Detection and isolation of homogeneous genotypes of *Citrus tristeza virus* for use in virus control through cross protection.

By

Kirsti Snyders

(28121903)

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae Microbiology (MSc.)

In the Faculty of Natural and Agricultural Science,
Department of Microbiology and Plant Pathology

University of Pretoria

Pretoria, South Africa
Submitted: April 2015

Supervisor: Prof. G. Pietersen

Declaration of originality

I, ……..... declare that the thesis/dissertation, which I hereby submit for the degree ……..... at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: ..…….. DATE: …….

Acknowledgements

I would like to extend my most sincere gratitude to the following people, as this research project would not have been possible without them:

First and foremost, to my heavenly father for giving me the talent and strength to pursue my dreams, for blessing me with amazing people in my life and carrying me through all the challenging times.

My family for their unconditional love and support during this chapter in my life. Thank you for the everyday encouragement, especially through the testing parts.

My supervisor, Prof. Gerhard Pietersen, for his continuous support and opportunities. Thank you for always making time to help in any possible way.

My lab colleagues; David Read, Ronel Viljoen, Jennifer Wayland, Jackie Lubbe and Megan Harris for the support and friendship. I am extremely grateful to have such amazing friends in my life.

Dr. Stephanus van Vuuren (Oom Fanie) for providing required information for the write up of my dissertation.

Glynnis Cook for all your help and support throughout my entire MSc study.

Jonathan Featherston at Onderstepoort, Agricultural Research Council (ARC) – Biotechnology platform for all your help and guidance with my Next-generation sequencing analysis.

Citrus Research International (CRI) and University of Pretoria for their financial support throughout this project.

The Department of Microbiology and Plant Pathology for the great memories, whether it was work related or a social event.

Last, but not the least, all authors of the respective research articles I made use of within my dissertation. Thank you for making it possible to use your research to back my work.

List of Abbreviations

AC	Average Coverage
Amp	Ampicilin
AMV-RT	*Avian myeloblastosis virus* - reverse transcriptase
ARC	Agricultural Research Council
ARC-ITSC	Agricultural Research Council – Institute for Tropical and Subtropical Crops
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BrCA	Brown citrus aphids
BSA	Bovine serum albumin
CA	California
CCPP	Citrus Clonal Protection Program
cDNA	Complementary Deoxyribonucleic acid
CL	Consensus Length
cm	centimetre
CP	Major coat protein
CPm	Minor coat protein
CREC	Citrus Research Education Center
CRI	Citrus Research International
CTV	*Citrus tristeza virus*
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotriphosphate

dsRNA	Double stranded Ribonucleic Acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ELISA	Enzyme-linked Immunosorbent Assay
FL	Florida
GFMS	Grapefruit Mild Strains
gRNA	Genomic Ribonucleic Acid
g/l	grams/litre
h	hour
HEL	Helicase
HSP	Heat shock protein
HSP70h	Homologue of Heat shock protein 70
IDR	Inter-domain region
IPTG	Isopropyl- ß-D-1 thiogalactopyranoside
KCl	Potassium Chloride
kDa	Kilodalton
KH_2PO_4	Potassium Dihydrogen Phosphate
km	kilometres
KPL	Kirkegaard & Perry Laboratories, Inc.
LB	Luria-Bertani
LMS	Lime Mild Strains
LNR-ITSG	Landbounavorsingsraad - Instituut vir Tropiese en Subtropiese Gewasse

M	Molar
MA	Massachusetts
Mab	Monoclonal antibody
MD	Maryland
MEGA	Molecular Evolutionary Genetics Analysis
mg	milligram
$MgCl_2$	Magnesium Chloride
min	minutes
ml	millilitre
mm	millimetre
mM	millimolar
MMM	Multiple Molecular Markers
MSCP	Mild strain cross-protection
MT	Methyltransferase
NaCl	Sodium Chloride
Na_2CO_3	Sodium Carbonate
Na_2HPO_4	Disodium Hydrogen Phosphate
$NaHCO_3$	Sodium bicarbonate
NCBI	National Center of Biotechnology Information
ng	nanogram
NGS	Next-generation sequencing
nm	nanometres
nt	Nucleotide

NTR	Non-translated region
ORF	Open reading frame
PBST	Phosphate-Buffered Saline/Tween
PCR	Polymerase Chain Reaction
PDR	Pathogen derived resistance
pH	Potential of Hydrogen
pmol	picomol
PRO	Papain-like proteases
PVP-40	Polyvinylpyrolidone-40
QD	Quick decline
TRC	Total Read Count
RdRp	RNA-dependent RNA polymerase
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Reverse transcription
RT-PCR	Reverse transcription-Polymerase Chain Reaction
SACIP	South African Citrus Improvement Program
SAT	Single aphid transmissions
sec	seconds
sgRNA	Sub-genomic Ribonucleic acid
siRNA	small/short interfering RNA
SP	Stem pitting
SSCP	Single strand conformational polymorphisms

ssRNA	Single stranded Ribonucleic acid
SY	Seedling yellows
Taq	Taq polymerase
TAS-ELISA	Triple Antibody Sandwich Enzyme-linked Immunosorbent Assay
tRNA	Transfer Ribonucleic acid
U	Units
UV	Ultra-violet
w/v	Weight/volume
wt	Wild-type
X-gal	Bromo-chloro-indolyl-galactopyranoside
µl	microliters
µM	micro Molar
µg	microgram
µg/µl	micrograms per microliter
µg/ml	microgram per millilitre

LIST OF FIGURES

FIGURE NUMBER **PAGE**

Review of literature (CHAPTER 1).

CHAPTER 2.

Appendix figures.

Appendix 2.

Appendix 4.

LIST OF TABLES

TABLE NUMBER **PAGE**

CHAPTER 3.

Appendix tables.

Appendix 1.

Appendix 5.

Table of Contents

Chapter 1

Summary and Review of Literature.

1.1. Summary.

One of the most serious diseases of citrus worldwide is caused by the viral pathogen *Citrus tristeza virus* (CTV) (family: *Closteroviridae*, genus: *Closterovrius*). This RNA virus is known to cause several economically important syndromes, namely “quick decline”, “stem pitting” and “seedling yellows”. The natural hosts for CTV include nearly all citrus species and cultivars, various citrus relatives such as *Aeglopsis, Afraegle, Fortunella* and *Pamburus*, and some intergeneric hybrids (Yokomi, 2009). The movement of CTV between different regions is mainly due to the dispersal of infected budwood (Bertolini *et al.,* 2008). However, natural transmission and spread of CTV can also occur with the help of several aphid species of which the most effective vector is the Brown citrus aphid (Kirkaldy) *Toxoptera citricida* (Tsai *et al.,* 2009). The severity of a CTV infection may vary depending on various contributing factors, such as environmental conditions, the rootstock combination used and lastly the presence of the CTV genotype predominating within the mixed infection. Currently, there are eight known and published CTV genotypes, each with their own degree of virulence, and up to date, forty-seven CTV genomes have been fully sequenced. It has been observed in numerous CTV population diversity studies, that a host is normally infected with a mixture of genotypic strains (Read and Pietersen, 2015; Roy *et al.*, 2010; Roy and Brlansky, 2004). This complicates biological indexing as well as diagnostic procedures in identifying homogenous genotypic sources. This pathogen is endemic to the citrus producing areas in South Africa; hence, a mild strain cross-protection scheme has been initiated during the year 1973; aiming to control the viral disease.

This cross-protection scheme entails creating CTV resistant citrus trees by pre-immunizing commercially grown citrus cultivars with a mild strain mixture of the virus, which will therefore aid in the exclusion of superinfections from more virulent strains. However, severe CTV infections may still occur on pre-immunized citrus hosts. This type of phenomenon has been explained by Folimonova *et al.* (2010), confirming why this infection process will occur. Within this publication the authors observed cross-protection occurring only against genotypes of which a mild form was present in the pre-immunizing mixture. Therefore, in order to improve future cross-protecting programs, true homogenous CTV genotypic sources of mild and

severe form must be obtained. Having both the mild and severe form of the specific genotypic strain can assist in strain/genotype resistance screening, in which analysis of different mild strain combinations can be tested to see whether cross-protection can be initiated against a variety of severe strain infections. Nevertheless, determining which genotypes are present within a host can be extremely challenging due to the CTV mixtures present within nature. Many techniques such as the reverse transcription-polymerase chain reaction (RT-PCR) analysis, using genotype specific molecular markers Hilf *et al.* (1999), have been developed to not only detect the virus, but to also identifying individual genotypes. However, characterization of the different CTV isolates based on selected genomic regions has been shown to produce conflicting results in phylogenetic studies (Scott *et al.,* 2012) due to the PCR bias detected within the different terminal ends (Read and Pietersen, 2015). To attempt in solving this issue, Illumina high-throughput next-generation sequencing (NGS) was introduced to the viral detection, characterisation and confirmation process (Zablocki and Pietersen, 2014). The addition of NGS, combined with RT-PCR analyses, would yield more conclusive results on CTV findings worldwide.

The objectives of this study are to 1) screen for presence of CTV in leaf samples (Rio Red grapefruit and other citrus cultivars) collected in the field, 2) isolate homogenous genotypes using single aphid transmissions (SAT's) and mechanical inoculations (bark-flap and stem slash) onto healthy Mexican Lime seedlings and lastly 3) to confirm any possible pure CTV sources obtained from all above analyses through NGS. The results obtained from this study will aid in future cross-protection programs against CTV as well as improving our understanding about this viral disease.

1.2. Review of literature

1.2.1. *Citrus tristeza virus* (CTV) – A general overview as a disease causing agent of citrus.

During the year 1836 a foot-rot epidemic caused by the genus group *Phytophthora* (Roistacher *et al.,* 2010; Moreno *et al.*, 2008), in which citrus plantings established from seeds, bud propagations on sweet orange [*Citrus sinensis* (L.) Osb.] seedlings or on rooted citron [*C. medica* (L.)] cuttings were detroyed. This epidemic forced citrus producers to produce budwood propagation of citrus cultivars on sour orange [*C. aurantium* (L.)], a then well known foot-rot-resistant rootstock (Moreno *et al.,* 2008). Sour orange was highly adaptable to all soil types and known to generate high quality fruits (Albiach-Marti, 2012), which was used almost exclusively in the Mediterranean area and later in America. This rootstock was critical in citrus production, but unknowingly citrus growers made use of a severely sensitive rootstock to protect their citrus trees against a then unknown citrus disease *Tristeza* (CTV) (Moreno *et al.,* 2008). The interaction of CTV with the new sensitive sour orange host [*C. aurantium* (L.)], commercially created after the propagation of sweet orange, mandarin (*C. reticulate* Blanco) and grapefruit (*C. paradisi* Macf.), initiated a new citrus disease epidemic (Moreno *et al.,* 2008).

CTV is an aphid-borne closterovirus which has been ranked as one of the most important citrus diseases for the last sixty years (Bar-Joseph *et al.*, 1989). CTV has caused numerous epidemics worldwide leading to severe decline as well as death of millions of citrus trees. CTV sequence variations have evolved through time by means of genetic drift, negative and positive selection (Albiach-marti, 2012), bottlenecks (Sanz-ramos *et al.,* 2008), recombination events, viral strain competitions and the effect of quasispecies (Domingo, 2002; Lauring and Andino, 2010). A quasispecies can be defined as diverse variants, of a (viral) family (like CTV), that are genetically linked through sequence mutations, in which all variants can interact together on a functional level, and contributing to the characteristics of the (viral) population (Lauring and Andino, 2010). These newly evolved sequence variants lead to the expression of different symptoms. The variation between the symptoms expressed can also be due to genomic differences in CTV strains by means of replication errors made by the RNA-dependent RNA polymerase (Albiach-Martí *et al.,* 2000; Cerni *et al.,* 2008; Hull, 2002).

The different CTV isolates can be classified as either mild or severe depending on the symptoms elicited. Mild CTV isolates typically show superficial symptoms or no symptoms at all (Hilf *et al.,* 2005), whereas severe isolates are able to produce economically devastating (field) symptoms (Figure 1), depending on the viral strain, the citrus host being infected by the specific aphid vector, the rootstock-scion combination and the environmental conditions (Ruiz-Ruiz *et al.,* 2007). However, the interactions that occur between these parameters, to produce varying symptoms, is still unknown (Lauring and Andino, 2010). Also, one should keep in mind that symptom severity can increase with time as continuous infections take place within orchards. The three symptoms induced are:

1) *Declining of trees (Quick decline, QD)*: CTV infections to citrus trees will trigger phloem cell necrosis in the sour orange rootstock, just below the bud union. This necrotic reaction will cause the tree canopies to wilt, eventually leading to complete collapse and tree death in only a few weeks (Roistacher *et al.,* 2010). From all the symptoms QD is the most dramatic and economically devastating (Moreno *et al.,* 2008).
2) *Stem pitting (SP) and stunting*: Second to tree death, stem pitting *tristeza* is one of the most evident and visually alarming symptom appearances of all the CTV symptoms, inducing SP in the scions and/or rootstocks of lime, grape fruit and sour orange rootstock (Roistacher *et al.,* 2010). Citrus trees affected by CTV stem pitting strains will produce a lower yield of citrus together with a lower product quality (Stewart, 2006). There are varying degrees of mild to severe stem pitting of trees. The trees infected are typically stunted, have a bushy appearance, the leaves are chlorotic, and the twigs are very brittle and break easily (Garnsey *et al.,* 2005).
3) *Seedling Yellows (SY)*: Certain CTV strains have the ability to induce seedling yellows reactions in seedlings of lemon, grapefruit and sour orange. This type of CTV symptom is known more as a “glasshouse disease” than as a field symptom, due to its ability to cause tree losses in nurseries (Stewart, 2006). The symptoms linked with SY include severe leaf chlorosis, reduction of leave size and seedling stunting (Garnsey *et al.,* 2005). Seedlings of Eureka lemon, grapefruit or sour orange are excellent indicator plants to detect SY (Roistacher *et al.,* 2010, 2004).

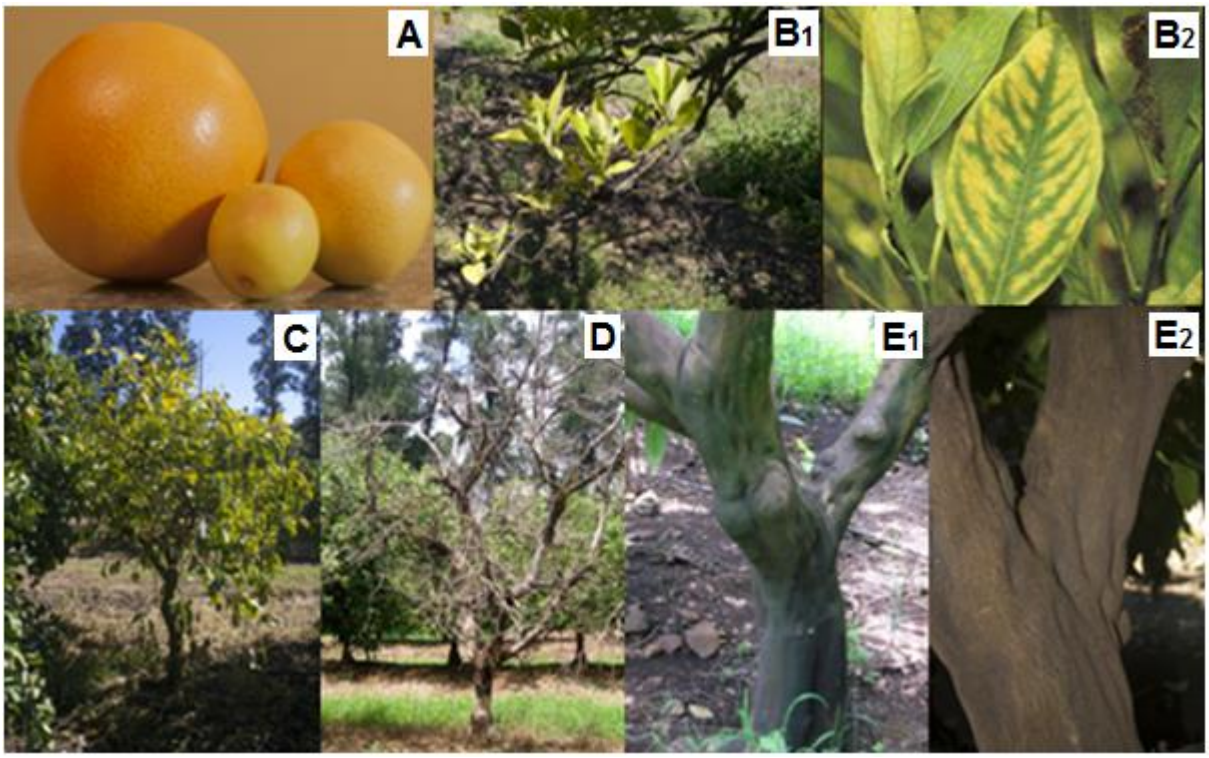


Figure 1: Schematic illustration CTV induced symptoms – [A] Reduction if fruit size, [B_1 – B_2] seedling yellows, [C] Tristeza decline (Quick decline), [D] tree death and [E_1 – E_2] stem pitting (Images; Jackie Lubbe [A]; http://ipm.ifas.ufl.edu/Agricultural_IPM/asian.shtml [B_2]).

The effect severe CTV strains have on the citrus industry worldwide is economically devastating. It is still not yet possible to link exact symptoms to a particular citrus cultivar or to a specific viral sequence, nor is it known which viral sequences influence the transmissibility of CTV by different aphid species (Roistacher *et al.,* 2007). This is because CTV isolates are commonly found as mixed populations (Albiach-martí *et al.,* 2000) of different gRNA populations, where some also contain multiple defective RNA's. The occurance of these mixed viral populations are usually due to CTV transmissions by means of infected budwood or aphid species. The propagation of citrus with CTV infected budwood is a well known manner to spread this viral pathogen locally and internationally (Costa *et al.,* 2010).

Aphid vectors are known to transmit CTV in a semi-persistent manner (Saponari *et al.,* 2008), especially when these aphid species are present on the citrus plants (Cerni *et al.,* 2008). The aphid species *Toxoptera citricida* (brown citrus aphid), *Aphis gossypii* and *Aphis spiraecola* are the most important CTV vectors, with the brown citrus aphid (BrCA) known as the most efficient CTV transmitter

(Bertolini *et al.,* 2008; Korkmaz *et al.*, 2008; Huang *et al.,* 2005). Despite the major economic impact of this aphid-virus association, only a small amount of information is available on the specific relationship between CTV and BrCA and the spread of the disease (Herron *et al.,* 2006). This limited information makes studying CTV epidemiology as well as the controlling of the disease very difficult.

CTV is a major concern in South Africa (SA) as most of the citrus produced are exported globally. Exporting of citrus, from CTV infected orchards, rarely occurs due to the low quality yield obtained. Grapefruit is one of the biggest citrus cultivars that is being exported worldwide form South African borders; however, CTV has caused severe losses in SA due to CTV epidemics (Scott *et al.,* 2012).

1.2.2. CTV molecular characterization

Citrus tristeza virus (CTV) belongs to the *Closterovirus* genus in the plant virus family *Closteroviridae* (Moreno *et al.,* 2008; Satyanarayana *et al.*, 2003). The virions produced are filamentous, flexuous (Soler *et al.*, 2012; Bar-Joseph *et al.,* 2002) and approximately 11nm x 2000nm in dimension (Albiach-marti, 2012), making CTV one of the largest plant viruses known up to date (Dolja *et al.,* 2006). The CTV particles are phloem restricted and can be detected in the leaves, stems, fruits and roots of the infected citrus, where they will occur as inclusion bodies (Brlansky and Lee, 1990; Folimonova *et al.,* 2008).

The CTV genomes contain a non-segmented, positive sense, single-stranded RNA genome (Karasev *et al.,* 1995) varying in length between 19226nt and 19302nt, depending on the CTV isolate (Albiach-Marti, 2012; Ayllon *et al.,* 2001). Each genome encodes for 12 open reading frames (ORF), which potentially expresses at least 19 proteins, and two non-translated regions (NTRs) of about 107 and 273nt (Roy and Brlansky, 2010; Moreno *et al.*, 2008) at the 5’ and 3’ termini of the genomic RNA (gRNA) (Ruiz-Ruiz *et al.,* 2006; Che *et al.*, 2001). The 3’-terminus of the CTV genome is highly conserved with 97% sequence identity, whereas the 5’-terminus is much less conserved as the sequence identity is less than 70% sequence identity, even as low as 44% in some cases (Ayllon *et al.,* 2001).

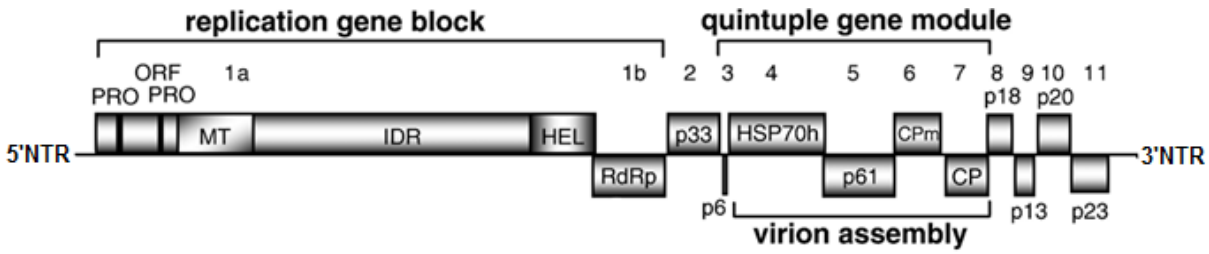


Figure 2: A schematic representation of the CTV genome organization. The proteins are represented as boxed and the protein products are numbered. PRO, MT, HEL and RdRp indicate protein domains of papain-like protease, methyltransferase, helicase and RNA-dependent RNA polymerase, respectively. HSP70h, CPm and CP indicate ORF's encoding a homologue of heat shock protein 70, the minor and the major coat proteins, respectively (Reproduced from the original figure published by Tatineni *et al.*, 2008).

The expression of the 12 ORF's includes at least three RNA expression mechanisms, which are widely used by positive-strand RNA viruses: 1) proteolytic processing of a polyprotein precursor, 2) + 1 translational frameshifting, producing two large polyproteins, one spanning MET–HEL, and the other one encompassing MET–HEL–POL (Dolja *et al.,* 2006), and 3) generating a nested set of 3'-coterminal sub-genomic RNA's (sgRNA) (Karasev *et al.,* 1997). ORF1a and 1b is known as the "*replication gene block*" (Figure 2), since the products produced by ORF 1a and 1b are required for viral replication (Tatineni *et al.,* 2008). This "*replication gene block*" also makes up the 5'-terminal half of the CTV genome. Both ORF 1a and 1b are translated directly from the positive stranded gRNA, producing a 400kDa polyprotein that would be proteolytically processed into at least nine protein products (Karasev *et al.,* 1995). Separately, ORF1a encodes a 349kDa polyprotein which contains two papain-like protease domains, a type 1 methyltransferase-like domain and a helicase-like domain (Yang *et al.,* 1999), while ORF 1b will encode a 54kDa protein, containing RNA-dependent RNA polymerase (RdRp) domains. This protein (RdRp) is occasionally expressed after OFR1a by means of a +1 ribosomal frameshifting reaction (Roy *et al.,* 2005).

The other 10 ORF's on the 3'-terminal end of the CTV genome are expressed during the synthesis of a set of 3' co-terminal sgRNA's (Karasev *et al.,* 1995), which encodes for the rest of the CTV proteins (p33, p6, p65, p61, p27, p25, p18, p13, p20, p23) (Soler *et al.*, 2012; Martın *et al.,* 2009; Hilf *et al.,* 1995). Each 3'-sgRNA acts as a messenger to initiate the translation of its 5'ORF (Hilf *et al.,* 1995), while the

expression of the ten 3' ORF's is regulated independently both in timing and amount present (Moreno *et al.*, 2008). Some part of the CTV 3' ORF's are also enclosed in a conserved gene area known as the "*quintuple gene block*" (Figure 2), which aids in virion assembly and trafficking (Dolja *et al.,* 2006). The "*quintuple gene block*" consists out of the major coat protein (CP of 25kDa), the minor coat protein (CPm of 27kDa) (Febres *et al.,* 1996), p61, HSP70h (heat-shock protein, 65kDa) and p6 (hydrophobic protein belonging to the single-span transmembrane proteins) (Albiach-Marti, 2012). CP, CPm, HSP70h and p61 are structural proteins, which is also proposed to play a role in affecting aphid transmission (Barzegar *et al.,* 2010). Approximately 97% of the CTV genome is covered by CP (p25), while the remaining 3% is encapsidated by CPm (p27), resulting in the production of a CTV virion containing an emblematic tail (Febres *et al.,* 1996; Albiach-marti, 2012). The harmonious action of HSP70h and p61 is required for proper CTV virion assembly (Satyanarayana *et al.,* 2000). During the assembly of the CTV virion, HSP70h and p61 will bind to the transition zone between CP and CPm (~630nt), which will restrict CPm to the virion tail (Satyanarayana *et al.,* 2004; Moreno *et al.*, 2008; Tatineni *et al.,* 2010). Although p6 is not required for virus assembly or replication, it is highly required for systemic invasion of the host plant (Tatineni *et al.,* 2008). Known as a transmembrane protein, p6 possibly also helps in virus movement (Folimonova *et al.,* 2010; Ambrós *et al.,* 2011).

The additional five CTV ORF's located at the 3-end of the CTV genome are the p20 ORF and four genes (p33, p18, p13 and p23) that encode for proteins that have no homologue in any other Closteroviruses (Dolja *et al.,* 2006). The p20 protein has been detected accumulating in the infected tissue of the plant, has a high affinity for itself and is mainly localized the infected cells within the anamorphous inclusion bodies (Gowda *et al.,* 2000). Furthermore, together with p6, p20 is required for CTV systemic infections, suggesting a possible role in CTV translocation in the citrus host (Satyanarayana *et al.,* 2001; Tatineni *et al.,* 2008). However, ORF's encoding for the proteins p33, p18 and p13 are not required for either replication, assembly or systemic infections, as the functions for these protein are still undefined (Folimonova *et al.,* 2010). According to the publication of Tatineni *et al.,* (2011), these proteins can be used as CTV host range determinants. The multifunctional p23 protein contains a Zn finger domain that will bind to both ssRNA and dsRNA molecules in a

non-sequence specific manner (Lopez *et al.,* 2000) and it controls the asymmetrical accumulation of both positive and negative RNA strands during viral replication, ensuring that there will be enough gRNA for virion assembly (Satyanarayana *et al.,* 2002).

CTV has evolved three proteins (p23, CP and p20) that act as RNA silencing suppressors in citrus and *Nicotiana benthamaina* and *N. tabacum* plants (Albiach-Marti, 2012; Amin and Falk, 2009). The p23 protein will inhibit intercellular RNA silencing, while CP will delay intracellular RNA silencing and p20 will block both intra and inter RNA silencing (Lu *et al.,* 2004). Furthermore, the 5' and 3' NTR's both contain cis-acting elements that are highly important for CTV replication (Albiach-marti, 2012). The 5'-NTR of the CTV genome is also protected with a cap structure (Karasev *et al.,* 1995) and comprises a secondary structure which is folded into two stem-loops separated by a short spacer region (Gowda *et al.,* 2003(a); Gowda *et al.,* 2003(b); Gowda *et al.,* 2009). This secondary structure contains all the necessary sequences for both replication and virion assembly (Ayllon *et al.,* 2001; Moreno *et al.,* 2008). The 3'-NTR lacks a poly-A-tail and does not fold into a tRNA structure (Karasev *et al.,* 1995), but consist in a secondary structure of 10 stem-loop structures, containing the required sequences for negative-strand RNA initiation (for the CTV gRNA and sgRNA's) (Satyanarayana *et al.,* 2002).

1.2.3. CTV vectors.

Aphid species are considered as one of the main culprits known to attack citrus worldwide (D'Onghia, 2009) based on their direct effects on the citrus and orchards self. However, the role aphids play in viral (CTV) transmission cannot be ignored. These species belong to the families *Aphidinea* and *Aphididae*, genera *Aphis* and *Toxoptera*. The four most important aphid species linked to CTV distribution and transmission are the Brown Citrus Aphid (BrCA) *Toxoptera citricidus* (Kirkaldy) (Figure 3 [A]), *Toxoptera aurantii* (Boyer de Fonscolombe), *Aphis gossypii* (Glover) (Figure 3 [B]) and *Aphis spiraecola* (*A. citricola*) (Patch) (Papayiannis *et al.,* 2007), however there are other minor aphid species also known to affect citrus orchards (Table 1). The species compositions and the seasonal occurrence vary in different countries and regions.

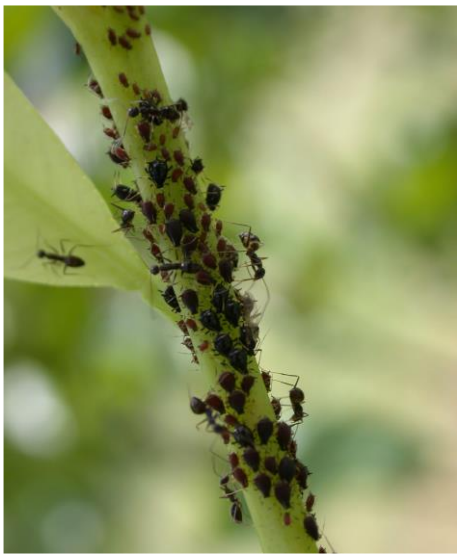
[A]  [B] 

Figure 3: [A] A colony of brown citrus aphids (*Toxoptera citricida*), herded by ants on Ladu tangerine (*C. reticulata* Ladu) foliage, and [B] a colony of melon aphids (*Aphis gossypii*) on Fairchild tangerine (*C. reticulata* Fairchild) foliage (Nelson *et al.*, 2011).

Table 1: Aphid species attacking citrus orchards worldwide (Komazaki, 1993).

Aphid species	Distribution
Major species	
Toxoptera citricidus	Asia, South Africa, Central and South America, Australia, New Zealand
Toxoptera aurantii	Worldwide
Aphis gossypii	Worldwide
Aphis spiraecola (*A. citricola*)	Worldwide
Minor species	
Aphis craccivora	Worldwide
Aphis fabae	Almost worldwide
Aphis nerii	Tropics of the Old and New World
Aulacorthum magnoriae	Japan, India, Korea
Aulacorthum solani	Worldwide
Brachyunguis harmalae	Israel, Sudan
Brachyunguis helichrysi	Worldwide
Myzus persicae	Worldwide
Macrosiphum euphorbiae	Worldwide
Toxoptera odinae	South and East Asia, South Africa
Ureleucon jaceae	Europe, the Middle East and Central Asia

Toxoptera citricida is anholocyclic (without sexual reproduction) and apterous (without wings), where the females produce nymphs. It feeds on newly expanding shoots, leaves and flower buds of its host plant which are fitting for growth and reproduction for a period of 3-4 weeks, depending on the surrounding environmental conditions (Qureshi, 2010; Michaud, 1998). When the source of food diminishes and overcrowding occurs apterae develop wings (alates) and fly away. Most alates do not fly far from their original colony (Michaud, 1998). BrCA outbreaks tend to be endemic in citrus orchards, surviving at low density on bits of asynchronous flush and root sprouts, mostly from the *Rutaceae* family (Pappu *et al.,* 1997; Michaud, 1998). Long-range transmission by alates is unusual and their movement is usually associated with human movement of infested material (Michaud, 1998). *T. citricida* nymphs and apterae transmits CTV more efficiently than the alatae; whereas, in the case of *A. gossypii* both the mature and immature aphids as well as the alatae transmits CTV at similar rates (Bar-Joseph *et al.*, 1989; Komazaki, 1994).

Costa and Grant (1954) showed that a single aphid of *T. citricida* could transmit CTV and has the ability to select and transmit the most severe CTV strains efficiently (Tsai *et al.,* 2009; Roistacher, 2004). *Toxoptera citricida* has been responsible for most of the natural spread of CTV in citrus-growing areas of South America, South Africa, Australia, USA and Asia (Solis-Garcia *et al.,* 2001; Harper *et al.,* 2010).

Apart from *T. citricida*, the aphid life cycle is generally defined by two criteria, the appearance of sexual forms and host alternation. Some aphid species, including *A. gossypii, A. spiraecola* and *T. citricidus*, have both holocyclic (sexual reproducing phase) and anholocyclic (reproducing excluding the sexual phase) strains (Komazaki 1984), while *T. aurantii* is almost completely anholocyclic. However, aphids can also reproduce and spread parthenogenetically (Michaud, 1998), linking the aphid parthenogenesis with viviparity. Sexual and parthenogenetic reproduction usually rotates in the aphid's life cycle. Sexual forms usually appear in fall and lay overwintering eggs on the primary host (holocycle). Eggs hatch in spring, and each hatched larva develops into a mother which reproduces sexual female aphids (fundatrix). Several aphid species and strains of species reproduce parthenogenetically all year round (anholocycle).

Host alternation is another characteristic of aphids (Komazaki, 1984). Most rotate between host plants belonging to different families. During the host rotation, the overwintering aphid eggs will be deposited on the primary host, while nymphs of the parthenogenetically reproducing generation are on the secondary host (usually in the summer). However, some aphid species do not change their hosts, and live on a single species of plant or a group of closely related species.

1.2.4. CTV transmission.

Aphids transmit CTV in a semi-persistent manner, with no latent period; and acquisition and inoculation periods being at least 30 minutes in some cases (Herron *et al.,* 2006). The time for the aphid to tap into the phloem is essential for viral transmission. Aphids can remain viruliferous for 24 to 28 hours after feeding on an infected plant (Yokomi, 2009) and can retain the availability to inoculate CTV during this period when transmitted to secondary hosts (Bar-Joseph *et al.*, 1989).

The transmission of CTV relies on the distribution of infected budwood, the aphid vectors and their abundance in a certain area, the viral strains present, the citrus cultivar infected and the occurring environmental conditions (temperatures) (Michaud, 1998). According to Albiach-marti (2012) and Monero *et al.,* (2008), CTV follows the parameters of a typical disease triangle, as shown below (Figure 4). Still, there is no understanding about how the interactions occur between these four factors (CTV, the host, the vector and the environment). The presence of different CTV strains in nature makes it particularly difficult to recognize the potential interactions mentioned.

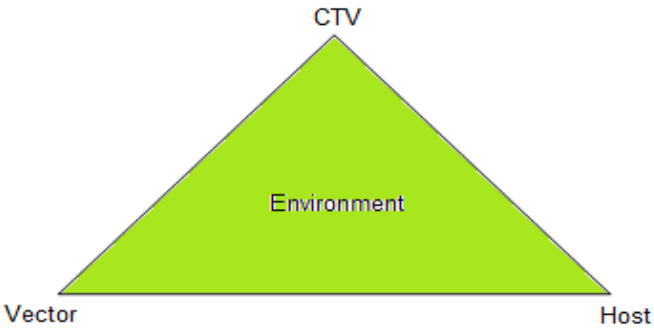


Figure 4: The classic disease triangle followed by CTV (Albiach-marti (2012), Monero *et al.,* (2008) and Stewart *et al.*, 2006).

Single aphid transmission experiments have been shown to be successful in transmitting CTV from an infected tree to a healthy tree (Huang *et al.*, 2005; Lee *et al.*, 2006; Tsai *et al.*, 2009). However this is a relatively inefficient process with low success rates obtained in several research studies, for example 0 – 4% (Komazaki, 1994; Huang *et al.*, 2005;), 25% (Herron *et al.*, 2006; Yokomi *et al.*, 1994) and 16.5 – 18.4% (Tsai *et al.*, 2000). Nevertheless, CTV transmission success rates do increase, for example 0 – 56.7% (Broadbent *et al.*, 1996; Barzegar *et al.*, 2010), when numerous aphids per seedling (10 – 100 aphids per source plant) are used (Broadbent *et al.*, 1996; Barzegar *et al.*, 2010). Still, these varying transmission percentages are a good indication that SAT's are not the most effective and reliable way in obtaining pure CTV sources.

An earlier study by Müller *et al.* (1974), made use of *Passiflora gracilis* to attenuate or reduce CTV genotypic mixtures before grafting budwood onto a variety of indicator hosts. During this experiment, aphids (*Aphis gossypii*) were collected from CTV infected plants (e.g. Sweet Orange) and transmitted onto *P. gracilis* or *P. caerulea*. Afterwards, CTV was transmitted through two different approaches: 1) by performing SAT's directly from the *Passiflora* species onto Mexican Lime or 2) by performing *Passiflora* budwood graft-inoculations onto other *Passiflora* species first, for a further viral genotype reduction, and then SAT's onto Mexican Lime. Thereafter, budwood of the inoculated Mexican Lime was grafted onto chosen indicator hosts. According to Müller *et al.* (1974) and Roistacher *et al.* (2010), attenuation or viral genotype reduction was observed through this process. By linking this process and molecular detection of CTV, several genotypic strains may be identified for use in the cross-protection of citrus.

1.2.5. CTV strain differentiation

There are generally two different kinds of criteria that can be used to differentiate between viral strains:

1. Structural criteria based on the properties of the virion and its components.
2. Biological criteria based on the interaction between the virus, the vector and the host plant (Hull *et al.,* 2002).

In this study the term “strain” is used where an aspect of biology of the virus is known e.g. either a mild or severe appearance. Traditionally, the term “strain” has only been used for isolates that induce a specific symptom such as stem-pitting, seedling yellows, quick decline (classified as a severe strain) or vein clearing (classified as a mild strain) (Stewart, 2006). Since no pathogenicity determinants have been found in CTV, it is not yet possible to relate pathogenic characteristics to a specific gene region. The term “isolate” is used interchangeably with the term “source” and refers to a collected or subsequent derived, plant or sample that is infected with CTV.

To designate CTV isolates as strains using the criteria of Hull *et al.,* (2002), it becomes clear that strain differentiation is a highly complex procedure. CTV strains are structurally identical, compiled out of the same genes, are similar in size and shape and are composed of genomes varying between 19.2kb to 19.3kb However, when focusing on the biological characteristics of CTV isolates, it is clear that these isolates do not induce the same symptoms. Factors such as environmental conditions and the aphid vectors are also included in this strain separation analysis.

Rapid CTV strain differentiation tests are required to aid in studies on CTV epidemiology, development of more effective cross-protection schemes and to prevent any further propagation of severe CTV strains through budwood sources. Various techniques have been developed for CTV strain differentiation and are constantly being revised and improved to provide more accurate outcomes. It has become essential to target the whole genome when designing methods to accurately differentiate CTV strains. Currently a whole range of PCR primers are designed to target the four main CTV genotypes (T3, T30, T36 and VT) (Brlansky *et al.*, 2011). Subsequently more genotypes have been screened for after the four main genotypes.

1.2.5.1. Indicator plants.

Indicator plants (Table 2) have been predominantly used to differentiate CTV isolates by inoculating the plants with the CTV virus and observing the symptoms (Huang *et al.,* 2005). This procedure indexes the symptoms produced by the virus, in so providing the specific isolate with a biological designation (Barzegar *et al.,* 2010);

although, the symptoms attained will be specific to the cultivar used, the strain mixture used, and the environmental conditions. Mexican lime (*Citrus aurantifola*) is the most commonly used indicator plant due to its high sensitivity to CTV (Hilf *et al.,* 1999; Niblett *et al.,* 2000).

The universal way of bio-indexing is through graft-inoculations using buds or bark patches from candidate trees. Thereafter, the trees are maintained in controlled glasshouse conditions (18-25˚C) (Targon *et al.,* 2000). Symptoms such as vein clearing, leaf cupping and stem pitting (Ghorbel *et al.*, 2001) are found to be common, but will vary between the trees, depending on the severity of the CTV isolates (Al-sadi *et al.,* 2012). Bio-indexing is very time consuming (12-15 months), very expensive and is not specific enough as a single method in strain differentiation. Symptoms expressed in field settings are more variable than in glasshouse conditions; therefore, any symptoms detected in the glasshouse must be tested in field conditions to confirm strain differentiation results (Garnsey *et al.,* 2005; Nelson *et al.,* 2011).

Table 2: Indicator plants generally used for CTV strain differentiation (Barzegar *et al.*, 2010).

Indicator plant	Scientific name
Mexican lime	*C. aurantifolia* (Christm.) Swingle
C. hystrix DC, alemow	*C. macrophylla*, Wester
Duncan grapefruit	*C. paradisi* Macf.
Sour orange	*C. aurantium* L.
Troyer citrange	*P. trifoliata* x *C. sinensis*

Adding non-citrus hosts to the already used indicator host list, can aid in isolating mild and severe CTV strains. An experimental procedure performed by Müller *et al.* (1974) can not only assist in isolating mild and severe strains, but can also help by using these isolated CTV strains in the mild strain cross-protecting scheme. By passing severe or mixed CTV isolates through non-citrus species *Passiflora gracilis* and *P. caerulea*, by means of single aphid transmissions, a possible strain neutralization process will occur (Müller *et al.*, 1974), i.e. different CTV strains present may be reduced to only a few or one strain. The transmission of CTV into *Passiflora* sp. was initiated in 1980, where it was detected that *Passiflora gracilis* can survive CTV infections. *P. caerulea* was found to be an excellent holding

plant for the severe isolates, showing only small-leaf symptoms. The following diagram illustrates the technique used to neutralize severe CTV strains, by means of passing the virus through *Passiflora* species (Figure 5).

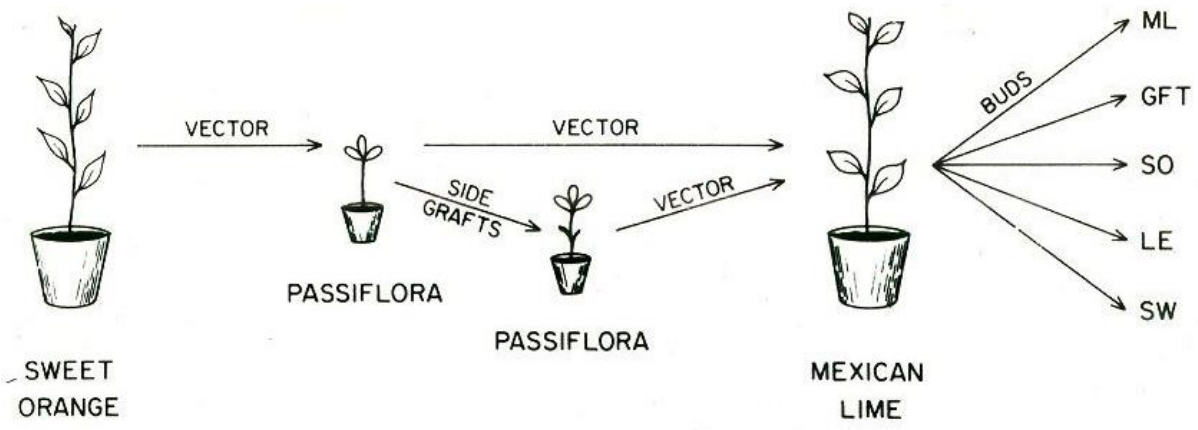


Figure 5: A diagram explaining the technique used to neutralize severe CTV strains using *Passiflora* species. The virus is vector transmitted from CTV infected sweet orange, by *Aphis gossypii* to *Passiflora caerulea* or *P. gracilis*. The virus can then be transmitted from *Passiflora* by vector to Mexican lime or by graft-transmission to other *Passiflora* species. From Mexican lime the virus is then transmitted by bud-graft to other index plants such as Mexican lime (ML), grapefruit (GFT), sour orange (SO), lemon (LE) and sweet orange (SW). In this way attenuation of CTV isolates was observed (Müller *et al.*, 1974).

1.2.5.2. Serological methods.

Serological methods using polyclonal and monoclonal antibodies have been valuable in detecting CTV. Monoclonal antibody (Mab), MCA13, was developed to differentiate between the different symptoms induced CTV isolates (Garnsey *et al.,* 2005). Serological methods used are limited to recognizing characteristics of the coat protein, which is often variable between different CTV strains (Saponari *et al.,* 2008). These methods alone however are not always sufficient enough to detect different variants of CTV (Iracheta-Cardenas *et al.,* 2009). Monoclonal and polyclonal antibodies used individually are not sufficient to provide reliable results, and a mixture of the two types of antibodies are normally used for rapid, economical and trustworthy results (Saponari *et al.,* 2008).

Several studies provided evidence that ELISA tests are amongst the least sensitive diagnostic procedures when compared to currently used DNA or RNA based assays such as real-time RT-PCR (Ruiz-Ruiz *et al*., 2007; Roy *et al*., 2010), reverse transcription PCR (Roy *et al*., 2010), conventional PCR, Single Strand Conformational Polymorphism (SSCP) (Rubio *et al*., 1996; van Vuuren *et al*., 2000), hybridization assays (Niblett *et al*., 2000), etc. (Bar-Joseph *et al.*, 1979; Bertolini & Moreno, 2008; Cervera *et al.*, 2010; Roy *et al.*, 2013). The CTV genotype specific antisera (Bar-Joseph *et al.* 1979) currently available and in use can also play a role in the sensitivity levels of this test. By creating more genotype specific antisera, ELISA analysis will become more species or genotype specific and therefore, aid in detecting unidentified pathogens during routine pathogen detection analysis. A study done by Dodds *et al.* (1987) provided evidence that different citrus hosts, temperature variances and the time of the year when samples are collected plays a role in the CTV concentration levels. Mathews *et al.* (1997) showed that CTV titres vary throughout the year and this can result in ELISA tests not detecting the virus consistently.

1.2.5.3. Molecular detection through Polymerase Chain Reactions.

Double-stranded RNA (dsRNA) patterns of CTV strains have been valuable in classifying differences in strains due to each specific nucleotide pattern (Stewart, 2006). However, the symptoms induced by CTV are not associated with given dsRNA patterns. For that reason, the Polymerase Chain Reaction (PCR) test was used to differentiate between different CTV strains. PCR analyses may be designed to specifically target variability within different isolates and to differentiate between genotypes (Yokomi *et al.,* 2010). Since the advent of PCR, numerous CTV isolates have been partially or fully sequenced, providing the necessary genetic information needed about the CTV genome (Palacio and Duran-Vila 1999; Harper *et al.,* 2010).

The development and use of multiplex RT-PCR's to detect, characterize and differentiate between CTV genotypes have been used in several studies (Adkar-purushothama *et al.*, 2010; Roy *et al.*, 2010). However, even though good results were obtained from these two studies, multiplex RT-PCR analysis can be extremely challenging. This analysis is well known for high contamination levels due to the

constant opening and closing of reaction tubes as this analysis requires several reagents, for a successful reaction. Also, reagent stocks can be easily contaminated throughout routine multiplex RT-PCR analysis. Accurate primer design is required to differentiate between the different genotypes or strains (Roy *et al.*, 2010), which can be extremely difficult especially within the 5”-half of the CTV genome. This can therefore also be very time consuming. Optimizations to multiplex RT-PCR analysis is a well known and continuous process to improve and reduce error introductions to this molecular detection technique (Adkar-purushothama *et al.* 2010).

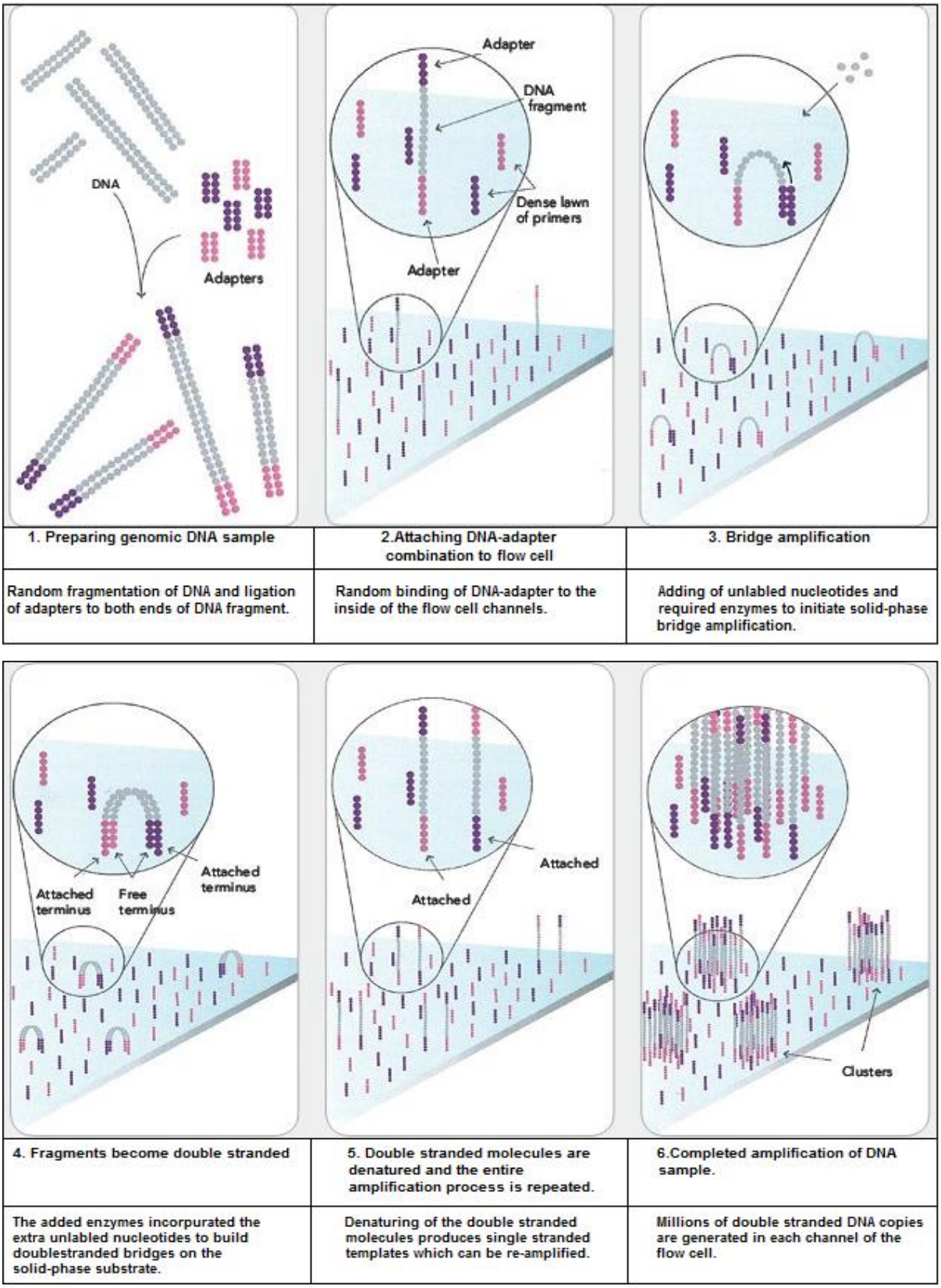
Real-time RT-PCR analysis has also been used for the detection of CTV (Saponari *et al.*, 2008). This analysis allows rapid detection, differentiation and quantification of the different CTV genotypes or strains (Ananthakrishnan *et al.*, 2010; Yokomi *et al.*, 2010; Saponari *et al.*, 2008; Ruiz-ruiz *et al.*, 2007). However, this analysis is extremely expensive and prone to contamination. Loading of samples into either capillaries or real-time plates can introduce contaminants. Cross-contamination between samples can occur if real-time analysis reactions are set up in a congested lab area where numerous PCR reactions are performed. Also, as with the multiplex RT-PCR analysis, accurate primer and now also probe design is required for high-quality end results (Ananthakrishnan *et al.* 2010). Saponari *et al.* (2008) performed an in-depth CTV screening study in which ELISA, conventional RT-PCR and real-time RT-PCR analysis was used. Citrus as well as aphid species were screened within this study. Of the three detection methods the real-time RT-PCR analysis was the most sensitive in detecting CTV, while ELISA was the least sensitive. However, all three detection techniques were able to detect CTV (Bertolini *et al.*, 2008).

1.2.5.4. Next-Generation high throughput Sequencing (NGS)

Next-generation high throughput sequencing, combined with bioinformatics, has been introduced and successfully used in plant virology since 2009 (Barba *et al.*, 2014). This new molecular approach has been used to determine the origin or cause of viral diseases, characterization of viral species, population genetics, taxonomy and diagnostic analysis on viruses (Voelkerding *et al.*, 2009; Metzker *et al.*, 2010; Wang, 2011; Fabre *et al.*, 2012; Knief, 2014; Neill *et al.*, 2014; Sebastien *et al.*, 2014), to provide a more clear answer to the genotypic homogeneity status for samples in question and many more applications. This new technology has developed a method in detecting and identifying organisms (pathogens) not only in plant material but also within insects (vectors) (Liu *et al.*, 2011).

When compared to Sanger sequencing, NGS produces millions of (Hall, 2007) sequence data much faster and provides more genomic or gene-related information (Barba, Czosnek, and Hadidi 2014). NGS has allowed complete genome sequencing of different organisms, metagenomics approaches to detect organisms present in collected samples and determining their role in the environment, to perform transcriptome sequencing, and sequencing of PCR amplicons in which several mutation detections can also be identified (Barba et al, 2014). Since the Illumina (Miseq platform) Next-Generation High-throughput sequencing platform was used within this research study, a more detailed explanation of the process is given (Figure 6).

This current electronic detection technology is based on three primary steps to produce high quality end results (Sebastien *et al.* 2014): firstly, the preparation of the specific DNA libraries for the amplification reaction; secondly, the amplification of these prepared libraries producing high amounts of DNA copies and then lastly, the massive parallel sequencing of the millions of DNA copies produced (Figure 6).



1. Preparing genomic DNA sample	2.Attaching DNA-adapter combination to flow cell	3. Bridge amplification
Random fragmentation of DNA and ligation of adapters to both ends of DNA fragment.	Random binding of DNA-adapter to the inside of the flow cell channels.	Adding of unlabled nucleotides and required enzymes to initiate solid-phase bridge amplification.

4. Fragments become double stranded	5. Double stranded molecules are denatured and the entire amplification process is repeated.	6.Completed amplification of DNA sample.
The added enzymes incorpurated the extra unlabled nucleotides to build doublestranded bridges on the solid-phase substrate.	Denaturing of the double stranded molecules produces single stranded templates which can be re-amplified.	Millions of double stranded DNA copies are generated in each channel of the flow cell.

Figure 6: Step-by-step explanation of Illumina Next-generation high throughput sequencing procedure, using bridge amplification to produce millions of double stranded DNA copies. Images: http://seqanswers.com/forums/images/content/ilmn-step1-6.jpg.

In a review article (Sebastien *et al.* 2014), several options for how NGS can be used in plant virus detection is provided. The first option was to extract DNA or RNA from a collected sample and sequencing it directly, making it a shotgun approach in detecting the targeted virus (Sebastien *et al.*, 2014). However, with this approach the plant material or other microflora-derived sequences will dominate in the data obtained. Thus, only a small portion of the sequences obtained will possibly contain the targeted sequence. If high quality sequences were obtained, this small pathogen-linked portion can be sufficient enough to confirm pathogen presence or even new viruses (Wylie and Jones, 2011). The second option discussed was by performing NGS on genomic material obtained from partial or complete virion purification techniques. This approach is said to be more challenging as counter-selection of viruses containing labile particles can occur. However, the genomic material obtained can contain sequences of enriched viral origin which can provide high numbers of specific sequence reads as an end product (Thapa *et al.*, 2012). The third option is by isolating and sequencing small interfering RNA's (siRNA). These intermediate molecules are produced by plant and animals during an antiviral defence reaction against pathogens (Barba etal., 2014; Albiach-marti, 2013, 2012; Moreno *et al.*, 2008). Due to the size (~22bp) of the RNA sequence produced, this option can also be used to detect viral presence within insect vectors (Liu *et al.*, 2011). This approach has the potential to be truly polyvalent and to allow detection of both DNA and RNA viruses as well as viroids (Itaya *et al.*, 2001), making it the most popular NGS technique by far (Sebastien *et al.*, 2014). However, this approach will be limited to insect viruses or viruses replicating or circulating in plants, as siRNA will only be produced when the virus replicate within the insect or plant. Alternatively, dsRNA, with high enough concentration values, can also be used. Lastly, the use of primer sets targeting a specific region within the genome can also be implemented as a NGS option. Primer specific RT-PCR amplicons will be produced prior to the sequencing reaction in which the structure of a specific viral population can be analyzed (Fabre *et al.*, 2012). With this type of approach one needs to make sure that the primers used have a broad polyvalence so that new viral species or in this case CTV genotypes (or genotypic strains) can be identified (James *et al.*, 2006).

A study performed by Zablocki and Pietersen (2014), made use of a random amplification (Roossinck *et al.* 2010) approach using full length dsRNA to detect CTV genotypes present in three GFMS 12 sub-isolates (12-7, 12-8 and 12-9). A novel CTV genotype (CT-ZA3) (KC333868.1) was detected through this NGS diagnostic approach. Double-stranded RNA has emerged as perhaps the most effective and efficient template for virus detection or discovery using NGS libraries.

However viruses in general have high mutation rates (especially RNA viruses), short generation times and large population sizes (Duffy *et al.*, 2008; Li *et al.*, 2012) and these factors can play a role in producing a high genetic diversity or even viral quasispecies.

Little is known regarding the sensitivity of NGS. Numerous factors can play a role in the sensitivity and quality of data obtained with NGS, namely; the availability of complete genomes or sequence information on the virus screened for, sample concentrations and amounts, different library preparations, different sequencing platforms and potential bias and contamination (Neill *et al.*, 2014; Sebastien *et al.*, 2014; Ross *et al.*, 2013; Solonenko *et al.*, 2013; Duhaime *et al.*, 2012; Linnarsson, 2010).

Ross *et al.* (2013) focused on characterizing and measuring biases within sequence data obtained during high-throughput sequencing, in which they found that biases are most commonly identified in particularly high- or low GC regions (Solonenko *et al.* 2013) within the genome as well as during long homopolymer runs. An in-depth discussion of bias (coverage bias) and comparison of different sequencing platforms was done within this publication, while exploring potential ways to decrease this problem. Other than the introduction of bias through sample or library construction contamination, inaccurate alignments as well as errors during the amplification and sequencing reactions (polymerase slippage) can also occur (Ross *et al.*, 2013; Solonenko *et al.*, 2013; Knief, 2014).

The type of sequencing platform, library preparations, amplification reactions, as well as the concentration of the input sample DNA, can also play a big role in preventing bias. Numerous factors can aid in deciding which platform to use, namely; (1) the size of the genome or gene studied, (2) the sequence complexity (e.g. GC

content) and (3) the depth of coverage and accuracy required to obtain high quality results (Barba *et al.*, 2014). Accuracy and skilled pipetting techniques plays a big role in library preparations as sample contamination must be prevented (Neill *et al.*, 2014). Also, the type of library preparation one will use depends on the type of sequencing platform used (Liu *et al.*, 2011). The standardized procedure for sample library preparation includes DNA or RNA sequence fragmentation, adding of adapters, RT-PCR or PCR amplification (Liu *et al.*, 2011). Since the NGS analysis still makes use of one or more PCR steps, contamination during the library preparation of the samples as well as the analysis process itself may still occur (Ross *et al.* 2013; Sebastien *et al.* 2014).

The amplification conditions, as well as the starting DNA concentrations play a role in the success of the library preparation (Solonenko *et al.* 2013). It is shown that the library amplification step is known to increase the levels of bias as the input DNA amounts are increased to balance out the losses during library preparation (Linnarsson, 2010). According to Duhaime *et al.*, 2012, viruses generally yield a DNA concentration of <1ng. In this study the authors made use of (*Pseudoalteromonas* phage H105/1: clonal virus lysate and environmental virus samples to explain the above statement (Duhaime *et al.*, 2012).

Therefore, it is extremely important to use the correct ratio of the adaptor used and the template DNA during the ligation step (Sambrook *et al.*, 1989). It was found that high DNA yields (1000ng) provides better overall sequencing results than low DNA amounts (100ng and less) (Solonenko *et al.* 2013). From the experiments performed by Solonenko *et al.* (2013), they found that unamplified DNA libraries are best for high DNA yields (e.g. >2µg) during the sequencing reaction. For low yield DNA (e.g. 1 – 100ng) libraries, a linker-amplified protocol (Duhaime *et al.*, 2012) can be used together with a specifically chosen template: adaptor ratio) and optimized sequencing technology. Adaptor-mediated amplification reactions has also been used on low DNA amounts, however, this step can alter the overall GC% of the DNA sequence (Aird *et al.*, 2011; Duhaime *et al.*, 2012; Oyola *et al.*, 2012).

Analysis of the NGS data can be divided into four basic steps, namely; (1) alignment of the sequence reads obtained, (2) base-calling and/or polymorphism detection, (3) *de novo* assembly of filtered sequence reads (Morey *et al.*, 2013) and

(4) genome browsing and annotation of the final results (Barba *et al.*, 2014). The analysis of sequence data obtained can be greatly simplified if complete sequences of reference genomes for that specific organism are available. The use of Basic Local Alignment Search Tool (BLAST) searched against reference sequences within National Center of Biotechnology Information (NCBI) databases (Bexfield and Kellam, 2011; Altschul *et al.*, 1990) prior to sequence assembly can also aid in the analysis procedure, as the sample sequences that meet the E-value cut-off will be selected for *de novo* assembly or reference mapping (Shendure *et al.*, 2008). *De novo* assembly has been considered the "golden standard" in assembling unknown sequences. This approach is also the best way to identify novel viral agents when no reference genomes are available (Sebastien *et al.*, 2014). However, NGS data interpretation, reproducibility, and data accessibility is still an ongoing learning curve for all scientists. Nekrutenko and Taylor (2012) made use of these unfamiliar aspects and published an article focusing on how one can approach each of these aspects and also, to improve and make data analysis easier.

However, there are some limitations linked with this NGS detection procedure. Detection of novel viruses is difficult if there is no sequence homology to known viruses. Secondly, if the virus is present in low titres within the host it is unlikely that the entire genome will be sequenced. Thirdly, since NGS is still a very new molecular detection method, no established parameters have been set and therefore non-standardized methods are currently used during data analysis (Liu *et al.*, 2011). It is still unknown what amount of total reads obtained from test samples can be considered sufficient confidently consider a sample positive or more importantly, negative for the specific pathogen (Sebastien *et al.* 2014). However, the sequence data coverage can also play a role in the results obtained (Barba *et al.*, 2014; Ross *et al.*, 2013). Too little sequence covered can easily give a false positive or negative result.

While there is to date, no other technique that can provide such broad-spectrum diagnostics (Sebastien *et al.* 2014), there are still numerous issues that will have to be addressed and solved before this procedure can fully l accepted in the plant virus diagnosis. As more partial or complete genome sequences become available, more broad-specificity primers can be designed and used during NGS to

detect specific pathogens, viruses or in our case CTV genotypes with more confidence in the results. Also, the use of this detection technology within plant certification and quarantine programs, such as the CTV cross-protection scheme, can aid in improving pathogen protection programs in which earlier pathogen detection and characterization will help in controlling the diseases being spread (Barba *et al.*, 2014).

1.2.5.5. Single strand conformational polymorphism analysis.

Single Strand Conformational Polymorphism (SSCP) analysis has also been used to detect single base pair mutations; however, differences in sequence patterns do not necessarily mean great differences in nucleotide sequences (Rubio *et al.*, 1996). Nucleic acid hybridization reactions will allow severe strain detections using specifically designed probes.

1.2.6. Studies of CTV variation.

Numerous CTV strains exist that are classified based on distinct biological and genetic characteristics (Folimonova, 2013; Roy *et al.*, 2013). The diverse range of symptoms expressed by CTV strains is indicative of the high variability within and between populations. Plant viruses are naturally heterogeneous (Garcia-Arenal *et al.,* 2001) and as such CTV hosts contain a “quasispecies” population, which most likely emerged due to the error-prone properties of the RNA-dependent RNA polymerase (Iglesias *et al.,* 2008). The dispersal of infected tissue within and between environmental regions and recombination events also plays a role in shaping CTV diversity (Moreno *et al.,* 2008). Therefore, numerous studies are focusing on CTV populations to determine the population structures of CTV within growing regions (Fisher *et al.*, 2010; Iglesias *et al.,* 2008; Rubio *et al.,* 2001). The information obtained can be used to determine population complexity as well as which genotypes are present within different regions. This knowledge will have an important influence on CTV control measures worldwide (Kong *et al.,* 2000).

Phylogenetic analysis of the 5'-UTR sequences (Ayllon *et al.,* 2001) or the p23 gene region (Sambade *et al.,* 2003) have been used to study CTV in general as well as population (genotype) diversity (Read and Pietersen, 2015; Melzer *et al.*, 2010; Gowda *et al.*, 2003). However, 3'-end coding regions have also been used in characterizing CTV populations (Iglesias *et al.,* 2008). Analysing the 5'-UTR gRNA sequences of CTV (isolates) populations revealed that the 5'-UTR is highly variable with as little as 44% sequence similarity between strains, while the 3'-UTR is highly conserved with sequence similarities of 97% and more (Ayllon *et al.,* 2001).

During a study done by Hilf *et al*. (1999), comparisons of the VT and T36 genomes with T3 and T30 complimentary sequence revealed an asymmetrical sequence identity distribution (Silva *et al.*, 2012; Hilf *et al.,* 1999) with a conserved level sequence divergence in the last 8kb in the 3'-end with an abrupt reduction of sequence conservation in the first 11kb of the 5'-end. Hilf *et al.* (1999) further suggested that the VT and T36 genotypes represent two major "lineages" within the CTV complex. In this study they showed that the T3 and T30 genotypes displayed a high degree of sequence similarity with VT in both the 5' and 3' regions and therefore, placed these three genotypes into a single group (Roy *et al.*, 2005).

Hilf *et al.* (1999) have proposed a "basic CTV classification scheme" where T3, T30, T36 and VT each represent specific "genotypes". In a follow up study by Hilf *et al.* (2005), a phylogenetic approach, using four genetic molecular markers supported the assignment of the four different genotypes. The main goal was to develop a simple procedure to classify CTV isolates and to establish more structure for current CTV classification. Roy and Brlansky (2010) continued with the CTV classification research by using multiple molecular marker (MMM) primers, where using nucleotide sequences from the ORF2 gene region, they discovered that the B165 (EU076703) isolate is a novel genotype that occupies a separate branch in a phylogenetic tree that was generated. These authors suggested that CTV isolates can therefore be grouped as five different genotypes, namely T3, T30, T36, VT (Hilf *et al.,* 2005) and B165 (Roy and Brlansky, 2010). Full genome sequences of these genotypes are mainly used as reference strains during 5'-UTR sequence analyses and aligning sequences to produce phylogenetic trees.

A review article written by Moreno *et al.* (2008), a section on genetic variation and CTV strain characterization is provided along with a Neighbor-Joining dendrogram (Figure 7) representing the full genome sequences and group classifications. Strain T318A (DQ151548), SY568 (AF001623), NUagA (AB046398) and the VT group (U56902) all fall within sequence type I (VT genotype, severe stem-pitting isolates). The T30 complete genome (AF260651) and CTV isolate T385 (Y18420) falls within sequence type II (T30 genotype, mild isolates). Lastly, the CTV complete genomes for T36 (U16304), Qaha from Egypt (AY340974) and the Mexican CTV isolate (DQ272579) grouped together as sequence type III (T36 genotype, intermediate characteristics). As mentioned above, Roy and Brlansky (2010) verified that the B165 isolate groups as a separate group on a phylogenetic tree and is now classified as a sequence type IV. However, apart from the 5'-UTR analyses, phylogenetic analysis of the p23 gene region showed the same isolates grouping together. Since the complete sequence of the T3 CTV isolate was not available at that point in time, the isolate was not included within the study of Roy and Brlansky (2010).

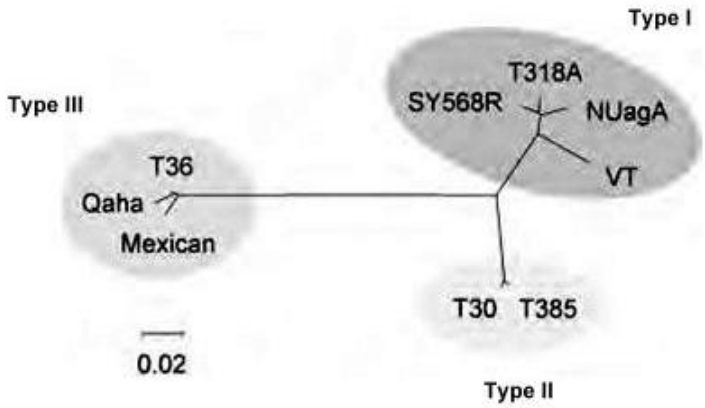


Figure 7: Grouping of the three genotypes obtained from 5'-UTR sequence analysis (Sequence type I, II and III). The branch lengths indicate genetic distance. The then newly discovered sequence type IV of B165 groups between the T30 and T36 groups (Moreno *et al.*, 2008). The B165 forms a unique grouping between T30 and T36.

In earlier years of CTV research, the RB genotype had not yet been identified and primer sets produced (Hilf and Garnsey, 2000; Hilf *et al.*, 2005; Roy *et al.*, 2005) were designed to detect the presence of the CTV genotypes known at the time. However, during a study by Roy *et al.* (2010) the authors determined that the T36

primer set designed not only picked up the T36 genotype, but also detected the up until then unknown RB genotype. The authors confirmed that the T36 primer set used within this study will detect both NZRB-M17 (FJ525435) and NZRB-TH30 (Harper *et al.*, 2010). However, the other three RB genotypic strains (NZRB-M12 (FJ525431), NZRB-G90 (FJ525432) and NZRB-TH28 (FJ525433)) together with the two then recently sequenced Hawaiian isolates (HA18-9 (GQ454869) and HA16-5 (GQ454870)) should not be picked up by this T36 primer set (Roy *et al.*, 2010). Subsequently it has been determined through phylogenetic analysis (Figure 8) that all five these RB strains are T36-like within the ORF1a region of the CTV genome. A detailed discussion of the relatedness of the RB strains with each other and the T36 CTV genotype was published by Harper *et al.* (2010). Even though these 5 RB isolates cluster into one monophyletic group, two distinct RB clades were detected. RB strains NZRB-M12, NZRB-G90 and NZRB- TH28 grouped together with a sequence similarity of 94.7% (referred to as RB Clade 1 in this dissertation), while NZRB-TH30 and NZRB-M17 formed the second clade, in which they share a sequence identity of 98.4% (referred to a RB Clade 2 in this dissertation). However, an overall sequence similarity of 96% is shared between these 5 RB strains (Harper *et al.*, 2010). Furthermore, it was also found that these 5 RB strains share a 90.3% sequence identity with the T36 CTV genotype, therefore, explaining the previously indiscriminate detection of these two genotypic groups. The close percentage of sequence identities between the 5 RB isolates and the T36 genotype can therefore specify the possible poly-specificity of the T36 primer set as well as the T36 RT-PCR analysis used by Roy *et al.* (2010).

