

Genotypic diversity of *Citrus tristeza virus* within red grapefruit, in a field trial site in South Africa.

D.A Read¹ and G. Pietersen^{1, 2}

¹Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa

²Agricultural Research Council—Plant Protection Research Institute, Pretoria 0002, South Africa

ABSTRACT

Grapefruit cultivars (*Citrus paradisi* Macfad.) are extremely sensitive to *Citrus tristeza virus* (CTV) infections and are pre-immunized with mild-strain cross-protecting sources not containing components that elicit symptoms such as stem-pitting and decline, to ensure longer periods of productivity. However, pre-immunizing sources often lose their efficiency and for this reason the previously commercially applied grapefruit cross-protecting source GFMS (grapefruit mild-strain) 12 has been replaced by GFMS 35. This study was undertaken to determine the diversity of CTV genotypes within trees that were inoculated with either GFMS 12 or GFMS 35. Samples were collected from a number of different trees of two red grapefruit cultivars (cv. Star Ruby and cv. Flame), planted 10 years prior to sampling in the Malelane production area of South Africa. Reverse transcription-polymerase chain reaction amplification of a 5' variable region (A-region) and a 3' conserved region (p23 gene) was followed by cloning, sequencing of multiple clones and phylogenetic analyses. The genotypic identities of clones were determined based on their relatedness to reference CTV strains. Sequence types within the VT genotypic group dominated in all of the samples, with T30-like sequence types being a minor component in some populations of the field collected samples. The original pre-immunising populations of GFMS 12 and GFMS 35 were characterised on greenhouse maintained plants and compared with the populations exposed to field infections by aphids. While the methodology employed only allows a coarse representation of the genotype composition of the CTV population, this study provides insight into which genotypes of CTV must be incorporated within a mild-strain cross-protecting source within the South African Citrus Improvement Scheme (SACIS).

KEYWORDS Cross-protection; cloning; Sanger sequencing; PCR bias; A-region; p23 gene

INTRODUCTION

Citrus tristeza virus (CTV) is the most severe viral pathogen of citrus and is responsible for one of the most economically important diseases of citrus (Bar-Joseph et al. 1989). Depending on the scion-rootstock combination and CTV population composition, symptoms can vary from being asymptomatic, to decline, stem-pitting, stunting (Niblett et al. 2000), quick decline (Moreno et al. 2008) and the nursery related syndrome, seedling yellows (Černi et al. 2008). The virus is spread semi-persistently by a number of aphid species. *Toxoptera citricida*, *Aphis gossypii*, and *A. spiraecola* are the most important vectors, with *T. citricida* being the most efficient (Moreno et al. 2008). Cross-protection is sometimes used in countries where severe stem-pitting CTV strains are endemic, such as South Africa and Brazil (Souza et al. 2002). In South Africa, CTV related stem-pitting was observed as early as the 1940's (Oberholzer et al. 1949). Cross-protection is the partial or complete resistance of a plant to infection of a severe strain of a virus after the intentional inoculation of a mild strain of the same virus (Gal-On and Shibolet 2005; Yoon et al. 2006).

The South African Citrus Improvement Scheme (SACIS) was established in 1972 but the implementation of cross-protection began when shoot-tip-grafting (STG) was employed to remove all graft-transmissible pathogens from budwood sources (von Broembsen and Lee 1988). Mild-strain populations were collected from interim (prior to STG) grapefruit (*Citrus paradisi* Macfad.) budwood sources that were older than 15 years and still producing large quantities of high quality fruit (van Vuuren and Collins 1993). These isolates were initially evaluated in greenhouse trials (van Vuuren and Moll 1987) and then field trials (van Vuuren and Collins 1993) where empirical evidence suggested that three CTV selections could be considered mild, namely GFMS 12, GFMS 27 and GFMS 35. The GFMS 12 source was used to pre-immunise all grapefruit cultivars within the SACIS, however later trials suggested that GFMS 12 and GFMS 35 performed poorly under field conditions (van Vuuren and van der Vyver 2000). Stem-pitting and reduced fruit size was also observed on GFMS 12 pre-immunised budwood mother trees at the Citrus Foundation Block at Uitenhage as early as 1993 (van Vuuren and Manicom, 2005). Single aphid transfers (SATs) were used to produce sub-isolates from GFMS 12, which alluded to the heterogeneity of strains within the original GFMS 12 population (van Vuuren et al. 2000). The GFMS 35 source was approved as the replacing cross-protecting source

for red grapefruit cultivars (Luttig *et al.* 2002) and has been used to pre-immunise red grapefruit budwood mother trees since 1998 (van Vuuren, personal communication). Various molecular techniques have been used to characterise cross-protecting sources and sub-isolates used within the SACIS, which include single stranded conformation polymorphism (SSCP) (Luttig *et al.* 2002) and restriction fragment length polymorphism (RFLP) (van Vuuren *et al.* 2000). Scott *et al.* (2012) used gene amplification and mass cloning to determine the genotype composition of the original GFMS 12 source maintained under glasshouse conditions. Very little data is available regarding the genotype compositions of CTV populations within trees under field conditions.

The unintentional introduction of “exotic” CTV strains to a production area places further pressure onto existing cross-protection schemes, since these strains have the potential to cause increased losses (Folimonova 2013). For this reason, it is essential to determine which CTV strains are present within a production area to allow for the targeted isolation of cross-protecting sources.

In the current study, the CTV population diversity of two different red grapefruit cultivars (cv. Flame and cv. Star Ruby) was determined through the reverse-transcription polymerase chain reaction (RT-PCR) and mass cloning of the A-region (located within the 5' variable region) (Rubio *et al.* 2001) and the p23 gene (located within the 3' conserved region) (Sambade *et al.* 2003) followed by the sequencing of multiple clones. These results were expected to provide insight into the diversity of CTV within these field grown pre-immunised South African grapefruit trees and will allow for a more targeted selection of populations for use within the mild-strain cross-protection program.

MATERIALS AND METHODS

Sample collection

Samples from trees pre-immunized with either GFMS 12 or GFMS 35, were collected in February, 2009 at the Agricultural Research Council (ARC) Experimental Station near Malelane in the Mpumalanga Province from a former trial of van Vuuren and Manicom (2005) with Star Ruby and Flame grapefruit cultivars. Trees had been randomly planted in 1998 amongst other red grapefruit cultivars that were pre-immunized with various CTV sources under test, including GFMS 12 and GFMS 35, in a block with 5 replicates of each scion/pre-immunizing source combination.

Samples were collected from three out of the five replicates, except for the combination of GFMS 12 on Star Ruby, where four trees were sampled due to the severe decline and stem-pitting symptoms associated with this combination. Cultivar and pre-immunizing source combinations and symptom descriptions are listed in *Table 1* together with the respective accession numbers. Individual samples will be referred hereafter by this number. The trees were approximately 10 years old at the time of sampling. Leaf material and green woody tissue were sampled at approximately equidistant points around the tree with 3-5 individual collection points per tree. Greenhouse maintained *Citrus aurantifolia* (Christm) Swing cv. Mexican lime and grapefruit cultivar, Star Ruby, previously inoculated with GFMS 12 or GFMS 35, were also sampled.

Table 1: List of samples collected from the ARC Experimental Station near Malelane.
SP = Stem-pitting; D = Decline.

Cross-protecting source	Cultivar	Sample accession numbers
GFMS 12	Star Ruby	301 (Severe SP + D); 302 (Mild SP); 303 (Mild SP); 313 (Severe SP + severe D)
GFMS 12	Flame	304 (Mild SP + D); 305 (Mild SP); 306 (Mild SP + severe D)
GFMS 35	Star Ruby	307 (Mild SP); 308 (Severe SP + D); 309 (Mild SP)
GFMS 35	Flame	310 (Mild SP); 311 (Severe SP + mild D); 312 (Severe SP + D)

Viral gene amplification.

