameters used by vMATCH, see Material & Methods. Domain borders of ASSVd [13] are marked by black lines; vertical gray bars mark TCR (Figure S20b), LCCR (Figure S24), UCCR, and HPI (Figure S23) of ASSVd.

The lack of match between recombination points and domain borders generally holds true. Further examples by JALI are shown in Figures S13, S14 and S16 for all members of genera *Pospiviroid*, *Apscaviroid*, and *Coleviroid*, respectively. Another example is CBCVd, a member of genus *Cocadviroid* infecting *Citrus* as well as hop; already Puchta et al. [74] described that CBCVd shares sequence stretches with members of genera *Pospiviroid* and *Hostuviroid*. Indeed, the sequence of CBCVd aligns best to upper TL, P, and V domains of HSVd, to C and lower V domain of HLVD, to TR domain of CEVd, and to lower TL domain of CCCVd (Figure 5). Note that CBCVd and CEVd have only one RY motif, while most other members of *Pospiviroid* have two RY motifs (see alignment in Figure S17). This similarity between CBCVd and CEVd is in support of recombination between an CBCVd ancestor and an CEVd-like sequence.

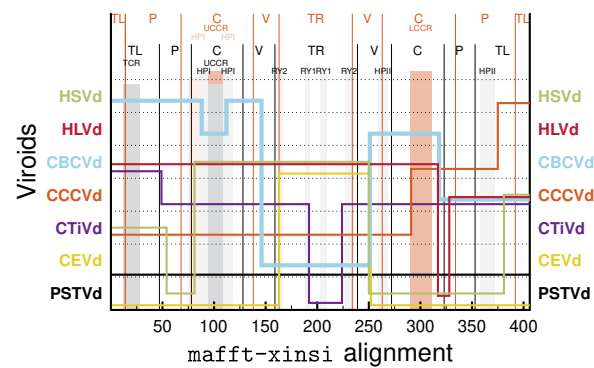


Figure 5. Similarity of CBCVd to other viroids. Domain borders and motifs of CCCVd and PSTVd are marked and labeled in orange and black, respectively, at the top. JALI output with a seed alignment of HSVd (genus *Hostuviroid*), HLVd, CBCVd, CCCVd, CTIVd (*Cocadviroid*), CEVd, and PSTVd (*Pospiviroid*). The seed alignment of the respective consensus sequences (Figure S17) was produced with MAFFT X-INS-I.

DLVd, a member of genus *Hostuviroid*, might be a recombinant between different viroid genera [75], in this respect comparable to CBCVd (see Figures S18 and S19). DLVd's TL domain is similar to that of PCFVd from genus *Pospiviroid*, includes a TCR like other members of *Pospiviroid* but not a TCH as HSVd, the type strain of *Hostuviroid*.

Similarly to the missing match of recombination points and domain borders, points of duplication in natural variants of CCCVd [61,62] and CEVd [58–60] do not generally coincide with domain borders. For a few examples, see flags in the respective structures of Figure 1; the flags enclose sequence stretches that occur duplicated in longer-than-unit length sequence variants. The exception, however, is the sequence duplication of a CCCVd variant (LOCUS CCC1SLOW), published in 1982 [61], that starts and ends exactly at the V/TR border.

4. Conclusions

Our analysis with the software tools NUCALN, MAFFT, JALI, and VMATCH of evolutionary relations between viroids of family *Pospiviroidae* support the original finding by Keese and Symons in 1985 that these viroids have evolved by the rearrangement of sequence stretches between different viroids infecting the same cell and subsequent mutations. Even recombination between viroids from the different genera of *Pospiviroidae* are likely. We were, however, not able to reproduce the exact borders of these recombined stretches and thus have doubts regarding the validity of the strict domain model as defined by Keese and Symons. For all ten species of *Pospiviroid*, we were able to improve the original domain borders to be consistent between species with respect to a MAFFT alignment. Nevertheless, the conserved sequence and structure motifs—especially those of *Pospiviroid* members—are sufficient to subdivide their rod-shaped structure into biologically functional sections.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/cells11020230/s1>. Modifications of CLUSTALΩ. Supplementary figures: Figure S1. Illustration of parameters in the Wilbur–Lipman algorithm. Figure S2. Examples for alignment between a candidate sequence and a seed alignment using JALI. Figure S3. JALI output with CSVd as the candidate sequence and all other viroids presented in Keese and Symons (1985) as the seed alignment. Figure S4. Pairwise alignments of PSTVd to the other sequences used by Keese and Symons (1985) with NUCALN. Figure S5. Dotplots between the sequences used by Keese and Symons (1985). Figure S6. Dotplots between consensus sequences of the species used by Keese and Symons (1985). Figure S7. MAFFT alignment between consensus sequences of *Pospiviroid* members. Figure S8. MAFFT alignment between sequences used in Keese and Symons (1987). Figure S9. Overlay of dotplots from NUCALN alignments for all *Pospiviroid* members. Figure S10. MAFFT alignment between consensus sequences of *Apscaviroid* members. Figure S11. Consensus sequences and secondary structures of *Apscaviroid* members. Figure S12. Alignment of CDVd, ASSVd, and PBCVd. Figure S13.

