Archives of Virology (2025) 170:12
https://doi.org/10.1007/s00705-024-06191-4

ORIGINAL ARTICLE

Characterisation of the genetic diversity of citrus viroid VII using amplicon sequencing

Grant A. Chambers[1,2] · Andrew D.W. Geering[2] · Daniel R. Bogema[1] · Paul Holford[3] · Georgios Vidalakis[4] · Nerida J. Donovan[1]

Received: 20 October 2024 / Accepted: 20 November 2024 / Published online: 12 December 2024

Abstract
Viroids occur in plants as swarms of sequence variants clustered around a dominant variant, leading to adoption of the term 'quasispecies' to describe the viroid population in an individual host. The composition of the quasispecies can potentially change according to the age of the infection, the position of the leaf or branch in the canopy, and the host species. The primary aim of this study was to investigate the quasispecies concept for citrus viroid VII (CVd-VII), a recently discovered member of the family *Pospiviroidae*. Three experiments were conducted to determine factors affecting viroid variability (i) within different tissues of a lemon plant, (ii) among different plants of the same species (citron), and (iii) among different species and hybrids of citrus. Using two primer sets to produce amplicons for high-throughput sequencing, viroid population profiles were generated for each sample. The number of variants that were identified with both primer sets ranged from 2 to 13 per sample, and each sample comprised 1 to 4 major (> 10% sample) variants. The composition of variants differed in samples from different plants and among tissue types of a single plant. Single-nucleotide polymorphisms (SNPs), mostly in the form of substitutions, were the primary source of variation; in this study, SNPs were observed in approximately 10% of the viroid genome. The results of the three experiments indicate that CVd-VII follows the quasispecies model as reported for other viroids and that variability occurs in viroid populations in different tissue types and host species.

Introduction

Members of the family *Pospiviroidae* are small (246–375 nucleotides [nt]), single-stranded, circular RNA molecules that adopt a rod-like or quasi-rod-like conformation due to internal base pairing. They contain two sites of sequence conservation: a central conserved region (CCR) and either a terminal conserved hairpin (TCH) or a terminal conserved region (TCR) [10, 16, 39]. The family *Pospiviroidae* is further divided into five genera (*Apscaviroid*, *Cocadviroid*, *Coleviroid*, *Hostuviroid*, and *Pospiviroid*) based on the type of CCR, the presence of a TCH or a TCR, and

Handling Editor Ioannis E. Tzanetakis

✉ Grant A. Chambers
grant.chambers@dpi.nsw.gov.au

Andrew D.W. Geering
a.geering@uq.edu.au

Daniel R. Bogema
daniel.bogema@dpi.nsw.gov.au

Paul Holford
P.Holford@westernsydney.edu.au

Georgios Vidalakis
vidalg@ucr.edu

Nerida J. Donovan
nerida.donovan@dpi.nsw.gov.au

[1] NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Private Bag 4008, Narellan, NSW 2567, Australia

[2] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, GPO Box 267, Brisbane, Queensland 4001, Australia

[3] School of Science, Western Sydney University, LB 1797, Penrith 2751, NSW, Australia

[4] Department of Microbiology and Plant Pathology, University of California, Riverside, Riverside 92521, CA, USA

the clades observed in phylogenetic analysis. Members of the *Pospiviroidae* also have five conserved structural elements: the terminal regions at the left and right extremities of the rod-shaped structure (T_L and T_R, respectively) and the pathogenic (P), variable (V), and central (C) domains [26]. However, the exact boundaries of these domains are not properly defined, particularly for members of the genus *Apscaviroid* [52]. Citrus viroid VII (CVd-VII, species *Apscaviroid etacitri*), the topic of this paper, belongs to the genus *Apscaviroid* and has a CCR that is identical to that of apple scar skin viroid (species *Apscaviroid cicatricimali*), the type member of the genus.

Members of the *Pospiviroidae* replicate in the nucleus of their host and utilise host-encoded DNA-dependent RNA polymerase II for this process [17]. This enzyme has a high error rate during transcription, reported to be $1.4\text{-}3 \times 10^{-4}$ errors per nucleotide position for potato spindle tuber viroid (PSTVd, species *Pospiviroid fusituberis*) [4, 29]. The low fidelity associated with this polymerase results in mutations and sequence variability, which has been observed in studies where viroid cDNA transcripts were inoculated into host plants, leading to the accumulation of genetic variants and the formation of quasispecies [19, 22–24].

The quasispecies theoretical model was first proposed by Eigen [15] to describe viruses that form self-sustaining populations of sequence variants due to imperfect replication in their host plant. This same concept applies to viroids, which exist as a suite of sequence variants present in an individual host with many single-nucleotide polymorphisms (SNPs) [2, 4, 9, 20]. For example, over a third of the genome of pear blister canker viroid (species *Apscaviroid pustulapyri*) is variable [35]. However, the SNPs are not randomly distributed in the genome; rather, they accumulate in locations where a mutation does not destabilise the secondary structure of the viroid or in core replication-associated motifs. Selection pressures are also exerted by the host plant, thus influencing which mutations become fixed in the quasispecies population and which are eliminated [13, 47]. The term 'quasispecies' in this study refers to the outcomes of the quasispecies theory, a population of closely related viroid variants observed in an individual host.

The sequencing of recombinant cDNA clones to determine viroid sequence diversity has been commonly used to analyse the quasispecies profiles of many viroid species [5, 19, 21, 24, 47, 48]. The development of high-throughput sequencing has enabled a more thorough analysis of the population composition of viroid sequence variants [1, 4, 22, 29, 46]. This technique permits the determination of nearly complete genome sequences of individual viroid molecules and their relative frequency. Tangkanchanapas et al. [46] determined that amplicon sequencing replicates displayed high reproducibility, with the vast majority of sequence variants observed in all polymerase chain reaction (PCR) replicates of a plant sample.

Citrus plants are susceptible to a range of viroids, including the highly pathogenic citrus exocortis viroid (species *Pospiviroid exocortiscitri*) and the cachexia-inducing variants of hop stunt viroid (HSVd, species *Hostuviroid impedihumuli*) [12, 41]. Sequence variants of citrus viroids have long been reported to differ in their biological properties and pathogenicity [33, 37, 38, 42, 49]. In HSVd, a single nucleotide change is responsible for the induction of severe or mild cachexia symptoms [43]. Citrus viroid VII is the most recently discovered of the viroids infecting citrus and was found in the asymptomatic lemon (*Citrus ×limon* (L.) Osbeck) cultivars, 'Lisbon' and 'Eureka' following routine biological indexing on 'Etrog' citron (*C. medica* L.) [6]. The viroid has only been detected in trees growing at a research station in Dareton, New South Wales, Australia. Little is known about the biology of this viroid, and knowledge of the genetic variability of CVd-VII is restricted, with only four sequence variants described [6].

The first aim of this study was to investigate the genome variability of CVd-VII. We hypothesised that the profile of sequence variants would vary according to host species infected and the tissue type that was sampled. Complex three-dimensional secondary structures play a crucial role in determining viroid activity. The loops or bulges that are found throughout the secondary structure of PSTVd are hypothesised to represent functional motifs that have distinct roles in replication and trafficking processes [18, 54]. Therefore, this study also aimed to determine locations of SNPs within the genome of CVd-VII and the effect of any sequence variability on its structure. We also sought to define the variable regions of the CVd-VII genome, which need to be considered when designing PCR-based diagnostic assays.