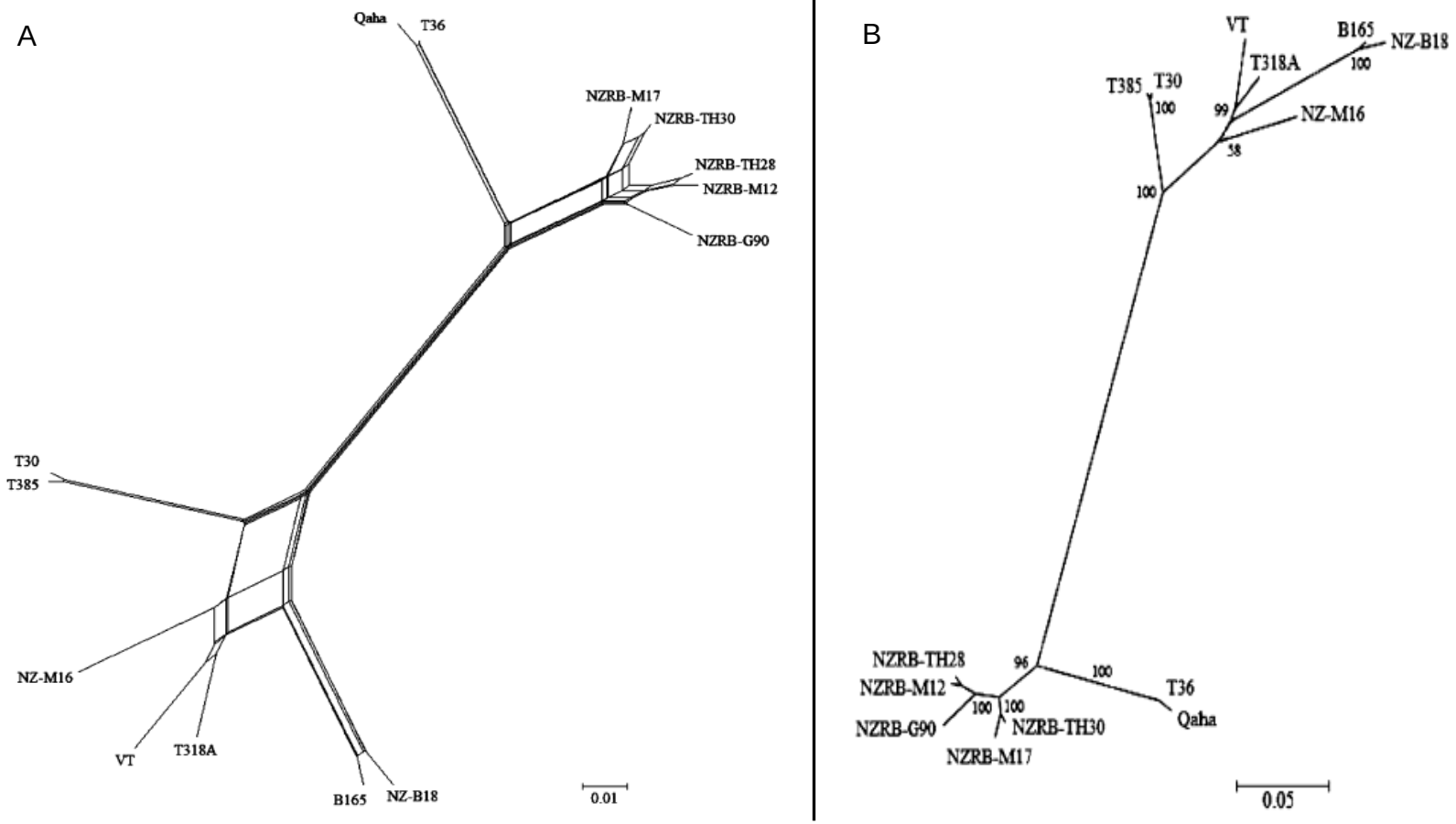


Figure 8: [A] Neighbor-joining analysis of the complete genomes of the RB isolates from New Zealand and currently available complete CTV genomes. **[B]** Maximum-likelihood dendrogram of the complete genomes of the RB isolates from New Zealand and currently available complete CTV genomes (Harper *et al.*, 2010).

In contrast to the T36 isolate, the RB isolates are fairly different from CTV genotypes T30 and T385, showing an average similarity of 82.2% (Harper *et al.*, 2010). Furthermore, a sequence similarity index determined with the more severe CTV isolates were as follow; T318A (81.2%), VT (80.2%), NZ-M16 (80.8%) and NZ-B18/B165 (80.8 and 80.5%, respectively) (Harper *et al.*, 2010). Furthermore, two CTV isolates (NZ-M16 and NZ-B18) from New Zealand were isolated, examined and fully sequenced by Harper *et al.* (2009). Each CTV isolate showed very distinct phenotypic characteristics; as isolate NZ-M16 is not transmittable through the aphid vector *T. citricida* and shows no symptoms, while isolate NZ-B18 is highly transmittable and present severe symptoms upon infection especially on *C. sinensis* and *C. aurantii* (Harper *et al.*, 2009). Nevertheless, both isolates shared a sequence similarity between 90 – 93% with the CTV isolates VT and T318. However, phylogenetic analysis between the two isolates showed that they only share an 89% sequence identity. Based on the sequence similarities, both of the CTV isolates are VT subtypes, with NZ-M16 being more T3-like and NZ-B18 is forming part of the B165 subtype from India (Harper *et al.*, 2009). Furthermore, another CTV isolate, B301, was isolated from Puerto Rico and confirmed as another RB isolate since it shares a sequence identity of 98% with the mentioned RB isolates (Roy *et al.*, 2013).

When analyzing individual ORFs within the CTV genome, a variable pattern of relatedness was found between the 5 mentioned RB isolates and other CTV isolates. ORF1a has shown to be more T36-like, while the rest of the CTV genome (ORF1b - p23) was found to show a lot of variation between the different CTV isolates present (Harper *et al.*, 2010; A. Roy *et al.*, 2010). Furthermore, Harper *et al.* (2010) discussed that several potential recombination events between different CTV isolates has already occurred or are taking place in a continuous manner. Evidence of multiple recombination events throughout the evolutionary history of these 5 RB isolates with each other as well as the CTV T36, T30 and VT-like isolates were detected and discussed within this article.

Melzer *et al.* (2010) published a CTV genotype differentiation study in which they found evidence of two CTV strains (HA16-5 and HA18-9) present in Hawaii. Within this study, extensive analysis was performed on these two strains to determine their relatedness with each other and whether a potential new CTV

genotype was detected. It was found that the sequence identity of HA18-9 was very similar (94.3 to 95.5%) to the five NZRB strains (NZRB-M12, -M17, -TH30, -TH28 and –G90). From the information gathered for HA16-5, it looked like this strain arose through a recombination reaction with the 3'-end of HA18-9 CTV strain. However, the 5'-end of the genome sequence of HA16-5 differed greatly from strain HA18-9 and had a relatively low sequence similarity to any other characterized CTV strains or genotypes. Therefore, strain HA16-5 was classified as a new CTV genotype (Melzer *et al.* 2010). A neighbor-Joining phylogenetic tree was drawn using complete genome sequences with the then available CTV genotypes and the two "new" strains, HA16-5 and HA18-9 (Figure 9).

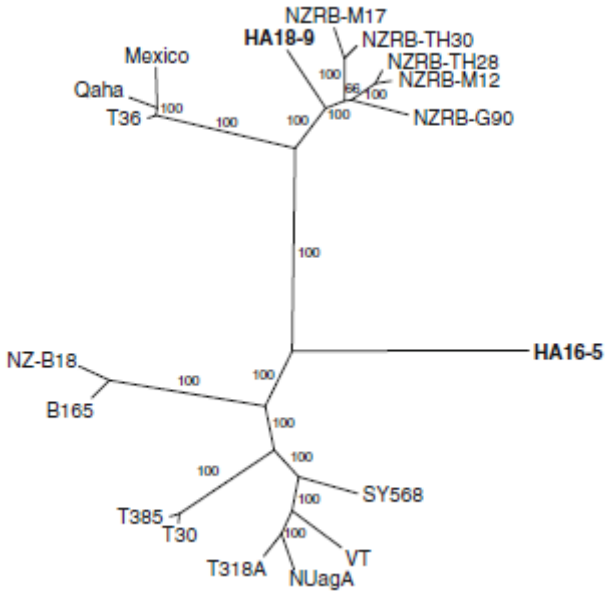


Figure 9: Neighbor-joining dendrogram of the currently available complete genomes of the CTV genotypes present, including the two CTV strains HA16-5 and HA18-9 (Melzer *et al.*, 2010).

A research study by Read and Pietersen (2015) was focused on the CTV genotype diversity within grapefruit cultivars in South African orchards. Two parts of the genome, the 5'- terminal half (A region) and the 3'-terminal half (p23 gene), were used for the analysis. Neighbor-Joining dendrograms (Figure 10 and Figure 11) were produced for each of these regions in which distinct CTV genotype separation can be seen (Read and Pietersen, 2015). These dendrograms produced are illustrating the currently available CTV genotypes being classified into eight major CTV genotypic groups namely; T3, T30, T36, VT, B165, HA16-5, RB1 and RB2 (Read and Pietersen, 2015; Matos *et al.*, 2013).

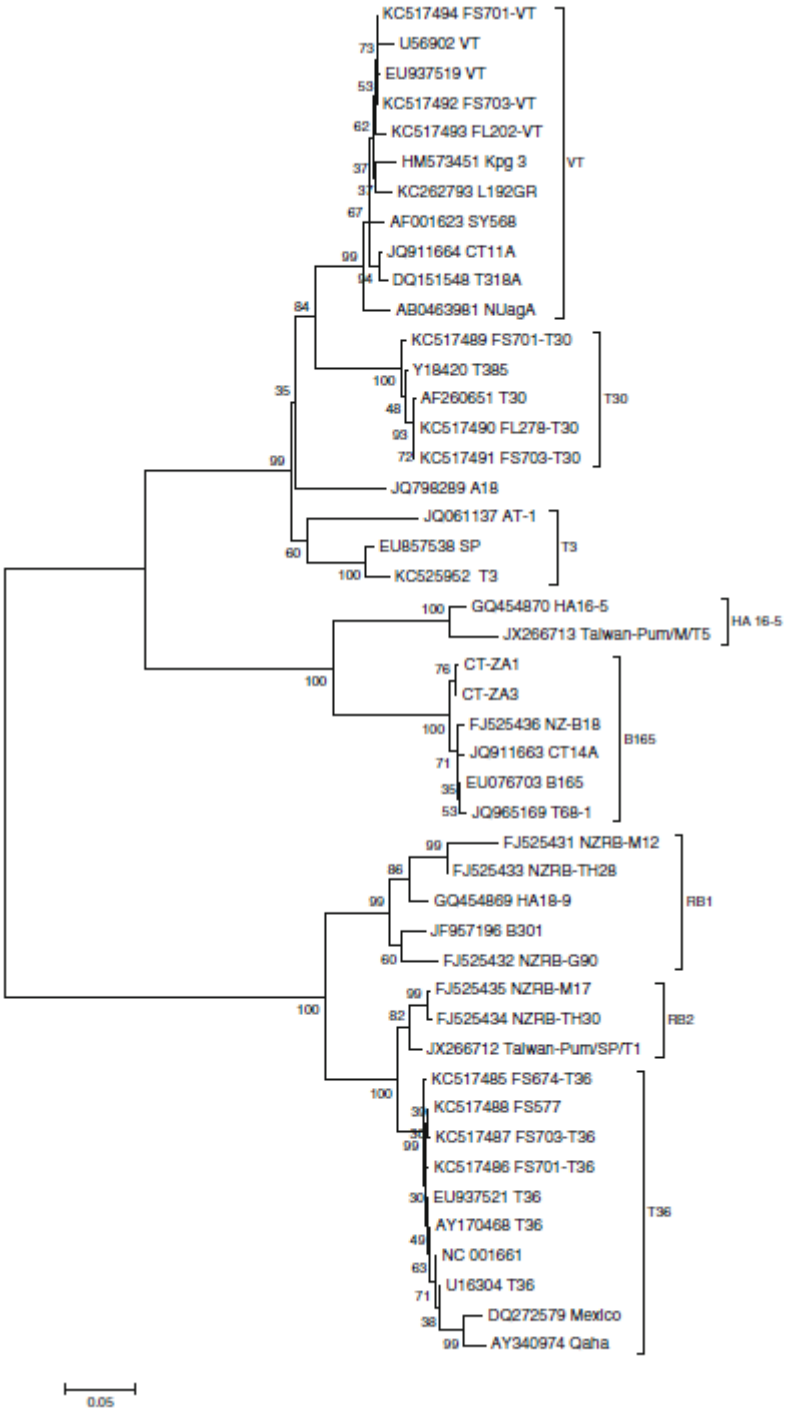


Figure 10: Neighbor-joining dendrogram based on the sequences of the A-region derived from full genome sequences of CTV, obtained from GenBank. The respective CTV genotypic groups are indicated by brackets and their corresponding labels (Read and Pietersen, 2015).

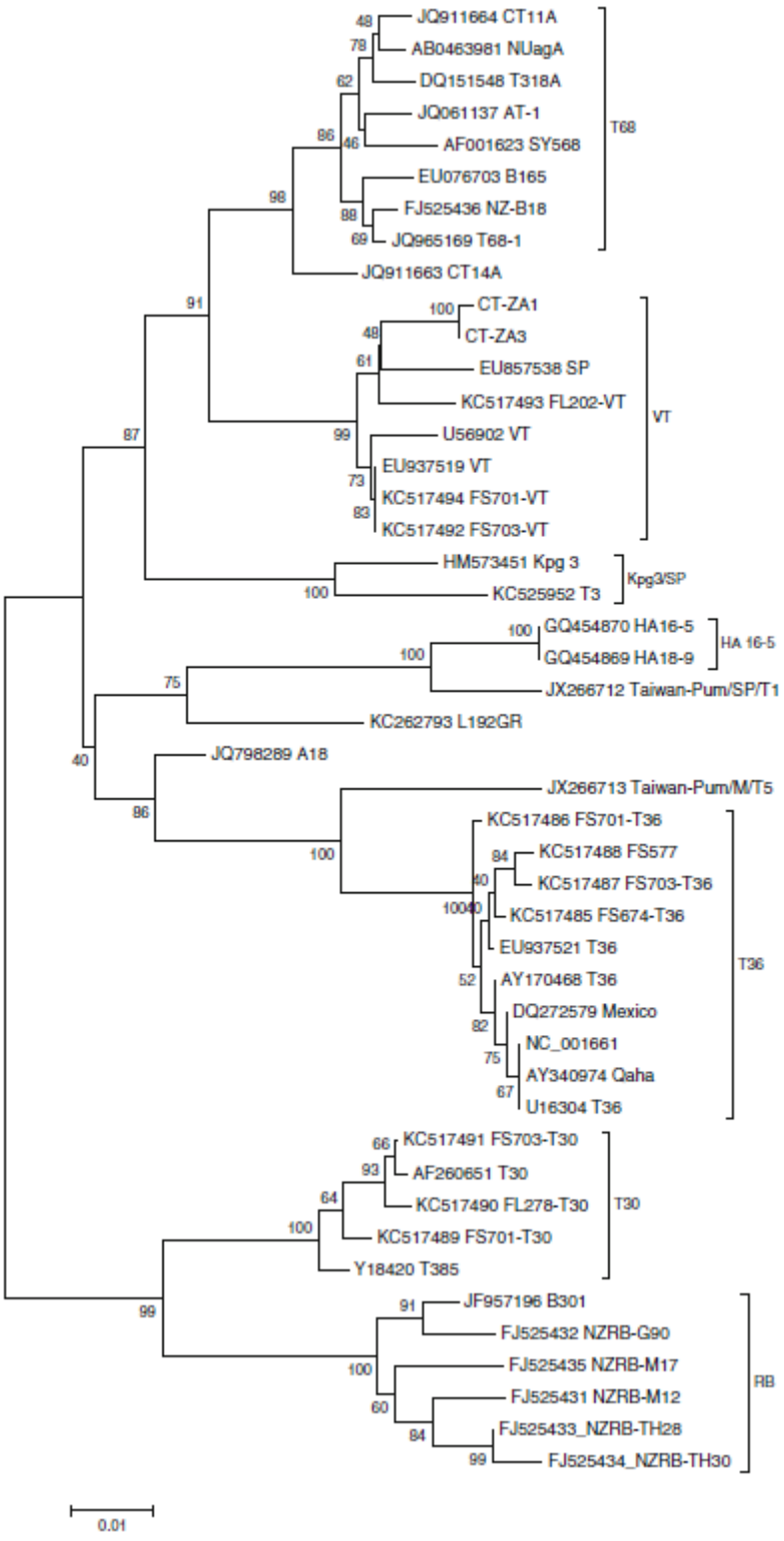


Figure 11: Neighbor-joining dendrogram based on the sequences of the p23 gene derived from full genome sequences of CTV, obtained from GenBank. The respective CTV genotypic groups are indicated by brackets and their corresponding labels (Read and Pietersen, 2015).

1.2.7. Control of CTV

The control of CTV remains a continuous challenge worldwide. General strategies in controlling CTV transmission depends significantly on the incidence and severity of the viral strains present, the rootstocks (D'Onghia, 2009; Grosser *et al.*, 2004) used and the citrus varieties present in the specific region (Moreno *et al.,* 2008).

1.2.7.1. *Quarantine and budwood certification programs*

Preventing a virus from being imported into a specific virus-free area means the virus will not be able to establish and no programs to prevent spread within that area is needed. Therefore, exporting countries must incorporate strict regulations regarding the movement of produce. A quarantine program (Yokomi, 2009) and scanning for pathogens, is one approach in preventing any introductions of viruses or diseases into countries, for example with citrus producing industries affected by CTV. Many epidemics can be prevented if certification programs, such as rapid indexing tests, are implemented and maintained at each checkpoint during product export.

1.2.7.2. *Elimination of infected trees*

Severely infected tree removal and vector control to prevent the spread of CTV is the main approach in field settings where *T. citricida* and/or severe CTV isolates are present. ,However, eradication of trees that already showed symptoms is rarely effective, as the virus has often already established itself in a number of symptomless infections especially if the environmental conditions are favourable for the aphid vectors (Abbas *et al.,* 2005).

1.2.7.3. *Biological control*

There are numerous natural enemies attacking aphids worldwide. The most common aphid predators include lady beetles (*Coelophora inaequalis* (Fab.)), *Cycloneda sanguine sanguine* (L.), parasitic wasps (*Lysiphlebus testaceipes* (Cresson.)), *Aphelinus gossypii* (Timberlake) (Tsai *et al.,* 2009). However, natural aphid predators seldom prevent CTV transmission, due to the active reproduction

cycles of aphids and their ability to recolonize nearby citrus trees (viruliferous winged aphids).

1.2.7.4. *Chemical control*

In Florida, two systemic insecticides Admire 2F (imidacloprid) and Temik (aldicarb) are currently registered for use on citrus. Chemical control of the insect vectors in citrus orchards throughout the year, especially prior to spring time, to aid in maintaining a good quality of produce in the new season (Perring *et al.*, 1999; Michaud, 1998). Insecticidal treatments can therefore also help in reducing the secondary spread of aphids in the orchard (Yokomi, 2009). Also, some citrus trees are more sensitive to chemicals than other; therefore, if it is decided to apply insecticides, a proper study will have to be done to see if chemical control can be applied to these citrus orchards.

1.2.7.5. *CTV resistant plants by means of genetic engineering*

Poncirus trifoliate (L.) and *P. trifoliate* hybrids, citranges and citrumelos, are immune to most CTV isolates (Harper *et al.,* 2010). Studies done in Florida and California attempted to map the resistance gene of *P. trifoliate* and insert it in citrus cultivars; thus, producing CTV immune citrus. Genetic engineering has been attempted various times to protect citrus from CTV. An experimental study was done incorporating a pathogen-derived resistance (PDR) gene through plant transformations, of CP and p23 or the 3'-NTR, producing transgenic plants (López *et al.,* 2010; Cervera *et al.,* 2010; Fagoaga *et al.,* 2006). The p23 protein will most likely act as a suppressor to RNA silencing. Variable results have been obtained from this type of CTV control; however, according to the numerous authors transgenic plants can be the future protection outline against CTV (Karasev *et al.,* 1995). Additionally, the p23 protein is also a RNA-binding protein which plays a role in controlling the asymmetrical accumulation of the positive and negative RNA during viral replication (Albiach-marti, 2013; Sambade *et al.*, 2003; Satyanarayana *et al.,* 2002). Introducing changes, such as mutations (point mutations), to the p23 protein sequence can not only assist in RNA silencing but also aid in decreasing the RNA amount during viral replication within the tree. In overall, introducing modifications to gene or protein

sequences can assist in future attempts to produce CTV-tolerant or resistant citrus trees.

1.2.7.6. *Tristeza-tolerant rootstocks*

When tree eradication is not possible or ineffective, an additional option of citrus propagation of CTV-tolerant rootstocks can be done. *Tristeza*-tolerant rootstocks have been used to grow citrus successfully following severe outbreaks (Roistacher *et al.,* 2010) and has allowed citrus industries to recover in places like Argentina. *Poncirus trifoliate* and its hybrids Carrizo and Troyer citrange (sweet orange x *Poncirus trifoliate*), Swingle citrumelo (grapefruit x *Poncirus trifoliate*) and Rangpur lime (*C. limonia* Osb.) are among the widely used CTV-tolerant rootstocks (Moreno *et al.,* 2008). Alternative CTV tolerant rootstocks to the above mentioned are Rough lemon, Volkamer lemon, Rangpur lime and Cleopatra mandarin (Roistacher and Moreno 2007; Roistacher, 2004). Sour orange has effectively been replaced as primary rootstock by these afore mentioned CTV tolerant rootstocks. However, alternative routes in producing somatic hybrids of mandarin and Pomelo, (to replace sour orange rootstock) are looked at, but future input is still required (van Vuuren, 2002; Grosser *et al.,* 2004). However, together with the struggle in finding an alternative CTV-tolerant rootstock, other pathogens and diseases such as nematodes, viroids and citrus blight are additional problems which will also assist in the decrease of resistance to CTV (Yokomi, 2009).

1.2.7.7. *Mild strain cross-protection (MSCP) programs*

One of the most effective approaches to reduce citrus crop losses, where severe CTV strains are prevalent and aphid populations are copious, is by cross-protection programs (D'Onghia, 2009). Cross-protection, also known as superinfection exclusion, is the process in which citrus cultivars are potentially protected by a primary viral infection of mild CTV strains in order to prevent secondary infections by severe CTV stains of the same or similar viral strain/genotype to take place (Folimonova *et al.*, 2010). MSCP has been largely successful in citrus production and has allowed citrus producers to continue to

produce citrus economical. The strain sources used in cross-protection programs are usually from citrus trees that showed little or no symptoms in areas where CTV is a serious problem. CTV cross-protection programs has been widely implemented on millions of citrus trees in Australia, South Africa and Brazil to protect against economic losses caused by stem pitting (SP) and quick decline (QD) (Albiach-marti, 2012; Abbas *et al.*, 2005).

In a landmark publication, Folimonova *et al.* (2010) determined the relationships between different CTV genotypes in terms of how they will aid in cross-protection of citrus cultivars. They showed that none of the CTV isolates from T30, VT, T68, or T3, was able to exclude superinfection by the GFP-tagged T36 source, but that T36 isolates were able. Furthermore they showed that even though phloem cells contains high amounts of the primary CTV source, superinfection will still occur if the primary infection strain and secondary challenging strain differs from each other. With this experiment the authors concluded that cross-protection is (viral) strain specific.

The strain specificity linked to cross-protection was further confirmed when Folimonova *et al.* (2010) attempted a further analysis in which they made use of a monoclonal antibody (MCA13) to monitor whether CTV infections could be prevented for both the primary (T30-1 and CTV9R-MCA13NR) and challenging isolates (T68-1, T3, FL202-1 and T36). Again the only superinfection exclusion event occurred between CTV9R-MCA13NR (primary isolate) and T36 (challenging isolate), since this primary isolate represent a virus of the same T36 strain (Folimonova *et al.*, 2010). A last attempt for superinfection exclusion was through CTV hybrids (T36/T68 and T36/T30), in which results also showed that even though homology was shared within the 3'-region of the genomes, superinfection exclusion was still not triggered between the two viruses used.

In several areas around the world, such as South Africa, Australia, Brazil, Reunion, India and Japan, mild CTV strains have been isolated from citrus trees that showed very little to no symptoms in the presence of endogenous severe CTV strains. Given that the mild strain cross-protection program is strain specific (Folimonova *et al.,* 2010) and that CTV sources used during this program may contain mixtures of different strains (Roistacher *et al.*, 2010); total protection against

new or severe CTV strains introduced into the area might not be possible (Yokomi, 2009; van Vuuren *et al.*, 1993). In addition to the above mentioned, recombination reactions between the different CTV strains will lead the production of new hybrid strains, where some hybrids produced can be more severe than the parental strains, in which cross-protection of our citrus orchards can become very unpredictable (Cambra *et al.*, 2000; Rubio *et al.*, 2001; Rubio *et al.*, 2013). Environmental conditions such as extreme temperature changes and seasonal changes may also reduce the protection efficiency (Stewart, 2006; Meyer *et al.*, 2005; van der Vyver *et al.*, 2002; Grisoni and Riviere, 1993).

Cross-protection breakdown is one of the main concerns when focusing on CTV protection. The reason for this “breakdown” reaction is not fully understood, but the following features have been suggested as possible reasons: super-infections (Folimonova *et al.*, 2010; Pelosi *et al.*, 2000; Rubio *et al.*, 2001; van der Vyver *et al.*, 2002), strains separation, strain dominance, recombination reactions between strains, an increase of viral replication (increased viral titre) (Abbas *et al.,* 2005), the cultivars infected and the effect of mix CTV infections (Zhou *et al.,* 2002).

1.2.8. Cross-protection scheme in South Africa.

The South African Citrus Improvement Program (SACIP) was initiated in 1973 (von Broembsen *et al.,* 1978). Since then, the SACIP has provided certified and cross-protected plants containing mixtures of mild CTV strains to protect against severe CTV strains (von Broembsen *et al.,* 1978). Following the initiation of the MSCP scheme, certification schemes were started during 1978 due to citrus tree deaths caused by CTV stem-pitting and other diseases (von Broembsen *et al.,* 1978). In a review done on CTV cross protection by da Graca and van Vuuren (2010), they mention that South Africa has produced competent cross-protection for all the budwood sources used. However, in addition to the citrus tree losses, another motivation to make use of both mentioned schemes was the then used primary rootstock “rough lemon”. Although rough lemon was tolerant to most graft-transmissible pathogens, the fruit produce from this rootstock were not of good

quality and it was unable to be used in replanting situations due to its high susceptibility to root pathogens (Roistacher *et al.,* 2010).

Two mild stain sources (consisting out of multiple strains each) have been used to cross-protect grapefruit in South Africa. These are GFMS12 (Grapefruit Mild Strains) and GFMS35 (van Vuuren *et al.,* 1993). Isolate LMS6 (Lime Mild Strains) (van Vuuren *et al.,* 1993) contains a mild form of seedling yellows and is mainly used to cross-protect sweet orange, mandarins and lime trees. These CTV mild sources were initially screened for symptom development on Mexican lime and Marsh grapefruit in greenhouse conditions (van Vuuren *et al.,* 1993). Thereafter, field trials were done where the results showed that these three sources provided good protection against CTV for years in Marsh, Nel Ruby and Star Ruby grapefruit as well as lime trees (van der Vyver *et al.,* 2002; van Vuuren *et al.,* 1993, 2000).

Since virus free control plants also became naturally infected in the field, but did not produce the expected small fruit, it was thought that the problem could be due to strain shifts in certain grapefruit cultivar-mild CTV source combinations (van Vuuren *et al.,* 2000). From these findings it became clear that the interactions between the citrus host and the CTV strain, infecting the host, are very specific.

A study done by van Vuuren *et al.* (2000) partially characterized the GFMS12 isolate by analysing the sub-isolates (12-1 to 12-9) that were obtained from single aphid transmissions. Two GFMS12 sub-isolates (12-2 and 12-5) showed lower virulence in Mexican lime and grapefruit hosts than the original GFMS12, while sub-isolate 12-3 was more virulent than the original isolate, having severe stem-pitting (van Vuuren *et al.,* 2000). Thereafter, van Vuuren *et al.* (2000) proposed that this sub-isolate could become dominant under specific environmental conditions, resulting in decline and small fruit production. Furthermore, Scott *et al.* (2012) performed an in-depth study on the genotype diversity within the GFMS 12 cross-protecting source, by using different citrus hosts as well as SAT's to confirm all possible findings. Within this study, the authors made use of the 5'-end (ORF1a) as well as the 3'-terminal half (p23 gene) for genotypic strain composition analysis. It was found that GFMS 12 consisted out of at least four CTV genotype-like sequences (RB, T30, VT and B165) amongst the Mexican lime, Marsh and Star Ruby, each with a different concentration percentage in the different hosts. Since then, it has been

said that without CTV cross-protection in South Africa, grapefruit production will be heavily affected, leading to a uneconomical trade product (van Vuuren *et al.,* 1993).

1.3. References

Abbas, M., Khan, M. M., Mughal, S. M. and Khan, I. A. 2005. “Prospects of Classical Cross Protection Technique Against Citrus Tristeza Closterovirus in Pakistan.” *Horticultural Science (Prague)* 32(2), 74–83.

Adkar-purushothama, C. R., Quaglino, F., Casati, P. and Bianco, P. A. 2010. "Reverse transcription-duplex polymerase chain reaction of simutanious detection of *Citrus tristeza virus* and *Candidatus Liberibacter* from citrus plants." *Journal of Plant Diseases and Protection*, 117(6), 241-243.

Aird, D., Ross, M. G., Chen, W. S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D. B., Nusbaum, C. and Gnirke, A. 2011. “Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries.” *Genome Biology* 12, R18. doi:10.1186/gb-2011-12-2-r18, http://genomebiology.com/2011/12/2/R18.

Al-sadi, A. M., Al-hilali, S. A., Al-yahyai, R. A., Al-said, F. A., Deadman, M. L., Al-Mahmooli, H. I. and Nolasco, G. 2012. “Molecular Characterization and Potential Sources of Citrus Tristeza Virus in Oman.” *Plant Pathology* 61, 632–640. doi:10.1111/j.1365-3059.2011.02553.x.

Albiach-martí, M. R., Mawassi, M., Gowda, S., Satyanarayana, T., Hilf, M. E., Shanker, S., Almira, E. C., Vives, M. C., López, C., Guerri, J., Flores, R., Moreno, P., Garnsey, S. M. and Dawson, W. O. 2000. “Sequences of Citrus Tristeza Virus Separated in Time and Space Are Essentially Identical Sequences of Citrus Tristeza Virus Separated in Time and Space Are Essentially Identical.” *Journal of Virology* 74(15), 6856–6865. doi:10.1128/JVI.74.15.6856-6865.2000.Updated.

Albiach-marti, M. R. 2012. “Molecular Virology and Pathogenicity of Citrus Tristeza Virus.” *Viral Genomes – Molecular Structure, Diversity, Gene Expression Mechanisms and Host-Virus Interactions* 14, 275–302. Retrieved from: www.intechopen.com.

Albiach-marti, M. R. 2013. “The Complex Genetics of Citrus Tristeza Virus.” *InTech, Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications* (Chapter 1), 1–26. http://dx.doi.org/10.5772/56122.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. 1990. “Basic Local Alignment Search Tool.” *Journal of Molecular Biology* 215, 403–410.

Ambrós, S., El-mohtar, C., Ruiz-ruiz, S., Peña, L., Guerri, J., Dawson, W. O. and Moreno, P. 2011. “Agroinoculation of Citrus Tristeza Virus Causes Systemic Infection and Symptoms in the Presumed Nonhost Nicotiana Benthamiana.” *Molecular Plant-Microbe Interactions* 24(10), 1119–1131.

Amin, H. A. and Falk, B. W. 2009. “Citrus Tristeza Virus–p23 Gene Correlated with the Pathogenicity in Non Citrus Hosts.” *Egypt. J. Phytopathol.* 37(1), 83–98.

Ananthakrishnan, G., Venkataprasanna, T., Roy, A. and Brlansky, R. H. 2010. "Characterization of the mixture of a Citrus tristeza virus isolate by reverse transcription-quantitative real-time PCR." *Journal of Virological Methods*, 164, 75-82.

Ayllon, M. A., Lopez, C., Navas-Castillo, J., Garnsey, S. M., Guerri, J., Flores, R. and Moreno, P. 2001. “Polymorphism of the 5 Terminal Region of Citrus Tristeza Virus (CTV) RNA : Incidence of Three Sequence Types in Isolates of Different Origin and Pathogenicity.” *Archives of Virology* 146, 27–40.

Barba, M., Czosnek, H. and Hadidi, A. 2014. “Historical Perspective, Development and Applications of Next-Generation Sequencing in Plant Virology.” *Viruses* 6, 106–136.

Bar-Joseph, M. Garsney, S. M., Gonsalves, D., Moscovitz, M., Purcifull, D. E., Clark, M. F. and Loebenstein, G. 1979. “The Use of Enzyme-Linked Immunosorbent Assay for Detection of Citrus Tristeza Virus.” *Phytopathology* 69(2), 190–194.

Bar-Joseph, M., Marcus, R. and Lee R. F. 1989. “The Continuous Challenge of Citrus Tristeza Virus Control.” *Annual Review of Phytopathology*, 27, 291-316. doi: 10.1146/annurev.py.27.090189.001451

Bar-Joseph, M., Che, X., Mawassi, M., Gowda, S., Satyanarayana, T., Ayllón, M. A., Albiach-martí, M. R., Garnsey, S. M. and Dawson, W. O. 2002. "The Continuous Challenge of Citrus Tristeza Virus Molecular Research." *Fifteenth IOCV Conference, 2002—Citrus Tristeza Virus*, 1-7

Barzegar, A., Rahimian, H. and Sohi, H. H. 2010. "Comparison of the Minor Coat Protein Gene Sequences of Aphid-Transmissible and -nontransmissible Isolates of Citrus Tristeza Virus." *Journal of Plant Pathology* 76, 143–151. doi:10.1007/s10327-009-0216-7.

Bertolini, E., Moreno, A., Capote, N., Olmos, A., de Luis, A., Vidal, E., Pérez-Panadés, J. and Cambra, M. 2008. "Quantitative Detection of Citrus Tristeza Virus in Plant Tissues and Single Aphids by Real-time RT-PCR." *European Journal of Plant Pathology* 120, 177–188. doi:10.1007/s10658-007-9206-9.

Bexfield, N. and Kellam, P. 2011. "Metagenomic and the molecular identification of novel viruses." The Veterinary Journal, 190(2), 191-198. doi: 10.1016/j.tvjl.2010.10.014.

Brlansky, H. R. and Lee, R. F. 1990. "Numbers of Inclusion Bodies Produced by Mild and Severe Strains of Citrus Tristeza Virus Hosts." *Plant Disease* 74(4), 297-299.

Brlansky, R. H., Roy, A. and Damsteegt, V. D. 2011. "Stem-Pitting Citrus tristeza virus Predominantly Transmitted by the Brown Citrus Aphid from Mixed Infections Containing Non-Stem-Pitting and Stem-Pitting Isolates." *Plant Disease*, 95, 913-920.

Broadbent, P., Brlansky, R. H. and Indsto, J. 1996. "Biological Characterization of Australian Isolates of Citrus Tristeza Virus and Separation of Subisolates by Single Aphid Transmissions." *Plant Disease* 80(3), 329–333.

Cambra, M., Marroquı, C., Roma, M. P., Olmos, A., Martınez, M. C., de Mendoza, A. H., Lopez, A. and Navarro, L. 2000. "Incidence and epidemiology of Citrus tristeza 6irus in the Valencian Community of Spain." *Virus Research* 71, 85–95.

Cerni, S., Ruscic, J., Nolasco, G., Gatin, Z., Krajacic, M. and 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. doi:10.1007/s11262-007-0183-z.

Cervera, M., Esteban, O., Gil, M., Gil, L., Gorris, T. M., Martınez, M. C., Pena, L. and Cambra, M. 2010. “Transgenic Expression in Citrus of Single-chain Antibody Fragments Specific to Citrus Tristeza Virus Confers Virus Resistance.” *Transgenic Research* 19, 1001–1015. doi:10.1007/s11248-010-9378-5.

Che, X., Piestun, D., Mawassi, M., Yang, G., Satyanarayana, T., Gowda, S., Dawson W. O. and Bar-Joseph, M. 2001. “5’-Coterminal Subgenomic RNAs in Citrus Tristeza Virus-Infected Cells.” *Virology*, 283(2), 374-381.

Costa, A. S., Grant, T. J. and Moreira, S. 1954. "Behavior of various citrus rootstock-scion combination following inoculation with Mild and Severe strains of Tristeza virus." *Florida State Horticultural Society*, 26–30.

Costa, A. T., de Carvalho Nunes, M. W., Zanutto, C. A. and Muller, G. W. 2010. “Stability of Citrus Tristeza Virus Protective Isolates in Field Conditions.” *Pesquisa Agropecuária Brasileira,* 45 (7), 693–700.

da Graca, J. V. and van Vuuren, S. P. 2010. "Managing *Citrus tristeza virus* losses using cross protection. In: *Citrus tristeza virus Complex and Tristeza Diseases*." (Karasev, A. V. and Hilf, M. E. eds.) APS Press, St. Paul MN, 247-260.

D’Onghia, A. M. 2009. “Natural Tolerance / Resistance of Citrus Plants to Citrus Tristeza Disease.” *Options Méditerranéennes*, 181–185.

Dodds, J. A., Jarupat, T., Lee, J. G. and Roistacher, C. N. 1987. “Effects of Strains, Host, Time of Harvest, and Virus Concentration on Double-Stranded RNA Analysis of Citrus Tristeza Virus.” *Phytopathology* 77(3): 442–447.

Dolja, V. V., Kreuze, J. F. and Valkonen, J. P. T. 2006. “Comparative and Functional Genomics of Closteroviruses.” *Virus Research* 117, 38–51. doi:10.1016/j.virusres.2006.02.002.

Domingo, E. 2002. “Quasispecies Theory in Virology.” *Journal of Virology* 76(1), 28–31. doi:10.1128/JVI.76.1.463.

Duffy, S., Shackelton, L. A. and Holmes, E. C. 2008. “Rates of evolutionary change in viruses: patterns and determinants.” *Nature Reviews Genetics* 9(4), 267– 76.

Duhaime, M., Deng, L., Poulos, B. and Sullivan M. B. 2012. "Towards quantitative metagenomics of wild viruses and other ultra-low concentration DNA samples: a rigorous assessment and optimization of the linker amplification method." *Environmental Microbiology* 14, 2526 – 2537.

Fabre, F., Montarry, J., Coville, J., Senoussi, R., Simon, V. and Moury, B. 2012. “Modelling the evolutionary dynamics of viruses within their hosts: a case study using high- throughput sequencing.” Public Library of Science *Pathogens* 8(4), e1002654, doi:10.1371/journal.ppat.1002654.

Fagoaga, C., Lopez, C., de Mendoza, A. F., Moreno, P., Navarro, L., Flores, R. and Pena, L. 2006. “Post-transcriptional Gene Silencing of the P23 Silencing Suppressor of Citrus Tristeza Virus Confers Resistance to the Virus in Transgenic Mexican Lime.” *Plant Molecular Biology* 60, 153–165. doi:10.1007/s11103-005-3129-7.

Febres, V. J., Ashoulin, L., Mawassi, M., Frank, A., Bar-Joseph, M., Manjunath, K. L., Lee, R. F. and Niblett, C. L. 1996. “The P27 Protein Is Present at One End of Citrus Tristeza Virus Particles.” *Molecular Plant Pathology* 86(12), 1331–1335.

Fisher, L. C., Tennant, P. F. and Mc Laughlin, W. A. 2010. “Detection and characterization of Citrus tristeza virus stem pitting isolates in Jamaica.” European Journal of Plant Pathology, 127(1), 1-6.

Folimonova, S. Y., Folimonov, A. S., Tatineni, S. and Dawson, W. O. 2008. “Citrus Tristeza Virus : Survival at the Edge of the Movement Continuum.” *Journal of Virology* 82(13), 6546–6556. doi:10.1128/JVI.00515-08.

Folimonova, S. Y., Robertson, C. J., Shilts, T., Folimonov, A. S., Hilf, M. E., Garnsey, S. M. and 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(3) 1314–25. doi:10.1128/JVI.02075-09.

Folimonova, S. Y. 2013. "Developing an understanding of cross-protection by Citrus tristeza virus." *Virology*, 4(76), 1-9.

Garcia-Arenal, F., Fraile, A. and Malpica, J. M. 2001. “Variability and Genetic Structure of Plant Virus Populations.” *Annual Review of Phytopathology* 39, 157–186.

Garnsey, S. M., Civerolo, E. L., Gumpf, D. J., Paul, C., Hilf, M. E., Lee, R. F., Brlansky, R. H., Yokomi, R. K. and Hartung, J. S. 2005. “Biological Characterization of an International Collection of Citrus Tristeza Virus (CTV) Isolates.” In *Sixteenth IOCV Conference, 2005—Citrus Tristeza Virus*, 75–93.

Ghorbel, R., López, C., Fagoaga, C., Moren,o P., Navarro, L., Flore, R. and Peña, L. 2001. “Transgenic citrus plants expressing the citrus tristeza virus p23 protein exhibit viral-like symptoms.” Molecular Plant Pathology, 2(1), 27-36.

Gowda, S., Satyanarayana, T., Davis, C. L., Navas-castillo, J., Albiach-martı, M. R., Mawassi, M., Valkov, N., Bar-joseph, M., Moreno, P. and Dawson, W. O. 2000. “The P20 Gene Product of Citrus Tristeza Virus Accumulates in the Amorphous Inclusion Bodies.” *Virology* 274, 246–254. doi:10.1006/viro.2000.0413.

Gowda, S., Ayllón, M. A., Satyanarayana, T., Bar-Joseph, M. and Dawson, W. O. 2003. “Transcription Strategy in a Closterovirus : a Novel 5 ′ -Proximal Controller Element of Citrus Tristeza Virus Produces 5 ′ - and 3 ′ -Terminal Subgenomic RNAs and Differs from 3 ′ Open Reading Frame Controller Elements Transcription Strategy in a Clostero.” *Journal of Virology* 77(1), 340–352. doi:10.1128/JVI.77.1.340.

Gowda, S., Satyanarayana, T., Ayllon, M. A., Moreno, P., Flores, R. and Dawson, W. O. 2003. "The Conserved Structures of the 5J Nontranslated Region of Citrus Tristeza Virus Are Involved in Replication and Virion Assembly." *Virology* 317, 50–64. doi:10.1016/j.virol.2003.08.018.

Gowda, S., Satyanarayana, T., Folimonova, S. Y., Hilf, M. E. and Dawson, W. O. 2009. "Accumulation of a 5′ Proximal Subgenomic RNA of Citrus Tristeza Virus Is Correlated with Encapsidation by the Minor Coat Protein." *Virology* 389(1-2), 122–131. doi:10.1016/j.virol.2009.04.009. http://dx.doi.org/10.1016/j.virol.2009.04.009.

Grisoni, M., and Riviere, C. 1993. "Analysis of Epidemics of Citrus Tristeza Virus (CTV) in Young Citrus Groves Exposed to Aphid Infestation under Different Climatic Conditions in Reunion Island." In *Proceedlings of the Twelfth IOCV Conference*, 62–68.

Grosser, J. W., Medina-urrutia, V., Ananthakrishnan, G. and Serrano, P. 2004. "Building a Replacement Sour Orange Rootstock : Somatic Hybridization of Selected Mandarin + Pummelo Combinations." *Journal of American Society of Horticultural Sciences* 129(4), 530–534.

Hall, N. 2007. "Advanced sequencing technologies and their wider impact in microbiology." *Journal of Experimental Biology* 210, 1518–1525.

Harper, S. J., Dawson, T. E. and Pearson, M. N. 2009. "Complete Genome Sequences of Two Distinct and Diverse Citrus Tristeza Virus Isolates from New Zealand." *Archives of Virology* 154: 1505–1510. doi:10.1007/s00705-009-0456-z.

Harper, S. J., Dawson, T. E. and Pearson, M. N. 2010. "Isolates of Citrus Tristeza Virus That Overcome Poncirus Trifoliata Resistance Comprise a Novel Strain." *Archives of Virology* 155, 471–480. doi:10.1007/s00705-010-0604-5.

Herron, C.M., Mirkov, T. E., da Graca, J. V. and Lee, R. F. 2006. “Citrus Tristeza Virus Transmission by the Toxoptera Citricida Vector : In Vitro Acquisition and Transmission and Infectivity Immunoneutralization Experiments.” *Journal of Virological Methods* 134, 205–211. doi:10.1016/j.jviromet.2006.01.006.

Hilf, M. E., Karasev, A. V., Pappu, H. R., Gumpf, D. J., Niblett, C. L. and Garnsey, S. M. 1995. “Characterization of Citrus Tristeza Virus Subgenomic RNAs in Infected Tissue.” *Virology*, 208, 576-582.

Hilf, M. E., Karasev, A. V., Albiach-marti, M. R., Dawson, W. O. and Garnsey, S. M. 1999. “Two Paths of Sequence Divergence in the Citrus Tristeza Virus Complex.” *Phytopathology* 89(4), 336–342.

Hilf, M. E., and Garnsey, S. M. 2000. “Characterization and Classification of Citrus Tristeza Virus Isolates by Amplification of Multiple Molecular Markers.” In *Fourteenth IOCV Conference, 2000—Citrus Tristeza Virus*, 19–27.

Hilf, M. E., Mavrodieva, V. A. and 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(8), 909–917.

Huang, Z., Rundell, P. A., Guan, X. and Powell, C. A. 2005. “Evaluation of the Transmission of Different Field Sources of Citrus Tristeza Virus and the Separation of Different Genotypes by Single Brown Citrus Aphids.” *Horticultural Science* 40(3), 687–690.

Hull, R. 2002. “Variation, Evolution and Origins of Plant Viruses.” Mathews’ Plant Virology, Chapter 17, 743-812, *Academic press*, doi: 10. 1016/B978-012361160-4/50068-2.

Iglesias, G, Gago-zachert, S. P., Robledo, G., Costa, N., Plata, M. I., Vera, O., Grau, O. and Semorile, L. C. 2008. “Population Structure of Citrus Tristeza Virus from Field Argentinean Isolates.” *Virus Genes* 36, 199–207. doi:10.1007/s11262-007-0169-x.

Iracheta-Cardenas, M. M., Metheney, P., Polek, M. L., Manjunath, K. L., Lee, R. F. and Rocha-Pena, M. A. 2009. “Serological Detection of Citrus Tristeza Virus with Antibodies Developed to the Recombinant Coat Protein.” *Plant Diseases* 93, 11–16.

Itaya, A., Folimonov, A., Matsuda, Y., Nelson, R.S. and Ding, B. 2001. "Potato spindle tuber viroid as inducer of RNA silencing in infected tomato." Molecular Plant Microbes Interactions 14(11), 1332–1334.

James, D., Varga, A., Pallas, V. and Candresse, T. 2006. “Strategies for simultaneous detec- tion of multiple plant viruses.” *Canadian Journal of Plant Pathology* 28(1), 16 – 29.

Karasev, A. V., Boyko, V. P., Gowda, S., Nikolaeva, O. V., Hilf, M. E., Koonin, E. V., Niblett, C. L., Clines, K., Gumpf, D. J., Lee, R. F., Garnsey, S. M., Lewandowski, D. J. and Dawson, W. O. 1995. “Complete Sequence of the Citrus Tristeza Virus RNA Genome.” *Virology* 208(2), 511-20.

Karasev, A. V., Hilf, M. E., Garnsey, S. M. and Dawson, W. O. 1997. “Transcriptional Strategy of Closteroviruses : Mapping the 5’ Termini of the Citrus Tristeza Virus Subgenomic RNAs.” *Journal of Virology* 71(8), 6233–6236.

Knief, C. 2014. “Analysis of plant microbe interactions in the era of next generation sequencing technologies.” Frontiers in Plant Science, 5(216), 1-23.

Komazaki, S. 1984. “Biology and Virus Transmission of Citrus Aphids.” *Fruit Tre Research Station, Ministry of Agriculture, Forestry and Fisheries, Akitsu, Hiroshima 729-24, Japan*, 1–9.

Komazaki, S. 1993. “Biology and virus transmission of citrus aphids.” *Akitsu Branch, Fruit Tree Research Station, Ministry of Agriculture, Forestry and Fisheries, Akitsu, Hiroshima 729-24, Japan*: 1–9.

Komazaki, S. 1994. “Ecology of Citrus Aphids and Their Importance to Virus Transmission.” Japan Agricultural Research Quarterly 28, 177–184.

Kong, P., Rubio, L., Polek, M. and Falk, B. W. 2000. “Population Structure and Genetic Diversity Within California Citrus Tristeza Virus (CTV) Isolates.” *Virus Genes* 21(3), 139–145.

Korkmaz, S., Cevik, B., Onder, S., Koc, K. and Bozan, O. 2008.”Detection of Citrus tristeza virus (CTV) from Satsuma Owari mandarins (Citris Unshiu) by direct tissue blot immunoassay (DTBIA), DAS-ELISA, and biological indexing.” *New Zealand Journal of Crop and Horticultural Science*, 36(4) 239-246.

Lauring, A. S. and Andino, R. 2010. “Quasispecies Theory and the Behavior of RNA Viruses.” Public Library of Science *Pathogens* 6(7), 1–8. doi:10.1371/journal.ppat.1001005. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2908548&tool=pmcentrez&rendertype=abstract.

Lee, R. F., Herron, C. M., Mirkov, T. E. and da Graca, J. V. 2006. “Citrus tristeza virus transmission by the Toxoptera citricida vector : In vitro acquisition and transmission and infectivity immunoneutralization experiments.” *Journal of Virological Methods* 134: 205–211.

Li, K., Shrivastava, S., Brownley, A., Katzel, D., Bera, J., Nguyen, A. T., Thovarai, V., Halpin, R. and Stockwell, T. B. 2012. “Automated degenerate PCR primer design for high- throughput sequencing improves efficiency of viral sequencing.” *Virology Journal* 9(261): 1–20.

Linnarsson, S. 2010 “Recent advances in DNA sequencing methods - general principles of sample preparation.” *Experimental Cell Research*, 316, 1339–1343.

Liu, S., Vijayendran, D. and Bonning, B. C. 2011. “Next Generation Sequencing Technologies for Insect Virus Discovery.” *Viruses* 3, 1849–1869. doi:10.3390/v3101849.

López, C., Navas-Castillo, J., Gowda, S., Moreno, P. and Flores, R. 2000. “The 23-kDa Protein Coded by the 3J-Terminal Gene of Citrus Tristeza Virus Is an RNA-Binding Protein.” *Virology* 269, 462–470. doi:10.1006/viro.2000.0235.

López, C., Cervera, M., Fagoaga, C., Moreno, P., Navarro, L., Flores, R. and Peña, L. 2010. “Accumulation of Transgene-derived siRNAs Is Not Sufficient for RNAi-mediated Protection Against Citrus Tristeza Virus in Transgenic Mexican Lime.” *Molecular Plant Pathology* 11(1), 33–41. doi:10.1111/J.1364-3703.2009.00566.X.

Lu, R., Folimonov, A., Shintaku, M., Li, W-X., Falk, B. W., Dawson, W. O. and Ding, S-W. 2004. “Three Distinct Suppressors of RNA Silencing Encoded by a 20-kb Viral RNA Genome.” Proceedings of the National Academy of Sciences 101(44), 15742–15747.

Martın, S., Sambade, A., Rubio, L., Vives, M. C., Moya, P., Guerri, J., Elena, S. F. and Moreno, P. 2009. “Contribution of Recombination and Selection to Molecular Evolution of Citrus Tristeza Virus Printed in Great Britain.” *Journal of General Virology* 90, 1527–1538. doi:10.1099/vir.0.008193-0.

Mathews, D. M., Riley, K. and Dodds, J. A. 1997. “Comparison of Detection Methods for Citrus Tristeza Virus in Field Trees During Months of Nonoptimal Titer.” *Plant Disease* 81: 525–529.

Matos, L. A., Hilf, M. E., Cayetano, X. A., Feliz, A. O., Harper, S. J. and Folimonova, S. Y. 2013. “Dramatic Change in Citrus tristeza virus Populations in the Dominican Republic.” *Plant Disease* 93, 339–345.

Melzer, M. J., Borth, W. B., Sether, D. M., Ferreira, S., Gonsalves, D. and Hu, J. S. 2010. “Genetic Diversity and Evidence for Recent Modular Recombination in Hawaiian Citrus Tristeza Virus.” *Virus Genes* 40, 111–118. doi:10.1007/s11262-009-0409-3.

Metzker, M. L. 2010. “Sequencing technologies - the next generation.” *Nature Reviews* Genetics 11, 31–46.

Meyer, J. B., van Vuuren, S. P., Luttig, M., Manicom, B. Q., Graça, J. V. and Ruby, S. 2005. “Strain Prevalence of Citrus tristeza virus Cross-Protecting Isolates Altered by Red Grapefruit Hosts citrus.” In *Sixteenth IOCV Conference, 2005—Citrus Tristeza Virus*, 205–212.

Michaud, J. P. 1998. “A Review of Literature on Toxoptera Citricida (Kirkaldy) (Homoptera: Aphididae).” *Florida Entomologist* 81(1), 37–61.

Moreno, P., Ambros, S., Albiach-Marti, M. R., Guerri, J. and Pena, L. 2008. “Plant Diseases That Changed the World Citrus Tristeza Virus : a Pathogen That Changed the Course of the Citrus Industry.” *Molecular Plant Pathology* 9(2), 251–268. doi:10.1111/J.1364-3703.2007.00455.X.

Morey, M., Fernandez-Marmiesse, A., Castineiras, D., Fraga, J. M., Couce, M. L. and Choco, J. A. 2013. “A glimpse into past, present, and future DNA sequencing.” *Molecular Genetics and Metabolism*, 110(1-2), 3-24. doi: 10.1016/j.ymgme.2013.04.024.

Müller, G. W., Costa, A. S., Kitajima, E. W. and Camargo, I. J. B. 1974. "Additional evidence that tristeza multiplies in Passiflora spp." *In Proceedlings of the 6th IOCV Conference*, 75-78.

Neill, J. D., Bayles, D. O. and Ridpath, J. F. 2014. “Simultaneous rapid sequencing of multiple RNA virus genomes.” *Journal of Virological Methods* 201, 68–72. http://dx.doi.org/10.1016/j.jviromet.2014.02.016.

Nekrutenko, A. and Taylor, J. 2012. “Next-generation sequencing data interpretation: enhancing reproducibility and accessibility.” *Nature Publishing Group* 13(9), 667–672. http://dx.doi.org/10.1038/nrg3305.

Nelson, S., Melzer, M. and Hu, J. 2011. “Citrus Tristeza Virus in Hawai’i.” *Plant Disease,* 1–16.

Niblett, C. L., Genc, H., Cevik, B., Halbert, S., Brown, L., Nolasco, G., Bonacalza, B.,Manjunath, K. L., Febres, V. J., Pappu, H. R. and Lee. R. F. 2000. “Progress on Strain Differentiation of Citrus Tristeza Virus and Its Application to the Epidemiology of Citrus Tristeza Disease.” *Virus Research* 71, 97–106.

Oyola, S. O., Otto, T. D., Gu, Y., Maslen, G., Manske, M., Campino, S., Turner, D. J., Macinnis, B., Kwiatkowski, D. P., Swerdlow, H. P. and Quail, M. A. 2012. "Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes." *BMC Genomics* 13(1), 1-12. doi:10.1186/1471-2164-13-1, http://www.biomedcentral.com/1471-2164/13/1.

Palacio, A., and Duran-Vila, N. 1999. "Single-strand Conformation Polymorphism (SSCP) Analysis as a Tool for Viroid Characterisation." *Journal of Virological Methods* 77, 27–36.

Papayiannis, L. C., Santos, C., Kyriakou, A., Kapari, T. and Nolasco, G. 2007. "Molecular Characterization of Citrus Tristeza Virus Isolates from Cyprus on the Basis of the Coat Protein Gene." *Journal of Plant Pathology* 89(2), 291–295.

Pappu, S. S., Febres, V. J., Pappu, H. R., Lee, R. F. and Niblett, C. L. 1997. "Characterization of the 3' Proximal Gene of the Citrus Tristeza Closterovirus Genome." *Virus Research* 47, 51–57.

Pelosi, R. R., Rundell, P. A., Cohen, M. and Powell, C. A. 2000. "Evaluation of a Sixteen-Year Citrus Tristeza Virus Cross-Protection Trial in Florida." In *Fourteenth IOCV Conference, 2000—Citrus Tristeza Virus*, 111–114.

Perring, T. M., Gruenhagen, N. M. and Farrar, C. A. 1999. "Management of plant viral diseases through chemical control of insect vectors." *Annual Review of Entomology* 44, 457–481.

Qureshi, J. A. 2010. "Implications of Climate Change for *Toxoptera citricida* (Kirkaldy), a Disease Vector fo Citrus in Florida." In: Aphid Biodiversity Under Environmental Change, Kindlmann, P.,Dixon, A. F. G and Michaud, J. P., Dordrecht : Springer Science+Business Media B.V., 91-106.

Read, D. A., and Pietersen, G. 2015. "Genotypic Diversity of Citrus Tristeza Virus within Red Grapefruit , in a Field Trial Site in South Africa." *European Journal of Plant Pathology*, 1–15. doi:10.1007/s10658-015-0631-x.

Rocha-Pena, M. A. and Lee, R. F. 1991. "Serological technique for detection of *Citrus tristeza virus*." *Journal of Virological Methods*, 34, 311-331.

Roistacher, C. N. 2004. “Diagnosis and Management of Virus and Virus like Diseases of Citrus.” Diseases of Fruits and Vegetables, Volume 1, 109-189.

Roistacher, C. N., and Moreno, P. 2007. “The Worldwide Threat from Destructive Isolates of Citrus Tristeza Virus-A Review.” In *Eleventh ZOCV Conference*, 7–19.

Roistacher, C. N., Graça, J. V. and Müller, G. W. 2010. “Cross Protection Against Citrus Tristeza Virus - a Review.” *Proceedings, 17th Conference, 2010 – Citrus Tristeza Virus*, 1–27.

Roossinck, M. J., Saha, P., Wiley, G. B., Quan, J., White, J. D., Lai, H., Chavarria, F., Shen, G. and Roe, B. A. 2010. “Ecogenomics: using massively parallel pyrosequencing to understand virus ecology.” *Molecular Ecology* 19(Suppl. 1), 81–88.

Ross, M. G., Russ, C., Costello, M., Hollinger, A., Lennon, N. J., Hegarty, R., Nusbaum, C. and Jaffe, D. B. 2013. “Characterizing and measuring bias in sequence data.” *Genome Biology* 14(5), R51. http://genomebiology.com/2013/14/5/R51.

Roy, A., Manjunath, K. L. and Brlansky, R. H. 2005. “Assessment of Sequence Diversity in the 5 -terminal Region of Citrus Tristeza Virus from India.” *Virus Research* 113, 132–142. doi:10.1016/j.virusres.2005.04.023.

Roy, A., and Brlansky, R. H. 2010. “Genome Analysis of an Orange Stem Pitting Citrus Tristeza Virus Isolate Reveals a Novel Recombinant Genotype.” *Virus Research* 151, 118–130. doi:10.1016/j.virusres.2010.03.017.

Roy, A., Ananthakrishnan, G., Hartung, J. S. and Brlansky, H. R. 2010. “Development and Application of a Multiplex Reverse-Transcription Polymerase Chain Reaction Assay for Screening a Global Collection of Citrus Tristeza Virus Isolates.” *Phytopathology* 100(10), 1077–1088.

Roy, A., Choudhary, N., Hartung, J. S. and Brlansky, R. H. 2013. “The Prevalence of the Citrus tristeza virus Trifoliate Resistance Breaking Genotype Among Puerto Rican Isolates.” *Plant Disease* 97, 1227–1234.

Rubio, L., Aylloni, M. A., Guerri, J., Pappu, H., Niblett, C. and Moreno, P. 1996. “Differentiation of citrus tristeza closterovirus (CTV) isolates by single-strand conformation polymorphism analysis of the coat protein.” Annals of Applied Biology, 129(3), 479–489.

Rubio, L., Ayllón, M. A., Kong, P., Fernandez, A., Polek, M., Guerri, J., Moreno, P. and 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(17), 8054–8062. doi:10.1128/JVI.75.17.8054.

Rubio, L., Guerri, J. and Moreno, P. 2013. “Genetic variability and evolutionary dynamics of viruses of the family Closteroviridae.” *Virology* 4(151), 1–15.

Ruiz-Ruiz, S., Moreno, P., Guerri, J. and Ambros, S. 2006. “The Complete Nucleotide Sequence of a Severe Stem Pitting Isolate of Citrus Tristeza Virus from Spain : Comparison with Isolates from Different Origins.” *Archives of Virology* 151, 387–398. doi:10.1007/s00705-005-0618-6.

Ruiz-Ruiz, S., Moreno, P. and Ambros, S. 2007. “A Real-time RT-PCR Assay for Detection and Absolute Quantitation of Citrus Tristeza Virus in Different Plant Tissues.” *Journal of Virological Methods* 145, 96–105. doi:10.1016/j.jviromet.2007.05.011.

Sambade, A., Lopez, C., Rubio, L., Flores, R., Guerri, J. and Moreno, P. 2003. “Polymorphism of a Specific Region in Gene P23 of Citrus Tristeza Virus Allows Discrimination Between Mild and Severe Isolates.” *Archives of Virology* 148, 2325–2340. doi:10.1007/s00705-003-0191-9.

Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989 "Molecular Cloning, a laboratory manual." Cold Harbor Spring Press. http://trove.nla.gov.au/work/13615226

Sanz-ramos, M., Díaz-san Segundo, F., Escarmís, C., Domingo, E. and Sevilla, N. 2008. "Hidden Virulence Determinants in a Viral Quasispecies In Vivo." 82(21), 10465–10476. doi:10.1128/JVI.00825-08.

Saponari, M., Manjunath, K. and Yokomi, R. K. 2008. "Quantitative Detection of Citrus Tristeza Virus in Citrus and Aphids by Real-time Reverse transcription-PCR (TaqMan®)." *Journal of Virological Methods* 147, 43–53. doi:10.1016/j.jviromet.2007.07.026.

Satyanarayana, T., Gowda, S., Mawassi, M., Albiach-Maria, M. R., Ayllon, M. A., Robertson, C., Garnsey, S. M. and Dawson, W. O. 2000. "Closterovirus Encoded HSP70 Homolog and P61 in Addition to Both Coat Proteins Function in Efficient Virion Assembly." *Virology* 278, 253–265. doi:10.1006/viro.2000.0638.

Satyanarayana, T., Bar-Joseph, M., Mawassi, M., Albiach-Martı, M. R., Ayllon, M. A., Gowda, S., Hilf, M. E., Moreno, P., Garnsey, S. M. and Dawson, W. O. 2001. "Amplification of Citrus Tristeza Virus from a cDNA Clone and Infection of Citrus Trees." *Virology* 280, 87–96. doi:10.1006/viro.2000.0759.

Satyanarayana, T., Gowda, S., Ayllon, M. A., Albiach-Martı, M. R., Rabindran, S. and Dawson, W. O. 2002. "The P23 Protein of Citrus Tristeza Virus Controls Asymmetrical RNA Accumulation." *Journal of Virology* 76(2), 473–483. doi:10.1128/JVI.76.2.473.

Satyanarayana, T., Gowda, S., Ayllón, M. A. and Dawson, W. O. 2003. "Frameshift mutations in infectious cDNA clones of Citrus tristeza virus: a strategy to minimize the toxicity of viral sequences to Escherichia coli." *Virology*, 313(2), 481-491.

Satyanarayana, T., Gowda, S., Ayllon, A. and Dawson, W. O. 2004. "Closterovirus Bipolar Virion : Evidence for Initiation of Assembly by Minor Coat Protein and Its Restriction to the Genomic RNA 5' Region." *Proceedings of the National Academy of Sciences*, 101(3), 799–804.

Scott, K. A., Hlela, Q., Zablocki, O., Read, D., van Vuuren, S. and 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,* 1-22. (Springer-Verlag). doi:10.1007/s00705-012-1450-4.

Sebastien, M., Olmos, A., Jijakli, H. and Candresse, T. 2014. "Current impact and future directions of high throughput sequencing in plant virus diagnostics." *Virus Research,* 1–7. http://dx.doi.org/10.1016/j.virusres.2014.03.029.

Shendure, J. A.; Porreca, G. J.; Church, G. M. 2008. "Overview of DNA sequencing strategies." Current Protocols *in* Molecular Biology*,* Chapter 7, Unit 7.1

Silva, G., Marques, N. and Nolasco, G. 2012. "The evolutionary rate of Citrus tristeza virus ranks among the rates of the slowest RNA viruses." *Journal of General Virology,*93(2), 419-429. doi: 10.1099/vir.0.036574-0.

Solis-Garcia, N., Kahlke, C. J., Herron, C. M., Graça, J. V., Essau, K. L., Miao, H. Q. and Skaria, M. 2001. "Surveys for Citrus Tristeza Virus in Texas 1991-2000." *Subtropical Plant Science* 53, 4–8.

Soler, N., Plomer, M., Fagoaga, C., Moreno, P., Navarro, L., Flores, R. and Pena, L. 2012. "Transformation of Mexican lime with an intron-hairpin construct expressing untranslatable versions of the genes coding for the three silencing suppressors of Citrus tristeza virus confers complete resistance to the virus." *Plant Biotechnology Journal* 10, 597–608.

Solonenko, S. A., Ignacio-espinoza, J. C., Alberti, A., Cruaud, C., Hallam, S., Konstantinidis, K., Tyson, G., Wincker, P. and Sullivan, M. B. 2013. "Sequencing platform and library preparation choices impact viral metagenomes." *BMC Genomics* 14(320), 1–12.

Stewart, K. A. 2006. "Strain Differentiation of Citrus Tristeza Virus Isolates from South Africa by PCR and Microarray." *Thesis*. University of Pretoria, South Africa, Pretoria, 1-122.

Targon, M. L. P. N., Machado, M. A., Carvalho, S. A., Souza, A. A. and Müller, G. W. 2000. “Differential Replication of a Mild and a Severe Citrus Tristeza Virus Isolate in Species and Varieties of Citrus.” In *Fourteenth IOCV Conference, 2000—Citrus Tristeza Virus*, 127–130.

Tatineni, S., Robertson, C. J., Garnsey, S. M., Bar-joseph, M., Gowda, S. and Dawson, W. O. 2008. “Three Genes of Citrus Tristeza Virus Are Dispensable for Infection and Movement Throughout Some Varieties of Citrus Trees.” *Virology* 376, 297–307. doi:10.1016/j.virol.2007.12.038.

Tatineni, S., Gowda, S. and Dawson, W. O. 2010. “Heterologous Minor Coat Proteins of Citrus Tristeza Virus Strains Affect Encapsidation , but the Coexpression of HSP70h and P61 Restores Encapsidation to Wild-type Levels.” *Virology* 402, 262–270. doi:10.1016/j.virol.2010.03.042. http://dx.doi.org/10.1016/j.virol.2010.03.042.

Tatineni, S., Robertson, C. J., Garnsey, S. M., and Dawson, W. O. 2011. "A plant virus evolved by acquiring multiple nonconserved genes to extend its host range." Proceedings of the National Academy of Sciences, 108 (42), 17366–17371. doi:10.1073/pnas.1113227108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1113227108

Thapa, V., Melcher, U., Wiley, G. B., Doust, A., Palmer, M. W., Roewe, K., Roe, B. A., Shen, G., Roossinck, M. J., Wang, Y. M. and Kamath, N. 2012. "Detection of members of the Secoviridae in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA." *Virus Research* 167(1), 34–42.

Tsai, J. H., Lui, Y. H., Wang, J. J. and Lee, R. F. 2000. "Recovering of organge stem pitting strains of *Citrus Tristeza Virus* (CTV) following single aphid transmissions with *Toxoptera citricida* from a Florida decline isolate of CTV." *Proceedings of the Florida State Horticultural Society* 113, 75–78.

Tsai, J. H., Lee, R. F., Liu, Y-h. and Niblett, C. L. 2009. *Tsai, J. H., Lee, R. F., Liu, Y-h. and Niblett, C. L. 2009, "Biology and Control of Brown Citrus Aphid (Toxoptera Citricida Kirkaldy) and Citrus Tristeza." In: Radcliffe, E. B., Hutchison, W. D. and R. E. Cancelado (eds.) "Radcliffe's IMP World Textbook."* Ed. E. B. Radcliffe, W. D. Hutchison, and R. E. Cancelado. St. Paul: University of Minnesota. http://ipmworld.umn.edu/chapters/tsaietal.htm

van Vuuren, S. P., Collins, R. P. and da Graca, J. V. 1993. "Evaluation of Citrus Tristeza Virus Isolates for Cross Protection of Grapefruit in South Africa." *Plant Diseases* 77, 24–28.

Van Vuuren, S. P., Collins, R. P. and Da Graca, J. V. 1993. "Growth and production of lime trees pre-immunized with different mild Citrus tristeza virus isolates in the presence of natural disease conditions." *Phytophylactica* 25, 49–52.

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." *Fourteenth IOCV Conference, 2000—Citrus Tristeza Virus*, 103–110.

van der Vyver, J. B., van Vuuren, S. P., Luttig, M. and Graça, J. V. 2002. "Changes in the Citrus Tristeza Virus Status of Pre-immunized Grapefruit Field Trees." *Fifteenth IOCV Conference, 2002—Citrus Tristeza Virus*, 175–185.

Voelkerding, K. V.; Dames, S. A.; Durtschi, J. D. 2009. "Next-generation sequencing: From basic research to diagnostics." *Clinical Chemistry* 55, 641–658.

von Broembsen, L. A., da Graca, J. V., Lee, A. T. C. and Waller, G. G. 1978. "South Africa's Citrus Improvement Programme after five years." *Citrus Subtropical Fruit Journal*, 532, 11-18.

Wang, L. F. 2011. "Discovering novel zoonotic viruses." New South Wales Public Health Bulletin 22, 113–117.

Wylie, S. J. and Jones, M. G., 2011. "The complete genome sequence of a Passion fruit woodiness virus isolate from Australia determined using deep sequencing, and its relationship to other potyviruses." *Archives of Virology* 156(3), 479–482.

Yang, Z-n., Mathews, D. M., Dodds, J. A. and Mirkov, T. E. 1999. “Molecular Characterization of an Isolate of Citrus Tristeza Virus That Causes Severe Symptoms in Sweet Orange.” *Virus Genes* 19(2), 131–142.

Yokomi, R. K., Lastra, R., Stroetzel, M. B., Damsteegt, D., Lee, R. F., Garsney, S. M., Gottwald, T. R., Rocha-Pena, M. A. and Niblett, C. L. 1994. "Establishment of the brown citrus aphid (Homoptera:Aphididae) in Central America and the Caribbian Basin and transmission of the citrus tristeza virus." *Journal of Economic Entomology* 87, 1078–1085.

Yokomi, R. K. 2009. “Citrus Tristeza Virus.” *Options Méditerranéennes,* 19–33.

Yokomi, R. K., Saponari, M. and Sieburth, P. J. 2010. “Rapid Differentiation and Identification of Potential Severe Strains of Citrus Tristeza Virus by Real-Time Reverse Transcription- Polymerase Chain Reaction Assays.” *Phytopathology* 100(4), 319–327.

Zablocki, O. and Pietersen, G. 2014. “Characterization of a Novel Citrus Tristeza Virus Genotype within Three Cross-Protecting Source GFMS12 Sub-Isolates in South Africa by Means of Illumina Sequencing.” *Archives of Virology* 107. doi:10.1007/s00705-014-2041-3.

