Virus Genes 19:2, 131–142, 1999

Molecular Characterization of an Isolate of Citrus Tristeza Virus that Causes Severe Symptoms in Sweet Orange

ZHONG-NAN YANG,[1] DEBORAH M. MATHEWS[2] J. ALLAN DODDS,[2] & T. ERIK MIRKOV[1]*

[1]*Dept. of Plant Pathology and Microbiology, The Texas A&M University Agricultural Experiment Station, Weslaco, TX 78596; and*
[2]*Dept. of Plant Pathology, University of California Riverside, Riverside, CA 92521*

Received February 8, 1999; Accepted March 25, 1999

Abstract. The complete sequence (19,249 nucleotides) of the genome of citrus tristeza virus (CTV) isolate SY568 was determined. The genome organization is identical to that of the previously determined CTV-T36 and CTV-VT isolates. Sequence comparisons revealed that CTV-SY568, a severe stem-pitting isolate from California, has more than 87% overall sequence identity with CTV-VT, a seedling yellows isolate from Israel. Although SY568 has an overall sequence identity of 81% with CTV-T36, a quick decline isolate from Florida, the sequence identity in the 3′ half of the genome is over 90% while the sequence identity in the 5′ half of the genome is as low as 56%. Based on the sequence alignments of these three isolates, sequences in the 3′ half of the genome are generally well conserved, while the sequences in the 5′ half are relatively divergent. Sequence data of independent overlapping clones from the CTV-SY568 genome revealed two regions with highly divergent sequences. In open reading frame 1b (RNA dependent RNA polymerase), there were 118 nucleotide differences that lead to 16 amino acid changes. In the open reading frame of the divergent coat protein gene, 5 amino acid changes result from 48 nucleotide differences. Most differences occurred in the third position of the codons, and resulted in silent amino acid substitutions. RNase protection assays demonstrated that most of the clones obtained are representative of the major RNA species of this isolate. Northern analysis indicated that CTV-SY568 accumulated more viral RNA including genomic and certain subgenomic RNAs than isolates VT or T36 in sweet orange.

Key words: citrus, tristeza, stem pitting, SY568, closterovirus, RNase protection

Introduction

Citrus tristeza virus (CTV) is one of the most economically important citrus pathogens worldwide (1). Isolates of CTV are traditionally distinguished by their reactivity in greenhouse grown plants of Mexican lime, grapefruit, sour orange, sweet orange and other citrus species (2). It is common for isolates to produce vein clearing and stem pitting in Mexican lime, but reactions in other hosts vary considerably between isolates and serve as a basis for strain discrimination. Chlorosis, stunting and stem pitting are the most common symptoms recorded. Isolates that cause chlorosis or pitting in one species may or may not cause these symptoms in other species.

Of the hosts mentioned above, sweet orange is the least likely to give a strong reaction to a CTV isolate. An unusually severe isolate of CTV was isolated from a Minneola tangelo in California in 1978 (3) and is now called SY568 in the University of California, Riverside collection of CTV strains. It causes severe reactions in all of the hosts mentioned above including sweet orange. This makes SY568 one of the most severe forms of CTV known, resembling the strains that cause severe field disease in sweet orange in Brazil and Peru (4–6). It has retained its severity through numerous passages in sweet orange or Mexican lime. Some biological variants have been obtained from SY568 by passage through *Passiflora gracilis* (5) indicating the possibility that this isolate,

*Dept. of Plant Pathology and Microbiology, The Texas A&M University Agricultural Experiment Station, Weslaco, TX 78596. E-mail: E-mirkov@tamu.edu

like most other CTV isolates, is a mixture of strains (7–9).

The 10–12 × 2000 nm virus particles of CTV contain a single-stranded, positive-sense genomic RNA and are the largest known single-component plant RNA virus (10). The complete sequence of the genome of two CTV isolates, T36 from Florida and VT from Israel, have previously been reported (11,12). Both genomes contain 11 open reading frames (ORFs) potentially encoding at least 17 protein products. The 5′ proximal ORF encodes a large polyprotein containing domains characteristic of two papain-like proteases (P-PRO), a methyl transferase (MT), a helicase (HEL), and an RNA-dependent RNA polymerase (RdRp) via a presumed +1 ribosomal frameshift from ORF1a to ORF1b. The downstream ORFs 2 to 11 encode potential protein products of 33 kDa (p33), 6 kDa (p6), a homologue of cellular heat shock proteins (HSP70), 61 kDa (p61), a diverged copy of the coat protein (dCP), coat protein (CP), 18 kDa (p18), 13 kDa (p13), 20 kDa (p20) and 23 kDa (p23) (11,13,14). Although the genome organizations of CTV-T36 and CTV-VT are identical, sequence comparisons of these two isolates revealed asymmetrical sequence conservation. The 3′ half of the genome is conserved, while identity in 5′ half of the genome is so low that these two isolates of CTV could be defined as separate viruses (12). Of the putative proteins, only dCP and CP are present in CTV particles (15); p20 has been detected in CTV infected tissue (14).

The severity of isolate SY568, and the threat it could present were it to be disseminated (6) make it a logical choice for molecular characterization, thereby permitting future mapping of the molecular determinants responsible for the appearance of severe symptoms, including stem pitting.

Methods

Biological Properties and Passage History of CTV Isolate SY568

This isolate was found initially in 1977 in a Minneola tangelo, but was also observed to be spreading to other experimental trees in field 12B at the University of California, Riverside (3). When indexed, budwood from declining trees induced a severe seedling yellows (SY) reaction (stunting and chlorosis) in seedlings of grapefruit, sour orange, lemon and sweet orange, plus severe stem pitting symptoms and vein corking of leaves in Mexican lime, grapefruit and sweet orange. It was highly transmissible by *Aphis gossypii* without loss of virulence (16). An isolate was maintained in sweet orange (*Citrus sinensis* Osbeck.), which was used as a source of graft inoculum for a single sweet orange plant which was maintained in a 5 gallon pot from 1988 to the present without further subculture. This plant is cut back once or twice a year, and each growth flush shows stunted poor growth with the erratic development of vein corking in some leaves. All stems are heavily pitted during all growth periods. With the exception of one cDNA clone obtained from SY568 infected grapefruit, viral RNA from this plant was used to obtain the complete genome sequence of SY568.