Similarity among *Pospiviroid* members. Figure S14. Similarity among *Apscaviroid* members. Figure S15. MAFFT alignment between the consensus sequences of *Coleviroid* members. Figure S16. Similarity among *Coleviroid* members. Figure S17. MAFFT alignment between members of the genera *Cocad-*, *Pospi-*, and *Hostuviroid*. Figure S18. MAFFT alignment of DLVd, HSVd, and selected members of *Pospiviroid*. Figure S19. Similarity among DLVd, HSVd, and selected members of *Pospiviroid*. Figure S20. Sequence logo of “Terminal Conserved Region” (TCR). Figure S21. Sequence logo of “Terminal Conserved Hairpin” (TCH). Figure S22. Part of the central domain including hairpin I (HPI) based on an alignment of 899 *Pospiviroid* sequences. Figure S23. Part of the central domain including hairpin I (HPI) based on an alignment of 612 *Apscaviroid* sequences. Figure S24. LCCR of *Apscaviroid* members. Figure S25. Hairpin II (HPII) of *Pospiviroid* sequences. Figure S26. RY motif of *Pospiviroid* sequences. Supplementary tables: Table S1. Classification of viroids, viroid names, and abbreviations. Table S2. Pairwise sequence identity of the sequences used by Keese and Symons (1985) after alignment with NUCALN ($k = 4$ and $g = 4$). Table S3. Pairwise sequence identity of the sequences used by Keese and Symons (1985) after alignment with NUCALN with parameters $k = 3$ and $g = 3$. Table S4. Average pairwise sequence identity of consensus sequences of *Pospiviroid* species after alignment with MAFFT X-INS-I. Table S5. Consistent domain borders of *Pospiviroid* members. Table S6. Nomenclature for incompletely specified bases.

Author Contributions: G.S. conceived the study and wrote the manuscript; K.-P.W. and G.S. analyzed the data. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Steffen Mielenbrink (IPB, HHU) for initial tests of software tools and Laura Rose (Institut für Populationsgenetik, HHU) and Detlev Riesner (IPB, HHU) for discussions. We acknowledge support by the Heinrich Heine University Düsseldorf.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sanger, H.; Klotz, G.; Riesner, D.; Gross, H.; Kleinschmidt, A. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 3852–3856. [[CrossRef](#)] [[PubMed](#)]
2. Davies, J.; Kaeseberg, P.; Diener, T. Potato spindle tuber viroid. XII. An investigation of viroid RNA as a messenger for protein synthesis. *Virology* **1974**, *61*, 281–286. [[CrossRef](#)]
3. Hadidi, A.; Randles, J.; Flores, R.; Palukaitis, P. (Eds.) *Viroids and Satellites*; Academic Press, Elsevier: Cambridge, MA, USA, 2017.
4. Harders, J.; Lukacs, N.; Robert-Nicoud, M.; Jovin, T.; Riesner, D. Imaging of viroids in nuclei from tomato leaf tissue by *in situ* hybridization and confocal laser scanning microscopy. *EMBO J.* **1989**, *8*, 3941–3949. [[CrossRef](#)]
5. Schindler, I.M.; Muhlbach, H.P. Involvement of nuclear DNA-dependent RNA polymerases in potato spindle tuber viroid replication: A reevaluation. *Plant Sci.* **1992**, *84*, 221–229. [[CrossRef](#)]
6. Qi, Y.; Ding, B. Differential subnuclear localization of RNA strands of opposite polarity derived from an autonomously replicating viroid. *Plant Cell* **2003**, *15*, 2566–2577. [[CrossRef](#)] [[PubMed](#)]
7. Dissanayaka Mudiyanse, S.; Qu, J.; Tian, N.; Jiang, J.; Wang, Y. Potato spindle tuber viroid RNA-templated transcription: Factors and regulation. *Viruses* **2018**, *10*, 503. [[CrossRef](#)] [[PubMed](#)]
8. Steger, G.; Perreault, J.P. Structure and associated biological functions of viroids. *Adv. Virus Res.* **2016**, *94*, 141–172. [[CrossRef](#)]
9. Giguere, T.; Adkar-Purushothama, C.; Perreault, J. Comprehensive secondary structure elucidation of four genera of the family *Pospiviroidae*. *PLoS ONE* **2014**, *9*, e98655. [[CrossRef](#)]
10. Lopez-Carrasco, A.; Flores, R. Dissecting the secondary structure of the circular RNA of a nuclear viroid *in vivo*: A “naked” rod-like conformation similar but not identical to that observed *in vitro*. *RNA Biol.* **2017**, *14*, 1046–1054. [[CrossRef](#)] [[PubMed](#)]
11. Keese, P.; Symons, R. Domains in viroids: Evidence of intermolecular RNA rearrangement and their contribution to viroid evolution. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 4582–4586. [[CrossRef](#)]
12. McInnes, J.; Symons, R. Comparative structure of viroids and their rapid detection using radioactive and nonradioactive nucleic acid probes. In *Viroids and Satellites: Molecular Parasites at the Frontier of Life*; Maramorosch, K., Ed.; CRC Press: Boca Raton, FL, USA, 1991; pp. 21–58.
13. Koltunow, A.; Rezaian, M. A scheme for viroid classification. *Intervirology* **1989**, *30*, 194–201. [[CrossRef](#)]