Materials and methods

Viroid source, inoculation, and sampling

The CVd-VII isolate used in this study originated from a lemon cv. Lisbon field tree [6]. This isolate was maintained in either a daughter lemon plant grafted onto a hybrid rootstock (Cox; mandarin (*C. ×aurantium* L.) × trifoliate orange (*C. trifoliata* L.)) and grown in a greenhouse under controlled environmental conditions (26°C/22°C) with natural light or in an 'etrog' citron slash-inoculated with CVd-VII RNA [6].

Factors affecting viroid variability

Three different experiments were performed to examine factors affecting viroid variability, as follows:

Intraplant variability of CVd-VII

Triplicate samples (~50 mg) of either leaf lamina or bark were taken from new growth (latest flush) or mature growth (second preceding flush) from a single CVd-VII-infected lemon plant grown in a temperature-controlled glasshouse maintained at 20–25°C for RNA extraction.

Interplant variability of CVd-VII

Eight healthy citron plants propagated from cuttings were grown in a warm-temperature glasshouse (30°C/26°C) and then graft-inoculated with CVd-VII using bark chips taken from a single budstick of a CVd-VII-infected citron source plant. Bark (~50 mg) was sampled from the single budstick of the CVd-VII-infected citron source plant at the time of inoculation of the recipient plants and from new growth of each of the eight inoculated citron plants seven months post-inoculation.

Intervarietal variability of CVd-VII

Healthy bergamot (*Citrus ×limon* (L.)), blood orange (*C. ×aurantium* L.), lemon (cv. 'Fino'), mandarin (cv. 'Oroval'), and Valencia orange (*C. ×aurantium* L., cv. 'Midknight') were grafted onto duplicate rough lemon (*C. ×otaitensis* (Risso & Poit.) Risso) rootstocks at the same time as bark inoculum from CVd-VII-infected citron was budded onto the rootstock. The scions were allowed to grow out in a temperature-controlled glasshouse maintained at 20–25°C. Four years post-inoculation, green bark tissue (~50 mg) was sampled from all scions for RNA extraction.

Table 1 Primers used in this study

Primer name	Sequence (5'-3')	Nucleotide position
CVdVII-F8	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG***GTCGACGAAGGGTTCTCCA***	101–119
CVdVII-R8	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG***GACGAGTCTCCAGGTGAGT***	100–82
CVdVII-F9	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG***GTTCCTATGGTGCAGCCC***	16–33
CVdVII-R9	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG***CAGAAGTGTCCTCGAGGGA***	15–365

RNA extraction and RT-PCR (reverse transcription polymerase chain reaction)

Total RNA was extracted using a MagMax Plant RNA Extraction Kit (Applied Biosystems, Australia) according to the manufacturer's instructions, either manually or using a KingFisher-96 Automated Extraction System (Thermo Fisher Scientific, Australia), and eluted using nuclease-free water.

To prepare cDNA, total RNA (2 µL) was reverse transcribed using a Tetro cDNA Synthesis Kit (Meridian Bioscience, Australia) with random hexamer primers in a 10-µL reaction volume, as per the manufacturer's instructions.

Two primer sets (Table 1) were designed using Geneious Prime (Biomatters, New Zealand) to amplify the entire CVd-VII genome. Each primer set amplified the regions where the alternate primer binding sites are located [28], ensuring that the complete viroid genome sequence was determined. Illumina overhang adaptor sequences (Table 1) were added to the 5' end of both forward and reverse primers for DNA sequence analysis on a MiSeq instrument (Illumina, United States).

Primers consisted of Illumina overhang adaptor sequences and CVd-VII-specific primer sequences, which are indicated in bold italics. The nucleotide position is in reference to CVd-VII (GenBank accession no. KX013551).

For each cDNA sample, triplicate PCR amplifications were performed using the CVdVII-F8/CVdVII-R8 (P8) and CVdVII-F9/CVdVII-R9 (P9) primer sets in separate reactions. The 10 µL reaction mixtures consisted of 1 µL of cDNA, 0.2 mM each primer, and 5 µL of 2× MyFi Mix (Meridian Bioscience, Australia). PCR cycling consisted of 95°C for 2 min followed by 35 cycles of 95°C for 20 s, 55°C for 20 s, and 72°C for 20 s, and then a final extension at 72°C for 5 min. To confirm the generation of an amplicon of the correct size, 2 µL of each PCR was electrophoretically separated in a 2% agarose gel in 1× TBE and visualised under UV light using Gel Red (Biotium Inc, USA).

Illumina library preparation and amplicon sequencing

Triplicate amplicons produced in separate PCRs were combined and purified using an Isolate II PCR and Gel Kit (Meridian Bioscience, Australia). Indexing PCRs were carried out with Nextera XT i7 and i5 indexes (Nextera XT index kit 24 indexes 96 samples, Illumina, USA), which were added to samples using an amplification step with MyFi Mix (Meridian Bioscience, Australia) in a 25-µL reaction volume. The indexing PCR cycling conditions were 95°C for 3 min, followed by 12 cycles of 95°C for 10 s, 55°C for 30 s, and 72°C for 30 s, and then a final

extension at 72°C for 5 min. Purification of individual indexed samples was completed using 45 µL of AmpureXP beads (Sigma Aldrich, Australia) and elution in 27.5 µL of 10 mM Tris-HCl (pH 8.5). Following quantification using a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Australia) and a 2200 TapeStation system (Agilent, Australia), a 4 nM pooled DNA library was produced for sequencing. Libraries were sequenced using an Illumina MiSeq platform with a 500-cycle MiSeq Reagent Kit Nano V2 (Illumina, USA).

Bioinformatics

The sequencing data generated from the pooled library were demultiplexed on the MiSeq instrument, and the quality of fastq files was assessed using fastp v0.23.2 [8]. Primer sequences were removed from read pairs individually using Cutadapt v4.1 [32], and read-pairs containing both correct primer pairs were then repaired using the repair.sh tool in BBMap v39.01 (https://sourceforge.net/projects/bbmap/). Reads were denoised using DADA2, and feature tables, export read counts, and rarefaction analysis [44] were generated using QIIME2 [3]. Paired consensus sequences were discarded if they were not 363–370 nt in length, to ensure that only true variants were analysed. Geneious Prime (Biomatters) was used to exclude primer-binding regions from the sequence variants generated with QIIME2 for each of the primer sets (P8 and P9), leaving approximately 290 nt of unique sequence for further analysis (Supplementary Fig. S1).

Only sequence variants that were amplified using both primer sets were analysed further [1, 46]. Major variants were classed as those having > 10% abundance within a sample, while the relative abundances of all sequence variants were determined in generating the viroid profile of each sample. A viroid profile in this study is described as the identity of the variants, and the estimations of relative frequency of those variants, within a single plant sample. Individual sequence variants were provided with a 'V' number.

Phylogenetic analysis

To enable alignment and structural analysis of unique variants, the primer sequences were restored to produce full-length viroid sequences. Alignment of complete sequences of individual sequence variants was performed using MUSCLE [14] with default parameters in Geneious Prime software (Biomatters). Neighbour-joining consensus trees were generated using Geneious Prime (Biomatters) with no outgroup, and trees were subsequently rooted with minimum variance using FastRoot [31]. Heat maps were generated using the Interactive Tree of Life (Version 6.8.1) online tool.

Sequence secondary structure predictions

To understand the possible effect of variants on viroid secondary structure, selected full-length sequences were analysed using the online RNA structure prediction tool Mfold version 2.3 (http://www.unafold.org/mfold/applications/rna-folding-form.php), using default parameters for circular RNA and a folding temperature of 25°C. To complete the genome sequence, primer sequences were reinserted prior to structural analysis.

