Virus Genes 20:3, 243–252, 2000

Both Point Mutation and RNA Recombination Contribute to the Sequence Diversity of Citrus Viroid III

R. A. OWENS,[1] G. YANG,[4] D. GUNDERSEN-RINDAL,[2] R. W. HAMMOND,[1] T. CANDRESSE,[3] & M. BAR-JOSEPH[4]*

[1]*Molecular Plant Pathology and* [2]*Insect Biocontrol Laboratories, Beltsville Agricultural Research Center. Beltsville, MD 20705*
[3]*Station de Pathologie Végétale, INRA, BP 81, 33883 Villenave d'Ornon Cedex, France*
[4]*S. Tolkowsky Laboratory, Department of Virology, Agricultural Research Organization, The Volcani Center, Bet Dagan 50250, Israel*

Received September 8, 1999; Accepted December 13, 1999

Abstract. Field-grown citrus trees often harbor complex mixtures of 4–5 different viroid species, and the presence of citrus viroid III (CVd-III) has been shown to reduce the rate of tree growth without inducing disease. To more fully define the structure of its quasi-species, we have examined nine citrus viroid complexes for the presence of previously undescribed sequence variants of CVd-III. Analysis of 86 full-length cDNAs generated from these nine viroid complexes by RT-PCR revealed the presence of 20 new CVd-III variants. Chain lengths ranged from 293–297 nucleotides, and sequence changes were confined largely to the lower portions of the central conserved region and variable domain. The previously described variants CVd-IIIa (297 nt) and CVd-IIIb (294 nt) were clearly predominant, but phylogenetic analysis indicated that certain isolates may contain representatives of two additional fitness peaks. At least one group of CVd-III variants appears to have arisen as a result of RNA recombination. Populations recovered from diseased/declining trees were the most diverse, but even dwarfing isolates originating from old line Shamouti trees showed considerable variability.

Key words: viroids, apscaviroids, phylogenetic analysis, RNA quasi-species, subviral RNAs

Introduction

Viroids—small, highly structured, single-stranded circular RNA molecules lacking both a protein capsid and detectable mRNA activity—are the smallest known agents of infectious disease. More than 20 individual viroid species have been described (Flores et al., 1998), nearly all of which were originally identified based on diseases induced in one or more species of crop plants. Much effort has been devoted to identifying the nature of the molecular signal(s) that mediate viroid-host interaction and symptom development (e.g. Schnölzer et al., 1985; Owens et al., 1996). Often overlooked in this emphasis on their disease-causing potential is the fact that some viroid infections can actually be beneficial to their hosts in an agricultural sense.

As described by Duran-Vila et al. (1988), individual field-grown citrus trees may harbor as many as 4–5 different species of viroids. Some of these viroids (e.g. citrus exocortis viroid; CEVd) cause specific disease symptoms, while the presence of others reduces tree growth without inducing specific symptoms (Gillings et al., 1991). The use of dwarf trees offers many potential economic and environmental benefits to the grower; indeed, several studies of the horticultural effects of so-called ''graft transmissible dwarfing factors'' predate the demonstration that citrus exocortis is a viroid disease (Mendel, 1968; Semancik and Weathers, 1972). Several different groups have described the use of

*Corresponding author: Molecular Plant Pathology Laboratory, Room 252 Building 011A, Beltsville Agricultural Research Center, Beltsville, MD 20705. E-mail: rowens@asrr.arsusda.gov
Disclaimer: Mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

different viroids for dwarfing and enhancing production of ''Valencia'' orange and grapefruit grafted on viroid-sensitive rootstocks (Broadbent et al., 1986; Bar-Joseph, 1993; Semancik et al., 1997). These dwarfing trials were carried out using naturally derived viroid isolates that may contain complex mixtures of sequence variants (Visvader and Symons, 1985; Ben Shaul et al., 1995). Citrus viroid III (CVd-III—a member of the genus Apscaviroid; Flores et al., 1998), in particular, shows great promise as a dwarfing agent (Albanese et al., 1996; Semancik et al., 1997; Hutton et al., 1999), but little is known about how its biological properties may be affected by specific sequence changes.

The rod-like ''native'' structure of potato spindle tuber viroid (PSTVd) and related viroids is divided into five domains whose boundaries are defined by sharp changes in sequence similarity (Keese and Symons, 1985). Mutational analysis of PSTVd and construction of intraspecific viroid chimeras have shown that sequences within at least three of these domains (i.e. the left terminal loop, pathogenicity domain, and variable domain/right terminal loop) play important roles in modulating symptom expression (e.g. Visvader and Symons, 1986; Sano et al., 1992). To date, the sequences of four different CVd-III variants have been published (Rakowski et al., 1994; Stasys et al., 1995; Semancik et al., 1997). Most of the sequence differences between CVd-IIIa, -IIIb, -IIIb (Australia), and -IIIc, are located in the lower portion of the central conserved region and variable domain. Their effects on either the dwarfing response or symptom expression in indicator hosts are unknown.

