Figure 1. In order to find nucleotide variability at both ends (3' and 5') of CTV California strain genome (highlighted in gray), the virus RNA has to be converted to a double-stranded DNA, first using a process called reverse transcription (going from RNA to DNA, rather than from DNA to RNA). In the second step, nucleotides (T, C, G, A or poly-A tail) are added to both ends of DNA strands using Rapid Amplification of cDNA Ends (RACE) technique. Addition of nucleotides will provide a basis for sequencing both ends on the two strands (blue and orange lines).

FIGHTING HLB WITH A *CITRUS TRISTEZA VIRUS*-BASED VECTOR

Heterogeneity in the genome ends of CTV is an important consideration

James Ng, Angel Chen and Raymond Yokomi

SUMMARY

As California prepares for a potential showdown with huanglongbing (HLB), contemporary strategies that use low inputs, yet produce high-value control, are needed to manage the disease. With biotechnology, we can develop Citrus tristeza virus (CTV) into a tool for the protection or treatment of citrus trees. Specifically, our goal is to engineer an infectious complementary (c)DNA clone of a mild California strain of CTV that can be developed for the control of HLB (and other disease-causing agents of citrus). Here, we performed a systematic analysis of the nucleotide variability located at the genome ends of California CTV strains with

the T30 and the T36 genotypes – an important consideration in the construction of infectious cDNA clones, given that nucleotides (nucleic acid components) in these locations serve essential biological functions for CTV. We are using knowledge gained from this analysis to construct the cDNAs for each of these CTV strains.

In the future, availability of a biologically active CTV full-length cDNA will provide a template for manipulation of the CTV genome, thereby facilitating the rational development and testing of therapeutic proteins and RNA silencing/interference against HLB-associated 'Candidatus Liberibacter' species and other citrus pests and pathogens.

THE BIOLOGY OF CTV

Citrus tristeza virus (CTV; a member of the virus family *Closteroviridae*) is the causal agent responsible for a number of citrus maladies characterized by mild to severe disease symptoms[1]. Symptomless infection causing no deleterious effects on citrus can also occur, which appears to be typical of the interactions between mild CTV strains and most commercial citrus varieties grown on CTV-resistant and CTV–tolerant rootstocks in California. CTV infects only phloem tissue located many layers beneath the plant surface. Experimental transmission can be achieved by grafting, bark peel inoculation (where a virus preparation is applied to the inner surface of a partially-peeled bark still attached to the stem) and slashing the stem of an uninfected citrus plant with a CTV-contaminated razor blade. The virus also can be transmitted from plant to plant by several different aphid species[2].

DNA technology has enabled the RNA genomes of many viruses to be converted to DNA, which is more stable than RNA and can be engineered for different purposes. For example, Bill Dawson, Ph.D., at the University of Florida has made the 19.3 kilo base (kb) RNA genome **(Figure 1)** of a Florida strain of CTV with the T36 genotype (T36-FL) into an infectious complementary (c)DNA, which made it possible to engineer mutations into CTV genes to study their functions[3,5].

CTV AS A VIRAL VECTOR

A more practical side to generating the infectious cDNA clone of CTV is that it also can be developed as a vector, i.e., a DNA molecule/genetic element with numerous advantages for applications in biotechnology that can benefit citrus production. For example, as with other existing plant virus-based vectors, including the *Tobacco mosaic* virus (TMV) and *Potato virus X* vectors (both of which infect a wide host range, but not citrus), a CTV vector is a self-replicating entity within the infected plant. An advantage of using a CTV-based vector is its stability within the inoculated plants. Existing data indicate that a CTV vector engineered to express the jellyfish green fluorescent protein (GFP) can be stably, but not permanently, maintained in citrus plants for many years[6,7].