Results

Detection of CVd-VII quasispecies populations

A combined total of 1,623,990 merged reads were generated using the P8 and P9 primer sets from the 34 samples that were used in the three experiments. Rarefaction analysis using the Shannon diversity index showed that a sequencing depth of 500 for the P8 primer set or 1800 for the P9 primer set was sufficient to uncover all possible sequence variation in each sample. The primer binding regions showed low variability (data not shown).

After filtering for length, 143 and 152 unique sequences were amplified using the P8 and P9 primer sets, respectively, which constituted 96% of all reads. Genome length was 363–370 nt, with a mode of 368 nt (29% of the variant population). However, genomes of 365 nt and 366 nt were also common, each constituting 24.2% of the variant population. A total of 62 unique CVd-VII variants (GenBank accession nos. PQ339054-PQ339115) were discovered after combining P8 and P9 amplicon sequences, and the 10 most abundant variants (V1 to V10) accounted for 64.6% of the merged reads. Only four variants (V1, V2, V20, and V40), all from the same 'clade' (Fig. 1), were found in each of three experiments.

Twenty-seven major variants were identified, each of which accounted for more than 10% of all reads from a sample (Fig. 1). Twenty-five polymorphic positions were observed in these major variants, which increased to 38 polymorphic positions (approximately 10% of the CVd-VII genome) when considering the 62 unique variants detected in this study. Thirty variants were only detected in a single sample, while variants V1 and V2 were found in 24 and 21 samples, respectively.

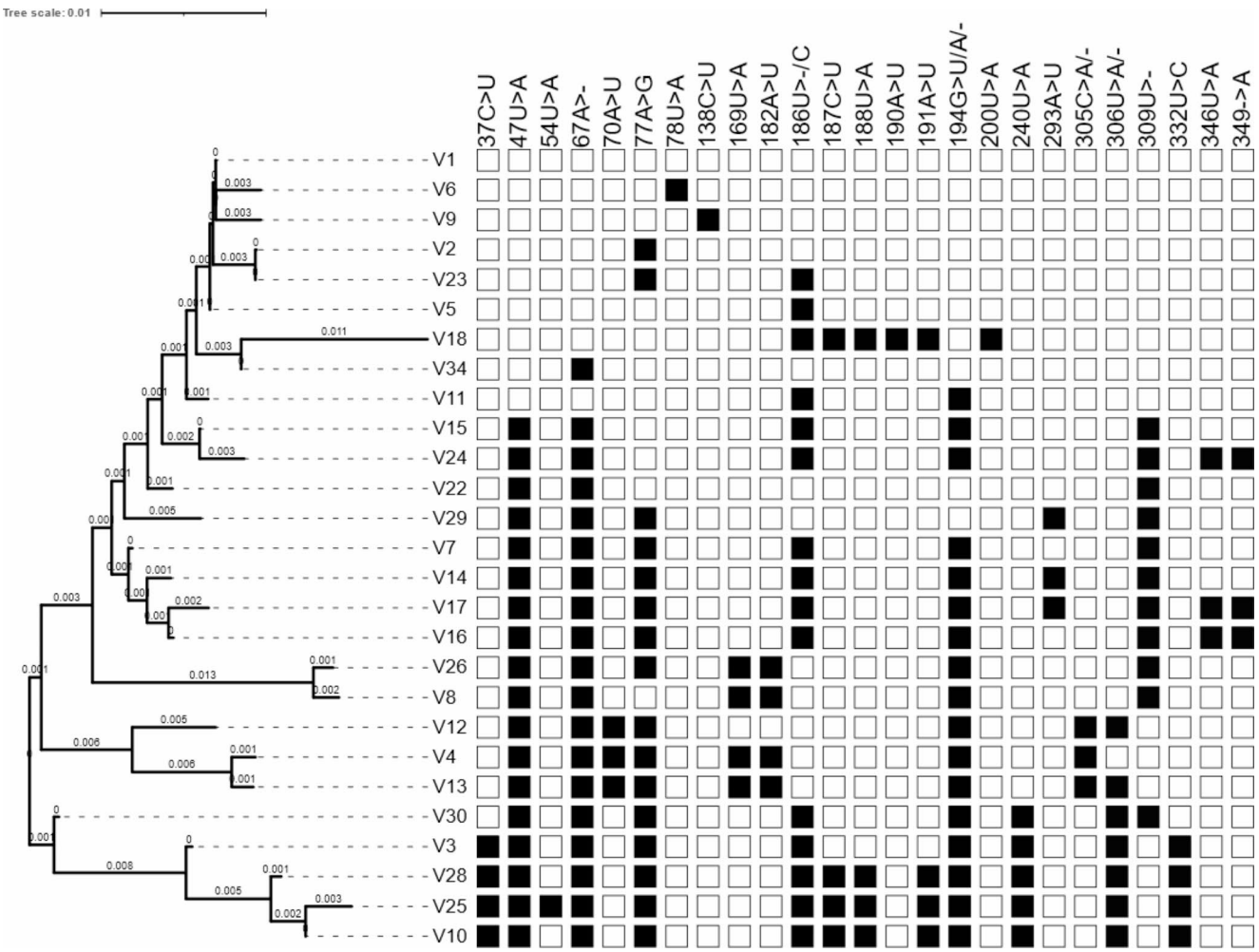

Fig. 1 Phylogenetic tree of the major variants (variants with a relative abundance of > 10% of the total variants found in a sample) of CVd-VII observed throughout the three experiments. The neighbour-joining consensus tree was generated using Geneious Prime software with no outgroup and was rooted with minimum variance using FastRoot. SNP locations are based on the reference sequence of CVd-VII (KX013551), which is identical to variant V1. Filled squares indicate that an SNP is present at a given location

Factors affecting viroid variability

Intraplant variability of CVd-VII

Fourteen unique variants of CVd-VII were detected in a single 'Lisbon' lemon tree, with the three most abundant variants (V1, V2, and V6) accounting for 94% of all sequence reads. These variants were nearly identical, differing by just single SNPs. Variant V1 is identical to the exemplar sequence variant for the viroid species (GenBank accession no. KX013551) reported by Chambers et al. [6]. When compared with V1, five variants contained a single SNP, six variants contained two SNPs, and two variants contained three SNPs. One SNP (77A > G) was detected in nine different variants. Most of the SNPs occurred in the putative pathogenic domain, and the remainder were in the putative T_L (2 SNPs) and T_R (1 SNP) domains. Variant V1 was the dominant variant in all samples except for new-leaf samples, where variant V2 was dominant (Fig. 2). Interestingly, variant V2 was more abundant in new growth (33–40% of the variant profile in bark tissue samples and 74–77% of the variant profile in leaf tissue samples) than in mature flush (0–2% or 1–2% of viroid profile bark and leaf tissues, respectively). The lowest number of variants was found in the new-bark samples, with either two or three variants in each sample.