Zhou, C. Y., Hailstones, D. L., Broadbent, P., Connor, R. and Bowyer, J. 2002. “Studies on Mild Strain Cross Protection Against Stem-pitting Citrus Tristeza Virus.” In *Fifteenth IOCV Conference, 2002—Citrus Tristeza Virus*, 151–157.

1.4. Website references:

http://ipm.ifas.ufl.edu/Agricultural_IPM/asian.shtml

http://seqanswers.com/forums/images/content/ilmn-step1-6.jpg

www.cdn.interchopen.com

www.independent.com.mt

www.eppo.org

www.ressources.ciheam.org

www.nappo.org

www.en.fffc.org.tw

Chapter 2

The isolation of homogeneous CTV genotypic sources by single aphid transmissions (SAT's) and "Bark-flap and stem-slash" mechanical inoculation approaches.

2.1. Abstract.

Citrus tristeza virus (CTV) is a well known viral pathogen to citrus worldwide, infecting nearly all citrus species and cultivars, various citrus relatives and some intergeneric hybrids. It is clear that CTV is capable of economically devastating the citrus exporting trade and economies worldwide. For that, CTV cross-protection programs were initiated to aid in crop protection and reducing yield losses in CTV infected areas. Identification and isolation of single genotypes and the evaluation of their symptom severity is critical for future cross-protection studies against CTV. The aim of this study is to obtain homogenous CTV genotypic sources by performing mechanical inoculation procedures as well as Single Aphid Transmissions (SAT's) onto healthy Mexican Lime seedlings and screening for CTV presence. Single aphid transmissions, using natural aphid infestations, together with bark-flap and stem slash inoculations were performed onto healthy Mexican Lime seedlings. All seedlings were screened for CTV presence using a CTV generic primed RT-PCR analysis. Of the three CTV inoculation procedures, six CTV positives were obtained from the SAT trials. These six positive seedlings were screened for genotypic presence using a genotype specific RT-PCR analysis. All six positives appeared to be potentially homogenous for the RB genotype. Amplification of the p33 gene region followed by Illumina next-generation sequencing (NGS) was performed on the six samples to detect, characterize and confirm this RB homogeneity. Reference mapping performed on each of the six filtered sample sequences confirmed that sample 14-7102 primarily contained Taiwan-Pum/SP/T1 (JX266712), an RB variant, while the other 5 samples (14-7103, 14-7104, 13-7105, 14-7107 and 14-7132) mainly contained the CTV isolate AT-1 (JQ061137), as detected in the p33 RT-PCR analysis. All six samples contained a mixed CTV genotypic presence with low levels of other genotypes being present as well as the dominant RB genotype.

2.2. Introduction.

To find and establish a pure *Citrus tristeza virus* (CTV) source is an empirical, extremely time consuming and difficult process. CTV consists out of numerous isolates, each having their own distinct biological and genetic traits (Folimonova *et al.*, 2010, 2013). Many citrus trees contain a mixed CTV infection (Zhou *et al.*, 2002) due to the variety of isolates circulating the environment by means of various viral transmission means. Mixed infections are often also transmitted within mild strain cross-protection schemes (Roistacher *et al.*, 2010). Single-aphid transmitted sub-isolates of the cross-protecting sources GFMS 12, GFMS 35 and LMS 6 are products of mixed infections (Roistacher *et al.*, 2010). The CTV vector, *Toxoptera citricida* (Brown citrus aphid) is the most efficient in transmitting this virus to citrus (Huang *et al.*, 2005; Lee *et al.*, 2006; Tsai *et al.*, 2009). The export of CTV infected budwood and the propagation of these materials on healthy citrus trees (Hilf *et al.*, 2005; Moreno *et al.*, 2008; Matos *et al.*, 2013) plays a big role in the worldwide spread of CTV. However, the type of citrus cultivar as well as the environmental conditions can also play a role in which CTV genotypic sources would be present or more severely expressed (van der Vyver *et al.*, 2002; Folimonova *et al.*, 2010).

Protection of citrus trees against CTV is still a major requirement and ongoing research process. Folimonova *et al.* (2010) focused on determining the relationships between different CTV genotypes in terms of how they will aid in cross-protection, by assisting in superinfection exclusion. Superinfection exclusion is defined as the process in which citrus cultivars are protected when primary viral infections (with mild strains) prevent secondary infections (severe stains) from a similar or the same viral strain, Within this publication it was shown that the same or similar CTV stain infections can indeed prevent superinfections from occurring. Both experimental attempts performed within the article supported the above statement. From these findings we now know the general approach to cross-protecting citrus cultivars is to maintain protection against severe isolates of a particular CTV genotype by first inoculating plants with mild isolates of the same genotype to successfully exclude superinfection (Folimonova, 2013). As CTV infections often contain more than one genotype, it is necessary to isolate each component separately in order to confirm the mild nature of the each. Several attempts to obtain pure CTV sources through single aphid transmissions (SAT's), onto indicator hosts, have been done.

Variable single aphid transmission rates have been reported over the years, ranging from extremely low to moderate percentages of success (Komazaki, 1994; Yokomi *et al.*, 1994; Broadbent *et al.*, 1996; Tsai *et al.*, 2000; Huang *et al.*, 2005; Barzegar *et al.*, 2010) However, effective CTV transmissions through SAT's, using hundreds of Brown citrus aphids (*T. citricida* (Kirkaldy)) (BrCA) per seedling, have been proven to be successful (Meneghini, 1946; Tsai *et al.*, 2009). Nevertheless, this approach can potentially transmits numerous CTV genotypes (Broadbent *et al.*, 1996; Barzegar *et al.*, 2010) in which continuous transfers will have to be performed to obtain a pure CTV genotypic strain. Still, single genotypes (mild and severe sources) can be isolated from mixed CTV infections by means of SAT's (Brlansky *et al.*, 2003; Huang *et al.*, 2005).

During this research chapter SAT's were performed by firstly, collecting naturally occurring *T. citricida* (Brown citrus aphid) in citrus orchards and then transferring one aphid per Mexican Lime seedling. Also, due to the low efficiency of SAT's, two alternative, mechanical inoculation approaches namely bark-flap and stem-slash inoculations were performed on Mexican Lime seedlings, using various dilutions, to obtain end-point transmissions. All three procedures were performed with the aim of obtaining pure CTV genotypic sources.

2.3. Materials and methods

2.3.1. *Collection of leaf material prior to bark-flap and stem-slash evaluations.*

Leaf samples were collected from citrus trees (Table 3) located at the University of Pretoria experimental farm (GPS co-ordinates 25°45'06.40"S; 28°15'35.43"E) (Figure 12), and a composite sample (13-4142) obtained from leaves from trees from Malelane and Schoemanskloof. All leaf material was processed by cutting out the midrib as well as the leaf petiole under sterile laboratory conditions. All processed samples were stored at -80°C until used.

Figure 12: Map of the experimental farm, University of Pretoria, showing the sites where random sampling of the different citrus trees was done.

Cara-cara sweet orange trees (25°45'02.92"S; 28°15'27.36"E); (25°45'03.58"S; 28°15'27.51"E)
GFMS 12 and GFMS 35 trees (25°45'06.17"S; 28°15'35.71"E); (25°45'06.44"S; 28°15'35.71"E)
Citrus trees behind greenhouse (25°45'05.07"S; 28°15'37.33"E); (25°45'05.00"S; 28°15'37.94"E)

Table 3: Citrus trees of which leaf samples were collected from located at the experimental farm, University of Pretoria and from the Schoemanskloof/Malelane area, South Africa.

Tree sample	Accession number
Cara-cara sweet orange	13-4138
Cara-cara sweet orange	13-4139
Unknown citrus tree behind greenhouse	13-4140
Unknown citrus tree behind greenhouse	13-4141
GFMS 12 inoculated tree located within the greenhouse▲	08-0049
GFMS 35 inoculated tree located within the greenhouse▲▲	08-0035
Unknown citrus leaf mixture from Schoemanskloof & Malelane	13-4142

▲: *Citrus x paradise* cultivar Star Ruby; ▲▲: *Citrus aurantifolia* cultivar Mexican Lime.

2.3.2. ***Collection of aphids prior to SAT's.***

Brown Citrus Aphids (*T. citricida* (Kirkaldy)) (BrCA) (Tsai *et al.*, 2009) were collected from two farms located near Brits in the North West province (Wolhuter kop area), South Africa. Sixty seven were collected and used from a Clementine (*Citrus clementina;* mandarin x sweet orange) tree (14-7000) at "Berg en Dal" farm and 100 from three Eureka Lemon (*Citrus limon* (L.) Burm. f.)) trees (14-7001 to 14-7003) at F19-21, WJ van Dis, Bokfontein (Agrelek). They were collected by sampling twigs and associated leaves containing the feeding aphids, placing them into prior labelled sampling bags and keeping them in a cooler box for the duration of the trip.

Aphids collected were removed from the twigs within the laboratory by agitating the aphids with a small paint brush causing them to withdraw their stylets from the plant tissue, and then placed in a petri dish (Merck KGaA, Darmstadt, Germany) in order to conduct the transmissions. Material from which the aphids had been derived were processed and stored as above and screened for CTV genotypes present.

2.3.3. ***Virus particle (Sucrose) partial purification from citrus leaf material (13-4142).***

The leaves of composite sample (13-4142) were pooled together and processed by cutting out the midrib and the petiole using a sterile scalpel blade. Partial Sucrose purification of the virus particles was done from this material using the procedure of Robertson *et al.* (2005) with slight modifications. Partially purified sample 13-4142 was used in the first bark-flap inoculation trail.

Briefly, the processed plant material was cut up to smaller pieces using a sterile scalpel blade and divided into two equal amount samples. To the first sample 1:5 (w/v) 40mM Potassium Phosphate buffer, pH7.8, with 5% Sucrose (buffer A) was added and macerated using a mortar and pestle, while the second sample was macerated with both buffer as well as liquid Nitrogen using a mortar and pestle. The macerated sample was filtered through cheese cloth pre-soaked in buffer A. The filtrate was centrifuged for 10 minutes at 10 000 x g in a Beckman-Coulter Optima L-100 XP ultracentrifuge (Beckman-Coulter, Fullerton, CA, USA), where the collected fraction (20ml) was transferred to ultracentrifuge tubes containing 1 ml 40mM Potassium Phosphate, pH 7.8, with 70% Sucrose (buffer B) and centrifuged for 90 minutes at 30 000 x g. Fractions (400µl) of the purified solution were collected and placed in separate microtubes, followed by visualizing each tube with a UV lamp to determine where the sucrose cushion is satiated so that the fractions could be collected. The fraction (400µl) closest to the sucrose cushion interface was discarded as it contained some of the sucrose cushioning layer, followed by the collection of three fractions from each of two tubes of 540µl, 250µl and 545µl (tube one) and 430µl, 410µl and 550µl (tube two). Three replicates of 10µl each of the collected fractions were used in a CTV TAS-ELISA analysis to determine the fraction containing CTV. The collected fractions were stored at 4°C until used.

2.3.4. ***Triple antibody sandwich (TAS)-ELISA analysis.***

Partially purified virion fractions were tested by CTV TAS-ELISA (Rocha-Pena *et al.*, 1991). Three replicates of 10µl each of the six collected fractions were added to 90µl of coating buffer (1.59g/l Na_2CO_3; 2.93g/l $NaHCO_3$; pH 9.6), each in separate 1.5ml eppendorf tubes. Three controls were used during the TAS-ELISA analysis: Healthy/negative control, 12-1013 ML5 [Mexican Lime 5] (CTV-), positive control, 12-1023 DG3 [Duncan Grapefruit 3] (CTV+) and a Buffer control (BC). The buffer control consists of 100µl of coating buffer with 0.2% Tween 20 with no plant sample added. All the incubation periods was done either in an incubator or fridge, depending on the required incubation temperature. The TAS-ELISA plate used was enclosed in a plastic container containing a damp cloth, which aids in preventing the test solutions from drying out.

A sterile flat bottom 96-well microtitre plate was coated with CTV 1052 (rabbit prepared anti-CTV) polyclonal antiserum (CREC, Lake Alfred, FL, USA), diluted (1:5000) in coating buffer (1.59g/l Na_2CO_3; 2.93g/l $NaHCO_3$; pH 9.6). 200µl aliquots of the diluted antiserum were added to each well on the plate. The plate was then incubated overnight at 4°C. The plate was washed three times, for three minutes, with deionized water and Phosphate-buffered saline with Tween 20 (8g/l NaCl; 0.2g/l KH_2PO_4 phosphate; 1.15g/l Na_2HPO_4 (anhydrous); 0.2g/l KCl; 1ml Tween 20; pH 7.4) between each incubation stage.

One-Hundred microlitres aliquots of the samples were pipetted into three test wells on the antibody coated microtitre plate. The plate was then incubated overnight at 4˚C, allowing antigens to bind to the available antibodies. Following washing as described above, the polyclonal detecting antibody, G604 (goat prepared anti-CTV) (CREC, Lake Alfred, FL, USA), was diluted (1:20 000) in conjugate buffer (500ml PBST; 10g/500ml; PVP-40; 1g/500ml BSA) and pipetted into the wells being tested, followed by an incubation period of 1 hours at 37˚C. After washing, rabbit-anti-goat antibody IgG conjugate with alkaline phosphatase (Sigma A-4187) (Sigma-Aldrich, St. Louis, MO, USA) was diluted (1:30 000) in the conjugate buffer, followed by pipetting 100µl aliquots to each test well. The test plate was then incubated for 2 hours at 37˚C. After washing, the substrate, 4-nitrophenyl phosphate hexahydrate (KPL, Gaithersburg, MD, USA) 0.6mg/ml in 10% diethanolaminee buffer, followed by

pipetting 200µl aliquots in each test well. The test plate was then incubated for the last time at room temperature (25˚C) for 30 to 60 minutes, allowing colour development to take place.

The hydrolyzed enzyme substrate extinction values were recorded at an absorbance of 405nm (OD_{405}) during the reaction, using a Multiskan™ Go Microplate Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). A sample was considered positive if the sample value obtained was higher than the threshold value calculated (mean of healthy control, plus 2x the standard deviation).

2.3.5. ***Bark-flap and stem-slash analysis on newly grown Mexican Lime seedlings.***

2.3.5.1. Bark-flap inoculation trial 1:

The Sucrose purified fraction with the highest CTV titre was used to inoculate 40 newly grown (5 months old) Mexican Lime seedlings. All seedlings were confirmed CTV negative by means of RT-PCR analysis prior to any of the inoculation trails. The bark-flap analysis was done according to the protocol of Robertson *et al.* (2005), with minor modification. Four five-fold serial dilutions were made and kept on ice until used. Ten Mexican Lime seedlings were bark-flap inoculated per dilution.

Plants were maintained under greenhouse conditions (temperature set at 20 - 30°C, watered daily). A single cut of approximate 1 to 2 cm long and 2 to 3 mm wide was made in the bark (exposing the phloem area) of the seedling, using a sterile scalpel blade. The bark was gently pulled away from the stem and 10-12µl of the diluted inoculum was applied to the phloem-exposed opening created. The flap was then put back into place and closed with grafting tape (Plastrip, Somerset West, South Africa).

2.3.5.2. Bark-flap inoculation trial 2:

Citrus leaves were collected from an unknown citrus tree (13-4140) in a citrus orchard (25°45'05.07" S; 28°15'37.33" E) situated at the Experimental farm of the University of Pretoria. The leaves were processed by cutting out the midrib and the petiole using a sterile scalpel blade and storing the material in sterile 1.5 ml eppendorf tubes at -80°C until used. Thereafter, the processed plant material and 10 ml of 40mM Potassium Phosphate buffer, pH7.8, with 5% Sucrose (Buffer A) were macerated using a Homex 6 (Bioreba, Reinach, Basel-Landschaft, Switzerland) tissue macerator. The macerated solution was collected and pipetted into a 15 ml tube (Greiner Bio-One GmbH, Frickenhausen, Germany) and kept on ice until used. A five-fold serial dilution of the macerated solution was made and kept on ice until used for bark-flap inoculation as described above.

2.3.5.3. Stem-slash inoculation trial:

Dilutions of the "crude" macerated (13-4140) solution which was prepared and used in the second bark-flap inoculation trail was also used for stem-slash inoculation onto 40 Mexican Lime seedlings using the protocol of Robertson *et al.* (2005) and Satyanarayana *et al.* (2001), with minor modifications. Ten Mexican Lime seedlings were stem-slash inoculated per dilution, with 10-20µl sap dilutions applied to a sterile scalpel blade and approximately 20 crosscuts made along the stem of each seedling holding the blade at a 30° angle to keep the droplet on the blade and to allow the solution to enter the stem openings made. This was repeated a second time with each seedling stem-slash inoculated on two sides. The inoculated areas were then closed with grafting tape (Plastrip, Somerset West, South Africa).

The bark-flap as well as the stem-slash inoculated seedlings was left for a 3 month post inoculation (PI) and 6 month PI intervals. After the respective periods PI, the seedlings were examined for any potential CTV symptoms and screened for CTV presence using CTV genotype specific RT-PCR analysis.

2.3.6. ***Single Aphid Transmissions onto newly grown Mexican Lime seedlings.***

Aphids collected were removed from the plant material as described above. Thereafter the aphids were starved for 60 minutes, followed by the transfer of a single aphid per seedling. The aphids were left for 24 hours to feed. Thereafter, all aphids were removed and killed with a contact insecticide to prevent any potential CTV spread. The seedlings were also treated with Confidor 200 SC systemic insecticide (Bayer CropScience AG, Monheim, Germany). Plants were maintained in the greenhouse situated at the experimental farm at the University of Pretoria, South Africa. The seedlings were left for 5 months PI where after, all seedlings were screened for CTV presence using a generic RT-PCR directed at all known genotypes. In the case of positive seedlings, these were screened for the individual CTV genotypes present, to determine whether a potential homogeneous source was obtained through SAT's.

TAS-ELISA analysis is a good alternative CTV screening method. In this current study TAS-ELISA analysis was not performed for this specific screening purpose as results would have been obtained much faster through a generic RT-PCR analysis.

2.3.7. ***RNA extractions.***

Leaf samples from all virus inoculation plants were collected after the respective periods PI. Up to 6 leaves were randomly sampled from each seedling. During sampling, the seedlings were also monitored for any potential CTV symptoms. Leaves were placed into labelled bags and kept at 4°C in the laboratory, until subjected to total RNA extraction. Total RNA extraction was performed on the plant material using a GeneJET Plant RNA Purification mini kit (Thermo Scientific, Waltham, MA, USA), according to the manufacturers protocol. This involved macerating petioles and leaf midribs using liquid Nitrogen and a sterile mortar and pestle. This was followed by adding 500µl Plant RNA Lysis Solution and 10µl of 2M Dithiothreitol (DTT) (Thermo Scientific, Waltham, MA, USA). Samples were mixed thoroughly by vortexing them for 10-20 second, then incubated for 3 minutes at 65°C

and centrifuged for 5 minutes at a maximum speed of 14,000 x g, at 10°C. The resultant supernatant was pipetted into a 1.5ml microtube, together with 250µl of 96% Ethanol. The solutions were mixed, and transferred to a purification column, inserted into a collection tube. The column was centrifuged for 1 minute at 11,000 x g, at 10°C. The flow-through was discarded and 700µl of Wash Buffer 1 (WB1) was added to the reassembled purification column. The column was centrifuged for 1 minute at 11,000 x g, at 10°C, followed by discarding both the flow-through and the collection tube. The purification column was reassembled with a new clean 2ml collection tube and 500µl of Wash Buffer 2 (WB2) added to the purification column. The column was centrifuged for 1 minute at 11,000 x g, at 10°C, subsequently discarding the flow-through and repeating the wash step one more time with WB2. The column was centrifuged for 1 minute at 14,000 x g, at 10°C. After the centrifugation period, the purification column was transferred to a RNase-free 1.5ml collection tube. To elute the required RNA, 50µl of nuclease-free water was added to each purification column membrane and centrifuged for 1 minute at 11,000 x g, at 10°C. The RNA obtained was stored at -80°C until used.

All extracted RNA were screened using 1% Agarose gel (SeaKem LE agarose, Lonza, Rockland, ME, USA) electrophoresis, to confirm a positive extraction. The prepared gel was pre-stained with Ethidium Bromide (10mg/ml) at 2.5µl per 50ml, (Promega, Madison, WI, USA), followed by the adding of 8µl of RNA together with 2µl of 6x loading dye (Thermo Scientific, Waltham, MA, USA) per well. The gel was run for 30 min at 100 Volts using a gel electrophoresis system (Bio-Rad, Hercules, California, USA) and 1x TAE buffer (40mM Tris-acetate, 1mM EDTA, pH 8.2).

2.3.8. Screening for CTV presence through genotype specific RT-PCR analysis.

A volume of 2µl of total RNA extract was used as template for the production of cDNA for the different CTV genotypes screened for. A denaturation reaction mixture of 10µl was made up for each CTV genotype, as follows: 2µl of total RNA, 1µl of the respective CTV genotypic reverse primers (10mM) (Table 4) and 7µl PCR

grade H_2O (Promega, Madison, WI, USA). The denaturation reaction conditions were as follows: 95°C for 3 minutes (to disrupt the RNA secondary structures) and 4°C for 1 minute to stop the reaction.

Reverse transcription was carried by adding 4µl 1x RT buffer (Roche, Mannheim, Germany), 2µl (1mM) dNTP mix (Kapa Biosystems Inc, Wilmington, MA, USA), 2µl (10mM) DTT (Thermo Scientific, Waltham, MA, USA), 0.125µl (1.25U/µl) *Avian myeloblastosis virus* reverse transcriptase (AMV-RT) (Roche, Mannheim, Germany), 0.25µl (0.5U/µl) Ribolock RNAse Inhibitor (Thermo Scientific, Waltham, MA, USA) and 1.625µl PCR grade H_2O (Promega, Madison, WI, USA) to the 10µl denatured RNA mixture, to make a final reaction volume of 20µl. The reverse transcription reaction conditions were as follow: 42°C incubation for 60 minutes, 85°C for 5 minutes and then maintained at 4°C until used.

PCR amplification was carried out on the cDNA in order to screen for individual CTV genotypes present. The amplification reaction was carried out by using 2µl of the cDNA as template with 10µl Promega GoTaq Hot Start Green Master Mix (1,000 reactions) (Promega, Madison, WI, USA), 0.75µl (0.375µM) forward primer, 0.75µl (0.375µM) reverse primer (Table 4) and 6.5µl PCR grade H_2O (Promega, Madison, WI, USA) to make a final volume of 20µl. The following PCR amplification reaction conditions were used: 1 cycle of 94°C for 2 minutes (denaturation step); 35 cycles of 94°C for 20 seconds, various temperatures for different primers (Table 2) for 30 seconds, 72°C for 20 seconds (primer annealing and DNA extension steps); 1 cycle of 72°C for 1 minute and 4°C until the PCR products are used. A 1.5% Agarose gel electrophoresis was done on all PCR reactions to confirm the presence of any potential CTV genotypes.

2.3.9. PCR amplification of the CTV p33 gene region.

PCR amplification was carried out on the CTV positive cDNA (obtained in section 2.3.8) by using 3µl of the 1kb Univ-p33-Reverse-primed cDNA as template, which was added to a reaction mixture containing 10µl Promega GoTaq Hot Start Green Master Mix (Promega, Madison, WI, USA), 0.75µl (0.375µM) Univ-p33-Forward primer 0.75µl (0.375µM) 1kb Univ-p33-Reverse primer and 5.5µl PCR

grade H_2O (Promega, Madison, WI, USA) to make a final volume of 20µl. The following PCR amplification reaction conditions were used: 1 cycle of 92°C for 2 minutes (initial denaturation); 35 cycles of 92°C for 30 seconds (denaturation), 55°C for 45 seconds (annealing), 72°C for 1 minute (elongation); 1 cycle of 72°C for 10 minutes (final extension) and 4°C until the PCR products are used. A 1.5% Agarose gel electrophoresis was done on all obtained PCR products to confirm the presence of the p33 gene region (±980bp). Illumina Next-Generation high-throughput sequencing was performed on the CTV p33-gene region positive SAT samples, to determine if any of the trees contained a potentially pure CTV source.

2.3.10. Agarose gel electrophoresis.

PCR's were analysed by means of Agarose gel (SeaKem LE agarose, Lonza, Rockland, ME, USA) electrophoresis. Twelve microlitres of each sample was loaded per well of a 1.5% Agarose gel, pre-stained with Ethidium Bromide (10mg/ml) at 2.5µl per 50ml, (Promega, Madison, WI, USA) and run for 60 min at 75 Volts using a gel electrophoresis system (Bio-Rad Hercules, California, USA) using 1x TAE buffer (40mM Tris-acetate, 1mM EDTA, pH 8.2). HyperLadder IV 100bp (Bioline, Celtic Molecular Diognostics (Pty) Ltd., South Africa) was used as a molecular size marker Results were viewed and recorded using the Gel Doc™ EZ Imager (Bio-Rad Hercules, California, USA).

2.3.11. Column purification of PCR products.

Gel band purifications were performed on all CTV positive samples prior to direct Sanger sequencing. The Promega, Wizard® SV Gel and PCR Clean-up system (Promega, Madison, WI, USA) was used, as per the manufacturer's instructions. In brief, the positive DNA bands were excised after electrophoresis by means of a sterile scalpel blade and transferred to a pre-weighed, sterile 1.5ml microtube, and the weight of the gel slice determined. Membrane binding solution was added in a ratio 10µl solution per 10mg of gel slice. Each tube was then vortexed incubated at 56°C for 10 min or until the gel slice was completely dissolved.

The tube was briefly centrifuged at 8, 000 x g for 2 seconds at room temperature and transferred to a SV Minicolumn assembly and incubated for 1 min at room temperature, followed by centrifugation at 10, 000 x g for 1 min. The sample flow-through was discarded, followed by a wash step by adding 700µl Membrane Wash Solution. The assemblies were then centrifuged at 10,000 x g for 1 min and the flow-through discarded and the wash step repeated with 500µl Membrane Wash Solution, followed by centrifugation of the SV Minicolumn assembly at 10,000 x g for 5 min. Collection tubes were emptied and re-centrifuged for 1 min. Each SV Minicolumn was transferred to a to a sterile 1.5ml eppendorf tube and fifty microlitres of Nuclease-free water (Promega, Madison, WI, USA) added directly onto the membrane and allowed to stand for 1 min at room temperature. The SV minicolumn and microtube assembly was then centrifugation at 10,000 x g for 1 min. The SV Minicolumns were discarded and all purified DNA samples were stored at -20°C until used. Agarose gel electrophoresis was done to confirm DNA presence as described above. Purified DNA was re-amplified by PCR prior to sequencing to increase the amount of DNA template available.

Table 4: CTV genotype specific primer sets used during RT-PCR analysis.

CTV genotype	Polarity *	Primer sequence (5' - 3')	Length (bp)	Annealing temp. (˚C)	Primer binding sites within CTV genotype genome	Product size (nt)	Reference
T36	S	TTC CCT AGG TCG GAT CCC GAG TAT A	25	59.9	1775...2610	836	Roy *et al.* 2010
	A	CAA ACC GGG AAG TGA CAC ACT TGT TA	26	59.2			
Generic	S	ATG GAC GAC GAR ACA AAG AAA TTG AAG A	28	57.8	16105…16776	672	Roy *et al.* 2010
	A	TCA ACG TGT GTT RAA TTT CCC AAG CT	26	58.4			
B165	S	GTT AAG AAG GAT CAC CAT CTT GAC GTT GA	29	58.4	2124...2633	510	Roy *et al.* 2010
	A	AAA ATG CAC TGT AAC AAG ACC CGA CTC	27	59.2			
T3	S	GTT ATC ACG CCT AAA GTT TGG TAC CAC T	28	58.7	4846...5254	409	Roy *et al.* 2010
	A	CAT GAC ATC GAA GAT AGC CGA AGC	24	57.7			
VT	S	TTT GAA AAT GGT GAT GAT TTC GCC GTC A	28	59.5	1945...2246	302	Roy *et al.* 2010
	A	GAC ACC GGA ACT GCY TGA ACA GAA T	25	60.2			
T30	S	TGT TGC GAA ACT AGT TGA CCC TAC TG	26	58.7	588...793	206	Roy *et al.* 2010
	A	TAG TGG GCA GAG TGC CAA AAG AGA T	25	60.1			
NZRB-28, -M12,G90 & HA18-9							
NZRB F1	S	AGT GGT GGA GAT TAC GTT G	19	60	1979...2626	646	G. Cook, *Unpublished*
NZRB R1	A	TAC ACG CGA CAA ATC GAG	18	60			
NZRB-30 & -17 (Should not detect T36)							
NZRB F2	S	CGG AAG GGA CTA CGT GGT	18	60	1981...2643	662	G. Cook, *Unpublished*
NZRB R2	A	CGT TTG CAC GGG TTC AAT G	19	60			
HA16-5 F	S	TAG GAA GGG TCA CTG CCC TGA CA	23	56	2154...2813	658	G. Cook, *Unpublished*
HA16-5 R	A	GTA AGT ATC TAA AAC CAG GAG	21	56			
CTV p33 gene region							
Univ-p33-F	S	GATGTTTGCCTTCGCGAGC	19	55	11020....12001	980	D. Read, *Unpublished*
▲Univ-p33-R	A	CCCGTTTAAACAGAGTCAAACGG	23	55			

Polarity *: Sense = S; Antisense = A. **▲:** 1kb sized Universal p33 reverse primer.

2.3.12. Sequencing.

PCR products were purified prior to Sanger sequencing. In brief, 2µl FastAP Alkaline Phosphatase (1U/µl) (Thermo Scientific, Waltham, MA, USA) and 0.5µl Exonuclease (20U/µl) (Thermo Scientific, Waltham, MA, USA) were added to 19µl of the purified PCR product or template DNA in individual PCR tubes. Samples were incubated for 15 min at 37°C followed by a 15 min incubation period at 85°C, to conclude the reaction.

Purified PCR products were sequenced using both sense and antisense primers using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) as per manufacturer's instructions. Sequencing reactions were as follows; 2.25µl 5x sequencing buffer (Applied Biosystems, Foster City, CA, USA), 0.75µl of 2µM CTV individual genotypic primers (Table 4) (Integrated DNA Technologies, Coralville, IA, USA), 1µl of 2.5x terminator mix v3.1 (Applied Biosystems, Foster City, CA, USA), nuclease free molecular grade water and template DNA to make it up to a final volume of 10µl. The amount of template DNA and hence the nuclease free molecular water added was determined by the intensity of the gel band obtained following column purification. The sequencing reaction conditions were as follows: 92°C for 1 min; 30 cycles each of 94°C for 10 seconds (sec), 50°C for 5 sec, and 60°C for 4 min.

For each of the 10µl reaction mixture a product precipitation was performed after the completion of the thermo-cycling to remove all unwanted elements and dry the DNA pellet before sending the sample in for direct Sanger sequencing. The precipitation process involved the addition of 1µl of 125Mm EDTA, 1µl of 3M Sodium Acetate and 25µl of 100% non-denature Molecular grade Ethanol (Sigma-Aldrich, St. Louis, MO, USA) to each of the reaction tubes containing the 10µl sequencing product. The solution was mixed by vortexing it for 2 seconds and then incubated for 15 minutes at room temperature (25°C). After incubation the solution was centrifuged at a maximum speed of 14,000 x g, at 4°C, for 20 to 30 minutes and the resulting supernatant was removed by pipette. One hundred microlitres of 70% Ethanol (Sigma-Aldrich, St. Louis, MO, USA) was immediately added to each reaction tube and centrifuged at a maximum speed of 14,000 x g, at 4°C, for 10 to 15 minutes. The

supernatant was once again removed without disturbing the pellet at the bottom of the reaction tube. After the removal of the supernatant the samples were allowed to air-dry, with the lid of the reaction tube open, for 20 minutes or alternatively at 94°C for no longer than 1 minute. The resulting precipitated reactions were submitted for DNA Sanger sequencing at the African Centre for Gene Technologies, Automated Sequencing Facility, Department of Genetics, University of Pretoria, South Africa, making use of an ABI Prism 3130XL Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

2.3.13. Phylogenetic analysis.

CLC Main Workbench 6 software package (CLC Bio, Aarhus, Denmark) was used to view the sequence chromatograms, evaluate the quality and completeness of sequences, correct any errors within the sequences and produce a consensus sequence for each sample. Screening of sequence identities was done using the National Centre for Biotechnology Information (NCBI) Basic Local Alignment Tool (BLAST) (http://www.ncbi.nlm.nih.gov/) (Altschul *et al.*, 1990). Complete consensus sequences, together with 45 complete genome CTV reference sequences, were aligned using MAFFT version 7 online software using the default settings (http://mafft.cbrc.jp/alignment/server/ ; Katoh *et al.*, 2002, 2005). The default setting were as follows; the “strategy” set to auto, “scoring matrix for amino acid sequences” set to BLOSUM62, “scoring matrix for nucleotide sequences” set to 200PAM/k=2, “align unrelated segments” set to try to align gap regions anyway, “number of homologs” set to 50, “threshold” set to 1a-10 and “plot last hit threshold” set to score=39 (E=8.4e11). Cognate sequences of all 45 CTV reference sequences were excised separately for each of the amplified regions of the specific CTV genotypes using BioEdit Sequence Alignment Editor software (Version 7.0.8, Hall, T.A. 1999, Ibis Bioscience, Carlsbad). The editing of the sequences was based on the different primer binding sites for each CTV genotype (Table 4). All edited and aligned sequences were saved in the necessary formats required (FASTA, Phylip4) for the different phylogenetic analyses. Three different phylogenetic analyses were done on all sequences generated namely; Neighbor-Joining, Bayesian and Maximum Likelihood phylogenetic analysis.

The reference sequences were retrieved from the NCBI database. The reference together with their respective Genbank accession number and original citation follows; NUagA [seedling yellows strain] (AB046398) (Suastika etal., unpublished data), SY568 (AF001623) (Yang *et al.*, 1999), T30 (AF260651) (Albiach-martı *et al.*, 2000), CTV (AY170468) (Satyanarayana *et al.*, 1999), Qaha from Egypt (AY340974) (Abdelmaksoud & Gamal El-din, unpublished data), T318A (DQ151548) (Ruiz-Ruiz *et al.*, 2006), Mexican CTV isolate (DQ272579) (Quiroz *et al.*, unpublished data), B165 (EU076703) (Roy & Brlansky, 2010), CTV strain SP (EU857538) (Harper *et al.*, 2009), CTV strain VT (EU937519) (Weng *et al.*, 2007), CTV strain T30 (EU937520) (Weng *et al.*, 2007), CTV strain T36 (EU937521) (Weng *et al.*, 2007), NZRB-M12 (FJ525431) (Harper *et al.*, 2010), NZRB-G90 (FJ525432) (Harper *et al.*, 2010), NZRB-TH28 (FJ525433) (Harper *et al.*, 2010), NZRB-TH30 (FJ525434) (Harper *et al.*, 2010), NZRB-M17 (FJ525435) (Harper *et al.*, 2010), NZ-B18 (FJ525436) (Harper *et al.*, 2010), HA18-9 (GQ454869) (Melzer *et al.*, 2010), HA16-5 (GQ454870) (Melzer *et al.*, 2010), CTV strain Kpg3 (HM573451) (Biswas *et al.*, 2012), CTV strain B301 (JF957196) (Roy *et al.*, unpublished data), CTV strain AT-1 (JQ061137) (Wu *et al.*, unpublished data), CTV strain A18 (JQ798289) (Sindhvajiva *et al.*, unpublished data), CTV strain CT14A (JQ911663) (Wang *et al.*, unpublished data), CTV strain CT11A (JQ911664) (Wang *et al.*, unpublished data), CTV strain T68-1 (JQ965169) (Hilf *et al.*, unpublished data), Taiwan-Pum/SP/T1 (JX266712) (Chen *et al.*, unpublished data), Taiwan-Pum/M/T5 (JX266713) (Chen *et al.*, unpublished data), L192GR (KC262793) (Varveri *et al.*, unpublished data), CT-ZA3 (KC 333868) (Zablocki & Pietersen, 2014), FS674-T36 (KC517485) (Harper & Hilf, unpublished data), FS701-T36 (KC517486) (Harper & Hilf, unpublished data), FS703-T36 (KC517487) (Harper & Hilf, unpublished data), FS577 (KC517488) (Harper & Hilf, unpublished data), FS701-T30 (KC517489) (Harper & Hilf, unpublished data), FL278-T30 (KC517490) (Harper & Hilf, unpublished data), FS703-T30 (KC517491) (Harper & Hilf, unpublished data), FS703-VT (KC517492) (Harper & Hilf, unpublished data), FL202-VT (KC517493) (Harper & Hilf, unpublished data), FS701-VT (KC517494) (Harper & Hilf, unpublished data), T3 (KC525952) (Hilf & Harper, unpublished data), CTV (NC001661) (Karasev *et al.*,1997), T36 (U16304) (Pappu *et al.*, 1994), T385 (Y18420) (Vives *et al.*, 1999). A more detailed table containing information for each of the 45 CTV reference genomes used is available in Appendix 1.

These CTV reference genomes were those available at the time of analysis and were used to confirm the results obtained from the RT-PCR as well as the BLAST analysis. The following CTV complete genomes were selected to represent the 9 different CTV genotypic clades within the phylogenetic trees; EU076703.3 (B165), EU937519.1 (VT), EU937520.1 (T30), EU937521.1 (T36), FJ525431.1 (RB Clade 1), FJ525434.1 (RB Clade 2), GQ454870.1 (HA16-5), JQ965169.1 (T68-1) and KC525952.1 (T3).

Nucleotide Sequence alignments were analysed in jModelTest version 2.1.4 (Darriba *et al.*, 2012) using the Akaike information criterion model. The models used for the sequences (FASTA format) are indicated in Table 5. The jModelTest analysis looks at all possible changes that can occur from one nucleotide to another, within sequences compared, along a phylogenetic tree. Thus, different substitution models are determined and used when different sequences are compared. The genomic region of where the specific sequences were taken from can also play a role, in which substitution model will be used, as some areas within a DNA sequence is more conserved than other. Each substitution model was calculated specifically for specific sequence data put together. This process will be repeated if combinations of different sequence data sets are put together; thus, a completely new substitution model will be calculated for these sequences. These models were used during Bayesian as well as Maximum Likelihood phylogenetic analysis.

Table 5: Nucleotide substitution models used for respective phylogenetic analysis data sets.

Sample	Nucleotide substitution model
Mexican Lime seedling CTV presence prior to bark-flap and stem-slash inoculations	
T36	GTR+I+G
NZRB F1R1	GTR+I+G
NZRB F2R2	GTR+I+G
HA16-5	TIM2+I+G
VT	TIM3+I+G
B165	TIM2+I+G
T3	TrN+I+G
Mexican Lime seedling CTV presence after the bark-flap and stem-slash inoculations	
T36	TIM2+I+G
NZRB F1R1	TIM2+I+G
B165	TIM2+I+G
VT	HKY+I+G
Collected citrus leaf material CTV presence prior to SAT's	
T36	TIM2+I+G
NZRB F1R1	TIM2+I+G
NZRB F2R2	TIM2+I+G
HA16-5	TIM3+I+G
VT	TIM3+I+G
B165	TIM2+I+G
T3	TrN+I+G
T30	TVMef+G
Mexican Lime seedling CTV presence after SAT's ▲	
T36, NZRB F1R1, NZRB F2R2, VT, HA16-5, B165, T3	
14-7102	TPM1uf+I+G
14-7103	
14-7104	
14-7105	
14-7107	
14-7132	

▲: CTV p33 gene region was used to determine the purity of the positive SAT's seedlings. One tree with all 6 samples was produced to determine where which sample genotypically groups. GTR: General Time Reversible (Rodriguez et. al., 1990); TIM: Transitional model; TrN: Tamura-Nei (Tamura and Nei 1993); HKY: Hasegawa-Kishino-Yano (Hasegawa et. al., 1985); TVMef: Transversional model with equal base frequencies. TPM1uf: Three-parameter model with unequal base frequencies. G: Gamma distribution; I: Proportion of invariable sites.

Maximum Likelihood phylogenetic analysis was performed for all sequences obtained (Phylip4 format) in PhyML version 3.0 (Guindon and Gascuel, 2003). All default parameters, together with the appropriate nucleotide substitution model, were used except for the branch support that was set on a 1000 bootstrap replicates. Additionally, the topology of tree searching was set on a "best of both" (NNI and SPR) approach.

Prior to the Bayesian phylogenetic analysis, MrBayes blocks were created using the JModel Test results obtained (Table 5) from the different sequences analysed as well as information gained for the following websites; http://www.molecularevolution.org/software/phylogenetics/mrbayes and http://hydrodictyon.eeb.uconn.edu/eebedia/index.php/Phylogenetics:_MrBayes_Lab . The MAFFT aligned sequences (FASTA format) were converted to a nexus data file using an online sequence format converter program (http://genome.nci.nih.gov/tools/reformat.html) and added to the MrBayes block created. Bayesian analysis was done for all obtained sequences using MrBayes version 3.2.1 software (Ronquist *et al.*, 2012). The appropriate substitution models were used together with 4 Markov chains and at least 3 million generations. Each analysis was run twice. P files obtain from the analysis were analysed in Tracer version 1.5 (Rambaut and Drummond, 2009) to determine the required burn-in values. The burn-in values were added to the analysis and tree files were completed and trees were viewed and edited in FigTree version 1.3.1 (Rambaut, 2009).

Neighbor-Joining analysis for all sequences generated was performed using MEGA version 5 (Tamura *et al.*, 2011). The Jukes-Cantor model was used with a bootstrap analysis of a 1000 replicates.

2.3.14. Illumina Next-Generation High-throughput sequencing of positive SAT samples.

The amplicons obtained from the above mentioned p33 PCR amplification reaction were purified using the Nucleospin® Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. The Wizard® SV Gel and PCR Clean-up System (Promega, Madison, WI, USA) was replaced by the above mentioned purification kit as higher concentrations and more pure DNA was obtained for Illumina sequencing. The concentrations of the purified products were determined using a Qubit-iT™ dsDNA BR Assay Kit (Invitrogen™, Carlsbad, CA, USA; Eugene, Oregon, USA), according to the manufacturer's instructions. The samples were submitted to the Agricultural Research Council (ARC) – Biotechnology platform, situated in Pretoria, South Africa, where the samples were prepared using

TruSeq3 – PE – 2 adapters (Illumina, San Diago, CA, USA) according to the manufacturer's specification, following by a sequencing reaction on 1/24th of a lane using Illumina (Miseq platform) Next-Generation High-throughput sequencing (Illumina, San Diago, CA, USA).

2.3.15. Illumina Next-Generation High-throughput sequencing data analysis.

Illumina paired-end data analysis was performed using CLC Genomics Workbench version 5.1.1. (CLC bio, QIAGEN company, Prismet, Aarhus, DK, Denmark) software. Briefly, the raw sequence data was imported as paired reads (distance of 180 – 600). Quality scores for each data set were generated using the FastQC function according to its default settings. The reads were filtered by removing low quality sequences (quality limit of 0.05), ambiguous nucleotides (maximum of 2 nucleotides allowed) as well as the adapter sequences used, Truseq3 – PE – 2 (PrefixPE/1: TACACTCTTTCCCTACACGACGCTCTTCCGATCT; PrefixPE/2: GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT; PE1: TACACTCTTTCCCTACACGACGCTCTTCCGATCT; PE1_rc: AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTA; PE2: GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT; PE2_rc: AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC).

The 45 CTV reference sequences trimmed to the cognate p33 gene region were imported and used during reference mapping. The following settings were used to perform reference mapping: mismatch cost of 2, insertion cost of 3, deletion cost of 3, length fraction of 0.9, similarity fraction of 0.9, conflict resolution by voting (A, T, C, G) and non-specific match handling was set to be ignored. Detailed reports were created after each of the steps discussed.

2.4. Results

2.4.1. Bark-flap and stem-slash inoculation results.

TAS-ELISA analysis was conducted on sample 13-4142 to find the sucrose gradient fraction with the highest CTV titre present, for the use in inoculation and is presented in Figure 13. Sample 13-4142 was chosen for this process due to the diverse number of CTV genotypes present (Table 6). The positive/negative threshold (0.552) was determined by using the mean of the healthy control plus twice the standard deviation obtained (Figure 13). Only the CTV positive control value (0.57) was greater than that of the threshold value calculated. Values lower than the negative and buffer control was obtained for all 6 partially purified fractions, suggesting that virus was lost during the partial purification, however virus transmission theoretically only needs a single virus particle and hence inoculations (first inoculation trial) were still conducted using fraction 1. For the second and third inoculation trial, crude plant sap of sample 13-4140 was used.

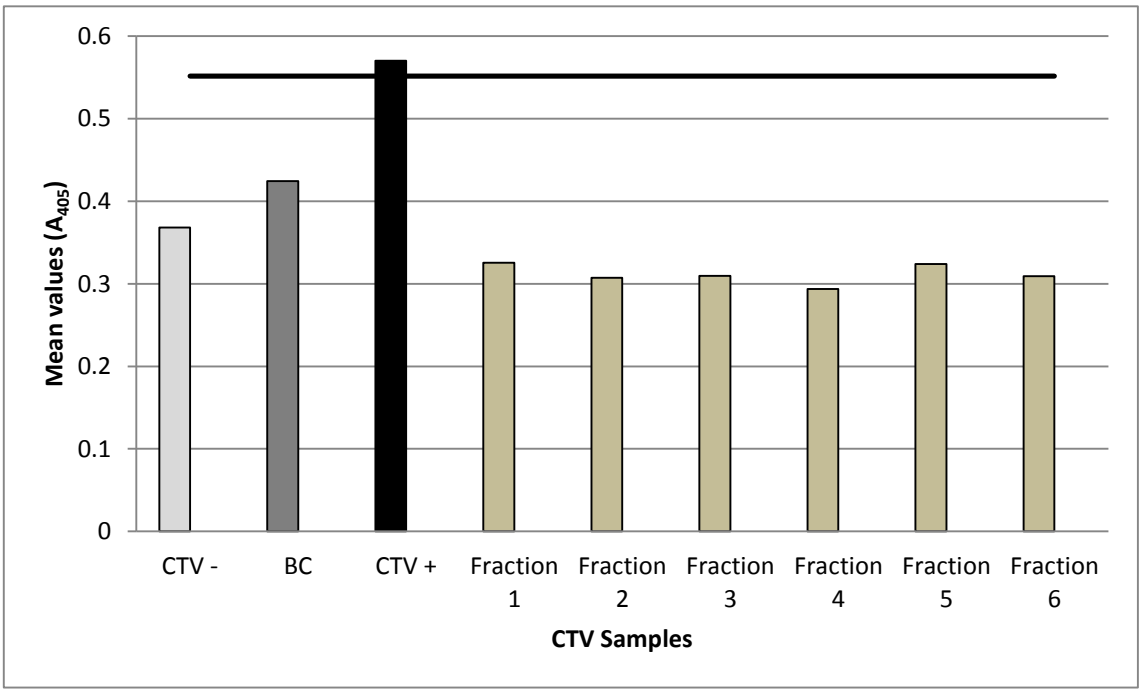


Figure 13: TAS-ELISA fraction concentration results obtained after partial CTV particle purification, by using polyclonal antibodies CTV 1052 and G604. The Y-axis represents the mean values obtained from each fraction evaluated. The X-axis represents the CTV samples (light Brown) screened during the process. Healthy (light Grey), buffer (dark Grey) and positive (Black) CTV controls were also included. Threshold (horizontal line) is based on the mean of the healthy control plus two times the standard deviation.

The results obtained from the three CTV inoculation trials are shown in Table 5. The table provides the symptom development detected on the inoculated Mexican Lime seedlings after the stated monthly intervals, as well as the CTV genotype presence. The CTV presence was determined using a generically primed RT-PCR analysis, followed by genotype-specific analysis on all CTV positive samples.

During the first bark-flap inoculation trial, only one Mexican Lime seedling (13-4158) appeared positive for CTV at 6 months PI (Table 7). No CTV was detected within the second bark-flap inoculation trail, while the stem-slash inoculation trail showed presence of CTV in four seedlings at 3 months PI inoculated at different inoculum dilutions (Table 7). Phylogenetic analysis was performed on all CTV positive results to differentiate the findings between each other and also already available CTV genomes (Table 8). However, during the 6 month PI. screening, none of the above positive seedlings tested CTV positive. All five CTV positive seedlings (Table 7) were screened again for CTV presence by means of a generically primed RT-PCR analysis. All results came back negative for CTV.

Table 6: CTV genotype status of samples collected and used in CTV isolation studies, determined using CTV genotype-specific RT-PCR analysis and comparison of sequences using Bayesian phylogenetic analysis.

Accession number	Generic	T36/RB	NZRB F2R2	NZRB F1R1	VT	B165	HA16-5	T3	T30
13-4142	+++	+++^{Y}	+++▲	+++^{Y}	+++▲	-	+++▲	+++■	-
13-4140	+++	+++^{X}	++▲	+++▲	+++▲	+++▲	-	-	-
13-4141	+++	+++^{X}	+++▲	+++▲	+++▲	+++▲	-	-	-
08-0049	+++	+++^{X}	+▲	-	-	+++▲	-	-	-
08-0035	+++	+++^{X}	+++▲	-	-	-	-	-	-
13-4138	-	-	-	-	-	-	-	-	-
13-4139	-	-	-	-	-	-	-	-	-

X = RB1; **Y** = RB2, i.e. no T36 was identified; **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Sample was negative in RT-PCR analysis; Generic positive samples indicates CTV presence only. **▲:** Confirmed by grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Appendix 2). ■: Sample 13-4142 (T3) will be discussed in text.

Table 7: CTV mechanical transmission (Bark-flap and Stem-slash) trail results obtained by means of CTV genotype specific RT-PCR analysis.

Inoculation method (trail number)▲	Form of sample material (tree accession number)	Symptoms	Positive (Accession number)/Number inoculated*	Dilution factor used	Months PI
Bark-flap (Inoculation trail 1)	Sucrose purified plant material (13-4142)	No apparent CTV-like symptoms	**1** (13-4158)/10	1:5	6 months PI
Stem-slash (Inoculation trail 3)	Crude plant sap (13-4140)	No apparent CTV-like symptoms	**1** (13-4236)/10	1:5	3 months PI
			1 (13-4243)/10	1:25	
			2 (13-4255; 13-4258)/10	1:125	

▲: Inoculation trail 2 did not provide any CTV positive results. *: 10 Seedlings were inoculated for each dilution made, in overall 40 seedlings were inoculated for each trail.

Table 8: CTV genotype status determined by specific RT-PCR analysis and sequences of the resultant amplicons using Bayesian phylogenetic analysis following isolation of CTV by stem-slash and bark- flap inoculation.

Accession number	Generic	T36/RB	NZRB F1R1	NZRB F2R2	VT	B165	HA16-5	T3	T30
13-4158■	+	-	-	-	-	-	-	-	-
13-4236Δ	+++	+++Y	-	-	++▲	+++▲	-	-	-
13-4243Δ	+	+Y	-	-	++▲	-	-	-	-
13-4255Δ	+++	+++X	+++▲	-	+▲	-	-	-	-
13-4258Δ	+++	+++Y	++▲	-	+++▲	-	-	-	-

T36 RT-PCR analysis samples: **X** = RB1; **Y** = RB2, i.e. no T36 was identified; **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Sample was negative in RT-PCR analysis; Generic positive samples indicates CTV presence only. **▲:** Confirmed by grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Appendix 2). ■: 6 months P.I. Δ: 3 months P.I.

2.4.2. Single aphid transmission trail results obtained.

All 4 trees from which aphids were collected in the field were positive for CTV, and all contained more than 3 CTV genotypes each (Table 9). Results obtained after the generic screening for CTV presence in the Mexican Lime seedlings 5 months PI. are shown in Table 10. Six Mexican Lime seedlings out of 167 seedlings tested positive for CTV presence. This was confirmed by CTV p33 gene region specific RT-PCR analysis (Table 11). Direct Sanger sequencing of p33 gene region (±980bp) identified the dominant genotype present (Figure 14), which was then confirmed by Illumina Next-Generation High-Throughput sequencing.

Table 9: Comparison of the results obtained between the CTV genotype specific RT-PCR analysis and sequences of the resultant amplicons using Bayesian phylogenetic analysis, prior to the Single Aphid Transmission (SAT) trail.

Accession number	Generic	T36/RB	NZRB F1R1	NZRB F2R2	VT	B165	HA16-5	T3	T30
14-7000	++	++^{X}	++▲	-	+++▲	+++▲	-	+++■	+▲
14-7001	+++	+++^{X}	+▲	-	+++▲	+++▲	+++▲	+++▲	++▲
14-7002	+++	+++^{Y}	++▲	+++▲	++■	+++▲	+++▲	+++▲	-
14-7003	+++	+++^{Y}	-	+++▲	+++▲	-	+++▲	+++▲	++▲

T36 RT-PCR analysis samples: **X** = RB1; **Y** = RB2, i.e. no T36 was identified; **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Sample was negative in RT-PCR analysis; Generic positive samples indicates CTV presence only. **▲:** Confirmed by grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Appendix 3). ■: Samples will be discussed in text.

Table 10: Generically-primed RT-PCR analysis results obtained from the Single Aphid Transmission (SAT) trails performed.

Sample	Number of SAT's performed	Symptoms	Positive (Accession number)
14-7000	67	No apparent CTV-like symptoms	0
14-7001	27		0
14-7002	40		**6** (14-7102, 14-7103, 14-7104, 14-7105, 14-7107, 14-7132)
14-7003	33		0

Table 11: Comparison of the results obtained between the CTV genotype specific RT-PCR analysis and sequences of the resultant amplicons using Bayesian phylogenetic analysis, after to the Single Aphid Transmission (SAT) trail.

Accession number	Generic	CTV p33 gene region	Potential CTV genotypic grouping
13-7102	+++	+++	RB
13-7103	++	+++	VT-like
13-7104	++	++	VT-like
13-7105	+	+++	VT-like
13-7107	++	+++	VT-like
13-7132	+	+++	VT-like

+++: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present). Bayesian phylogenetic dendrogram (Figure 14) shows confirmed grouping of sequenced amplicons based on analysis.

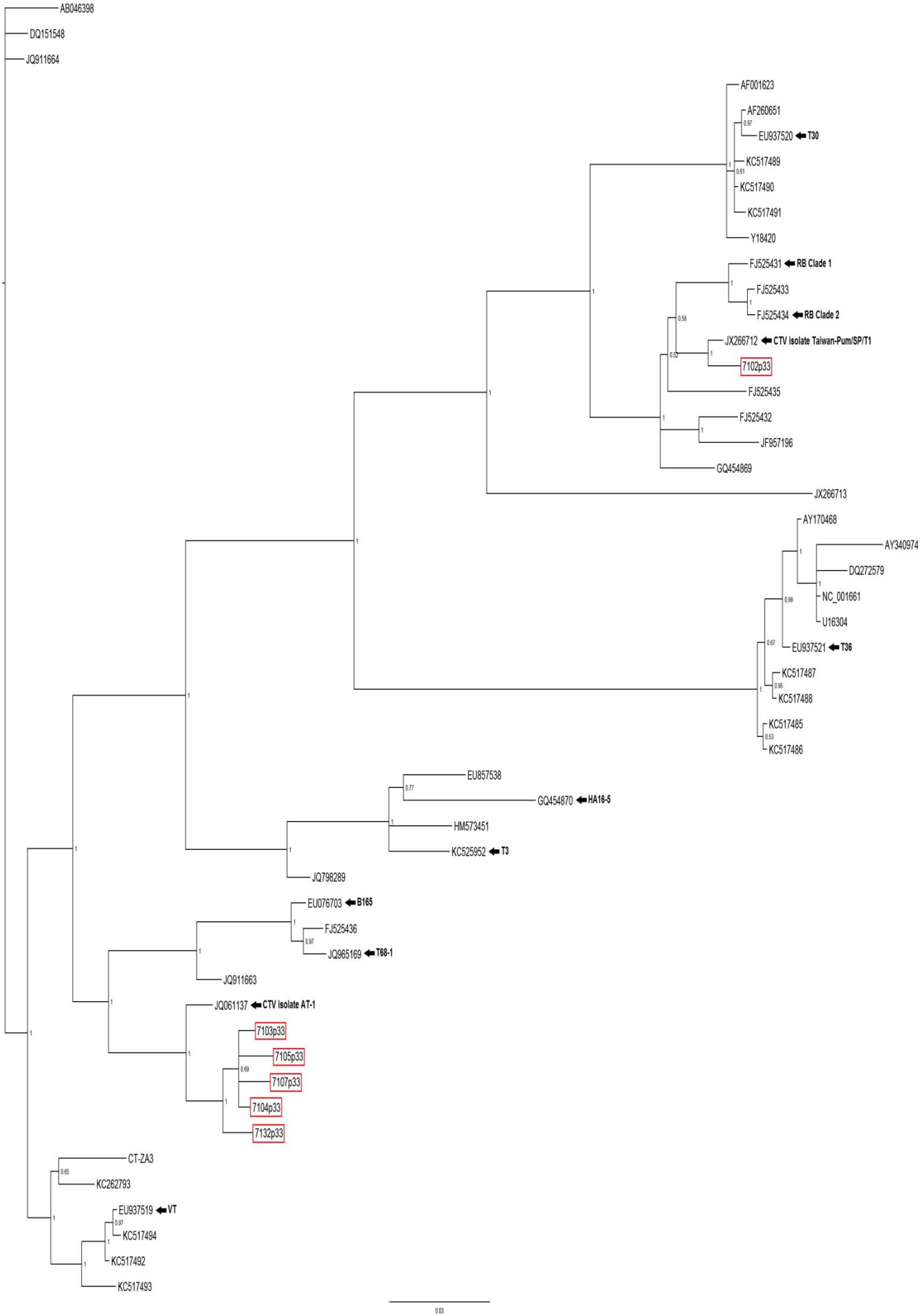


Figure 14: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the p33 RT-PCR of the 6 SAT positive Mexican Lime seedlings 14-7102, 14-7103, 14-7104, 14-7105, 14-7107 and 14-7132 and the 45 CTV reference genomes (red rectangles). The confidence levels are shown as posterior probability values at each node.

2.4.3. *Illumina Next-Generation High-throughput sequencing data analysis results of the six SAT positive Mexican Lime seedlings.*

A total of 396, 070 raw paired-end reads were obtained from the 6 CTV positive samples following SAT's (14-7102, 14-7103, 14-7104, 14-7105, 14-7107 [1 and 2], 14-7132). After all the necessary filtering processes, 394, 338 paired-end reads remained, which indicated an overall good quality dataset. Each of the datasets obtained showed high sequence duplication levels (>80%) as well as PHRED scores of 37 (Appendix 4), as confirmed by the FastQC analysis performed. Results of Illumina Next-Generation High-Throughput sequencing on SAT sub-isolates are shown in Table 12 and Figures 15 – 22 with each CTV sub-isolate discussed below in greater detail. Data on which this is based is presented in Appendix 5.

Table 12: Illumina data statistics, surrounding dataset quality scores (PHRED) and raw reads processing.

Sample name	Average PHRED score (/40)	Total reads (unfiltered)	Average read length (unfiltered)	Total reads (filtered)	% reads filtered	Average read length (filtered)
14-7102	37	88, 632	301	88, 272	99.59%	232.6
14-7103	37	50, 388	301	50, 137	99.50%	250.1
14-7104	37	69, 264	301	68, 917	99.50%	248.9
14-7105	37	131, 754	301	131, 227	99.60%	214.2
14-7107 [1]	37	11, 156	301	11, 101	99.51%	210.5
14-7107 [2]	37	4, 098	301	4, 081	99.59%	214.1
14-7132	37	40, 778	301	40, 603	99.57%	226.2

Two independent sequencing runs were conducted on sample 14-7107, as indicated in Table 12 as well as in Figure 19 and Figure 20. These sequencing results were used to determine the total raw pair-end reads as well as the pair-end reads obtained after the filtering step (Table 12). Final results of this sample was produced by combining the consensus lengths and the total read counts obtained for each of the sequence data sets obtained for 14-7107. The combination process is indicated in Appendix 5. The final result obtained (Figure 21) is very similar to that obtained when analyzing the separate sequencing results (Figure 19 and Figure 20).

2.4.3.1. Sample 14-7102.

The metrics obtained for the read mapping of this sub-isolate against a set of 45 cognate CTV reference sequences is shown in Figure 15. The majority of reads (27,072) generated from sub-isolate 14-7102 mapped against the stem-pitting severe CTV strain Taiwan-Pum/SP/T1 (JX266712) (Chen *et al.*, unpublished data), a RB genotype-like member (Appendix 5) and these almost fully mapped consensus length, 977bp, for the p33 gene region (±980bp) and the most significant coverage (7527,56). We also found that some sequence reads mapped to several more CTV strains were when an in depth analysis of the mapping data was done. However, very low total read counts and coverage were obtained for all of these (Appendix 5) with the following metrics observed for reads mapping to AT-1 [JQ061137] (Total read count = 1041; Consensus length = 977; Average coverage = 277.77), NZRB-M12 [FJ525431] (TRC = 620; CL = 475; AC = 181.94) and HA18-9 [GQ454869] (TRC = 426; CL = 933; AC = 129.33). A number of reads mapped over relatively large portions of other CTV references (e.g. NZRB-G90, NZRB-M17, B301 and Kpg3) but were present at extremely low read counts.

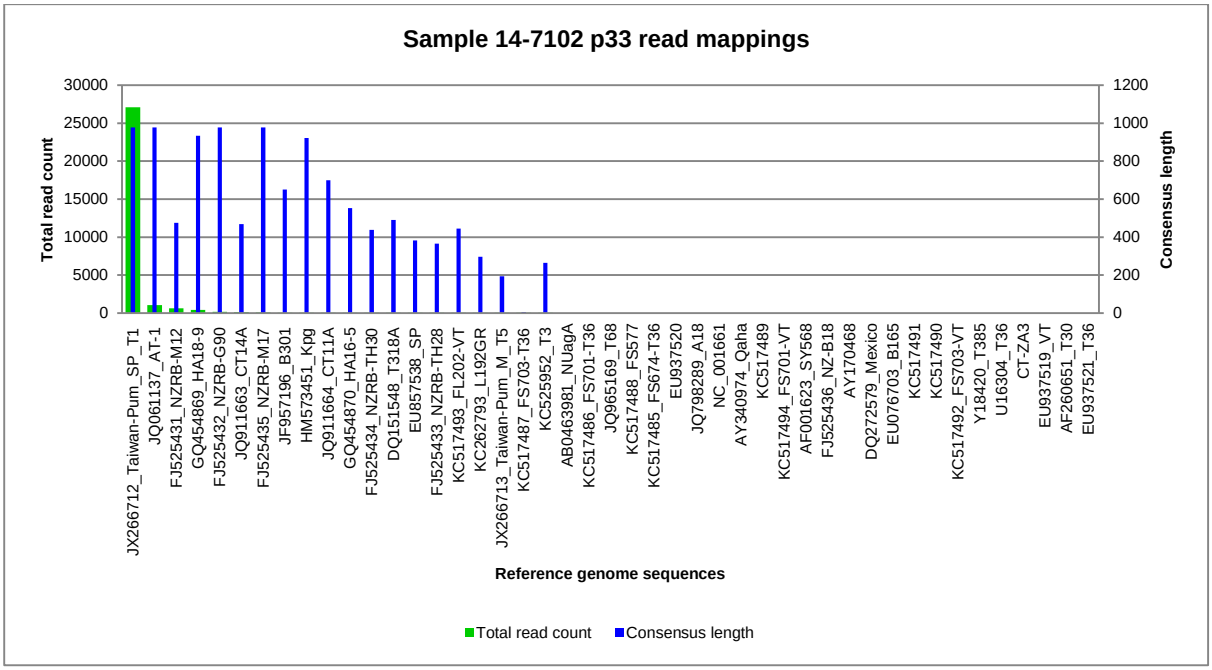


Figure 15: Distribution of reads of sub-isolate 14-7102 following the mapping against the p33 gene regions of 45 CTV reference genomes. Reference CTV genotypes are ranked by number of read counts obtained. The primary vertical axis (left) represents the number of reads (Green) while the secondary vertical axis represents the scale of consensus lengths (Blue) in nucleotides and is kept constant at 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.2. *Sample 14-7103.*

The mapping metrics of sequence reads obtained using CTV sub-isolate 14-7103 p33 gene region as template, against the cognate region of 45 CTV reference sequences is shown in Figure 16. The largest number of reads (16411) mapped to CTV strain AT-1 (JQ061137), (Wu *et al.*, unpublished data) a member of the VT-like genotype, and also almost fully mapping (977bp) its ±980bp cognate p33 gene region (±980bp) (Appendix 5). Furthermore, the average coverage over this region was very high (4397,32 reads). A more in depth look at the mapping data obtained showed sequence reads mapping to several more CTV strains, often at high consensus lengths but these generally had low total read counts and coverage, for example: Taiwan-Pum/SP/T1 [JX266712] (TRC = 2250; CL = 977; AC = 629.08), CT14A [JQ911663] (TRC = 1674; CL = 491; AC = 477.14) and CT11A [JQ911664] (TRC = 319; CL = 532; AC = 95.36) (Appendix 5). Extremely low total reads were present for several further CTV references (e.g. T318A, NZRB-M12, HA18-9 and NZRB-G90).

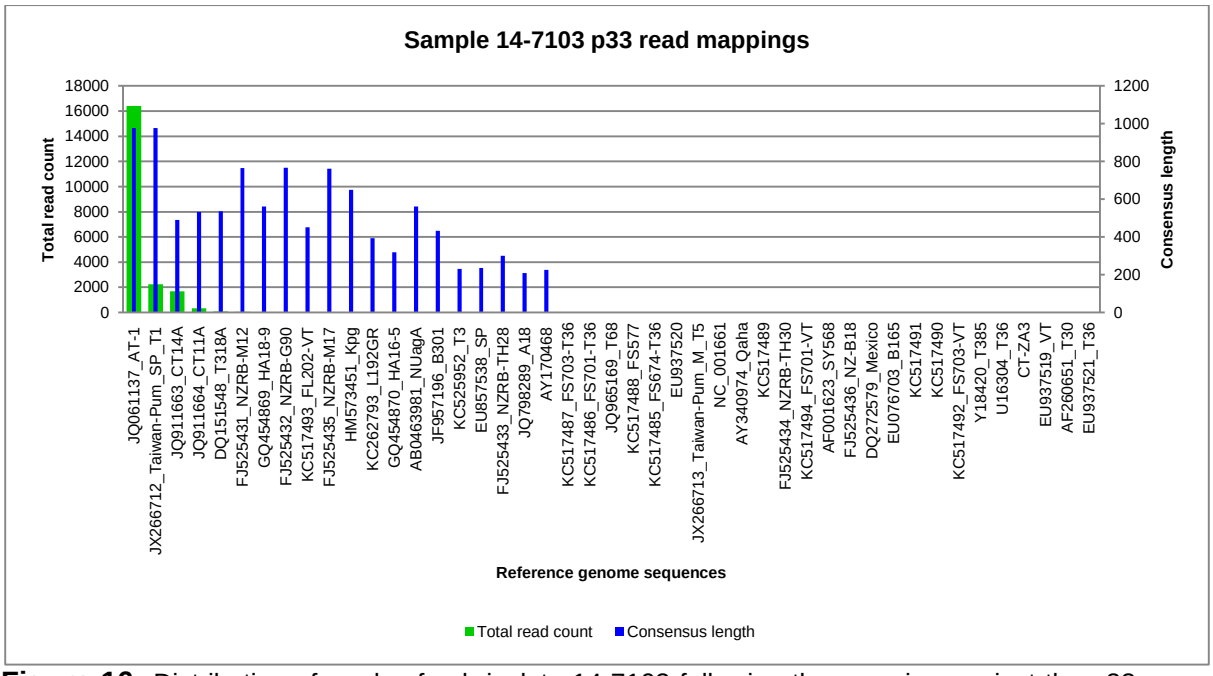


Figure 16: Distribution of reads of sub-isolate 14-7103 following the mapping against the p33 gene region of 45 CTV reference genomes. Reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.3. Sample 14-7104.

The metrics obtained for the read mapping of sub-isolate 14-7104 against a set of 45 p33 gene region edited CTV reference sequences is indicated in Figure 17. A total read count of 11,150 was generated against the CTV strain AT-1 (JQ061137) (Wu *et al.*, unpublished data) a member of the VT-like genotype (Appendix 5). An almost fully mapped consensus length, 977bp, for the p33 gene region (±980bp) was obtained together with an average coverage of 2985,90 reads. Apart from the distinct mapping result obtained for AT-1, we also found that sequence reads were also mapped to several more CTV strains during the mapping process. However, they generally had very low total read counts and coverage (Appendix 5) (CTV [AY170468]; TRC = 8203; CL = 981; AC = 2242.09;, Taiwan-Pum/SP/T1 [JX266712] TRC = 1901; CL = 977; AC = 535.94 and SP [EU857538]; TRC = 1691; CL = 786; AC = 492.1). A number of reads mapped over relatively large portions of other CTV references, showing low total read counts and varying consensus lengths. Amongst the sub-isolates screened, 14-7104 generated reads mapping to the largest number of the reference sequences used, as total read counts were obtained up to the CTV strain Taiwan-Pum/M/T5 (JX266713).

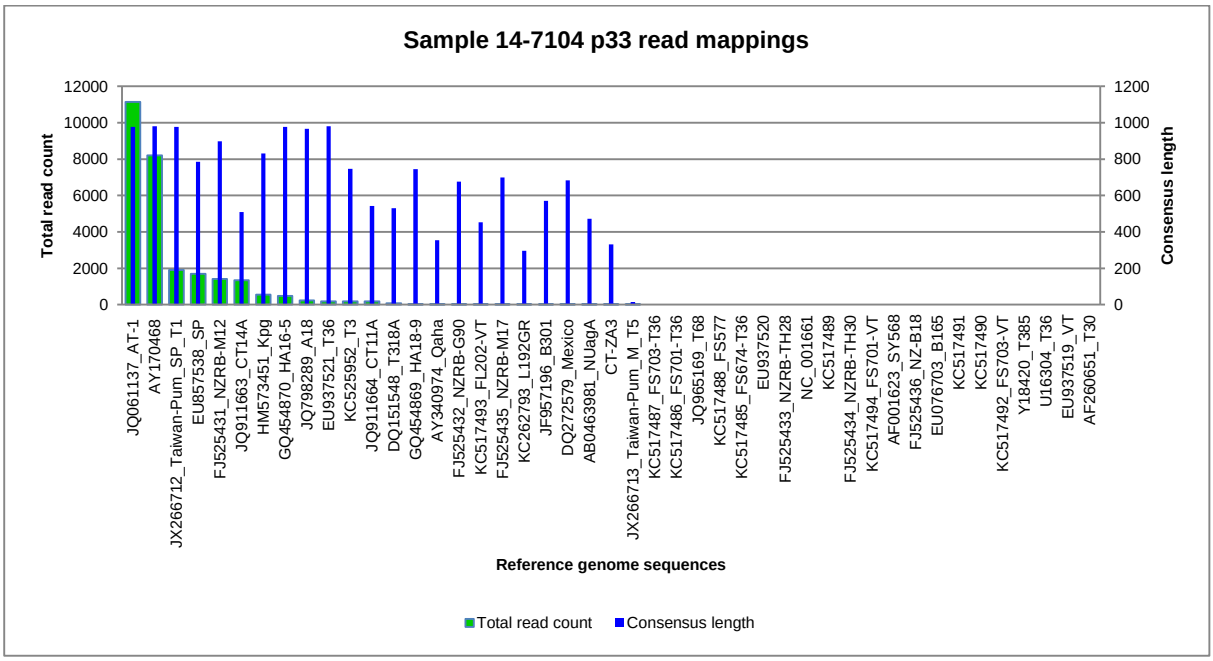


Figure 17: Distribution of reads of sub-isolate 14-7104 following the mapping against the p33 gene regions of 45 CTV reference genomes. Reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.4. Sample 14-7105.