We are interested in the development of improved citrus viroid dwarfing agents, especially sequence variants of CVd-III that are specifically adapted to certain rootstock-scion combinations. Recent advances in understanding the role of quasi-species in the rapid evolution of RNA genomes (Holland and Domingo, 1998) suggest that it may be possible to adapt methods developed for the selection and amplification of rare functional nucleic acids *in vitro* (see Joyce, 1994) for use with viroids *in vivo*. To better define the structure of the CVd-III quasi-species, we have examined nine citrus viroid complexes for the presence of previously underscribed sequence variants of CVd-III. CVd-IIIa and -IIIb are clearly the predominant natural variants, but our analysis suggests that certain complexes may contain CVd-III variants representing additional fitness peaks.

Materials and Methods

Origin of CVd-III Field Isolates

Nine citrus viroid field isolates known to contain CVd-III were examined for the presence of sequence variation. Seven isolates (i.e. GTD 225-S, Alumot, Minneola decline, Navalate, Shamouti Shaar Efraim, Shamouti Nir Galim, and Shamouti Zofit) were from Israel, and two isolates (i.e. EEFB 3 and EEFB 6; Villalobos et al., 1997) were from Costa Rica. The origins of Nir Galim and GTD 225-S have been described previously (Ben-Shaul et al., 1995). Dwarfing isolates Shaar Efraim and Zofit were derived from Shamouti sweet orange, while Alumot was isolated from Eureka lemon. The Minneola isolate was obtained from declining scions grafted on Volkameriana rootstocks, and the Navalate isolate was obtained from budwood originating in Spain. In addition to CVd-III, most of these isolates contain several other viroid species (Ben-Shaul et al., 1995; MB-J, unpublished data).

Amplification and Sequence Analysis of CVd-III cDNAs

Full-length double-stranded CVd-III cDNAs were synthesized by RT-PCR using antisense primer C2′ (5′-ACTCTCCGTCTTTACTCCAC-3′, complementary to nucleotides 138-119) and sense primer H2′ (5′-CTCCGCTAGTCGGAAAGACT-3′, homologous to nucleotides 139–158) essentially as previously described (Yang et al., 1992; Rakowski et al., 1994). AmpliTaq DNA polymerase (Perkin Elmer) was used for all PCR amplifications. Following cloning in plasmid vector pCR II (Invitrogen, Carlsbad, CA), the nucleotide sequences of individual CVd-III cDNAs were determined by automated sequence analysis using a Model 373A sequencer (Applied Biosystems) and dye-labeled M13 forward and reverse sequencing primers. Approximately 10 randomly selected clones from each isolate were analyzed, and all clones were sequenced in both directions.

Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignments were generated with the Clustal W program (Thomson et al., 1994) implemented through the MegAlign (Version 1.10) function of the Lasergene software package from

DNASTAR (Madison, WI). Minor adjustments were manually introduced in the final alignment to maximize sequence homology. Phylogenetic relationships among CVd-III sequence variants were evaluated using PAUP Version 3.1.1 (Swofford, 1993) and SplitsTree (Huson, 1998). For analysis using PAUP, a phylogenetic tree was constructed by heuristic search via random stepwise addition implementing the tree bisection and reconnection branch-swapping algorithm. The analysis was replicated 100 times, producing an unrooted strict consensus tree. Bootstrap analysis (100 replicates; Felsenstein, 1987) was performed to estimate support for inferred clades. Possible recombination events were identified using the test developed by Sawyer (1989) as implemented in the program VTDIST2 (provided by S. Sawyer, Washington University, St. Louis).

Structural Calculations and GenBank Accession Numbers

Lowest free-energy structures for sequence variants of CVd-III were calculated using the Zuker-Turner *mfold* RNA folding package available on the World Wide Web (http://www.ibc.wustl.edu/∼zucker/rna/node3.html#MATZURA). Nucleotide sequences of variants III(4)–III(29) together with the alignment of 24 CVd-III sequences used for phylogenetic analysis have been deposited in GenBank (Accession numbers AF123858-60 and AF123863-79). This information has also been added to a catalog of viroid sequences maintained on the World Wide Web (http://www.callisto.si.usherb.ca/∼jpperra).

Results

Identification of CVd-III Sequence Variants

Nucleotide sequences of 86 independent CVd-III cDNA clones derived from nine citrus viroid isolates were compared to those of the four known CVd-III sequence variants; i.e. CVd-IIIa, CVd-IIIb, CVd-IIIb(Australia), and CVd-IIIc (Semancik et al., 1997). As shown in Table 1, a total of 20 previously undescribed CVd-III sequence variants were detected. Fourteen of these were recovered only once, and recovery frequencies for the other variants ranged between two [i.e. III(26)] and ten [i.e. III(04)] times each. CVd-IIIa and IIIb were clearly the most prevalent sequence variants—representing, respectively, 21/86 (24.4%) and 22/86 (25.6%) of the clones examined—but these variants were not always the predominant species in individual field isolates. Other sequence variants appeared to comprise $\geq$ 50% of the CVd-III population in the isolates from Alumot, Nir Galim, Minneola, Navalate, and EEFB6. Results from previous analyzes (e.g. Semancik et al., 1997) suggested that the sequences covered by primers C2′ and H2′ were unlikely to exhibit significant variation;