Leaf midrib and green bark tissue was macerated using liquid nitrogen in a mortar and pestle. Total RNA extraction was conducted on all samples using a Promega® SV Total RNA extraction kit (Promega, Madison, WI, USA) according to the manufacturer's instructions.

Amplification of the viral templates was carried out using a two-step RT-PCR protocol. cDNA synthesis of the A-region and p23 gene region was initiated using A-R (5' GTCGATAACTCGACAAACGAGC 3') (Rubio et al. 2001) or PM51 (5' AACTTATTCCGTCCACTTC 3') (Sambade et al. 2003) primers respectively. The cDNA reaction consisted of 12µl of total RNA extracts, 50pmol of reverse primer,

10U Avianmyeloblastosis virus (AMV) reverse transcriptase (Roche, Mannheim, Germany), 5U RNase inhibitor (Roche, Mannheim, Germany), 1x RT buffer (Roche, Mannheim, Germany) and 0.1 mM dNTP mix (Promega, Madison, WI, USA). Reactions were held at 42°C for 1 hour.

PCR was conducted using 10µl of cDNA as template, 2µM A-F (5' ACGTGTTCTGTGAAACGCGG 3') (Rubio et al., 2001) and A-R primers to yield the "A-region" amplification, or PM50 (5' ACTAACTTTAATTCGAACA 3') (Sambade et al. 2003) and PM51 primers for the p23 gene region amplification, 5µl of 1x PCR reaction buffer (Bioline, London, United Kingdom), 2.5mM MgCl₂, 0.14mM dNTP mix, 2.5 U standard fidelity Taq DNA polymerase (Bioline, London, United Kingdom), 2µl of 20µg/µl Bovine serum albumin (BSA) and PCR grade water to a total volume of 50µl. PCR cycling conditions were 1 cycle of 92°C for 2min, 40 cycles of 92°C for 30s, 55°C for 45s, 72°C for 1min and 1 cycle of 72°C for 10min. Amplification products were resolved using 1% agarose gels and stained with ethidium bromide.

Cloning and sequencing.

Amplified products were purified from the agarose gels using a Promega, Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA). DNA concentrations were determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Purified products were cloned into the p-GEM T Easy® vector (Promega, Madison, WI, USA), followed by the transformation of competent E. coli JM109. Putative recombinant clones were selected using blue/white selection, followed by an alkaline lysis plasmid extraction protocol (Sambrook 2001). Sixty A-region and thirty p23 gene clones were chosen per field sample. Between 40 and 120 A-region and p23 gene clones were selected per greenhouse maintained sample. Plasmid insert sizes were verified using a PCR protocol making use of the vector specific primers T7 (5' TAATACGACTCACTATAGGG 3') and SP6 (5' ATTTAGGTGACACTATAG 3') (Promega, pGEM®-T and pGEM®-T Easy Vector Systems, technical manual). Amplified products were resolved on a 1% agarose gel, stained with ethidium bromide. After confirming the band amplicon size, 19µl of the PCR product was purified using 2µl of Fast Alkaline Phosphatase and 0.5µl Exonuclease I (Thermo Scientific, Vilnius, Lithuania) and held at a constant temperature of 37°C for 15 minutes followed by 85°C for 15 minutes. The reverse vector specific primer, SP6, was used to uni-directionally sequence each of these purified PCR products.

Sequencing mixtures contained 1µl BigDye[®] Terminator mix v3.1 (Applied Biosystems, Foster City, CA, USA), 2.25µl 5x BigDye[®] v3.1 sequencing buffer, 0.75µl SP6 primer (2µM) and molecular grade water to a total volume of 10µl. Cycling conditions were: one cycle of 94°C for 1 minute followed by 30 cycles of 94°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes. After the completion of the sequencing reaction, the sequencing products were purified using an EDTA (Ethylenediaminetetraacetic acid disodium salt dihydrate) and sodium acetate precipitation. The purified sequencing products were then submitted to the African Centre for Gene Technologies (ACGT), Automated Sequencing Facility, Department of Genetics, University of Pretoria, South Africa and sequenced using an ABI Prism[®] 3130XL Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

Phylogenetic analysis

The CLC Main Workbench 5 software package (CLC Bio, Aarhus, Denmark) was used to correct errors in chromatograms. Sequences with multiple peaks in their chromatograms and ambiguous nucleotides were discarded from the study, hence cloned sequences analysed (*Table 3 and 4*) were less than clones isolated. Reverse complements of the sequences were produced, also using CLC Main Workbench 5 and alignments of sequences were carried using the CLUSTAL W alignment software (EBI, Cambridgeshire, England) within the BioEdit Sequence alignment editor 7.1.3 (Hall 1999). A total of 44 full-genome reference sequences were accessed from GenBank (www.ncbi.nlm.nih.gov/genbank/) and have the following accession numbers (strain names are in brackets): NC_001661 (T36); AY_340974 (Qaha); U16304 (T36); DQ272579 (Mexico); AY170468 (T36); EU937521 (T36); KC517485 (FS674-T36); KC517486 (FS701-T36); KC517487 (FS703-T36); KC517488 (FS577); JX266713 (Taiwan-Pum/M/T5); AF001623 (SY568); AF260651 (T30); Y18420 (T385); KC517489 (FS701-T30); KC517490 (FL278-T30); KC517491 (FS703-T30); JF957196 (B301); FJ525432 (NZRB-G90); GQ454869 (HA 18-9); FJ525435 (NZRB-M17); JX266712 (Taiwan-Pum/SP/T1); FJ525431 (NZRB-M12); FJ525433 (NZRB-TH28); FJ525434 (NZRB-TH30); JQ798289 (A18); KC525952 (T3); HM573451 (Kpg3); EU857538 (SP); GQ454870 (HA 16-5); DQ151548 (T318A); AB0463981 (NUagA); JQ911664 (CT11A); KC517493 (FL202-VT); U56902 (VT); KC517492 (FS703-VT); EU937519 (VT); KC517494 (FS701-VT); KC262793 (L192GR); JQ911663 (CT14A); FJ525436 (NZ-B18); JQ965169 (T68); EU076703 (B165); JQ061137 (AT-1). An additional three reference sequences from South

Africa were used in the alignments, namely CTZA1, CTZA2 and CTZA3 (Zablocki and Pietersen 2014).

Two different phylogenetic trees were produced and examined to determine which provided the greatest branch support for the sequence alignment of each sample and the references. Neighbour-joining phylogenetic trees were constructed for each alignment, using MEGA 4.1 (Tamura et al, 2007) and Maximum Composite Likelihood substitution model with a 1000 bootstrap replicates. Maximum likelihood phylogenetic trees were constructed by aligning group of sequences, using MAFFT (Kato and Standley 2013). jModel test 2.1.4 (Darriba et al. 2012) was used to determine the AIC (Akaike information criterion). Maximum likelihood trees were produced using the PhyML 3.0 (Guindon and Gascuel 2003) using the best fit NNI and SPR tree searching models with a bootstrap value of 1000.