14. Goodman, T.; Nagel, L.; Rappold, W.; Klotz, G.; Riesner, D. Viroid replication: Equilibrium association constant and comparative activity measurements for the viroid-polymerase interactions. *Nucleic Acids Res.* **1984**, *12*, 6231–6246. [[CrossRef](#)] [[PubMed](#)]
15. Kolonko, N.; Bannach, O.; Aschermann, K.; Hu, K.H.; Moors, M.; Schmitz, M.; Steger, G.; Riesner, D. Transcription of potato spindle tuber viroid by RNA polymerase II starts in the left terminal loop. *Virology* **2006**, *347*, 392–404. [[CrossRef](#)] [[PubMed](#)]
16. Bojić, T.; Beeharry, Y.; Zhang, D.; Pelchat, M. Tomato RNA polymerase II interacts with the rod-like conformation of the left terminal domain of the potato spindle tuber viroid positive RNA genome. *J. Gen. Virol.* **2012**, *93*, 1591–1600. [[CrossRef](#)] [[PubMed](#)]
17. Dissanayaka Mudiyanselage, S.; Wang, Y. Evidence Supporting That RNA Polymerase II Catalyzes De Novo Transcription Using Potato Spindle Tuber Viroid Circular RNA Templates. *Viruses* **2020**, *12*, 371. [[CrossRef](#)]
18. Koltunow, A.; Rezaian, M. Grapevine yellow speckle viroid: Structural features of a new viroid group. *Nucleic Acids Res.* **1988**, *16*, 849–864. [[CrossRef](#)]
19. Puchta, H.; Ramm, K.; Sängner, H. The molecular structure of hop latent viroid (HLV), a new viroid occurring worldwide in hops. *Nucleic Acids Res.* **1988**, *16*, 4197–4216. [[CrossRef](#)]
20. Di Serio, F.; Li, S.F.; Palls, V.; Pwens, R.; Randles, J.; Sano, T.; Verhoeven, J.; Vidalakis, G.; Flores, R. Viroid taxonomy. In *Viroids and Satellites*; Hadidi, A., Randles, J., Flores, R., Palukaitis, P., Eds.; Academic Press, Elsevier: Cambridge, MA, USA 2017; pp. 135–146.
21. Schnölzer, M.; Haas, B.; Ramm, K.; Hofmann, H.; Sängner, H. Correlation between structure and pathogenicity of potato spindle tuber viroid (PSTV). *EMBO J.* **1985**, *4*, 2181–2190. [[CrossRef](#)]
22. Visvader, J.; Symons, R. Eleven new sequence variants of citrus exocortis viroid and the correlation of sequence with pathogenicity. *Nucleic Acids Res.* **1985**, *13*, 2907–2920. [[CrossRef](#)] [[PubMed](#)]
23. Kitabayashi, S.; Tsushima, D.; Adkar-Purushothama, C.; Sano, T. Identification and Molecular Mechanisms of Key Nucleotides Causing Attenuation in Pathogenicity of Dahlia Isolate of Potato Spindle Tuber Viroid. *Int. J. Mol. Sci.* **2020**, *21*, 7352. [[CrossRef](#)]
24. Sano, T.; Candresse, T.; Hammond, R.; Diener, T.; Owens, R. Identification of multiple structural domains regulating viroid pathogenicity. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10104–10108. [[CrossRef](#)]
25. Rodriguez, M.; Randles, J. Coconut cadang-cadang viroid (CCCVd) mutants associated with severe disease vary in both the pathogenicity domain and the central conserved region. *Nucleic Acids Res.* **1993**, *21*, 2771. [[CrossRef](#)] [[PubMed](#)]
26. Chaffai, M.; Serra, P.; Gandía, M.; Hernández, C.; Duran-Vila, N. Molecular characterization of CEVd strains that induce different phenotypes in *Gynura aurantiaca*: Structure-pathogenicity relationships. *Arch. Virol.* **2007**, *152*, 1283–1294. [[CrossRef](#)]
27. Li, R.; Padmanabhan, C.; Ling, K. A single base pair in the right terminal domain of tomato planta macho viroid is a virulence determinant factor on tomato. *Virology* **2016**, *500*, 238–246. [[CrossRef](#)]