Table 1. Sequence heterogeneity within CVd-III isolates

	Sequence Variants																					
Isolate	*IIIa*	*IIIb*	04	05	06	09	10	13	14	15	16	18	19	20	21	22	23	24	25	26	27	29
GTD complexes																						
Alumot			7	1	2																	
Shaar Efraim	9					1																
Nir Galim	2	2	2		1		1	1	1													
GTD 225-S	9				1																	
Zofit	1	6	1		1													1				
Other Sources																						
Minneola		3					2			4	1											
Navalate		2										1	1	1	1	1	1					
EEFB#3		6																	1	2		
EEFB#6		3																	4		1	1
*Frequency**	24.4	25.6	11.6		5.8		3.5			4.7									5.8	2.3		

*Expressed as percentage of total clones examined.

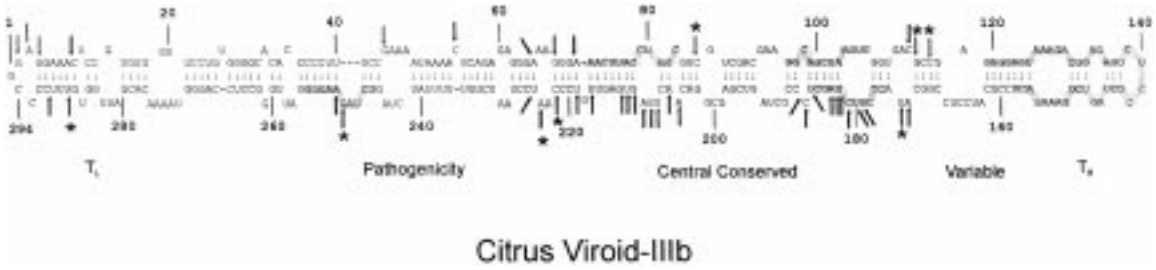


Fig. 1. Lowest free energy structure of CVd-IIIb. Nucleotides involved in the formation of secondary hairpins I (upper strand) and II (lower strand) as well as those complementary to primers C2′ and H2′ (right terminal loop) have been shaded. Variable positions are indicated by arrows. Asterisks denote variation detected in only a single PCR-amplified cDNA.

thus, we did not feel it necessary to repeat our amplification and sequencing studies with a second set of oligonucleotides designed to prime outside the right terminal loop.

To facilitate analysis of the possible functional significance of the variation observed, sequences of all 24 variants were aligned using CVd-IIIb (Rakowski et al., 1994) as the reference sequence. After minor manual adjustment to maximize homology, differences were found to occur at 41 of 299 positions in the aligned sequences. As shown in Fig. 1, these changes were clustered in four portions of the molecule; i.e. positions 288–291 (left terminal loop), positions 44–71 (upper portion of putative pathogenicity domain), and positions 173–187a plus 205–225 (lower portion of the central conserved region). Two other isolated positions (i.e. nucleotides 113 and 251) were also highly variable. Our analysis did not include positions 119–158 targeted by primer pair C2′ + H2′; thus, the calculated value for overall CVd-III sequence variability (i.e. 41/299 = 13.7%) provides only a minimum estimate.

As described by Semancik and colleagues (Rakowski et al., 1994; Semancik et al., 1997), the sequences of CVd-IIIa and CVd-IIIc differ from that of CVd-IIIb at 13 and 16 positions, respectively. The newly described variants CVd-III(04)–(29) differ from CVd-IIIb at as few as one and as many as 21 positions. A total of 47 sequence changes (i.e. 19 transitions, 15 transversions, 5 insertions, and 8 deletions) have now been detected in CVd-III, 36 of which were present in two or more independent cDNA clones. Changes observed in only a single clone could result from nucleotide mis-incorporation by the comparatively error-prone *Taq* DNA polymerase used for PCR; thus, the origin of variants 5, 9, 13, 20, 21, 23, and 24 is uncertain. Overall lengths of the new CVd-III variants range from 293–297 nt.

Phylogenetic Analysis of CVd-III

Phylogenetic analysis of the 24 aligned CVd-III sequences as carried out using three approaches: First, maximum parsimony (as implemented by the PAUP package); second, split decomposition (as implemented by the SplitsTree program); and third, Sawyer's test for recombination (as implemented by VTDIST2 program). All three analyses revealed similar patterns of sequence relationships.

As shown in Fig. 2A, the strict consensus of 25 heuristic trees produced by maximum parsimony analysis indicated that CVd-III sequence variants comprise two major clusters. One cluster of variants contains CVd-IIIb, CVd-IIIb(Aus), and 12 newly described variants. The second cluster appears to contain two subgroups: i.e. CVd-IIIa plus variants 9 and 14 as well as a more diverse grouping containing CVd-IIIc and six new variants. The bootstrap value for the initial step separating CVd-IIIb from these second and third lineages was comparatively low (i.e. 58%), reflecting the small number of sequence changes involved. Within the large cluster containing CVd-IIIb, newly described variants 24, 26, and 27 occupy a short weakly supported branch.