Determination of primer bias in the A-region and the p23 gene

The A-region and p23 gene (PM50 and PM51) forward and reverse primer binding sites of the 44 CTV full-genome sequences available on GenBank were aligned using CLUSTAL W alignment software (EBI, Cambridgeshire, England) within the BioEdit Sequence alignment editor 7.1.3 (Hall 1999). Relative to the 44 known CTV sequences a total of six and eight possible mismatch sites were observed for the A-region forward and reverse primer binding targets respectively. No mismatches were observed for the forward or the reverse primer binding sites of the p23 gene and no genotype specific primers were designed for this conserved region (*Table 2*). Eight plasmid clones containing phylogenetically distinct inserts of the A-region were selected to represent the following CTV strains: B165, VT, HA 16-5, T3, T30, RB TH-28 (RB-clade 1), RB TH-30 (RB-clade 2). Five clones containing inserts of the p23 gene were selected to represent the following CTV strains: T36, T30, NZRB TH-28, SP and Kpg3. The A-region inserts of each of these plasmid clones were amplified using their respective genotype specific primer pairs (i.e. creating amplicons amplified using the original primer pair (Rubio *et al.* 2001) and the p23 gene template was amplified using the PM50 and PM51 primer pair (Sambade et al. 2003), (*Table 2*) and the same reagents and PCR reaction conditions described earlier. The amplified products were analysed on 1% agarose gels and the single bands were purified using a NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany), according to the manufacturers specifications.

Table 2: Genotype specific primer sequences used to amplify the insert of each specific plasmid insert. Bases that differ from the published “standard” primers are shown in red, with the last column indicating the total number of differences between genotype specific primers and the previously published primer pair (Rubio et al, 2001, Sambade et al, 2003)

Gene/region	Primer orientation	Primer name	Strain represented	Primer sequence (5' to 3')	Number of differences
A-region	Forward	A-F standard	Consensus	AGCTGTTCTGTGAAACGCGG	-
		A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
	Reverse	A-R standard	Consensus	GTCGATAACTCGACAAACGA	-
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
		A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
p23 gene	Forward	A-F M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
	Reverse	A-R M16	HA16-5	ACGTGTTTGTGAAACGCGG	1
p23 gene	Forward	PM50	Consensus	ACTAACTTTAATTCGAACA	-

The purified products were quantified and cloned according to the same protocol described earlier. A total of 131 clone inserts for the A-region and 118 for the p23 gene were sequenced and aligned using CLUSTAL W alignment software (EBI, Cambridgeshire, England) within the BioEdit Sequence alignment editor 7.1.3 (Hall 1999), using the original clone insert sequences as references. Neighbour-joining phylogenetic trees was constructed for the alignments, using MEGA 4.1 (Tamura et al. 2007) and Maximum Composite Likelihood substitution model with a 1000 bootstrap replicates. The number of clones obtained for each genotype was determined and significance of the deviation from an expected equivalence of numbers was determined using the Chi-squared test.

RESULTS

Defining CTV groups within phylogenetic trees

In order to define the phylogenetic groupings for the A-region and p23 gene, the cognate sequences of each reference were aligned and subjected to two different phylogenetic analyses. The neighbour joining phylogenetic trees yielded the greatest bootstrap support among the phylogenetic clusters. Generally, branches with bootstrap support of > 75% were considered to represent CTV genotypes. The naming of clades was based on that of the established CTV genotype they contained (VT, T68, T3, T30, T36 and RB) (Harper 2013). The sequence of the A-region allows the resolution of 8 different clades, namely VT, T30, T3, HA 16-5, T68, T36, and two clades of RB called RB1 and RB2 (*Figure 1*). The p23 gene resolves the same clades but does not separate the RB genotype into two clades (*Figure 2*). Bootstrap values supported these clades with the exception of the VT and T30 branches in the A-region phylogeny and the T36 and HA 16-5 branches in the p23 gene phylogenies which have bootstrap values of 43 and 32 respectively. However to retain conformity with previous genotype classification schemes (Hilf et al. 2005; Harper 2013) we treat these groups as separate.

A-region phylogenetic analysis and strain prevalence

A measure of the prevalence of the genotypes within the CTV population of each sample was determined by determining the number of cloned sequences grouping within each of the defined clades (*Table 3*). The A-region phylogeny indicated that all of the populations from field-collected samples comprised of one dominant genotype with various minor sequence types. The dominant sequence types were VT-like and ranged from 79% to 100% of the total sequences identified in each sample. T30-like sequences made up the minor strain component of all of the samples except for 312 and 313 and ranged from between 2% to 21% of the total sequences identified in each sample. Sample 312 also contains a sequence with a unique grouping within the A-region phylogeny. The GFMS 12 pre-immunising population from plants grown under greenhouse conditions was dominant for CTZA/T68-like genotypes when maintained on *C. aurantifolia* cv. Mexican lime with 100% of variants grouping with this genotype out of a total of 48 clones. However, a more diverse population was observed on *C. paradisi* cv. Star Ruby with 53% T30-like, 42% CTZA/T68-like and 5% VT/Kpg3-like variants. A total of 38 clones were isolated for the population on Star Ruby. Greenhouse maintained populations of GFMS 35 were also

characterised. A total of 100 clones were isolated for the population on Mexican lime and of these 95% were CTZA/T68-like and 5% were Kpg3/VT-like. A total of 117 clones were sequenced for the population on Star Ruby. These sequences grouped within 5 different groups, with 74% VT/Kpg3-like, 15% T30-like, 7% RB clade2-like, 2% CTZA/T68-like and 2% falling within a unique group branching close to VT/Kpg3 branch. This group falls between VT and T30 with strong bootstrap support of 77. Sample 313 shows a unique grouping falling between T30 and T3. The significance of this group is uncertain since it only shows a bootstrap support value of 65.