Mapping distribution metrics of sequence reads for this sub-isolate against 45 cognate p33 gene CTV references are indicated in Figure 18. A largest total read count of 28,398 was obtained mapping against the CTV strain AT-1 (JQ061137) a member of the VT-like genotype (Wu *et al.*, unpublished data) (Appendix 5). An almost fully mapped consensus length, 977bp, for the p33 gene region (±980bp), together with an average coverage of 7632,12 reads are obtained for this sub-isolate. The mapping data collected from the analysis showed that several sequence reads also mapped to several more CTV strains, apart from AT-1. However, very low total read counts and coverage were obtained (Appendix 5) with the following metrics;CT14A [JQ911663] (TRC = 2767; CL = 520; AC = 790.64), CT11A [JQ911664] (TRC = 429; CL = 546; AC = 128.79) and T318A [DQ151548] (TRC = 56; CL = 533; AC = 16.31). Extremely low total read counts are obtained for several CTV reference sequences (e.g. FL202-VT to CT-ZA3) together with highly varying consensus lengths. Reference sequences AY170468 and Kpg have consensus lengths longer than the predominantly mapped CTV strain AT-1 at 981bp and 978bp respectively, however total reads obtained for these two sequences are extremely low.

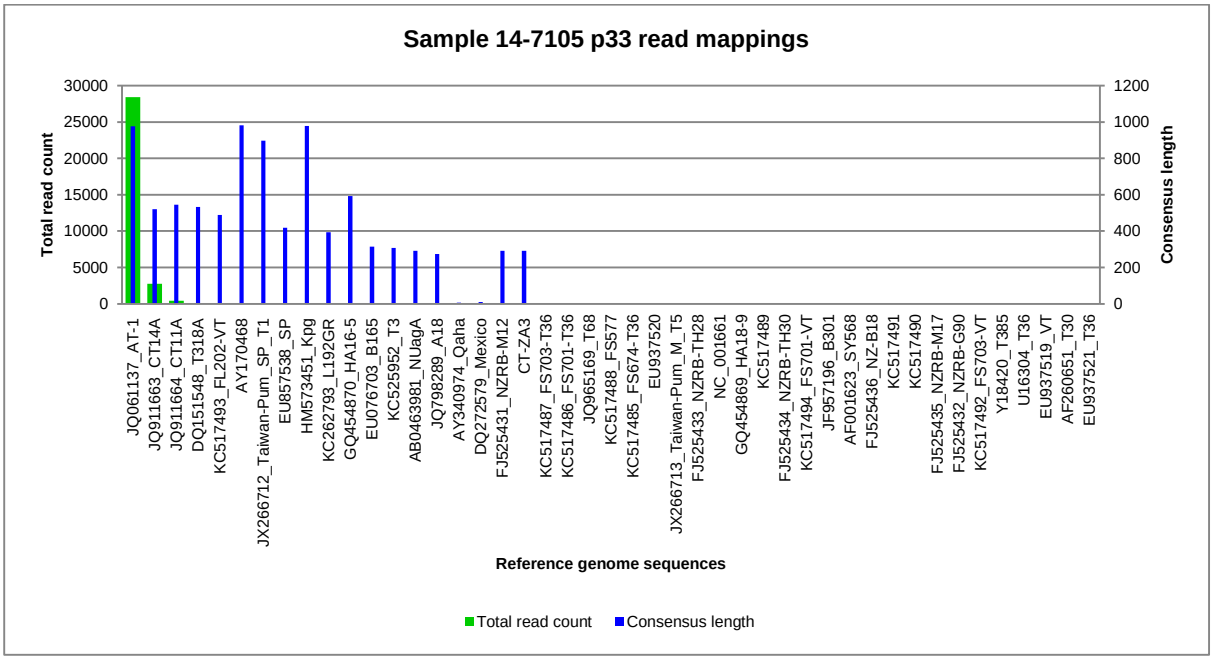


Figure 18: Distribution of reads of sub-isolate 14-7105 following the mapping against the p33 gene regions of 45 CTV reference genomes. Reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.5. Sample 14-7107 (1).

The metrics obtained for sub-isolate 14-7107 [1] reads against a set of 45 cognate p33 gene CTV reference sequences is indicated in Figure 19. The largest number of total read counts (1277) mapped to the CTV strain AT-1 (JQ061137), a member of the VT-like genotype (Wu *et al.*, unpublished data) (Appendix 5). An almost fully mapped consensus length, 977bp, for the p33 gene region (±980bp), and an average coverage of 339,51 reads are obtained for 14-7107 [1]. Sequence reads obtained through this analysis not only mapped to the CTV AT-1 strain, as it was also shown that some of the reads mapped to several other CTV strains too. Still, very low total read counts and coverage were obtained for these specific CTV strain sequences (Appendix 5). The following metrics are observed for reads mapping to Kpg [HM573451] (TRC = 313; CL = 977; AC = 85.04), CT14A [JQ911663] (TRC = 284; CL = 470; AC = 81.2) and SP [EU857538] (TRC = 85; CL = 942; AC = 24.72). Again, consensus lengths of varying heights are obtained for numerous reference sequences during reference mapping together with extremely low total reads varying from 32 to 2 reads.

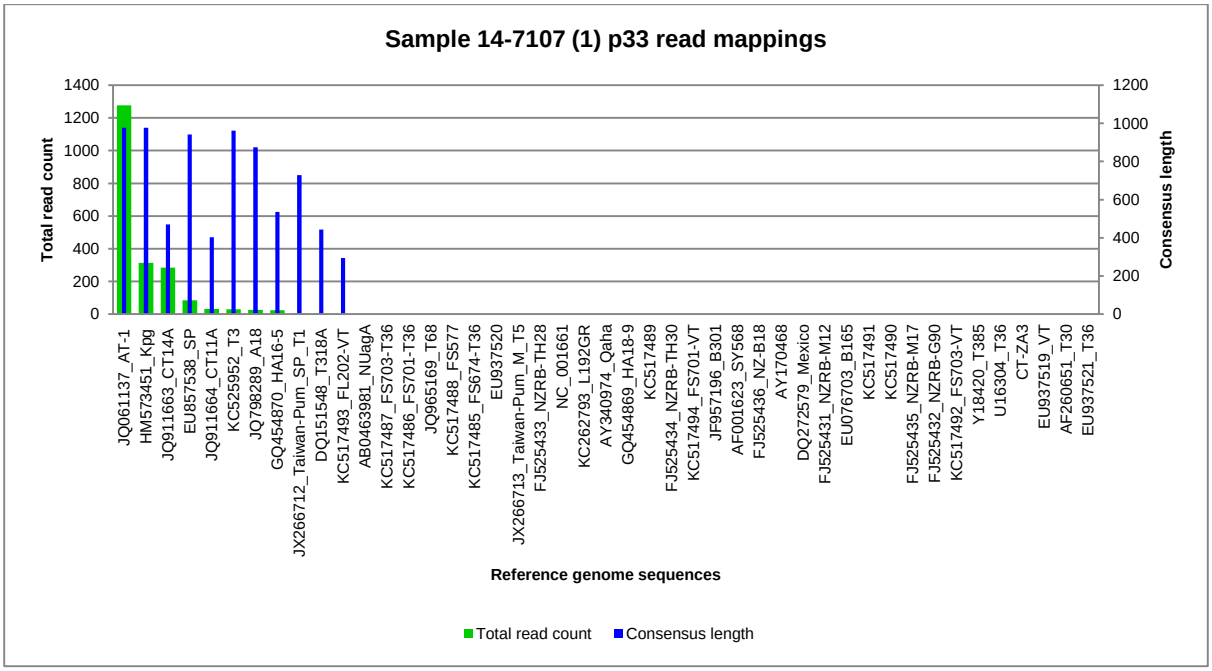


Figure 19: Distribution of reads of sub-isolate 14-7107 [1] following the mapping against the p33 gene regions of Read 45 CTV reference genomes. The mapped reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.6. Sample 14-7107 (2).

Mapping metrics of sequence reads obtained from CTV sub-isolate 14-7107 [2] against a set of 45 cognate p33 gene region CTV reference sequences are indicated in Figure 20. A total read count of 512 was generated against the CTV strain AT-1 (JQ061137), a member of the VT-like genotype (Wu *et al.*, unpublished data) (Appendix 5). A nearly complete mapped consensus length, 977bp, for the p33 gene region (±980bp) is obtained together with an average read coverage of 136,36 reads. It was also found that the sequence reads obtained mapped to several other CTV strains, apart from the dominantly mapped AT-1 CTV strain., Very low total read counts and coverage were obtained for these mappings (Appendix 5) in which the following metrics are observed for reads mapping to Kpg [HM573451] (TRC = 105; CL = 977; AC = 28.08), SP [EU857538] (TRC = 41; CL = 461; AC = 12.18) and CT14A [JQ911663] (TRC = 38; CL = 427; AC = 10.89). A number of reads mapped over relatively large portions of other CTV reference sequences, but were present at extremely low read counts.

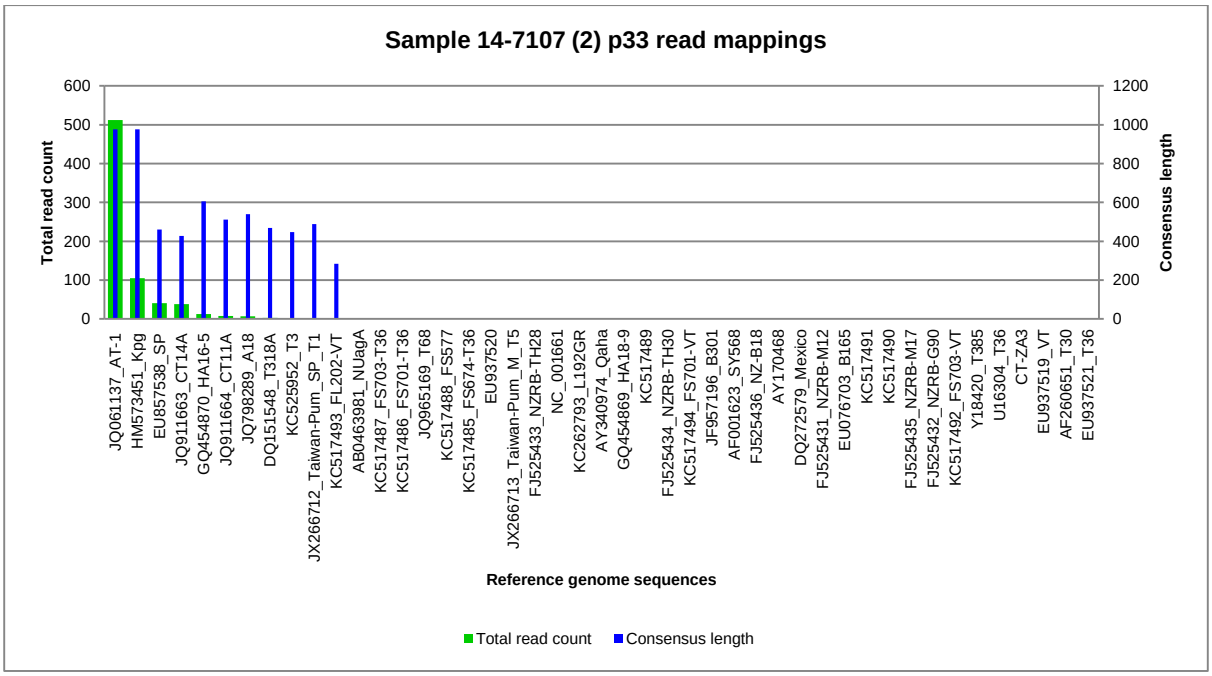


Figure 20: Distribution of reads of sub-isolate 14-7107 [2] following the mapping against the p33 gene regions of 45 CTV reference genomes. Reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.7. Sample 14-7107; combined results.

The mapping metrics obtained when combining the reads of the two sequencing runs of sub-isolate 14-7107 and mapping them against a set of 45 cognate p33 gene region CTV reference sequences is indicated in Figure 21. The consensus lengths and total read counts obtained for each sample was used to construct a combined graph for sub-isolate 14-7107. A total read count of 1789 was generated against the CTV strain AT-1 (JQ061137), a VT-like genotype (Wu *et al.*, unpublished data) (Appendix 5) with an almost fully mapped consensus length of 977bp for the p33 gene region (±980bp) is obtained. Additional reads mapped to several more CTV strains, although low total read counts and coverage were obtained (Appendix 5). The following metrics are observed for reads mapping to Kpg [HM573451] (TRC = 418; CL = 977), CT14A [JQ911663] (TRC = 322; CL = 448.5) and SP [EU857538] (TRC = 126; CL = 701.5).

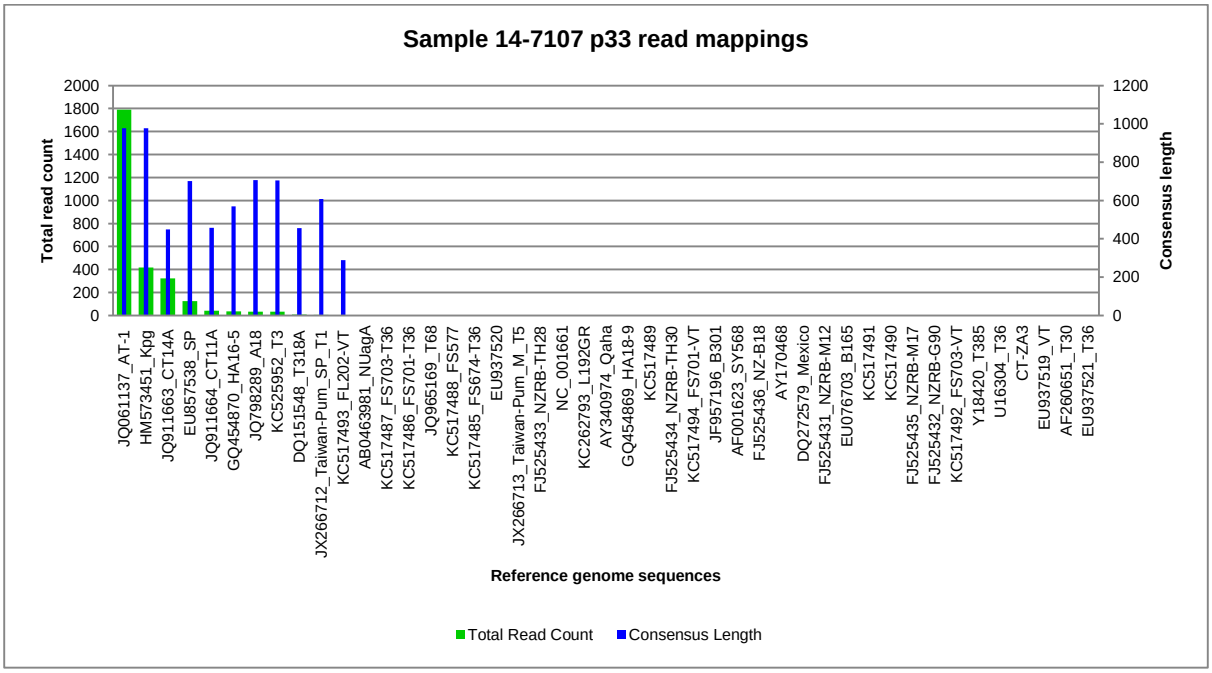


Figure 21: Distribution of reads of sub-isolate 14-7107 (Combined results) following the mapping against the p33 gene regions of 45 CTV reference genomes. Reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.4.3.8. *Sample 14-7132.*

Metrics obtained for the read mapping of sub-isolate 14-7132 against a set of 45 cognate p33 gene region CTV reference sequences is indicated in Figure 22. The majority of reads (11,378) mapped against the CTV strain AT-1 (JQ061137), a VT-like genotype (Wu *et al.*, unpublished data) (Appendix 5). Furthermore a nearly complete mapped consensus length of 977bp the p33 gene region (±980bp) and average read coverage of 3042,84 reads was obtained. Sequence reads mapping to several other CTV strains were also identified, all with low total read counts and average coverage's (Appendix 5). The following metrics are observed for reads mapping to CT14A [JQ911663] (TRC = 1030; CL = 491; AC = 296.37), CT11A [JQ911664] (TRC = 135; CL = 734; AC = 40.28) and T318A [DQ151548] (TRC = 44; CL = 531; AC = 13.3).

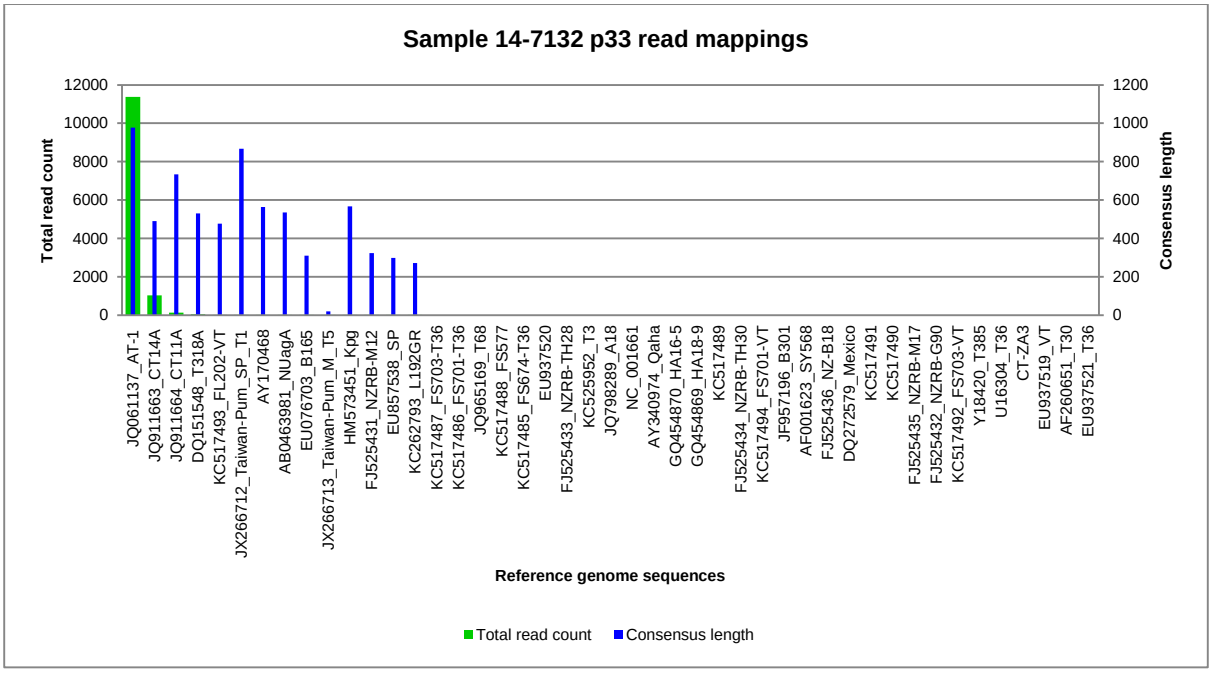


Figure 22: Distribution of reads of sub-isolate 14-7132 following the mapping against the p33 gene regions of 45 CTV reference genomes. Reference CTV genotypes are ranked from largest to smallest total read counts obtained. The primary vertical axis (left) represents the total reads (Green). The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1200 base pairs (i.e. larger than the ±980bp region sequenced and expected). The horizontal axis represents the 45 CTV reference genome sequences.

2.5. Discussion

As stated in the introduction, SAT's do not have a high success rate (Komazaki, 1994; Yokomi *et al.*, 1994; Broadbent *et al.*, 1996; Huang *et al.*, 2005; Tsai *et al.*, 2000; Barzegar *et al.*, 2010). CTV transmission rates do increase when high amounts of aphids are used per seedling (10 – 100 aphids per source plant) (Broadbent *et al.*, 1996; Barzegar *et al.*, 2010). Still, the varying transmission percentages are a good indication that SAT's are not the most effective and reliable way in obtaining pure CTV sources. Within this current study, the SAT's performed within this current study provided 6 positive CTV infections amongst 176 attempts (3.6%), confirming the relative inefficiency of the transmission but at the same time affirming its usefulness for virus isolation.

The disadvantage with this current transmission trail (using naturally infected sources) is that one does not know in advance which specific CTV genotype sources are being transferred by the naturally occurring viruliferous aphids (*T. citricida*) directly onto healthy seedlings. The most commonly used approach in obtaining single CTV genotypes or pure sources through SAT's is by maintaining an aphid colony on a healthy seedling, followed by allowing these aphids to feed on characterized CTV infected trees and then lastly, performing SAT's onto healthy seedlings. This last step usually aid in transferring specific CTV genotypic sources (Scott *et al.*, 2012; Müller *et al.*, 1974; Lubbe, *unpublished data*). This approach was not performed within this study as this procedure was confirmed to be highly insufficient in obtaining pure CTV sources. Also, the focus was not to isolate specific CTV genotypic strains, but to isolate any possible CTV genotype that was present within the sample. Therefore, the "shotgun" approach (*Materials and methods section 2.3.6.)* in transmitting naturally occurring aphids, directly from the orchards; to newly grown seedlings were attempted. However, for future work, the above mentioned approach can be included and compared to the approach used within this study.

In order to determine whether a homogenous CTV source was obtained in any of the SAT sub-isolates, amplification of the CTV p33 gene by RT-PCR using primers to a conserved region flanking genotype specific sequences (Read, *personal communication*), Sanger sequencing of product, and Bayesian phylogenetic analysis

were conducted. Results suggest that seedling 14-7102 is positive for the RB genotype, grouping with the CTV strain Taiwan-Pum/SP/T1 (JX266712) (Chen *et al.*, unpublished data). We were unable to determine in which clade of RB 14-7102 grouped as the p33 gene region is highly conserved between the two RB clades, and hence doesn't resolve them. This was confirmed by aligning references of the two RB clade reference sequences edited for the p33 gene region. Almost no sequence differences within that region, explaining the inability to distinguish between the two clades. The rest of the 5 seedlings (14-7103, 14-7104, 14-7105, 14-7107 and 14-7132) grouped in the dendrogram with CTV isolate AT-1 (JQ061137) (Wu *et al.*, unpublished data), which according to the Genbank information released by Wu *et al.* (unpublished data) forms part of the aphid-transmissible VT genotype from China. Further analysis of the six samples was done through Illumina Next-generation sequencing to give a more clear answer to the genotypic homogeneity status for each of these samples.

As Sanger sequencing suggested that the six sub-isolates contained a clear dominant genotype, these samples were all subjected to NGS sequencing which is better at determining the presence of minor variants within a virus population (Read and Pietersen, 2015). The NGS data obtained provided a more in-depth description of the CTV genotypes, and their strains, present. Reference mapping performed on each of the seedling sequences (filtered) confirmed that seedling 14-7102 primarily contained Taiwan-Pum/SP/T1 (JX266712) an RB variant, while the other 5 seedlings (14-7103, 14-7104, 13-7105, 14-7107 and 14-7132) again primarily contained the CTV isolate AT-1 (JQ061137), as detected in the p33 RT-PCR analysis. However, even though Taiwan-Pum/SP/T1 (JX266712) and the CTV isolate AT-1 (JQ061137) represented the two dominant CTV sources for the respective seedlings, additional reads mapping to other references were obtained during the sequence analysis of each seedling.

In the graphs produced for the 6 seedlings, consensus lengths of varying percentage of the total length together with extremely low read counts were obtained for a number of CTV reference sequence, other than the two dominant CTV sources, Taiwan-PUM/SP/T1 (JX266712) and AT-1 (JQ061137). Close examination of the results indicated that in the case of 14-7102 or 14-7103 the additional mappings observed were mainly to CTV reference sequences (strains) of the other dominant

CTV sources (for example if AT-1 was dominant, then the next most prevalent reads were to Taiwan-Pum/Sp/T1). In the case of sub-isolates 14-7104, 14-7105 and 14-7107 where AT-1 was the dominant strain there were a fair amount of reads to T36, CT-14A, and Kpg3 respectively. Often however additional reads were present of strains closely related to the dominant strain. While this may indicate that these CTV strains are present at extremely low concentrations, the relatedness suggests that is more likely that this reflects the quasispecies nature of the virus population or minor errors introduced through PCR amplification or during the sequencing reaction. RNA-dependent RNA Polymerase slippage) is a known source of sequencing error (Albiach-marti, 2013; Ross *et al.*, 2013). This may alter the nucleotide sequences sufficiently that they resemble closely related CTV strains over relatively short sequence areas. Therefore, using a polymerase enzyme with proofreading capabilities in future analysis will provide more accurate results on PCR analysis performed well as Illumina analysis results. The 3'-terminal half of the CTV genome, where the p33 gene is located, is well known for being a highly conserved region (Moreno *et al.*, 2008) with a sequence identity ranging between 89 – 94.8% (Folimonova, 2013), explaining why this region is not used usually used for CTV genotype screening. In this study however the association of p33 gene sequences with specificity of super-infection exclusion (Folimonova, 2013) made it imperative that we characterize the variability of CTV sources with regards this genome region in South Africa. The results from the CTV p33 RT-PCR analysis and Bayesian phylogenetic analysis confirm the conserved nature of this region as the clustering of the different CTV genotypes were less defined. A single nucleotide difference in this area may result in it appearing more like a different strain of CTV. In the case of sub-isolate 14-7104 it is evident that even though the CTV isolate AT-1 (JQ061137) is the dominant CTV source, a mixture of several more distantly related CTV genotypic sources are also present.

Some of the low read counts obtained in samples may be attributable to inter sample contamination as parallel library preparations were done prior to the NGS reactions or to index leaching (Solonenko *et al.*, 2013; Oyola *et al.*, 2012) a phenomenon observed in our laboratory. Studies to determine confidence thresholds are currently being developed in our laboratory (Read, personal communication). Also, routine NGS analysis can introduce cross-contamination between samples as

reagents used during the analysis can become contaminated with time. Thus, making aliquots of all reagents used as well as allocating certain lab equipment for only NGS analysis can aid in decreasing cross-contamination,

In conclusion, the NGS analysis suggests that none of the 6 sub-isolates contain a completely homogenous CTV source. Repeating one or all of the CTV inoculation trials performed within the study, using plant material from these 6 seedlings, can help create an artificial genetic bottleneck and aid in a further potential reduction of CTV genotypes present.

While theoretically possible, the attempt to isolate pure CTV genotypic sources through bark-flap and stem-slash inoculations proved to be extremely difficult. The Sucrose purification procedure (Robertson *et al.*, 2005) used to purify sample 13-4142 (bark-flap inoculation trail 1) provided partially purified (virion) fractions. As indicated in Figure 13, very low CTV concentrations, for the partially purified fractions, were confirmed through TAS-ELISA analysis. Nevertheless, 5 positive CTV infections were obtained after the respective inoculation trials. However, we believe that extreme temperature increases within the greenhouse on 2014-02-22 (40°C) and 2014-02-23 (48°C), one month prior to the 6 months P.I. sampling date (2014-03-23), caused a massive reduction in titer of CTV present or potentially even a thermotherapy event, as confirmed with the 6 month RT-PCR screening process. Numerous attempts in screening for CTV presence within these formerly positive seedlings, with the old and newly extracted RNA, were performed. All analyses yielded negative results for CTV. Evidence of CTV concentration reduction by high temperatures within different citrus cultivars has be found within research studies focusing on monthly temperature variations in the field over an approximate 1 to 2 year screening process (Dodds *et al.*, 1987; Mathews *et al.*, 1997; Zhou *et al.*, 2002), A recent study on season variation of CTV in Pakistan by Abbas *et al.* (2013) confirmed that viral titres do decrease during seasonal changes. It was also found that the viral titre differs within different parts of the citrus trees during these seasonal changing months. Temperature changes can also play a role in altering the composition of CTV isolates as well as strain presence as well as CTV occurrences within certain citrus cultivars (Dodds *et al.*, 1987; Broadbent *et al.*, 1991; Mathews *et al.*, 1997). However, in both studies performed by Mathews *et al.* (1997) and Zhou *et al.* (2002), they were able to detect CTV by means of PCR analysis

even though the virus titre was extremely low. Although some CTV strains are able to adapt to high temperatures, some prefer cooler environmental conditions (Abbas *et al.*, 2005). Therefore, environmental changes will not only have an effect on the tree itself but also on the virus present. In our case, factors such as the size of the seedlings prior to inoculations, root development and potential biotic and abiotic stresses unique to greenhouses can also play a role in the results obtained when high temperature peaks occur.

The use of genotype specific primers in order to determine which CTV genotypes are present in the sub-isolates obtained during this study yielded two interesting results. Firstly, samples apparently T36 positive by genotype specific primers all yielded amplicons with sequences grouping with either one of two RB genotypic clades (RB clade 1 or 2). It was confirmed by Roy *et al.* (2010) that regions of identical sequences were shared between the T36 primer set developed and some of the RB strains (NZRB-M17 (FJ525435) and NZRB-TH30 (FJ525434)) within ORF1a. This suggests that the T36 primer is in fact poly-specific within ORF1a, detecting both T36-like strains as well as RB-like strains and therefore, requires direct Sanger sequencing of the resultant amplicon to differentiate the two.

Furthermore, Mexican Lime seedlings that tested positive for the CTV B165 genotype prior (08-0049, 13-4140 and 13-4141) and after (13-4236) the inoculation trails showed to specifically group with CTV strain CT-ZA3 (KC333868) (Zablocki & Pietersen, 2014) which groups with both the B165 and T68-1 genotypic clades (Appendix 2). From these two phylogenetic trees, one could also see that the sequences amplified by the B165 primer set used cannot differentiate between the two genotypes, B165 and T68-1. Furthermore, by aligning the whole genomes of B165 and T68-1 and looking at the specific primer binding area (2,214 – 2,633), very few differences were detected hence the primers will also not differentiate between them. Lastly, contrasting results were obtained for seedling 13-4142 which could have been due to experimental errors made during the analysis of this sample. Possible sequence editing errors could have been made as seedling 13-4142 tested positive for RB Clade 1, however, grouped within RB Clade 2 instead of Clade 1 (Appendix 2,) Also, even though 13-4142 tested positive for T3 through RT-PCR analysis, the sequence of the resulting amplicon grouped on its own, outside the T3 genotypic cluster (Appendix 2), potentially indicating a new CTV strain. Unfortunately

the source plant material collected for sample 13-4142 was a mixture of different citrus trees collected in the Schoemanskloof/Malelane region and hence it was not possible to return to a specific tree to confirm this result.

2.6. References

Abbas, M., Khan, M. M., Mughal, S. M. and Khan, I. A. 2005. “Prospects of classical cross protection technique against Citrus tristeza closterovirus in Pakistan.” *Horticultural Science (Prague)* 32(2): 74–83.

Abbas, M., Khan, M. M., Mughal, S. M. and Ji, P. 2013. “Characterization and assessment of seasonal variation of Citrus tristeza closterovirus (CTV) in citrus in Pakistan.” *Journal of Food, Agriculture and Environment* 11(3 & 4): 1063–1068.

Abdelmaksoud,H.M. and Gamal El-din, A.S. "Complete Nucleotide Sequence of Citrus Tristeza Virus Qaha isolate." *Unpublished.* Genbank Submitted (07-JUL-2003) (AY340974.1) Plant Virology Research Department, Plant Pathology Research Institute, 9 Gamma Str, Giza 12619, Egypt.

Albiach-martı, M. Mawassi, R., M., Gowda, S., Satyanarayana, T. , Hilf, M. E., Shanker, S., Almira, E. C., Vives, M. C., Lopez, C., Guerri, J., Flores, R., Moreno, P., Garnsey, S. M. and Dawson, W. O. 2000. “Sequences of Citrus Tristeza Virus Separated in Time and Space Are Essentially Identical.” *Journal of Virology* 74 (15): 6856–6865. doi:10.1128/JVI.74.15.6856-6865.2000.Updated.

Albiach-marti, M. R. 2013. “The Complex Genetics of Citrus tristeza virus.” *InTech, Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications* (Chapter 1): 1–26. http://dx.doi.org/10.5772/56122.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. 1990. “Basic Logal Alignment Search Tool.” *Journal of Molecular Biology* 215: 403–410.

Barzegar, A., Rahimian, H. and Sohi, H. H. 2010. “Comparison of the Minor Coat Protein Gene Sequences of Aphid-Transmissible and -nontransmissible Isolates of Citrus Tristeza Virus.” *Journal of Plant Pathology* 76, 143–151. doi:10.1007/s10327-009-0216-7.

Biswas, K. Tarafdar, K., A. and Sharma. S. K. 2012. “Complete Genome Sequence of Mandarin Decline Citrus Tristeza Virus of the Northeastern Himalayan Hill Region of India : Comparative Analyses Determine Recombinant.” *Archives of Virology* 157: 579–583. doi:10.1007/s00705-011-1165-y.

Brlansky, R. H., Damsteegt, V. D., Howd, D. S. and Roy, A. 2003. “Molecular Analyses of Citrus tristeza virus Subisolates Separated by Aphid Transmission.” *Plant Disease* 87: 397–401.

Broadbent, P., Bevington, K.B. and Coote, B.G. 1991. Control of stem-pitting of grapefruit in Australia by mild strain protection. In “Proceedings of the 11th Conf. of the IOCV.” (R.H. Brlansky, R.F. Lee, L.W. Timmer, Eds), pp. 64-70, IOCV, Riverside, California.

Broadbent, P., Brlansky, R. H. and Indsto, J. 1996. “Biological Characterization of Australian Isolates of Citrus Tristeza Virus and Separation of Subisolates by Single Aphid Transmissions.” *Plant Disease* 80(3), 329–333.

Chen, Y-H., Feng, Y-C., Hung, T-H. and Su, H-J. "Genomic Analysis for Different Strains of Citrus tristeza virus in Taiwan." *Unpublished.* Genbank Submitted (03-JUL-2012) (JX266713.1) Plant Pathology and Microbiology, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Da'an Dist.d., Taipei City 106, Taiwan.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. 2012. jModelTest 2: More Models, New Heuristics and Parallel Computing. *Nature Methods* 9(8), pp. 772-772.

Dodds, J. A., Jarupat, T., Lee, J. G. and Roistacher, C. N. 1987. “Effects of Strains, Host, Time of Harvest, and Virus Concentration on Double-Stranded RNA Analysis of Citrus Tristeza Virus.” *Phytopathology* 77(3): 442–447.

Folimonova, S. Y., Robertson, C. J., Shilts, T., Folimonov, A. S., Hilf, M. E., Garsney, S. M. and 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(3): 1314–1325.

Folimonova, S. Y. 2013. “Developing an understanding of cross-protection by Citrus tristeza virus.” *Virology* 4(76): 1–9.

Guindon, S. and Gascuel, O. 2003. “A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood.” *Systematic Biology* 52 (5): 696–704. doi:10.1080/10635150390235520.

Hall, T. A. 1999. “BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT.” *Nucleic Acid Symposium Series* 41: 95–98.

Harper, S. J., Dawson, T. E. and Pearson, M. N. 2009. “Complete Genome Sequences of Two Distinct and Diverse Citrus Tristeza Virus Isolates from New Zealand.” *Archives of Virology* 154: 1505–1510. doi:10.1007/s00705-009-0456-z.

Harper, S. J., Dawson, T. E. and Pearson, M. N. 2010. “Isolates of Citrus Tristeza Virus That Overcome Poncirus Trifoliata Resistance Comprise a Novel Strain.” *Archives of Virology* 155: 471–480. doi:10.1007/s00705-010-0604-5.

Harper, S. J. and Hilf, M. E. "Citrus tristeza virus: Evolution of complex and varied genotypic groups." *Unpublished.* Genbank Submitted (23-JAN-2013) (KC517485.1 - KC517494.1) CREC-IFAS, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA.

Hasegawa, M., Kishino, H. and Yano, T. 1985. “Dating of the human-ape splitting by a molecular clock of mitochondrial DNA.” Journal of Molecular Evolution, 22(2), 160-74.

Hilf, M. E., Mavrodieva, V. A. and 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(8): 909–917.

Hilf, M. E., Harper, S. J. and Dawson, W. O. "Citrus tristeza virus: strains, genotypes and isolates." *Unpublished*. Genbank Submitted (23-APR-2012) (JQ965169.1) CREC-IFAS, University of Florida, 700 Experiment Stn Road, Lake Alfred, FL 33850, USA.

Hilf, M. E. and Harper, S. J. "Citrus tristeza virus: Evolution of complex and varied genotypic groups." *Unpublished.* Genbank Submitted (24-JAN-2013) (KC525952.1) CREC-IFAS, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA.

Huang, Z., Rundell, P. A., Guan, X. and Powell, C. A. 2005. "Evaluation of the Transmission of Different Field Sources of Citrus Tristeza Virus and the Separation of Different Genotypes by Single Brown Citrus Aphids." *HortScience* 40(3): 687–690.

Karasev, A. V., Hilf, M. E., Garnsey, S. M. and Dawson, W. O. 1997. "Transcriptional Strategy of Closteroviruses : Mapping the 5J Termini of the Citrus Tristeza Virus Subgenomic RNAs." *Journal of Virology* 71 (8): 6233–6236.

Katoh, K., Kuma, K-I., Toh, H. and Miyata, T. 2005. "MAFFT Version 5: Improvement in Accuracy of Multiple Sequence Alignment." *Nucleic Acids Research* 33 (2): 511–518. doi:10.1093/nar/gki198.

Katoh, K., Misawa, K., Kuma, K-I. and Miyata, T. 2002. "MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform." *Nucleic Acids Research* 30 (14): 3059–3066.

Komazaki, S. 1994. "Ecology of Citrus Aphids and Their Importance to Virus Transmission." *Japan Agricultural Research Quarterly* 28: 177–184.

Lee, R. F., Herron, C. M., Mirkov, T. E. and da Graca, J. V. 2006. "Citrus tristeza virus transmission by the Toxoptera citricida vector : In vitro acquisition and transmission and infectivity immunoneutralization experiments." *Journal of Virological Methods* 134: 205–211.

Mathews, D. M., Riley, K. and Dodds, J. A. 1997. “Comparison of Detection Methods for Citrus Tristeza Virus in Field Trees During Months of Nonoptimal Titer.” *Plant Disease* 81: 525–529.

Matos, L. A., Hilf, M. E., Cayetano, X. A., Feliz, A. O., Harper, S. J. and Folimonova, S. Y. 2013. “Dramatic Change in Citrus tristeza virus Populations in the Dominican Republic.” *Plant Disease* 93: 339–345.

Melzer, M. J., Borth, W. B., Sether, D. M., Ferreira, S., Gonsalves, D. and Hu, J. S. 2010. “Genetic Diversity and Evidence for Recent Modular Recombination in Hawaiian Citrus Tristeza Virus.” *Virus Genes* 40: 111–118. doi:10.1007/s11262-009-0409-3.

Meneghini, M. 1946. "Sôbre a natureza e transmissibilidade da doença "tristeza" dos citrus." *O Biológico* 12: 285–287.

Moreno, P., Ambrós, S., Albiach-martí, M. R., Guerri, J. and Peña, L. 2008. “Review: Plant diseases that changed the world. Citrus tristeza virus: a pathogen that changed the course of the citrus industry.” *Molecular Plant Pathology* 9(2): 251–268.

Müller, G. W., Costa, A. S., Kitajima, E. W. and Camargo, I. J. B. 1974. "Additional evidence that tristeza multiplies in Passiflora spp." In: Proc. 6th Conf. IOCV, 75-78. IOCV, Riverside, CA.

Oyola, S. O., Otto, T. D., Gu, Y., Maslen, G., Manske, M., Campino, S., Turner, D. J., Macinnis, B., Kwiatkowski, D. P., Swerdlow, H. P. and Quail, M. A. 2012. “Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes.” *BMC Genomics* 13:1.

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. 1994. "Nucleotide sequence and organization of eight 3' open reading frames of the citrus tristeza closterovirus genome." *Virology.* 199 (1), 35-46.

Quiroz, J. D. C., Mendoza, A., Cruz, M. A., Fernandez, S. and Pena, A. "The Complete RNA genome sequence of Citrus Tristeza Virus from Mexico." *Unpublished.* Genbank Submitted (01-NOV-2005) (DQ272579.1) Centro de Biotecnologia Genomica, Instituto Politecnico Nacional, Blvd. del Maestro SN, Reynosa, Tamaulipas 88710, Mexico.

Rambaut, A. 2009. FigTree: Tree Figure Drawng Tool. University of Edinburgh: Institute of Evolutionary Biology.

Rambaut, A. and Drummond, A. 2009. Tracer: MCMC Trace Analysis Tool. University of Edinburgh: Institute of Evolutionary Biology and University of Auckland: Department of Computer Science.

Read, D. A., and Pietersen, G. 2015. "Genotypic Diversity of Citrus Tristeza Virus within Red Grapefruit , in a Field Trial Site in South Africa." *European Journal of Plant Pathology*, 1–15. doi:10.1007/s10658-015-0631-x.

Robertson, C. J., Garsney, S. M., Satyanarayana, T., Folimonova, S. and Dawson, W. O. 2005, "*Efficient Infection of Citrus Plants with Different Cloned Constructs of Citrus tristeza virus Amplified in Nicotiana benthamiana Protoplasts."* Sixteenth IOCV Conference, 2005 – Citrus Tristeza Virus. Department of Plant Pathology, Citrus Research and Education Center, University of Florida, Lake Alfred, Florida 33850, USA, pp 187-194.

Rocha-Peña, M. A. and Lee, R. F. 1991, "Serological techniques for detection of citrus tristeza virus." J*ournal of Virological Methods*. 34(3), pp. 311-31.

Rodríguez, F., Oliver, J. L., Marín, A.and Medina, J. R. 1990. "The general stochastic model of nucleotide substitution." Journal of *Theoretical Biology*, 142(4), 485-501.

Roistacher, C. N., da Graça, J. V. and Müller, G. W. 2010. "Cross Protection Against Citrus tristeza virus - a Review." In *Proceedings, 17th Conference, 2010 – Citrus Tristeza Virus*, , p. 1–27.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., Larget, B., Lui, L., Suchard, M. A. and Huelsenbeck, J. P. 2012. “MrBayes 3 . 2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space.” *Systematic Biology* 61 (3): 539–542. doi:10.1093/sysbio/sys029.

Ross, M. G., Russ, C., Costello, M., Hollinger, A., Lennon, N. J., Hegarty, R., Nusbaum, C. and Jaffe, D. B. 2013. “Characterizing and measuring bias in sequence data.” *Genome Biology* 14(5), R51. http://genomebiology.com/2013/14/5/R51.

Roy, A., Ananthakrishnan, G., Hartung, J. S. and Brlansky, R. H. 2010. “Development and Application of a Multiplex Reverse-Transcription Polymerase Chain Reaction Assay for Screening a Global Collection of Citrus Tristeza Virus Isolates.” *Phytopathology* 100 (10): 1077–1088.

Roy, A. and Brlansky, R. H. 2010. “Genome Analysis of an Orange Stem Pitting Citrus Tristeza Virus Isolate Reveals a Novel Recombinant Genotype.” *Virus Research* 151: 118–130. doi:10.1016/j.virusres.2010.03.017.

Roy, A., Choudhary, N., Hartung, J. H. and Brlansky, R. H. "Trifoliate resistance breaking (RB) Citrus tristeza virus genotype specific RT-PCR and complete genome sequence analysis of an unassigned isolate discovered the presence of RB genotype from Puerto Rico." *Unpublished.* Genbank Submitted (12-MAY-2011) (JF957196) Plant Pathology, Citrus Research and Education Center, University of Florida, 700, Experiment Station Road, Lake Alfred, FL 33850, USA.

Ruiz-Ruiz, S., Moreno, P., Guerri, J. and Ambros, S. 2006. "The complete nucleotide sequence of a severe stem pitting isolate of Citrus tristeza virus from Spain: comparison with isolates from different origins." Archives of Virology. 151 (2), 387-398.

Satyanarayana, T., Gowda, S., Boyko, V. P., Albiach-Martı, M. R., Mawassi, M., Navas-castillo, J., Karasev, A. V., Dolja, V., Hilf, M. E., Lewandowski, D. J., Moreno, P., Bar-Joseph, M., Garsney, S. M. and Dawson, W. O. 1999. “An Engineered Closterovirus RNA Replicon and Analysis of Heterologous Terminal Sequences for Replication.” *Proceedings of the National Acadamy of Science USA* 96: 7433–7438.

Satyanarayana, T., Bar-Joseph, M., Mawassi, M., Albiach-Martı, M. R., Ayllon, M. A., Gowda, S., Hilf, M. E., Moreno, P., Garnsey, S. M. and Dawson, W. O. 2001. "Amplification of Citrus Tristeza Virus from a cDNA Clone and Infection of Citrus Trees." *Virology*, *280*, 87–96. doi:10.1006/viro.2000.0759

Scott, K. A., Hlela, Q., Zablocki, O., Read, D., van Vuuren, S., and 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.

Sindhvajiva, K., Paradornuwat, A., Chowpongpang, S. and Vichukit, V. "Genome of Citrus Tristeza Virus Isolate A18 from Mandarin Citrus in Thailand." *Unpublished*. Genbank Submitted (19-MAR-2012) (JQ798289.1) Department of Plant Pathology, Faculty of Agriculture, Kasetsart University, 50 Ngam Wong Wan Rd., Ladyaow, Chatuchak, Bangkok 10900, Thailand.

Solonenko, S. A., Ignacio-espinoza, J. C., Alberti, A., Cruaud, C., Hallam, S., Konstantinidis, K., Tyson, G., Wincker, P. and Sullivan, M. B. 2013. “Sequencing platform and library preparation choices impact viral metagenomes.” *BMC Genomics* 14(320): 1–12.

Suastika,G., Natsuaki,T., Kano,T., Ieki,H. and Okuda,S. "Nucleotide sequence of citrus tristeza virus seedling yellows strain." *Unpublished.* Genbank Submitted (19-JUL-2000) (AB046398.1) Tomohide Natsuaki, Utsunomiya University, Faculty of Agriculture; Mine-machi 350, Utsunomiya, Tochigi 321-8505, JAPAN (E-mail:natsuaki@cc.utsunomiya-u.ac.jp).

Tamura, K. and Nei, M. 1993. “Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.” *Molecular Biology* and *Evolution,* 10(3), 512-26.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. “MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood , Evolutionary Distance , and Maximum Parsimony Methods.” *Molecular Biology and Evolution* 28 (10): 2731–2739. doi:10.1093/molbev/msr121.

Tsai, J. H., Lee, R. F., Liu, Y-h. and Niblett, C. L. 2009, “Biology and Control of Brown Citrus Aphid (Toxoptera Citricida Kirkaldy) and Citrus Tristeza.” In: Radcliffe, E. B., Hutchison, W. D. and R. E. Cancelado (eds.) “Radcliffe’s IMP World Textbook.” Ed. E. B. Radcliffe, W. D. Hutchison, and R. E. Cancelado. St. Paul: University of Minnesota. http://ipmworld.umn.edu.

Tsai, J. H., Lui, Y. H., Wang, J. J. and Lee, R. F. 2000. "Recovering of organge stem pitting strains of *Citrus Tristeza Virus* (CTV) following single aphid transmissions with *Toxoptera citricida* from a Florida decline isolate of CTV." *Proceedings of the Florida State Horticultural Society* 113: 75–78.

Van der Vyver, J. B., van Vuuren, S. P., Luttig, M. and Graça, J. V. 2002. “Changes in the Citrus tristeza virus Status of Pre-immunized Grapefruit Field Trees.” In *Fifteenth IOCV Conference*, , p. 175–185.

Varveri, C., Olmos, A. and Cambra, M. "Full genome of CTV L192GR." *Unpublished.* Genbank Submitted (05-DEC-2012) (KC262793.1) Phytopathology, Benaki Phytopathological Institute, 8, St.Delta str., Kifissia, Attika 14561, Greece.

Vives, M. C., Rubio, L., Lopez, C., Navas-castillo, J., Albiach-martí, M. R., Dawson, W. O., Guerri, J., Flores, R. and Moreno, P. 1999. “The Complete Genome Sequence of the Major Component of a Mild Citrus Tristeza Virus Isolate.” *Journal of General Virology* 80: 811–816.

Wang, Y. J., Zhou, C. Y., Li, Z. A., Zhou, Y., Cao, M. J., Li, L. D., Yang, F. Y., Tang, K. Z. and Liu, J. X. "The complete nucleotide sequence analysis of two citrus tristeza virus isolates from China." *Unpublished.* Genbank Submitted (09-APR-2012) (JQ911663.1 and JQ911664.1) National Citrus Virus-free Seedling Center, Citrus Research Institute, Chinese Academy of Agricultural Sciences, Citrus Village NO15, Xiema, Chongqing 400712, China.

Weng, Z., Barthelson, R., Gowda, S., Hilf, M. E., Dawson, W. O., Galbraith, D. W. and Xiong, Z. 2007. "Persistent Infection and Promiscuous Recombination of Multiple Genotypes of an RNA Virus within a Single Host Generate Extensive Diversity." *PLoS ONE* 2 (9): e917. doi:10.1371/journal.pone.0000917. doi:10.1371/journal.pone.0000917.

Wu, G. W., Hong, N. and Wang, G. P. "Complete genome sequence analysis of an aphid-transmissible isolate from China." *Unpublished*. Genbank Submitted (21-NOV-2011) (JQ061137) Huazhong Agricultural University, College of Plant Science & Technology, Shizishan Street, Hongshan District, Wuhan, Hubei 430070, China.

Yang, Z-N., Mathews, D. M., Dodds, J. A. and Mirkov, T. E. 1999. "Molecular Characterization of an Isolate of Citrus Tristeza Virus That Causes Severe Symptoms in Sweet Orange." *Virus Genes* 19 (2): 131–142.

Yokomi, R. K., Lastra, R., Stroetzel, M. B., Damsteegt, D., Lee, R. F., Garsney, S. M., Gottwald, T. R., Rocha-Pena, M. A. and Niblett, C. L. 1994 "Establishment of the brown citrus aphid (Homoptera:Aphididae) in Central America and the Caribbian Basin and transmission of the citrus tristeza virus." *Journal of Economic Entomology* 87: 1078–1085.

Zablocki, O. and Pietersen, G. 2014. "Characterization of a Novel Citrus Tristeza Virus Genotype within Three Cross-Protecting Source GFMS12 Sub-Isolates in South Africa by Means of Illumina Sequencing." *Archives of Virology*: 107. doi:10.1007/s00705-014-2041-3.

Zhou, C. Y., Broadbent, P., Hailstones, D. L., Bowyer, J., Connor, R. 2002. "Movement and Titer of Citrus tristeza virus (Pre-immunizing isolate PB61) Within Seedlings and Field Trees." In *Fifteenth IOCV Conference*, p. 39–47.

Zhou, C. Y., Broadbent, P., Hailstones, D. L., Bowyer, J., Connor, R., Broadbent, P., Connor, R. and Bowyer, J. 2002. "Studies on Mild Strain Cross Protection Against Stem-pitting Citrus tristeza virus." In *Fifteenth IOCV Conference*, p. 151–157.

2.7. Website references

http://www.ncbi.nlm.nih.gov/

http://mafft.cbrc.jp/alignment/server/

http://www.molecularevolution.org/software/phylogenetics/mrbayes

http://hydrodictyon.eeb.uconn.edu/eebedia/index.php/Phylogenetics:_MrBayes_Lab

http://genome.nci.nih.gov/tools/reformat.html

Chapter 3

Screening for CTV T36 genotype in a Rio Red Grapefruit orchard purportedly harbouring this genotype.

3.1. Abstract.

Citrus tristeza virus (CTV) is one of the most devastating viruses to citrus producers worldwide. CTV is capable of severe economic losses infecting nearly all citrus species and cultivars. Therefore, CTV mild strain cross-protection programs are extremely important in reducing yield and tree losses within areas where CTV is endemic. Therefore, identification and isolation of mild strains of each individual genotype is critical for establishing efficient pre-immunizing sources. The aim of this study is to confirm the presence of the CTV T36 genotype in Southern Africa found in a survey performed during 2012 and to establish it within our greenhouse. Instead, this study revealed that the identification of T36 during the previous study was a misdiagnosis, and that T36 was not present in the original sample or in other grapefruit trees within that orchard. A large numbers of grapefruit trees were found to be infected with the CTV RB genotype; hence, we aimed to find a pure RB population and establish it in our greenhouses. To achieve this RB infected samples were also screened for other known genotypes using genotype specific RT-PCR. A single 2012 graft-inoculated Mexican Lime tree (12-7006) appeared to contain a potential homogenous RB source, and was therefore further assessed for its homogeneity of infection by progressively more robust, but also increasingly more expensive methods of virus population characterization. Direct sequencing of the p33 gene using primers capable of amplifying this gene from all genotypes confirmed the dominant nature of RB within the population. Sequencing of 20 clones of these amplicons further supported the homogeneity of this source. Finally Illumina Next-Generation High-throughput sequencing of the p33 gene amplification products confirmed the purity of this source. In this latter technique hundreds of thousands of sequence reads generated from the amplicons of the p33 gene were mapped to the cognate p33 gene sequences of CTV sources for which whole genome sequences are available.

3.2. Introduction.

In 2012, a survey was carried out where different commercial grapefruit cultivars were screened to identify which CTV genotypes dominated within commercial grapefruit orchards in Southern African, with a view to identifying genotypes for which cross protecting mild strains are required. During this survey the T36 CTV genotype, which had not previously been found in any of the South African grapefruit cultivars, was detected within a foreign grapefruit cultivar named “Rio Red” not commonly grown in the South African citrus industry, in an orchard in Swaziland. Due to the preliminary finding of this new CTV genotype, the specific farm was revisited to confirm the results obtained, to establish the CTV genotype in the greenhouse and to determine whether the T36 genotype had spread from the original tree in which it was identified to others in the same orchard The study involved the screening of “Rio Red” grapefruit trees for the presence of the CTV T36 genotype, including the original tree in which the CTV T36 genotype was detected, as well as collecting budwood samples to establish this genotype within our greenhouse in South Africa. In addition to the routine screening for the CTV T36 genotype, we also planned to obtain and establish a pure T36 source, and to determine the virulence of this source. If mild, the source could potentially be used within the Mild Strain Cross-protection Program (MSCP) to protect against T36-like sources,

The T36 genotype is not a commonly found CTV source within South African citrus orchards. Studies performed on determining the genotype composition and diversity of CTV populations present within grapefruit cultivars in Southern Africa confirmed the rareness of this T36 genotype (Read and Pietersen, 2015; Zablocki and Pietersen, 2014; Scott *et al.*, 2012). Wang *et al.* (2013) published and research article in which they provided information about the CTV collection maintained since 1914 in the Citrus Clonal Protection Program (CCPP), Riverside, California. T36 in the collection was only present in mixed infections with T30 or VT or both, with T36 always present in the lowest percentage. Similar findings of the presence of T36 in a mixture with T30 and VT was published by Roy and Brlansky (2009) and Stewart, 2006.

The “Rio Red” cultivar was developed during a variety improvement program at Texas A & I University Citrus Center (Hensz, R.A., 1985). This cultivar originated in 1953 from a “Ruby Red” seedlings which were propagated onto sour orange rootstocks (Waibel, C., 1953). “Ruby Red” seedlings which showed promising growth and fruit production were selected and replanted in a larger study area. In 1963, budwood samples of the second planting were irradiated either through thermal neutrons or though X-rays and grafted onto sour orange rootstock (Hensz, R. A., 1977). It was found that one of the grapefruit trees grown from a thermal neutron irradiated budwood graft produced fruit with flesh three times brighter than those of a “Ruby Red” tree. In 1971, this specific tree, entitled A & I-1-48, was propagated and together with ten more trees placed in a cultivar specific test field. In 1976, a natural mutation was observed on a limb of the test tree A & I-1-48 (Hensz, R.A., 1981). This natural mutation initiated the production of fruit of which the color was five times redder than that of “Ruby Red”, twice as red as the “Ray Red” cultivar (Hensz, R.A., 1978) and almost as red as the “Star Ruby” cultivar (Hensz, R.A., 1971). This newly produced grapefruit cultivar was named “Rio Red” in 1984 (Hensz, R.A., 1985). Even though the flesh textures, sugars and acid content of the fruit produced by this “Rio Red” cultivar (A & I-1-48) is very similar to the “Ruby Red” and “Ray Ruby” grapefruit cultivars, this cultivar lacks the characteristic red coloured cambium cells found under the bark of the “Star Ruby” cultivar. Also, it was observed that the “Rio Red” trees had a more open tree structure compared to the denser and bushy growth structure of the “Star Ruby” cultivar (Hensz, R.A., 1985). The first “Rio Red” budwood was released during 1984, where the samples were kept safe at the Rio Grande Valley nurserymen, USA. The “Rio Red” grapefruit cultivar was made commercially available in California through the CCPP.

Limited information about the history as well as the introduction of the “Rio Red” Grapefruit in South Africa is available since this cultivar is not native to our country. The following information was obtained through personal communications with Dr. Fanie van Vuuren from Agricultural Research Council – Institute for Tropical and Subtropical Crops (ARC-ITSC) in Nelspruit, Mpumalanga, South Africa. The Rio Red Grapefruit cultivar was introduced to South Africa by Dr. Etienne Rabe during 1987 (14-12-1987) from the University of California, Riverside. In support of the citrus cross-protection scheme which was implemented in 1973 (von Broembsen *et*

al., 1978), quarantine and security provisions were put into practice as soon as the material entered the country. All citrus collected or accepted within the country to date is listed and documented at the Agricultural Research Council (ARC) under a specific code, for example the Star Ruby grapefruit cultivar information is kept under the following code, ARC1179. The information for the Rio Red Grapefruit cultivar can be obtained from the citrus quarantine list (ARC) under the following code, ARC1283 (Ncango, D. and van Vuuren, F., personal communication).

3.3. Materials and methods.

3.3.1. Origin of samples.

In 2012, 49 Grapefruit trees were sampled from different geographical regions throughout South Africa, including an orchard (Block 26) in Swaziland in which CTV T36 was identified using a one-step RT-PCR analysis (Smocilac, S., *unpublished*) (Table 13). The analysis performed by Smocilac was focussed on a amplifying a gene fragment within ORF1a, using the above mentioned analysis. These were graft inoculated onto Mexican Lime seedlings which were maintained within a greenhouse situated at the experimental farm (GPS co-ordinates 25°45'06.40"S; 28°15'35.43"E) of the University of Pretoria. Leaf material was collected from 45 of these graft-inoculated trees in 2013 for this current study as well as to confirm the results obtained in the 2012 survey.

Subsequently, 100 Rio Red Grapefruit trees were sampled by collecting leaf material from block 26 (Figure 23), Tambuti Estate, Swaziland (GPS co-ordinates 26°43'56.95"S; 31°45'45.97"E) using a systematic sampling strategy where every fourth tree was sampled in every third row. Budwood was also collected from twenty-five representative (Table 14) sampled trees, to establish sources within the plant virology greenhouse situated at the experimental farm (GPS co-ordinates 25°45'06.40"S; 28°15'35.43"E) at the University of Pretoria. The collected leaf and budwood samples were kept in a cooler box for the duration of the trip and at 4°C in the laboratory, until used.

All leaf material collected throughout the study was processed by cutting out the midrib as well as the leaf petiole under sterile laboratory conditions. All processed samples were stored at -80°C until used. The collected budwood samples were stored at 4°C for no longer than 2 weeks and graft-inoculated onto newly grown Mexican Lime seedlings.

Table 13: List of Grapefruit trees sampled in 2012 throughout South Africa (Smocilac S., *unpublished*).

Accession number	Grafted onto	Date collected	District/Farm	Scion	Rootstock	Date planted	Date grafted	Symptoms
12-7000	Mexican Lime	2012/04/17	Hoedspruit (Redeberg Estates)	Star Ruby	Unknown	1996 Duroi	19/04/2012	
12-7001	Mexican Lime	2012/04/17	Hoedspruit (Redeberg Estates)	Star Ruby	Unknown	1997 Duroi	19/04/2012	
12-7002	Mexican Lime	2012/04/17	Hoedspruit (Redeberg Estates)	Star Ruby	Unknown	1998 Duroi	19/04/2012	
12-7003	Mexican Lime	2012/04/17	Hoedspruit (Redeberg Estates)	Star Ruby	Unknown	1999 Duroi	20/04/2012	
12-7004	Mexican Lime	2012/04/17	Hoedspruit (Redeberg Estates)	Star Ruby	Unknown	2000 Duroi	19/04/2012	
12-7005	Mexican Lime	2012/04/17	Hoedspruit (Redeberg Estates)	Star Ruby	Unknown	2001 Duroi	19/04/2012	Severe stem pitting
12-7006	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Marsh	Carizzo	1995	19/04/2012	
12-7007	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Marsh	Carizzo	1995	19/04/2012	Severe stem pitting
12-7008	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Marsh	Carizzo	1995	20/04/2012	
12-7009	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Marsh	Carizzo	1995	20/04/2012	
12-7010	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Marsh	Carizzo	1995	20/04/2012	Severe stem pitting
12-7011	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Star Ruby	C-35	2005	19/04/2012	
12-7012	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Star Ruby	C-36	2005	19/04/2012	
12-7013	Mexican Lime	2012/04/17	Hoedspruit (MCE Estates)	Star Ruby	C-37	2005	19/04/2012	

12-7014	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Marsh	Troyer	1989	19/04/2012	Severe stem pitting
12-7015	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Marsh	Troyer	1989	19/04/2012	
12-7016	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Marsh	Troyer	1989	20/04/2012	
12-7017	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Star Ruby	Troyer	1991	20/04/2012	Severe stem pitting
12-7018	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Star Ruby	Troyer	1991	20/04/2012	Severe stem pitting
12-7019	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Star Ruby	Troyer	1991	19/04/2012	Severe stem pitting
12-7020	Mexican Lime	2012/04/17	Malelane (Heno Radley)	Star Ruby	Troyer	1991	19/04/2012	Severe stem pitting
12-7021	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Marsh	Swingle	1999	20/04/2012	
12-7022	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Marsh	Swingle	1999	20/04/2012	Severe stem pitting
12-7023	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Marsh	Swingle	1999	20/04/2012	Severe stem pitting
12-7024	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Rose	Rough Lemon	1998	20/04/2012	
12-7025	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Rose	Rough Lemon	1998	19/04/2012	
12-7026	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Rose	Rough Lemon	1998	19/04/2012	
12-7027	Mexican Lime	2012/04/17	Malelane (Golden Frontiers)	Rose	Rough Lemon	1998	20/04/2012	
12-7028	Mexican Lime	2012/04/18	Swaziland (SwaziCam)	Star Ruby	Swingle	1984	19/04/2012	
12-7029	Mexican Lime	2012/04/18	Swaziland (SwaziCam)	Troyer	Troyer	1984	20/04/2012	
12-7030	Mexican Lime	2012/04/18	Swaziland (SwaziCam)	Swingle	Swingle	1984	20/04/2012	Severe stem pitting
12-7031	Mexican Lime	2012/04/18	Swaziland (SwaziCam)	Swingle	Swingle	1981	20/04/2012	Severe stem pitting
12-7032		2012/04/18	Swaziland (SwaziCam)	Swingle	Swingle	1981	19/04/2012	

12-7033	Mexican Lime	2012/04/18	Swaziland (SwaziCam)	Trifoliate	Trifoliate	1981	19/04/2012	
12-7034		2012/04/18	Swaziland (SwaziCam)	Trifoliate	Trifoliate	1981	19/04/2012	
12-7035	Mexican Lime	2012/04/18	Swaziland (SwaziCam)	Trifoliate	Trifoliate	1981	19/04/2012	Severe stem pitting
12-7036		2012/04/18	Swaziland (SwaziCam)	Trifoliate	Trifoliate	1981	19/04/2012	
12-7037	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1994	19/04/2012	
12-7038	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1994	19/04/2012	
12-7039	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1994	19/04/2012	
12-7040	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1994	19/04/2012	
12-7041	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1999	19/04/2012	
12-7042	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1999	19/04/2012	
12-7043	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1999	19/04/2012	
12-7044	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Marsh	Swingle	1999	19/04/2012	
12-7045	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Rio Red	Rio Red	1994	19/04/2012	
12-7046	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Rio Red	Rio Red	1994	19/04/2012	
12-7047	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Rio Red	Rio Red	1994	19/04/2012	
12-7048	Mexican Lime	2012/04/18	Swaziland (Tambuti Estates)	Rio Red	Rio Red	1994	19/04/2012	

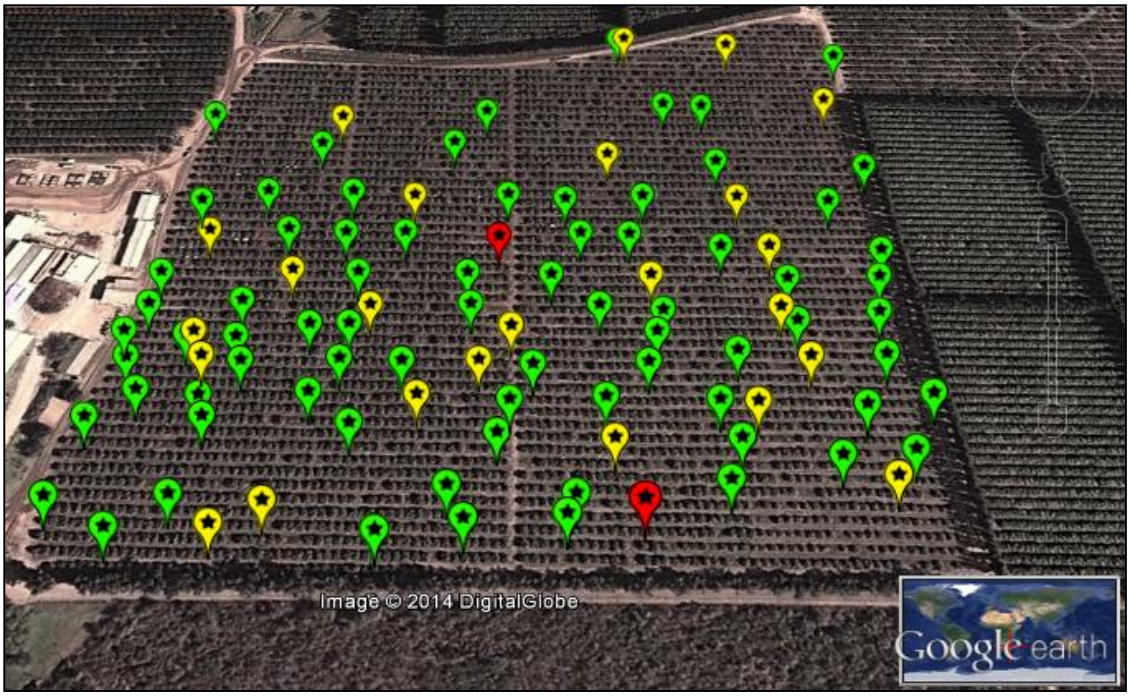


Figure 23: Rio Red Grapefruit orchard, block 26, showing the systematic sampling that took place. Position within Rio Red Grapefruit orchard where leaf samples were collected; Position within Rio Red Grapefruit orchard where both leaf and budwood samples were collected; Position of budwood samples 13-4040 and 13-4112. (13-4040 replaced in analysis for CTV with sample 13-4112 after having tested negative within the CTV T36 RT-PCR.

3.3.2. Establishment of virus sources.

Rio Red Grapefruit budwood was graft-inoculated onto newly grown (3 months old) Mexican Lime seedlings, under insect-free greenhouse conditions (20 - 30°C, watered daily, and entirely sheltered/protected from outside environmental factors). The stem of each seedling was cut vertically at an angle (about 30°) using a sterile scalpel blade, creating a grafting site for the budwood sample to be inserted. Buds of the collected budwood samples were also excised at a 30° angle, using a new sterile scalpel blade, to ensure a suitable fit to the grafting site made. The buds were inserted into their respective grafting sites on the Mexican Lime seedlings and closed with grafting tape (Plastrip, Somerset West, South Africa). The seedlings were left for 4 months, allowing the virus to multiply and spread throughout the entire seedling. After 4 months the seedlings were examined for any distinct CTV symptoms and leaf samples were collected, followed by a generically primed CTV RT-PCR analysis, where all the inoculated seedlings were screened for CTV presence.

Table 14: Representative Rio Red Grapefruit trees from which budwood were collected at Tambuti Estate, Swaziland for graft-inoculations on Mexican Lime Seedlings.

Tree	GPS co-ordinates		Altitude above sea level	Accession number	Tree symptoms
	South	East			
1	31.76111052	-26.73149834	203.08	13-4012	Decline with stunted fruit (Smocilac 12-7047)
3	31.7613776	-26.7309686	202.309	13-4014	Severe decline, sparse canopy, small fruit
5	31.76208034	-26.7306933	207.564	13-4016	Severe decline and stunting with small fruit
9	31.76131085	-26.73321468	204.25	13-4020	Healthy at the bottom of the tree, decline and small fruit at the top of the tree
13	31.76212728	-26.73197977	206.756	13-4024	Healthy at the bottom of the tree, decline and small fruit at the top of the tree
17	31.7626997	-26.73149086	208.967	13-4028	Mostly healthy, some decline at the top of the canopy
21	31.76211439	-26.73302313	206.671	13-4032	Healthy looking, no obvious symptoms
25	31.7620274	-26.73405736	202.826	13-4036	Healthy looking, but has mild decline at the top of the tree
29	31.7625707	-26.73271582	205.592	13-4040	Healthy looking, no obvious symptoms
33	31.76311226	-26.73150277	210.986	13-4044	Mostly healthy looking, has some mild decline
37	31.76310446	-26.73211953	209.993	13-4048	Mild decline throughout the tree, small fruit at the top of the tree
41	31.76246288	-26.73372865	207.125	13-4052	Mostly healthy looking with large fruit, mild decline at the top of the tree
45	31.76283998	-26.73343593	207.46	13-4056	Healthy looking with large fruit yield
49	31.76353344	-26.73163893	210.656	13-4060	Mostly healthy looking with some mild decline

53	31.76320923	-26.73285635	209.25	13-4064	Mild decline throughout the tree
57	31.76273009	-26.73424828	207.599	13-4068	Mild decline throughout the tree
61	31.76290277	-26.7342415	206.032	13-4072	Mild decline near the top of the tree
65	31.76337396	-26.73306324	207.644	13-4076	Mild to severe decline at the top of the tree canopy
69	31.76388016	-26.73165573	211.173	13-4080	Healthy looking, no obvious symptoms
73	31.76405519	-26.73199502	211.465	13-4084	Mild to severe decline towards the top of the tree canopy
77	31.76348538	-26.73338343	207.401	13-4088	Healthy looking with large fruit, no obvious symptoms
85	31.76403524	-26.73265331	211.459	13-4096	Healthy looking with large fruit, no obvious symptoms
89	31.76465005	-26.73163577	213.724	13-4100	Mild to severe decline on certain branches of the tree
93	31.76385237	-26.73414245	205.976	13-4104	Healthy looking with large fruit yield, no obvious symptoms
97	31.76389357	-26.73437175	205.461	13-4108	Healthy looking with large fruit yield, no obvious symptoms

3.3.3. RNA extractions.

All collected leaf samples were subjected to total RNA extraction. Petioles and leaf midribs were macerated using liquid Nitrogen in a sterile mortar and pestle. Total RNA extraction was performed on the macerated plant material using a GeneJET Plant RNA Purification mini kit (Thermo Scientific, Waltham, MA, USA), according to the manufacturers protocol.