28. Steger, G.; Hofmann, H.; Förtisch, J.; Gross, H.; Randles, J.; Sängner, H.; Riesner, D. Conformational transitions in viroids and virusoids: Comparison of results from energy minimization algorithm and from experimental data. *J. Biomol. Struct. Dyn.* **1984**, *2*, 543–571. [[CrossRef](#)] [[PubMed](#)]
29. Shen, L.; Cai, Z.; Tinoco, I. RNA structure at high resolution. *FASEB J.* **1995**, *9*, 1023–1033. [[CrossRef](#)] [[PubMed](#)]
30. Zhong, X.; Leontis, N.; Qian, S.; Itaya, A.; Qi, Y.; Boris-Lawrie, K.; Ding, B. Tertiary structural and functional analyses of a viroid RNA motif by isostericity matrix and mutagenesis reveal its essential role in replication. *J. Virol.* **2006**, *80*, 8566–8581. [[CrossRef](#)] [[PubMed](#)]
31. Branch, A.; Benenfeld, B.; Robertson, H. Ultraviolet light-induced crosslinking reveals a unique region of local tertiary structure in potato spindle tuber viroid and HeLa 5 S RNA. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 6590–6594. [[CrossRef](#)]
32. Wang, Y.; Zhong, X.; Itaya, A.; Ding, B. Evidence for the existence of the loop E motif of *Potato spindle tuber viroid* in vivo. *J. Virol.* **2007**, *81*, 2074–2077. [[CrossRef](#)]
33. Eiras, M.; Kitajima, E.; Flores, R.; Darós, J. Existence in vivo of the loop E motif in potato spindle tuber viroid RNA. *Arch. Virol.* **2007**, *152*, 1389–1393. [[CrossRef](#)]
34. Henco, K.; Sängner, H.; Riesner, D. Fine structure melting of viroids as studied by kinetic methods. *Nucleic Acids Res.* **1979**, *6*, 3041–3059. [[CrossRef](#)]
35. Diener, T. Viroid processing: A model involving the central conserved region and hairpin I. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 58–62. [[CrossRef](#)] [[PubMed](#)]
36. Steger, G.; Tabler, M.; Brüggemann, W.; Colpan, M.; Klotz, G.; Sängner, H.; Riesner, D. Structure of viroid replicative intermediates: Physico-chemical studies on SP6 transcripts of cloned oligomeric potato spindle tuber viroid. *Nucleic Acids Res.* **1986**, *14*, 9613–9630. [[CrossRef](#)] [[PubMed](#)]
37. Baumstark, T.; Schröder, A.; Riesner, D. Viroid processing: Switch from cleavage to ligation is driven by a change from a tetraloop to a loop E conformation. *EMBO J.* **1997**, *16*, 599–610. [[CrossRef](#)] [[PubMed](#)]
38. Schrader, O.; Baumstark, T.; Riesner, D. A mini-RNA containing the tetraloop, wobble-pair and loop E motifs of the central conserved region of potato spindle tuber viroid is processed into a minicircle. *Nucleic Acids Res.* **2003**, *31*, 988–998. [[CrossRef](#)]
39. Gas, M.; Hernández, C.; Flores, R.; Darós, J. Processing of nuclear viroids in vivo: An interplay between RNA conformations. *PLoS Pathog.* **2007**, *3*, e182. [[CrossRef](#)] [[PubMed](#)]
40. Loss, P.; Schmitz, M.; Steger, G.; Riesner, D. Formation of a thermodynamically metastable structure containing hairpin II is critical for infectivity of potato spindle tuber viroid RNA. *EMBO J.* **1991**, *10*, 719–727. [[CrossRef](#)]
41. Qu, F.; Heinrich, C.; Loss, P.; Steger, G.; Tien, P.; Riesner, D. Multiple pathways of reversion in viroids for conservation of structural elements. *EMBO J.* **1993**, *12*, 2129–2139. [[CrossRef](#)]