Fig. 2B presents the relationships between the same CVd-III sequence variants as determined by the SplitsTree algorithm. As described by Huson (1998), the split decomposition analysis strategy does not attempt to ''force'' data onto a phylogenetic tree, rather, the results of such an analysis can provide a

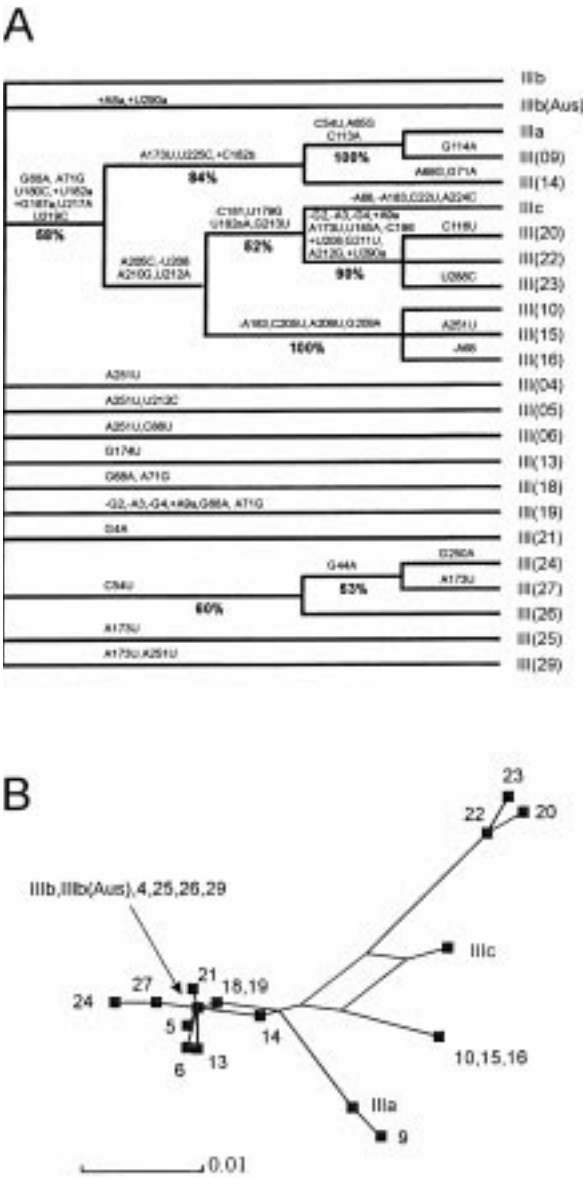


Fig. 2. Phylogenetic relationships among CVd-III sequence variants. (A) Strict consensus of 25 heuristic trees obtained by maximum parsimony analysis. Nature and location of sequence changes are shown to the right of each node; e.g. G68A indicates an A/G transition at CVd-IIIb position 68. Percentage values indicate statistical significance as determined by bootstrap analysis (100 replicates). (B) Split decomposition analysis as implemented by the SplitsTree algorithm (Huson, 1998).