Interplant variability of CVd-VII

Twenty-two different variants were detected among the eight replicate-inoculated citron plants, seven of which were not observed in the citron source inoculum (Fig. 3). Variants V3 and V4, which have 97.8% nt sequence identity, were the most abundant in citron, accounting for more than 71%

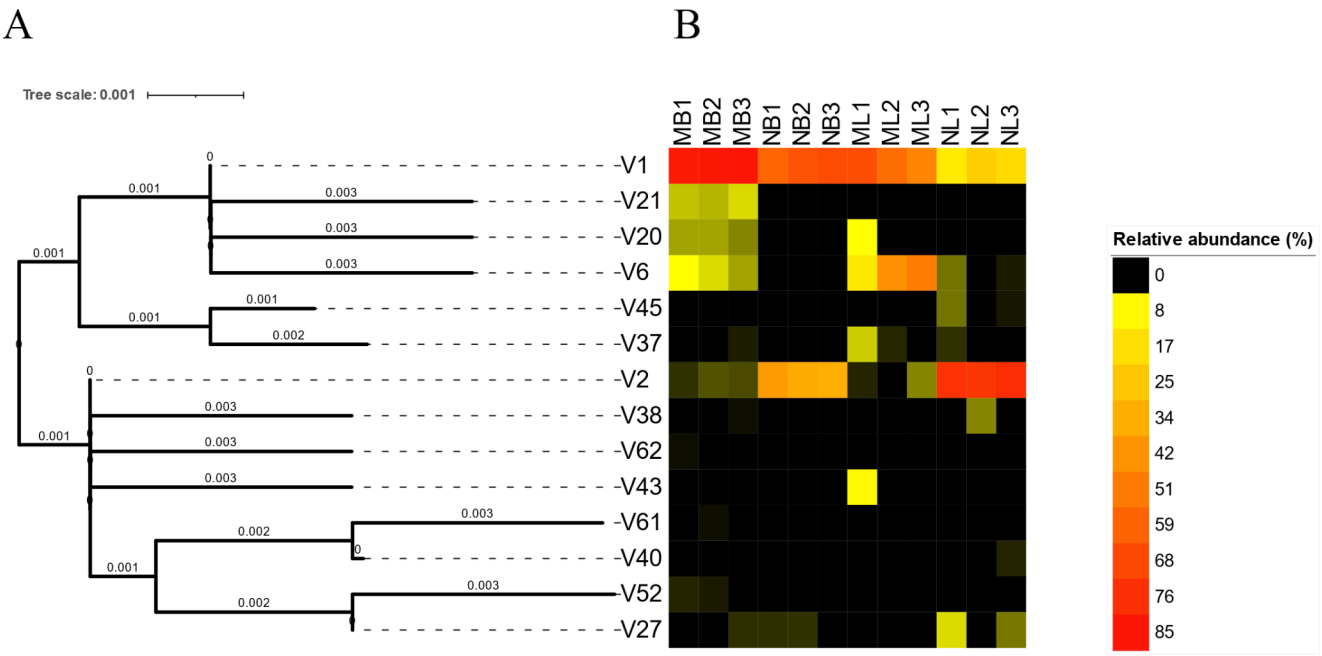

Fig. 2 Analysis of CVd-VII quasispecies profiles in different tissues of a single lemon tree. (**A**) Phylogenetic tree showing the sequence similarity between CVd-VII variants found in different lemon tissues. The neighbour-joining consensus tree was generated using Geneious Prime software with no outgroup and was rooted with minimum variance using FastRoot. (**B**) Relative abundance (%) of each variant in each triplicate sample of different tissues (MB, mature bark samples; ML, mature leaf samples; NB, bark of the most recent flush; NL, leaves of the most recent flush). Red indicates highly abundant variants, yellow indicates low-abundance variants, and black indicates the absence of variants

of all reads. Neither variant was observed in any other host species. Two of the eight inoculated citron plants (replicates P6 and P8) did not contain variant V3, which was the most abundant variant in the other replicate plants and the source citron. However, P6 and P8 contained a unique variant, V10, present at 35.1% and 50.6%, respectively, which was absent from the source and other inoculated citron plants analysed. Plants P6 and P8 also contained two other variants, V25 and V28, that were not seen in other source or inoculated citron plants. All plants in this experiment showed similar disease symptoms, including leaf bending, epinasty, midvein necrosis, stem grooving, and mild dwarfing.

Intervarietal variability in CVd-VII

Thirty-nine variants were observed in the duplicate plants from five different asymptomatic host species that were inoculated with CVd-VII four years prior to sampling, with up to 10 variants detected in a single plant (Fig. 4). Different variants were dominant in the different host varieties. Variants V1 and V2 were detected in five and four host plant samples, respectively, and other variants detected in this experiment were found in one or two samples only. Variant V5 was detected only in a single lemon plant in this experiment and has a sequence almost identical to V1 but with a single nucleotide deletion (SNP105U>-). Major differences were observed between quasispecies profiles of different plants of a single host, in addition to variation among the five host species studied. In the sequence variants identified in this experiment, 29 different SNP sites were observed, primarily located in the putative terminal and pathogenic domains.

Effects of quasispecies variation on the predicted secondary structure of CVd-VII

The secondary structure of all variants that were detected in this study was predicted using the RNA folding program, Mfold, and differences in secondary structure of major variants (> 10% of the relative abundance in a sample) are shown in Fig. 5. Twenty-five loops are present in the secondary structure of the most abundant variant, V1; between 24 and 28 loops are present in each of the variants. The major structural changes in the CVd-VII variants occurred in the putative terminal right domain (Fig. 6). Twenty-six variants, including 13 of the major variants, have a rod-shaped terminal right region, whereas the remaining variants have a bifurcated terminal right region. Primarily, the SNPs 194G > U or 194G > A are associated with the rod-shaped right terminal region for these variants. Variants V3, V4, V10, V12, V13, V25, and V28 have a reduced loop 8 present in the pathogenic domain, whereas V4, V12, and