Furthermore, because the HLB-associated agent '*Candidatus* Liberibacter asiaticus' (*C*Las) is delivered by its insect vector, the Asian citrus psyllid (ACP), into (and propagates in) the phloem, the coincident cohabitation of CTV and *C*Las in this location makes it ideal to target the latter for control by using an engineered CTV vector. CTV/T36-FL already has been developed and deployed as an RNA interference (RNAi) vector designed to inhibit the ACP[8], as well as for the production of therapeutic proteins both as a curative and a prophylactic against *C*Las and other phloem-residing pathogens.

In light of the emergence of HLB in citrus production areas across the United States, the breakthrough in constructing a functional CTV vector is a remarkable achievement that demonstrates the concerted efforts of the citrus scientific community in dealing with the problems in the wake of this devastating disease.

OBJECTIVES

To use a CTV vector for RNAi-mediated silencing of the ACP or for the expression of therapeutics against *C*Las, the initial step involves development of an infectious cDNA clone of the virus. Because of current EPA, APHIS and CDFA restrictions, use of CTV-FL in California is not being considered. Rather, we are developing infectious cDNA clones derived from CTV strains endemic to California. Our research on constructing the cDNA clone of CTV has focused on using mild strains that are widespread but restricted to California – e.g., California strains of CTV with the T30 genotype or its equivalent.

As with many viruses, CTV exists in a host as a population with one predominant nucleotide (nt) sequence accompanied by a pool of nt variants[9,10]. Interestingly, the frequency of nt variations at the proximal (also called the five-foot) end of the CTV genome is generally higher than those found at the distal (also called the three-foot) end[9]. Because the 5' and 3' ends of viruses are known to play important roles – such as replication, virion assembly and pathogenicity – during viral infection[11], the identity of the nts at these locations are essential considerations for generating the full-length CTV cDNA.

There are three objectives in this study:

1) Construct the cDNA clone of a mild California strain of CTV.
2) Determine the biological activity/infectivity of the cDNA

clone in *Nicotiana benthamiana* plants by *Agrobacterium tumefaciens* agro-mediated inoculation (agro-inoculation). **3)** Prepare CTV from infected *N. benthamiana* plants following the agro-inoculation in Objective 2, and test its infectivity in citrus plants, as well as transmission by aphids. This progress report focuses on Objective 1 and includes most of the work done to determine the variability in the 5' and 3' ends of CTV populations with the T30 (T30-CA strain) and the T36 (T36-CA strain) genotypes.

VARIABILITY IN GENOME ENDS OF CTV POPULATIONS

Replication of the CTV RNA genome produces abundant double-stranded (ds)RNA consisting of a plus (+)-strand and a complementary minus (-)-strand form of RNA **(Figure 1)**. We purified the dsRNAs of T30-CA and T36-CA and converted them to (+)- and (-)-strands DNA by two procedures:

1) reverse transcription, and

2) Rapid Amplification of cDNA-Ends [RACE] **(Figure 1)**.

Subsequently, the DNA products were cloned and sequenced to determine their nt identities. The identities of the most commonly seen nt variants and the frequency with which these variant nts appeared in the cloned DNA are shown in **Figures 2 - 5**.