PCR Amplification and Cloning Strategy

Total RNA was purified from infected leaves with a RNeasy Plant Total RNA Kit (QIAGEN Inc., Chatsworth, CA). CTV dsRNA was isolated from bark tissue collected from twigs as described (17). Primers used for polymerase chain reaction (PCR) amplification are listed in Table 1. The 3′ half of the genome of CTV-SY568 (clones pTEM24, pZY16A, pZY16B, pZY16C, pZY15A, pZY15B, pZY18A, pZY18B, pZY23A, pZY23B, pZY20A, pZY20B, pZY19A and pZY19B; Fig. 1) was cloned using a one tube reverse-transcriptase-PCR (RT-PCR) primer walking strategy essentially as described (18). Briefly, 10 μg total RNA in 20 μl H_2O plus 200 ng each primer was incubated at 70°C for 5 min. This was added to a RT-PCR reaction containing 20 mM Tris, pH 8.5, 50 mM KCl, 2.5 mM $MgCl_2$, 0.075% Triton X-100, 0.001% Gelatin, 10 mM DTT, 0.2 mM each dNTP, 10 units RNasin (Promega Corp., Madison, WI), 15 units AMV reverse transcriptase (Promega), and 2.5 units AmpliTaq polymerase (Perkin-Elmer Corp., Norwalk, CT) in a final volume of 100 μl. Reverse transcription and PCR were carried out in a Perkin-Elmer Cetus 480 thermal cycler programmed to give one cycle at 42°C (60 min), two cycles at 94°C (5 min), 37°C (3 min), 72°C (3 min), and 35 cycles at 94°C (1 min), 52°C (1 min), 72°C (1–3 min), with a final cycle of 72°C (7 min).

The 5′ half of the genome of CTV-SY568 (clones pZY5A, pZY5B, pZY30A, pZY30B, pZY40A, pZY40B, pZY80A and pZY80B; Fig. 1) was cloned

Table 1. Primers used for cloning and corresponding clones obtained

Clone	Primer	Sequences[a]	Source of Sequences
pTEM24	F214	ATCGATAAACTTAAAGCATCATT	Sekiya et al., 1991
	R215	CCGAGTCTATGTTAGCTAGACGTC	pTJ555
pZY19	CTVF4	CGTCGCTTTGTTTAGCGCG	SY568 (pTEM24)
	CTV3′	CTATGGGGGGCCAACATAGGTCCAT	T36
pZY23	F1700	GCGAGTACTCAAAGAAGTACCCGAGC	T36
	P61/27R1	CCCAAATCGCGTAGGTTA	SY568 (pTEM24)
PZY20	p3444	GTCGGTAGATGCGTTAAT	SY568 (pZY23A)
	pCla	AGTCGAATTTAGCCAACTCGTAAC	SY568 (pZY19A)
PZY18	P65, 5′	GAAGTGCTTCCCCTTACGGCACTC	T36
	R1920	GCGCATTCAAAAGACTCTGCCTAC	SY568 (pZY23A)
PZY15	p3′pol	CGCTATTCATTGTGTGCGATCGAA	T36
	p14.3′	CCAGAAGCCGCTAAATATTCCGCA	SY568 (pZY18A)
PZY16	pol5′	TGAGGTCGTCGTTGGTTCTTCTGT	T36
	p16.3′	ACGTTGGCTCTCAGCGAGAATTGA	SY568 (pZY15A)
PZY80	F7920	ACTTCACGGGAMKTTGRG	T36, BYV and LIYV
	R9430	GGACGCTTTTCTTGGCGGAATTGC	SY568 (pZY16A)
pZY5	F20	CAAATTCACCCGTACCCTCCGGA	T36 and VT
	R1550	TGATGGTTAGGAAYAGATC	T36 and VT
pZY30	R3040	TCTTCACCACCCATCTTATG	T36 and VT
	F1490	CACTTTTGGCAGTGTATTTCA	SY568 (pZY5A)
pZY40	F3040	CATAAGATGGGTGGTGAAGA	SY568 (pZY30A)
	R4050	GGTAAACAACTACGAGTCTTA	T36 and VT
pZY70	F4050	TAAGACTCGTAGTTGTTTACC	SY568 (pZY40A)
	R7990	TCATCGTCYAAATCCAT	SY568 (pZY80A)
pZYN	dT16	AATTCGGATCCTTTTTTTTTTTTTTTN	
	5R1	CGAACGTAGCAATGGCCG	SY568 (pZY5A)
	5R2	GATATCATCCAACGCTGC	SY568 (pZY5A)
pZYC	F716	TGGTTATGCACACCAAAG	SY568 (pZY19A)
	dT16	AATTCGGATCCTTTTTTTTTTTTTTTN	

[a]: N = G + A + T + C, M = A + C, K = T + G, R = A + G, Y = C + T

using a two step RT-PCR strategy. First strand cDNA was synthesized with SuperScript II reverse transcriptase (Gibco BRL, Gaithersburg, MD) as described in Yang & Mirkov (18). One µl of synthesized first strand cDNA was used as a template for the PCR step. The PCR reaction mix (50 µl final volume) consisted of 250 ng each primer, 60 mM Tris-HCl (pH 8.5), 75 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 2.5 units AmpliTaq DNA polymerase and PCR cycles were the same as with the one tube RT-PCR except that the reverse transcription step was omitted.

A 10 µl aliquot of the PCR products was analyzed by electrophoresis in a 1% agarose gel (19). Reaction products of the expected sizes were excised from the gel and eluted using a GENECLEAN II Kit (BIO 101 Inc., La Jolla, CA), and the DNA fragments were cloned directly into plasmid vector pCRII (Invitrogen, San Diego, CA).

Clone pZY70 was obtained using the LA PCR Kit (Takara Shuzo Co Ltd, Japan). Two µl of the reverse transcription product were used as template. PCR conditions were: 98°C (20 sec), 52°C (1 min), 68°C (5 min), 40 cycles; 72°C (7 min). The 3.9 kb PCR product was digested into three fragments with *Hind*III (Fig. 1), and the two end fragments were cloned into *Hind*III-*Sma*I digested pGEM-7Z(+) while the middle fragment was cloned into the *Hind*III site of pGEM-7Zf(+).