42. Gozmanova, M.; Denti, M.; Minkov, I.; Tsagris, M.; Tabler, M. 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.* **2003**, *31*, 5534–5543. [[CrossRef](#)]
43. Martínez de Alba, A.; Sägesser, R.; Tabler, M.; Tsagris, M. A bromodomain-containing protein from tomato specifically binds potato spindle tuber viroid RNA in vitro and in vivo. *J. Virol.* **2003**, *77*, 9685–9694. [[CrossRef](#)]
44. Steger, G. Modelling the three-dimensional structure of the right-terminal domain of pospiviroids. *Sci. Rep.* **2017**, *7*, 711, [[CrossRef](#)] [[PubMed](#)]
45. Kalantidis, K.; Denti, M.; Tzortzakaki, S.; Marinou, E.; Tabler, M.; Tsagris, M. Virp1 is a host protein with a major role in *Potato spindle tuber viroid* infection in *Nicotiana* plants. *J. Virol.* **2007**, *81*, 12872–12880. [[CrossRef](#)] [[PubMed](#)]
46. Chaturvedi, S.; Kalantidis, K.; Rao, A. A bromodomain-containing host protein mediates the nuclear importation of a satellite RNA of *Cucumber mosaic virus*. *J. Virol.* **2013**, *88*, 1890–1896. [[CrossRef](#)]
47. Wilbur, W.; Lipman, D. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 726–730. [[CrossRef](#)]
48. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539, [[CrossRef](#)]
49. Spang, R.; Rehmsmeier, M.; Stoye, J. Sequence database search using jumping alignments. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2000**, *8*, 367–375.
50. Spang, R.; Rehmsmeier, M.; Stoye, J. A novel approach to remote homology detection: Jumping alignments. *J. Comput. Biol.* **2002**, *9*, 747–760. [[CrossRef](#)] [[PubMed](#)]
51. Abouelhoda, M.; Kurtz, S.; Ohlebusch, E. Replacing suffix trees with enhanced suffix arrays. *J. Discr. Algorithms* **2004**, *2*, 53–86. [[CrossRef](#)]
52. Keese, P.; Symons, R. Molecular structure (primary and secondary). In *The Viroids*; Diener, T., Ed.; Plenum Press: New York, NY, USA, 1987; pp. 37–62.
53. Katoh, K.; Toh, H. Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinform.* **2008**, *9*, 212, [[CrossRef](#)]
54. Wilm, A.; Linnenbrink, K.; Steger, G. ConStruct: Improved construction of RNA consensus structures. *BMC Bioinform.* **2008**, *9*, 219, [[CrossRef](#)] [[PubMed](#)]
55. Weinberg, Z.; Breaker, R. R2R—Software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinform.* **2011**, *12*, 3, [[CrossRef](#)]
56. Weinberg, Z.; Barrick, J.; Yao, Z.; Roth, A.; Kim, J.; Gore, J.; Wang, J.; Lee, E.; Block, K.; Sudarsan, N.; et al. Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res.* **2007**, *35*, 4809–4819. [[CrossRef](#)] [[PubMed](#)]
57. Rivas, E.; Clements, J.; Eddy, S. A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods* **2017**, *14*, 45–48. [[CrossRef](#)] [[PubMed](#)]
58. Semancik, J.; Szychowski, J.; Rakowski, A.; Symons, R. Isolates of citrus exocortis viroid recovered by host and tissue selection. *J. Gen. Virol.* **1993**, *74*, 2427–2436. [[CrossRef](#)] [[PubMed](#)]
59. Semancik, J.; Szychowski, J.; Rakowski, A.; Symons, R. A stable 463 nucleotide variant of citrus exocortis viroid produced by terminal repeats. *J. Gen. Virol.* **1994**, *75*, 727–732. [[CrossRef](#)]
60. Fadda, Z.; Darós, J.; Flores, R.; Duran-Vila, N. Identification in eggplant of a variant of citrus exocortis viroid (CEVd) with a 96 nucleotide duplication in the right terminal region of the rod-like secondary structure. *Virus Res.* **2003**, *97*, 145–149. [[CrossRef](#)] [[PubMed](#)]
61. Haseloff, J.; Mohamed, N.; Symons, R. Viroid RNAs of cadang-cadang disease of coconuts. *Nature* **1982**, *299*, 316–321. [[CrossRef](#)]
62. Vadamalai, G.; Hanold, D.; Rezaian, M.; Randles, J. Variants of Coconut cadang-cadang viroid isolated from an African oil palm (*Elaeis guineensis* Jacq.) in Malaysia. *Arch. Virol.* **2006**, *151*, 1447–1456. [[CrossRef](#)]
63. Boratyn, G.; Camacho, C.; Cooper, P.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T.; Matten, W.; McGinnis, S.; Merezchuk, Y.; et al. BLAST: A more efficient report with usability improvements. *Nucleic Acids Res.* **2013**, *41*, W29–W33. [[CrossRef](#)]
64. Graphics Layout Engine. Available online: <http://glx.sourceforge.net/> (accessed on 1 November 2021).
65. Katoh, K.; Kuma, K.; Miyata, T.; Toh, H. Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Inform. Ser.* **2005**, *16*, 22–33. [[CrossRef](#)]
66. Katoh, K.; Rozewicki, J.; Yamada, K. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, 1160–1166. [[CrossRef](#)] [[PubMed](#)]
67. Lorenz, R.; Bernhart, S.; Höner Zu Siederdissen, C.; Tafer, H.; Flamm, C.; Stadler, P.; Hofacker, I. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **2011**, *6*, 26, [[CrossRef](#)] [[PubMed](#)]
68. Viral Genome Browser (*Pospiviroidae*). Available online: <https://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=185751> (accessed on 1 November 2021).
69. Brister, J.; Ako-Adjei, D.; Bao, Y.; Blinkova, O. NCBI viral genomes resource. *Nucleic Acids Res.* **2015**, *43*, D571–D577, [[CrossRef](#)] [[PubMed](#)]
70. Klein, R.; Eddy, S. RSEARCH: Finding homologs of single structured RNA sequences. *BMC Bioinform.* **2003**, *4*, 44, [[CrossRef](#)]