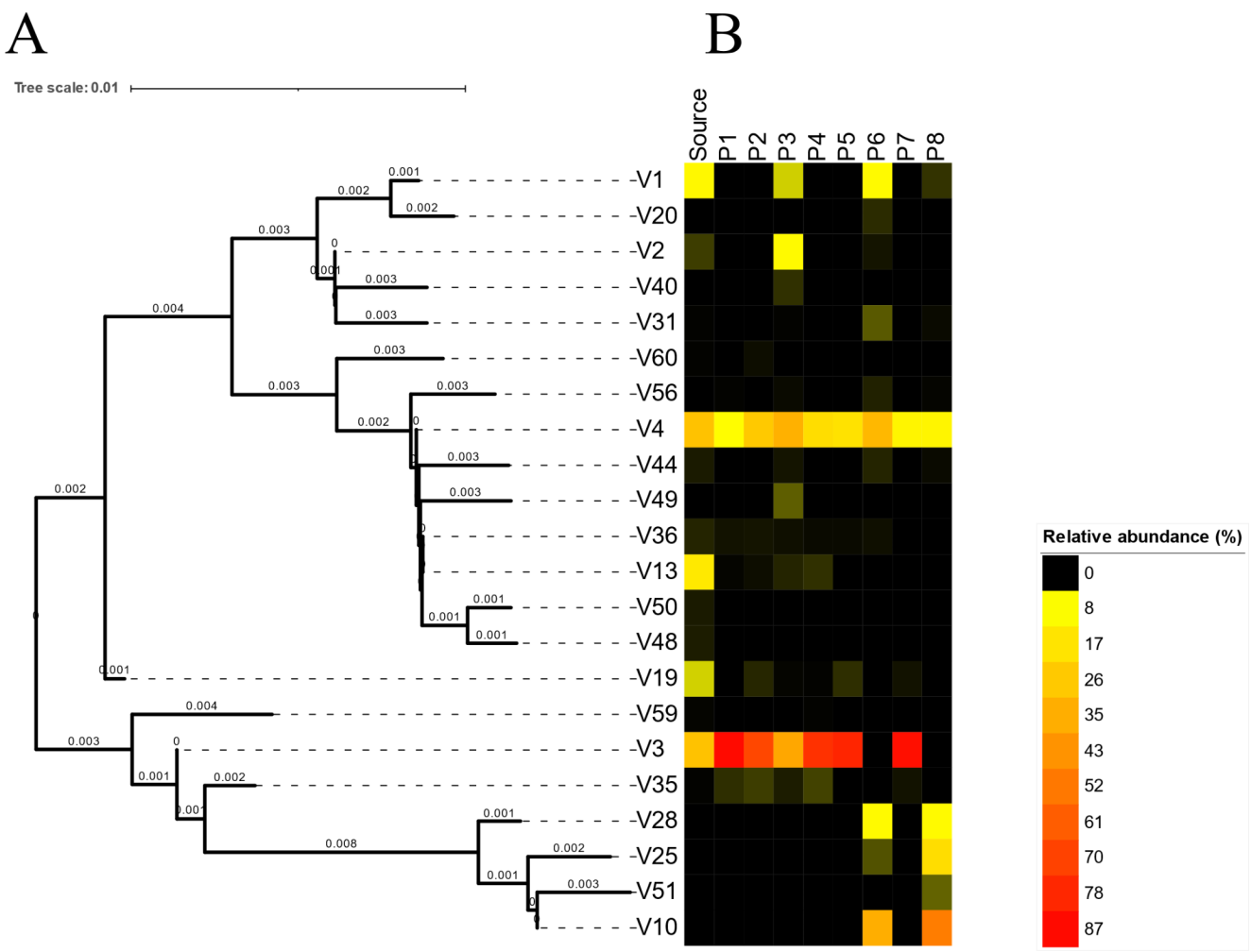

Fig. 3 Analysis of CVd-VII quasispecies profiles in the citron parent (inoculum source) and eight inoculated progeny plants (P1-P8). (**A**) Phylogenetic tree showing the genetic relationships among variants found in the source and inoculated plants. The neighbour-joining consensus tree was generated using Geneious Prime software with no outgroup and was rooted with minimum variance using FastRoot. (**B**) Relative abundance (%) of variants per sample in eight different inoculated plants and the source plant, with red indicating highly abundant variants, yellow indicating low-abundance variants, and black indicating the absence of variants

V13 have an additional small loop in this region. V14, V17, and V29 are missing the large loop 10 that is present in all other variants, due to the 293A > U change in those variants. SNP240U > A reduces the size of loop 18 in the variable region and occurs in the major variants V3, V10, V25, V28, and V30 (Fig. 6).

Discussion

Citrus viroid VII is the most recently discovered and least understood of the suite of viroids that infect citrus. In this study, we investigated factors affecting the evolution of the viroid and identified the sites in the genome that are most prone to mutation. Sixty-two different sequence variants of CVd-VII were observed, but two to four major variants accounted for over 80% of the population. Likewise, two to six sequence variants make up 65–75% of the population of columnea latent viroid (species *Pospiviroid latenscolumneae*), with the remainder being low-frequency variants [46]. At least some of the low-frequency variants may be RT-PCR artifacts or 'dead-end' progeny from viroid replication that are rapidly culled from the population through purifying selection. To resolve these questions, the replication competency of each low-frequency sequence variant would need to be studied by generating infectious clones of CVd-VII.

Variable positions in the viroid genome were primarily located in the putative pathogenic and terminal right domains, with some occurring in the terminal left region; few occurred in the variable or central regions. A single SNP, 240U > A, occurred in the metastable hairpin II (HPII) 5' stem of 10 variants and was predominantly found in citron plants, except for variants V19 and V30, which were

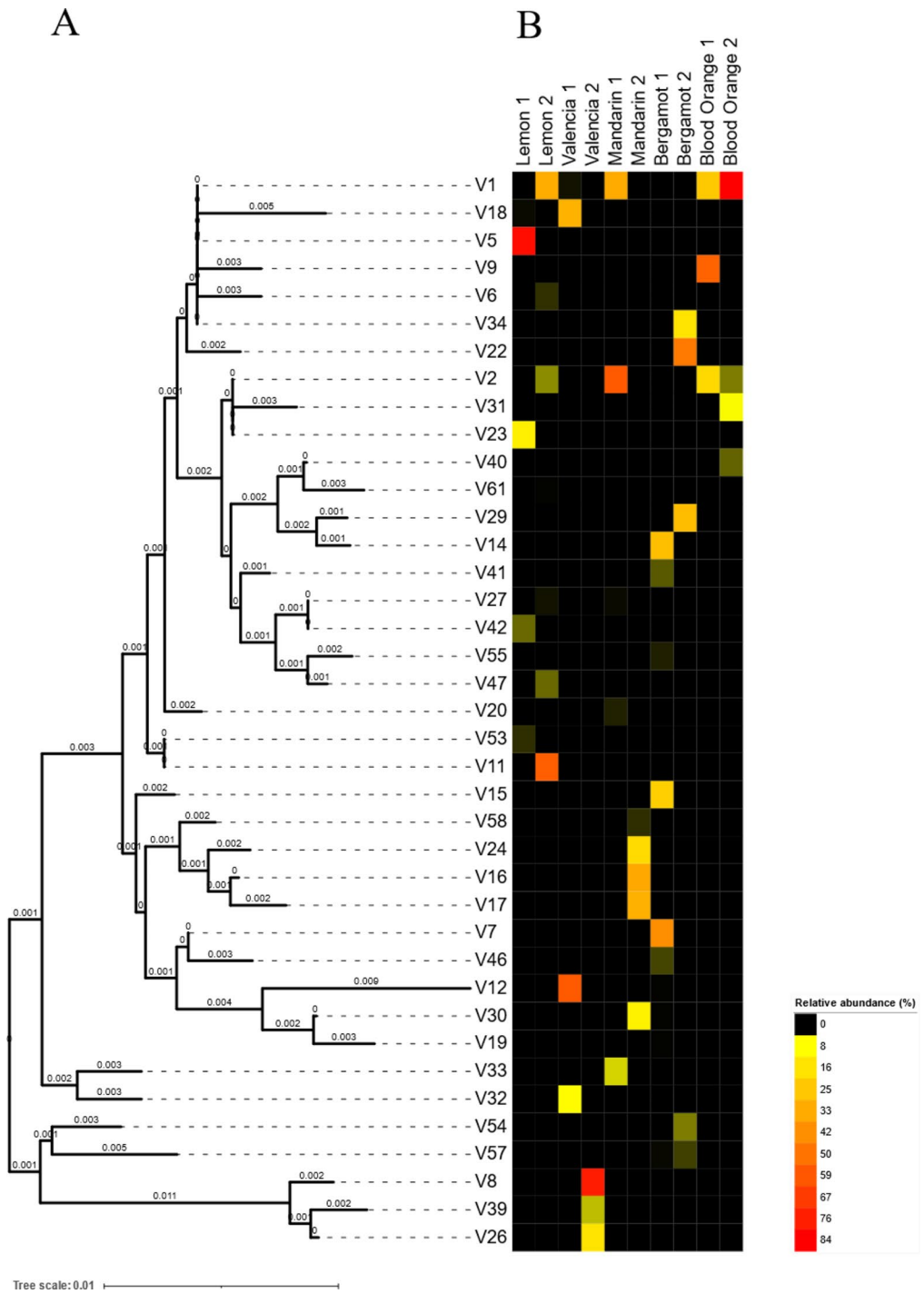

Fig. 4 Analysis of CVd-VII quasispecies profiles in the various citrus host plants. (**A**) Phylogenetic tree showing the sequence similarity between variants found in the duplicate plants of the various host species: lemon (cv. 'Fino'), Valencia orange (cv. 'Midknight'), mandarin (cv. 'Oroval'), bergamot, and blood orange. The neighbour-joining consensus tree was generated using Geneious Prime software with no outgroup and was rooted with minimum variance using FastRoot. The tree scale is shown, and branch lengths are displayed on tree branches. (**B**) Relative abundance (%) of CVd-VII variants per sample in duplicate host plants, with red indicating highly abundant variants, yellow indicating low-abundance variants, and black indicating the absence of variants