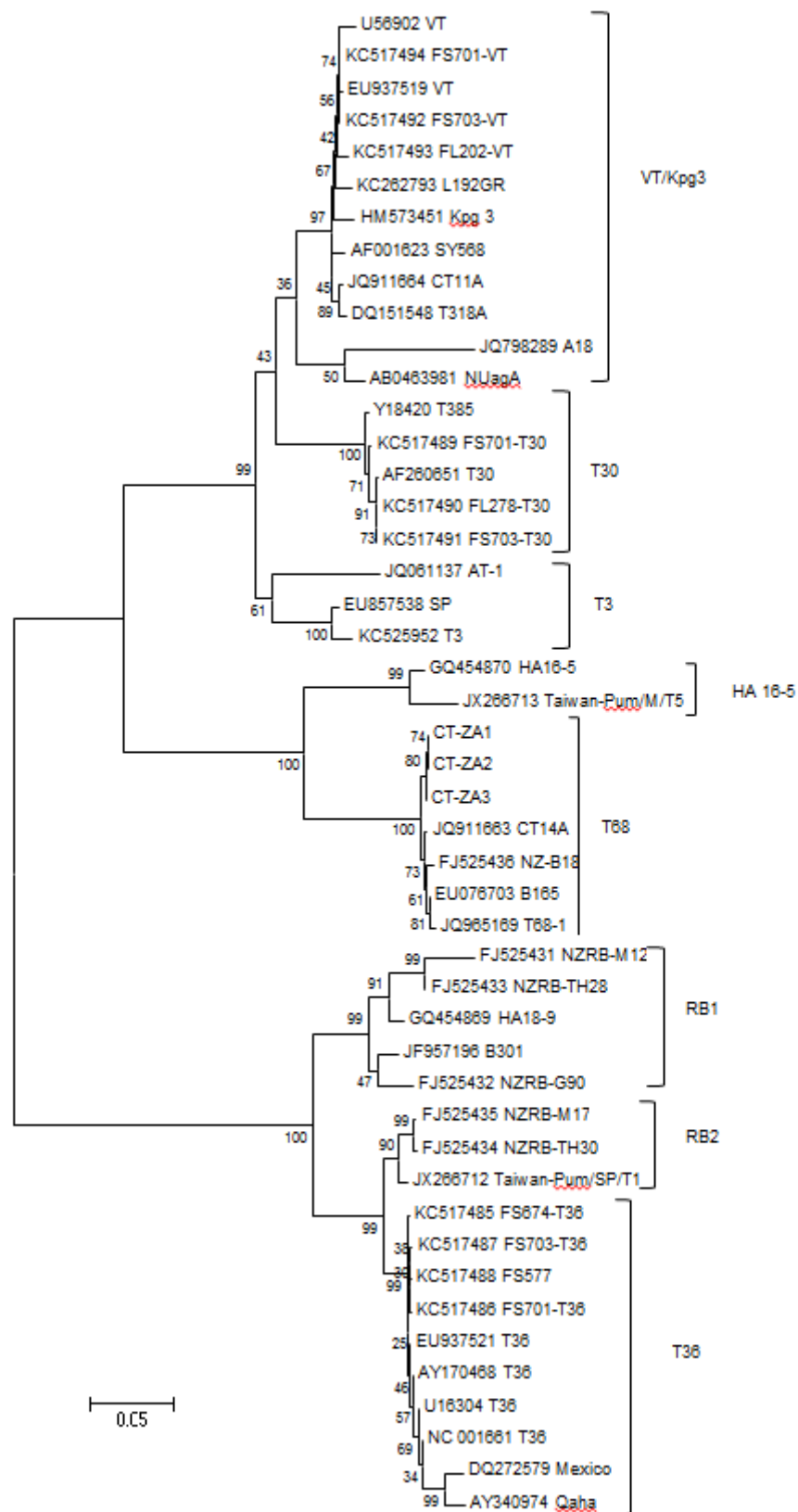


Figure 1: Neighbour-joining phylogenetic tree based on the sequences of the A-region derived from full genome sequences of CTV, obtained from GenBank. Genotypic groups are indicated by brackets and their corresponding labels. Genotypic groups are supported by a bootstrap value of 75 and above. The exception to this constraint is the VT-T30 branch split, which has a bootstrap support of 43 but has been separated for the purposes of conventional CTV nomenclature. Cloned sequences grouping within these genotypic groups were given the name of the corresponding group.

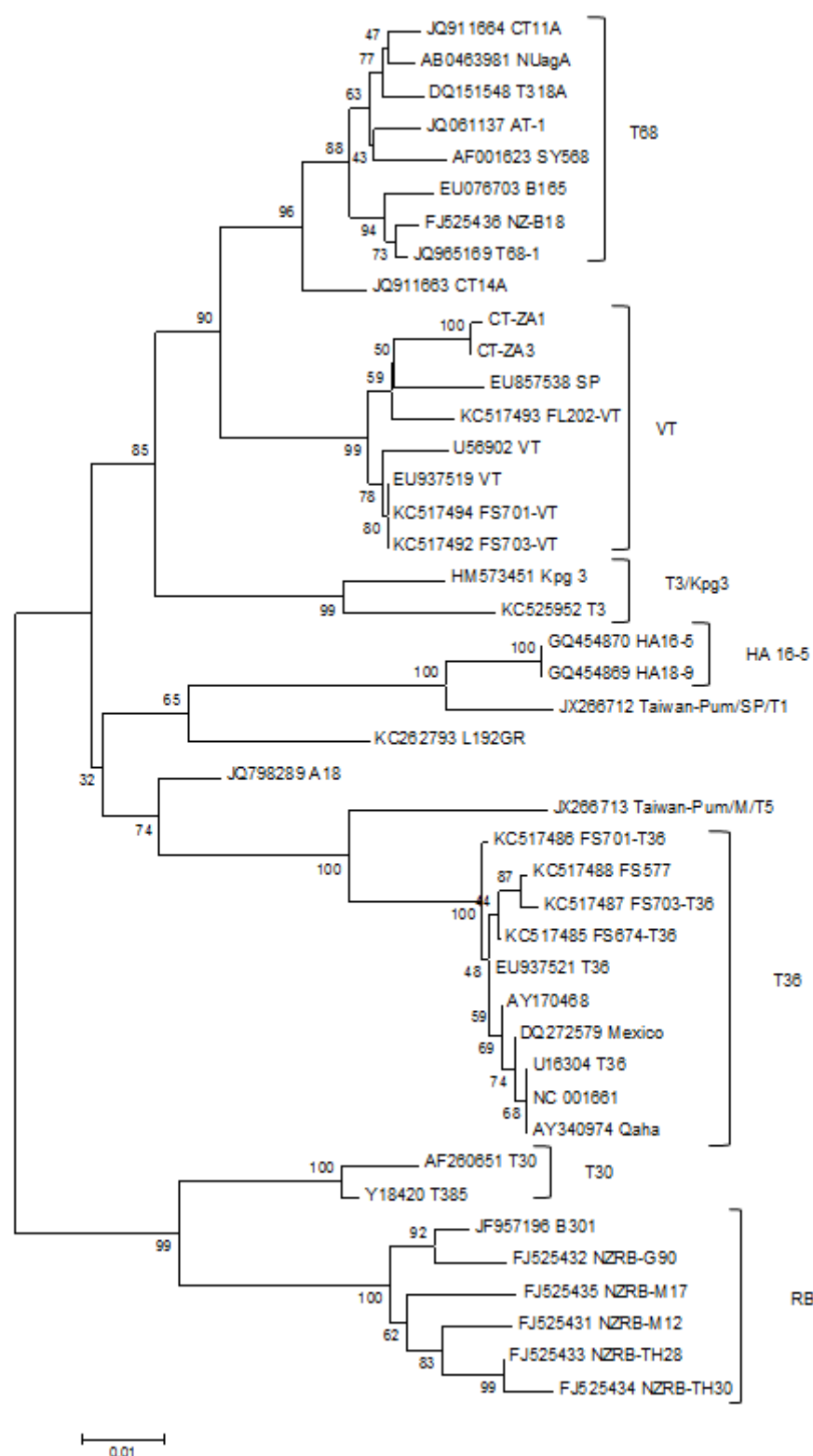


Figure 2: Neighbour-joining phylogenetic tree based on the the p23 gene derived from full genome sequences of CTV obtained from GenBank. Genotypic groups are indicated by brackets and their corresponding labels. Genotypic groups are supported by a bootstrap value of 75 and above. The exception to this constraint is the HA16-5-T36 branch split, which has a bootstrap support of 32 but has been separated for the purposes of conventional CTV nomenclature. Cloned sequences grouping within these genotypic groups were given the name of the corresponding group.

p23 gene phylogenetic analysis and genotype

Analysis of the p23 gene phylogenies also suggest that the CTV population of each sample is made up of a mixed infection of a single dominant genotype together with other genotypes occurring in small amounts. In the case of the p23 gene T3-like sequences dominate each population with a range of 55% to 93% of the total sequences identified in each sample. The minor sequence types grouped with; VT: (4% to 19%), T30:(4% to 16%), HA 16-5: (4% to 23%), T68:(21% to 29%), RB: (3% to 21%) and one unique branch from sample 306 making up 4% of the total sequences (*Table 4*).

The p23 gene analysis of the greenhouse maintained GFMS 12 pre-immunising population from Star Ruby showed that 100% of a total of 60 sequenced clones grouped with VT/CTZA. The GFMS 12 population maintained in Mexican lime was dominant for RB-like variants with 53% of the 58 clones grouping with this genotype. The remaining sequence variants sequenced for this group were 45% VT/CTZA and 2% of the clones falling within a unique group branching close to the VT/CTZA group. The GFMS 35 populations from Mexican Lime and Star Ruby were very similar in terms of their population structures. From the Star Ruby and Mexican lime populations, 69 and 91 clones were sequenced respectively, with 32% and 47% grouping with HA 16-5, 67% and 52% grouping with RB and 1% each grouping within a unique branch close to T30.