Briefly, plant material collected was macerated followed by the adding of 500µl Plant RNA Lysis Solution and 10µl of 2M dithiothreitol (DTT) (Thermo Scientific, Waltham, MA, USA). Samples were mixed thoroughly by vortexing them for 10-20 second. Thereafter, the samples were incubated for 3 minutes at 65°C and centrifuged for 5 minutes at a maximum speed of 14,000 x g, at 10°C. The resultant supernatant was pipetted into a 1.5ml microtube, together with 250µl of 96% ethanol. The solutions were mixed by pipetting, and transferred to a purification column, inserted into a collection tube. The column was centrifuged for 1 minute at 11,000 x g, at 10°C. The flow-through was discarded and 700µl of Wash Buffer 1 (WB1) (210.5ml total volume: 200ml concentrated wash solution containing H2O and guanidinium chloride; 10.5ml 96-100% ethanol; pH at 25°C 7.5) was added to the reassembled purification column. . The column was centrifuged for 1 minute at 11,000 x g, at 10°C, followed by discarding both the flow-through and the collection tube. The purification column was reassembled with a new clean 2ml collection tube and 500µl of Wash Buffer 2 (WB2) (200ml total volume: 100ml concentrated wash solution containing H2O; 100ML 96-100% ethanol and 100ml 96-100% ethanol; pH at 25°C 5.4) was added to the purification column. The column was centrifuged for 1 minute at 11,000 x g, at 10°C, subsequently discarding the flow-through and repeating the wash step one more time with WB2. The column was centrifuged for 1 minute at 14,000 x g, at 10°C. After the centrifugation period, the purification column was transferred to an RNase-free 1.5ml collection tube. To elute the required RNA, 50µl of nuclease-free water was added to each purification column membrane and centrifuged for 1 minute at 11,000 x g, at 10°C. The quality of the RNA extract was assessed by agarose gel electrophoresis. All RNA obtained was stored at -80°C until used.

3.3.4. Screening for CTV presence through RT-PCR analysis.

RNA obtained from the 45 2012 graft-inoculated Mexican Lime trees, 100 Rio Red Grapefruit trees, and 25 2013 graft-inoculated Mexican Lime seedlings, were screened for CTV presence firstly, by means of a RT-PCR analysis capable of detecting all known CTV genotypes ("Generic RT-PCR"). Thereafter, all CTV positive samples were screened by PCR for known CTV genotypes and amplicons directly Sanger sequenced to confirm the identity. Both a healthy control as well as a CTV positive control was included in all tests.

Total RNA extract was used as template for the production of cDNA for the different CTV genotypes screened for. A final denaturation reaction mixture of 10µl was made up in a 0.2µl PCR tube, for each CTV genotype, as follows: 2µl of total RNA, 1µl of the respective CTV genotypic reverse primers (10mM) (Table 15) and 7µl PCR grade H_2O (Promega, Madison, WI, USA). The denaturation reaction conditions were 95°C for 3 minutes (to disrupt the RNA secondary structures) and 4°C for 1 minute to stop the reaction.

Reverse transcription was carried out on the denatured RNA by adding 4µl 1x reverse transcriptase (RT) buffer (Roche, Mannheim, Germany), 2µl (1mM) deoxynucleotide triphosphate (dNTP) mix (Kapa Biosystems Inc, Wilmington, MA), 2µl (10mM) dithiothreitol (DTT) (Thermo Scientific, Waltham, MA, USA), 0.125µl (1.25U/µl) *Avian myeloblastosis virus* reverse transcriptase (AMV-RT) (Roche, Mannheim, Germany), 0.25µl (0.5U/µl) RNAse Inhibitor (Ribolock, Thermo Scientific, Waltham, MA, USA) and 1.625µl PCR grade H_2O (Promega, Madison, WI, USA) to the 10µl denatured RNA mixture, to make a final reaction volume of 20µl. The reverse transcription reaction conditions were 42°C incubation for 60 minutes, 85°C for 5 minutes and then maintained at 4°C until used.

PCR amplification was carried out on the cDNA by using 2µl of the cDNA as template, which was added to a reaction mixture containing 10µl Taq polymerase master mix (Promega GoTaq Hot Start Green Master Mix, Promega, Madison, WI, USA), 0.75µl (0.375µM) forward primer, 0.75µl (0.375µM) reverse primer (Table 15) and 6.5µl PCR grade H_2O (Promega, Madison, WI, USA) to make a final volume of 20µl. The following PCR amplification reaction conditions were used: 1 cycle of 94°C for 2 minutes (denaturation step); 35 cycles of 94°C for 20 seconds, various

temperatures for different primers (Table 2) for 30 seconds, 72°C for 20 seconds (primer annealing and DNA extension steps); 1 cycle of 72°C for 1 minute and 4°C until the PCR products are used. A 1.5% Agarose gel electrophoresis was done on all PCR reactions to confirm the presence of any potential CTV genotypes.

3.3.5. Agarose gel electrophoresis.

Positive PCR products were analysed by means of agarose gel (SeaKem LE agarose, Lonza, Rockland, ME, USA) electrophoresis. Twelve microlitres of each sample was loaded per well of a 1.5% agarose gel, pre-stained with 2.5µl (10mg/ml) per 50ml ethidium bromide, (Promega, Madison, WI, USA) and run for 60 min at 75 Volts using a Bio-Rad gel electrophoresis system (Bio-Rad Laboratories LTD, South Africa, Parkland, Johannesburg) using 1x TAE buffer (40mM Tris-acetate, 1mM EDTA, pH 8.2). HyperLadder IV 100bp (Bioline, Celtic Molecular Diagnostics (Pty) Ltd., South Africa) was used to differentiate the different fragment sizes of the PCR products. Results were viewed using the Bio Rad Gel Doc™ EZ Imager (Bio-Rad Laboratories LTD, South Africa, Parkland, Johannesburg).

Table 15: Primer sets used in this research study.

CTV genotype	Polarity *	Primer sequence (5' - 3')	Length (bp)	Annealing temp. (˚C)	Primer binding sites within CTV genotype genome	Product size (nt)	Reference
T36#	S	TTC CCT AGG TCG GAT CCC GAG TAT A	25	59.9	1775...2610	836	Roy *et al.* 2010
	A	CAA ACC GGG AAG TGA CAC ACT TGT TA	26	59.2			
Generic	S	ATG GAC GAC GAR ACA AAG AAA TTG AAG A	28	57.8	16105…16776	672	Roy *et al.* 2010
	A	TCA ACG TGT GTT RAA TTT CCC AAG CT	26	58.4			
B165#	S	GTT AAG AAG GAT CAC CAT CTT GAC GTT GA	29	58.4	2124...2633	510	Roy *et al.* 2010
	A	AAA ATG CAC TGT AAC AAG ACC CGA CTC	27	59.2			
T3#	S	GTT ATC ACG CCT AAA GTT TGG TAC CAC T	28	58.7	4846...5254	409	Roy *et al.* 2010
	A	CAT GAC ATC GAA GAT AGC CGA AGC	24	57.7			
VT#	S	TTT GAA AAT GGT GAT GAT TTC GCC GTC A	28	59.5	1945...2246	302	Roy *et al.* 2010
	A	GAC ACC GGA ACT GCY TGA ACA GAA T	25	60.2			
T30	S	TGT TGC GAA ACT AGT TGA CCC TAC TG	26	58.7	588...793	206	Roy *et al.* 2010
	A	TAG TGG GCA GAG TGC CAA AAG AGA T	25	60.1			
NZRB-28, -M12,G90 & HA18-9							
NZRB F1#	S	AGT GGT GGA GAT TAC GTT G	19	60	1979...2626	646	G. Cook, *Unpublished*
NZRB R1#	A	TAC ACG CGA CAA ATC GAG	18	60			
NZRB-30 & -17 (Should not detect T36)							
NZRB F2#	S	CGG AAG GGA CTA CGT GGT	18	60	1981...2643	662	G. Cook, *Unpublished*
NZRB R2#	A	CGT TTG CAC GGG TTC AAT G	19	60			
HA16-5 F#	S	TAG GAA GGG TCA CTG CCC TGA CA	23	56	2154...2813	658	G. Cook, *Unpublished*
HA16-5 R#	A	GTA AGT ATC TAA AAC CAG GAG	21	56			

CTV p33 gene region							
Univ-p33-F[#]	S	GATGTTTGCCTTCGCGAGC	19	55	11020....12001	980	D. Read, *Unpublished*
▲Univ-p33-R[#]	A	CCCGTTTAAACAGAGTCAAACGG	23	55			
CTV p33 gene region cloning primers							
T7 forward[#]	S	TAA TAC GAC TCA CTA TAG GG	20	55	11020....12001	980	D. Read, *Unpublished*
SP6 reverse[#]	A	TAT TTA GGT GAC ACT ATA G	19	55			

Polarity *: Sense = S; Antisense = A. **▲:** 1kb sized Universal p33 reverse primer. [#]: Primer sets used during sequencing analysis.

3.3.6. Column purification of PCR products.

Due to the small amount of DNA template left after gel electrophoresis, gel band purifications were performed on all CTV positive samples prior to direct Sanger sequencing. The Promega, Wizard® SV Gel and PCR Clean-up system (Promega, Madison, WI, USA) was used, as per the manufacturer's instructions. In brief, the positive DNA bands were excised after electrophoresis by means of a sterile scalpel blade and transferred to pre-weighed sterile 1.5ml microtubes. Each tube containing the excised band was then re-weighed, to determine the final weight of the product. Membrane binding solution was added in a ratio of 10µl of membrane binding solution per 10mg of agarose gel slice. Each tube was then vortexed incubated at 56°C for 10 minutes (min) or until the gel slice was completely dissolved.

This was briefly centrifuged at 8, 000 x g for 2 seconds at room temperature and transferred to their respective SV Minicolumn assembly (in a collection tube). This was allowed to incubate for 1 min at room temperature, followed by centrifugation at 10, 000 x g for 1 min. The sample flow through of each SV Minicolumn assembly was discarded, followed by a wash step by adding 700µl Membrane Wash Solution to each of the SV minicolumns. The assemblies were then centrifuged at 10,000 x g for 1 min. At the end of the centrifugation the flow through was discarded and the wash step was repeated with 500µl Membrane Wash Solution, followed by centrifugation of the SV Minicolumn assembly at 10,000 x g for 5 min. Each collection tube was emptied of all flow through and re-centrifuged for 1 min, allowing any residual ethanol to disperse. Each SV Minicolumn was transferred to a to a sterile 1.5ml microtube. Fifty microlitres of nuclease-free water (Promega, Madison, WI, USA) was added directly onto the membrane and allowed to stand for 1 min at room temperature. The SV minicolumn and microtube assembly was then centrifugation at 10,000 x g for 1 min. The SV Minicolumns were discarded and all purified DNA samples were stored at -20°C until used. A 1% Agarose gel electrophoresis was done on all column purified DNA to confirm DNA presence. Purified DNA was subjected to another round of PCR amplification prior to sequencing to increase the DNA concentration, to obtain better quality sequences.

3.3.7. Sanger sequencing.

Purified PCR products were Sanger sequenced to confirm positive results. In brief, 2µl FastAP Alkaline Phosphatase (1U/µl) (Thermo Scientific, Waltham, MA, USA) and 0.5µl Exonuclease (20U/µl) (Thermo Scientific, Waltham, MA, USA) was added to 19µl of the purified PCR product or template DNA in individual PCR tubes. Samples were incubated for 15 min at 37°C followed by a 15 min incubation period at 85°C, to conclude the reaction.

Purified PCR products were sequenced in both directions using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) as per manufacturer's instructions. Sequencing reactions were as follows; 2.25µl 5x BigDye® v3.1 sequencing buffer (Applied Biosystems, Foster City, CA, USA), 0.75µl of 2µM CTV individual genotypic primers (Table 15) (Integrated DNA Technologies, Coralville, IA, USA), 1µl of 2.5x terminator mix v3.1 (Applied Biosystems, Foster City, CA, USA), nuclease free molecular grade water and template DNA to make it up to a final volume of 10µl. The amount of template DNA and hence the nuclease free molecular water added was determined by the intensity of the gel band obtained (concentration of the DNA) after the PCR products were column purified. The brighter the band obtained the less template was added to the reaction mixture, for example 1µl of template was added when a strong band was obtained. Two to three microlitres of template was added for a medium intensity band and 5-6µl template was added to the reaction mixture for a very faint band. The sequencing reaction conditions were as follows: 92°C for 1 min; 30 cycles each of 94°C for 10 seconds (sec), 50°C for 5 sec, and 60°C for 4 min.

For each of the 10µl reaction mixture a product precipitation was performed after the completion of the thermo-cycling. The precipitation process involved the addition of 1µl of 125Mm EDTA, 1µl of 3M Sodium Acetate and 25µl of 100% non-denature Molecular grade Ethanol (Sigma-Aldrich, St. Louis, MO, USA) to each of the reaction tubes containing the 10µl sequencing product. The solution was mixed by vortexing it for 2 seconds and then incubated for 15 minutes at room temperature (25°C). After incubation the solution was centrifuged at a maximum speed of 14,000 x g, at 4°C, for 20 to 30 minutes and the resulting supernatant was removed by pipette. One hundred microlitres of 70% Ethanol (Sigma-Aldrich, St. Louis, MO,

USA) was immediately added to each reaction tube and centrifuged at a maximum speed of 14,000 x g, at 4°C, for 10 to 15 minutes. The supernatant was once again removed without disturbing the pellet at the bottom of the reaction tube. After the removal of the supernatant the samples was allowed to air-dry, with the lid of the reaction tube open, for 20 minutes or alternatively at 94°C for no longer than 1 minute.

The resulting precipitated reactions were submitted for DNA Sanger sequencing at the African Centre for Gene Technologies, Automated Sequencing Facility, Department of Genetics, University of Pretoria, South Africa, making use of an ABI Prism 3130XL Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

3.3.8. Phylogenetic analysis.

The CLC Main Workbench 6 software package (CLC Bio, Aarhus, Denmark) was used to 1) view the sequence chromatograms obtained, 2) evaluate the quality and completeness of sequences, 3) correct any errors within the sequences and 4) produce a consensus sequence for each sample. Screening of sequence identities was done using the National Centre for Biotechnology Information (NCBI) Basic Local Alignment Tool (BLAST) (http://www.ncbi.nlm.nih.gov/) (Altschul *et al.,* 1990). Complete consensus sequences, together with cognate regions from 45 complete genome CTV reference sequences, were aligned using online MAFFT version 7 software, with the default settings (http://mafft.cbrc.jp/alignment/server/ ; Katoh *et al.*, 2002, 2005). The default setting were as follow; the “strategy” set to auto, “scoring matrix for amino acid sequences” set to BLOSUM62, “scoring matrix for nucleotide sequences” set to 200PAM/k=2, “align unrelated segments” set to try to align gap regions anyway, “number of homologs” set to 50, “threshold” set to 1a-10 and “plot last hit threshold” set to score=39 (E=8.4e11). The reference sequences together with their respective Genbank accession number and original citation were retrieved from the NCBI database; NUagA [seedling yellows strain] (AB046398) (Suastika etal., unpublished data), SY568 (AF001623) (Yang *et al.*, 1999), T30 (AF260651) (Albiach-martı *et al.* 2000), CTV (AY170468) (Satyanarayana *et al.*, 1999), Qaha from Egypt (AY340974) (Abdelmaksoud & Gamal El-din, unpublished data), T318A

(DQ151548) (Ruiz-Ruiz *et al.*, 2006), Mexican CTV isolate (DQ272579) (Quiroz *et al.*, unpublished data), B165 (EU076703) (Avijit Roy and Brlansky 2010), CTV strain SP (EU857538) (Harper *et al.*, 2009), CTV strain VT (EU937519) (Weng *et al.*, 2007), CTV strain T30 (EU937520) (Weng *et al.*, 2007), CTV strain T36 (EU937521) (Weng *et al.*, 2007), NZRB-M12 (FJ525431) (Harper *et al.*, 2010), NZRB-G90 (FJ525432) (Harper *et al.*, 2010), NZRB-TH28 (FJ525433) (Harper *et al.*, 2010), NZRB-TH30 (FJ525434) (Harper *et al.*, 2010), NZRB-M17 (FJ525435) (Harper *et al.*, 2010), NZ-B18 (FJ525436) (Harper *et al.*, 2010), HA18-9 (GQ454869) (Melzer *et al.*, 2010), HA16-5 (GQ454870) (Melzer *et al.*, 2010), CTV strain Kpg3 (HM573451) (Biswas *et al.*, 2012), CTV strain B301 (JF957196) (Roy *et al.*, unpublished data), CTV strain AT-1 (JQ061137) (Wu *et al.*, unpublished data), CTV strain A18 (JQ798289) (Sindhvajiva *et al.*, unpublished data), CTV strain CT14A (JQ911663) (Wang *et al.*, unpublished data), CTV strain CT11A (JQ911664) (Wang *et al.*, unpublished data), CTV strain T68-1 (JQ965169) (Hilf *et al.*, unpublished data), Taiwan-Pum/SP/T1 (JX266712) (Chen *et al.*, unpublished data), Taiwan-Pum/M/T5 (JX266713) (Chen *et al.*, unpublished data), L192GR (KC262793) (Varveri *et al.*, unpublished data), CT-ZA3 (KC 333868) (Zablocki & Pietersen, 2014), FS674-T36 (KC517485) (Harper & Hilf, unpublished data), FS701-T36 (KC517486) (Harper & Hilf, unpublished data), FS703-T36 (KC517487) (Harper & Hilf, unpublished data), FS577 (KC517488) (Harper & Hilf, unpublished data), FS701-T30 (KC517489) (Harper & Hilf, unpublished data), FL278-T30 (KC517490) (Harper & Hilf, unpublished data), FS703-T30 (KC517491) (Harper & Hilf, unpublished data), FS703-VT (KC517492) (Harper & Hilf, unpublished data), FL202-VT (KC517493) (Harper & Hilf, unpublished data), FS701-VT (KC517494) (Harper & Hilf, unpublished data), T3 (KC525952) (Hilf & Harper, unpublished data), CTV (NC001661) (Karasev *et al.*,1997), T36 (U16304) (Pappu *et al.*, 1994), T385 (Y18420) (Vives *et al.*, 1999). A more detailed table containing information for each of the 45 CTV reference genomes used is available in Appendix 1. These CTV reference genomes were those available at the time of analysis and were used to confirm the results obtained from the RT-PCR as well as the BLAST analysis.

Cognate regions of all 45 CTV reference sequences were excised separately for each of the specific CTV genotype amplicon regions based on the different primer binding sites for each CTV genotype using BioEdit Sequence Alignment Editor

software (Version 7.0.8, Hall, T.A. 1999, Ibis Bioscience, Carlsbad) (Table 15). All edited and aligned sequences were saved in the necessary formats required (FASTA, Phylip4) for the different phylogenetic analyses that were done. Neighbor-Joining, Bayesian and Maximum Likelihood phylogenetic analyses were done on all sequences obtained. The following CTV complete genomes were selected to represent the 9 different CTV genotypic groups within the phylogenetic trees produced; EU076703.3 (B165), EU937519.1 (VT), EU937520.1 (T30), EU937521.1 (T36), FJ525431.1 (RB Clade 1), FJ525434.1 (RB Clade 2), GQ454870.1 (HA16-5), JQ965169.1 (T68-1) and KC525952.1 (T3).

Nucleotide Sequence alignments were analysed in jModelTest version 2.1.4 (Darriba *et al.*, 2012) using the Akaike information criterion model. The models used for the sequences (FASTA format) are indicated in Table 16. A detail explanation for the different substitution models obtained is discussed in Chapter 2 (*section 2.3.13. Phylogenetic analysis*). These models were used during Bayesian as well as Maximum Likelihood phylogenetic analysis.

Table 16: Nucleotide substitution models used for respective phylogenetic analysis data sets.

Sample	Nucleotide substitution model
T36*	TrN+G
NZRB F1R1*	GTR+I+G
NZRB F2R2*	TIM2+I+G
HA16-5*	TIM2+I+G
VT*	TVM+I+G
B165*	TIM2+I+G
T3*	TIM1+I+G
12-7006 p33	TPM1uf+G
12-7006 p33 clones	TVM+I+G
T36 Rio Red leaf samples	TPM1uf+I+G
13-4118 T36**	GTR+I+G
13-4118 NZRB F1R1**	GTR+I+G
13-4118 NZRB F2R2**	TIM2+I+G
13-4118 HA16-5**	TIM2+I+G
13-4118 VT**	TIM3+I+G

*Graft-inoculated Mexican Lime trees (2012 survey) situated within the greenhouse; **Rio Red budwood graft-inoculated Mexican Lime seedling. GTR: General Time Reversible (Rodriguez et. al., 1990); TIM: Transitional model; TrN: Tamura-Nei (Tamura and Nei 1993); HKY: Hasegawa-Kishino-Yano (Hasegawa et. al., 1985); TVMef: Transversional model with equal base frequencies. TPM1uf: Three-parameter model with unequal base frequencies. G: Gamma distribution; I: Proportion of invariable sites.

Maximum Likelihood phylogenetic analysis was performed for all sequences generated (Phylip4 format) in PhyML version 3.0 (Guindon and Gascuel, 2003). All default parameters, together with the appropriate nucleotide substitution model, were used except for the branch support that was set on a 1000 bootstrap replicates. Additionally, the topology of tree searching was set on a “best of both” (NNI and SPR) approach.

Prior to the Bayesian phylogenetic analysis, MrBayes blocks were created using the JModel Test results obtained (Table 16) from the different sequences analysed as well as information gained for the following websites; http://www.molecularevolution.org/software/phylogenetics/mrbayes and http://hydrodictyon.eeb.uconn.edu/eebedia/index.php/Phylogenetics:_MrBayes_Lab . The MAFFT aligned sequences (FASTA format) were converted to a nexus data file using an online sequence format converter program (http://genome.nci.nih.gov/tools/reformat.html) and added to the MrBayes block created. Bayesian analysis was done for all obtained sequences using MrBayes version 3.2.1 software (Ronquist *et al.*, 2012). The appropriate substitution models were used together with 4 Markov chains and at least 3 million generations. Each analysis was run twice. P files obtain from the analysis were analysed in Tracer version 1.5 (Rambaut and Drummond, 2009) to determine the required burn-in values. The burn-in values were added to the analysis and tree files were completed and trees were viewed and edited in FigTree version 1.3.1 (Rambaut, 2009).

Neighbor-Joining analysis for all sequences generated was performed using MEGA version 5 (Tamura *et al.*, 2011). The Jukes-Cantor model was used with a bootstrap analysis of a 1000 replicates.

3.3.9. ***Homogeneity of CTV source 12-7006 determined by cloning of the p33 gene region amplicons.***

RT-PCR products obtained from the p33 gene region of sample 12-7006 were run on a 1.5% Agarose gel and purified by using the Nucleospin® Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. As high DNA concentrations are required for cloning purposes as well as Illumina Next-Generation High-throughput sequencing, we decided to replace the Wizard® SV Gel and PCR Clean-up System (Promega, Madison, WI, USA) with the above mentioned purification kit as higher concentrations and more pure DNA was obtained. This decision was based on a comparative analysis performed between the two purification kits prior to the cloning and NGS analysis performed.

In brief, positive DNA bands were excised after electrophoresis by means of a sterile scalpel blade and transferred to a sterile 1.5ml eppendorf tube of known weight. Tubes containing the excised band were weighed again, to determine the final weight of the product. Two volumes of binding buffer NT1 was added to one volume of the sample being purified, for e.g. for each 100mg of gel 200µl of buffer NT1 was added. Samples were incubated for 10 minutes at 50°C, vortexing every 2 minutes until the gel was completely dissolved. The tube containing the dissolved gel solutions was briefly centrifuged at 8, 000 x g for 2 seconds at room temperature to ensure that all the solution is at the base of the tube. The solution (700µl at a time) was transferred to a NucleoSpin Column and collection tube assembly at centrifuged at 11, 000 x g for 30 seconds, allowing the DNA to bind to the membrane. The sample flow-through was discarded and 700µl of Wash Buffer NT3 was added to the assembly and centrifuged at 11, 000 x g for 30 seconds. At the end of the centrifugation the flow-through was discarded and the wash step was repeated. Ann additional centrifugation step was done at 11, 000 x g for 1 minute, to remove excess Buffer NT3. Thereafter, the NucleoSpin Column was placed into a new 1.5ml Eppendorf tube and incubated for 5 minutes at 70°C, allowing excess Ethanol to evaporate. Thirty microlitres of Elution Buffer NE was added to the assembly and incubated for 1 minute at room temperature (18 - 25°C). The assembly was centrifuged for 1 minute at 11, 000 x g, followed by the removal of the NucleoSpin Column and storing of all purified DNA samples at -20°C until used. A 1% agarose

gel electrophoresis was done on all column purified DNA sample to confirm DNA presence.

The concentration of each purified DNA product was determined by means of UV spectroscopy using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Purified PCR products of the p33 gene were cloned into the pGEM®-T Easy vector (Promega, Madison, WI, USA) according to the manufacturer's instructions. Thereafter, the cloned products were transformed into competent *E.coli* JM109 cells (Promega, Madison, WI, USA) according to the manufacturer's instructions.

The transformants were grown on Luria-Bertani (LB) agar plates, containing (1µl/1ml LB agar) 50µg/ml Ampicilin (Amp) (United States Biochemical, Cleveland, OH, USA), 40µl of 25mg/ml bromo-chloro-indolyl-galactopyranoside (X-gal) solution (Fermentas, Vilnius, Lithuania) and 100µl of 100mM isopropyl-ß-D-1 thiogalactopyranoside (IPTG) solution (Sigma-Aldrich, St. Louis, MO, USA). Prospective recombinant colonies were selected by means of blue/white screening, where positively identified colonies were grown in LB broth, containing 50µg/ml Amp, overnight at 37˚C in a shaking incubator (200rpm). A total of 100 recombinant colonies were selected to be screened for the presence of the p33 gene region.

Plasmid extractions were performed for each of the selected recombinant colonies by using the alkaline lysis plasmid miniprep test (Sambrook, J. and Russell, D. W., 2001). Each of the extracted recombinant plasmids was tested for the presence of the expected insert size (p33 gene region targeted: ±980bp). A PCR reaction, making use of vector specific primers T7 (5'- TAA TAC GAC TCA CTA TAG GG -3') and SP6 (5'- ATT TAG GTG ACA CTA TAG AA -3') (Promega, pGEM®-T and pGEM®-T Easy Vector Systems, technical manual), was done to amplify the p33 insert, between these two vector specific promoter regions. The PCR conditions used in this reaction is the same as described for the p33 gene region. The PCR products were run on a 1.5% Agarose gel and the plasmids whose PCR products were the correct size was submitted for sequencing to specify the insert presence. Only 20 extracted plasmids were chosen for sequencing due to the high amount of positives obtained (Figure 26).

3.3.10. ***Homogeneity determination of CTV source 12-7006 by Illumina Next-Generation High-throughput sequencing.***

The amplicons obtained from the p33 PCR amplification reaction were purified and the concentration of the purified product was determined using a Qubit-iT™ dsDNA BR Assay Kit (Invitrogen™, Carlsbad, CA, USA; Eugene, Oregon, USA), according to the manufacturer's instructions. Samples were submitted to the Agricultural Research Council (ARC) – Biotechnology platform, situated in Pretoria, South Africa, where libraries were prepared using TruSeq3 – PE – 2 adapters (Illumina, San Diago, CA, USA) according to the manufacturer's specification, following by a sequencing reaction on 1/24th of a lane using the Illumina Miseq platform (Illumina, San Diago, CA, USA).

Illumina paired-end data analysis was performed using CLC Genomics Workbench version 5.1.1. (CLC bio, QIAGEN company, Prismet, Aarhus, DK, Denmark) software. Briefly, the raw sequence data was imported as paired reads (distance of 180 – 600). Quality scores for the obtained data set were generated using the FastQC function according to its default settings. The sample reads were filtered by removing low quality sequences (quality limit of 0.05), ambiguous nucleotides (maximum of 2 nucleotides allowed) as well as the adapter sequences used, Truseq3 – PE – 2 (PrefixPE/1: TACACTCTTTCCCTACACGACGCTCTTCCGATCT; PrefixPE/2: GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT; PE1: TACACTCTTTCCCTACACGACGCTCTTCCGATCT; PE1_rc: AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTA; PE2: GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT; PE2_rc: AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC). A set of 45 CTV reference sequences edited for the p33 gene region, was imported and used during reference mapping. The following settings were used to perform reference mapping: mismatch cost of 2, insertion cost of 3, deletion cost of 3, length fraction of 0.9, similarity fraction of 0.9, conflict resolution by voting (A, T, C, G) and non-specific match handling was set to be ignored. Detailed reports were created after each of the steps discussed.

3.4. Results

3.4.1. Genotype characterization of greenhouse maintained, graft transmitted CTV sources from a 2012 survey of grapefruit.

Forty-four of the 45 grapefruit graft-inoculated Mexican Lime trees tested positive for CTV following RT-PCR analysis using primers capable of binding to all known CTV genotypes (CTV generic RT-PCR). Only Mexican Lime seedling 12-7010 was negative, and was not studied further. No CTV-like symptoms were detected during the sampling of the seedlings. The 44 CTV-infected trees were tested for various known CTV genotypes using genotype specific RT-PCR and the amplicons from each of these tests sequenced and subjected to Bayesian phylogenetic analysis to confirm the results (Appendix 6).

Most of the Mexican Lime trees screened contained a mixed infection of 2 or more CTV genotypes (Table 17). The CTV genotype specific PCR analysis yielded amplicons of the expected size in all 44 sources. However, as illustrated in Table 17, all amplicon sequences of the T36 PCR grouped within the Resistance Breaking (RB) genotypic group (Clade 1 or 2) during Bayesian phylogenetic analysis. Thus, indicating that the T36 RT-PCR analysis is unable to distinguish between these two CTV genotypes. The majority of these samples also tended to yield amplicons to both the RB Clade 1 and 2, while only a small percentage of the trees screened yielded amplicons to only one of the two RB clades. CTV B165- , VT- and HA16-5 specific PCR's also yielded amplicons in 28, 25 and 20 out of the 44 sources respectively. Four of the Mexican Lime trees (12-7014, 12-7017, 12-7023 and 12-7025) yielded amplicons in the T3 CTV genotype specific PCR, while the T30 genotype specific PCR did not yield amplicons with any of the sample templates.

Of the 44 CTV positive Mexican Lime trees, 10 trees (12-7014, 12-7015, 12-7019, 12-7020, 12-7023, 12-7025, 12-7029, 12-7033, 12-7037 and 12-7038) yielded amplicons in 6 of the 8 CTV PCRs used. Six of the 10 trees were grafted with samples collected from the Malelane region, while the other 4 trees were grafted with samples collected in Swaziland (Table 13).

Bayesian phylogenetic analysis performed on sequenced amplicons, did not always confirm the genotype-specific PCR results (Appendix 6), as amplicons sometimes did not cluster within the clades of the expected CTV genotype. We have confirmed here that the T36 RT-PCR is poly-specific and detects not only T36, but also RB Clade 1 and RB Clade 2 sources. It is due to this poly-specific nature of this PCR that sequencing of amplicons followed by phylogenetic analysis was essential as a method to confirm all CTV positive findings, especially for the T36 CTV genotype. Sequence analysis of the T36 PCR derived amplicons using the Mexican Lime samples as templates suggested that the T36 CTV genotype was not detected in any of the 44 samples, with all sequences clustering with either/or both of the RB 1 or RB 2 clades. Amplicons from the T3 PCR also clustered, in 2 of the 4 instances, with the cognate regions of CTAF001623 CTV complete genome; strain SY568 (AF001623), which is not a T3-like genotype. Furthermore, phylogenetic analysis of the amplified region of the B165 RT-PCR failed to differentiate between B165 and the T68-1 genotypes. Sequences of amplicons from the RB Clade 1, RB Clade 2, VT and HA16-5 RT-PCRs clustered as expected in their respective CTV genotypic clusters.

Based on the genotype specific PCR analysis it appeared as though only a single graft-inoculated Mexican Lime tree (12-7006) appeared to contain only one CTV genotype (Table 17). This finding is based on the fact that none of the genotype specific RT-PCR tests other than T36 and RB Clade 1 specific RT-PCR tests yielded amplicons for this sample, and yet the sequence of each of these amplicons clustered in the RB Clade 1 CTV genotypic cluster.

Table 17: CTV genotype specific RT-PCR analysis and alignment of sequences of resultant amplicons using Bayesian phylogenetic analysis on 45 grapefruit budwood graft-inoculated Mexican Lime trees maintained at the greenhouse.

Accession number	Generic	T36/RB	NZRB F1R1	NZRB F2R2	B165	VT	HA16-5	T3	T30
12-7000	+++	+++ X	+++▲	-	+++▲	-	-	-	-
12-7001	+++	+++ Y	-	+++▲	+++▲	-	-	-	-
12-7002	+++	+++ X	+++▲	+++▲	+++▲	+++▲	-	-	-
12-7003	+++	+++ X	+++▲	-	+++▲	+++▲	-	-	-
12-7004	+++	+++ X	+++▲	-	+++▲	-	-	-	-
12-7005	+++	+++ X	+++▲	-	+▲	-	+++▲	-	-
12-7006	+++	+++ X	+++▲	-	-	-	-	-	-
12-7007	+++	+++ X	+++▲	-	+++▲	-	-	-	-
12-7008	+++	+++ X	+++▲	+++▲	+++▲	-	-	-	-
12-7009	+++	+++ X	+++▲	-	+++▲	-	-	-	-
12-7010	-	-	-	-	-	-	-	-	-
12-7011	+++	+++ Y	+++▲	+++▲	-	-	-	-	-
12-7012	+++	+++ Y	+++▲	+++▲	-	-	-	-	-
12-7013	+++	+++ Y	+++▲	+++▲	-	-	-	-	-
12-7014	+++	+++ Y	+++▲	+++▲	-	+++▲	+++▲	+++▲	-
12-7015	+++	+++ Y	+++▲	+++▲	++▲	+++▲	+++▲	-	-
12-7016	+++	+++ Y	+++▲	+++▲	-	+++▲	+++▲	-	-
12-7017	+++	+++ X	+++▲	-	-	+++▲	+++▲	+++▲	-
12-7018	+++	+++ X	+++▲	-	+++▲	+++▲	-	-	-
12-7019	+++	+++ Y	+▲	+++▲	+++▲	+++▲	+++▲	-	-
12-7020	+++	+++ X	+++▲	+++▲	+++▲	+++▲	+++▲	-	-
12-7021	+++	+++ X	+++▲	-	+++▲	+++▲	-	-	-
12-7022	+++	+++ Y	+++▲	+++▲	+++▲	+++▲	-	-	-
12-7023	+++	+++ X	+++▲	+++▲	+++▲	-	+++▲	+++▲	-

12-7024	+++	+++ Y	+++▲	+++▲	+++▲	+++▲	-	-	-
12-7025	+++	+++ Y	+++▲	+++▲	+++▲	+++▲	-	+++▲	-
12-7026	+++	+++ Y	+++▲	+++▲	-	-	-	-	-
12-7027	+++	+++ X	+++▲	+++▲	+++▲	+++▲	-	-	-
12-7028	+++	+ X	+++▲	+++▲	-	-	-	-	-
12-7029	+++	+++ Y	+++▲	+++▲	+++▲	++▲	+++▲	-	-
12-7030	+++	+++ Y	+++▲	+++▲	+++▲	+++▲	-	-	-
12-7031	+++	+ X	+++▲	-	+++▲	+++▲	+▲	-	-
12-7033	+++	+++ Y	+++▲	+▲	+++▲	+++▲	+++▲	-	-
12-7035	+++	+++ Y	+++▲	+++▲	-	-	+▲	-	-
12-7037	+++	+++ Y	+++▲	+++▲	+++▲	++▲	+++▲	-	-
12-7038	+++	+++ Y	+++▲	+++▲	+++▲	+++▲	+++▲	-	-
12-7039	+++	++ Y	+++▲	+++▲	-	+++▲	+++▲	-	-
12-7040	+++	++ Y	+++▲	++▲	-	+++▲	+++▲	-	-
12-7042	+++	+++ Y	+++▲	++▲	-	+++▲	-	-	-
12-7043	+++	+++ Y	-	+++▲	+++▲	-	-	-	-
12-7044	+++	+++ Y	-	+++▲	+++▲	-	-	-	-
12-7045	+++	+++ Y	+++▲	+++▲	-	+++▲	+++▲	-	-
12-7046	+++	+++ Y	+++▲	+++▲	-	+++▲	++▲	-	-
12-7047	+++	+++ Y	+++▲	+++▲	-	-	+++▲	-	-
12-7048	+++	+++ Y	+++▲	+++▲	+++▲	-	+++▲	-	-

T36 RT-PCR analysis samples: **X** = Amplicon clusters with RB1 clade; **Y** = Amplicon clusters with RB2 clade i.e. T36 not found; **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Sample was negative in RT-PCR analysis; **▲:** Confirmed by expected grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Appendix 6).

3.4.2. ***Analysis of the CTV genotype homogeneity of Mexican Lime tree 12-7006.***

In order to confirm that sample 12-7006 is a homogeneous source of a RB Clade 1 member further analysis of this sample was done by targeting the p33 gene region situated within the 3'-teminal half of the CTV genome. Genotype specific RT-PCR followed by sequencing was mainly directed at the 5' terminal half of the CTV genome. The screening of 12-7006 in both ends of the CTV genome yields data with greater credence.

The CTV p33 region-primed RT-PCR analysis performed on sample 12-7006 yielded an amplification product of the expected size (±980bp) for the p33 gene region (Figure 24). Direct Sanger sequencing and Bayesian phylogenetic analysis (Figure 25) produced a dendrogram indicating that the dominant sequence within the amplicon population clustered within the clade incorporating the CTV RB genotype. However the resolution of this genomic region for genotyping is poor as the CTV p33 gene region used is relatively conserved amongst CTV genotypes. It is not possible, based on this region of the genome to differentiate between RB1 and the RB2 (Figure 25), however this does not discount the possibility that sample 12-7006 is a homogenous source (Table 17) of one of the RB clade members.

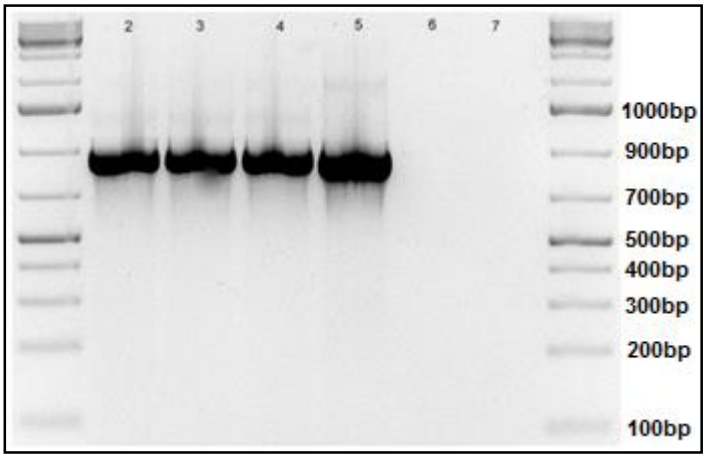


Figure 24: CTV p33 gene region-primed RT-PCR analysis results, indicating the correct product size (±980bp). Lane 2, 3 and 4 represents triplicates of sample 12-7006. Lane 5 represents the positive control, a p33 gene containing plasmid CTV9ΔGFP (supplied by University of Florida, Gainesville, FL 32611, United States), Lane 6 represents the healthy Mexican Lime control and lastly, lane 7 represents the negative nuclease-free water control. Lane 1 and 8 contain HyperLadder IV 100bp (Bioline, Celtic Molecular Diagnostics (Pty) Ltd., South Africa).

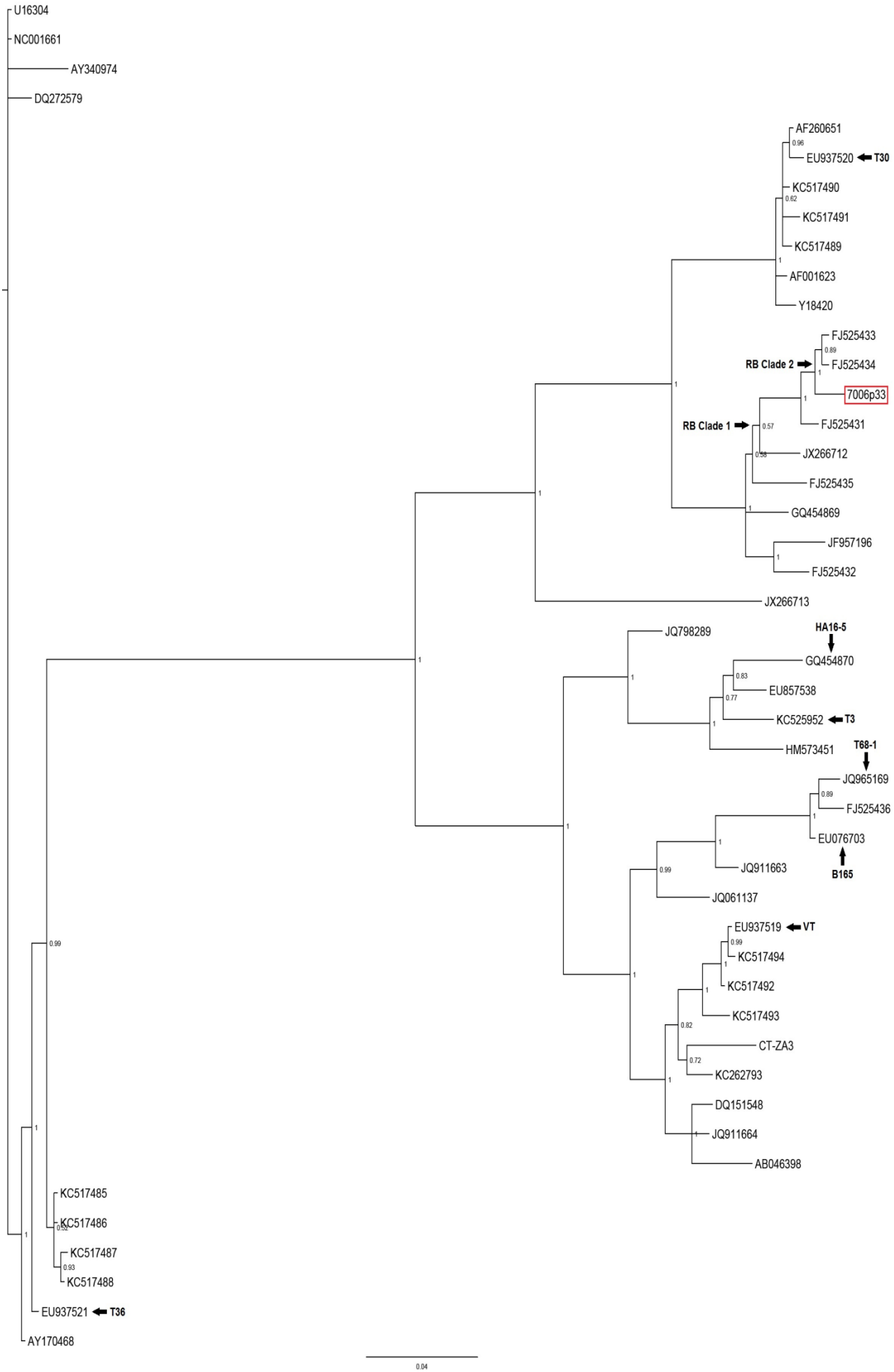


Figure 25: Bayesian dendrogram showing the phylogenetic relationship between the sequence of amplicons to the p33 gene region of sample 12-7006 and cognate sequences of the 45 CTV references, especially the RB genotypic clade. The confidence levels are shown as posterior probability values at each node.

3.4.3. ***Analysis of homogeneity of Mexican Lime tree 12-7006 by sequencing multiple clones of the p33 gene region.***

The cloning and sequencing of multiple clones of the p33 gene region (±980bp) was used to assess the probability that sample 12-7006 represents a pure form of the RB genotype. Of the 100 recombinant plasmids selected (Figure 26), the 12-7006 p33 insert (±980bp) was present in 97. Twenty 12-7006p33 recombinant plasmids were selected for sequencing and phylogenetic analysis using a Bayesian approach. All 20 sequences grouped within the RB Clade 2 (Figure 27) albeit that differentiation from RB1 is poor in this region of the genome. This lends further support that the Mexican Lime tree 12-7006 source represents a potential CTV RB homogeneous source. These preliminary results were followed up by an in depth analysis of the CTV genotype composition in this sample using next-generation high-throughput sequencing on an Illumina platform.

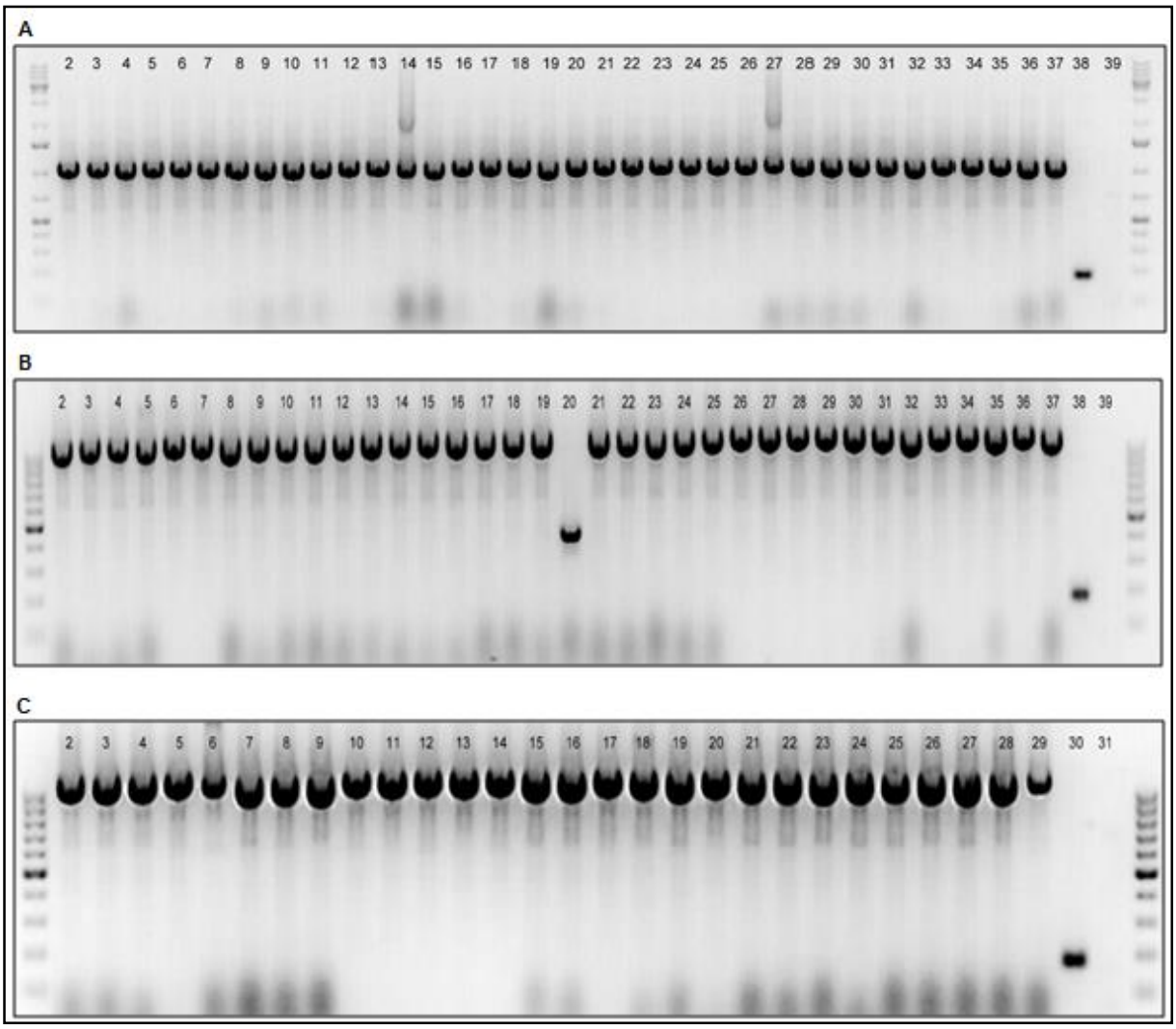


Figure 26: CTV 12-7006 p33 region insert (±980bp) cloning results. [A] Plasmids 1 to 36. Lane 2 to 37 represents 12-7006 p33 plasmids 1 to 36. Lane 14 (plasmid 13) and land 27 (plasmid 26) contained an extra band and therefore was not used for further analysis. Lane 38 represents the CTV positive control, non-recombinant p33 vector, produced and provided by David Read. Lane 39 represents the negative nuclease-free water control. Lane 1 and 40 contain GeneRuler 1kb DNA ladder (Thermo Scientific, Waltham, MA, USA). [B] Plasmids 37 to 72. Lane 2 to 37 represents 12-7006 p33 plasmids 37 to 72. Lane 20 (plasmid 55) contained a smaller insert band and therefore was not used for further analysis. Lane 38 represents the CTV positive control, non-recombinant p33 vector, produced and provided by David Read. Lane 39 represents the negative nuclease-free water control. Lane 1 and 40 contain HyperLadder IV 100bp (Bioline, Celtic Molecular Diagnostics (Pty) Ltd., South Africa). [C] Plasmids 73 to 100. Lane 2 to 29 represents 12-7006 p33 plasmids 73 to 100. Lane 30 represents the CTV positive control, non-recombinant p33 vector, produced and provided by David Read. Lane 31 represents the negative nuclease-free water control. Lane 1 and 32 contain HyperLadder IV 100bp (Bioline, Celtic Molecular Diagnostics (Pty) Ltd., South Africa).

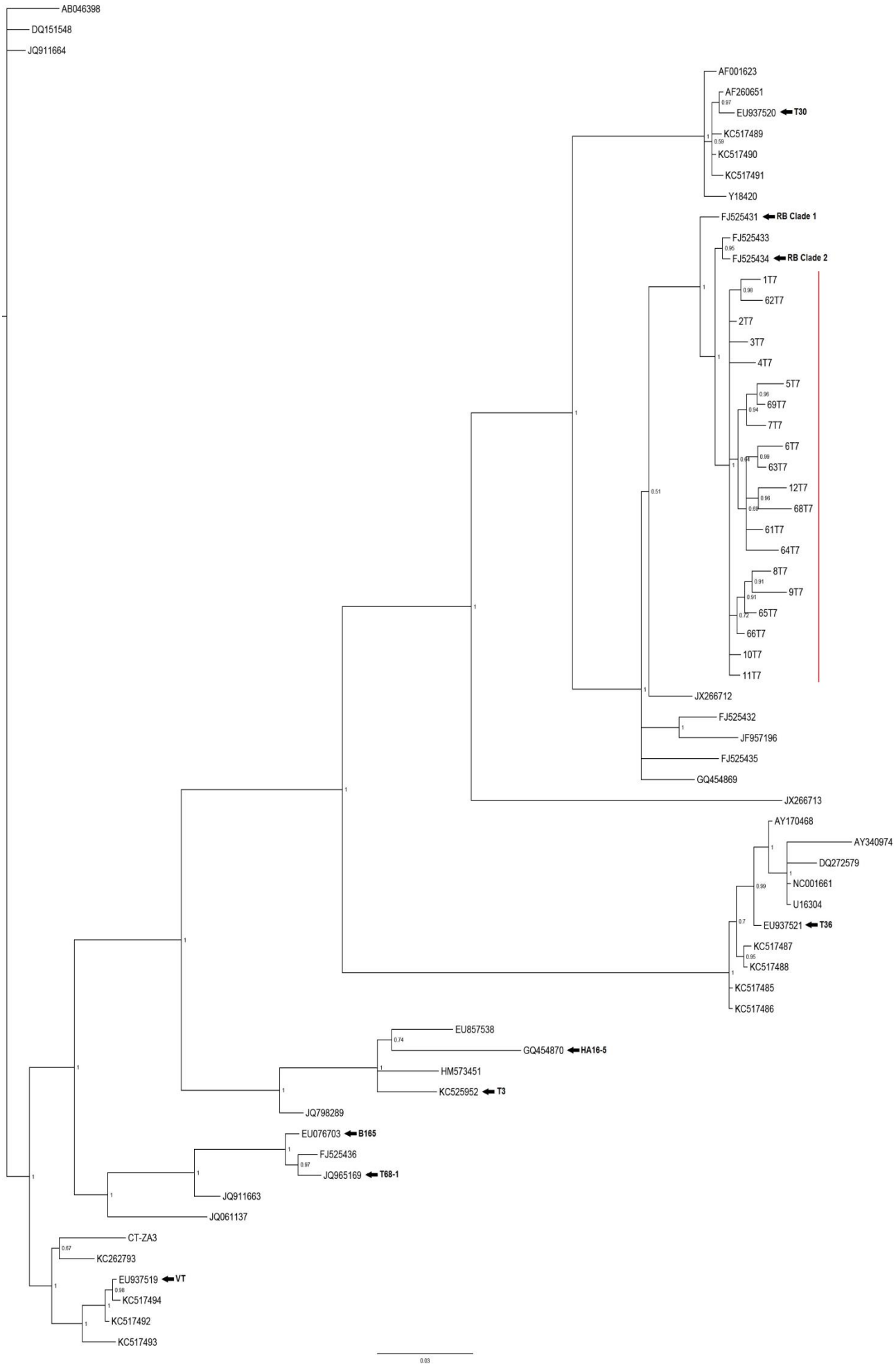


Figure 27: Bayesian dendrogram showing the phylogenetic relationship between the p33 gene sequence of 20 selected 12-7006 recombinant plasmids and the cognate sequence of 45 CTV references. In the dendrogram, the 20 clones groups within the RB genotypic group. The confidence levels are shown as posterior probability values at each node.

3.4.4. Homogeneity analysis of sample 12-7006 with Illumina Next-Generation High-throughput sequencing of the p33 gene.

A total of 169,574 raw paired-end reads were obtained following a sequencing run on the Illumina platform using CTV p33 gene specific amplicons from sample 12-7006 (Table 18). The dataset obtained showed high sequence duplication levels (>80%) (Figure 28a) as well as a good PHRED score of 38 (Figure 28b). The results of reference mapping of the trimmed reads against the sequences of 45 CTV references are presented in Figure 29. The vast majority of reads (102500) mapped to NZRB-TH30 (FJ52534), a member of the CTV RB genotype (Clade 2). This was followed by negligible numbers (8 to a maximum of 1355 reads) mapping to six other isolates of the RB genotype (Appendix 5). This is taken as definitive results suggesting that source 12-7006 is a pure NZRB-TH30 (Accession number: FJ525434) (RB clade 2) source.

Table 18: Illumina data statistics, surrounding dataset quality scores (PHRED) and raw reads processing.

Sample name	Average PHRED score (/40)	Total reads (unfiltered)	Average read length (unfiltered)	Total reads (filtered)	% reads filtered	Average read length (filtered)
12-7006	38	169,574	301	168,281	99.24%	251.5

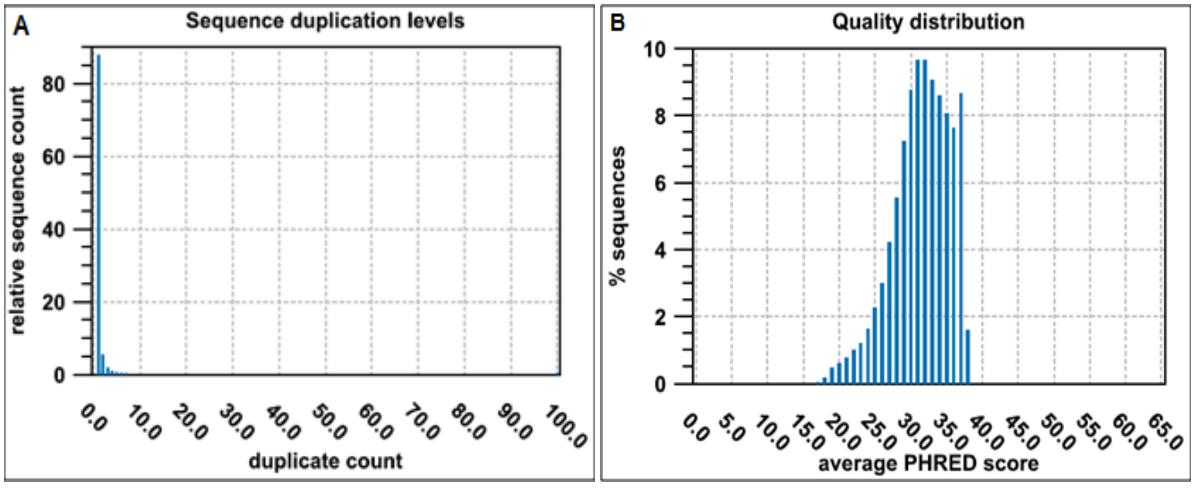


Figure 28: [A] Sequence duplication levels attained for sample 12-7006 during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 12-7006. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

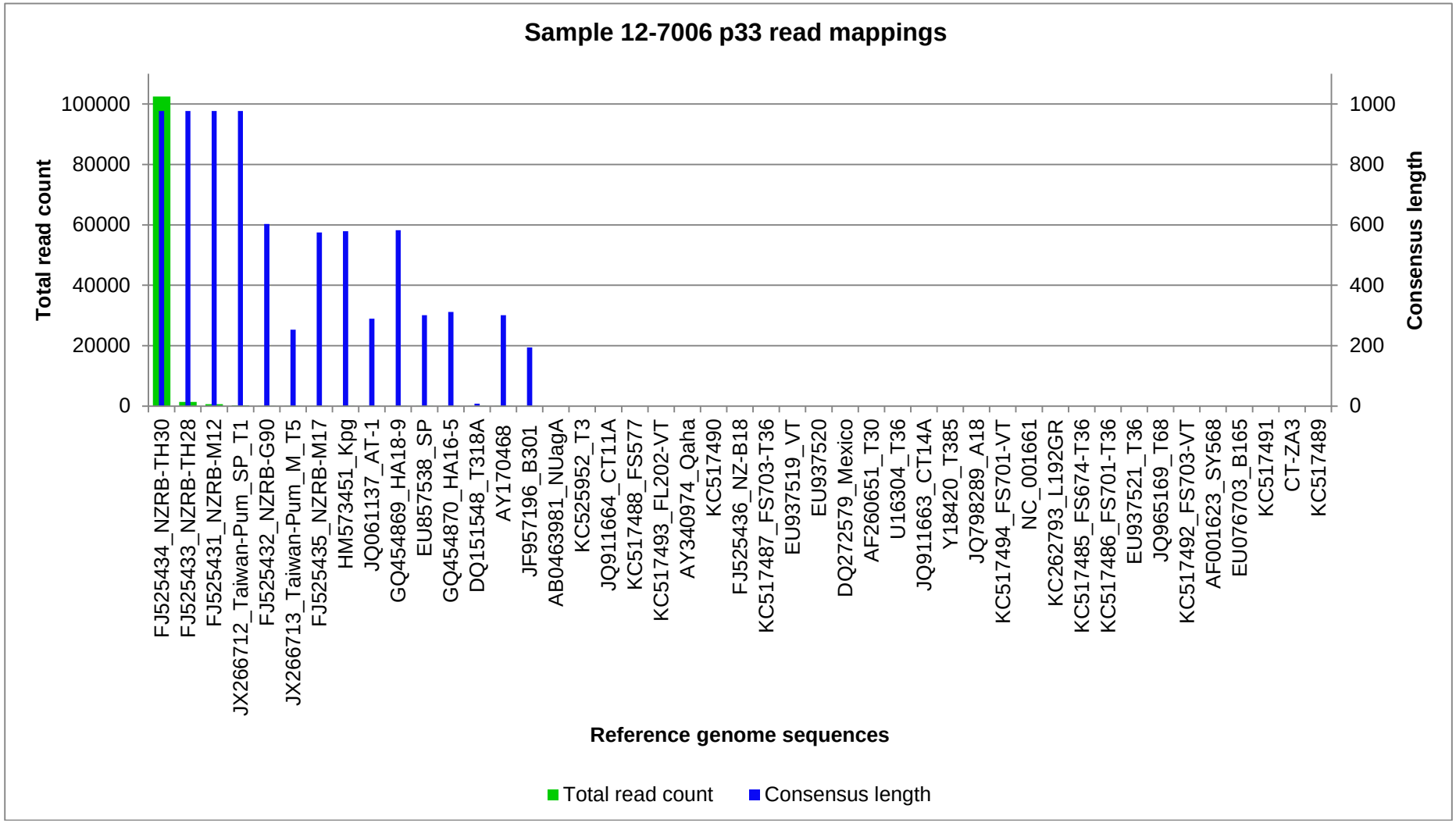


Figure 29: Read mapping results spanning 45 CTV reference genomes for the Mexican Lime tree 12-7006. The mapped reference CTV genotypes were ranked from largest to smallest in terms of the total read counts attained. The primary vertical axis (left) represents the total reads (Green) accounted for and the scale is kept constant to 110 000 throughout the graph. The secondary vertical axis represents the consensus lengths (Blue) in nucleotides for the reads and the scale is kept constant to 1100 base pairs, a few base pairs more than the p33 gene region focused on (±980bp). The horizontal axis represents the 45 CTV reference genome sequences.

3.4.5. Results obtained from field collected Rio-Red samples (13-4012 to 13-4112).

Results obtained from the CTV T36-targetted RT-PCR analysis shows that 95 of the 100 collected Rio Red leaf samples yielded amplicons in this test. In view of this large number of positive samples, only the 25 samples for which budwood had been collected and established on Mexican Lime seedlings (Table 19) were selected for further analysis for the presence of the CTV T36 genotype. Amongst the 25 field collected leaf samples chosen, 24 had tested positive in the CTV T36 test. As sample 13-4040 had tested negative for CTV T36 it was replaced with T36 positive sample 13-4112 for sequencing even though this sample had not been established in the greenhouse by grafting. Bayesian phylogenetic analysis performed on the sequence of the amplicons of these 25 Rio Red samples showed that even though they all yielded amplicons in the CTV T36 RT-PCR, none of the sequences represented those of the CTV T36 genotype. One of the Rio Red samples (13-4048) grouped within the RB Clade 1 genotypic group while the remaining 24 grouped within the RB Clade 2 group (Figure 30).

Table 19: Comparison of the results obtained between the CTV T36-targetted RT-PCR analysis and sequences of the resultant (green blocks) amplicons by using Bayesian phylogenetic analysis. T36 CTV screening was done on the 100 Rio Red grapefruit leaf samples (13-4012 to 13-4112). Twenty-five representatives (X and Y superscripts) were selected for sequencing.

Sample	Positive / Negative	Sample	Positive / Negative	Sample	Positive / Negative	Sample	Positive / Negative
13-4012	+++ Y	13-4038	-	13-4064	+++ Y	13-4090	+++
13-4013	+++	13-4039	-	13-4065	+++	13-4091	+++
13-4014	+++ Y	13-4040	-	13-4066	+++	13-4092	+++
13-4015	++	13-4041	+++	13-4067	+++	13-4093	+++
13-4016	+++ Y	13-4042	+++	13-4068	+++ Y	13-4094	+++
13-4017	-	13-4043	+++	13-4069	+++	13-4095	+++
13-4018	+++	13-4044	+++ Y	13-4070	+++	13-4096	+++ Y
13-4019	++	13-4045	+++	13-4071	+++	13-4097	+++
13-4020	+++ Y	13-4046	+++	13-4072	+++ Y	13-4098	+++
13-4021	+++	13-4047	+++	13-4073	+++	13-4099	+++
13-4022	+++	13-4048	+++ X	13-4074	+++	13-4100	+++ Y
13-4023	+++	13-4049	+++	13-4075	+++	13-4101	+++
13-4024	+++ Y	13-4050	+++	13-4076	+++ Y	13-4102	+++
13-4025	+++	13-4051	+++	13-4077	+++	13-4103	+++
13-4026	+++	13-4052	+++ Y	13-4078	+++	13-4104	+++ Y
13-4027	+++	13-4053	+++	13-4079	+++	13-4105	+++
13-4028	+++ Y	13-4054	+++	13-4080	+++ Y	13-4106	+++
13-4029	+++	13-4055	+++	13-4081	+++	13-4107	+++
13-4030	+++	13-4056	+++ Y	13-4082	+++	13-4108	+++ Y
13-4031	+++	13-4057	+++	13-4083	+++	13-4109	+++
13-4032	+++ Y	13-4058	+++	13-4084	+++ Y	13-4110	+++
13-4033	+++	13-4059	+++	13-4085	+++	13-4111	+++

13-4034	**+++**	13-4060	**+++** [Y]	13-4086	**+++**	13-4112	**+++** [Y]
13-4035	**+++**	13-4061	**+++**	13-4087	**+++**		
13-4036	**+++** [Y]	13-4062	**+++**	13-4088	**+++** [Y]		
13-4037	**-**	13-4063	**+++**	13-4089	**+++**		

T36 RT-PCR analysis samples: **X** = Positive in RB1 RT-PCR; **Y** = Positive in RB2 RT-PCR, i.e. T36 not found; Confirmed by grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Figure 32); **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Negative.

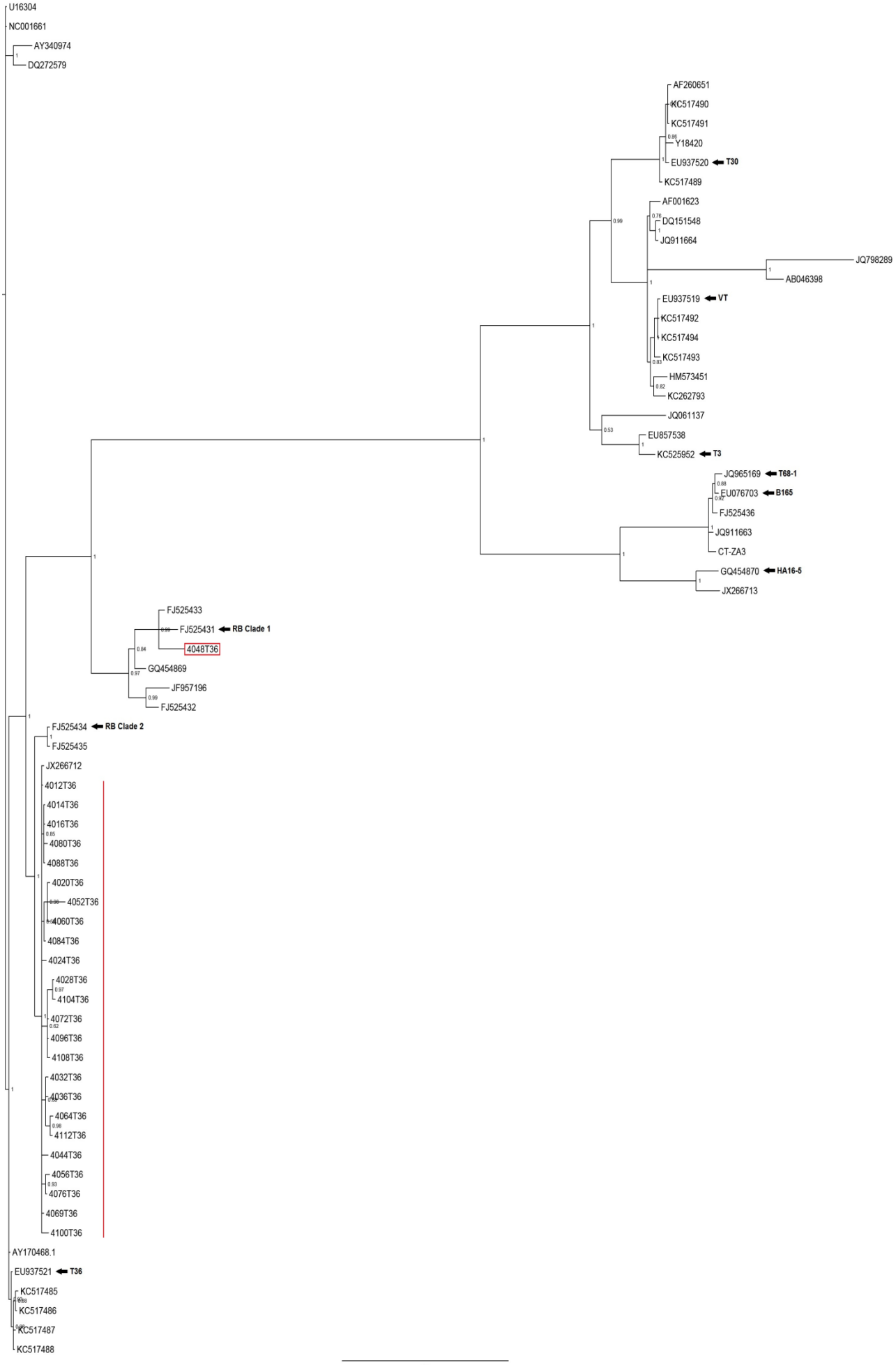


Figure 30: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T36 RT-PCR of the 25 T36 CTV positive Field-collected Rio Red Grapefruit samples and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

3.4.6. ***Results obtained from Rio Red Grapefruit graft-inoculated Mexican lime, greenhouse maintained plants.***

No CTV-like symptoms were detected on any of the 25 Rio Red Grapefruit graft-inoculated Mexican Lime seedlings at 4 months P.I. with only one seedling (13-4118) (Table 20) yielding amplicons in the CTV T36 CTV RT-PCR test (Figure 31). CTV genotype RT-PCR analysis was performed on seedling 13-4118, with the aim of determining whether this source consists of a single homogeneous CTV genotype. Amplicons were obtained from T36, RB, HA16-5 and VT-RT-PCRs (Table 21); hence we can conclude that seedling 13-4118 is infected with multiple CTV genotypes.

Bayesian phylogenetic analysis (Figure 32 - 36) on the sequence of amplicons of the genotype “specific” T36 RT-PCR with seedling 13-4118 as template indicated that the amplicon sequence clustered with RB Clade 2 ([CTV isolate Taiwan-Pum/SP/T1 (JX266712)]. Amplicons of sample 13-4118 from the RB 2 specific RT-PCR also clustered within RB Clade 2 [CTV isolate Taiwan-Pum/SP/T1 (JX266712)]. Amplicons from the VT specific RT-PCR for this sample clustered within the VT genotype [CTV isolate Kpg3 (HM573451)]. Sequences of amplicons from the HA16-5 specific RT-PCR clustered as expected with the HA16-5 CTV genotype.

Since no T36 sources were obtained amongst any of the 2013 graft-inoculated Mexican Lime seedlings or the field collected leaf material collected, no further analyses were done on these plants.

Table 20: Additional information, and the comparison of the CTV T36-targetted RT-PCR analysis results and Bayesian phylogenetic analysis of the resultant 13-4118 (In bold) amplicon. T36 CTV screening was done on the 25 Rio Red budwood graft-inoculated Mexican Lime seedlings, 4 months P.I.