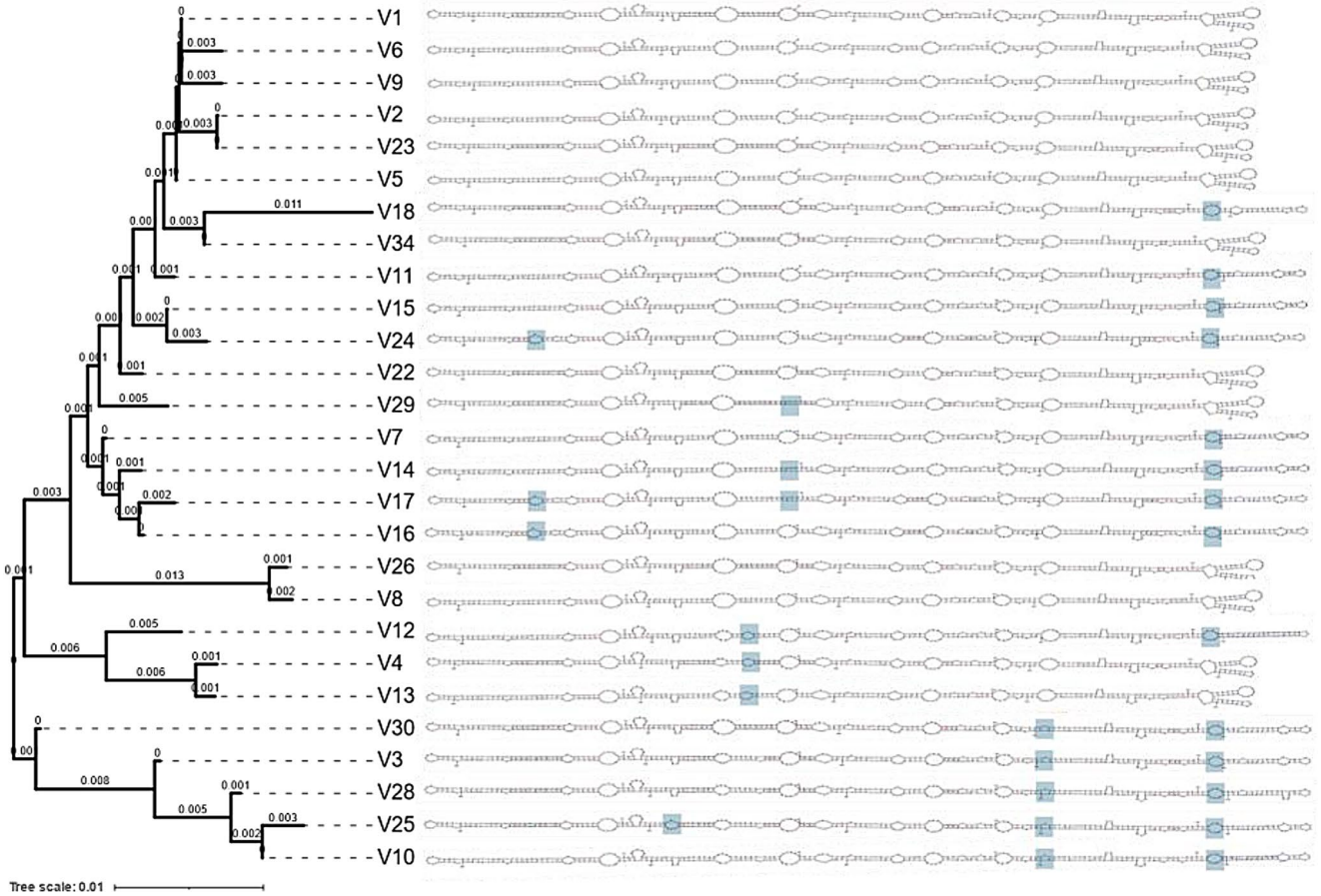

Fig. 5 Phylogenetic tree and predicted secondary structure of CVd-VII major variants. The neighbour-joining consensus tree was generated using Geneious Prime software with no outgroup and was rooted with minimum variance using FastRoot. The tree scale is shown, and branch lengths are displayed on tree branches. The secondary structure of major variants determined in this study was predicted using Mfold. Structural differences from variant V1 are highlighted

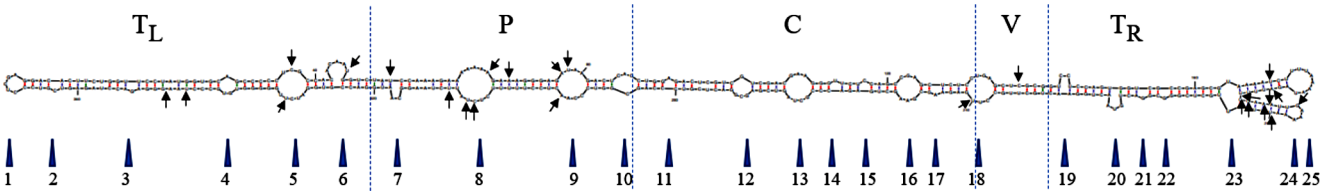

Fig. 6 Location of SNP sites and loops and bulges found in major variants of CVd-VII. SNP locations are indicated by black arrows pointing to the nucleotide involved in the secondary structure of the CVd-VII V1 variant. Putative domains based on [52] terminal left (T_L) pathogenic (P), central (C), variable (V) and terminal right (T_R) domains are delineated by blue dashed lines, while loop/bulge identifying numbers are shown with blue triangles at the position of the feature

also found in bergamot samples and bergamot and mandarin samples, respectively. This SNP leads to a shorter stem of the HPII structure, an element that is critical for infectivity in PSTVd [30]. The secondary structure of hairpin I (HPI) was stable in all variants observed in this study. With the exception of V16, the terminal left and central conserved regions were identical in all variants, consistent with previous findings that these regions play essential roles in PSTVd replication [53]. Two models of the secondary structure of the terminal right region of CVd-VII were predicted: an elongated rod-shaped model and a bifurcated model. Duplicate reverse-complemented repeat regions of nucleotides forming a branched terminus predicted in CVd-VII were similarly predicted in some PSTVd variants in the terminal left region [11]. Most loops that are found in the secondary structure of PSTVd are thought to play a role in the replication and/or trafficking of the viroid [54]. The numerous loops predicted by the secondary structure analysis of CVd-VII are possibly involved in the above processes, but

no clear relationship between secondary structure and host adaptation was revealed in our study.