Determination of primer bias in the A-region and p23 gene

Table 5 shows the number of clones obtained for specific genotypes, following amplification from a template containing equimolar concentrations of the amplicons derived from clones representing each genotype, as well as the number of differences (mismatches) between the published A-region and p23 gene primer pairs utilised and the forward and reverse primer binding sites of each respective genotype. Significant differences ($P = 0.05$) in numbers of clones obtained of the genotypes was observed with both primer sets utilised (*Table 2*). While both the A-F and A-R primers have some differences in sequence between various CTV genotypes and the primer at the primer-binding site (some occurring in the critical 3' end of the primer), only one to two nucleotide differences were observed between 11 of the 44 reference sequences and the p23 gene forward primer (PM50) and no differences exist within the p23 gene reverse primer which was totally conserved with respect to all of the reference sequences used. VT sequences were significantly

overrepresented ($P = 0.05$) amongst the A-region clones, being 44 (34%) of the total number of clones sequenced rather than the expected 12.5% if no bias had occurred. This was expected since the VT primer binding sites have the least number of nucleotide differences ($n = 1$) between the primer and the genotype sequence in the forward and the reverse primers combined. This is the lowest among all of the variable primer binding sites. However, T3 sequences also represented 34% of the total cloned sequences while three nucleotide differences between primer and genotype sequence exist. Conversely, the T30 primer binding sites have single nucleotide differences between primer and the genotype in the forward and reverse directions, yet, clones with this sequence are underrepresented at only 8 (6.1%) of the total number of clones. Interestingly, representatives of B165 were present in 19 (15%) clones, similar to the expected 16 (12.5%) in spite of the four nucleotide difference in both the forward and reverse directions. Based on the low number of nucleotide difference within the primer binding sites of these strains and the primers, it was expected that T30 sequences would be more abundant and with a number of clones similar to VT and T3. The remainder of the strains each had between five and nine nucleotide differences within their primer binding sites and clones representing these, made up less than 6% each. In spite of the bias observed, only the T36 genotype was not represented by any clones, while all others were detected.

In contrast, T36 sequences were significantly overrepresented ($P=0.05$) amongst the p23 gene clones, being 41 (35%) of the total clones sequenced. When excluding the overrepresented T36 clones, no significant ($P=0.05$) bias towards particular genotypes among the remainder of the clones occurred with SP, Kpg3, RB TH-28 and T30 representing 21 (18%), 19 (16%), 19 (16%) and 18 (15%) of the total clones respectively.

Table 3: Number of cloned sequences that clustered within each of the defined A-region genotypic groups.

Pre-immunizing source	Cultivar	Sample name	Percentage of total clones corresponding to genotypes									Number of clones analysed
			VT/Kpg3	T30	T3	HA 16-5	T68/CTZA	RB1	RB2	T36	Unique group	
GFMS 12	Star Ruby	301	90.7	9.3	-	-	-	-	-	-	-	54
		302	100	-	-	-	-	-	-	-	-	47
		303	98	2	-	-	-	-	-	-	-	50
		313	87	-	-	-	-	-	-	-	13	54
	Flame	304	98	2	-	-	-	-	-	-	-	48
		305	94	6	-	-	-	-	-	-	-	54
		306	87.5	12.5	-	-	-	-	-	-	-	48
	Star Ruby	Pre-immunizing source	5	53	-	-	42	-	-	-	-	38
	Mexlime		-	-	-	-	100	-	-	-	-	48
GFMS 35	Star Ruby	307	79	21	-	-	-	-	-	-	-	47
		308	100	-	-	-	-	-	-	-	-	55
		309	100	-	-	-	-	-	-	-	-	56
	Flame	310	79	21	-	-	-	-	-	-	-	53
		311	98	2	-	-	-	-	-	-	-	54
		312	88.5	-	-	-	-	-	-	-	11.5	52
	Star Ruby	Pre-immunizing source	74	15	-	-	2	-	7	-	2	117
	Mexlime		5	-	-	-	95	-	-	-	-	55

Table 4: Number of cloned sequences that clustered within each of the defined p23 gene genotypic groups.

Pre-immunizing source	Cultivar	Sample name	Percentage of total clones corresponding to genotypes								Number of clones analysed
			VT/CTZA	T30	T3/Kpg 3	HA 16-5	T68	RB	T36	Unique group	
GFMS 12	Star Ruby	301	-	16	72	-	-	12	-	-	25
		302	12	-	88	-	-	-	-	-	25
		303	4	4	92	-	-	-	-	-	27
		313	-	-	91	-	-	9	-	-	23
	Flame	304	-	-	75	4	-	21	-	-	28
		305	-	8	69	23	-	-	-	-	13
		306	-	-	85	7	-	4	-	4	28
	Star Ruby	Pre-immunizing source	100	-	-	-	-	-	-	-	60
	Mexlime		45	-	-	-	-	53	-	2	58
GFMS 35	Star Ruby	307	-	-	100	-	-	-	-	-	25
		308	-	3	69	-	24	3	-	-	29
		309	19	4	65	-	-	12	-	-	26
	Flame	310	-	-	75	9	-	16	-	-	24
		311	-	-	85	-	-	15	-	-	26
		312	-	14	55	-	21	10	-	-	29
	Star Ruby	Pre-immunizing source	-	-	-	32	-	67	-	1	69
	Mexlime		-	-	-	47	-	52	-	1	91

Table 5: Number of clones representing the eight sequence variant strains used to determine the degree of PCR primer bias that exists for the published A-region primer pair and the five sequence variants to determine the degree of PCR primer bias that exists for the p23 gene primer pair. The fifth column indicates how many nucleotide differences exist between the published and the strain specific primer binding sites for each forward and reverse primer and the total number of differences for both forward and reverse directions in brackets. For the Chi-squared analysis 14.067 critical value for df=7 ($P=0.05$); 9.488 critical value for df=4 ($P=0.05$)

Gene region represented	Strain name	Number of clones representing their corresponding genotype/expected number of clones in the absence of bias	$(O-E)^2/E$ Chi-squared analysis/Critical value for $P=0.05$	Number of differences between the published and strain specific primer binding sites. Forward / Reverse (Total)
A-region	VT	44/16.4	46.5	- / 1 (1)
	T3	44/16.4	46.5	1 / 2 (3)
	B165	19/16.4	0.4	4 / 4 (8)
	T30	8/16.4	4.3	1 / 1 (2)
	HA 16-5	7/16.4	5.4	2 / 7 (9)
	RB TH-30	6/16.4	6.6	3 / 2 (5)
	RB TH-28	3/16.4	11	3 / 3 (6)
	T36	0/16.4	16.4	4 / 3 (7)
			137.1/14.067	
p23 gene	T36	41/23.6	12.83	-/-
	SP	21/23.6	0.3	-/-
	Kpg3	19/23.6	0.9	-/-
	RB-TH28	19/23.6	0.9	-/-
	T30	18/23.6	1.3	-/-
			16.23/9.488	

DISCUSSION

In this study, replicates of two different red grapefruit cultivars, Star Ruby and Flame, pre-immunised with either the GFMS 12 or GFMS 35 CTV sources were analysed for the genotype composition of the viral population. Viral populations were determined by using the well-established method of cloning viral gene amplicons and performing Sanger sequencing (Beerenwinkel and Zagordi 2011).

Populations from either Star Ruby or Flame, pre-immunised with either GFMS 12 or GFMS 35 yielded similar population compositions, regardless of the cultivar. Within the A-region, the largest numbers of clones of amplicons from these trees consisted of sequences of the VT genotype. Within the A region the VT genotype includes the CTV Kpg3 sequence variant, the relevance of which is discussed below. T30-like sequences were detected in some of these populations, mostly in low percentages. A unique cluster was observed in the dendrograms of one GFMS 35 and one GFMS 12 population. When considering the p23 gene, all field populations

seemed to be composed of T3-like sequence types, which in this region of the CTV genome has a sequence very similar to that of the Kpg3 source. This therefore supports the predominance of this genotype found by the study of the A-region. Clones of this T3/Kpg3 sequence variant comprised more than 50% of sequences from all of the p23 gene sequences from the field. Only one tree sample yielded only T3/Kpg3-like clones whereas all the other samples yielded one to three additional genotypic sequences, occurring as various combinations of VT-, T30, HA16-5, T68-, RB-like genotypes, as well as a novel sequence.