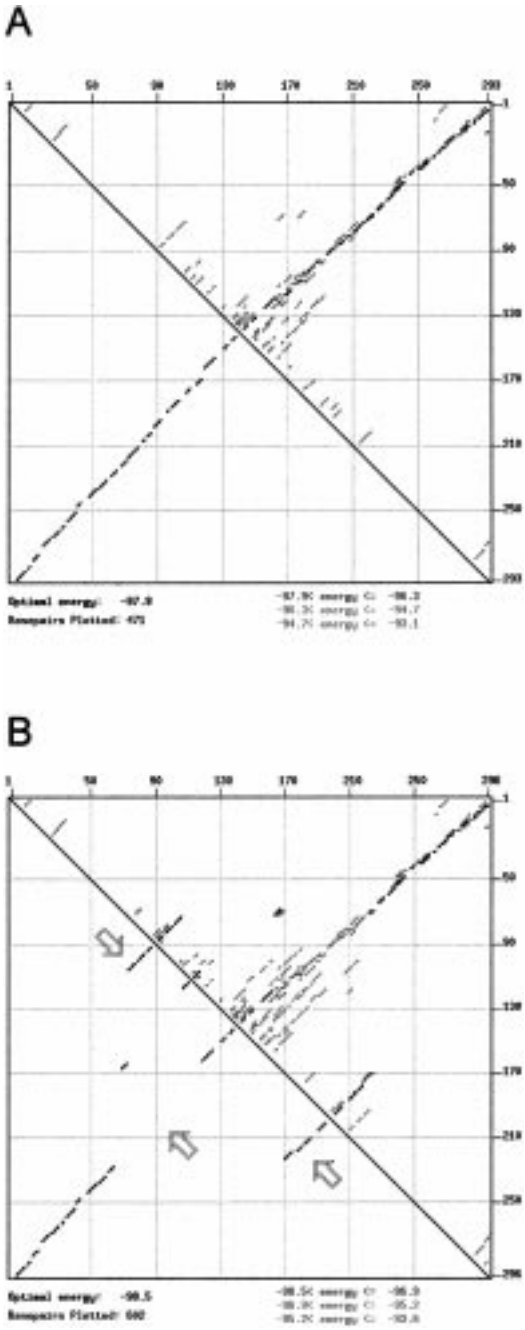


Fig. 3. Secondary structures of (A) CVd-IIIb and (B) hypothetical phylogenetic intermediate H2 from the lineage leading to CVd-IIIa. The dark dots in lower left portion of each panel indicate the lowest free energy structure(s), while the lighter gray dots in upper right portion define alternative, suboptimal structures whose calculated free energies differ by $\leq$ 5%. Arrows denote alternative secondary structures for the upper and lower portions of the central conserved region that are available only to the hypothetical phylogenetic intermediate.

good indication of how tree-like a given data set is. CVd-IIIb and related variants (including 24, 26 and 27) can be seen to group toward the left side of Fig. 2B. On the right side of the diagram, relationships between CVd-IIIa, CVd-IIIc, and the other variants comprising the second and third lineages appear much the same as those predicted by maximum parsimony analysis.

Neither maximum parsimony nor split decomposition explicitly considers the possible effect of recombination on the predicted phylogeny. Because RNA recombination is thought to play an important role in viroid evolution (e.g. Keese and Symons, 1985; Kofalvi et al., 1997; Spieker, 1996), we examined our collection of CVd-III sequence variants for evidence of recombination using a statistical test devised by Sawyer (1989) as implemented by the programs VTDIST and VTDIST2. This test allows a statistical evaluation of the probability of genetic rearrangement (recombination) based on imbalances in the size distributions of regions in which pairs of sequence within a data set are identical. For CVd-III, the calculated probability for the *absence* of recombination ranged between $< 1/10{,}000$ (VTDIST) and $< 7/1{,}000$ (VTDIST2). Thus, as far as Sawyer's test is concerned, the probability that recombination had taken place was quite high.

As mentioned above, alignment of all 24 CVd-III sequences revealed four major regions of sequence variation. Mapping these sequence changes on the phylogenetic tree shown in Fig. 3A revealed that the changes affecting positions 173–187a and 217–224 exhibit near-constant covariation. Five different sequence patterns could be identified. From top to bottom, these patterns are found in: CVd-IIIa plus variants 9 and 14; CVd-IIIc; variants 20, 22, and 23; variants 10, 15, and 16; and, finally, CVd-IIIb together with the remaining 12 variants. If this apparent covariation were real, the fact that the areas involved lie on the same rather than opposite strands of the native structure would call into question the validity of a rod-like structure for CVd-III (see Fig. 1). This point is dealt with in more detail in the Discussion.

Alternatively, these two regions could be co-inherited as a block. In such a situation, three groups of sequence could be considered to be recombinants: Variant 14 could be the result of recombination between a parental 18-like variant and CVd-IIIa (source of the 173–224 region). CVd-IIIc and variants 10, 15, and 16 could be the result of a similar process involving replacement of the 173–224 region in CVd-IIIa by the corresponding portion of an as-yet-unidentified variant. The final and clearest example of a possible recombinant is variant 19. This molecule could have arisen as the result of an 18-like variant having acquired the left terminal loop region from a sequence variant like 20, 22, or 23. Because recombination is a symmetrical process, variant 19 can also be considered to be a 20, 22, or 23-like variant that has acquired the 173–224 region from an 18-like variant. Consistent with such a scenario, variants 18–23 were all recovered from a single, geographically distinct isolate (i.e. Navalate; see Table 1).

Structural Effects of CVd-III Sequence Variation

Each of the sequence changes detected in our collection of variants was mapped on the consensus phylogenetic tree shown in Fig. 2A. Because most of these changes were not obviously compensatory, we decided to evaluate their possible structural effects by comparing the lowest free-energy structures of the naturally-occurring CVd-III variants with those of the intervening evolutionary intermediates. The results of these analyses are summarized in Fig. 3.

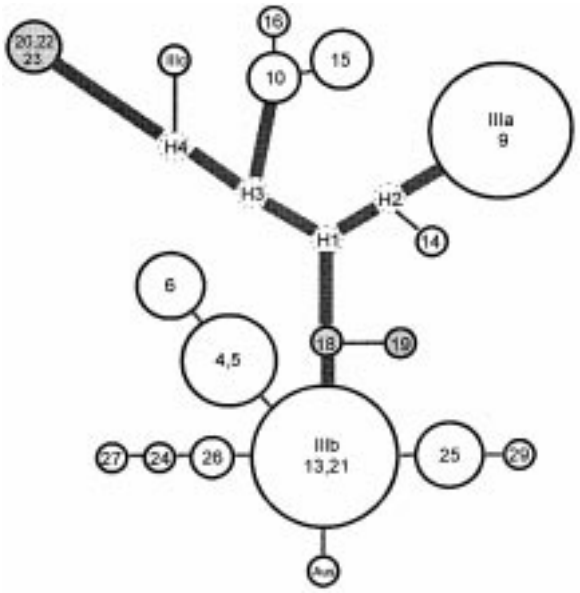
Fig. 3A shows the rod-like lowest free-energy structure predicted for CVd-IIIb. In this ''dot plot'' representation, each base pair shown in Fig. 1 is represented by a single dot along the diagonal in the lower left half of the panel. The upper right half of Fig. 3A compares this lowest free energy structure with a series of suboptimal, less energetically favored structures. With only one exception, all naturally occurring variants and hypothetical intermediates were predicted to have very similar structures. The single exception was a hypothetical intermediate in the lineage containing variants 9, 14, and CVd-IIIa. As shown in Fig. 3B, the central portion of its lowest free energy dot plot is empty. Inspection of the upper right portion of Fig. 3B reveals that its predicted secondary structure has begun to shift from rod-like to partially cruciform. Sequence changes at positions 173, 182b, and 225 have weakened the central portion of rod-like native structure sufficiently that nucleotides in the upper and lower portions of the central conserved region are able to form small hairpin loops (marked by arrows) rather than interact only with nucleotides on the opposing strand. Consistent with these theoretical predictions, the results of tempera-