The terminal 3′ and 5′ sequences were determined by obtaining clones pZYN and pZYC. DsRNA of CTV-SY568 was denatured in 10 mM methyl mercuric hydroxide (Alfa AESAR, Ward Hill, MA). A poly(A) tail was added to the 3′ end of this denatured RNA with yeast poly(A) polymerase according to the manufacturer's instructions (USB, Cleveland, Ohio). First-strand cDNA was synthesized

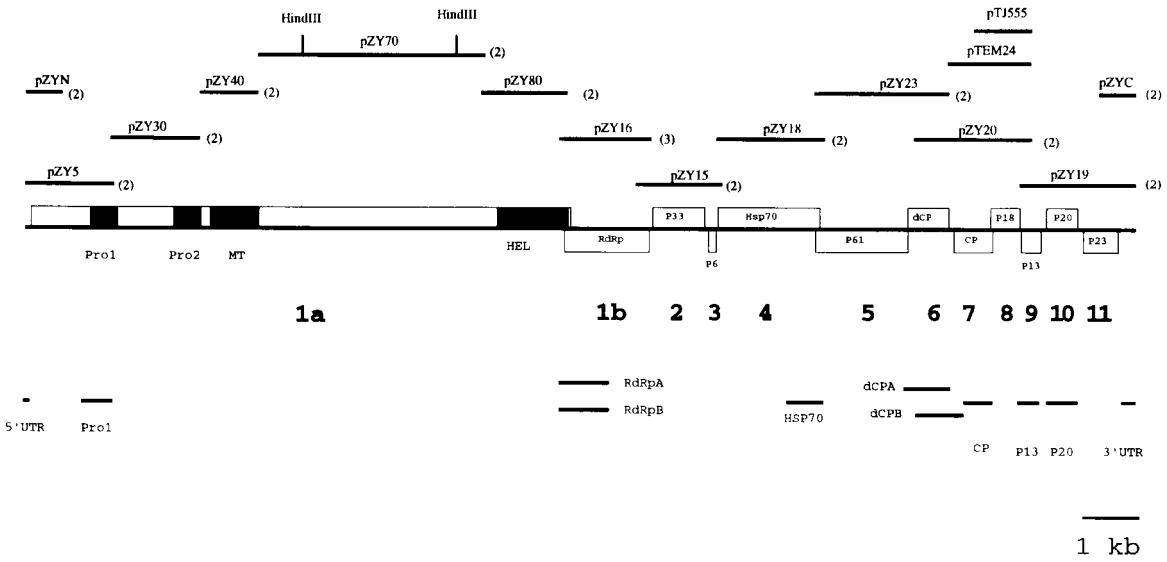

Fig. 1. Schematic representation of cDNA clones of CTV isolate SY568. Clones pZY80, pZY16, pZY15, pZY18, pZY23, pZY20 and pZY19 were obtained by RT-PCR using degenerate or CTV-T36 specific primers; Clones pZY5, pZY30 and pZY40 were obtained with primers based on sequences of CTV-T36 and CTV-VT; Clone pZY70 was obtained using primers based on the sequences of clones pZY40 and pZY80. Numbers in parentheses indicate the number of independent clones sequenced for each region. The numbers below the schematic representation of the genome organization represent the ORFs. Lines below the ORF description represent the probes used for RNase protection assays targeting the following regions: 5′UTR, Pro1, RdRp, HSP70, dCP, CP, p13, p20 and 3′UTR.

using the two step RT-PCR described above, with poly(A) tailed dsRNA as template and oligo dT_{16} as primer. One µl of the cDNA was used for PCR amplification. For the 3′ terminal cloning, we used primer pair dT16 and F716 (Table 1). For the 5′ terminal cloning, we used a nested PCR approach. Primer pair dT16 and 5R1 (Table 1) was used for the primary PCR. Then 2 µl of this 50 µl PCR product was used as template for the next PCR with primer pair dT16 and 5R2 (Table 1). The secondary PCR products were purified and cloned as above.

DNA Sequence

The cloned cDNAs were sequenced using the Sequenase 2.0 dideoxy chain termination method (USB) as detailed by the manufacturers protocol using external T7, SP6, or M13 reverse primers. Internal oligonucleotides based on the previously determined sequence were used for primer walking. Both strands of at least two independent clones were sequenced. Independent clones were designated with ''A'', ''B'' and ''C''. Sequences were analyzed with the Genetics Computer Group (GCG) sequence analysis software from the University of Wisconsin (20). Sequences were edited with the program ''SEQED'', and were assembled with ''FRAGMENT ASSEMBLY''. The program ''PLOTSIMILARITY'' was used for the complete nucleotide sequence comparisons. The program ''GAP'' was used for the nucleotide and amino acid sequence comparisons of ORFs and UTRs. The complete sequence of CTV-SY568 was deposited in the Genbank under accession number AF001623.

RNase Protection Assays and dsRNA Analysis

DsRNAs (17) of six CTV isolates maintained in sweet orange were analyzed: T505, T516, SY560, SY563, SY565, and SY568. DsRNAs were electrophoresed through 6% polyacrylamide mini-gels and visualized by staining with ethidium bromide.

In total, 11 probes (Fig. 1) were used for the RNase protection assays (RPAs) which correspond to the following nucleotide positions: 5′UTR (8–130), Prol (996–1553), RdRpA (9230–10115), RdRpB (9230–10115), HSP70 (13225–13878), dCPA (15202–16006), dCPB (15403–16257), CP (16256–16616), P13 (17224–17609), P20 (17737–18288) and 3′UTR (18992–19249). Clones used for transcription tem-

plates are p5m, pZY5A, pTEM16A, pTEM16B, pZY18A, pTEM23, pTEM20, pTEMcp, p1904, pGem/p20 and pZY19A. pZY5A, pZY18A, pZY19A are the original clones used for sequencing. Clone p5m is derived from pZY5A in which the *Ecl*136II fragment was deleted. Clone p1904 is derived from pZY19A in which the *Bam*HI fragment was deleted. Clone pGem/p20 is also a subclone of pZY19A in which the *Hpa*I to *Afl*III fragment (552 bps) was made blunt and ligated into the *Sma*I site of pGEM-7Zf(+). pTEMcp is a clone of the coat protein gene from which the *Kpn*I to *Eag*I fragment isolated from pZY20A was ligated into the *Eco*RV and *Not*I sites of pGEM5Zf(+). Clone pTEM16A and pTEM16B are subclones of pZY16A and pZY16B in which the fragments of *Sal*I to *Bam*HI were deleted. pTEM20 is a subclone of pZY20A from which the *Kpn*I fragment was deleted. pTEM23 is a subclone of pZY23A from which the *Xba*I fragment was deleted. Preparation of plasmid templates, RNA transcript probes, complementary control RNA target, and RPAs have been detailed previously (21,22). RPA products were electrophoresed through 6% polyacrylamide gels with 7 M urea and were dried at 80°C under vacuum prior to autoradiography. Experimental targets were SY568 dsRNA which had been denatured by boiling in water for 3 min then placed immediately on ice. The dsRNA used as target was extracted from a single harvest of SY568 tissue from the parental sweet orange plant which was used for the production of the cDNA clones.