To determine whether specific genotypes are preferentially amplified by the primer pairs utilised in this study and to make sense of incongruences between the two regions sequenced, a study was conducted whereby an equimolar mixture of a number of A-region or p23 gene amplicons derived from clones and containing genotype-specific primer binding sites was prepared. This experiment was conducted in view of the occurrence of some nucleotide variability amongst CTV genotypes at the primer binding sites for the A-region (Rubio et al. 2001), resulting in mismatches between the primers and their corresponding binding sites. The VT and T3 genotypes appeared to be preferentially amplified when using the A-region primers. These sequences were significantly overrepresented at ~ 3 times the expected level, which was probably due to the low number of mismatches between the primers and targets. T30 sequences were significantly under-represented, probably due to the mismatches within the critical last nucleotides on the 3' end of the primer (Dieffenbach et al. 1993). B165, which has amongst the highest number of mismatches within these primers, yielded a slightly higher than expected representation despite having exactly the same mismatches in the last 3' nucleotides of the primer as the T30 genotype in addition to additional mismatches. HA 16-5, RB1, RB2 and T36 sequences were underrepresented as expected. This was probably due to the increasing number of mismatches between the consensus primer sequences and the genotype specific primer binding sites.

Based on these results it is evident that deviations from the expected representation of genotypes within the clones is not just due to the number and position of mismatches of the primer but possibly also due to the number of clones sequenced being inadequate to account for the population variability. While generally 10-20 clones are sequenced to gain some information on plant viral populations (Kim et al. 2005; Ge et al. 2007; Acosta-Leal et al. 2008) this is inadequate when dealing

with viral populations consisting of various genotypes such as the case with CTV. In the case of the experiment where bias associated with the A-region primers was determined, 131 clones were sequenced from a mixed population of 8 different genotypes. Based on probability theory, hundreds of clones would have had to be sequenced in order to detect all eight genotypes equally, with $P=0.01$, given that no bias exists in the system. While the 131 clones analysed in this study falls far short of the number required to equally represent each genotype within a theoretically unbiased system, it is considerably more than that the number usually isolated for population studies. Given the labour intensive and expensive nature of having to sequence so many clones it is evident that an improved method for analysing CTV populations, possible with the use of next generation sequencing (Beerenwinkel and Zagordi 2011) needs to be developed. Studies are currently underway to achieve this.

In the case of the p23 gene, with no mismatches between primers and their binding sites, only clones of T36 were significantly overrepresented. This cannot be ascribed to differential amplification based on variability of primer binding but is probably due to the fact that “only” 118 clones from the population were sequenced.

While flawed when too few clones are analysed, the use of 25-50 sequenced clones per sample of the p23 gene is still useful as a coarse indicator of the relative composition of the sequence variation found in CTV infected citrus trees but is likely not to detect sequence variants comprising low percentages of the population. From this study, it appears that the Star Ruby and Flame trees, pre-immunized 10 years prior with GFMS 12 and kept in the hot Malelane region, contain CTV populations in which T3/Kpg3 sequence variants predominate, with other sequence variants belonging to the VT-, T30-, HA16-5, T68-, and RB genotypes forming relatively minor components (less than 20%) of the viral population. While the apparent bias of the A-region primers towards VT (the same genotypic group containing Kpg3) and T3 suggests that the relative number of clones obtained for the A-region may not accurately reflect the relative abundance of the genotypes in the plant, the data further supports that of the p23 gene region data in that they both suggest the presence of the dominant Kpg3 sequence variants.

The original Kpg3 sequence variant was sequenced from Mandarin (*Citrus reticulata*) trees displaying decline symptoms in the Darjeeling hills in India (Biswas et al. 2012). Recombination analysis (Biswas et al. 2012) indicated that Kpg3 is a

recombinant and originated through multiple recombination events between divergent CTV sequences. Based on its whole genome, Kpg3 groups with other VT-like sequences, which is also reflected in the current study by the A-region sequences. However based on the p23 gene analysis it does not cluster with any of the accepted genotypes but appears most closely related to the CTV T3 genotype. As only the A-region and p23 gene sequences were determined in the current study and the fact that large portions of the genome of the dominant sequence variant observed within the samples trees were not characterised, this variant should not be referred to as Kpg3, but must be described as “Kpg3-like”, based on A-region and p23 gene sequences. In order to fully describe the local sequence variant, whole genome sequencing is required. T30 sequences were the second most numerous variants derived from A-region and p23 gene clones. However, due to the same limitations that are described above, these variants will also be referred to as “T30-like”.

In this study, both GFMS 12 and GFMS 35 pre-immunising populations maintained on *C. aurantifolia* cv. Mexican lime and *C. paradisi* cv. Star Ruby were characterised according to their A-region and p23 gene sequences. According to the A-region analysis of GFMS 12 on Star Ruby, 53% of the clones grouped within the T30 group, 42% within the T68/CTZA group and 5% within the VT/Kpg3 group. This data compares well with the A-region analysis of Scott et al. (2012), where T30-like variants were found to be dominant within the GFMS 12 population on Star Ruby. This analysis also suggests that the VT/Kpg3-like and T30-like variants found within the field trees could have been components of the original GFMS 12 cross-protecting source. The prevailing climatic conditions within the Malelane production area could have favoured the proliferation of the VT/Kpg3 component while repressing the replication of the T30-like component. The p23 gene analysis of GFMS 12 on Star Ruby suggests that the original GFMS 12 population is homogenous for VT/CTZA-like variants. The p23 gene analysis depicts a population that is less diverse than that suggested by the A-region. This could be due to the fact that recombinant variants were present within this GFMS 12 population. The A-region analysis of the GFMS 12 population on Mexican lime suggests that the population is homogenous for variants from the T68/CTZA group. The p23 gene analysis of this population suggests that it is more diverse, with an almost 1:1 ratio of VT/CTZA and RB-like variants. The presence of these RB-like variants could once again suggest the

presence of recombinants or that the A-region primers preferentially amplified CTZA-like variants.

The GFMS 35 populations collected from Star Ruby appear to be more diverse than GFMS 12 when considering the analysis of the A-region, with 74% of the clones being VT/Kpg3-like, 15% T30-like and minor of components of T68/CTZA and RB. The detection of these minor variants may have been due to the large number of clones that were isolated and sequenced. Little agreement exists between the A-region and p23 gene analyses for GFMS 35 on Star Ruby and Mexican Lime. The differences observed between the analyses of these two gene regions could once again suggest the presence of recombinant strains. However, until whole-genome sequences of the individual components of these populations exist, this will remain speculative. The possibility of PCR bias towards certain sequence types, especially in the amplification of the A-region, remains a potential hindrance to the accurate determination of genotype compositions.

Despite these potential shortcomings, a number of important conclusions can be drawn from the analyses of the original pre-immunising sources. Firstly, the resulting differences observed between populations on Mexican lime and Star Ruby, pre-immunised with the same cross-protecting source suggests that cultivars could play an important role in determining which variants of a population are supported. Secondly, the A-region analysis of both the GFMS 12 and GFMS 35 populations suggest the presence of both VT/Kpg3-like and T30-like variants, which were also found to be the dominant variants in the populations analysed from the field. VT-like variants were also found to be the major components of the GFMS 12 populations from Star Ruby and Mexican lime. The components within the field populations could therefore be present due to the original GFMS 12 pre-immunisation of the trees. The p23 gene analysis of GFMS 35 however, suggests that it is quite different from the population of GFMS 12 and therefore the components present within the field grown trees, pre-immunised with GFMS 35, may have been introduced through secondary inoculations while other components of the original population could have been lost through cultivar selection.