Spatial differences in the profile of sequence variants of CVd-VII were also observed in a single 'Lisbon' lemon tree. Variant V1 was dominant in all samples taken from the bark and leaves of previous season flushes, and from the bark of the current season flush. However, in the leaf samples of the latest flush, variant V2 dominated over variant V1. In the remainder of the samples, Variant V2 was in relatively low abundance in the samples from the previous season's flush and was a secondary major variant in the bark of the current season's flush. In a study of PSTVd in tomato seedlings, Adkar-Purushothama et al. [1] showed that viroid profiles change over time (1–4 weeks). Wu and Bisaro [51] suggested that some variants of PSTVd are more likely to infect new cells, establishing dominance, as they may have a higher movement capacity than other variants. However, the profile differences may also be attributed to the exclusion of a variant from mature tissue, assuming that variant was present in the mature tissue previously or that variant may only be able to infect new tissue.

In the interplant variability experiment, only one inoculated plant had all of the variants that were present in the source citron plant, indicating that the viroid profile stability among citron plants is low. The inoculated plants P6 and P8 did not contain the most abundant variant observed in this experiment, V3, but the same symptoms were observed in all inoculated trees. This suggests that the genetic changes that had occurred in the quasispecies, particularly in plants P6 and P8, were not associated with symptom development or that the change in variant V3 was not associated with symptom expression. Wu and Bisaro [51] suggested that the evolution of PSTVd quasispecies could be influenced by unpredictable and unknown factors leading to biological replicates evolving distinct profiles or that the generation of quasispecies is a stochastic process. Whatever the cause, such variation was clearly observed in the viroid profiles of the citron plants.

The greatest variation in viroid sequence profiles was observed in our intervarietal variability experiment, with 39 different variants found in the various citrus host species. Host-dependent variability in the genome sequence and structure of viroids is well documented [2, 40, 45, 47], and this was also observed in our study of CVd-VII. The variability of the viroid profiles seen in different hosts may be attributed to the individual host-encoded DNA-dependent RNA polymerase II enzymes used for viroid replication having different error rates, different selection pressures in different plant species, or a combination of these factors. The variation may also be due to a founder effect in which, by chance, the bark sample used for inoculation contains rare sequence variants, which are amplified when a new plant is inoculated.

Analysis of the SNP data provides insight into how quasispecies genomic variation potentially affects published diagnostic assays. The SNP 240U>A is at the location of the 5' end of the reverse primer (VII-R6) used in the RT-qPCR assay for detection of CVd-VII [7]. Viroid variants containing this SNP were primarily found in citron. These include V3 (up to 87% of a sample), V10 (50%), V19 (9%), V25 (20%), V28 (10%), V30 (12%), V35 (2%), V51 (3%), and V59 (0.2%). RT-qPCR analysis of these citron samples revealed that the mismatch does not appear to hinder the amplification of the CVd-VII variants present (data not shown). Therefore, this study provides confidence in the diagnostic assay developed for CVd-VII [7], given its ability to detect multiple variants.

Analysis of viroid quasispecies profiles in this study using high-throughput amplicon sequencing technologies showed that there were multiple variants of the viroid present in all of the plants sampled. This raises the possibility that the findings of earlier studies were misinterpreted [34, 36, 43], where Sanger sequencing of RT-PCR amplicons led to the conclusion that one variant was present in a plant or where a limited number of clones were used to examine genetic variability [6, 25, 27]. Newer technologies provide deeper insight into the viroid populations within a plant; therefore, caution must be used when interpreting studies that employed older techniques.

This study confirms that CVd-VII exists as a quasispecies of closely related variants, as found previously in studies with other viroids [1, 2, 4, 20, 50]. In the current study, we did not investigate the symptom expression of different CVd-VII variants, instead focusing primarily on the composition of the viroid profile in a range of circumstances. Future experiments involving specific variants of CVd-VII would provide insights into the molecular mechanisms of pathogenicity of specific variants. High-throughput amplicon sequencing technologies could be used further in investigations of CVd-VII variant genomic stability over time, variant populations in symptomatic vs. non-symptomatic leaves, and the biological interactions of different variants.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00705-024-06191-4.

Acknowledgements The authors acknowledge the Dharawal people as the traditional custodians of the land on which the experimental work was completed. The authors thank Dr. T. Berg for the critical review of the manuscript.

Author contributions GC, AG, ND contributed to the design and conception of the study. Experiments were conducted by GC, and data analysis was performed by DB and GC. The first draft of the manuscript was written by GC, and all authors edited and commented on it.

Funding was acquired by ND. All authors read and approved the final manuscript.

Funding Funding support from Hort Innovation is acknowledged, using the citrus research and development levy and contributions from the Australian Government under project CT21005 'Improving Australia'a ability to respond to graft transmissible citrus diseases'. Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture.

Data availability The nucleotide sequence data reported here are available in the public GenBank database under the accession numbers PQ339054-PQ339115.

Declarations

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors have no competing interests to disclose.

References

1. Adkar-Purushothama CR, Bolduc F, Bru P, Perreault J-P (2020) Insights into potato spindle tuber viroid quasi-species from infection to disease. Front Microbiol 11:1235
2. Bernad L, Duran-Vila N, Elena SF (2009) Effect of citrus hosts on the generation, maintenance and evolutionary fate of genetic variability of citrus exocortis viroid. J Gen Virol 90:2040–2049
3. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857
4. Brass JR, Owens RA, Matoušek J, Steger G (2017) Viroid quasispecies revealed by deep sequencing. RNA Biol 14:317–325
5. Cao M, Wu Q, Yang F, Wang X, Li R, Zhou C, Li Z (2017) Molecular characterization and phylogenetic analysis of Citrus viroid VI variants from citrus in China. Eur J Plant Pathol 149:885–893
6. Chambers GA, Donovan NJ, Bodaghi S, Jelinek SM, Vidalakis G (2018) A novel citrus viroid found in Australia, tentatively named citrus viroid VII. Arch Virol 163:215–218
7. Chambers GA, Geering AD, Holford P, Vidalakis G, Donovan NJ (2022) Development of a one-step RT-qPCR detection assay for the newly described citrus viroid VII. J Virol Methods 299:114330
8. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890
9. Codoñer FM, Darós J-A, Solé RV, Elena SF (2006) The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. PLoS Pathog 2(12):e136
10. Di Serio F, Owens RA, Li S-F, Matoušek J, Pallás V, Randles JW, Sano T, Verhoeven JTJ, Vidalakis G, Flores R (2021) ICTV Virus Taxonomy Profile: Pospiviroidae. J Gen Virol 102:001543
11. Dingley AJ, Steger G, Esters B, Riesner D, Grzesiek S (2003) Structural characterization of the 69 nucleotide potato spindle tuber viroid left-terminal domain by NMR and thermodynamic analysis. J Mol Biol 334:751–767
12. Duran-Vila N, Roistacher C, Rivera-Bustamante R, Semancik J (1988) A definition of citrus viroid groups and their relationship to the exocortis disease. J Gen Virol 69:3069–3080
13. Duran-Vila N, Elena SF, Daròs J-A, Flores R (2008) Structure and Evolution of Viroids. Origin and Evolution of Viruses. Elsevier, pp 43–64
14. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797
15. Eigen M (1993) Viral Quasispecies. Sci Am 269:42–49
16. Flores R, Hernández C, de Alba AEM, Daròs JA, Di Serio F (2005) Viroids and viroid-host interactions. Annual Rev Phytopathol 43:117–139
17. Flores R, Gas M-E, Molina-Serrano D, Nohales M-Á, Carbonell A, Gago S, De la Peña M, Daròs J-A (2009) Viroid replication: rolling-circles, enzymes and ribozymes. Viruses 1:317–334
18. Flores R, Serra P, Minoia S, Di Serio F, Navarro B (2012) Viroids: from genotype to phenotype just relying on RNA sequence and structural motifs. Front Microbiol 3:217
19. Gago S, Elena SF, Flores R, Sanjuán R (2009) Extremely High Mutation Rate of a Hammerhead Viroid. Science 323:1308–1308
20. Gandía M, Duran-Vila N (2004) Variability of the progeny of a sequence variant Citrus bent leaf viroid (CBLVd). Arch Virol 149:407–416
21. Gandía M, Rubio L, Palacio A, Duran-Vila N (2005) Genetic variation and population structure of a Citrus exocortis viroid (CEVd) isolate and the progenies of infectious haplotypes. Arch Virol 150:1945–1957
22. Glouzon J-PS, Bolduc F, Wang S, Najmanovich RJ, Perreault J-P (2014) Deep-sequencing of the peach latent mosaic viroid reveals new aspects of population heterogeneity. PLoS ONE 9(1):e87297
23. Hajeri S, Ramadugu C, Manjunath K, Ng J, Lee R, Vidalakis G (2011) In vivo generated Citrus exocortis viroid progeny variants display a range of phenotypes with altered levels of replication, systemic accumulation and pathogenicity. Virology 417:400–409
24. Hataya T (2024) Genetic diversity of apple fruit crinkle viroid populations in Japanese persimmons and the infectivity of a predominant sequence variant to tomato plants. Eur J Plant Pathol 169:273–285
25. Ito T, Ieki H, Ozaki K (2000) A population of variants of a viroid closely related to citrus viroid-I in citrus plants. Arch Virol 145:2105–2114
26. Keese P, Symons RH (1985) Domains in viroids: evidence of intermolecular RNA rearrangements and their contribution to viroid evolution. Proc Natl Acad Sci 82:4582–4586
27. Kitabayashi S, Tsushima D, Adkar-Purushothama CR, Sano T (2020) Identification and molecular mechanisms of key nucleotides causing attenuation in pathogenicity of dahlia isolate of potato spindle tuber viroid. Int J Mol Sci 21:7352
28. Leichtfried T, Dobrovolny S, Reisenzein H, Steinkellner S, Gottsberger RA (2019) Apple chlorotic fruit spot viroid: a putative new pathogenic viroid on apple characterized by next-generation sequencing. Arch Virol 164:3137–3140