ture gradient gel electrophoresis also indicated that CVd-IIIb has a rod-like, highly base-paired structure (results not shown).

A variety of evidence (e.g. Qu et al., 1993) indicates that viroid replication is dependent upon the formation of certain alternative structural interactions upon disruption of the rod-like native structure. Locations of CVd-III sequences involved in the most thoroughly studied of these interactions (i.e. the so-called ''secondary hairpins'' I and II) are indicated by shading in Fig. 1. Note that none of the variation affects positions 72–83 or 96–107, sequences involved in formation of secondary hairpin I. The situation with respect to secondary hairpin II, however, is less clear. Firstly, the pattern of sequence changes affecting secondary hairpin II is not obviously complementary; i.e. changes between positions 176–185 are much more extensive than those between positions 246–255. Secondly, only the U > C transition at position 180 and a G > A transition at position 250 lead to ''silent'' base-pairing changes within the context of the (−)strand of CVd-III. Finally, the frequently observed A > U transversion at position 251 would be expected to disrupt secondary hairpin II within the context of either the (+) or (−) strand.

Discussion

When our studies began, only four sequence variants of CVd-III had been described. Comparison of the sequences of CVd-IIIa, IIIb, and IIIb(Australia), and IIIc (Rakowski et al., 1994; Stasys et al., 1995; Semancik et al., 1997) suggested that variability was confined largely in the lower portions of the central conserved and variable domains, but the full extent of natural sequence variability was unclear. Sequences comprising the right terminal loop of CVd-III and targeted by the primer set used in our analyzes appeared to be highly conserved. Our studies have shown that CVd-III is considerably more variable in sequence than previously indicated. Variability remains concentrated within the lower portions of the central conserved and variable domains, but several changes affecting the left terminal loop of CVd-III were also identified. Most importantly, phylogenetic analysis of a total of 24 sequence variants provides new insight into the structure of the CVd-III quasispecies.



Composition of individual CVd-III populations

Isolate	Variant clusters IIIa	IIIb	10,15,16	20,22,23	Other
Alumot	−	+++	−	−	−
Shaar Efraim	+++	−	−	−	−
Nir Galim	+	++	−	−	14
GTD 225	++	+	−	−	−
Minnesia	−	++	++	−	−
Navalate	−	++	−	+	18,19
Zafit	+	++	−	−	−
EEFB #3	+++	−	−	−	−
EEFB #6	+++	−	−	−	−

Fig. 4. Quasispecies structure and composition of individual CVd-III populations. Sequence relationships between naturally-occurring CVd-III variants were determined by maximum parsimony bootstrap analysis, and variants differing by only a single unique sequence change have been grouped. (*Above*) In this schematic representation, lengths of lines connecting the different variants in the quasispecies are directly proportional to the number of sequence changes observed. To maintain overall proportions, diameters of the circles enclosing each variant (or group of variants) are proportional to the square root of their observed frequency. Variants possibly arising as a result of RNA recombination *in vivo* are indicated by shading. H1–H4, hypothetical CVd-III variants predicted by maximum parsimony analysis but not recovered from infected trees. (*Below*) Genetic diversity of individual CVd-III populations. +++, only variant(s) detected: ++, majority variant; +, minority variant; −, not detected.

As shown in Fig. 4, the fitness landscape occupied by the CVd-III quasispecies appears to contain two major and two minor peaks. Here, data on relative prevalence of each sequence variant (Table 1) has been combined with the phylogenetic relationships identified by maximum parsimony and split decomposition methodologies (Fig. 2). Lengths of the connecting segments are proportional to the number of underlying sequence changes whereas, to maintain an appropriate scale, the diameters of the circles representing the different variants are proportional to the square root of their frequency. Variants containing only a single, unique sequence change have been combined with their nearest neighbors to reduce the effect of possible PCR-related nucleotide misincorporation.