Mild-strain cross-protection has become an essential component for ensuring that the southern African citrus industry can remain productive and competitive. The presence of stem pitting strains, especially on grapefruit cultivars, means that cross-protection has become the only viable option for extending the productive periods of

many citrus cultivars (Lee et al. 2013). Cross-protection sources have previously been assessed by using symptom evaluation on indicator hosts (Moreno et al. 2008), usually with little information regarding the genetic composition of the populations. Despite this, cross-protection has in some cases been successful in reducing the negative effects of severe strain CTV infections on red grapefruit in southern Africa (van Vuuren et al. 2005). Folimonova et al. (2010) showed that strains provide an absolute exclusion against secondary introductions of the same strain. In light of this, it has been suggested that mild-strains of each strain, present in a region, will need to be obtained and used as cross-protecting sources, in order to elicit effective protection against severe strains present in the field (Scott et al. 2012). The process of selecting effective cross-protecting sources is complicated by factors that can influence CTV populations, such as climate (Broadbent *et al.* 1996), varying replicative capacities of strains in different hosts (Targon et al. 2000) and variable transmission of certain strains depending on the regional vector population (Roy and Brlansky 2009). Molecular characterisation of CTV pre-immunising sources as well as populations that exist in the field will be essential for the improvement of CTV cross-protection schemes.

The phylogenetic relationships between single genes and even full-genomes amongst CTV isolates are complicated by potential recombination events (Harper et al. 2010). Therefore, relating potential symptom expressions of previously, biologically characterised isolates with the sequences of single or even multiple genes should be done with caution, especially in the case of multiple infections of genotypes. Due to the labour and resources required to perform large numbers of SATs to isolate individual genotypes, the more pragmatic approach of gene sequencing was employed as a means of identifying homogeneous South African CTV populations. All of the trees sampled exhibited varying degrees of stem-pitting, which is most often associated with grapefruit cultivars in South Africa (van Vuuren and Manicom 2005). However, only some of the trees exhibited symptoms of decline, which varied from mild to severe and it is not possible to associate these symptoms with any particular strains present in the CTV populations.

A variety of molecular typing methods have been previously used to characterize CTV populations. These include single-strand conformation polymorphism (Rubio *et al.* 1996), restriction fragment length polymorphism (RFLP), bi-directional polymerase chain reaction (BD-PCR) (Jiang et al, 2008) and nucleotide sequence

analysis (Iglesias et al. 2008, Scott et al. 2012, Wang et al. 2013). While techniques such as SSCP and RFLP can indicate differences between CTV populations, the identity of CTV strains can be difficult to determine without actual nucleotide sequencing data. Therefore sequencing analyses are favoured for determining the strain composition of CTV populations.

Sequencing experiments also suffer from a number of potential shortcomings. Among these is the potential for PCR primers to show a bias toward certain variants during amplification, especially those that target consensus sequences that are highly variable between variants that are present within a mixed population. This has been shown to be true for the A-region primer pair that targets a sequence within the highly variable 5' half of the genome. The incorporation of incorrect nucleotides during the PCR reaction through the use of a non-proofreading DNA polymerase may also have introduced false mutations at a very low level (Cline et al. 1996) but was required to keep the costs of this large scale sequencing experiment down. This study has provided insight into the variant composition of 13 CTV populations that were exposed to field conditions for a period of 10 years, as well as the original pre-immunising sources of GFMS 12 and GFMS 35 maintained under greenhouse conditions and has provided the industry with valuable information regarding which variants are potentially circulating within cross-protecting sources, as well as in the field. . The trees sampled for this study were in close proximity to one another, within a single block. Large scale surveys, with sampling from all major citrus production areas will be required for the determination of the regional diversity of CTV within southern Africa's grapefruit production areas. In addition to this, PCR primers targeting highly conserved binding sites that flank a phylogenically informative region of the genome will also need to be developed to prevent the bias of data toward particular variants. These tools will enable the industry to target important strains within the cross-protection scheme, which should allow for the improvement of grapefruit production in the region.

ACKNOWLEDGEMENTS

We gratefully acknowledge funding from Citrus Research International, Agricultural Research Council-Plant Protection Research Institute and the National Research Foundation-THRIP program. We also wish to thank Fanie van Vuuren for his guidance, patience and support during these studies.

REFERENCES

Acosta-Leal, R., Fawley, M. W., & Rush, C. M., (2008). Changes in the intra-isolate genetic structure of Beet necrotic yellow vein virus populations associated with plant resistance breakdown. *Virology*, 376, 60-68

Bar-Joseph, M., Marcus, R., & Lee, R.F., (1989). The continuous challenge of Citrus Tristeza Virus Control. *Annual Review of Phytopathology*, 27, 291-316

Beerenwinkel, N., & Zagordi, O. (2011). Ultra-deep sequencing for the analysis of viral populations. *Current Opinion in Virology*, 1, 413-418

Biswas, K.K., Tarafdar, A., Diwedi, S., & Lee, R.F., (2012). Distribution, genetic diversity and recombination analysis of Citrus tristeza virus of India. *Virus Genes*, 45, 139-148

Broadbent, P., Bransky, R.H., & Indsto, J., (1996). Biological Characterization of Australian Isolates of Citrus Tristeza Virus and Separation of Subisolates by Single Aphid transmissions. *Plant Disease*, 80, 329-333

Černi, S., Ruscic, J., Nolasco, G., Gatin, Z., Krajacic, M., & Skoric, D., (2008). 'Stem pitting and seedling yellows symptoms of *Citrus tristeza virus* infection may be determined by minor sequence variants'. *Virus Genes*, 36, 241-249

Cline, J., Braman, J.C., & Hogrefe, H.H., (1996). PCR fidelity of *Pfu* polymerase and other thermostable DNA polymerases. *Nucleic Acids Research*, 24(18), 3546-3551

Darriba, D., Taboada, G.L., Doallo, R., & Posada, D., (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9, 772

Dieffenbach, C.W., Lowe, T.M., & Dveksler, G.S., (1993). General concepts for PCR primer design. *Genome Research*, 3, S30-37

Folimonov, A.S., Folimonova, S.Y., Bar-Joseph, M., & Dawson, W.O., (2007). A stable RNA virus-based vector for citrus trees. *Virology*, 368, 205–216

Folimonova, S.Y., Folimonov, A.S., Tatineni, S., & Dawson, W.O., (2008). *Citrus Tristeza Virus*: Survival at the Edge of the Movement Continuum. *Journal of Virology*, 82, 6546–6556

Folimonova, S.Y., Robertson, C.J., Shilts, T., Folimonov, A.S., Hilf, M.E., Garnsey, S.M., & Dawson, W.O., (2010). Infection with Strains of *Citrus Tristeza Virus* Does Not Exclude Superinfection by Other Strains of the Virus. *Journal of Virology*, 84, 1314–1325

Folimonova, S.Y., (2013). Developing an understanding of cross-protection by *Citrus tristeza virus*. *Frontiers in Microbiology*, 4, 1-9

Gal-On, A., & Shibolet, Y.M., (2005). Cross-protection. In G. Loevenstein and J.P. Carr (Ed.), *Natural resistance mechanisms of plants to viruses* (pp. 261-288). Dordrecht, The Netherlands: Springer.

Ge, L., Zhang, J., Zhou, X., & Li, H., (2007). Genetic Structure and Population Variability of Tomato Yellow Leaf Curl China Virus. *Journal of Virology*, 81, 5902-5907.