Northern Hybridization

Leaf tissues of sweet orange infected with CTV-T36 and CTV-VT were obtained from Stephen M. Garnsey (USDA-ARS Horticultural Research Laboratory, 2120 Camden Road, Orlando, Florida). Total RNA was purified as described (23). Eight μg of RNA from each sample (T36, SY568 and VT) was denatured and separated as described (24). After electrophoresis, the gel was washed and the RNA denatured by soaking for 30 min in 1% glycine buffer, 20 min in 50 mM NaOH solution and finally equilibrated for 40 min in 20 X SSC (25). The RNA was transferred to a positively charged nylon membrane (Hybond-N+, Amersham International, Little Chalfont, Buckinghamshire, UK) by blotting for 5 h in a downward direction (26). After transfer, the RNA was fixed by incubating the membrane for 5 min in a 50 mM NaOH solution on a rotary shaker, briefly washed in 2 X SSC, air-dried and stored at room temperature until further use. Hybridization, washing and autoradiography were done as described (27). The probe was a riboprobe corresponding to the 3′UTR.

Results and Discussion

Cloning Strategy

The complete genome of CTV-SY568 was cloned into 13 overlapping fragments (Fig. 1). pTJ555 was the original clone used in this work. It was obtained from a PCR product using random primers with gel purified dsRNA of CTV-SY568. The second clone obtained was pTEM24. It was obtained from a PCR product using CTV-SY568 total RNA isolated from infected grapefruit with primers R214 and F215. These two primers were designed based on the sequence of pTJ555 and the published sequence of T36 (28). After sequencing clone pTEM24, we began to clone the other regions of the genome by a ''walking procedure'' in both directions. A primer complementary to the known sequence of SY568 and a T36 specific primer upstream or downstream were used to generate clones corresponding to the genomic regions of SY568. To ''walk'' to the 3′ terminus of the genome, primers CTVF4 and CTV3′, based on the sequences of pTEM24 and the 3′ end of the CTV-T36 genome (14), were designed. The corresponding clone obtained by RT-PCR was pZY19. By walking toward the 5′ terminus of the genome, we obtained clones pZY23, pZY18, pZY15 and pZY16. Clone pZY20 was obtained with primers based on sequences from pZY23 and pZY19. However, no PCR products were obtained using this method to obtain clones upstream of pZY16. Based on conserved nucleotide sequences of CTV-T36, beet yellows virus (BYV) and lettuce infectious yellows virus (LIYV) (11,29,30), we designed a degenerate primer F7920 (Table 1). Using primer F7920 and CTV-SY568 specific primer R9430, clone pZY80 was obtained. Clones upstream of pZY80 were obtained using degenerate primers based on sequences from the 5′ terminus of CTV-VT (31) and T36. Degenerate primers F20 and R1550 (Table 1) were used to obtain clone pZY5. With the CTV-SY568 specific primer F1490 and primer R3040 (Table 1) specific to conserved regions between T36 and VT, we obtained clone pZY30.

Using the same strategy, clone pZY40 was obtained. To clone the region between clones pZY40 and pZY80, we used primers based on sequences from pZY40 and pZY80 to clone this 3.9 kb region with the LA PCR kit.

The exact 5′ and 3′ terminal sequences were confirmed through clones pZYN and pZYC. Purified dsRNA of SY568 was denatured and polyA-tailed. This polyA-tailed RNA was reverse-transcribed and used as a PCR template. The expected 0.53 kb product could be amplified successfully with primer pair F716 and dT16 from the 3′ end of the genome. As CTV has a nested set of 3′-coterminal subgenomic RNAs (32), there are more 3′ end fragments of the genome than 5′ end fragments of the genome. A PCR product of the expected size from the 5′ end of the genome could not be detected using primers dT16 and 5R1. But using nested PCR, and this primary PCR product as template, we obtained a 0.7 kb product with primers dT16 and 5R2.

The complete genome of CTV-SY568 is 19,249 nucleotides (nt), which is larger than that of CTV-VT (19,226 nt) but smaller than that of CTV-T36 (19,296 nt). It encompasses 11 ORFs (Fig. 1) and the genome organization is identical to that of T36 and VT. ORF1 comprises the 5′ half of the genome and encodes a polyprotein from ORF1a to ORF1b through a presumed +1 ribosomal frameshift, while the 3′ half of the genome encompasses 10 ORFs (Fig. 1). Amino acid sequence alignments of the predicted polyproteins of CTV-SY568 and VT (12) allowed us to locate the putative catalytic residues of two leader papain-like protease domains (Cys404, Cys889, His465 and His949) and predicted cleavage sites (Gly485–Gly486 and Gly969–Gly970).

Sequence and Subgenomic Comparisons among CTV-SY568, CTV-VT and CTV-T36

At the nucleotide level, CTV-SY568 has 81% average identity with T36 over the entire genome (Fig. 2). The nucleotide identity between CTV-SY568 and T36 in ORF1a is below the average and identities from ORF3 to the 3′ end of the genome are above the average. The transition occurs in ORF1b and ORF2. The 5′UTR and 3′UTR of CTV-SY568 were 56.0% and 96.3% identical to that of CTV-T36 respectively. The predicted polyprotein of ORF1a has an overall 70.8% amino acid sequence identity. ORFs 3–11 in the 3′ half have more than 90% overall amino acid sequence identities (Table 2). The polyprotein ORF contains domains characteristic of two papain-like proteases (Pro1, Pro2), a methyltransferase (MT), and a helicase (HEL). These domains are more conserved than sequences in other regions of the polyprotein. The sequence between Pro2 and MT is also fairly well conserved (82.54%). Amino acid sequence alignments revealed many gaps between CTV-SY568 and T36 coding regions. There is one gap of 6 amino acids in the RdRp and one gap of one amino acid in the p33 protein, and there are 10 gaps in the polyprotein of ORF1a. Five of these gaps are 3–6 amino acids in lengths and the rest of them are only one amino acid in length. In the polyprotein, the regions of Pro1, Pro2, MT and HEL are more conserved than other regions. There are two gaps in these conserved regions with one amino acid each. There are no gaps in the amino acid alignments of the 3′ terminal ORFs. These comparisons show asymmetrical genomic conservation with the 3′ half of the genome being more conserved than the 5′ half of the genome.

At the nucleotide level, CTV-SY568 has 92% average identity with VT over the entire genome (Fig. 2). Values for most regions are around the average. The region from the 3′ end of ORF1a to ORF5 is slightly lower than the average. The 5′UTR and 3′UTR of CTV-SY568 were 94.9% and 97.4% identical to that of CTV-VT respectively. The nucleotide sequence identities of ORF4-6 are approximately 90% (87.4%, 87.9% and 89.2% respectively; Table 2). The nucleotide sequence identities in all other regions are higher than 90%. However, all ORFs except p33 have more than 90% amino acid sequence identities. In p33, both amino acid and nucleotide sequence identities were about 84–86% between all three isolates (Table 2; 12). The asymmetrical genomic conservation could still be observed when comparing gaps in the amino acid sequences between CTV-SY568 and CTV-VT. In the 3′ half of the genome, there is only one gap of one amino acid in length located in the p61 protein. However, there are 8 gaps of one amino acid each in the polyprotein, although these gaps are not located in the Prol, Pro2, MT or HEL regions of the polyprotein.