The fitness of a virus or viroid variant is not easy to determine (e.g. Miralles et al., 1999), and we have not yet attempted to compare the relative fitness of individual CVd-III sequence variants. Nevertheless, the previously described variants CVd-IIIa and IIIb are clearly the predominant sequences in the quasispecies. CVd-IIIa appears to occupy a rather isolated position in this landscape, and our studies detected only a single closely related sequence variant. CVd-IIIb, in contrast, lies at the center of a large cluster of closely related variants. Each of the short lineages originating from CVd-IIIb begins with a single, frequently observed sequence change; i.e. $A_{251} \rightarrow T$ for variants 4–6; $C_{54} \rightarrow T$ for variants 24, 26, and 27; and $A_{173} \rightarrow T$ for variants 25 and 29. CVd-IIIb(Aus) contains two opposing single nucleotide insertions in the left terminal loop. Two groups of less common variants (i.e. variants 10, 15, and 16 and variants 20, 22, and 23) complete this landscape.

Phylogenetic analysis distributes these groupings along a Y-shaped pathway containing four hypothetical intermediates. The unrooted nature of this analysis means that neither CVd-IIIa nor CVd-IIIb can be considered as the ancestral state. Interestingly, this pathway contains only one extant sequence variant. CVd-III(18) differs from CVd-IIIb at only two positions (i.e. 68 and 71) and could arise spontaneously at a relatively high frequency. Hypothetical intermediates H1–H4, in contrast, each contain 3–11 sequence changes.

Computer analysis revealed that introduction of just two of the three changes required to convert H1 into H2 is sufficient to disrupt the central portion of the rod-like structure of CVd-III, thereby allowing the molecule to assume a cruciform conformation (see Fig. 3B). Efforts are currently underway to use information about the hemaglutinin gene of human influenza virus gathered by phylogenetic analysis to predict the epidemic potential of circulating virus strains (Fitch et al., 1997), and it is tempting to speculate about the possible relationship between the structural properties of phylogenetic intermediate H2 and the comparatively isolated position of CVd-IIIa within the quasispecies. Much additional work will be required to determine the number and timing of sequence changes in CVd-III required to move from one fitness peak to another, however.

In a broader context, their small size and ability to induce often-dramatic disease symptoms have made viroids and plant viral satellite RNAs favored subjects for studies of molecular evolution. Recently, Ambros et al., 1998 have shown that the biological properties of three phenotypically different isolates of peach latent mosaic viroid are correlated with both population complexity and the presence of specific sequence variants. Genetic diversity of viroid and satellite RNA populations is limited by the need to maintain a functional secondary/tertiary structure (Ambros et al., 1998; Aranda et al., 1997). CVd-III appears to be more tolerant of sequence variation than PSTVd but less tolerant than either citrus exocortis or grapevine yellow speckle viroid (Visvader and Symons, 1985; Gora et al., 1994; Syzchowski et al., 1998).

As summarized in Fig. 4 (lower portion), five of the nine CVd-III populations that we examined proved to be genetically diverse. The most diverse populations, Navalate and Minneola, were isolated from diseased or declining trees, and in one case (i.e. Navalate), RNA recombination appears to be responsible for some of the diversity observed. The source of variability in the other populations—competition among the sequence variants that continuously arise or superinfection—is unclear. Kearney et al. (1993) have shown that rate of genetic drift for foreign sequences replicating in an RNA plant virus vector (TMV) is very low (i.e. $\leq 10^{-4}$ mutations per base per passage), but studies of PSTVd and PLMVd population dynamics suggest that viroids may be considerably more mutable (Góra et al., 1994; Góra-Sochacka et al., 1997; Ambros et al., 1998). Individual cDNA clones derived from severe isolates of PSTVd or PLMVd are quite variable in their pathogenicity, and propagation of most individual

PSTVd variants in Rutgers tomato quickly leads to the appearance of complex populations—sometimes in as little as one plant passage (Góra-Sochacka et al., 1997). Interestingly, some of these new variants may be defective in late functions such as movement within the plant.

Finally, we note that the three CVd-III populations maintained in Shamouti sweet orange by grafting differed markedly in their degree of homogeneity. Multiple infection events provide one possible explanation for the differences observed between the Shaar Efraim, Nir Galim, and Zofit, isolates, but other explanations are also possible. Mutations can become fixed in a population via either competitive selection or ''bottlenecking'' where the inoculum is not representative of the population under study. Long assumed to be best described by deterministic evolutionary models because of the large number of circulating virus particles, HIV-1 evolution has recently been shown to be strongly influenced by stochastic processes during suboptimal protease-inhibitor therapy (Nijhuis et al., 1998). Thus, the *effective population size* of a pathogen may be much smaller than previously recognized. Each of these dwarfing complexes contains a different mixture of viroids. Although cross protection among viroids is usually considered to be group-specific (Niblett et al., 1978), the presence of unrelated viroids may affect the population structure of CVd-III in more subtle ways.

Acknowledgments

We thank S.M. Thompson and Y. Nemchinova for skilled technical assistance in the collection of sequence data. The comments of S. Scheffer (USDA/ARS, Beltsville) regarding the evolutionary implications of our data were particularly helpful in preparing the manuscript.