Gowda, S., Satyanarayana, T., Davis, C.L., Navas-Castillo, J., Albiach-Marti, M.R., Mawassi, M., et al. (2000). The p20 Gene Product of *Citrus Tristeza Virus* Accumulates in the Amorphous Inclusion Bodies. *Virology*, 274, 246-254

Guindon, S., & Gascuel, O., (2003). A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52, 696-704

Hall, T.A., (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98

Harper, S.J., Dawson, T.E., & Pearson, M.N., (2010). Isolates of *Citrus tristeza virus* that overcome *Poncirus trifoliata* resistance comprise a novel strain. *Archives of Virology*, 155, 471-480

Harper, S.J., (2013). *Citrus tristeza virus*: evolution of complex and varied genotypic groups. *Frontiers in Microbiology*, 4 (93), 1-18

Hilf, M.E., Mavrodieva, V.A., & Garnsey, S.M., (2005). Genetic Marker Analysis of a Global Collection of Isolates of *Citrus tristeza virus*: Characterization and Distribution of CTV Genotypes and Association with Symptoms. *Phytopathology*, 95, 909-917

Iglesias, N.G., Gago-Zachert, S.P., Robledo, G., Costa, N., Plata, M.I., Vera, O., et al. (2008). Population structure of *Citrus tristeza virus* from field Argentinean isolates. *Virus Genes*, 36, 199-207

Jiang, B., Hong, N., Wang, G., Hu, J., Zhang, J., Wang, C., et al. (2008). Characterization of *Citrus tristeza virus* strains from southern China based on analysis of restriction patterns and sequences of their coat protein genes. *Virus Genes*, 37, 185-192

Karasev, A.V., Boyko, V.P., Gowda, S., Nikolaeva, O.V., Hilf, M.E., Koonin, E.V., et al. (1995). 'Complete Sequence of the Citrus Tristeza Virus RNA Genome'. *Virology*, 208, 511-520

Katoh, K., & Standley, D.M., (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30, 772-780

Kim, T., Youn, M. Y., Min, B. E., Choi, S. H., Kim, M., & Ryu, K. H., (2005). Molecular analysis of quasispecies of Kyuri green mottle mosaic virus. *Virus Research*, 110, 161-167

Lee, R.F., & Keremane, M.L., (2013). Mild strain cross-protection of tristeza: a review of research to protect against decline on sour orange in Florida. *Frontiers in Microbiology*, 4 (259), 1-11

Luttig, M., van Vuuren, S.P., & van der Vyver, J.B., (2002). Differentiation of single aphid cultured sub-isolates of two South African *Citrus tristeza virus* isolates from grapefruit by single-strand conformation polymorphism. In: *Proc 15th Conf. IOCV*, 186-196, IOCV, Riverside, California.

Moreno, P., Ambrós, S., Albiach-Martí, M.R., Guerri, J. & Peña, L., (2008). '*Citrus tristeza virus*: A pathogen that changed the course of the citrus industry'. *Molecular Plant Pathology*, 9, 251-268

Niblett, C.L., Genc, H., Cevik, B., Halbert, S., Brown, L., Nolasco, G., et al. (2000). Progress on strain differentiation of *Citrus tristeza virus* and its application to the epidemiology of citrus tristeza disease. *Virus Research*, 71, 97-106

Oberholzer, P.C., Mathews, I., & Stiemie, S.F., (1949). The decline of grapefruit trees in South Africa. A preliminary report on so-called "Stem-Pitting". *Science. Bulletin Department of Agriculture South Africa*, 287, 17

Roy, A., & Bransky R.H., (2009). Population Dynamics of a Florida Citrus tristeza virus Isolate and Aphid-Transmitted Subisolates: Identification of Three Genotypic Groups and Recombinants After Aphid Transmission. *Phytopathology*, 99, 1297-1306

Rubio, L., Ayllón, M.A., Guerri, J., Pappu, H., Niblett, C., & Moreno, P., (1996). Differentiation of citrus tristeza closterovirus (CTV) isolates by single-strand conformation polymorphism analysis of the coat protein gene. *Annals of Applied Biology*, 129, 479-489

Rubio, L., Ayllón, M.A., Kong, P., Fernandez, A., Polek, M., Guerri, J., Moreno, P., & Falk, B.W., (2001). Genetic Variation of *Citrus Tristeza Virus* Isolates from California and Spain: Evidence for Mixed Infections and Recombination. *Journal of Virology*, 75, 8054-8062

Sambade, A., Lopez, C., Rubio, L., Flores, R., Guerri, J., & Moreno, P., (2003). Polymorphism of a specific region in gene *p23* of *Citrus tristeza virus* allows for the discrimination between mild and severe isolates. *Archives of Virology*, 148, 2325-2340

Sambrook J. (2001). In *Molecular Cloning: 'A laboratory manual'*. Cold Spring Harbor, New York. Cold Spring Harbor Laboratory Press.

Scott, K.A., Hlela, Q., Zablocki, O., Read, D., van Vuuren, S., & Pietersen G., (2012). Genotype composition of populations of grapefruit-cross-protecting citrus tristeza virus strain GFMS12 in different host plants and aphid-transmitted sub-isolates. *Archives of Virology*, 158, 27-37

Souza, A.A., Müller, G.W., Targon, M.L.P.N., Takita, M.A., & Machado, M.A., (2002). Stability of the Mild Protective 'PIAC' Isolate of *Citrus tristeza virus*. In: *Proc. 15th Conf. IOCV*, 131-135. IOCV, Riverside, CA

Tamura, K., Dudley, J., Nei, M., & Kumar, S., (2007). *MEGA4*: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596-1599

Targon, M.L.P.N., Machado, M.A., Carvalho, S.A., Souza, A.A., & Müller, G.W., (2000). Differential Replication of a Mild and a Severe Citrus Tristeza Virus Isolate in Species and Varieties of Citrus. In: *Proc. 14th Conf. IOCV*, 127-130. IOCV, Riverside, CA

van Vuuren, S.P & Moll, J.N., (1987). Glasshouse evaluation of citrus tristeza virus isolates. *Phytophylactica*, 19, 219-221

van Vuuren, S.P., & Collins, R.P., (1993). Evaluation of Citrus Tristeza Virus Isolates for Cross Protection of Grapefruit in South Africa. *Plant Disease*, 77, 24-28

van Vuuren S.P., & van der Vyver, J.B., (2000). Comparison of South African Pre-Immunizing Citrus Tristeza Virus Isolates with Foreign Isolates in Three Grapefruit Selections. In: *Proc. 14th Conf. IOCV*, 51-56 IOCV, Riverside, California

van Vuuren, S.P, van der Vyver, J.B., and Luttig, M., (2000). Diversity Among Sub-Isolates of Cross-Protecting Citrus Tristeza Virus Isolates in South Africa. In: *Proc. 14th Conf. IOCV*, 103-110 IOCV, Riverside, California

van Vuuren, S.P., & Manicom, B.Q., (2005). The Response of Star Ruby Grapefruit to Different *Citrus tristeza virus* isolates. In: *Proc 16th Conf. IOCV*, 112-116 IOCV, Riverside, California

von Broembsen, L. A., & Lee, A.T.C., (1988). South Africa's Citrus Improvement Program. In: *Proc. 10th Conf. IOCV*, 407-416 IOCV, Riverside, California

Wang, J., Bozan, O., Kwon, S., Dang, T., Rucker, T., Yokomi, R.K., et al. (2013). Past and future of a century old *Citrus tristeza virus* collection: a California citrus germplasm tale. *Frontiers in Microbiology*, 4 (366), 1-9

Yoon, J.Y., Ahn, H.I., Kim, M., Tsuda, S., & Ryu, K.I., (2006). 'Pepper mild mottle virus' pathogenicity determinants and cross protection effect of attenuated mutants in pepper'. *Virus Research*, 118, 23-30

Zablocki, O., & Pietersen, G., (2014). Characterization of a novel citrus tristeza virus genotypewithin three cross-protecting source GFMS12 sub-isolatesin South Africa by means of Illumina sequencing. *Archives of Virology*, 158 (8), 2133-2139