We have also compared the genomic and subgenomic RNA patterns for these 3 isolates by

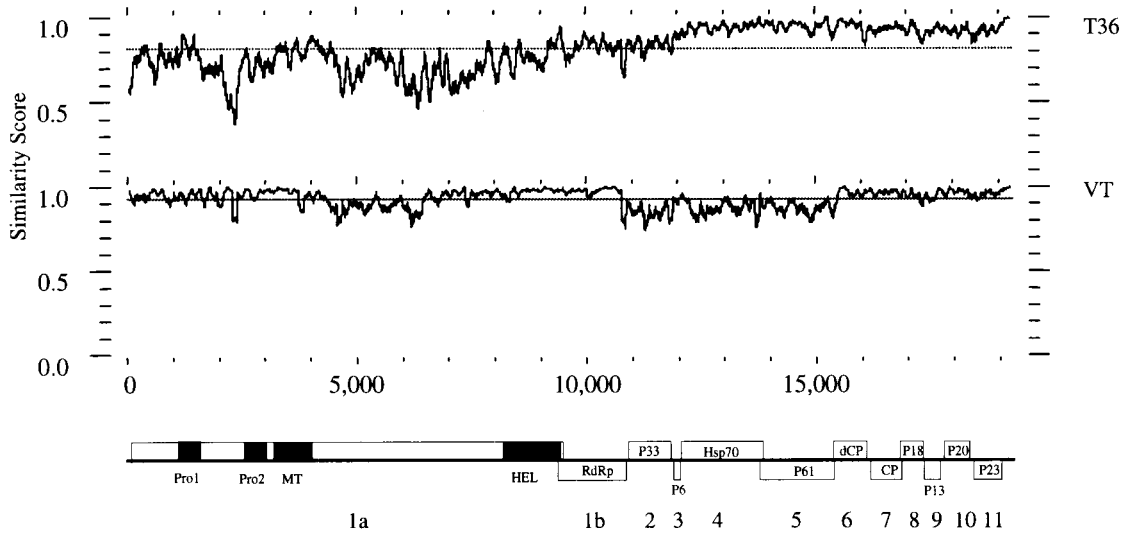

Fig. 2. Nucleotide sequence identity comparisons of CTV isolate SY568 with either T36 or VT. The similarity scores were averaged over a 100 nt window with the GCG program PLOTSIMILARITY (20). The dotted lines represent the average identities. The bottom half of the figure is the schematic representation of the genome organization. ORFs are as indicated in Fig. 1.

Northern hybridization using a probe corresponding to the 3′ UTR of SY568. Fig. 3 shows that these patterns are very similar among these isolates. The subgenomic RNAs for the p23 and the p20 are the most prominent and of similar intensities. However, some major differences could be observed. The subgenomic RNA that corresponds to the coat protein ORF (CP) of CTV-SY568 is stronger than that of CTV-T36 and VT. For CTV-T36 and VT, the genomic band is very weak, as compared to CTV-

Table 2. Sequence comparisons of the CTV isolate SY568 with the corresponding regions of isolates VT and T36. The sources of sequence data were GenBank accession numbers AF001623 (CTV-SY568), U56902 (CTV-VT) and U16304 (CTV-T36). Sequence identities were calculated by making pairwise comparisons using the GAP program (20)

				T36[b]		VT[b]	
Region	Product	Start[a]	Stop[a]	Nucleotide (%)	Amino Acid (%)	Nucleotide (%)	Amino Acid (%)
5′UTR	–	1	107	56.0		94.9	
ORF 1a	pol	108	9455	72.8	70.8	94.5	92.5
ORF 1b	RdRp	9328	10812	79.7	89.9	90.4	94.8
ORF 2	p33	10861	11772	84.5	86.5	84.9	85.8
ORF 3	p6	11842	11997	88.5	92.3	90.9	92.2
ORF 4	p65	12003	13787	93.8	95.6	87.4	91.2
ORF 5	p61	13711	15318	95.0	95.9	87.9	90.5
ORF 6	p27	15293	16015	93.2	95.4	89.2	93.3
ORF 7	p25 (cp)	16107	16778	93.5	95.5	96.6	96.9
ORF 8	p18	16744	17247	93.0	94.0	95.2	94.0
ORF 9	p13	17281	17640	90.8	90.8	93.3	94.1
ORF 10	p20	17716	18264	92.0	97.3	96.7	98.4
ORF 11	p23	18347	18976	91.3	90.0	95.2	95.7
3′UTR	–	18977	19249	96.3		97.4	

[a]Numbering begins at the 5′ end of the genome of CTV-SY568.
[b]Percent identity.

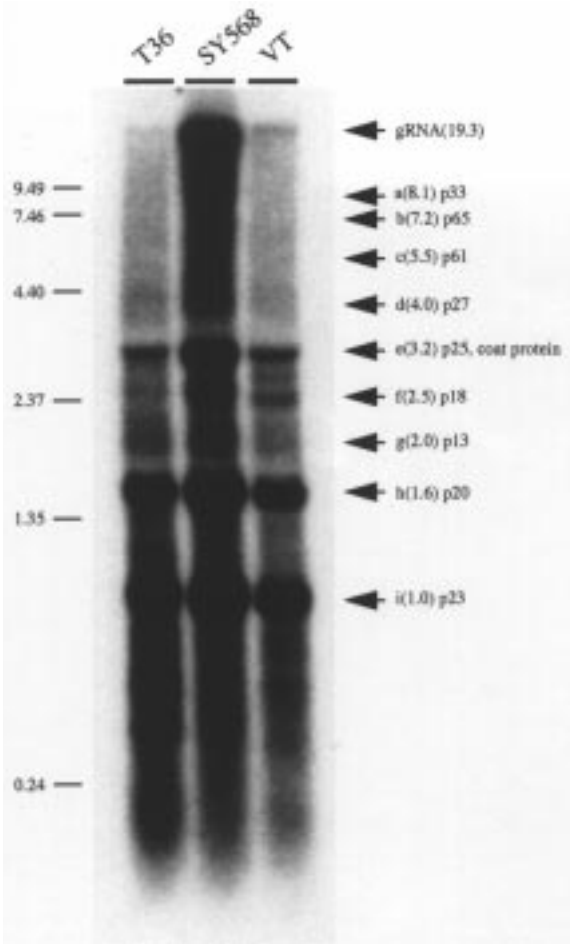

Fig. 3. Comparison of genomic and subgenomic RNAs by Northern hybridization using a probe corresponding to the 3′UTR of SY568 (pZY19; Fig. 1). Each lane was loaded with 8 μg of total RNA isolated from sweet orange infected leaves with either CTV-T36 (T36), CTV-SY568 (SY568) or CTV-VT (VT). Molecular weight markers in kilobases are indicated to the left of the figure. The genomic RNA and subgenomic RNAs are indicated with arrows to the right of the figure. The designation of subgenomic RNAs are as shown in Hilf et al. (32). The sizes in kb of the genomic and subgenomic RNAs are indicated in parentheses.