References

Albanese G., LaRosa R., Tessitori M., Fuggetta E., and Catara A., Proceedings of 13th Conference IOCV. (Riverside), pp. 367–369, 1996.

Ambros S., Hernández C., Desvignes J.C., and Flores R., J Virol *72*, 7397–7406, 1998.

Aranda M.A., Fraile A., Dopazo J., Malpica J.M., and Garcia-Arenal F., J. Mol Evol *44*, 81–88, 1997.

Bar-Joseph M., Acta Horticulturae *349*, 271–276, 1993.

Ben-Shaul A., Guang Y., Mogilner N., Hadas R., Mawassi M. Gafny R., and Bar-Joseph M., Phytopathology *85*, 359–364, 1995.

Broadbent P., Forsyth J.B., Hutton R.J., and Bevington K.B., Calif Citrograph *71*, 8–10, 1986.

Duran-Vila N., Roistacher C.N., Rivera-Bustamente R., and Semancik J.S., J Gen Virol *69*, 3069–3080, 1988.

Felsenstein J., Evolution *39*, 783–791, 1987.

Fitch W.M., Bush R.M., Bender C.A., and Cox N.J., Proc Natl Acad Sci USA *94*, 7712–7718, 1997.

Flores R., Randles J.W., Bar-Joseph M., and Diener T.O., Arch Virol *143*, 623–629, 1998.

Gillings M.R., Broadbent, and Gollnow B.I., Aus J., Plant Physiol *18*, 559–570, 1991.

Góra A., Candresse T., and Zagórski W., Arch. Virol *138*, 233–245, 1994.

Góra-Sochacka A., Kierzek A., Candresse T., and Zagórski W., RNA *3*, 68–74, 1997.

Holland J. and Domingo E., Virus Genes *16*, 13–21, 1998.

Huson D.H., Bioinformatics *14*, 68–73, 1998.

Hutton R.J., Broadbent P., and Bevington K.B., Hort Rev *24*, 277–317, 1999.

Joyce G.F., Curr Opinion Struc Biol *4*, 331–336, 1994.

Kearney C.M., Donson J., Jones G.E. and Dawson W.O., Virology *192*, 11–17, 1993.

Keese P. and Symons R.H., Proc Natl Acad Sci USA *82*, 4582–4586, 1985.

Kofalvi S.A., Marcos J.F., Cañizares M.C., Pallás V., and Candresse T., J Gen Virol *78*, 3177–3186, 1997.

Mendel K., Proc Inter Org Citrus Virol *4*, 310–313, 1968.

Miralles R., Gerrish P.J., Moya A., and Elena S.F., Science *285*, 1745–1747, 1999.

Niblett C.L., Dickson E., Fernow K.H., Horst R.K., and Zaitlin M., Virology *91*, 198–203, 1978.

Nijhuis M., Boucher C.A.B., Schipper P., Leitner T., Schurman R., and Albert J., Proc Natl Acad Sci USA *95*, 14441–14446, 1998.

Owens R.A., Steger G., Hu Y., Fels A., Hammond R.W. and Riesner D., Virology *222*, 144–158, 1996.

Polivka H., Staub U., and Gross H.J., I J Gen Virol *77*, 155–161, 1996.

Qu F., Heinrich C. Loss P. Steger G., Tien P., and Riesner D., EMBO J. *12*, 2129–2139, 1993.

Rakowski A., Szychowski J.A., Avena Z.S., and Semancik J.S., J Gen Virol *75*, 3581–3584, 1994.

Sano T., Candresse T., Hammond R.W., Diener T.O., and Owens R.A., Proc Natl Acad Sci USA *89*, 10104–10108, 1992.

Sawyer S., Mol Biol. Evol *6*, 526–538, 1989.

Schnölzer M., Haas B., Ramm K., Hofmann H., and Sänger H.L., EMBO J *4*, 2181–2190, 1985.

Semancik J.S. and Weathers L.G., Nature *235*, 242–244, 1972.

Semancik J.S., Rakowski A.G., Bash J.A., and Gumpf D.J., J Hort Sci *72*, 563–570, 1997.

Spieker R.L., J Gen Virol *77*, 2839–2846, 1996.

Stasys R.A., Dry I.B., and Rezaian M.A., FEBS Letters *358*, 182–184, 1995.

Swofford D.L., Phylogenetic analysis using parsimony (PAUP), version 3.1.1. University of Illinois, Champaign, IL, 1993.

Szychowski J.A., Credi R., Reanwarakorn K., and Semancik J.S., Virology *248*, 432–444, 1998.

Thomson J.D., Higgins D.G., and Gibson T.J., CLUSTAL W. Nucleic Acids Res *22*, 4673–4680, 1994.

Villalobos W., Rivera C., and Hammond R.W., Rev Biol Trop *45*, 983–987, 1997.

Visvader J.E. and Symons R.H., Nucleic Acids Res *13*, 2907–2920, 1985.

Visvader J.E. and Symons R.H., EMBO J *5*, 2051–2055, 1986.

Yang X., Hadidi A., and Garnsey S.M., Phytopathology *82*, 279–285, 1992.