Point	Budwood samples collected	Date graft-inoculated	Citrus tree grafted onto	Mexican Lime accession number	T36 Genotype present
1	13-4012	2013/05/03	Mexican Lime	13-4113	-
3	13-4014	2013/05/03	Mexican Lime	13-4114	-
5	13-4016	2013/05/03	Mexican Lime	13-4115	-
9	13-4020	2013/05/03	Mexican Lime	13-4116	-
13	13-4024	2013/05/03	Mexican Lime	13-4117	-
17	**13-4028**	**2013/05/03**	**Mexican Lime**	**13-4118**	**+++ Y**
21	13-4032	2013/05/03	Mexican Lime	13-4119	-
25	13-4036	2013/05/03	Mexican Lime	13-4120	-
Negative	13-4040	2013/05/03	Mexican Lime	13-4121	-
33	13-4044	2013/05/03	Mexican Lime	13-4122	-
37	13-4048	2013/05/03	Mexican Lime	13-4123	-
41	13-4052	2013/05/03	Mexican Lime	13-4124	-
45	13-4056	2013/05/03	Mexican Lime	13-4125	-
49	13-4060	2013/05/03	Mexican Lime	13-4126	-
53	13-4064	2013/05/03	Mexican Lime	13-4127	-
57	13-4068	2013/05/03	Mexican Lime	13-4128	-
61	13-4072	2013/05/03	Mexican Lime	13-4129	-
65	13-4076	2013/05/03	Mexican Lime	13-4130	-
69	13-4080	2013/05/03	Mexican Lime	13-4131	-
73	13-4084	2013/05/03	Mexican Lime	13-4132	-
77	13-4088	2013/05/03	Mexican Lime	13-4133	-
85	13-4096	2013/05/03	Mexican Lime	13-4134	-
89	13-4100	2013/05/03	Mexican Lime	13-4135	-
93	13-4104	2013/05/03	Mexican Lime	13-4136	-
97	13-4108	2013/05/03	Mexican Lime	13-4137	-
101	13-4112				

Negative: CTV negative seedling. T36 RT-PCR analysis samples: **X** = RB1; **Y** = RB2, i.e. no T36 was identified; Confirmed by grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Figure 34); **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Negative.

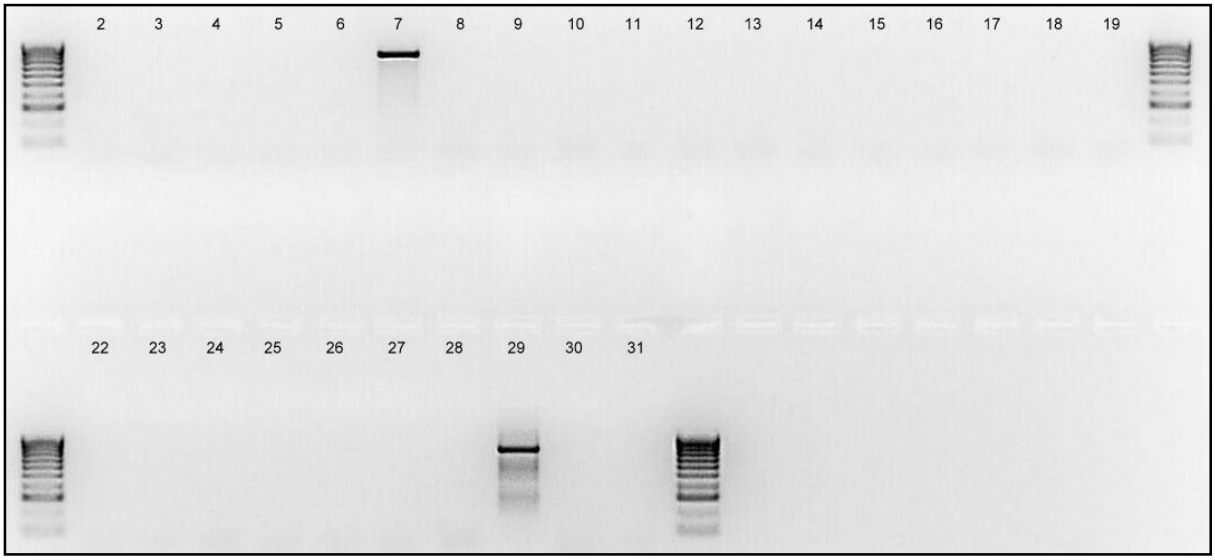


Figure 31: CTV T36-targetted RT-PCR analysis results of the 25 Rio Red budwood graft-inoculated Mexican Lime seedlings 4 months P.I. Lane 2 to 19 represents samples 13-4113 to 13-4131, lane 22 to 28 represents samples 13-4132 to 13-4137. Lane 7 represents sample 13-4118, the only seedling that tested positive for the CTV T36 genotype. Lane 29 represents the CTV T36 positive control. Lane 30 represents the healthy Mexican Lime control and lastly, lane 31 represents the negative nuclease-free water control. Lane 1, 20, 21 and 32 contain HyperLadder IV 100bp (Bioline, Celtic Molecular Diagnostics (Pty) Ltd., South Africa).

Table 21: Results of CTV genotype specific RT-PCR analysis and sequences of the resultant amplicons using Bayesian phylogenetic analysis. CTV screening was done on the budwood graft-inoculated Mexican Lime Seedling 13-4118 to determine the CTV genotypic presence.

Graft-inoculated sample	Genotype	RT-PCR results
13-4118	Generic**	+++**
	T36	+++ Y
	NZRB F1R1	+++ ▲
	NZRB F2R2	+++ ▲
	HA16-5	+++ ▲
	VT	++ ▲
	B165	-
	T30	-
	T3	-

T36 RT-PCR analysis samples: **X** = RB1; **Y** = RB2, i.e. no T36 was identified; **+++**: Definite positive, bright band; **++**: Positive, clear band present; **+**: Positive, extremely faint band (low concentrations of the specific CTV genotype present); **-**: Sample was negative in RT-PCR analysis; **▲:** Confirmed by grouping of sequenced amplicons in dendrograms based on Bayesian phylogenetic analysis (Figure 34 – 38). **Generic RT-PCR analysis was only performed to determine if the seedlings are CTV positive, no phylogenetic analysis was done on the generic positive amplicons.

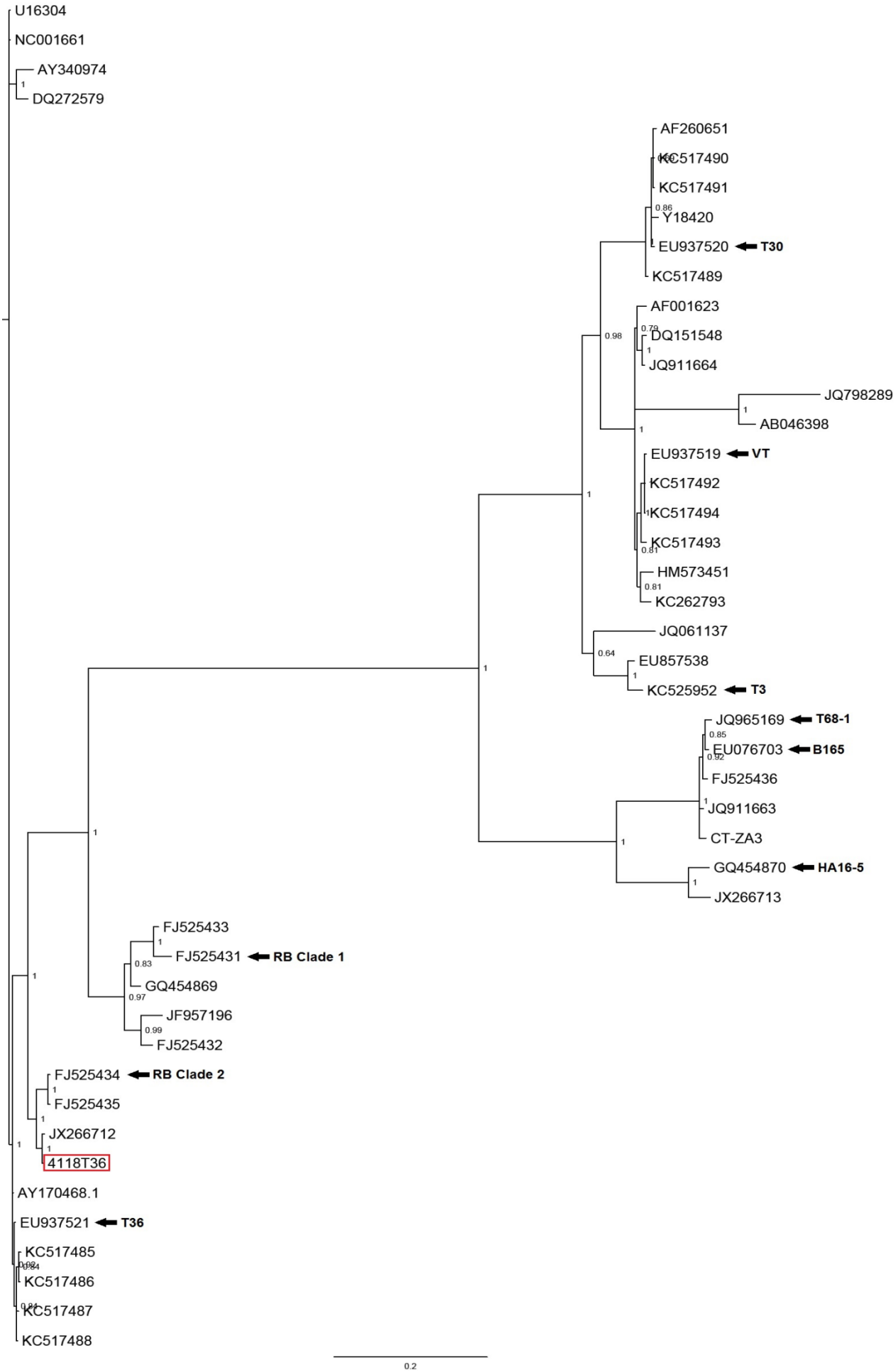


Figure 32: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T36 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime seedling 13-4118 and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

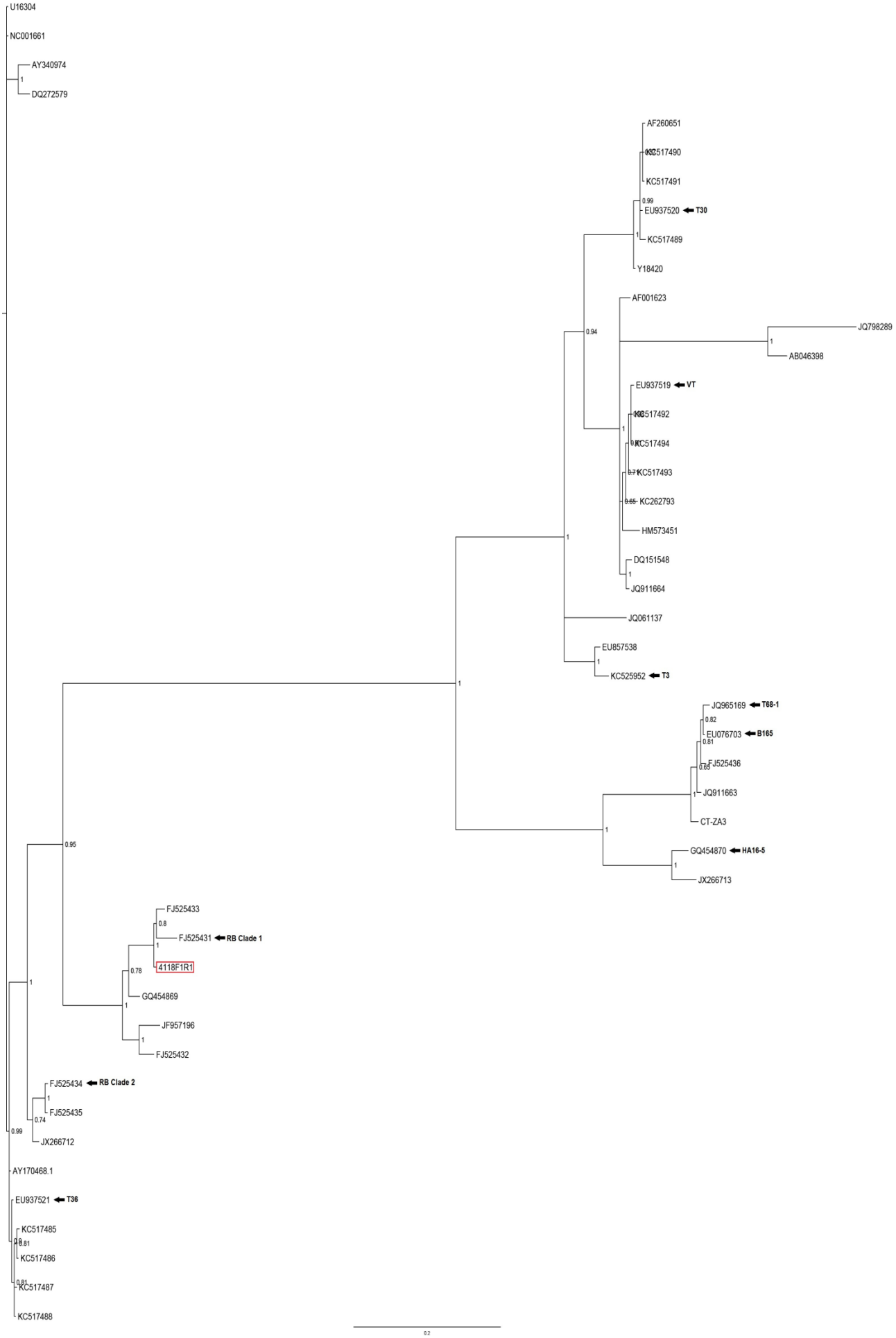


Figure 33: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F1R1 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime seedling 13-4118 and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

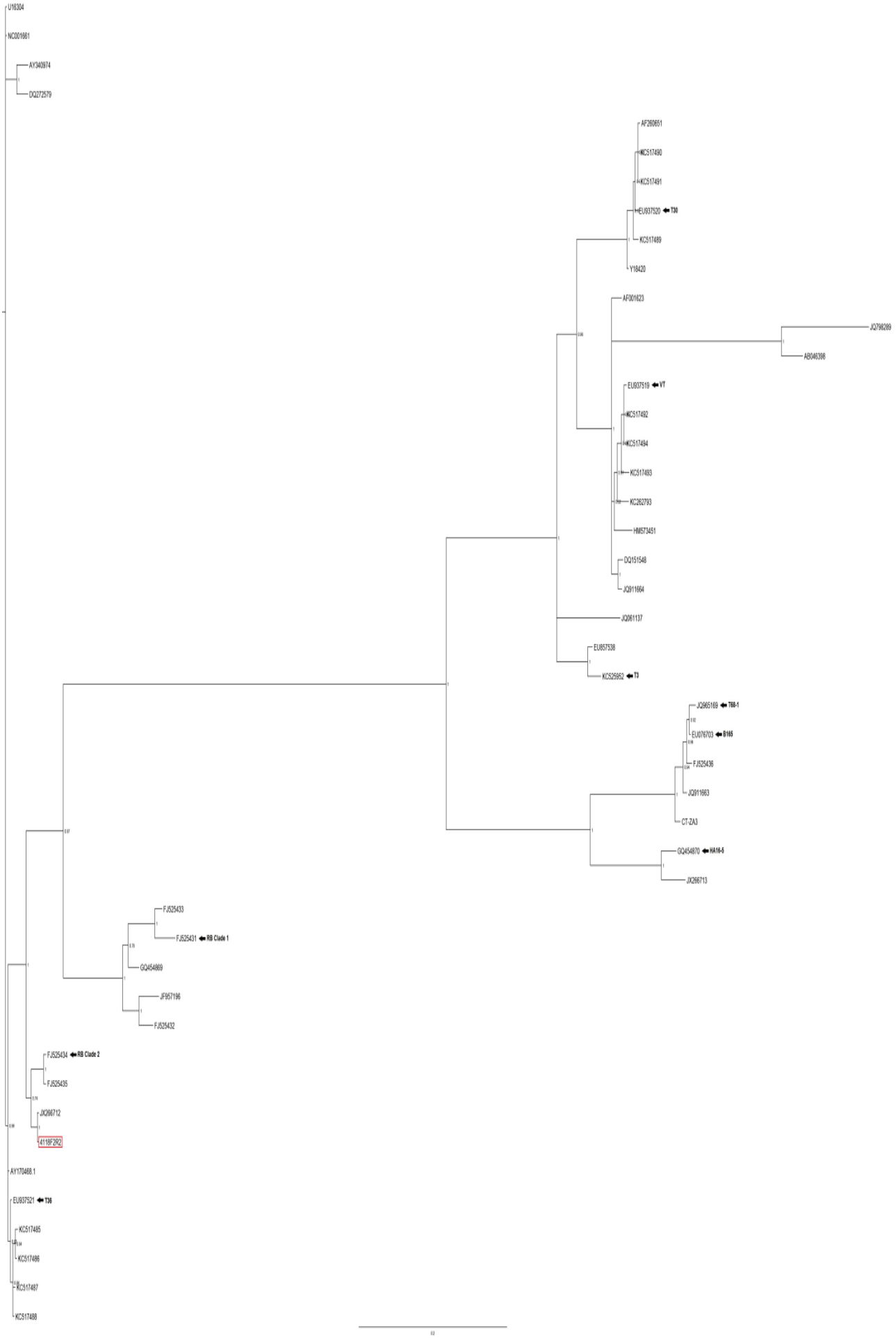


Figure 34: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F2R2 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime seedling 13-4118 and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

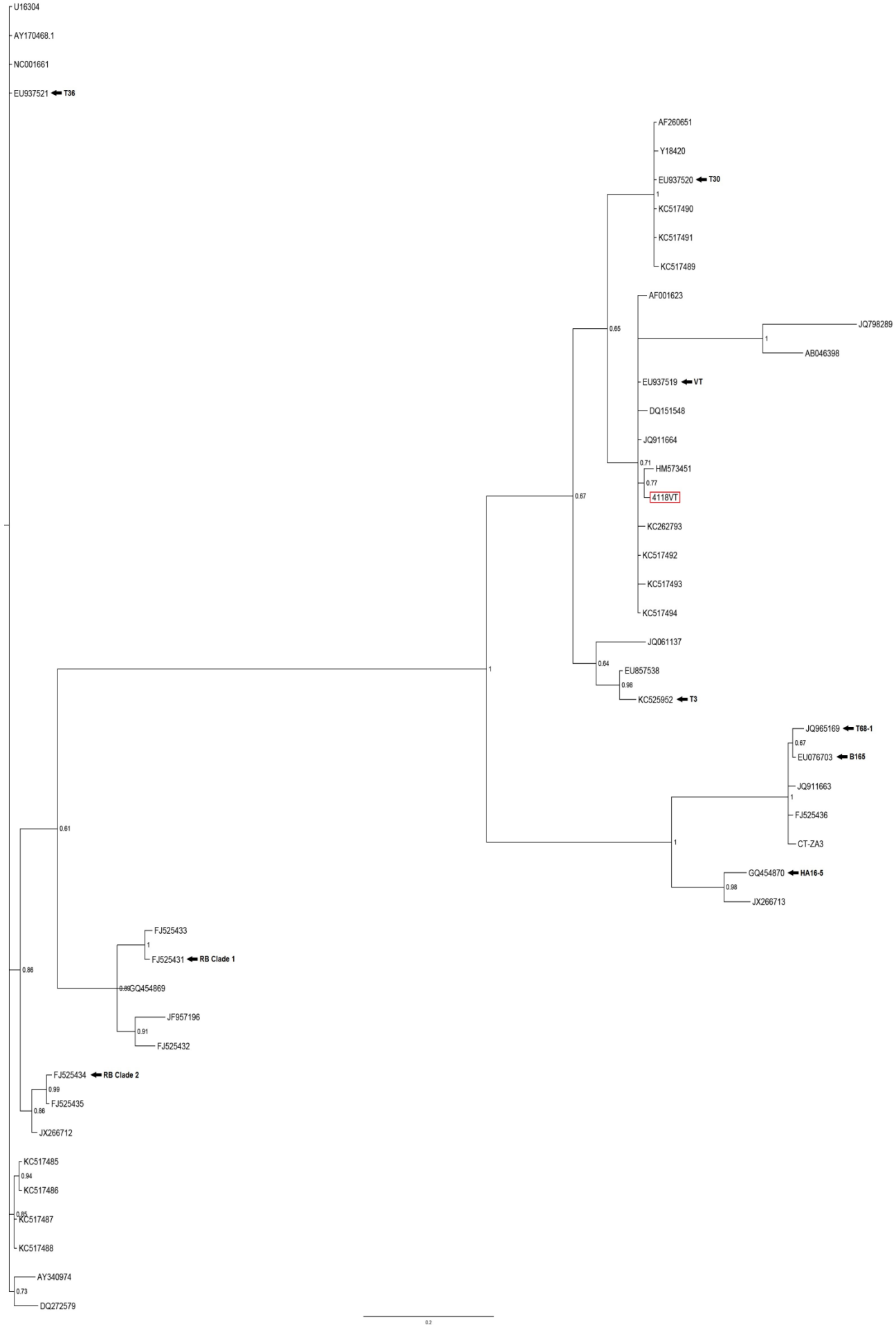


Figure 35: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the VT RT-PCR of grapefruit budwood graft-inoculated Mexican Lime seedling 13-4118 and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

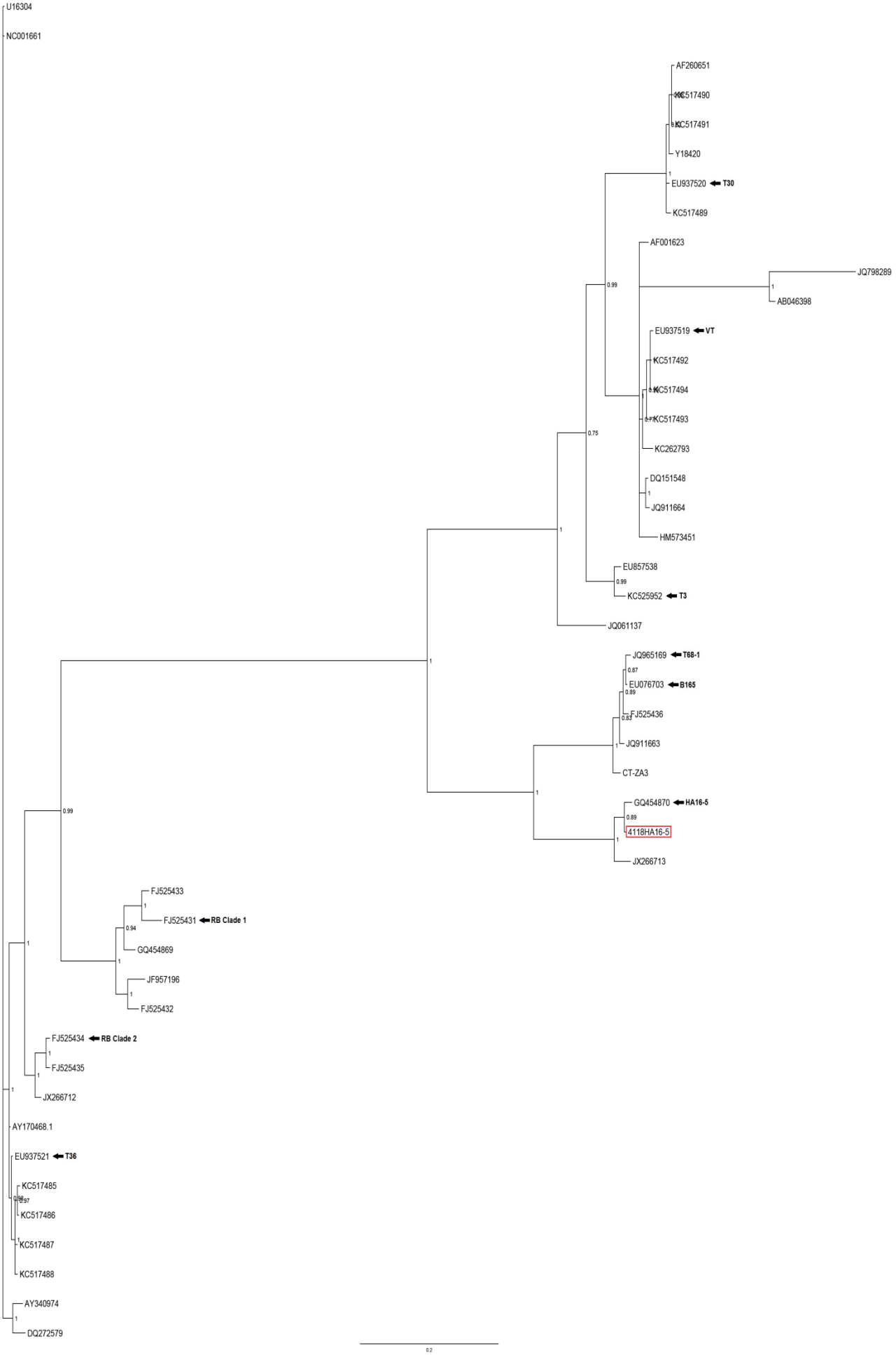


Figure 36: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the HA16-5 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime seedling 13-4118 and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

3.5. Discussion.

The aim of this study was to confirm the presence in Southern Africa of the CTV T36 genotype as found in a 2012 preliminary survey by our group, and to establish it within our greenhouse. However this study revealed that identification of T36 during the previous study was a misdiagnosis, and that T36 was not present in the original sample or in other grapefruit trees within that orchard. This necessitated us to modify our objectives. As we had found large numbers of grapefruit trees infected with CTV of either of the two RB genotypes we decided to find (or create by isolation techniques) a pure population of this genotype and to establish it in our greenhouses in order to assess its suitability as a mild strain cross protecting source. To determine whether the RB infected samples contained only this genotype they were also screened for other known genotypes using genotype specific RT-PCR. A single 2012 graft-inoculated Mexican Lime tree (12-7006) did not appear to have mixed infections of RB with other genotypes, and was therefore further assessed for its homogeneity of infection by progressively more robust, but also increasingly more expensive methods of virus population characterization. Direct sequencing of the p33 gene using primers capable of amplifying this gene from all genotypes confirmed the dominant nature of RB within the population. Sequencing of 20 clones of these amplicons further supported the homogeneity of this source. Finally Illumina Next-Generation High-throughput sequencing of the p33 gene amplification products confirmed the purity of this source. In this latter technique hundreds of thousands of sequence reads generated from the amplicons of the p33 gene were mapped to the cognate p33 gene sequences of CTV sources for which whole genome sequences are available. The mapping of a small number of sequences to other CTV reference sequences is discussed in more detail later within this section.

CTV was shown to be successfully transmitted in 44 of the 45 samples collected and grafted during the preliminary survey of grapefruit by our laboratory in 2012 (Smocilac, *personal communication*). This included the Rio Red sample in which T36 was originally identified. All 44 grafted plants yielded amplicons in the T36 specific (Roy *et al.*, 2005) RT-PCR test. However, upon sequencing the amplicons from these reactions all contained sequences clustering within one of the two RB genotype clades rather than with T36 strains. Furthermore, of a further 100 samples collected from the same orchard in which the original T36 result was obtained, 95

again tested positive in T36 specific PCR but yielded amplicons with sequences related to one of the two RB genotype clades . These results suggested that the T36 “specific” PCR was in fact poly-specific and was capable of amplifying members of the CTV RB genotype as well.

As stated in an earlier section (*1.2.6. Studies of CTV variation*), the RB genotype had not yet been discovered (Harper *et al.*, 2010; Roy *et al.*, 2010; Roy *et al.*, 2013) at the time the CTV-genotype specific primers were designed (Hilf and Garnsey, 2000; Hilf *et al.*, 2005; Roy *et al.*, 2005). However, Roy *et al.* (2010) confirmed that regions of identical sequences were shared, within ORF1a, between the T36 primer set developed and the two RB strains mentioned above [NZRB-M17 (FJ525435) and NZRB-TH30 (FJ525434)]. For further confirmation, an alignment of all available RB genotypic strains together with the T36 primer set was done. Within this alignment very little sequence differences were present in the area between the primer binding sites and respective sequences (results not shown), giving a clear indication why RB was detected through the T36 RT-PCR analysis.

This suggested that the original identification of a T36 CTV source in Rio Red tree 12-7047 was possibly due to misidentification of an RB source. The Smocilac, 2012 study (*unpublished data*) made use of a modified A-fragment primer set (Read, *unpublished*), which amplifies a portion of the open reading frame 1a (ORF1a) in the 5’-terminal half of the CTV genome. Close analysis of the Neighbor-Joining phylogenetic dendrogram created during that study (Appendix 7) revealed that the sequence for a clone of the amplicon from sample 12-7047 clusters within a clade comprising strains of both RB and T36 genotypes, and was most closely related to Taiwan-Pum/Sp/T1, a strain now also known to be a member of the RB genotype. At the time of analysis however this cluster was interpreted as being T36-like (Smocilac, *unpublished results*). Harper *et al.*, (2010) demonstrated that five RB strains were T36-like within the ORF1a region of the CTV genome where they shared a 90.3% sequence identity with the T36 CTV genotype. The identification of a T36 source from 12-7047 Rio Red grapefruit is thus erroneous and it should be identified as a RB infected source.

As none of the samples collected in searching for a T36 source yielded amplicon sequences related to T36, but contained sources of RB it was decided that the objective of the study be altered and that it may be worthwhile screening them in order to possibly find pure sources of the CTV RB genotype in order to assess its usefulness as a cross protecting source.

Genotype specific RT-PCRs were therefore conducted on 24 representative Rio-Red samples collected in 2013, which had also been established in the greenhouse through graft-inoculations onto Mexican Lime, along with the 45 samples inoculated in 2012. Sanger sequencing was conducted on amplicons generated from these genotype specific RT-PCR's, followed by Bayesian phylogenetic analysis. In addition to the Roy *et al.* (2010) primers, two RB "specific" primer sets (Cook, *unpublished*) were added to the study.

As different results are possible when using different phylogenetic models (Darriba *et al.,* 2012), three different phylogenetic software programs were assessed and utilized in this study viz.; Mega version 5 (Tamura *et al.*, 2011) for Neighbor-Joining dendrograms, PhyML version 3.0 (Guindon and Gascuel, 2003) for Maximum Likelihood dendrograms and MrBayes version 3.2.1 (Ronquist *et al.*, 2012) for Bayesian phylogenetic trees, of which the latter provided the most robust results. The Neighbor-Joining algorithm provided the weakest bootstrap support for the differentiation of the known CTV genotypes. When comparing Maximum Likelihood and Bayesian algorithm derived phylogenetic trees, the Bayesian dendrograms, with parameters utilized in this study allowed all the different CTV genotypes to cluster in such a way that a distinct genotypic separation can be seen, hence using this phylogenetic analysis program to confirm all "genotype specific" RT-PCR positive results.

Rio Red samples collected and analysed directly or grafted are derived from trees planted in 1994, and in all probability had been pre-immunized with GFMS12. The South African Citrus Improvement Program (SACIP) was initiated in 1973 (Scott *et al.*, 2012; von Broembsen & Lee,1988), when citrus trees started being pre-immunized against CTV with the then cross-protecting source GFMS 12. The first budwood protected by GFMS 12 was released during 1984 (Marais, 1994). Grapefruit mild strain 12 was initially chosen as the cross-protecting source for all

citrus, especially grapefruit. The results obtained inthe current study emphasizes the importance of knowing the different CTV genotypes present and having a a specific cross-protecting source against each genotype (Scott *et al.*, 2012; Read *et al.*, 2015). VT and RB were detected within most of these samples and, aphid transmission of these CTV genotypes within citrus orchards could have played a role in the specific genotypic dominance (Scott *et al.*, 2012). Scott *et al*. (2012) showed that the CTV population in grapefruit inoculated with GFMS 12 consisted of VT, T30, RB and B165 genotypes. This was followed by Read *et al.* (2015) using slightly different methodology and focusing mainly on GFMS 12 and GFMS 35. The results obtained compared very well to the data published by Scott *et al.* (2012), especially within the A-region of the CTV genome as T30, the T68/CTZA group and VT/Kpg group were the dominant sources. From these findings it has been suggested that VT/Kpg-like and T30-like variants, detected in both studies are components of the original GFMS 12 cross-protecting source. These findings acts as a further confirmation on earlier year's detection of the mixed CTV composition of the GFMS 12 source (Stewart, 2006; van Vuuren *et al.*, 2000).

The RB source was successfully graft-transmitted from the 25 Rio Red grafted samples onto only one (13-4118) of the Mexican Lime seedlings. The low success rate for the graft-transmissions may be due to lack of callus tissue formation, grafts being too old or dried out or with too little phloem exposed on both surfaces to allow proper binding to take place. As this seedling, along with the other grafted established sources, all were shown to contain mixed infections of various genotypes while we were trying to obtain single genotype infections no further analysis on them were conducted. This was offset by the fact that sample 12-7006 was shown to be a homogenous RB source. The sequencing of multiple clones of this sample confirmed that the RB genotype was indeed the predominate source. Cloning is a well-known procedure used to determining whether samples contained mixed CTV infection or pure sources of the virus (Al-sadi *et al.*, 2012; Roy and Brlansky, 2004). This procedure has also been used in several studies to aid in genetic differentiation and structure of CTV populations (Melzer *et al.*, 2010; Yokomi *et al.*, 2011; Read and Pietersen, 2015). However, cloning is seen as a more old conventional approach in viral discovery, while the newly introduced Next-generation sequencing procedure is more advanced, sensitive and faster in detection and analysis of samples (Liu *et al.*,

2011). In short, NGS allows enormous volumes of sequence data to be produced; in some cases more than one billion short reads per instrument run, in a very short period of time (Barba *et al*., 2014). Highly accurate (genomic) results have also been obtained through NGS analysis. Additionally, bioinformatics software tools and algorithms are continuously developed and improved to assist in the ever-growing nature of NGS technologies. These improvements will assist in future NGS studies, in which longer DNA reads will be produced, where the identification of novel pathogens (especially viral diagnostics) can be expedited and increased (Sebastien *et al*., 2014). At this point in time a DNA read length of 750bp can be produced (Barba *et al*., 2014). NGS analysis has already been used in disease etiology as well as viral ecology (Sebastien *et al*., 2014). NGS is a better and more reliable diagnostic and detection method for future studies, compared to sequencing multiple clones.

The Illumina sequencing results obtained for sample 12-7006 clearly demonstrated the quasispecies nature of the viral population (Eigen, 1971; Eigen and Schuster, 1977). The quasispecies concept is described as a well-defined distribution of mutants generated through a specific mutation-selection process (Domingo, 1992). While the RB genotypic strain NZRB-TH30 sequences were the dominant reads mapped, with high total read count (102500 reads) and with a consensus sequence of 977bp only 3 base pairs shy of the complete expected p33 gene region (±980bp), some reads were obtained for several other CTV genotypic strains. Of these, the 5 reference genome sequences with the next highest read counts or lengthiest consensus lengths were all other strains of the RB genotype. As the differences between strains in this part of the CTV genome are often only a few nucleotides, the low read counts to these strains are interpreted as representing minor point mutations in those nucleotide positions, as is typical of viral quasispecies populations. Similarly, when analyzing the consensus lengths obtained for the other RB strains to which sequence reads were obtained, it is evident that an almost complete (977bp) p33 gene region consensus length was obtained for strains NZRB-TH28 (FJ525433), NZRB-M12 (FJ525431) and Taiwan-Pum/SP/T1 (JX266712), while NZRB-G90 (FJ525432) only obtained a consensus length of 603bp, much shorter than the other four RB strains. This reflects the sequence similarity between these RB strains within this genome segment. While some of the single nucleotide

differences observed may be due to small errors in sequences introduced during amplification or sequencing of the templates, some variation within even a genotypically homogeneous CTV population can be expected due to the error threshold during the replication cycle of RNA viruses (Eigen and Schuster, 1977; Nowak, 1992). Error-free replication cycles would have led to an evolutionary halt, since no mutants or variants would have been produced. However, an overly high replication error rate can also prevent evolution from occurring, as new progeny might be too genetically diverse to the "parental" strand (Lauring and Andino, 2010). This high genetic diversity between the different variants produced can lead to a reduced fitness level or even deterioration. The fitness of a strain can be defined as how well the strain is able to "fit in" with a new environment, for example how well it adapts and reproduce in different conditions (Holmes and Moya, 2002; Orr, 2009). Also, the viral pathogenicity of each strain can also change throughout this entire process, making disease control strategies extremely difficult (Domingo, 1992; Domingo *et al.*, 2001; Lauring and Andino, 2010; Domingo *et al.*, 2012).

Having confirmed the existence of an essentially homogenous RB CTV source, future studies are required to determine whether this source has potential for cross protection by inducing only mild symptoms on a range of citrus hosts including grapefruit. In addition to the above, no CTV specific symptoms were identified on the source plant (12-7006). Folimonova *et al.* (2010) published a study in which they focused on determining the relationships between different CTV genotypes in terms of how they will aid in cross-protection of citrus cultivars. Cross-protection, thought to be based on a superinfection exclusion phenomenon, can be defined as the process in which citrus cultivars are potentially protected in which primary viral infections (mild CTV strains) prevent secondary infections (severe CTV stains) from a similar or the same viral strain to take place (Folimonova *et al.*, 2010).

For cross-protection of South Africa citrus cultivars we need to obtain mild strains of all genotypes circulating locally in order pre-immunize citrus trees, and thereby lead to super-infection exclusion of any related genotypes (Folimonova, 2013). However, most CTV infections consist of a mixture of genotypes (or strains) (Roy and Brlansky, 2004). Studies performed on potential cross-protecting isolates have also shown that the composition of these isolates is of major concern

(Roistacher *et al.*, 2010; van Vuuren *et al.*, 2000). The evolution of new viral strains can occur through complex CTV mixtures interacting with each other within the environment (Roy and Brlansky, 2004). Following biological indexing of SAT isolates, three cross-protecting isolates LMS 6, GFMS 12 and GFMS 35 were shown to consist out of mixed CTV infections, with a composition of severe and mild components (van Vuuren *et al.*, 2000; Luttig *et al.*, 2002). Thus, the detection of isolates that contain all the necessary CTV genotypes (and strains), in both mild and severe form, is of great importance to obtain a 100% cross-protection success rate (Roy and Ramachandran, 2002). Several attempts to obtain pure CTV sources, as well as determining the source severity, through single aphid transmissions (SAT's) or graft-inoculations, onto indicator hosts, have been performed in several studies (Roistasher, 1991; Lbida *et al.*, 2005; Zhou *et al.*, 2007; Korkmaz *et al.*, 2008; Moreno *et al.*, 2008; Folimonova, 2013; Albiach-marti, 2013; Abbas *et al.*, 2013).

An earlier study by Müller *et al.* (1974), made use of single aphid transmissions and graft inoculations onto *Passiflora gracilis* and Mexican Limes, to obtain attenuated or homogenous CTV source. According to Müller *et al.* (1974) and Roistacher *et al.* (2010), genotypic attenuation or potential homogenous CTV source were observed through this process. Therefore, by combining this inoculation process and molecular CTV detection, several genotypic strains can be identified for the cross-protection of citrus. Thus, since we confirmed that we obtained a pure RB CTV source, this procedure can also be performed to determine the severity of the strain.

In addition to the difficulty level and the lengthiness of each procedure in obtaining pure CTV sources from mixed infections, the introduction as well as the spread of different CTV genotypes within citrus orchards is of great importance. Citrus trees within orchards are continuously exposed and challenged with viruses such as CTV, which leads to the presence of more than one genotypic strain within these trees (Roy and Brlansky, 2004). Using indicator hosts such as Mexican Lime, Sour Orange, Duncan Grapefruit and Madam Vinous Sweet Orange cultivars, can be used to assist in reducing the amount of CTV genotypes present within a source plant. This process may occur by the specific indicator hosts selecting for specific CTV genotypes. Grafting plant material collected from naturally occurring citrus trees onto indicator hosts can 1) aid in lowering the genotypic strain presence to a more

pure or homogenous CTV source, which can be repeated to obtain a pure form, and 2) can assist in developing new variants as CTV strains will interact with each other during the graft-inoculation process (Albiach-marti, 2012). Single aphid transmissions together with graft-inoculations with above mentioned indicator hosts have been used by Müller *et al*. (1974) to obtain possible homogenous CTV sources. Mexican Lime trees, known as the universal indicator host, have shown to be the best indicator host for CTV, showing clear CTV symptoms shortly after infection trials (Garnsey *et al.* 1991; Roistacher *et al.*, 1991; Targon *et al.*, 2000; D'Onghia *et al.*, 2009; Hancevic *et al.*, 2013). Lastly, it has also been shown that the severity of the CTV genotypes or strains found can also be determined through indicator hosts.

In conclusion, a number of samples screened yielded amplicons within the T3 genotype specific RT-PCR analysis. However, the sequences obtained from the T3 amplicons did not all cluster within the T3 genotypic group during phylogenetic analysis. While this may partly be due to small sequencing errors, it is also possible that new CTV T3 strains (12-7023 and 12-7025) have been found. This requires further investigation.

3.6. References.

Abbas, M., Khan, M. M., Mughal, S. M. and Ji, P. 2013. “Characterization and assessment of seasonal variation of Citrus tristeza closterovirus (CTV) in citrus in Pakistan.” *Journal of Food, Agriculture and Environment* 11(3 & 4): 1063–1068.

Abdelmaksoud, H.M. and Gamal El-din, A.S. "Complete Nucleotide Sequence of Citrus Tristeza Virus Qaha isolate." *Unpublished.* Genbank Submitted (07-JUL-2003) (AY340974.1) Plant Virology Research Department, Plant Pathology Research Institute, 9 Gamma Str, Giza 12619, Egypt.

Albiach-martı, M. R., Mawassi, M., Gowda, S., Satyanarayana, T., Hilf, M. E., Shanker, S., Almira, E. C. Vives, M. C., Lopez, C., Guerri, J., Flores, R., Moreno, P., Garnsey, S. M. and Dawson, W. O. 2000. “Sequences of Citrus Tristeza Virus Separated in Time and Space Are Essentially Identical.” *Journal of Virology* 74 (15): 6856–6865. doi:10.1128/JVI.74.15.6856-6865.2000.Updated.

Albiach-Marti, M. R. 2012. "Molecular Virology and Pathogenicity of Citrus tristeza virus, Viral Genomes - Molecular Structure, Diversity, Gene Expression Mechanisms and Host-Virus Interactions," Prof. Maria Garcia (Ed.), ISBN: 978-953-51-0098-0, InTech, Available from: http://www.intechopen.com/books/viral-genomes- molecular-structure-diversity-gene-expression-mechanisms-and-host-virus-interactions/molecular-virology- and-pathogenicity-of-citrus-tristeza-virus.

Albiach-marti, M. R. 2013. “The Complex Genetics of Citrus Tristeza Virus.” *InTech, Current Issues in Molecular Virology - Viral Genetics and Biotechnological Applications* (Chapter 1): 1–26. http://dx.doi.org/10.5772/56122.

Al-sadi, A. M., Al-hilali, S. A., Al-yahyai, R. A., Al-said, F. A. and Deadman, M. L. 2012. “Molecular Characterization and Potential Sources of Citrus Tristeza Virus in Oman.” *Plant Pathology*, 632–640. doi:10.1111/j.1365-3059.2011.02553.x.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. 1990. “Basic Logal Alignment Search Tool.” *Journal of Molecular Biology* 215: 403–410.

Biswas, K. K., Tarafdar, A. and Sharma, S. K. 2012. “Complete Genome Sequence of Mandarin Decline Citrus Tristeza Virus of the Northeastern Himalayan Hill Region of India : Comparative Analyses Determine Recombinant.” *Archives of Virology* 157: 579–583. doi:10.1007/s00705-011-1165-y.

Chen, Y-H., Feng, Y-C., Hung, T-H. and Su, H-J. "Genomic Analysis for Different Strains of Citrus tristeza virus in Taiwan." *Unpublished.* Submitted (03-JUL-2012) Plant Pathology and Microbiology, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Da'an Dist.d., Taipei City 106, Taiwan.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. 2012. jModelTest 2: More Models, New Heuristics and Parallel Computing. *Nature Methods* 9(8), pp. 772-772.

D'Onghia, A. M., El Sayed, T., Djelouah K. and Brandonisio R. 2009. “Improved biological indexing of Citrus tristeza virus (CTV).” In: D'Onghia, A. M. (ed.), Djelouah, K. (ed.), Roistacher C. N. (ed.). “*Citrus Tristeza virus and Toxoptera citricidus: a serious threat to the Mediterranean citrus industry.* Bari CIHEAM, 2009. 143–146. (Options Méditerranéennes: Série B. Etudes et Recherches; n.65).

Domingo, E. 1992. “Genetic Variation and Quasi-Species.” *Annual Review of Genetics*: 61–63.

Domingo, E., Biebricher, C., Eigen, M. and Holland, J. J. 2001. “Quasispecies and RNA virus evolution: principles and consequences.” Landes Bioscience, Austin, Texas.

Domingo, E., Sheldon, J. and Perales, C. 2012. “Viral Quasispecies Evolution.” *Microbiology and Molecular Biology Reviews* 76(2): 159–215.

Eigen, M. 1971. "Self-organization of matter and the evolution of biological macromolecules." *Naturwissenschaften* 58, 465–523.

Eigen, M., and Schuster, P. 1977. "A principle of natural self-organization." *Naturwissenschaften* 64, 541–565.

Folimonova, S. Y., Robertson, C. J., Shilts, T., Folimonov, A. S., Hilf, M. E., Garsney, S. M. and 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(3): 1314–1325.

Folimonova, S. Y. 2013. "Developing an Understanding of Cross-Protection by Citrus Tristeza Virus." *Virology* 4 (76): 1–9. doi:10.3389/fmicb.2013.00076.

Garnsey, S. M., Civerolo, E. L., Gumpf, D. J., Yokomi, R. K. and Lee, R. F. 1991. "Development of a Worldwide Collection of Citrus Tristeza Virus Isolates." In *Eleventh ZOCV Conference*, 113–120.

Guindon, S. and Gascuel, O. 2003. "A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood." *Systematic Biology* 52 (5): 696–704. doi:10.1080/10635150390235520.

Hall, T. A. 1999. "BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT." *Nucleic Acid Symposium Series* 41: 95–98.

Hancevic, K., Cerni, S., Nolasco, G., Radic, T., Djelouah, K. and Skoric, D. 2013. "Biological Characterization of Citrus Tristeza Virus Monophyletic Isolates with Respect to p25 Gene." *Physiological and Molecular Plant Pathology* 81: 45–53. doi:10.1016/j.pmpp.2012.10.005.

Harper, S. J., Dawson, T. E. and Pearson, M. N. 2009. "Complete Genome Sequences of Two Distinct and Diverse Citrus Tristeza Virus Isolates from New Zealand." *Archives of Virology* 154: 1505–1510. doi:10.1007/s00705-009-0456-z.

Harper, S. J., Dawson, T. E. and Pearson, M. N. 2010. "Isolates of Citrus Tristeza Virus That Overcome Poncirus Trifoliata Resistance Comprise a Novel Strain." *Archives of Virology* 155: 471–480. doi:10.1007/s00705-010-0604-5.

Harper, S. J. and Hilf, M. E. "Citrus tristeza virus: Evolution of complex and varied genotypic groups." *Unpublished.* Submitted (23-JAN-2013) CREC-IFAS, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA.

Hasegawa, M., Kishino, H. and Yano, T. 1985. "Dating of the human-ape splitting by a molecular clock of mitochondrial DNA." Journal of *Molecular Evolution*, 22(2), 160-74.

Hensz, R. A. 1971. "'Star Ruby' a new deep-red-fleshed grapefruit variety with distinct tree characteristics." *Journal of the Rio Grande Valley Horticultural Society.* 25: 54–58.

Hensz, R. A. 1977. "Mutation breeding and development of the 'Star Ruby' grapefruit." *Proceedings* of the *International Society* of *Citriculture.* 2: 582-585.

Hensz, R. A. 1978. "'Star Ruby' Grapefruit a mutant of 'Ruby Red' with redder flesh and peel color." *Journal of the Rio Grande Valley Horticultural Society.* 32: 39–41.

Hensz, R. A. 1981. Bud mutations in citrus cultivars in Texas. *Proceedings of* the *International Society* of *Citriculture.* 1: 89–91.

Hensz, R. A. 1985. "'Rio Red', a New Grapefruit with a Deep-Red Color." *Journal of the Rio Grande Valley Horticultural Society.* 38: 75–76.

Hilf, M. E., and Garnsey, S. M. 2000. "Characterization and Classification of Citrus Tristeza Virus Isolates by Amplification of Multiple Molecular Markers." In *Fourteenth IOCV Conference, 2000—Citrus Tristeza Virus Characterization*, 18–27.

Hilf, M. E., Mavrodieva, V. A. and 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 (8): 909–917.

Hilf, M. E., Harper, S. J. and Dawson, W. O. "Citrus tristeza virus: strains, genotypes and isolates." *Unpublished*. Submitted (23-APR-2012) CREC-IFAS, University of Florida, 700 Experiment Stn Road, Lake Alfred, FL 33850, USA.

Hilf, M. E. and Harper, S. J. "Citrus tristeza virus: Evolution of complex and varied genotypic groups." *Unpublished.* Submitted (24-JAN-2013) CREC-IFAS, University of Florida, 700 Experiment Station Road, Lake Alfred, FL 33850, USA.

Holmes, E. C., and Moya, A. 2002. “Is the Quasispecies Concept Relevant to RNA Viruses?” *Journal of Virology* 76 (1): 460–462. doi:10.1128/JVI.76.1.460.

Karasev, A. V., Hilf, M. E., Garnsey, S. M. and Dawson, W. O. 1997. “Transcriptional Strategy of Closteroviruses : Mapping the 5J Termini of the Citrus Tristeza Virus Subgenomic RNAs †.” *Journal of Virology* 71 (8): 6233–6236.

Katoh, K., Misawa, K., Kuma, K-I. and Miyata, T. 2002. “MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform.” *Nucleic Acids Research* 30 (14): 3059–3066.

Katoh, K., Kuma, K-I., Toh, H. and Miyata, T. 2005. “MAFFT Version 5: Improvement in Accuracy of Multiple Sequence Alignment.” *Nucleic Acids Research* 33 (2): 511–518. doi:10.1093/nar/gki198.

Korkmaz, S., Cevik, B., Onder, S., Koc, K. and Bozan, O. 2008. “Detection of Citrus tristeza virus (CTV) from Satsuma mandarin (Citrus unshiu) by direct tissues blot immunoassay (DTBIA), DAS-ELISA and biological indexing.” *New Zealand Journal for Crop. Hort.* 36:239-246.

Lauring, A. S. and Andino, R. 2010. “Quasispecies Theory and the Behavior of RNA Viruses.” *PLoS Pathogens* 6 (7): 1–8. doi:10.1371/journal.ppat.1001005.

Lbida, B., Bennani, A., Serrhini, M. N. and Zemzami, M. 2005. “Biological, serological and molecular characterization of three isolates of Citrus tristeza closterovirus introduced into Morocco.” *EPPO Bulletin* 35:511- 517.

Liu, S., Vijayendran, D. and Bonning, B. C. 2011. “Next Generation Sequencing Technologies for Insect Virus Discovery.” *Viruses* 3: 1849–1869. doi:10.3390/v3101849.

Luttig, M., van Vuuren, S. P. and 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-stranded conformation polymorphism. In: Proc. 15th Conf. IOCV, 186- 196. IOCV, Riverside, CA.

Marais, L. J. 1994. Citrus tristeza virus and its effect on the southern Africa citrus industry. Citrus Ind. 75(6): 58-60.

Melzer, M. J., Borth, W. B., Sether, D. M., Ferreira, S., Gonsalves, D. and Hu, J. S. 2010. “Genetic Diversity and Evidence for Recent Modular Recombination in Hawaiian Citrus Tristeza Virus.” *Virus Genes* 40: 111–118. doi:10.1007/s11262-009-0409-3.

Moreno, P., Ambrós, S., Albiach-martí, M. R., Guerri, J. and Peña, L. 2008. “Review: Plant Diseases That Changed the World. Citrus Tristeza Virus: A Pathogen That Changed the Course of the Citrus Industry.” *Molecular Plant Pathology* 9 (2): 251–268. doi:10.1111/J.1364-3703.2007.00455.X.

Müller, G. W., Costa, A. S., Kitajima, E. W. and Camargo, I. J. B. 1974. "Additional evidence that tristeza multiplies in Passiflora spp." In: Proc. 6th Conf. IOCV, 75-78. IOCV, Riverside, CA.

Nowak, M. A. 1992. “What Is a Quasispecies?” *Trens in Ecology and Evolution, Elsevier Trends Journals*.

Orr, H. A. 2009. “Fitness and its role in evolutionary genetics.” *Nature Reviews Genetics* 10, 531 – 539.

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. 1994. "Nucleotide sequence and organization of eight 3' open reading frames of the citrus tristeza closterovirus genome." *Virology.* 199 (1), 35-46.

Quiroz, J. D. C., Mendoza, A., Cruz, M. A., Fernandez, S. and Pena, A. "The Complete RNA genome sequence of Citrus Tristeza Virus from Mexico." *Unpublished.* Submitted (01-NOV-2005) Centro de Biotecnologia Genomica, Instituto Politecnico Nacional, Blvd. del Maestro SN, Reynosa, Tamaulipas 88710, Mexico.

Rambaut, A. 2009. FigTree: Tree Figure Drawng Tool. University of Edinburgh: Institute of Evolutionary Biology.

Rambaut, A. and Drummond, A. 2009. Tracer: MCMC Trace Analysis Tool. University of Edinburgh: Institute of Evolutionary Biology and University of Auckland: Department of Computer Science.

Read, D. A. and Pietersen, G. 2015. “Genotypic Diversity of Citrus Tristeza Virus within Red Grapefruit , in a Field Trial Site in South Africa.” *European Journal of Plant Pathology*: 1–15. doi:10.1007/s10658-015-0631-x.

Rodríguez, F., Oliver, J. L., Marín, A.and Medina, J. R. 1990. “The general stochastic model of nucleotide substitution.” Journal of *Theoretical Biology*, 142(4), 485-501.

Roistacher, C. N. 1991. "Graft-Transmissable Diseases of Citrus." *Handbook for Detection and Diagnosis.* Rome: FAO.

Roistacher, C. N., da Graça, J. V. and Müller, G. W. 2010. “Cross Protection Against Citrus Tristeza Virus - a Review.” In *Proceedings, 17th Conference, 2010 – Citrus Tristeza Virus*, 1–27.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., Larget, B., Lui, L., Suchard, M. A. and Huelsenbeck, J. P. 2012. “MrBayes 3 . 2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space.” *Systematic Biology* 61 (3): 539–542. doi:10.1093/sysbio/sys029.

Roy, A. and Ramachandran, P. 2002. “Bidirectional PCR – a tool for identifying citrus tristeza virus isolates.” *Indian Phytopath*ology 55: 182–186.

Roy, A. and Brlansky, R. H. 2004. “Genotype classification and molecular evidence for the presence of mixed infections in Indian Citrus tristeza virus isolates.” *Archives of Virology* 149: 1911–1929.

Roy, A., Manjunath, K. L., and Brlansky, R. H. 2005. "Assessment of sequence diversity in the 5 -terminal region of Citrus tristeza virus from India." *Virus Research 113*, 132 – 142. doi:10.1016/j.virusres.2005.04.023

Roy, A. and Brlansky, 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(11): 1297–1306.

Roy, A., Ananthakrishnan, G., Hartung, J. S. and Brlansky, R. H. 2010. “Development and Application of a Multiplex Reverse-Transcription Polymerase Chain Reaction Assay for Screening a Global Collection of Citrus Tristeza Virus Isolates.” *Phytopathology* 100 (10): 1077–1088.

Roy, A., Choudhary, N., Hartung, J. S. and Brlansky, R. H. 2013. “The Prevalence of the Citrus Tristeza Virus Trifoliate Resistance Breaking Genotype Among Puerto Rican Isolates.” *Plant Disease* 97: 1227–1234.

Roy, A., Choudhary, N., Hartung, J. H. and Brlansky, R. H. "Trifoliate resistance breaking (RB) Citrus tristeza virus genotype specific RT-PCR and complete genome sequence analysis of an unassigned isolate discovered the presence of RB genotype from Puerto Rico." *Unpublished.* Submitted (12-MAY-2011) Plant Pathology, Citrus Research and Education Center, University of Florida, 700, Experiment Station Road, Lake Alfred, FL 33850, USA.

Ruiz-Ruiz, S., Moreno, P., Guerri, J. and Ambros, S. 2006. "The complete nucleotide sequence of a severe stem pitting isolate of Citrus tristeza virus from Spain: comparison with isolates from different origins." Archives of Virology. 151 (2), 387-398.

Sambrook, J., and Russell, D. W. 2001. "*Molecular Cloning: "A Laboratory Manual"."* (3rd Ed.). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. http://trove.nla.gov.au/version/45739910

Satyanarayana, T., Gowda, S., Boyko, V. P., Albiach-Martı, M. R., Mawassi, M., Navas-castillo, J., Karasev, A. V., Dolja, V., Hilf, M. E., Lewandowski, D. J., Moreno, P., Bar-Joseph, M., Garsney, S. M. and Dawson, W. O. 1999. "An Engineered Closterovirus RNA Replicon and Analysis of Heterologous Terminal Sequences for Replication." *Proceedings of the National Acadamy of Science USA* 96: 7433–7438.

Scott, K. A., Hlela, Q., Zablocki, O., Read, D., van Vuuren, S. and 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* (Springer-Verlag): 1–22. doi:10.1007/s00705-012-1450-4.

Sindhvajiva, K., Paradornuwat, A., Chowpongpang, S. and Vichukit, V. "Genome of Citrus Tristeza Virus Isolate A18 from Mandarin Citrus in Thailand." *Unpublished*. Submitted (19-MAR-2012) Department of Plant Pathology, Faculty of Agriculture, Kasetsart University, 50 Ngam Wong Wan Rd., Ladyaow, Chatuchak, Bangkok 10900, Thailand.

Smocilac, S. 2012. "Survey of CTV in Southern African Grapefruit Orchards." *Unpublished* University of Pretoria, Department of Microbiology and Plant Pathology. 1–113.

Stewart, K. A. 2006. "Strain Differentiation of Citrus Tristeza Virus Isolates from South Africa by PCR and Microarray." *Africa*.

Suastika,G., Natsuaki,T., Kano,T., Ieki,H. and Okuda,S. "Nucleotide sequence of citrus tristeza virus seedling yellows strain." *Unpublished.* Submitted (19-JUL-2000) Tomohide Natsuaki, Utsunomiya University, Faculty of Agriculture; Mine-machi 350, Utsunomiya, Tochigi 321-8505, JAPAN (E-mail:natsuaki@cc.utsunomiya-u.ac.jp).

Tamura, K. and Nei, M. 1993. “Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.” *Molecular Biology* and *Evolution,* 10(3), 512-26.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. “MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood , Evolutionary Distance , and Maximum Parsimony Methods.” *Molecular Biology and Evolution* 28 (10): 2731–2739. doi:10.1093/molbev/msr121.

Targon, M. L. P. N., Machado, M. A., Carvalho, S. A., Souza, A. A. and Müller, G. W. 2000. “Differential Replication of a Mild and a Severe Citrus Tristeza Virus Isolate in Species and Varieties of Citrus.” In *Fourteenth IOCV Conference, 2000—Citrus Tristeza Virus*, 127–130.

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 *Fourteenth IOCV Conference*, 103–110.

Varveri, C., Olmos, A. and Cambra, M. "Full genome of CTV L192GR." *Unpublished.* Submitted (05-DEC-2012) Phytopathology, Benaki Phytopathological Institute, 8, St.Delta str., Kifissia, Attika 14561, Greece.

Vives, M. C., Rubio, L., Lopez, C., Navas-castillo, J., Albiach-martí, M. R., Dawson, W. O., Guerri, J., Flores, R. and Moreno, P. 1999. “The Complete Genome Sequence of the Major Component of a Mild Citrus Tristeza Virus Isolate.” *Journal of General Virology* 80: 811–816.

von Broembsen, L. A., da Graça, J. V., Lee, A. T. C. and Waller G. G. 1978. “South Africa’s Citrus Improvement Programme after five years.” *Citrus and Subtropical Fruit Journal* 532, 11 – 18.

von Broemsen, L. and Lee, A. T. C. 1988. “South Africa’s citrus improvement programme.” In: Timmer LW, Garnsey SM, Navarro L (eds) Proc 10th Conf Int Organization Citrus Virol, IOCV, Riverside, pp 407 – 416.

Waibel, C. 1953. “Varieties and strains of citrus originating in the Lower Rio Grande Valley of Texas.” Proc. *Rio Grande Valley Horticultural Inst.* 7: 18–24.

Wang, Y. J., Zhou, C. Y., Li, Z. A., Zhou, Y., Cao, M. J., Li, L. D., Yang, F. Y., Tang, K. Z. and Liu, J. X. "The complete nucleotide sequence analysis of two citrus tristeza virus isolates from China." *Unpublished.* Submitted (09-APR-2012) National Citrus Virus-free Seedling Center, Citrus Research Institute, Chinese Academy of Agricultural Sciences, Citrus Village NO15, Xiema, Chongqing 400712, China.

Wang, J., Bozan, O., Kwon, S-J., Dang, T., Rucker, T., Yokomi, R. K., Lee, R. F., Folimonova, S. Y., Krueger, R. R., Bash, J., Greer, G., Diaz, J., Serna, R. and Vidalakis, G. 2013. “Past and future of a century old Citrus tristeza virus collection : a California citrus germplasm tale.” *Virology* 4(366): 1–13.

Weng, Z., Barthelson, R., Gowda, S., Hilf, M. E., Dawson, W. O., Galbraith, D. W. and Xiong, Z. 2007. “Persistent Infection and Promiscuous Recombination of Multiple Genotypes of an RNA Virus within a Single Host Generate Extensive Diversity.” *PLoS ONE* 2(9): e917. doi:10.1371/journal.pone.0000917. doi:10.1371/journal.pone.0000917.

Wu, G. W., Hong, N. and Wang, G. P. "Complete genome sequence analysis of an aphid-transmissible isolate from China." *Unpublished*. Submitted (21-NOV-2011) Huazhong Agricultural University, College of Plant Science & Technology, Shizishan Street, Hongshan District, Wuhan, Hubei 430070, China.

Yang, Z-N., Mathews, D. M., Dodds, J. A. and Mirkov, T. E. 1999. “Molecular Characterization of an Isolate of Citrus Tristeza Virus That Causes Severe Symptoms in Sweet Orange.” *Virus Genes* 19 (2): 131–142.

Yokomi, R. K., Saponari, M., Metheney, P. and Vidalakis, G. 2011. “Genetic Differentiation and Biology of Citrus Tristeza Virus Populations Spreading in California.” In *18th IOCV Conference*, 1–8.

Zablocki, O. and Pietersen, G. 2014. “Characterization of a Novel Citrus Tristeza Virus Genotype within Three Cross-Protecting Source GFMS12 Sub-Isolates in South Africa by Means of Illumina Sequencing.” *Archives of Virology*: 107. doi:10.1007/s00705-014-2041-3.

Zhou, Y., Zhou, C.Y., Song, Z., Liu, K. H. and Yang, F. Y. 2007. “Characterization of Citrus tristeza virus isolates by indicators and molecular biology methods.” *Agric. Sci. China.* 6:573-579.

3.7. Website references.

http://www.ncbi.nlm.nih.gov/

http://mafft.cbrc.jp/alignment/server/

http://www.molecularevolution.org/software/phylogenetics/mrbayes

http://hydrodictyon.eeb.uconn.edu/eebedia/index.php/Phylogenetics:_MrBayes_Lab

http://genome.nci.nih.gov/tools/reformat.html

Chapter 4

Concluding remark.

4.1. Concluding remark.

Citrus tristeza virus (CTV) is a major viral pathogen to citrus worldwide. The quasispecies nature of CTV is one of the major reasons for the complexity of this viral pathogen worldwide, and is linked with its error-prone replication as well as continuous recombination reactions between the different viral strains within mixed infections in the field, which aid in producing multiple closely related genomic variants. This increased genetic variability between strains, makes it extremely difficult to initiate a successful CTV cross-protection scheme. In the current study our main objective was to obtain homogenous CTV genotypic sources to aid in the protection of citrus cultivars in Southern Africa.

In Chapter 2 we focused on isolating potential homogenous genotypic source using bark-flap, stem-slash inoculation, as well as SAT's. Both the bark-flap and stem-slash procedures were followed as published by Robertson *et al.* (2005) and Satyanarayana *et al.* (2001), with minor modifications. The SAT trail performed was a shotgun approach, in which the collected aphids were directly transferred to healthy Mexican Limes, without knowing the CTV genotypic presence within the source. All three procedures yielded some transmission of CTV. Even though low success rates were obtained, further optimisation of each approach may possibly lead to a larger numbers of CTV isolations.

Since all six SAT positive seedlings contained a mixed CTV infection, future work in decreasing the genotypic heterogeneity within inoculated seedlings can be done. For example, inoculating plant material such as budwood or plant sap, of each source onto indicator hosts can be the first approach in reducing the CTV genotypic presence, as these hosts might select stronger for specific genotypes. To aid in reducing the genotypic presence, different options prior and after the inoculation step can be included; for example, serially diluting the plant sap prior to the inoculation step and keeping the inoculated indicator hosts at different incubation temperatures. Single aphid transmissions can also be serially repeated, by allowing a clean and maintained aphid colony to feed on the infected seedlings and then transfer them to indicator hosts in two passages at least. During both these processes, one can also determine the strain severity of the specific genotype present. Symptom evaluation on these indicator hosts by means of biological indexing can provide an indication of

the strain severity, which can therefore aid in determining what experimental approach may follow. However, the inoculation procedure may need to be repeated as too many genotypes might still be present. One can also attempt the strain attenuation procedure performed by Müller *et al.* (1974) in which not only strain isolations can be done, but also strain severity can be determined. The detailed basis of this approach is explained in section *1.2.4. CTV trainsmission*.

Within Chapter 3, a T36 genotype focussed screening was performed due to confirm previous findings by Smocilac (2012) and establish the T36 source within our greenhouse. Findings showed that the previous T36 detection was a misdiagnosis as no T36 was detected in the grafted Mexican Lime trees of the 2012 survey. Furthermore, T36 was also not detected in the plant material collected from the same site as well as the newly graft-inoculated Mexican Lime seedlings. However, we determined that the RB genotype was the predominant genotypic source present, and confirmed the presence of a homogenous RB source (NZRB-TH30 (FJ525434)) within Mexican Lime tree 12-7006. In addition, the Illumina NGS results obtained for this RB source also demonstrated the viral quasispecies nature of this population as RB variants NZRB-TH28 (FJ525433), NZRB-M12 (FJ525431), Taiwan-Pum/SP/T1 (JX266712) and NZRB-G90 (FJ525432) provided reads with mapping results close to that of NZRB-TH30. As consensus lengths similar to NZRB-TH30 were obtained for these RB variants but with very low read counts, we can assume that reads with sequences related to these variants were present in extremely low concentrations. Thus, having confirmed the presence of the homogenous RB source NZRB-TH30, the severity of the strain can be determined by using the same approach discussed above for Chapter 2, only using plant material from tree 12-7006.

In conclusion, the viral isolation procedures used within this study can act as the first step detecting CTV in future studies, followed by more in-depth molecular diagnostic approaches to differentiate and characterise the different CTV strains available. Molecular detection by means of CTV genotype specific RT-PCR analysis, phylogenetic analysis as well as Illumina Next-generation high throughput sequencing are useful procedures to use to obtain accurate results. The genotype specific RT-PCR analysis performed within this study provided accurate and reliable results with RT-PCR and Illumina NGS results yields more data regarding the composition of virus populations Future work however is required to establish

thresholds with regards reads counts in order to confidently state the presence or not of a given genotype.

Phylogenetic analysis to confirm all PCR results as well as to determine whether potential new strains or genotypes, where shown during this study to be very useful additional CTV genotype or strain detection methods as well as confirmation procedures.

Results from this study provide useful information for the South African citrus industry for the expansion of the mild strain cross-protection scheme.

4.2. References.

Müller, G. W., Costa, A. S., Kitajima, E. W. and Camargo, I. J. B. 1974. "Additional evidence that tristeza multiplies in Passiflora spp." In: Proc. 6th Conf. IOCV, 75-78. IOCV, Riverside, CA.

Robertson, C. J., Garsney, S. M., Satyanarayana, T., Folimonova, S. and Dawson, W. O. 2005, "*Efficient Infection of Citrus Plants with Different Cloned Constructs of Citrus tristeza virus Amplified in Nicotiana benthamiana Protoplasts.*" Sixteenth IOCV Conference, 2005 – Citrus Tristeza Virus. Department of Plant Pathology, Citrus Research and Education Center, University of Florida, Lake Alfred, Florida 33850, USA, pp 187-194.

Satyanarayana, T., Bar-Joseph, M., Mawassi, M., Albiach-Martı, M. R., Ayllon, M. A., Gowda, S., Hilf, M. E., Moreno, P., Garnsey, S. M. and Dawson, W. O. 2001. "Amplification of Citrus Tristeza Virus from a cDNA Clone and Infection of Citrus Trees." *Virology*, *280*, 87–96. doi:10.1006/viro.2000.0759.

Smocilac, S. 2012. "Survey of CTV in Southern African Grapefruit Orchards." *Unpublished* University of Pretoria, Department of Microbiology and Plant Pathology. 1–113.

Appendix 1

Citrus tristeza virus reference genomes.

Table 22: List of CTV reference genomes used during phylogenetic analysis.