SY568. These results may indicate that CTV-SY568 could replicate and accumulate more viral RNAs than the other two isolates in sweet orange. CTV-SY568 is a very severe isolate and causes strong reactions in sweet orange which is the least likely host to give a strong reaction to a CTV isolate. This high degree of virus accumulation may be partly responsible for the severity of symptoms seen in sweet orange.

Sequence Variation within CTV-SY568

Sequence variations between independent clones and regions of overlap between clones were analyzed with the ''GAP'' program. There were a total of 198 nucleotide variations in the 19,249 nt sequence of CTV-SY568 based on our sequence data of 2 or 3 independent clones from each of the 13 overlapping cloned regions that span the genome. These nucleotide sequence variations are mainly distributed in three regions. In ORF1a which encodes a putative polyprotein, there are 31 nucleotide differences between positions 487 and 9375 in the genome, with 14 amino acid changes. The second region of sequence variation is the ORF1b which encodes the RdRp gene. There are 118 nucleotide sequence variations in this region resulting in only 16 amino acid changes. The ratio of silent mutations to amino acid altering mutations is 6.5:1. The third region is inside ORF6 (dCP), which encodes a diverged copy of coat protein gene. There are 48 nucleotide variations with only 5 leading to an amino acid change. The ratio of silent mutations and amino acid altering mutations is 9.6:1. There is one nucleotide difference in the region between ORF1b and ORF2. Of the 198 nucleotide differences 35 lead to amino acid changes with 24 of these being nonconservative amino acid substitutions (Table 3).

The complete genomic sequences of BYV, LIYV, two isolates of CTV, and the 3′-terminal half of beet yellow stunt virus genome have been determined (11,12,29,30,33). However, very little sequence variation was detected in these studies. In our work, 198 nt sequence variations were identified. The variations were distributed in the regions of the polyprotein, RdRp and p27. The diversity of the RdRp clones (pZY16A and pZY16B) and that found in the overlapping region of clones pZY20 and pZY23 (dCP), suggests that the virus population of CTV-SY568 is a mixture of different genomes. Since our sequence data are based on PCR products and the sequence variations resulting from polymerase errors are captured, in addition to actual sequence variants, it was necessary to determine if our sequence data were representative of the major RNA populations and

Table 3. Nucleotide sequence variation between independent SY568 cDNA clones leading to amino acid changes

Region	Position[a]	Nucleotide Differences	Amino Acid Changes	Similarity[b]
ORF1a	487	T-C	Ile-Thr	−
(Polyprotein)	909	T-C	Ser-Pro	−
	1963	G-A	Arg-Gln	+
	2449	T-G	Leu-Arg	−
	3429	G-T	Asp-Tyr	−
	4300	C-T	Ala-Val	−
	4315	C-A	Thr-Asn	−
	6456	T-C	Phe-Leu	−
	6537	A-G	Lys-Glu	+
	6747	G-A	Gly-Ser	−
	7056	A-G	Ile-Val	+
	7293	A-G	Arg-Gly	−
	9318	A-G	Ile-Val	+
	9324	G-A	Val-Ile	+
ORF1b (RdRp)	9495	A-T	Asn-Cys	−
	9496	A-G		
	9541	G-A	Arg-Lys	+
	9639	T-G	Ser-Ala	+
	9670	A-C	Asp-Ala	−
	9750	A-C	Lys-Gln	+
	9933	A-G	Asn-Asp	+
	9935	T-C		
	9990	G-A	Val-Ile	+
	10018	C-T	Ser-Phe	−
	10480	T-A	Ile-Asn	−
	10537	C-T	Thr-Ile	−
	10755	C-A	Pro-Thr	−
	10761	G-A	Ala-Thr	−
	10773	G-A	Ala-Thr	−
	10790	G-A	Ala-Met	−
	10791	C-T		
	10802	G-A	Ala-Met	−
	10803	C-T		
	10809	A-G	Asp-Gly	−
p27	15493	C-A	His-Gln	−
	15521	G-A	Ala-Thr	−
	15603	C-T	Thr-Ile	−
	15666	C-T	Ser-Leu	−
	15878	T-A	Ser-Thr	+

[a]Nucleotide position starting with numbering from the 5′ end of the genome as in Table 2.
[b]The similarity is obtained by comparing peptide sequences using the GCG ''GAP'' program; '' + '' the two amino acids are similar and '' − '' non-conservative change.

which sequence variants represented the consensus sequence in the variable regions.

RNase Protection Assays and dsRNA Analysis

The dsRNA profiles of 6 different isolates of CTV all had a major segment representing the genomic replicative form of the virus (Fig. 4, Rf band). Each isolate had a different electrophoretic pattern with varying complexity. A band was present in all 4 SY (seedling yellows) isolates that was not present in the two T (typical mild) isolates (Fig. 4, see 0.5 band). This band has been described before and can be correlated to those isolates which cause seedling yellows and/or stem pitting (32,34,35). Based on the

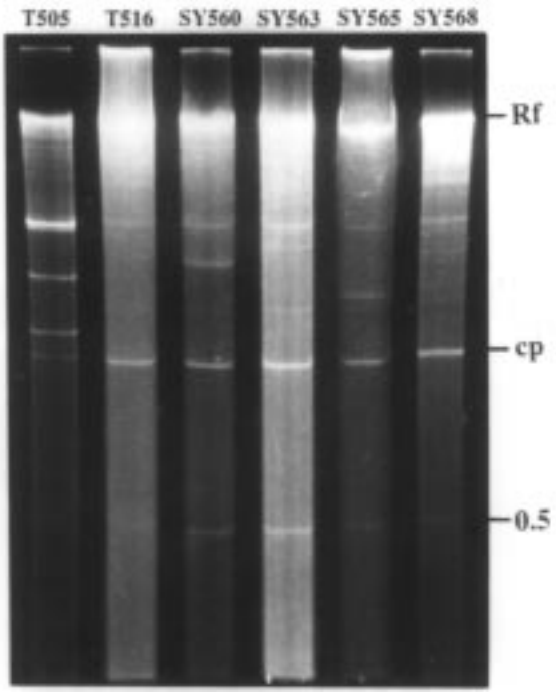

Fig. 4. CTV dsRNAs electrophoresed through 6% polyacrylamide gel and stained with ethidium bromide. Isolate designations are above each lane. Positions of the genomic RNA (Rf), coat protein subgenomic RNA (CP) and seedling yellows and/or stem pitting indicator (0.5) are shown to the right.

dsRNA patterns observed in this study, isolate SY568 did not contain obvious defective RNAs as are commonly found in many CTV isolates (36). Further, since SY568 had a relatively simple dsRNA pattern (Fig. 4, Lane SY568), it could be argued that it is not likely to be a mixed infection. Despite this conclusion from dsRNA analysis, evidence for a mixed infection was obtained by RPA (see below).