29. López-Carrasco A, Ballesteros C, Sentandreu V, Delgado S, Gago-Zachert S, Flores R, Sanjuán R (2017) Different rates of spontaneous mutation of chloroplastic and nuclear viroids as determined by high-fidelity ultra-deep sequencing. PLoS Pathog 13(9):e1006547
30. Loss P, Schmitz M, Steger G, Riesner D (1991) Formation of a thermodynamically metastable structure containing hairpin II is critical for infectivity of potato spindle tuber viroid RNA. EMBO J 10:719–727
31. Mai U, Sayyari E, Mirarab S (2017) Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction. PLoS ONE 12(8):e0182238
32. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:3
33. Murcia N, Bernad L, Serra P, Hashemian SB, Duran-Vila N (2009) Molecular and biological characterization of natural variants of Citrus dwarfing viroid. Arch Virol 154:1329–1334
34. Murcia N, Bernad L, Duran-Vila N, Serra P (2011) Two nucleotide positions in the Citrus exocortis viroid RNA associated with symptom expression in Etrog citron but not in experimental herbaceous hosts. Mol Plant Pathol 12:203–208
35. Navarro B, Ambrós S, Di Serio F, Hernández C (2023) On the early identification and characterization of pear blister canker viroid, apple dimple fruit viroid, peach latent mosaic viroid and chrysanthemum chlorotic mottle viroid. Virus Res 323:199012
36. Oksal HD, Aydin S, Baran M, Emir C, Karanfil A, Bozdogan O, Sipahioglu HM (2021) Persimmon (*Diospyros kaki* L.) and Johnsongrass (*Sorghum halepense* (L.) Pers.) Are New Natural Hosts of peach latent mosaic viroid. Appl Ecol Env Res 19:4629–4639
37. Owens R, Thompson S, Feldstein P, Garnsey S (1999) Effects of natural sequence variation on symptom induction by citrus viroid III. Ann Appl Biol 134:73–80
38. Reanwarakorn K, Semancik J (1999) Correlation of hop stunt viroid variants to cachexia and xyloporosis diseases of citrus. Phytopathology 89:568–574
39. Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK (1976) Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. Proceedings of the National Academy of Sciences 73:3852–3856
40. Sano T, Kashiwagi A (2022) Host Selection-producing variations in the genome of hop stunt viroid. Virus Res 311:198706
41. Semancik J, Roistacher C, Rivera-Bustamante R, Duran-Vila N (1988) Citrus cachexia viroid, a new viroid of citrus: relationship to viroids of the exocortis disease complex. J Gen Virol 69:3059–3068
42. Semancik J, Szychowski J, Rakowski A, Symons R (1993) Isolates of citrus exocortis viroid recovered by host and tissue selection. J Gen Virol 74:2427–2436
43. Serra P, Gago S, Duran-Vila N (2008) A single nucleotide change in Hop stunt viroid modulates citrus cachexia symptoms. Virus Res 138:130–134
44. Shannon CE (1948) A Mathematical Theory of Communication. Bell Syst Tech J 27:379–423
45. Suzuki T, Fujibayashi M, Hataya T, Taneda A, He Y-H, Tsushima T, Duraisamy GS, Siglová K, Matoušek J, Sano T (2017) Characterization of host-dependent mutations of apple fruit crinkle viroid replicating in newly identified experimental hosts suggests maintenance of stem–loop structures in the left-hand half of the molecule is important for replication. J Gen Virol 98:506–516
46. Tangkanchanapas P, Haegeman A, Ruttink T, Höfte M, De Jonghe K (2020) Whole-Genome Deep Sequencing Reveals Host-Driven in-planta Evolution of Columnea Latent Viroid (CLVd) Quasi-Species Populations. Int J Mol Sci 21:3262
47. Tessitori M, Rizza S, Reina A, Causarano G, Di Serio F (2013) The genetic diversity of Citrus dwarfing viroid populations is mainly dependent on the infected host species. J Gen Virol 94:687–693
48. Vidalakis G, Davis J, Semancik J (2005) Intra-population diversity between citrus viroid II variants described as agents of cachexia disease. Ann Appl Biol 146:449–458
49. Visvader JE, Symons RH (1985) Eleven new sequence variants of citrus exocortis viroid and the correlation of sequence with pathogenicity. Nucleic Acids Res 13:2907–2920
50. Walia Y, Dhir S, Ram R, Zaidi A, Hallan V (2014) Identification of the herbaceous host range of Apple scar skin viroid and analysis of its progeny variants. Plant Pathol 63:684–690
51. Wu J, Bisaro DM (2024) Cell-cell communication and initial population composition shape the structure of potato spindle tuber viroid quasispecies. Plant Cell 36:1036–1055
52. Wüsthoff K-P, Steger G (2022) Conserved Motifs and Domains in Members of Pospiviroidae. Cells 11:230
53. Zhong X, Leontis N, Qian S, Itaya A, Qi Y, Boris-Lawrie K, Ding B (2006) Tertiary structural and functional analyses of a viroid RNA motif by isostericity matrix and mutagenesis reveal its essential role in replication. J Virol 80:8566–8581
54. Zhong X, Archual AJ, Amin AA, Ding B (2008) A genomic map of viroid RNA motifs critical for replication and systemic trafficking. Plant Cell 20:35–47