Name	Acc. Number	Description
AB046398	AB046398.1	CTV genomic RNA, complete genome, seedling yellows strain
AF001623	AF001623.1	CTAF001623 CTV, complete genome
AF260651	AF260651.1	CTV T30, complete genome
AY170468	AY170468.1	CTV, complete genome
AY340974	AY340974.1	CTV isolate Qaha from Egypt, complete genome
DQ151548	DQ151548.1	CTV strain T318A, complete genome
DQ272579	DQ272579.1	CTV from Mexico, complete genome
EU076703	EU076703.3	CTV isolate B165, complete genome
EU857538	EU857538.1	CTV strain SP, complete genome
EU937519	EU937519.1	CTV strain VT, complete genome
EU937520	EU937520.1	CTV strain T30, complete genome
EU937521	EU937521.1	CTV strain T36, complete genome
FJ525431	FJ525431.1	CTV isolate NZRB-M12, complete genome
FJ525432	FJ525432.1	CTV isolate NZRB-G90, complete genome
FJ525433	FJ525433.1	CTV isolate NZRB-TH28, complete genome
FJ525434	FJ525434.1	CTV isolate NZRB-TH30, complete genome
FJ525435	FJ525435.1	CTV isolate NZRB-M17, complete genome
FJ525436	FJ525436.1	CTV isolate NZ-B18, complete genome
GQ454869	GQ454869.1	CTV strain HA18-9, complete genome
GQ454870	GQ454870.1	CTV strain HA16-5, complete genome
HM573451	HM573451.1	CTV isolate Kpg 3, complete genome
JF957196	JF957196.1	CTV isolate B301, complete genome
JQ061137	JQ061137.1	CTV isolate AT-1, complete genome
JQ798289	JQ798289.1	CTV isolate A18, complete genome
JQ911663	JQ911663.1	CTV isolate CT14A, complete genome
JQ911664	JQ911664.1	CTV isolate CT11A, complete genome
JQ965169	JQ965169.1	CTV isolate T68-1, complete genome
JX266712	JX266712.1	CTV isolate Taiwan-Pum/SP/T1, complete genome
JX266713	JX266713.1	CTV isolate Taiwan-Pum/M/T5, complete genome
KC262793	KC262793.1	CTV isolate L192GR, complete genome
KC333868	KC333868.1	CTV isolate 12-8 (CT-ZA3), complete genome
KC517485	KC517485.1	CTV isolate FS674-T36, complete genome
KC517486	KC517486.1	CTV isolate FS701-T36, complete genome
KC517487	KC517487.1	CTV isolate FS703-T36, complete genome
KC517488	KC517488.1	CTV isolate FS577, complete genome
KC517489	KC517489.1	CTV isolate FS701-T30, complete genome
KC517490	KC517490.1	CTV isolate FL278-T30, complete genome
KC517491	KC517491.1	CTV isolate FS703-T30, complete genome
KC517492	KC517492.1	CTV isolate FS703-VT, complete genome
KC517493	KC517493.1	CTV isolate FL202-VT, complete genome
KC517494	KC517494.1	CTV isolate FS701-VT, complete genome
KC525952	KC525952.1	CTV isolate T3, complete genome
NC001661	NC001661.1	CTV, complete genome
U16304	U16304.1	CTU16304 CTV complete genome
Y18420	Y18420.1	CTV complete genome, isolate T385

Appendix 2

Bayesian phylogenetic trees produced for the CTV analysis of the bark-flap and stem-slash inoculated Mexican Lime trees.

Prior to the inoculation trials [Figure 37 – 43].

After the inoculation trails [Figure 44 – 47].

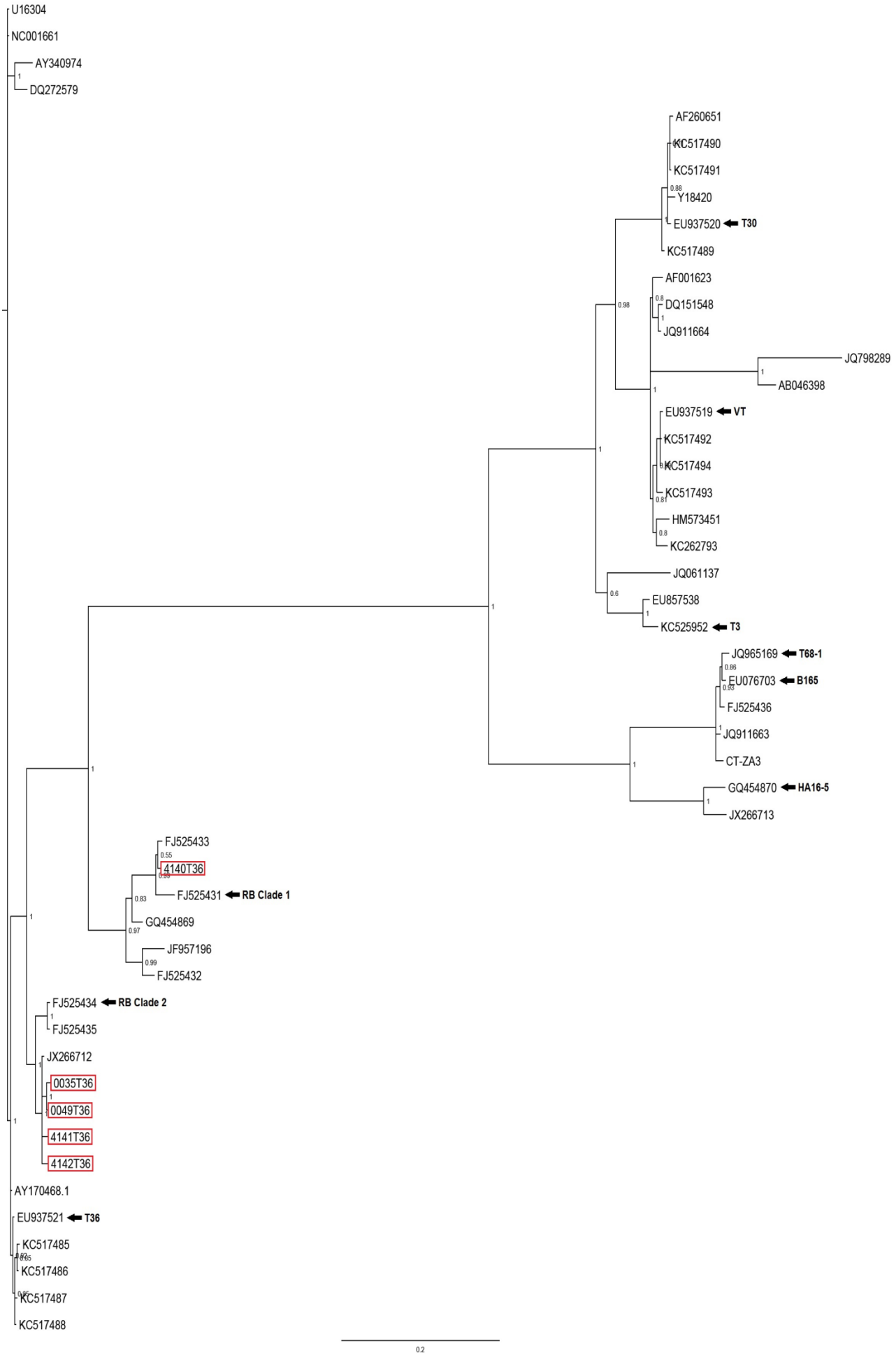


Figure 37: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T36 RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

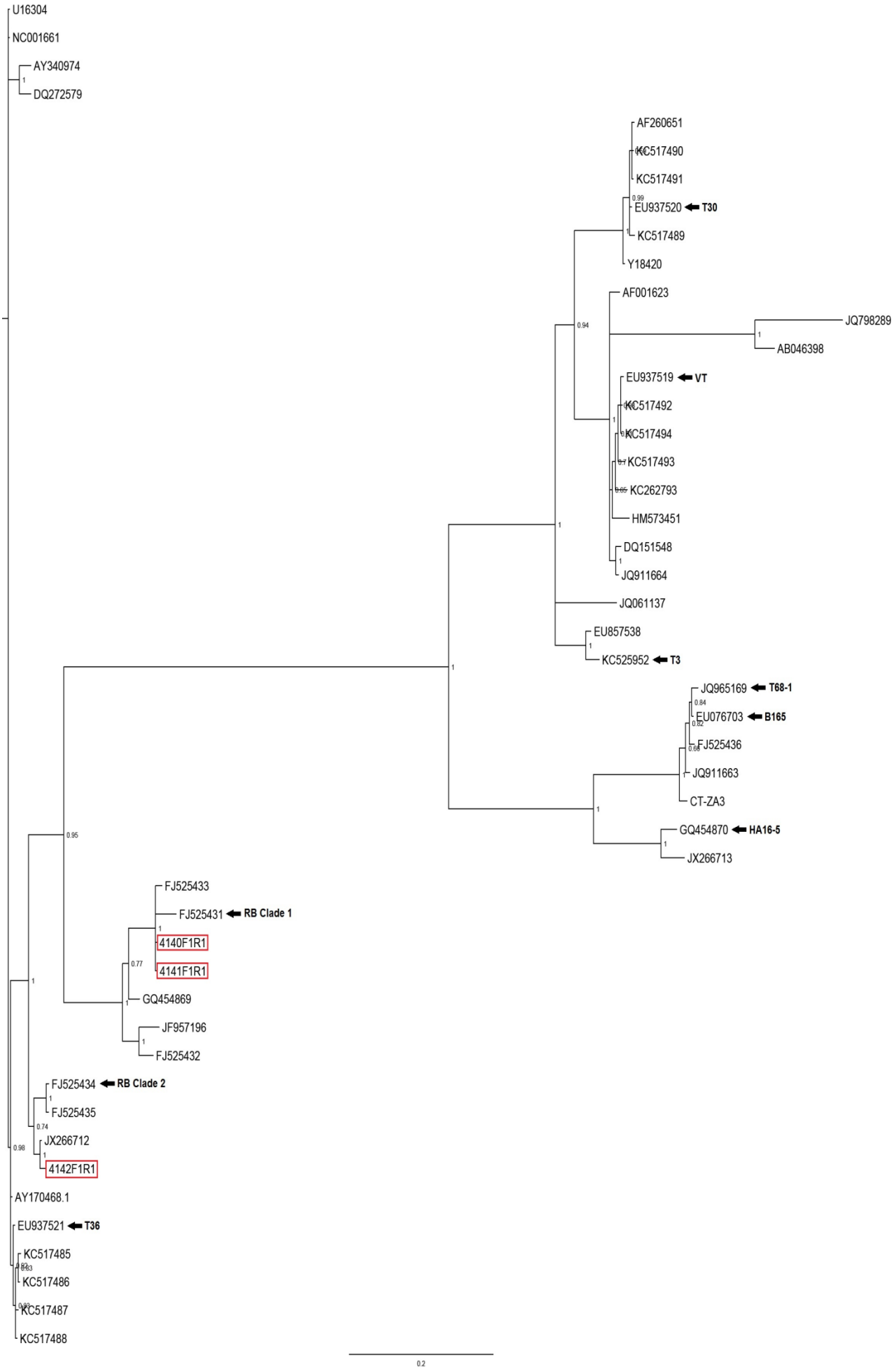


Figure 38: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F1R1 RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes (red rectangle). The confidence levels are shown as posterior probability values at each node.

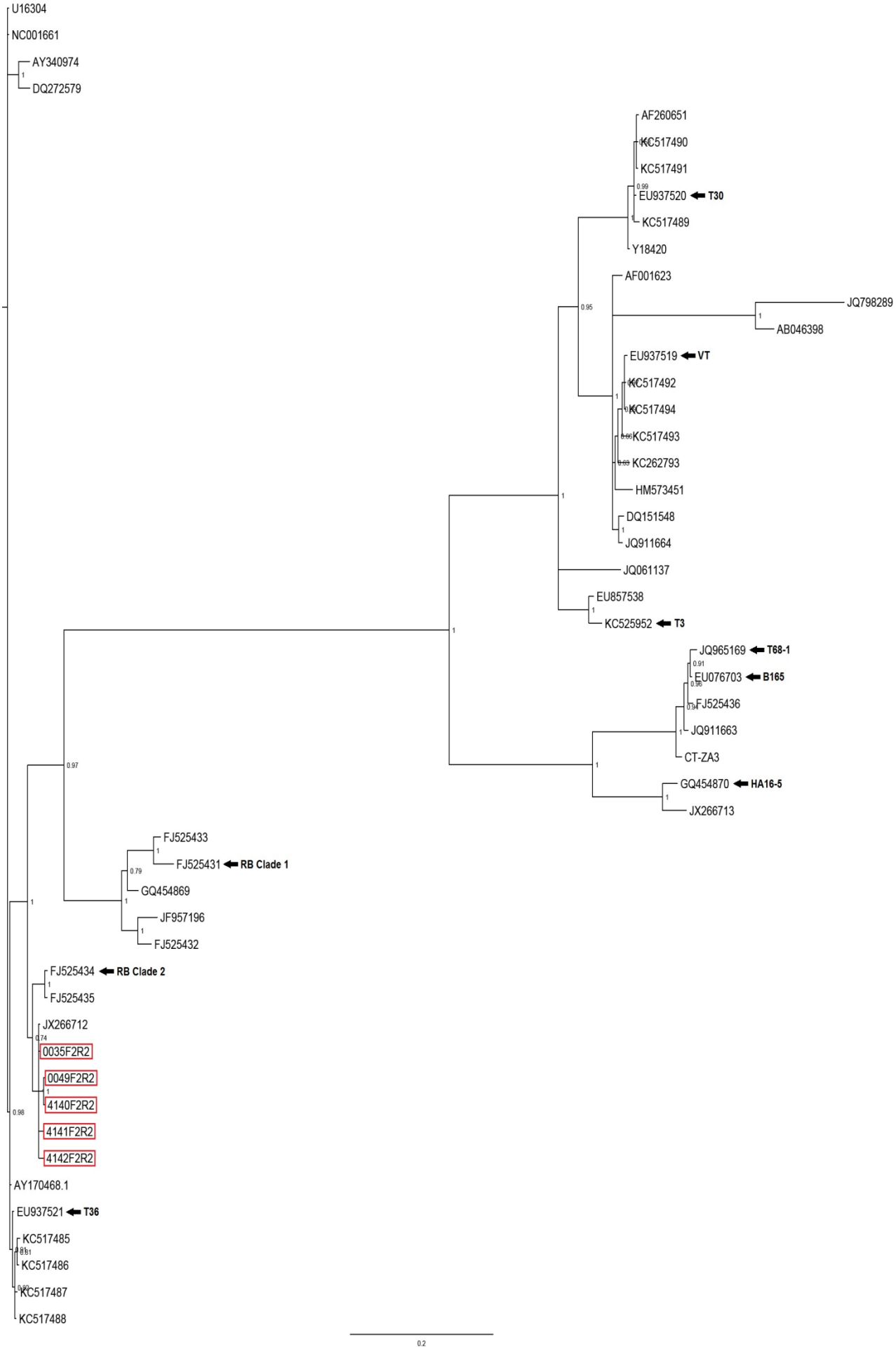


Figure 39: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F2R2 RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

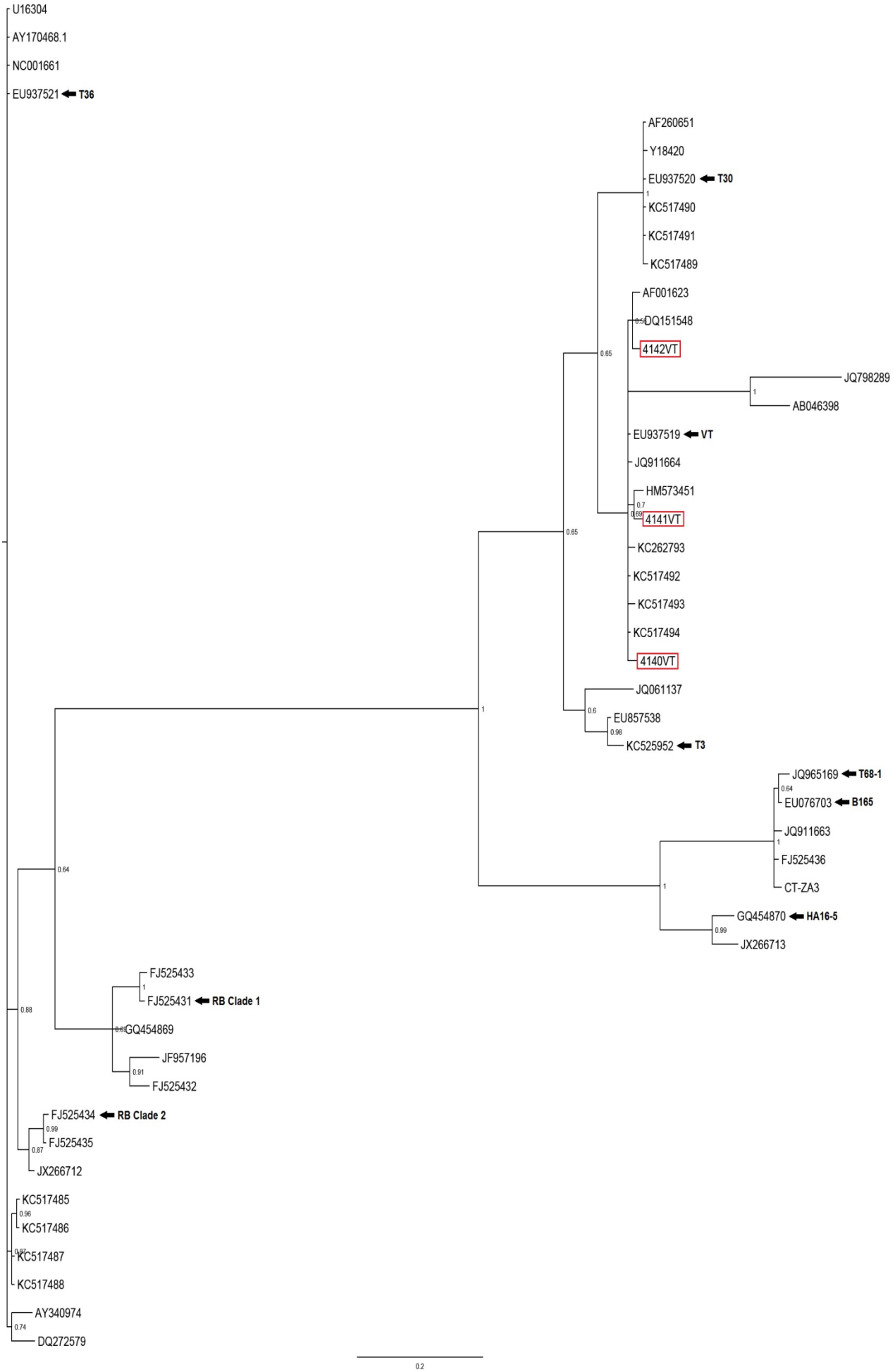


Figure 40: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the VT RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

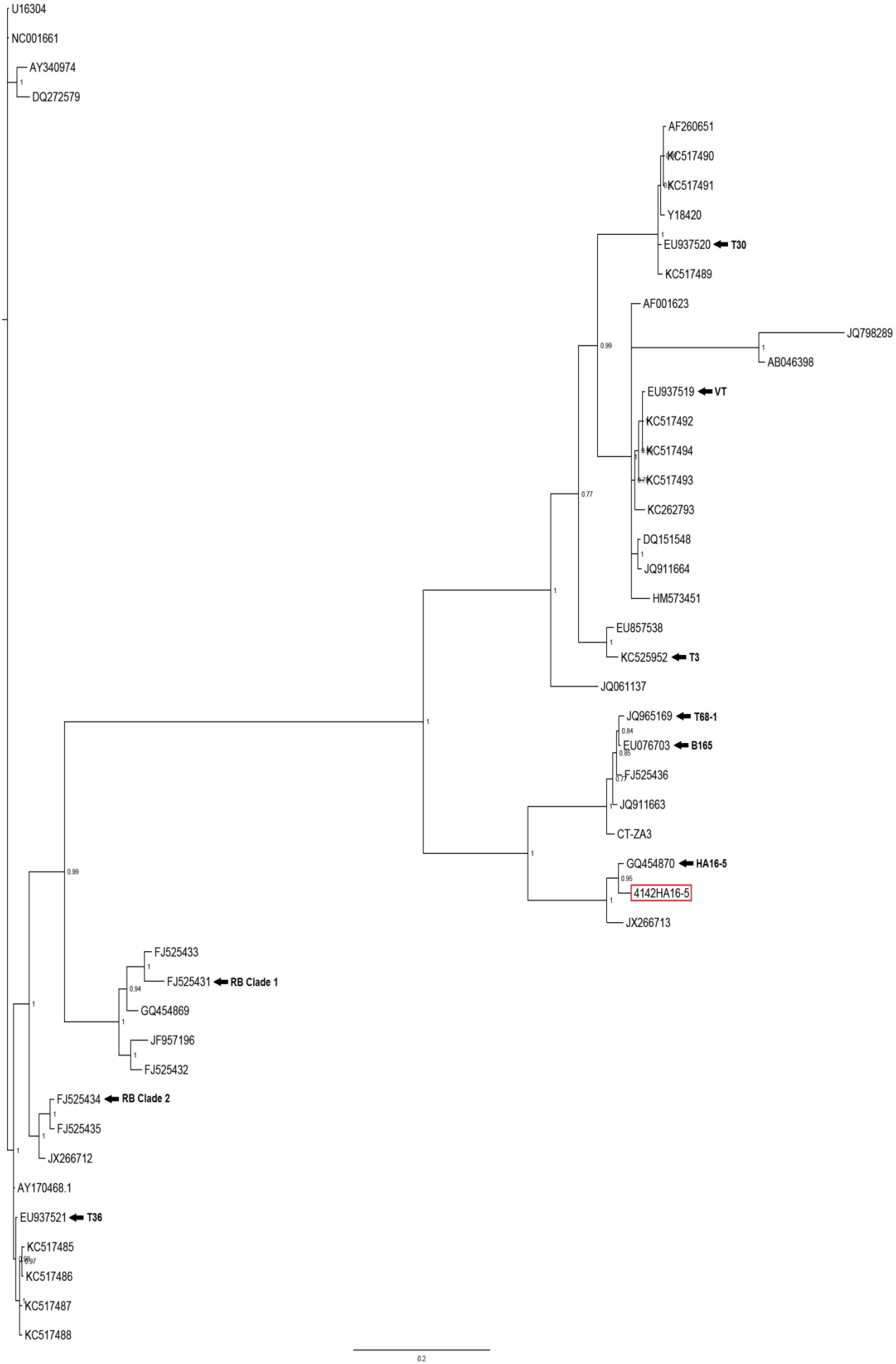


Figure 41: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the HA16-5 RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

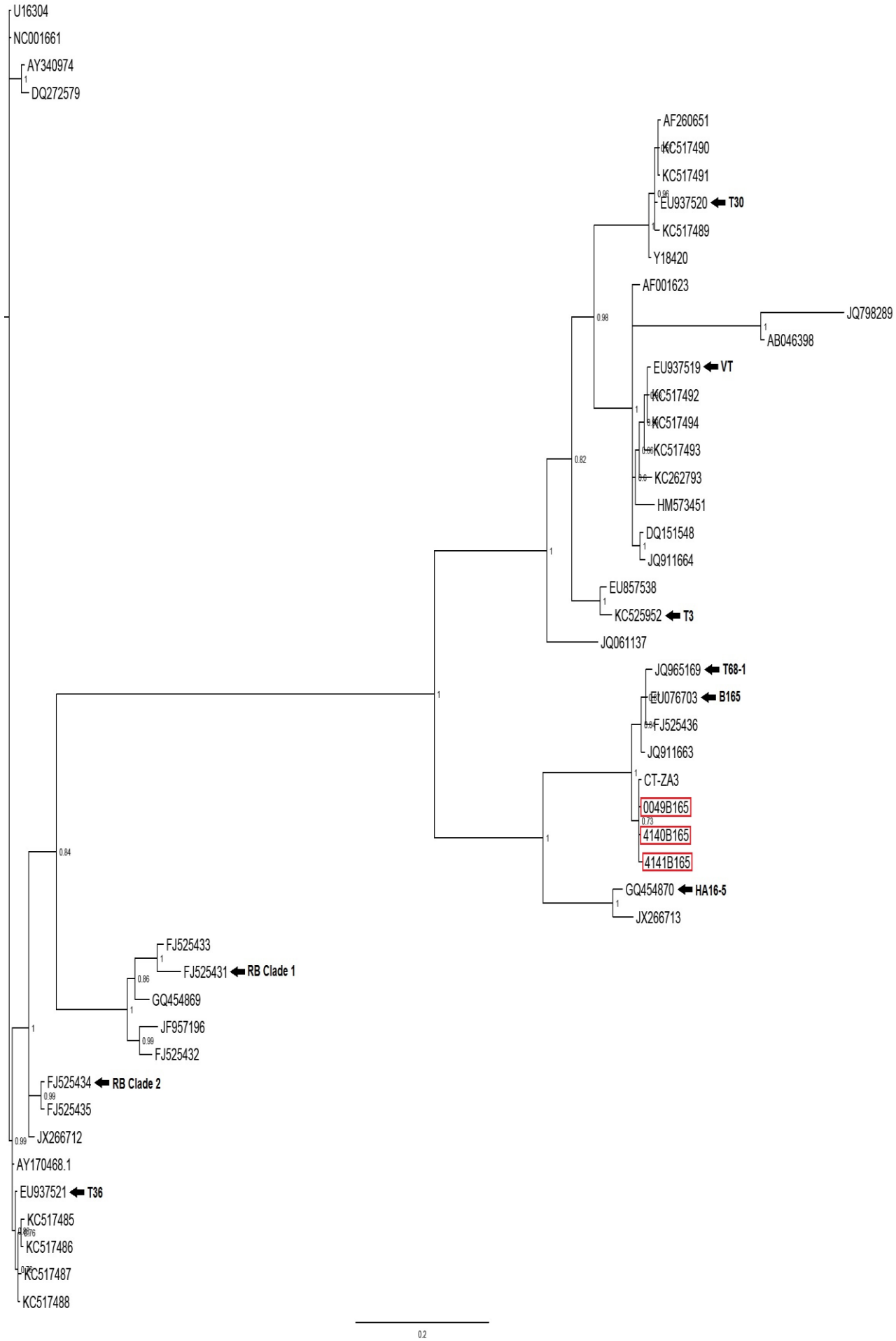


Figure 42: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the B165 RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

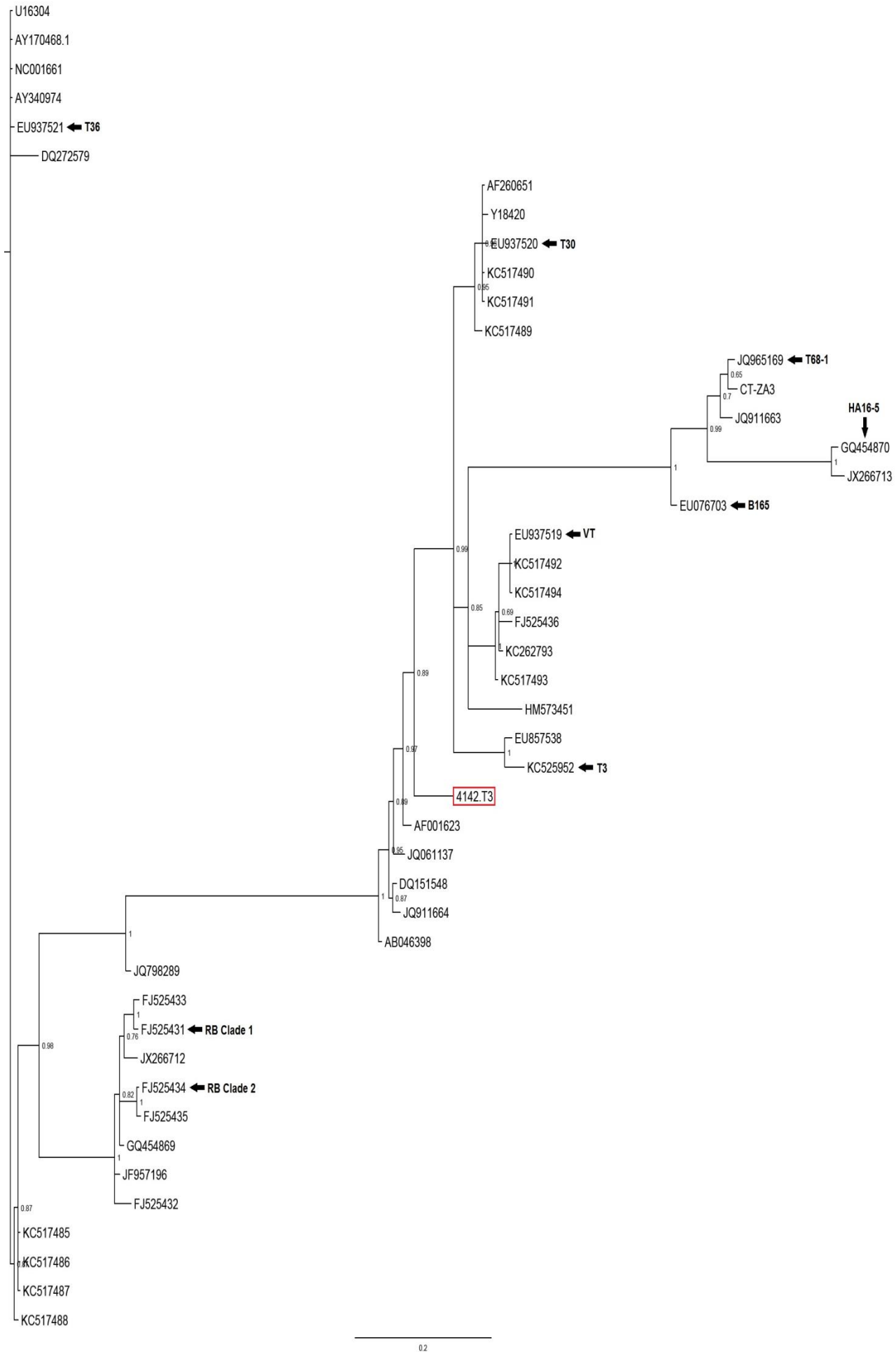


Figure 43: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T3 RT-PCR of citrus trees sampled from, prior to the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

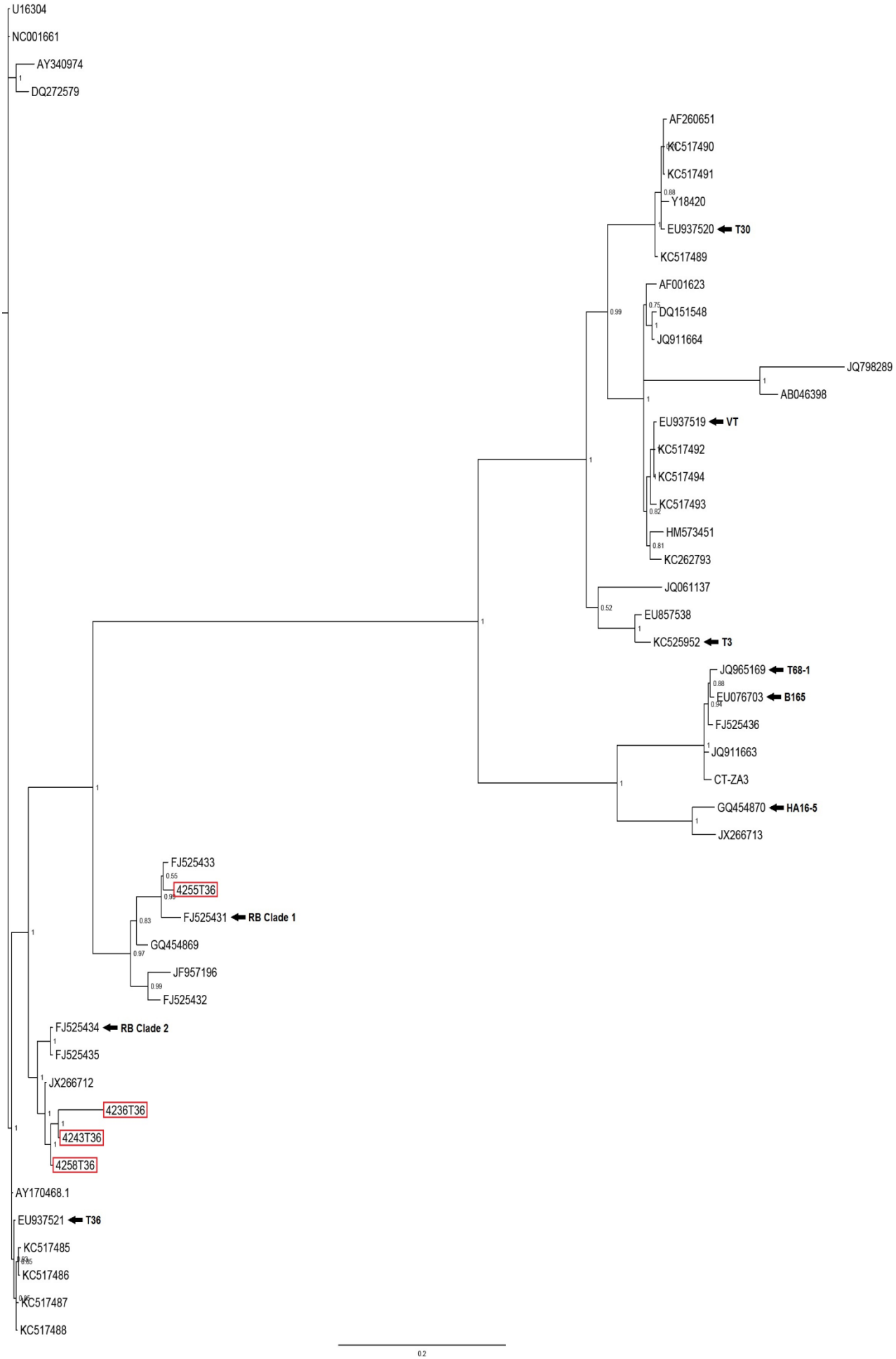


Figure 44: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T36 RT-PCR of the positive Mexican Lime seedlings, after the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

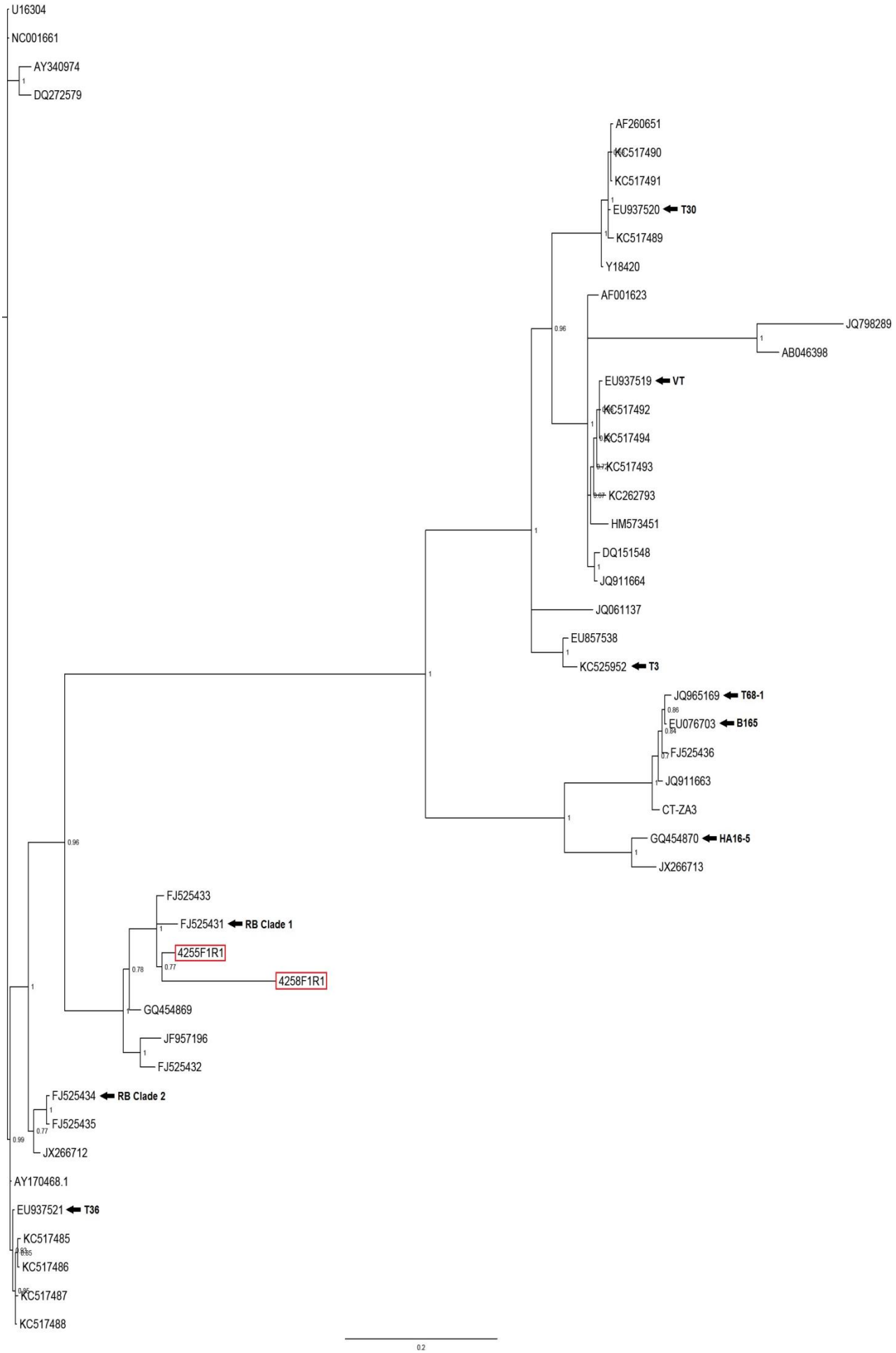


Figure 45: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F1R1 RT-PCR of the positive Mexican Lime seedlings, after the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

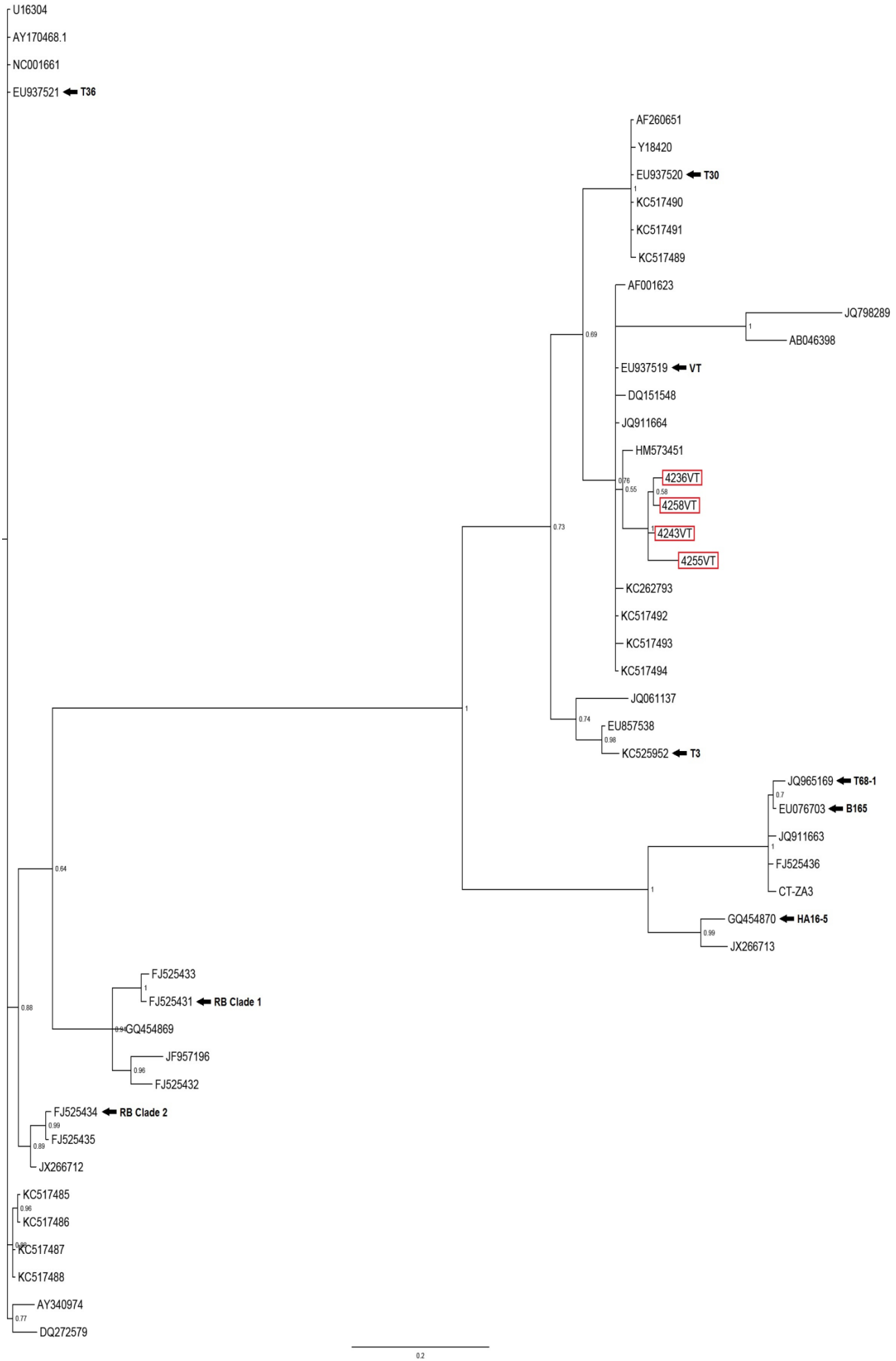


Figure 46: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the VT RT-PCR of the positive Mexican Lime seedlings, after the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

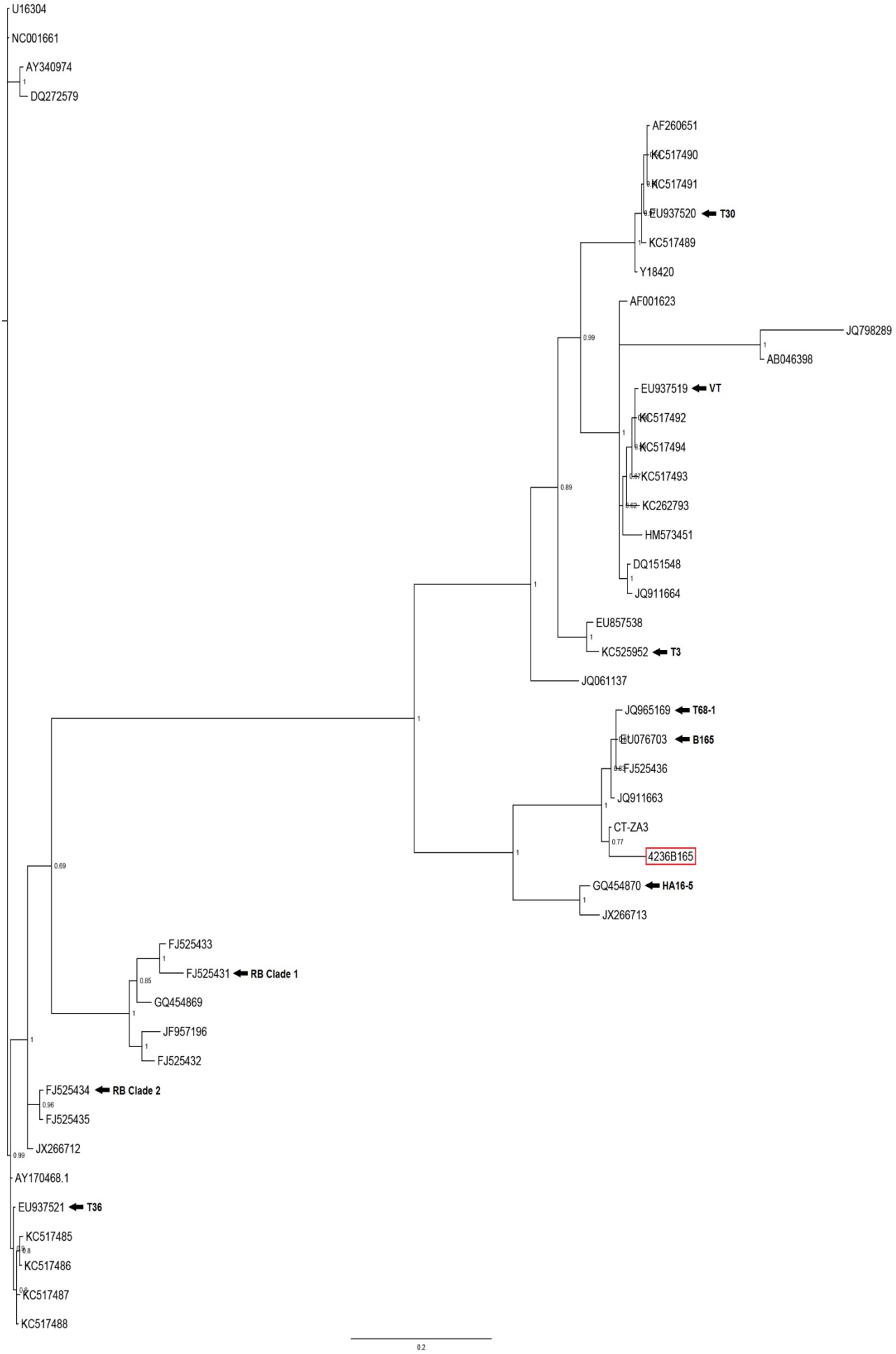


Figure 47: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the B165 RT-PCR of the positive Mexican Lime seedlings, after the CTV inoculation trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

Appendix 3

Bayesian phylogenetic trees produced for the analysis of the CTV genotype status of samples collected for use in the Single Aphid Transmission trail.

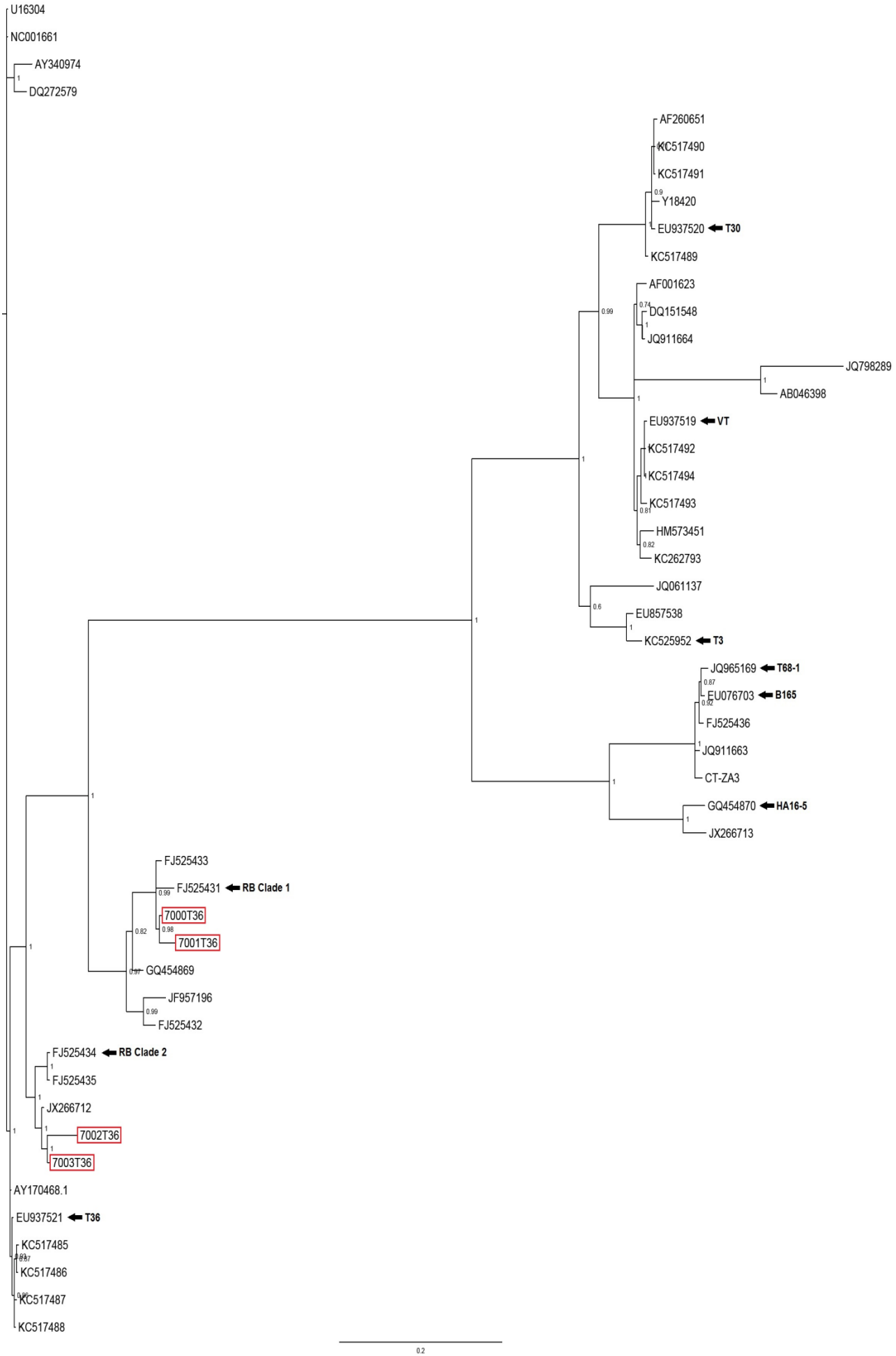


Figure 48: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T36 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

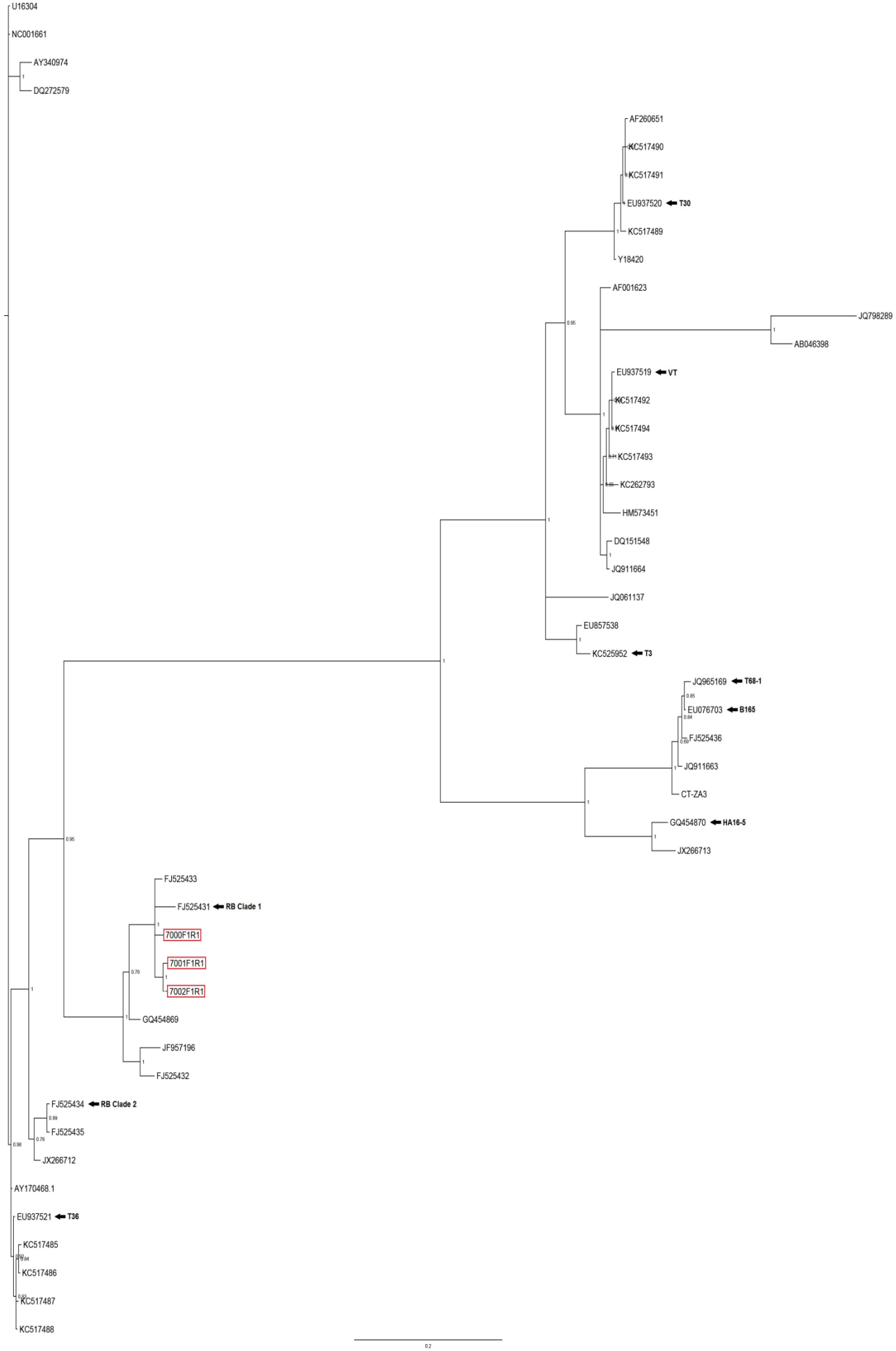


Figure 49: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F1R1 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

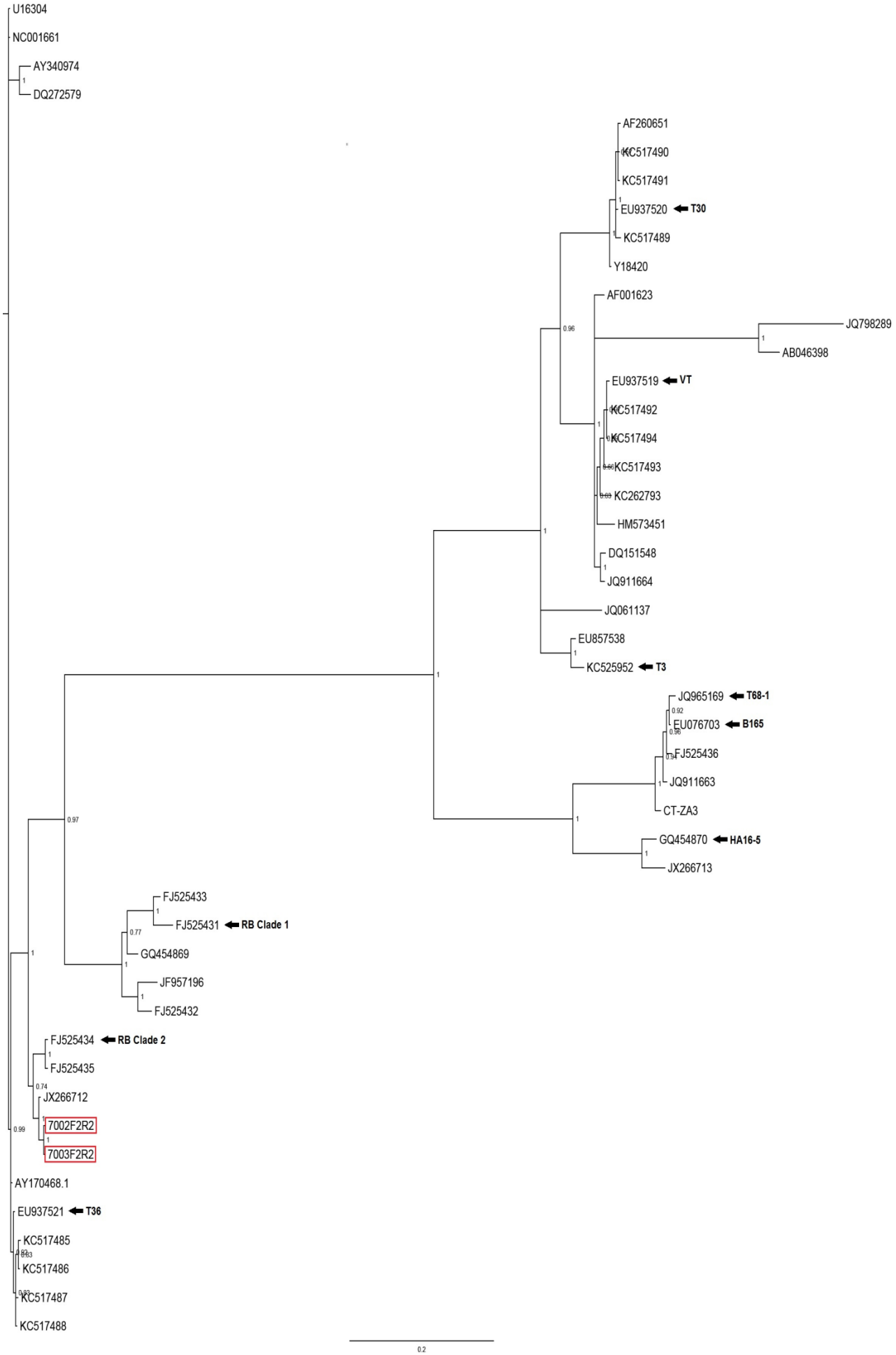


Figure 50: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F2R2 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

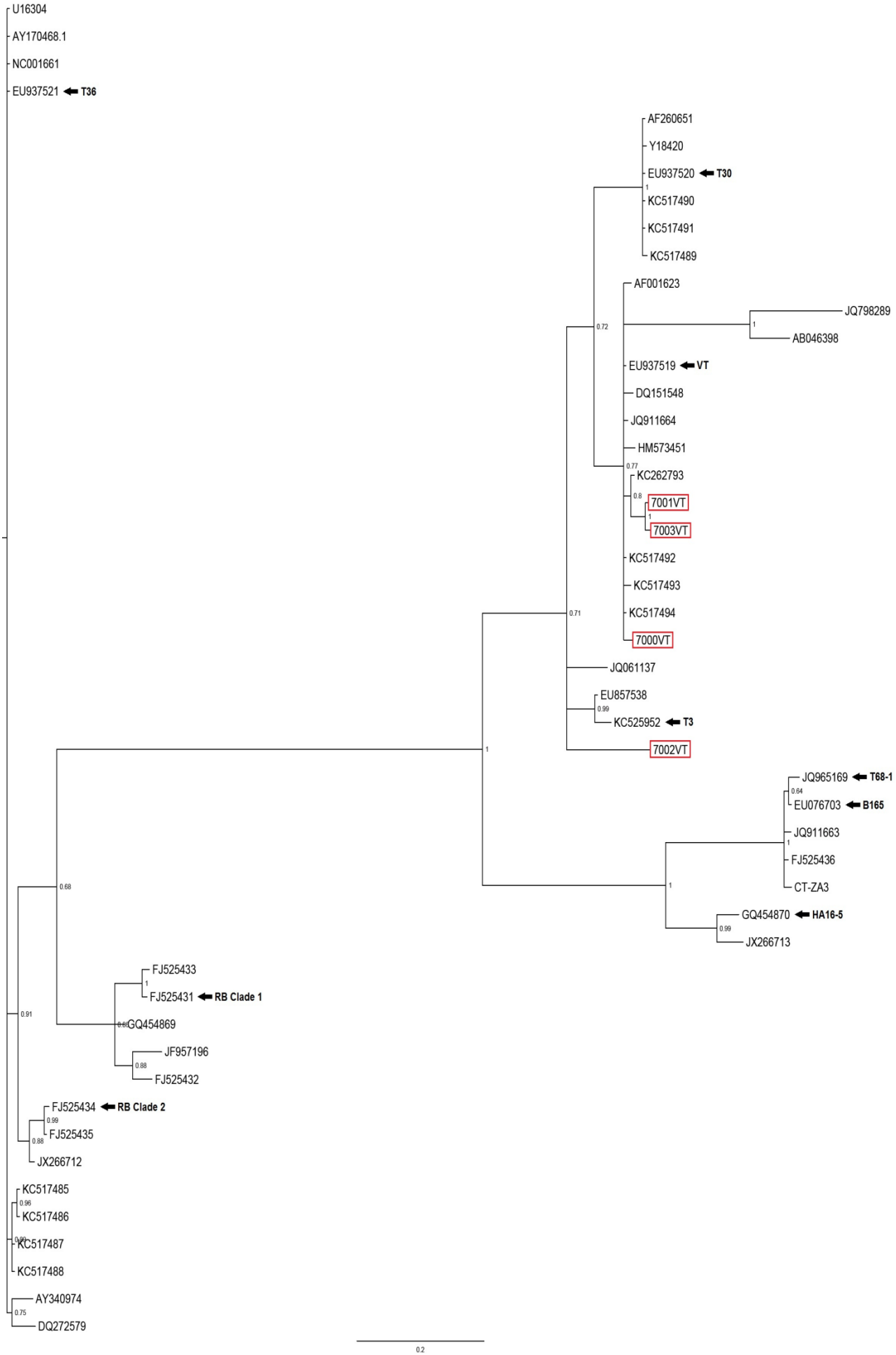


Figure 51: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the VT RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

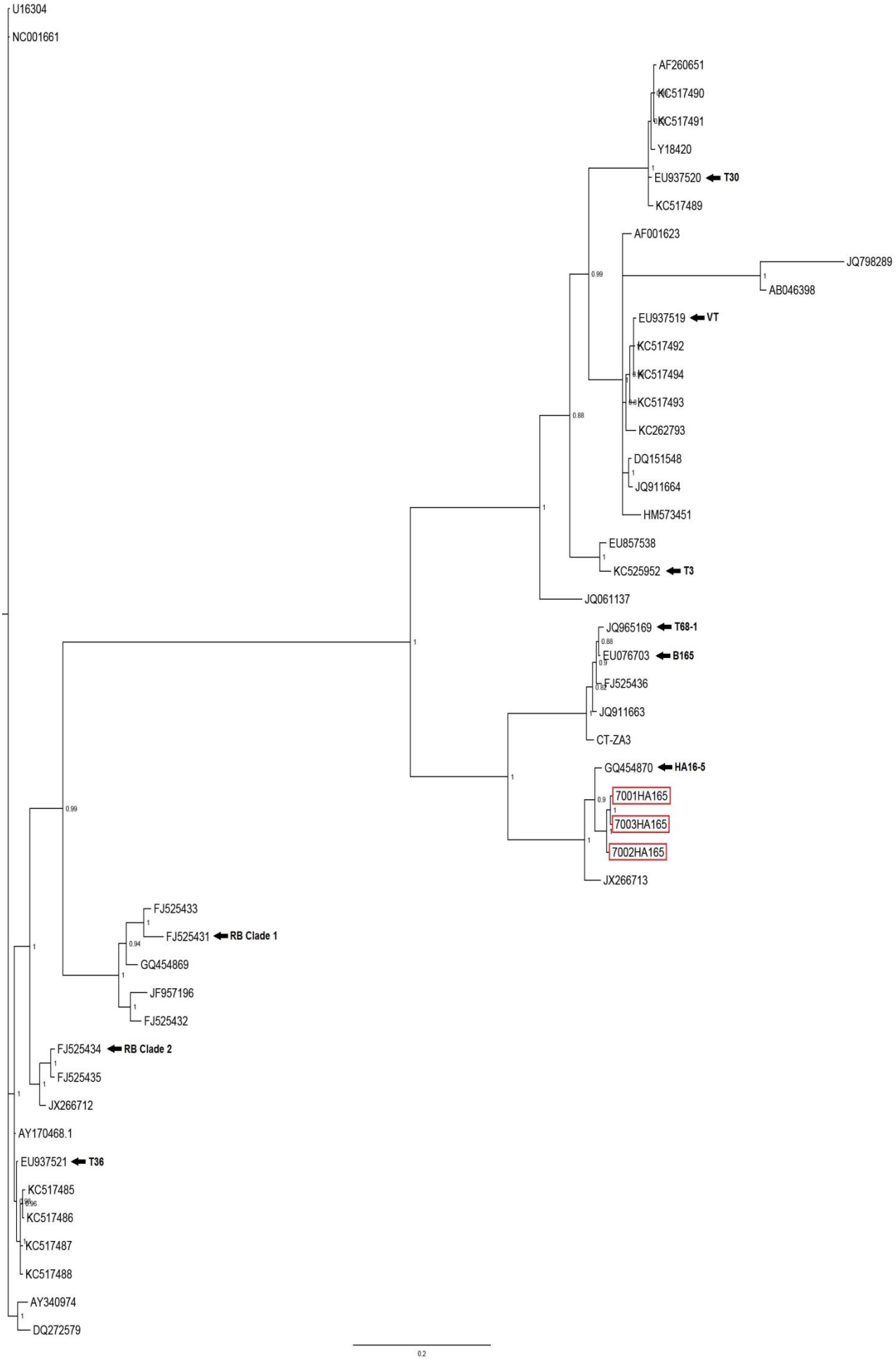


Figure 52: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the HA16-5 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

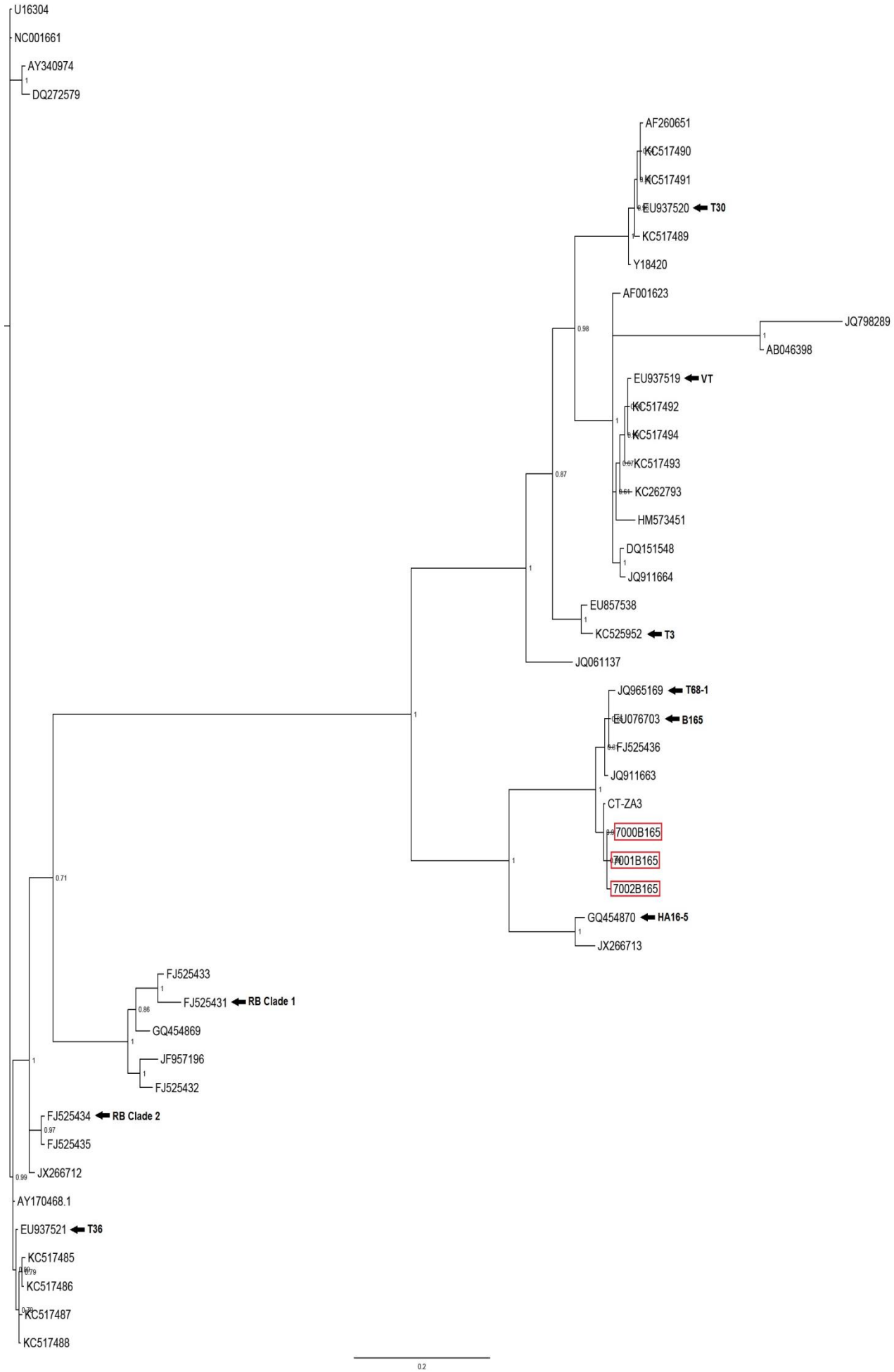


Figure 53: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the B165 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

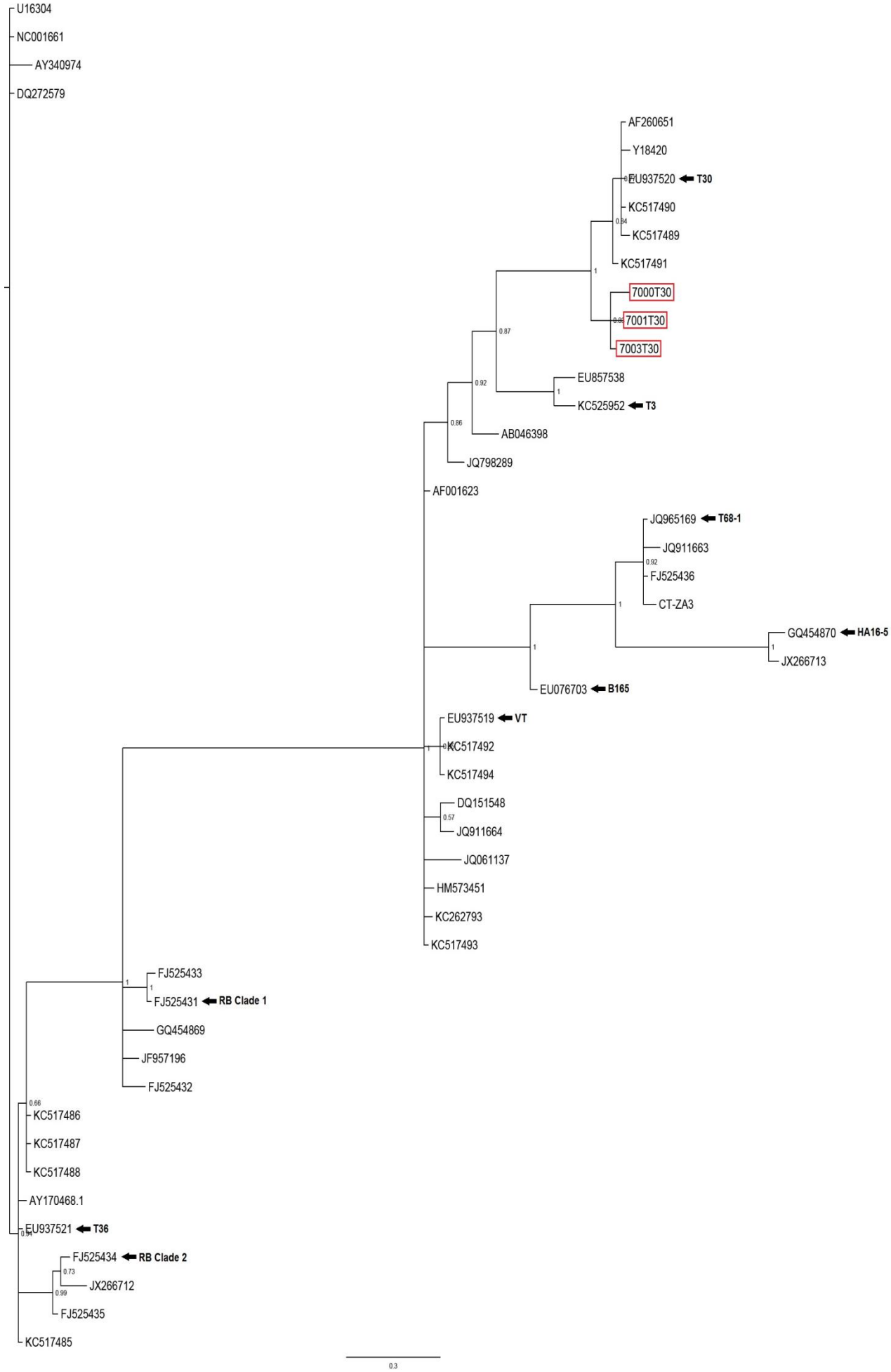


Figure 54: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T30 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

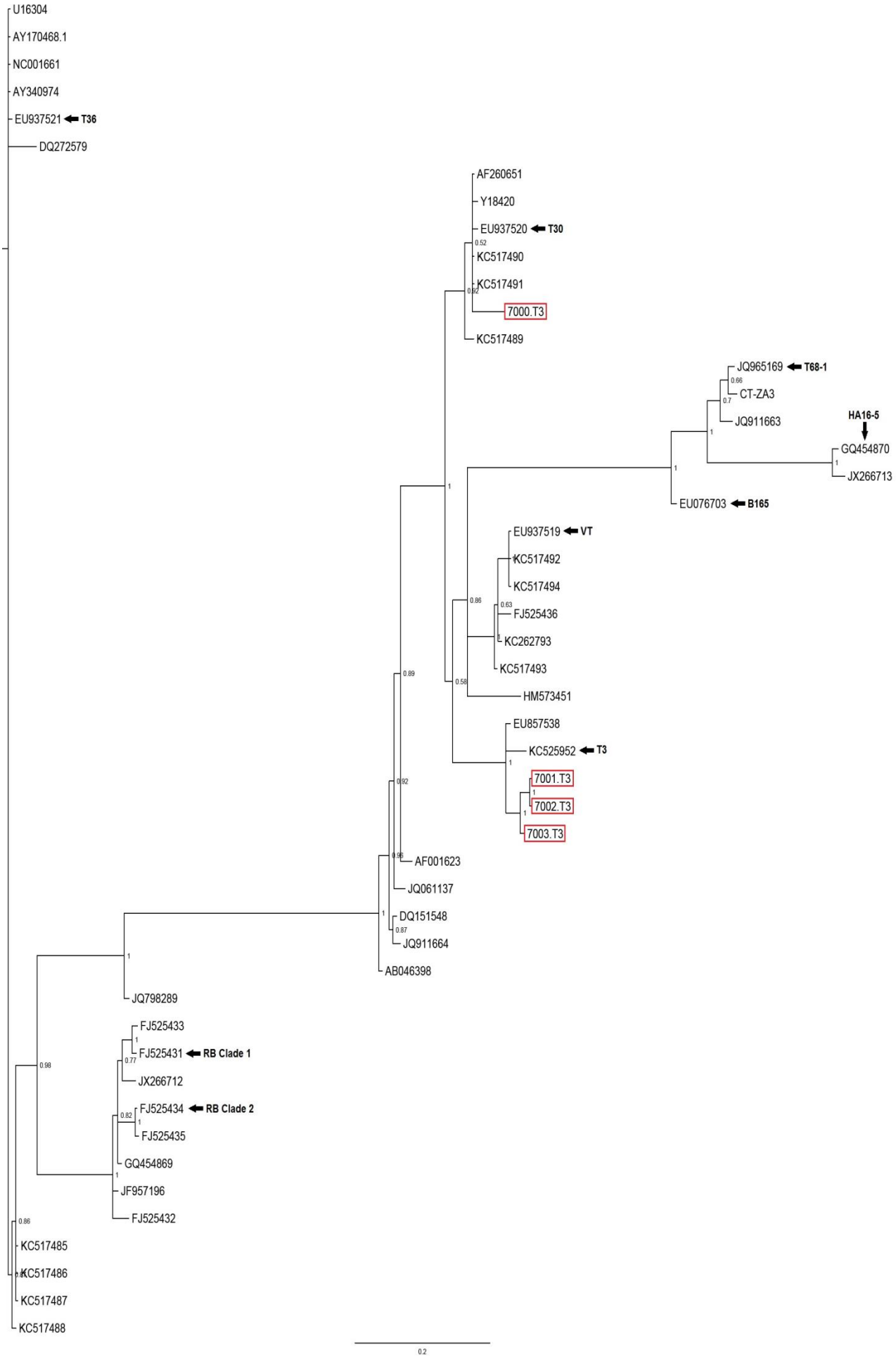


Figure 55: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T3 RT-PCR of the citrus trees sampled, prior to the SAT trails, and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

Appendix 4

Sequence duplication levels and PHRED scores for the Illumina data obtained of the six CTV SAT sub-isolates.