In order to confirm that our sequence data were representative of the major CTV genomic RNA population, RPAs were employed. A total of 11 probes covering 9 regions of the genome were used (Fig. 1). Minus-sense RNA probes from the following clones were able to protect primarily full length molecules of denatured dsRNA from RNase digestion: 5′UTR, Prol, RdRpB, dCPB, CP, P13, P20 and 3′UTR (Fig. 5). Minus-sense RNA probes that protected virtually no full length molecules, with the majority of the dsRNA cleaved into smaller fragments were RdRpA and dCPA (Fig. 5). The HSP70 probe fully protected approximately 25% of the total dsRNA while the rest was cleaved into relatively small fragments. All plus-sense complement control targets fully protected each respective probe and unprotected probes (no target) were completely digested by RNase treatment (data not shown). Minor fragmentation patterns of probes Prol, RdRpB, dCPB, P13, P20, and 3′UTR when protected by CTV-SY568 dsRNA were similar to those produced by protection with the plus-sense transcript controls (data not shown).

These results suggest that most of our clones represent the major component of the SY568 RNA population. In the variable regions, the probe from pZY16B protected the major population fully, while the probe from pZY16A fully protected a minor population with most of the dsRNA being cleaved into approximately 13 smaller fragments. This indicates that the sequence from pZY16B is the consensus sequence of SY568 in the RdRp region. A similar result was obtained for probes dCPB and dCPA respectively.

All RPA analyses were repeated at least twice and patterns obtained were identical for each experiment. While theoretically RPAs can detect single base mismatches, in practice it probably does not. Using the same protocol employed here, Kurath et al. (21) found that when isolates of satellite tobacco mosaic virus were sequenced and compared to the RPA predictions of mismatches, only 50% were detected by RPA. In this study, however, our concern was not with whether we could pinpoint a specific mismatch, but only to look at a probe's ability to protect the dominant population of viral dsRNA in the plant. We have used these data to aid in the assembly of a full length clone of SY568, preferentially using those clones that fully protect the major RNA population.

Sweet orange industries throughout the world are usually able to tolerate infection by CTV unless the isolates are of the SY568 type. Concerns about stem pitting, stunting and associated small fruit size in sweet orange increases once an efficient vector, the brown citrus aphid (*Toxoptera citricidus*), becomes established in new areas such as Central America, Florida (now present) and eventually Mexico, western USA including California and perhaps the Mediterranean. An expansion of the distribution of the brown citrus aphid is going on at the present time. The molecular characterization of SY568 will assist in the attempt to create infectious clones with

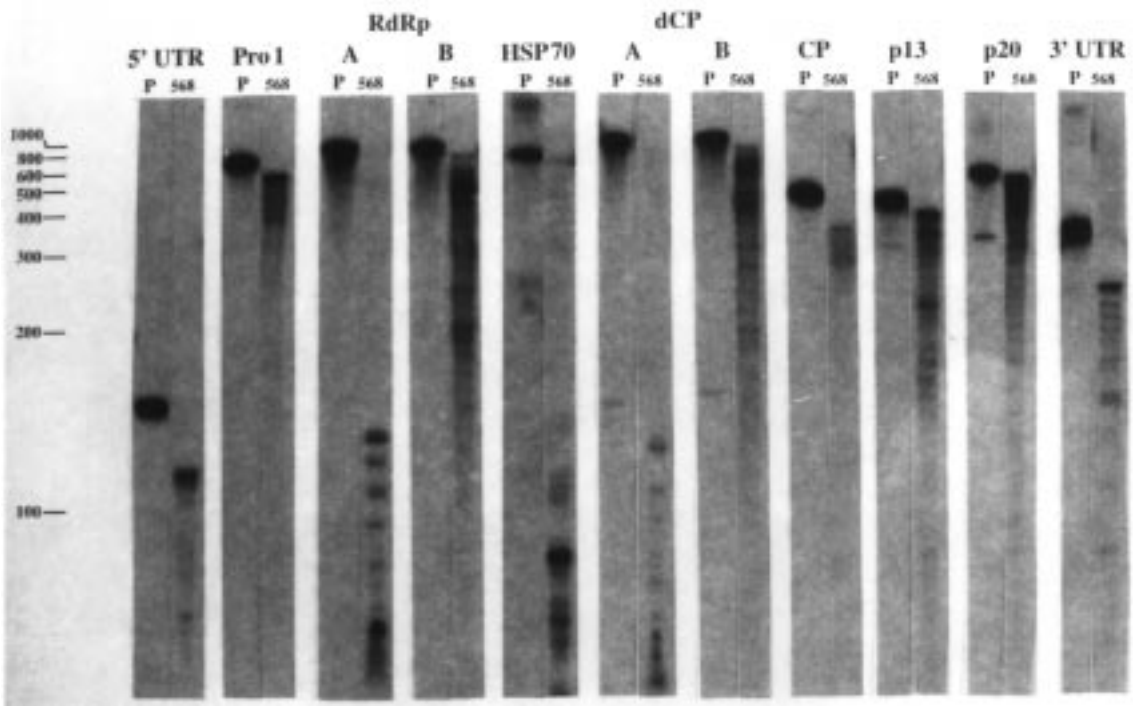

Fig. 5. RNase protection assays using 11 probes covering 9 regions of the SY568 genome (Fig. 1). The target RNA in each assay was CTV-SY568 dsRNA that was extracted from the parental sweet orange plant which was used for the production of the cDNA clones (568). Molecular weight markers in bases are indicated to the left of the figure. Note that all probes contain from 40 to 90 non-viral nucleotides from the vectors which accounts for the shifts from probe alone, not treated with RNase (P) to ''probe + SY568 target'' (568) in each case.

different symptom inducing potentials. By ensuring that the clone is representative of the RNA population of SY568 the probability of reproducing the severe biology in the cDNA clone should be increased. The identification of alternate sequences for variable regions of the genome, which have been shown to be present as minor variants in the RNA population, proves that SY568 is actually a mixture of strains. Biological subcultures of SY568 that show stable symptom diversity are available for further study (Dodds, unpublished results) and the availability of alternate sequences of variable regions will be valuable in efforts to determine the role of different strains in the overall biology of this severe isolate.