71. Keese, P.; Osorio-Keese, M.; Symons, R. Coconut tinangaja viroid: Sequence homology with coconut cadang-cadang viroid and other potato spindle tuber viroid related RNAs. *Virology* **1988**, *162*, 508–510. [[CrossRef](#)]
72. Koltunow, A.; Rezaian, M. Grapevine viroid 1B, a new member of the apple scar skin viroid group contains the left terminal region of tomato planta macho viroid. *Virology* **1989**, *170*, 575–578. [[CrossRef](#)]
73. Stasys, R.; Dry, I.; Rezaian, M. The termini of a new citrus viroid contain duplications of the central conserved regions from two viroid groups. *FEBS Lett.* **1995**, *358*, 182–184. [[CrossRef](#)]
74. Puchta, H.; Ramm, K.; Luckinger, R.; Hadas, R.; Bar-Joseph, M.; Sanger, H. Primary and secondary structure of citrus viroid IV (CVd IV), a new chimeric viroid present in dwarfed grapefruit in Israel. *Nucleic Acids Res.* **1991**, *19*, 6640, [[CrossRef](#)]
75. Verhoeven, J.; Meeke, E.; Roenhorst, J.; Flores, R.; Serra, P. Dahlia latent viroid: A recombinant new species of the family *Pospiviroidae* posing intriguing questions about its origin and classification. *J. Gen. Virol.* **2012**, *94*, 711–719. [[CrossRef](#)] [[PubMed](#)]