Sample 14-7102

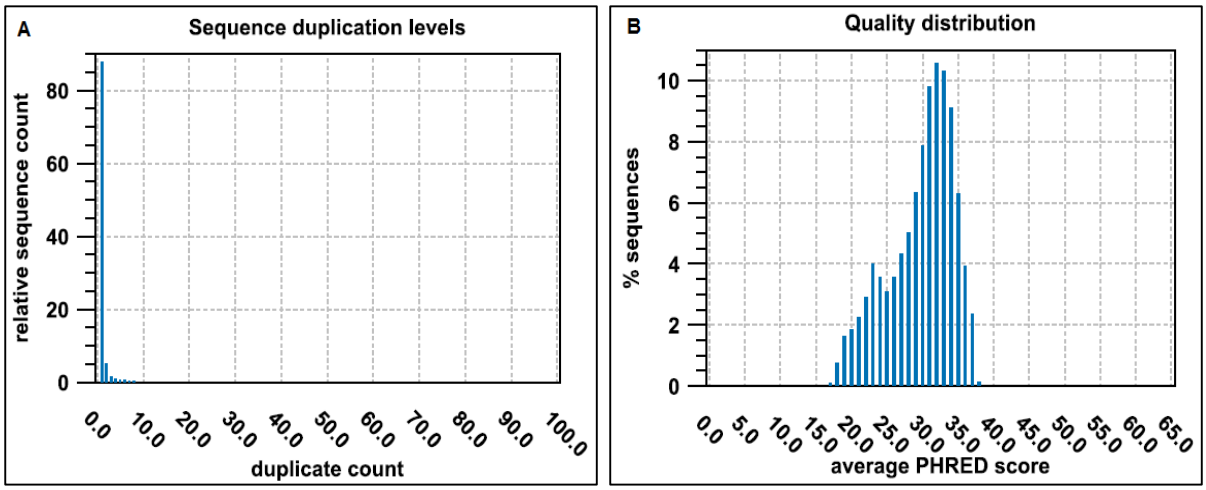


Figure 56: [A] Sequence duplication levels attained for sample 14-7102 during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7102. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Sample 14-7103

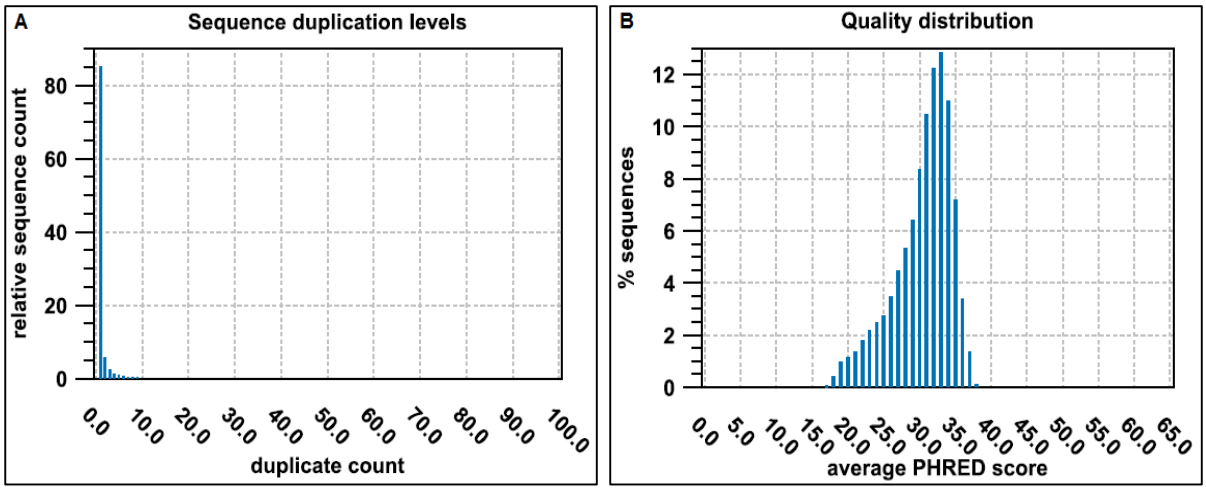


Figure 57: [A] Sequence duplication levels attained for sample 14-7103 during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7103. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Sample 14-7104

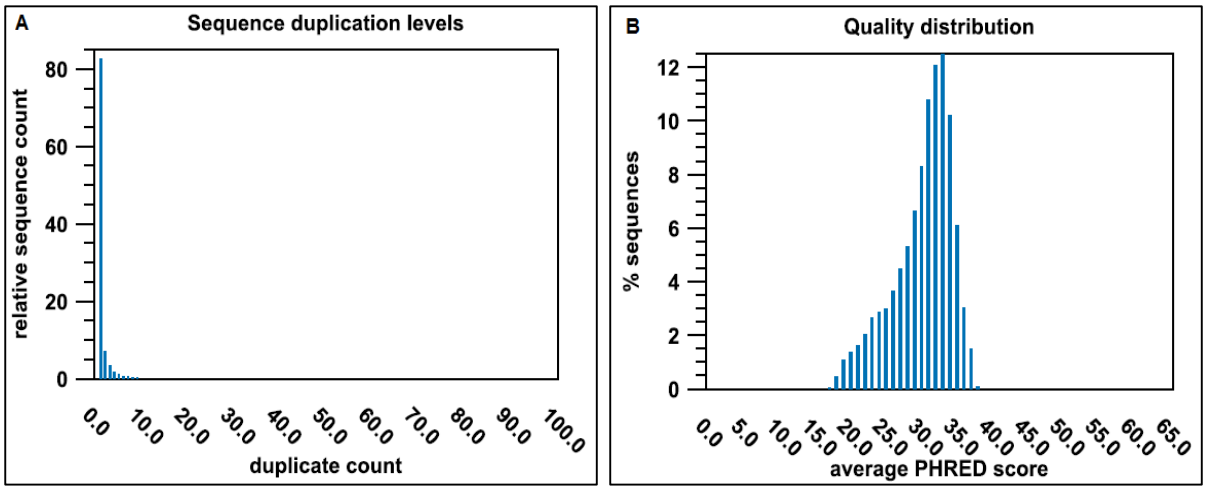


Figure 58: [A] Sequence duplication levels attained for sample 14-7104 during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7104. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Sample 14-7105

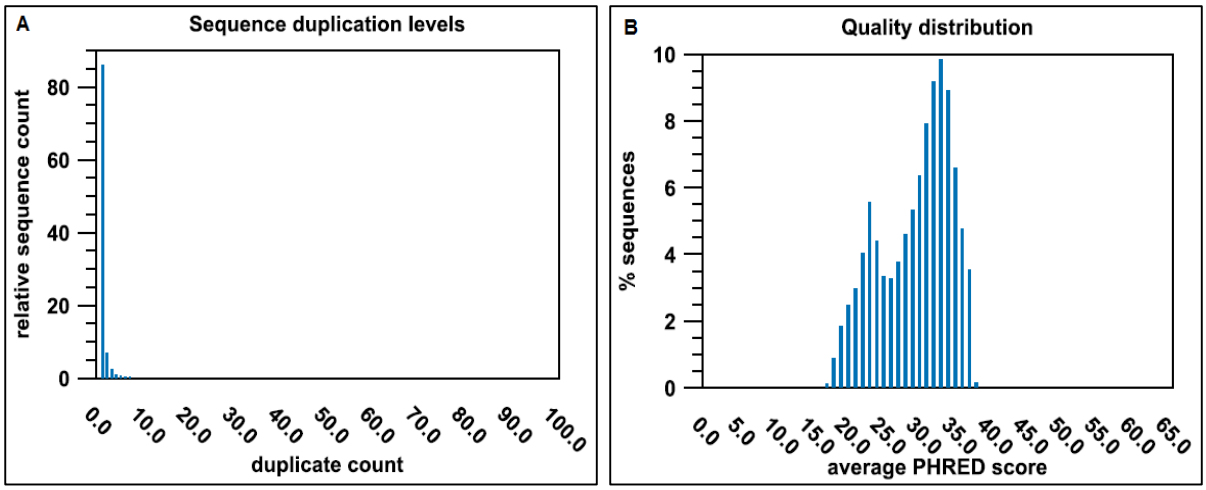


Figure 59: [A] Sequence duplication levels attained for sample 14-7105 during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7105. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Sample 14-7107 [1]

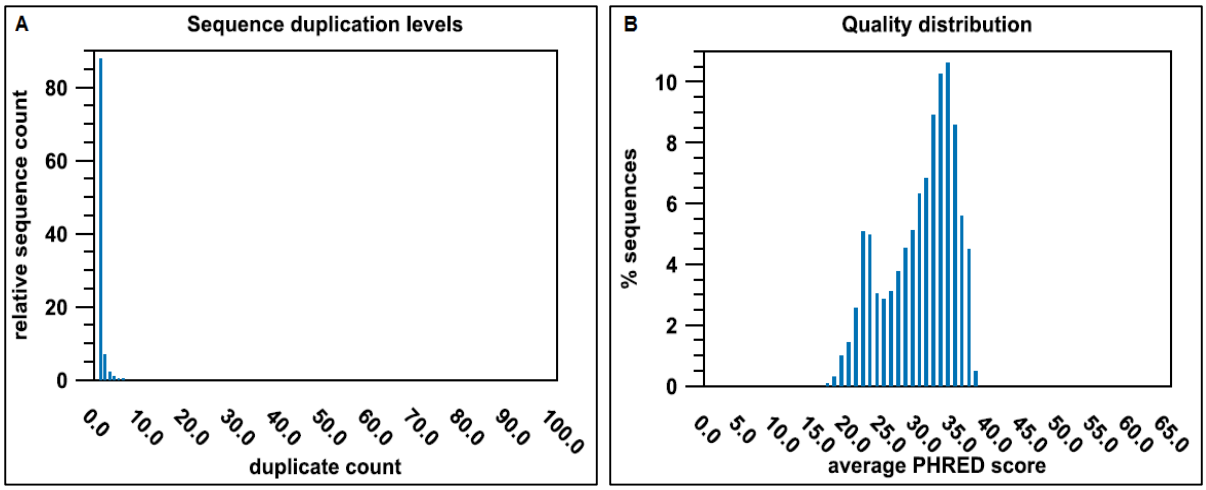


Figure 60: [A] Sequence duplication levels attained for sample 14-7107 [1] during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7107 [1]. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Sample 14-7107 [2]

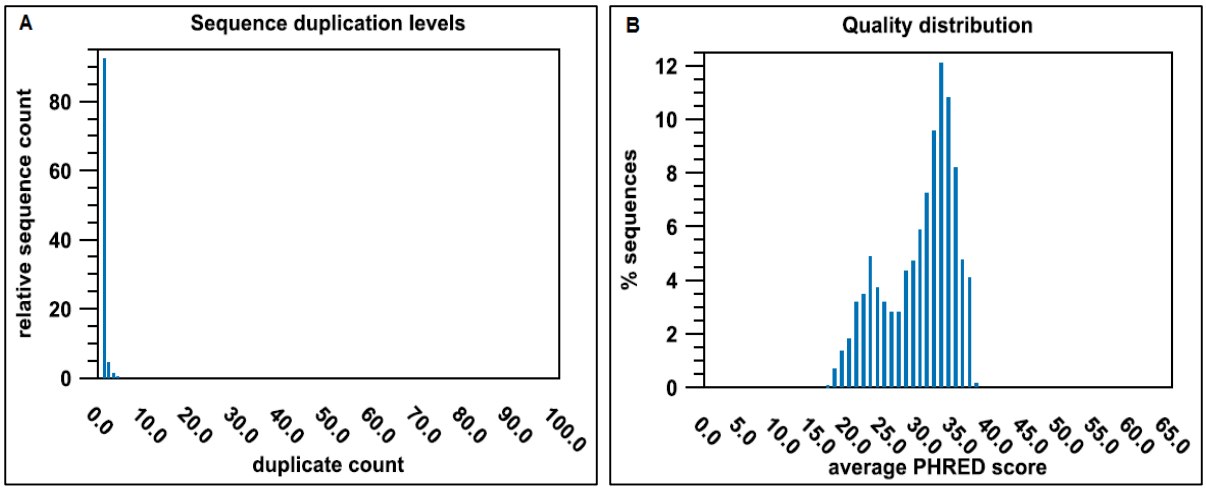


Figure 61: [A] Sequence duplication levels attained for sample 14-7107 [2] during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7107 [2]. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Sample 14-7132

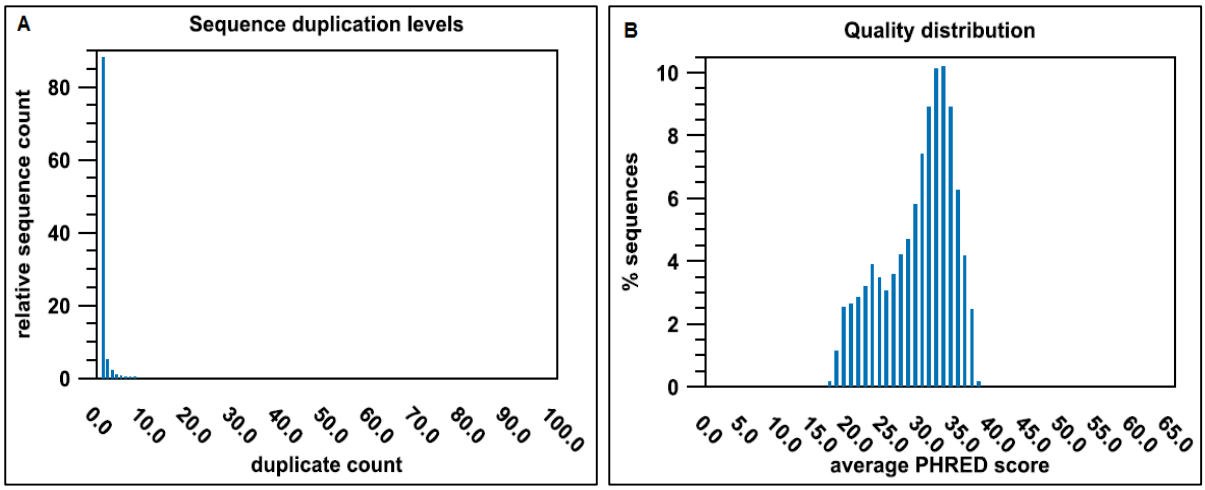


Figure 62: [A] Sequence duplication levels attained for sample 14-7132 during Illumina Next-Generation High-throughput sequencing data analysis. X-axis: Duplicate count. Y-axis: Number of sequences that have been found that many times normalized to the number of unique sequences. **[B]** Distribution of the average sequence quality scores expressed as PHRED scores for sample 14-7132. The quality of the sequence is calculated as the arithmetic mean of its base qualities. X-axis: Average PHRED score. Y-axis: Number of sequences observed at that quality score normalized to the total number of sequences.

Appendix 5

Illumina reference mapping data statistics obtained for the Mexican Lime sample 12-7006 and the six SAT CTV sub-isolates.

Mexican Lime sample 12-7006

Table 23: Illumina reference mapping data statistics for CTV subisolate 12-7006, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
FJ525434_NZRB-TH30	977	102500	3182	99318	27760.13	977
FJ525433_NZRB-TH28	977	1355	413	942	375.402	977
FJ525431_NZRB-M12	977	650	290	360	181.067	977
JX266712_Taiwan-Pum_SP_T1	977	225	165	60	59.9	977
FJ525432_NZRB-G90	603	53	53	0	11.946	977
JX266713_Taiwan-Pum_M_T5	253	12	12	0	2.169	977
FJ525435_NZRB-M17	575	8	8	0	1.911	977
HM573451_Kpg	579	7	5	2	1.976	977
JQ061137_AT-1	289	4	4	0	0.825	977
GQ454869_HA18-9	582	3	3	0	0.831	977
EU857538_SP	301	3	3	0	0.919	977
GQ454870_HA16-5	312	2	2	0	0.32	977
DQ151548_T318A	8	2	2	0	0.014	977
AY170468	301	1	1	0	0.307	981
JF957196_B301	194	1	1	0	0.199	977
AB0463981_NUagA	0	0	0	0	0	977
KC525952_T3	0	0	0	0	0	977
JQ911664_CT11A	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517493_FL202-VT	0	0	0	0	0	977
AY340974_Qaha	0	0	0	0	0	978
KC517490	0	0	0	0	0	977

FJ525436_NZ-B18	0	0	0	0	0	977
KC517487_FS703-T36	0	0	0	0	0	981
EU937519_VT	0	0	0	0	0	977
EU937520	0	0	0	0	0	977
DQ272579_Mexico	0	0	0	0	0	978
AF260651_T30	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
JQ911663_CT14A	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
JQ798289_A18	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
KC262793_L192GR	0	0	0	0	0	977
KC517485_FS674-T36	0	0	0	0	0	981
KC517486_FS701-T36	0	0	0	0	0	981
EU937521_T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
EU076703_B165	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
CT-ZA3	0	0	0	0	0	977
KC517489	0	0	0	0	0	977

SAT positive Mexican Lime seedlings

Table 24: Illumina reference mapping data statistics for CTV subisolate 14-7102, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JX266712_Taiwan-Pum_SP_T1	977	27072	8492	18580	7527.556807	977
JQ061137_AT-1	977	1041	363	678	277.765609	977
FJ525431_NZRB-M12	475	620	580	40	181.9447288	977
GQ454869_HA18-9	933	426	410	16	129.3265097	977
FJ525432_NZRB-G90	977	158	158	0	41.92221085	977
JQ911663_CT14A	469	127	111	16	36.67860798	977
FJ525435_NZRB-M17	977	73	73	0	20.10235415	977
JF957196_B301	650	31	31	0	7.969293756	977
HM573451_Kpg	922	22	14	8	5.819856704	977
JQ911664_CT11A	700	22	22	0	6.524053224	977
GQ454870_HA16-5	553	17	15	2	5.179119754	977
FJ525434_NZRB-TH30	439	13	13	0	3.876151484	977
DQ151548_T318A	491	7	7	0	2.118730809	977
EU857538_SP	383	6	6	0	1.786079836	977
FJ525433_NZRB-TH28	366	6	6	0	1.673490276	977
KC517493_FL202-VT	445	4	4	0	1.064483112	977
KC262793_L192GR	297	3	3	0	0.811668373	977
JX266713_Taiwan-Pum_M_T5	194	2	2	0	0.396110542	977
KC517487_FS703-T36	6	1	1	0	0.006116208	981
KC525952_T3	265	1	1	0	0.271238485	977
AB0463981_NUagA	0	0	0	0	0	977
KC517486_FS701-T36	0	0	0	0	0	981

JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981
EU937520	0	0	0	0	0	977
JQ798289_A18	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
AY340974_Qaha	0	0	0	0	0	978
KC517489	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
AY170468	0	0	0	0	0	981
DQ272579_Mexico	0	0	0	0	0	978
EU076703_B165	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
CT-ZA3	0	0	0	0	0	977
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977
EU937521_T36	0	0	0	0	0	981

Table 25: Illumina reference mapping data statistics for CTV subisolate 14-7103, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JQ061137_AT-1	977	16411	3951	12460	4397.318321	977
JX266712_Taiwan-Pum_SP_T1	977	2250	910	1340	629.0839304	977
JQ911663_CT14A	491	1674	1222	452	477.1432958	977
JQ911664_CT11A	532	319	319	0	95.35619243	977
DQ151548_T318A	536	87	85	2	25.93346981	977
FJ525431_NZRB-M12	765	62	58	4	17.96724667	977
GQ454869_HA18-9	561	32	32	0	9.659160696	977
FJ525432_NZRB-G90	766	16	16	0	4.409416581	977
KC517493_FL202-VT	451	13	13	0	3.713408393	977
FJ525435_NZRB-M17	762	9	9	0	2.674513818	977
HM573451_Kpg	650	6	2	4	1.726714432	977
KC262793_L192GR	393	4	4	0	1.169907881	977
GQ454870_HA16-5	319	4	4	0	1.226202661	977
AB0463981_NUagA	561	3	3	0	0.817809621	977
JF957196_B301	433	3	3	0	0.917093142	977
KC525952_T3	231	2	2	0	0.250767656	977
EU857538_SP	235	1	1	0	0.240532242	977
FJ525433_NZRB-TH28	301	1	1	0	0.308085977	977
JQ798289_A18	208	1	1	0	0.212896622	977
AY170468	225	1	1	0	0.229357798	981
KC517487_FS703-T36	0	0	0	0	0	981
KC517486_FS701-T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981

EU937520	0	0	0	0	0	977
JX266713_Taiwan-Pum_M_T5	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
AY340974_Qaha	0	0	0	0	0	978
KC517489	0	0	0	0	0	977
FJ525434_NZRB-TH30	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
DQ272579_Mexico	0	0	0	0	0	978
EU076703_B165	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
CT-ZA3	0	0	0	0	0	977
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977
EU937521_T36	0	0	0	0	0	981

Table 26: Illumina reference mapping data statistics for CTV subisolate 14-7104, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JQ061137_AT-1	977	11150	3092	8058	2985.903787	977
AY170468	981	8203	1397	6806	2242.099898	981
JX266712_Taiwan-Pum_SP_T1	977	1901	753	1148	535.9355169	977
EU857538_SP	786	1691	1335	356	492.1003071	977
FJ525431_NZRB-M12	898	1419	887	532	421.4513818	977
JQ911663_CT14A	509	1345	1029	316	388.7993859	977
HM573451_Kpg	832	553	485	68	161.2507677	977
GQ454870_HA16-5	977	474	412	62	134.6458547	977
JQ798289_A18	967	233	231	2	69.49437052	977
EU937521_T36	981	184	154	30	48.42813456	981
KC525952_T3	747	175	173	2	51.6591607	977
JQ911664_CT11A	543	174	172	2	52.39099284	977
DQ151548_T318A	531	69	69	0	20.78607984	977
GQ454869_HA18-9	745	24	24	0	7.271238485	977
AY340974_Qaha	354	20	20	0	5.478527607	978
FJ525432_NZRB-G90	676	16	16	0	4.430910952	977
KC517493_FL202-VT	452	15	15	0	4.34493347	977
FJ525435_NZRB-M17	700	8	8	0	2.30296827	977
KC262793_L192GR	297	6	4	2	1.518935517	977
JF957196_B301	571	6	6	0	1.636642784	977
DQ272579_Mexico	683	6	6	0	1.461145194	978
AB0463981_NUagA	473	3	3	0	0.90174002	977
CT-ZA3	332	3	3	0	0.852610031	977
JX266713_Taiwan-Pum_M_T5	15	1	1	0	0.01432958	977
KC517487_FS703-T36	0	0	0	0	0	981

KC517486_FS701-T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981
EU937520	0	0	0	0	0	977
FJ525433_NZRB-TH28	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
KC517489	0	0	0	0	0	977
FJ525434_NZRB-TH30	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
EU076703_B165	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977

Table 27: Illumina reference mapping data statistics for CTV subisolate 14-7105, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JQ061137_AT-1	977	28398	6584	21814	7632.116684	977
JQ911663_CT14A	520	2767	2107	660	790.6397134	977
JQ911664_CT11A	546	429	419	10	128.7871034	977
DQ151548_T318A	533	56	56	0	16.30603889	977
KC517493_FL202-VT	489	36	36	0	10.5506653	977
AY170468	981	16	2	14	4.387359837	981
JX266712_Taiwan-Pum_SP_T1	898	13	9	4	3.515864893	977
EU857538_SP	419	11	7	4	3.178096213	977
HM573451_Kpg	978	11	7	4	2.902763562	977
KC262793_L192GR	393	4	4	0	0.993858751	977
GQ454870_HA16-5	592	4	4	0	0.842374616	977
EU076703_B165	314	3	3	0	0.610429448	978
KC525952_T3	307	2	2	0	0.578300921	977
AB0463981_NUagA	292	1	1	0	0.298874104	977
JQ798289_A18	274	1	1	0	0.280450358	977
AY340974_Qaha	7	1	1	0	0.007157464	978
DQ272579_Mexico	11	1	1	0	0.011247444	978
FJ525431_NZRB-M12	291	1	1	0	0.297850563	977
CT-ZA3	292	1	1	0	0.298874104	977
KC517487_FS703-T36	0	0	0	0	0	981
KC517486_FS701-T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981
EU937520	0	0	0	0	0	977

JX266713_Taiwan-Pum_M_T5	0	0	0	0	0	977
FJ525433_NZRB-TH28	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
GQ454869_HA18-9	0	0	0	0	0	977
KC517489	0	0	0	0	0	977
FJ525434_NZRB-TH30	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
JF957196_B301	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
FJ525435_NZRB-M17	0	0	0	0	0	977
FJ525432_NZRB-G90	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977
EU937521_T36	0	0	0	0	0	981

Table 28: Illumina reference mapping data statistics for CTV subisolate 14-7107 [1], showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JQ061137_AT-1	977	1277	387	890	339.5056295	977
HM573451_Kpg	977	313	159	154	85.03889458	977
JQ911663_CT14A	470	284	220	64	81.19959058	977
EU857538_SP	942	85	65	20	24.71954964	977
JQ911664_CT11A	404	32	32	0	9.536335722	977
KC525952_T3	961	30	30	0	8.403275333	977
JQ798289_A18	874	26	24	2	7.761514841	977
GQ454870_HA16-5	535	24	24	0	6.629477994	977
JX266712_Taiwan-Pum_SP_T1	729	4	4	0	1.062436029	977
DQ151548_T318A	443	4	4	0	1.152507677	977
KC517493_FL202-VT	294	2	2	0	0.575230297	977
AB0463981_NUagA	0	0	0	0	0	977
KC517487_FS703-T36	0	0	0	0	0	981
KC517486_FS701-T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981
EU937520	0	0	0	0	0	977
JX266713_Taiwan-Pum_M_T5	0	0	0	0	0	977
FJ525433_NZRB-TH28	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
KC262793_L192GR	0	0	0	0	0	977
AY340974_Qaha	0	0	0	0	0	978
GQ454869_HA18-9	0	0	0	0	0	977
KC517489	0	0	0	0	0	977

FJ525434_NZRB-TH30	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
JF957196_B301	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
AY170468	0	0	0	0	0	981
DQ272579_Mexico	0	0	0	0	0	978
FJ525431_NZRB-M12	0	0	0	0	0	977
EU076703_B165	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
FJ525435_NZRB-M17	0	0	0	0	0	977
FJ525432_NZRB-G90	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
CT-ZA3	0	0	0	0	0	977
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977
EU937521_T36	0	0	0	0	0	981

Table 29: Illumina reference mapping data statistics for CTV subisolate 14-7107 [2], showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JQ061137_AT-1	977	512	122	390	136.3561924	977
HM573451_Kpg	977	105	63	42	28.08495394	977
EU857538_SP	461	41	33	8	12.17502559	977
JQ911663_CT14A	427	38	30	8	10.89662231	977
GQ454870_HA16-5	605	13	13	0	3.654042989	977
JQ911664_CT11A	511	8	8	0	2.362333675	977
JQ798289_A18	540	7	7	0	1.946775844	977
DQ151548_T318A	469	3	3	0	0.907881269	977
KC525952_T3	447	2	2	0	0.45752303	977
JX266712_Taiwan-Pum_SP_T1	488	2	2	0	0.499488229	977
KC517493_FL202-VT	284	1	1	0	0.290685773	977
AB0463981_NUagA	0	0	0	0	0	977
KC517487_FS703-T36	0	0	0	0	0	981
KC517486_FS701-T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981
EU937520	0	0	0	0	0	977
JX266713_Taiwan-Pum_M_T5	0	0	0	0	0	977
FJ525433_NZRB-TH28	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
KC262793_L192GR	0	0	0	0	0	977
AY340974_Qaha	0	0	0	0	0	978
GQ454869_HA18-9	0	0	0	0	0	977
KC517489	0	0	0	0	0	977

FJ525434_NZRB-TH30	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
JF957196_B301	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
AY170468	0	0	0	0	0	981
DQ272579_Mexico	0	0	0	0	0	978
FJ525431_NZRB-M12	0	0	0	0	0	977
EU076703_B165	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
FJ525435_NZRB-M17	0	0	0	0	0	977
FJ525432_NZRB-G90	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
CT-ZA3	0	0	0	0	0	977
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977
EU937521_T36	0	0	0	0	0	981

Table 30: Illumina reference mapping data statistics for CTV subisolate 14-7107, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences. Results obtained from sample 14-7107 [1] and sample 14-7107 [2] were combined to get an overall results for SAT sample 14-7107.

14-7107 [1]			14-7107 [2]			Combined		
Reference sequence name	**CL**	**TRC**	**Reference sequence name**	**CL**	**TRC**	**Reference sequence name**	**CL**	**TRC**
AB0463981_NUagA	0	0	AB0463981_NUagA	0	0	JQ061137_AT-1	977	1789
KC517487_FS703-T36	0	0	KC517487_FS703-T36	0	0	HM573451_Kpg	977	418
KC517486_FS701-T36	0	0	KC517486_FS701-T36	0	0	JQ911663_CT14A	448.5	322
JQ965169_T68	0	0	JQ965169_T68	0	0	EU857538_SP	701.5	126
EU857538_SP	942	85	EU857538_SP	461	41	JQ911664_CT11A	457.5	40
KC517488_FS577	0	0	KC517488_FS577	0	0	GQ454870_HA16-5	570	37
KC517485_FS674-T36	0	0	KC517485_FS674-T36	0	0	JQ798289_A18	707	33
EU937520	0	0	EU937520	0	0	KC525952_T3	704	32
JX266713_Taiwan-Pum_M_T5	0	0	JX266713_Taiwan-Pum_M_T5	0	0	DQ151548_T318A	456	7
FJ525433_NZRB-TH28	0	0	FJ525433_NZRB-TH28	0	0	JX266712_Taiwan-Pum_SP_T1	608.5	6
JQ911663_CT14A	470	284	JQ911663_CT14A	427	38	KC517493_FL202-VT	289	3
KC525952_T3	961	30	KC525952_T3	447	2	AB0463981_NUagA	0	0
JQ798289_A18	874	26	JQ798289_A18	540	7	KC517487_FS703-T36	0	0
NC_001661	0	0	NC_001661	0	0	KC517486_FS701-T36	0	0
KC262793_L192GR	0	0	KC262793_L192GR	0	0	JQ965169_T68	0	0
AY340974_Qaha	0	0	AY340974_Qaha	0	0	KC517488_FS577	0	0
GQ454870_HA16-5	535	24	GQ454870_HA16-5	605	13	KC517485_FS674-T36	0	0
GQ454869_HA18-9	0	0	GQ454869_HA18-9	0	0	EU937520	0	0
JQ061137_AT-1	977	1277	JQ061137_AT-1	977	512	JX266713_Taiwan-Pum_M_T5	0	0
KC517489	0	0	KC517489	0	0	FJ525433_NZRB-TH28	0	0
FJ525434_NZRB-TH30	0	0	FJ525434_NZRB-TH30	0	0	NC_001661	0	0

KC517494_FS701-VT	0	0	KC517494_FS701-VT	0	0	KC262793_L192GR	0	0
JF957196_B301	0	0	JF957196_B301	0	0	AY340974_Qaha	0	0
JX266712_Taiwan-Pum_SP_T1	729	4	JX266712_TaiwanPum_SP_T1	488	2	GQ454869_HA18-9	0	0
AF001623_SY568	0	0	AF001623_SY568	0	0	KC517489	0	0
FJ525436_NZ-B18	0	0	FJ525436_NZ-B18	0	0	FJ525434_NZRB-TH30	0	0
HM573451_Kpg	977	313	HM573451_Kpg	977	105	KC517494_FS701-VT	0	0
AY170468	0	0	AY170468	0	0	JF957196_B301	0	0
DQ272579_Mexico	0	0	DQ272579_Mexico	0	0	AF001623_SY568	0	0
DQ151548_T318A	443	4	DQ151548_T318A	469	3	FJ525436_NZ-B18	0	0
FJ525431_NZRB-M12	0	0	FJ525431_NZRB-M12	0	0	AY170468	0	0
EU076703_B165	0	0	EU076703_B165	0	0	DQ272579_Mexico	0	0
KC517491	0	0	KC517491	0	0	FJ525431_NZRB-M12	0	0
KC517490	0	0	KC517490	0	0	EU076703_B165	0	0
FJ525435_NZRB-M17	0	0	FJ525435_NZRB-M17	0	0	KC517491	0	0
FJ525432_NZRB-G90	0	0	FJ525432_NZRB-G90	0	0	KC517490	0	0
JQ911664_CT11A	404	32	JQ911664_CT11A	511	8	FJ525435_NZRB-M17	0	0
KC517492_FS703-VT	0	0	KC517492_FS703-VT	0	0	FJ525432_NZRB-G90	0	0
Y18420_T385	0	0	Y18420_T385	0	0	KC517492_FS703-VT	0	0
U16304_T36	0	0	U16304_T36	0	0	Y18420_T385	0	0
KC517493_FL202-VT	294	2	KC517493_FL202-VT	284	1	U16304_T36	0	0
CT-ZA3	0	0	CT-ZA3	0	0	CT-ZA3	0	0
EU937519_VT	0	0	EU937519_VT	0	0	EU937519_VT	0	0
AF260651_T30	0	0	AF260651_T30	0	0	AF260651_T30	0	0
EU937521_T36	0	0	EU937521_T36	0	0	EU937521_T36	0	0

Consensus Length [CL] was calculated by adding up the CL values of sample 14-7107 [1] and sample 14-7107 [2] together and dividing the result by 2 to get an average length for the specific CTV reference genome. The Total Read Count was calculated by adding up the TRC values of sample 14-7107 [1] and sample 14-7107 [2], to obtain a final result.

Table 31: Illumina reference mapping data statistics for CTV subisolate 14-7132, showing detailed information of the results obtained for each of the CTV p33-edited reference genome sequences.

Reference sequence name	Consensus length	Total read count	Single reads	Reads in pairs	Average coverage	Reference length
JQ061137_AT-1	977	11378	2486	8892	3042.843398	977
JQ911663_CT14A	491	1030	800	230	296.3725691	977
JQ911664_CT11A	734	135	135	0	40.28454452	977
DQ151548_T318A	531	44	44	0	13.30194473	977
KC517493_FL202-VT	477	18	18	0	5.202661208	977
JX266712_Taiwan-Pum_SP_T1	867	13	5	8	3.303991812	977
AY170468	564	4	2	2	1.067278287	981
AB0463981_NUagA	535	3	3	0	0.729785056	977
EU076703_B165	311	3	3	0	0.886503067	978
JX266713_Taiwan-Pum_M_T5	20	2	2	0	0.021494371	977
HM573451_Kpg	568	2	2	0	0.616171955	977
FJ525431_NZRB-M12	323	2	0	2	0.580348004	977
EU857538_SP	299	1	1	0	0.306038895	977
KC262793_L192GR	272	1	1	0	0.278403275	977
KC517487_FS703-T36	0	0	0	0	0	981
KC517486_FS701-T36	0	0	0	0	0	981
JQ965169_T68	0	0	0	0	0	977
KC517488_FS577	0	0	0	0	0	981
KC517485_FS674-T36	0	0	0	0	0	981
EU937520	0	0	0	0	0	977
FJ525433_NZRB-TH28	0	0	0	0	0	977
KC525952_T3	0	0	0	0	0	977
JQ798289_A18	0	0	0	0	0	977
NC_001661	0	0	0	0	0	978
AY340974_Qaha	0	0	0	0	0	978

GQ454870_HA16-5	0	0	0	0	0	977
GQ454869_HA18-9	0	0	0	0	0	977
KC517489	0	0	0	0	0	977
FJ525434_NZRB-TH30	0	0	0	0	0	977
KC517494_FS701-VT	0	0	0	0	0	977
JF957196_B301	0	0	0	0	0	977
AF001623_SY568	0	0	0	0	0	977
FJ525436_NZ-B18	0	0	0	0	0	977
DQ272579_Mexico	0	0	0	0	0	978
KC517491	0	0	0	0	0	977
KC517490	0	0	0	0	0	977
FJ525435_NZRB-M17	0	0	0	0	0	977
FJ525432_NZRB-G90	0	0	0	0	0	977
KC517492_FS703-VT	0	0	0	0	0	977
Y18420_T385	0	0	0	0	0	977
U16304_T36	0	0	0	0	0	978
CT-ZA3	0	0	0	0	0	977
EU937519_VT	0	0	0	0	0	977
AF260651_T30	0	0	0	0	0	977
EU937521_T36	0	0	0	0	0	981

Appendix 6

Bayesian phylogenetic trees produced for the CTV analysis of the 45 Rio Red budwood 2012 graft-inoculated Mexican Lime trees maintained within the greenhouse.

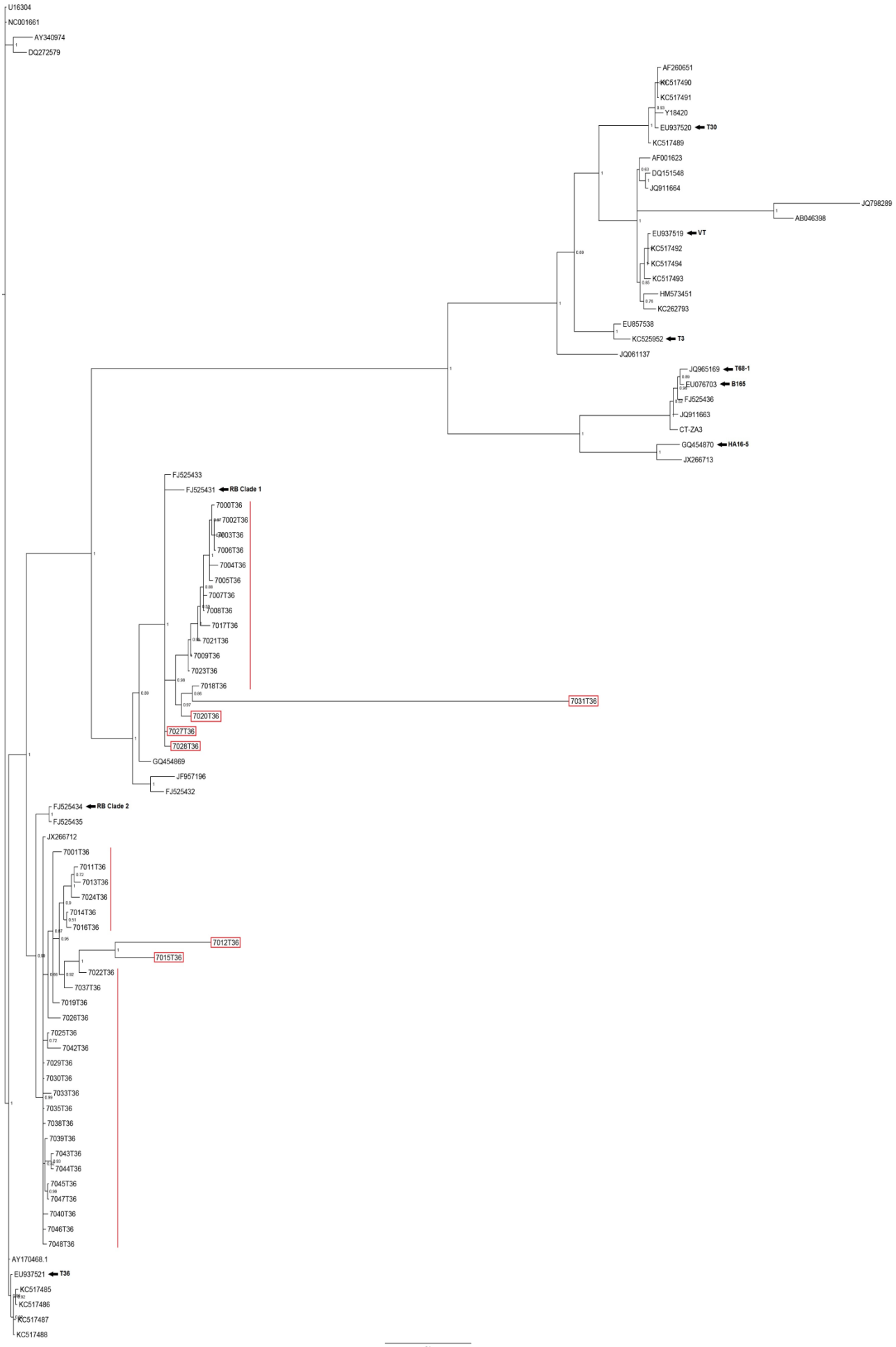


Figure 63: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T36 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

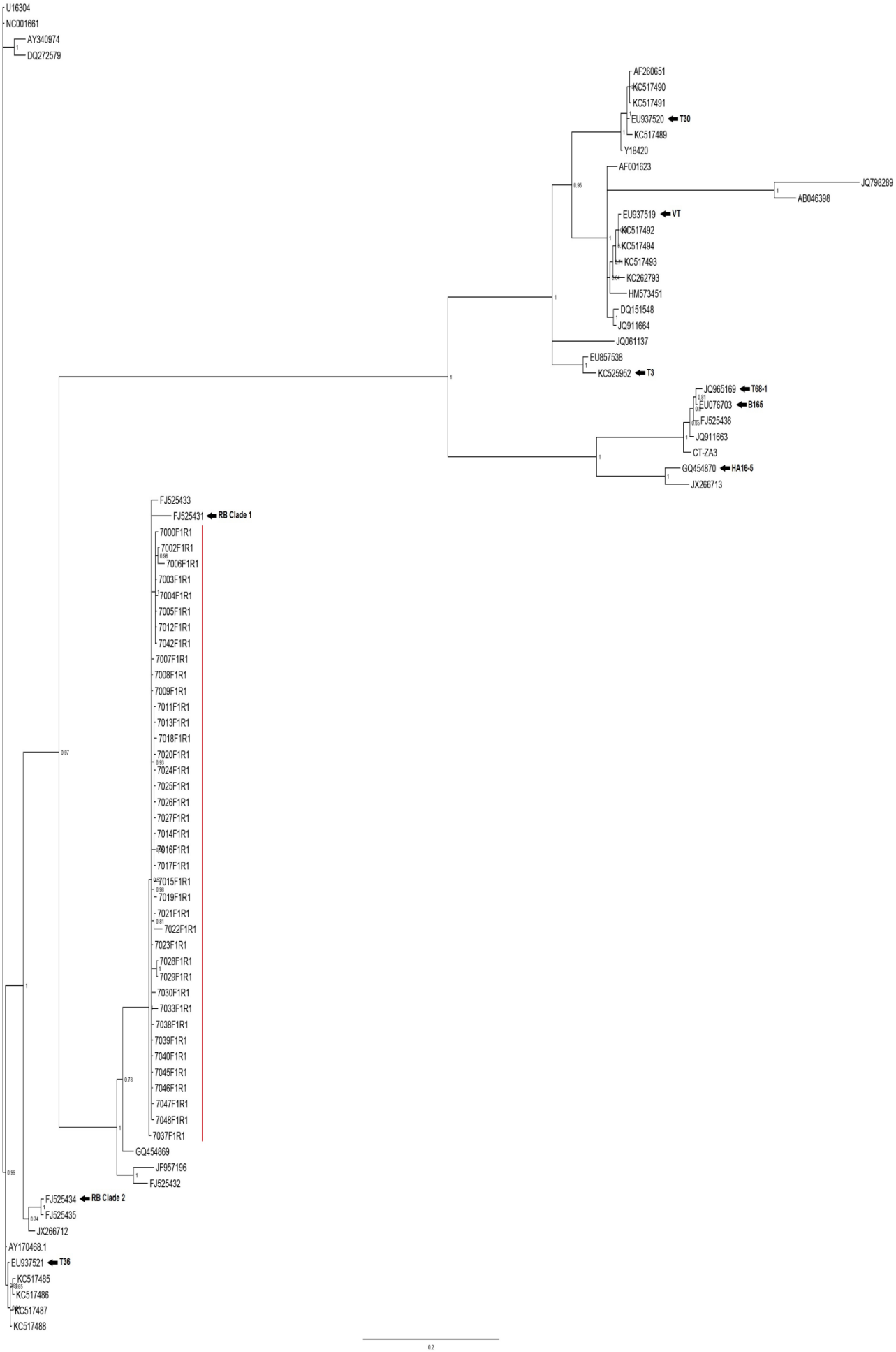


Figure 64: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F1R1 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

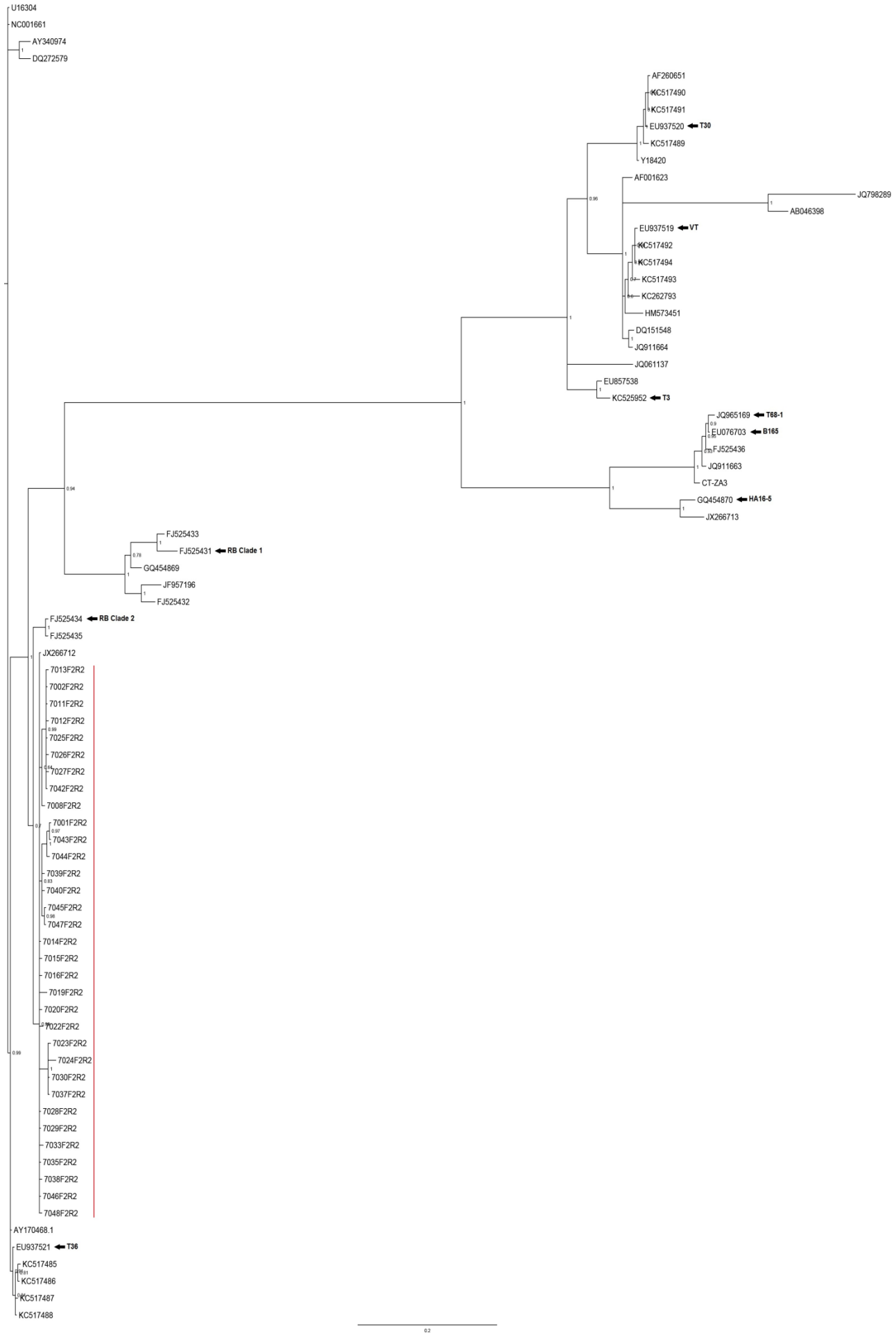


Figure 65: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the NZRB F2R2 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

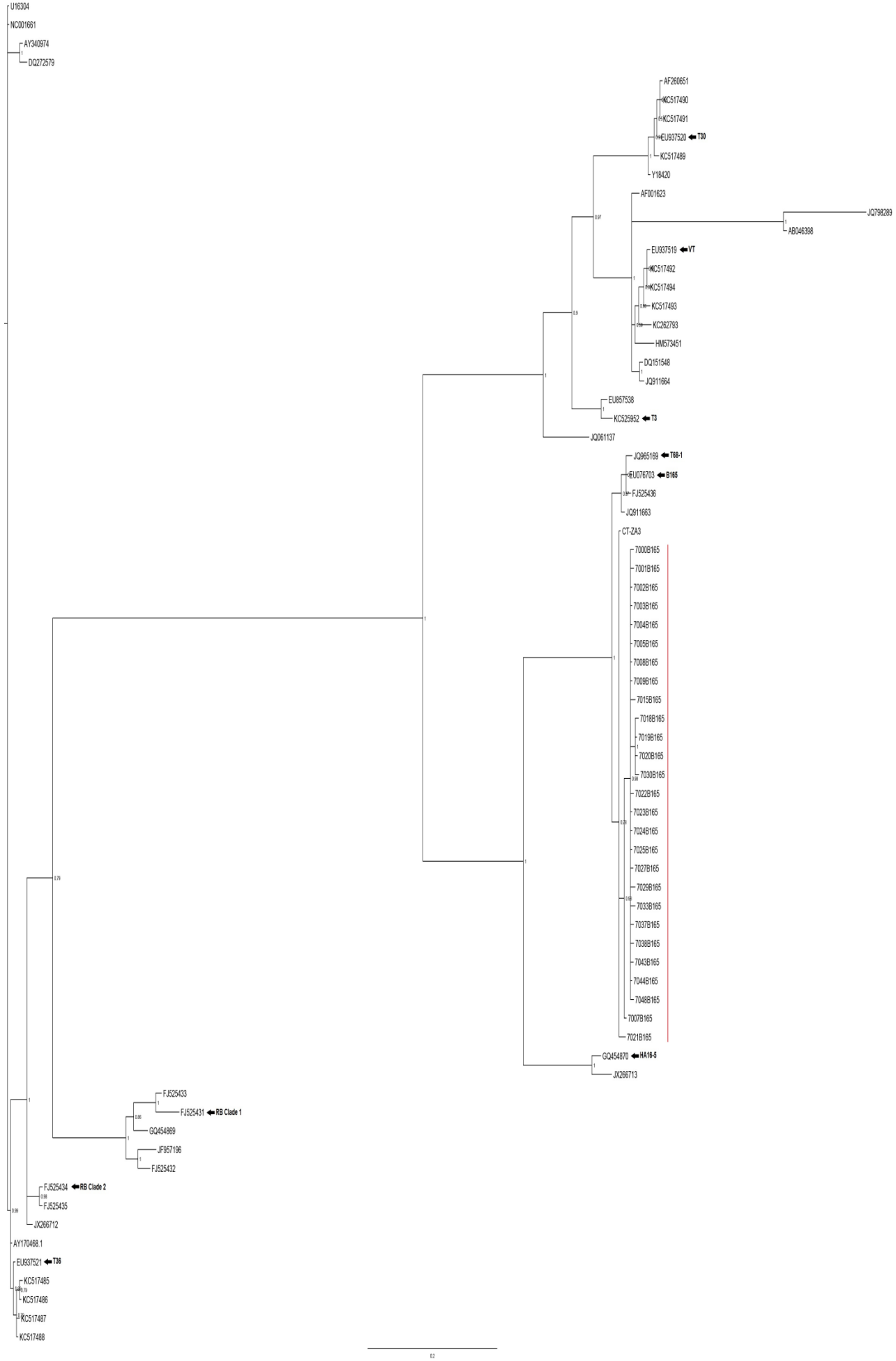


Figure 66: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the B165 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

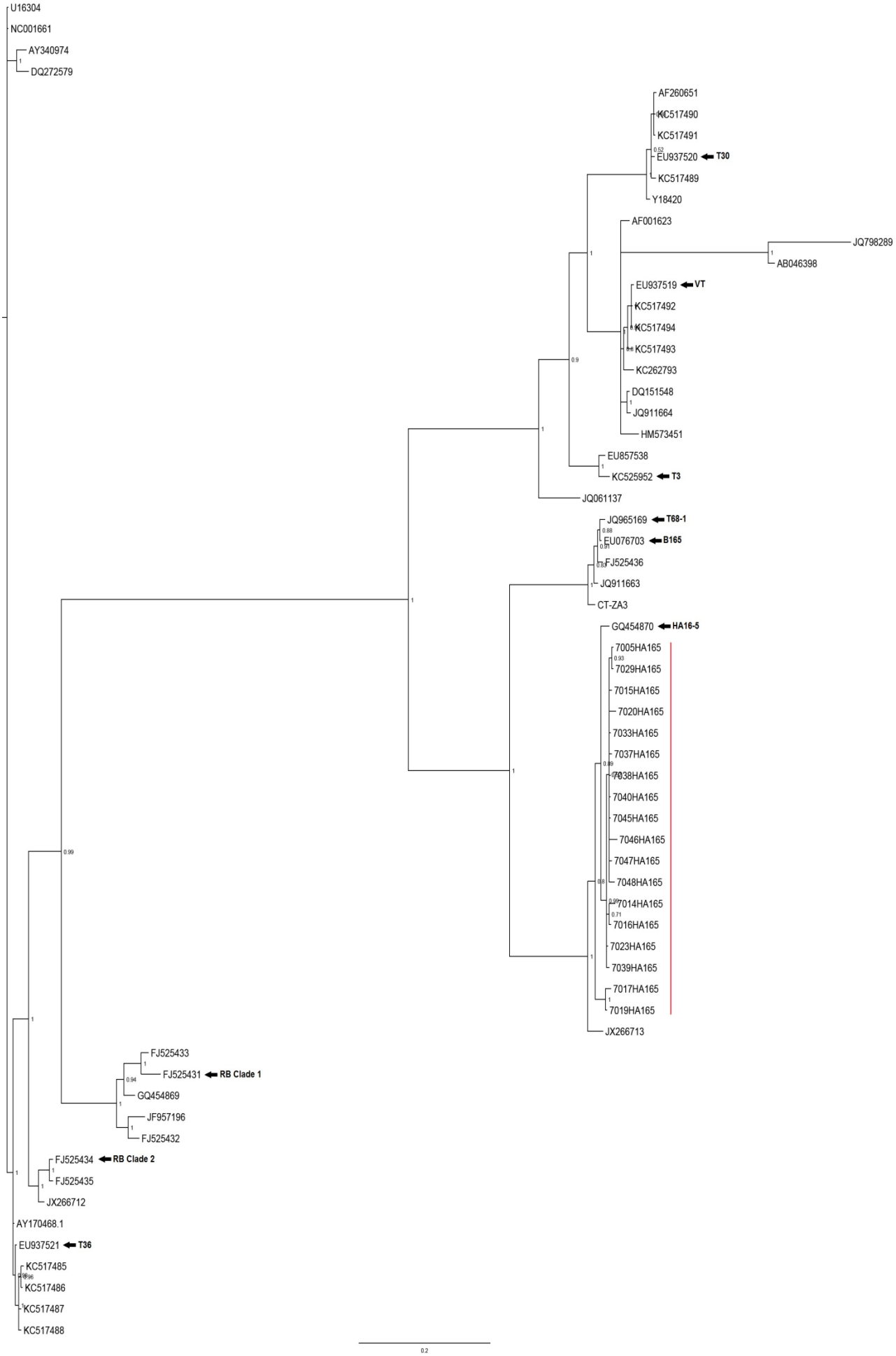


Figure 67: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the HA16-5 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

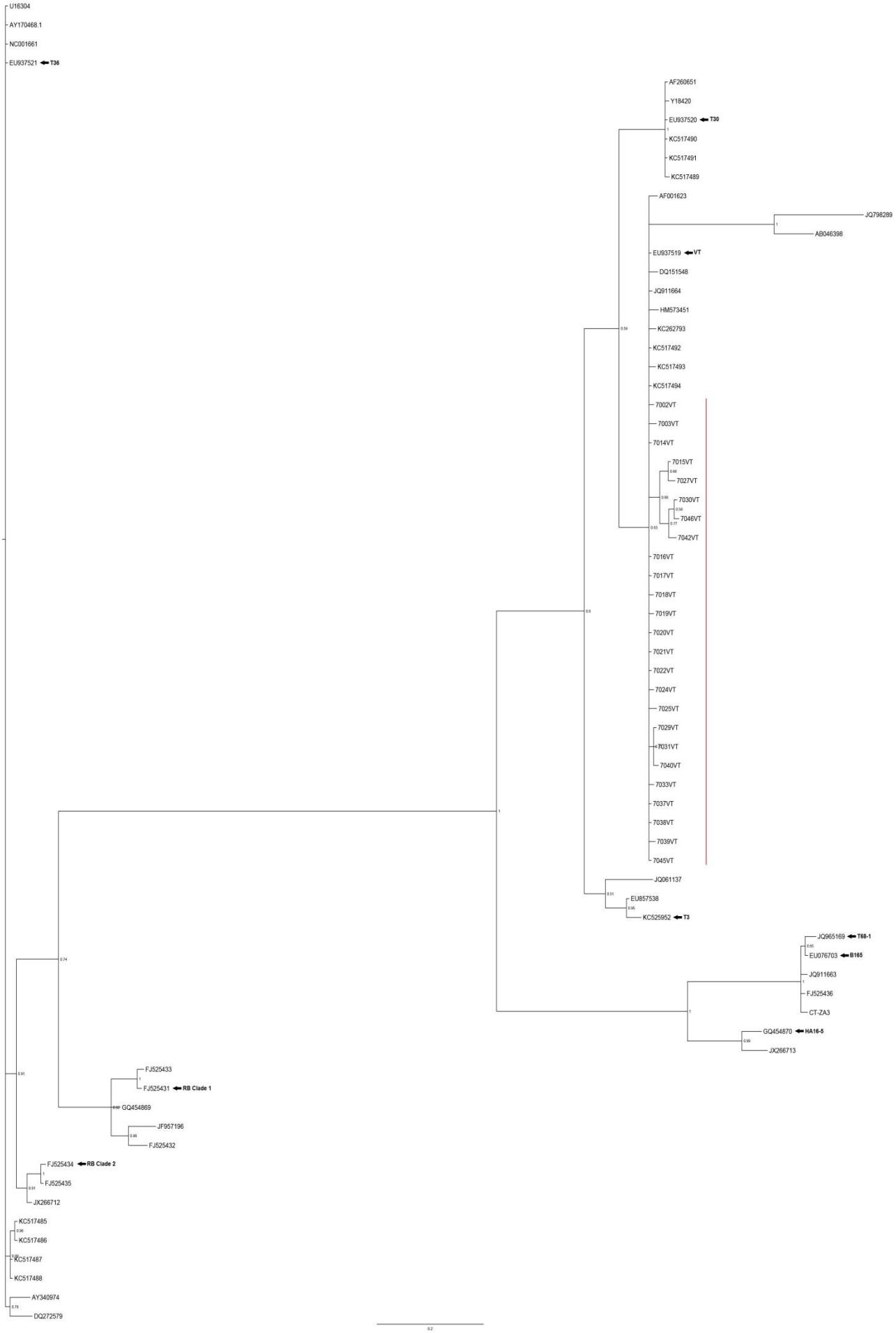


Figure 68: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the VT RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

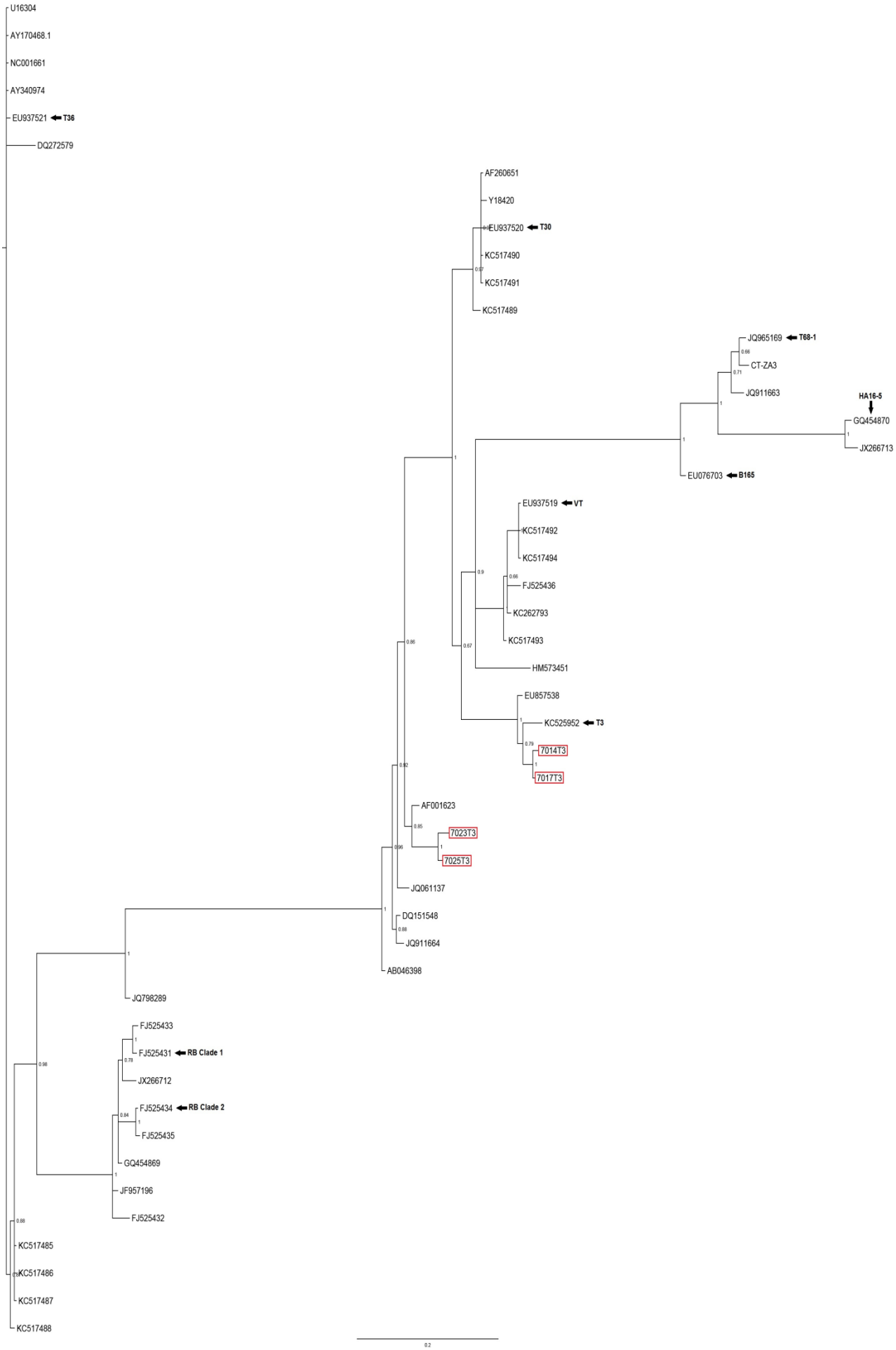


Figure 69: Bayesian dendrogram showing the phylogenetic relationship between the sequences of amplicons from the T3 RT-PCR of grapefruit budwood graft-inoculated Mexican Lime trees and the 45 CTV reference genomes. The confidence levels are shown as posterior probability values at each node.

Appendix 7

Phylogenetic Neighbor-Joining dendrogram from the 2012 survey performed by Savo Smocilac (unpublished).

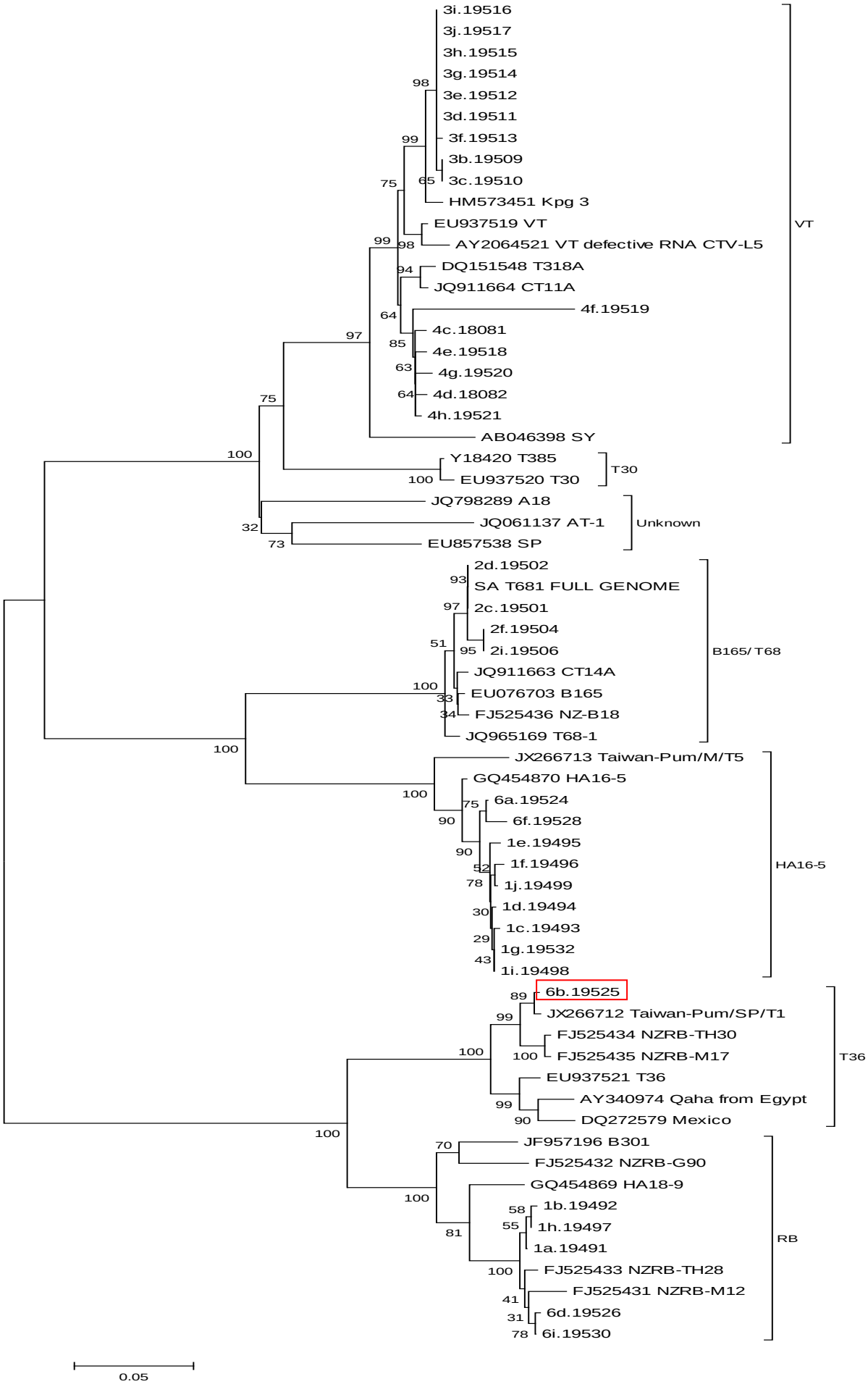


Figure 70: Phylogenetic Neighbor-Joining dendrogram based on the ORF 1a gene region of *Citrus tristeza virus*. Samples 1 through 6 are represented by 1-6 and clones are represented by a-j (Smocilac, *unpublished*). Tree 12-7047 T36 positive clone is represented in the red rectangle. (Sample 1 represents 6 clones: a to i. Sample 2 represents 4 clones: c, d, f and i. Sample 3 represents 9 clones: b to j. Sample 4 represents 6 clones: c, d, e, g, h and f. Sample 6 represents 5 clones: a, b, d f and i)