Acknowledgments

The authors thank Dr. Ivan L. Ingelbrecht for critical review and useful comments, Drs. Francis Moonan and William O. Dawson for their helpful discussions, and Mercedes O. Campos for excellent technical help. We thank Dr. Stephen M. Garnsey for providing CTV-T36 and VT infected tissues of sweet orange. This work was supported by grants from the Texas Citrus Producers Board, the California Citrus Research Board, and the USDA-California CTV Coalition. We acknowledge the contribution of Thanakorn Jarupat, who obtained the first cDNA clone of SY568 (pTJ555).

Note

The nucleotide sequence data reported in this paper has been deposited in Genbank under Accession No. AF001623

References

1. Bar-Joseph M., Marcus R., and Lee R.F., Ann Rev Phytopath *27*, 291–316, 1989.

2. Garnsey S.M., Gumpf D.J., Roistacher C.N., Civerolo E.L., Lee R.F., Yokomi R.K., and Bar-Joseph M., Phytophylactica *19*, 151–157, 1987.
3. Calavan E.C., Harjung M.K., Blue R.L., Roistacher C.N., Gumpf D.J., and Moore P.W., *Proc 8th Conf IOCV*. Riverside, USA. 69–75, 1980.
4. Costa A.S. and Muller G.W., Plant Dis *64*, 538–541, 1980.
5. Roistacher C.N., Dodds J.A., and Bash J.A., *Proc 10th Conf IOCV*. Valencia, Spain. 91–100, 1988.
6. Roistacher C.N. and Moreno P., *Proc 11th Conf IOCV*. Riverside, USA. 7–19, 1991.
7. Jarupat T., Dodds J.A., and Roistacher C.N., *Proc 10th Conf IOCV*. Riverside USA, 39–45, 1988.
8. Moreno P., Guerri J., Ballesterolmos J.F., Albiach R., and Martinez M.E., Plant Path *42*, 35–41, 1993.
9. Broadbent P., Brlansky R.H., and Indsto J., Plant Dis *80*, 329–333, 1996.
10. Bar-Joseph M., Gumpf D.F., Dodds J.A., Rosner A., and Ginzburg I., Phytopathology *75*, 195–198, 1985.
11. Karasev A.V., Boyko V.P., Gowda S., Nikolaeva O.V., Hilf M.E., Koonin E.V., Niblett C.L., Cline K., Gumpf D.J., Lee R.F., Garnsey S.M., Lewandowski D.J., and Dawson W.O., Virology *208*, 511–520, 1995.
12. Mawassi M., Mietkiewska E., Gofman R., Yang G., and Bar-Joseph M., J Gen Virol *77*, 2359–2364, 1996.
13. Karasev A.V., Nikolaeva O.V., Koonin E.V., Gumpf D.J., and Garnsey S.M., J Gen Virol *75*, 1415–1422, 1994.
14. Pappu H.R., Karasev A.V., Anderson E.J., Pappu S.S., Hilf M.E., Febres V., Eckloff R.M.G., McCaffery M., Boyko V., Gowda S., Dolja V.V., and Koonin E.V., Virology *199*, 35–46, 1994.
15. Febres V.J., Ashoulin L., Mawassi M., Frank A., Bar-Joseph M., Manjunath K.L., Lee R.F., and Niblett C.L., Phytopathology *86*, 1331–1335, 1996.
16. Roistacher C.N., Bar-Joseph M., and Gumpf D.J., Plant Dis *68*, 494–496, 1984.
17. Morris T.J. and Dodds J.A., Phytopathology *69*, 854–858, 1979.
18. Yang Z.N. and Mirkov T.E., Phytopathology *87*, 932–939, 1997.
19. Sambrook J., Fritsch E.F., and Maniatis T., *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, p6.1–6.21, 1989.
20. Devereux J., Haeberli P., and Smithies O., Nucleic Acids Res *12*, 387–395, 1984.
21. Kurath G., Rey M.E.C., and Dodds J.A., Virology *189*, 233–244, 1992.
22. Winter E., Yamamoto F., Almoguera C., and Perucho M., Proc Natl Acad Sci USA *82*, 7575–7579, 1985.
23. Jones J.D.G., Dunsmuir P., and Bedbrook J., EMBO J *4*, 2411–2418, 1985.
24. Tsang S.S., Yin X., Guzzo-Arkuran C., Jones V.S., and Davison A.J., BioTechniques *14*, 380–381, 1993.
25. Beckers T., Schmidt P., and Hilgard P., BioTechniques *16*, 1074–1078, 1994.
26. Koetsier P.A., Schorr J., and Doerfler W., BioTechniques *15*, 260–262, 1993.
27. Ingelbrecht I.L., Mandelbaum C.I., and Mirkov T.E., BioTechniques *25*, 420–423, 1998.
28. Sekiya M.E., Lawrence S.D., McCaffery M., and Cline K., J Gen Virol *72*, 1013–1020, 1991.
29. Agranovsky A.A., Koonin E.V., Boyko V.P., Maiss E., Frotschl R., Lunina N.A., and Atabekov J.G., Virology *198*, 311–324, 1994.
30. Klaassen V.A., Boeshore M.L., Koonin E.V., Tian T., and Falk B.W., Virology *208*, 99–110, 1995.
31. Mawassi M., Mietkiewska E., Hilf M.E., Ashoulin L., Karasev A.V., Gafny R., Lee R.F., Garnsey S.M. Dawson W.O., and Bar-Joseph M., Virology *214*, 264–268, 1995.
32. Hilf M.E., Karasev A.V., Pappu H.R., Gumpf D.J., Niblett C.L., and Garnsey S.M., Virology *208*, 576–582, 1995.
33. Karasev A.V., Kikolaeva O.V., Mushegian A.R., Lee R.F., and Dawson W.O., Virology *221*, 199–207, 1996.
34. Dodds J.A., Jordan R.L., Roistacher C.N., and Jarupat, T., Intervirology *27*, 177–188, 1987.
35. Dodds J.A., Jarupat T., Roistacher C.N., and Lee J.G., Phytophylactica *19*, 131–137, 1987.
36. Mawassi M., Karasev A.V., Mietjiewska E., Gafny R., Lee R.F., Dawson W.O., and Bar-Joseph M., Virology *208*, 383–387, 1995.