VARIABILITY AT GENOME ENDS OF T30-CA

The consensus (conserved) RNA sequence at the 5' end of selected T30 sequences retrieved from the NCBI GenBank database is "AAUUUC." Sequencing of the (+)-strand DNA and the (-)-strand DNA revealed that the consensus "AATTTC" (in DNA, Uracil is replaced by Thymine) sequence was present in many of the clones sequenced **(Figure 2)**. In several cloned sequences, one or more extra (variant) nts were observed upstream of this consensus sequence **(Figure 2)**. There were
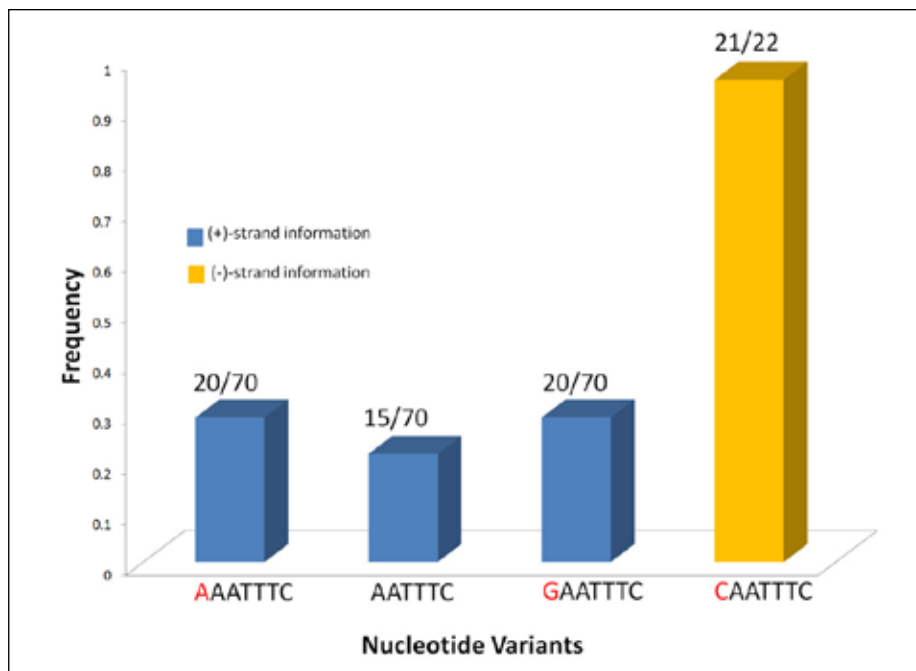

Figure 2. Nucleotide (nt) variations at 5' end of cDNA clones of CTV strain T30-CA. Only nt variants at frequencies greater than 20 percent among all the clones sequenced are shown. Numbers at the top of each bar represent the number of clones with nt sequence shown at the base of that bar over the total number of clones sequenced. Consensus (conserved) or common nucleotides (nts) are shown in black letters, variant nts are in red.
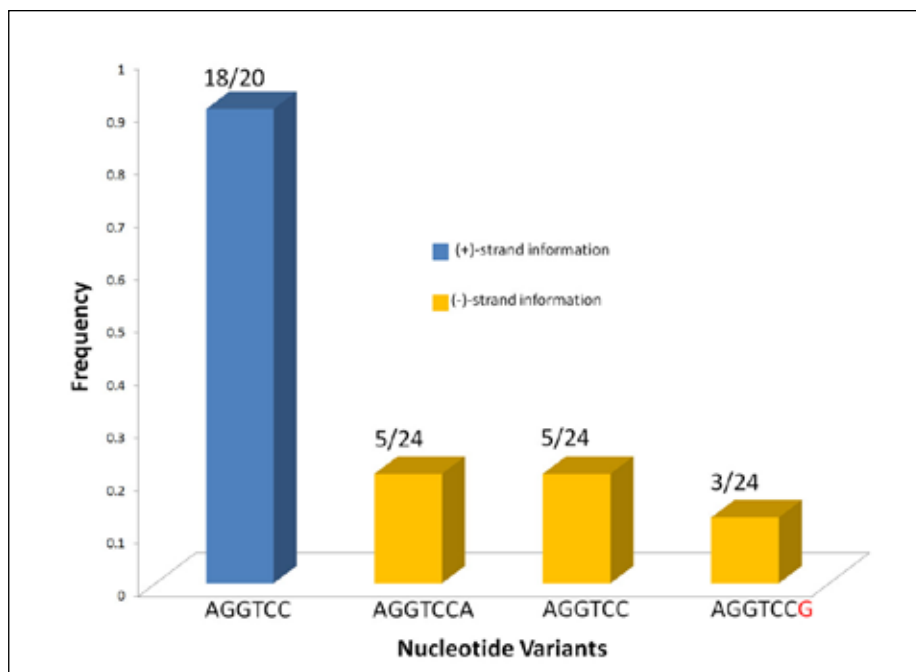

Figure 3. Nucleotide variations at 3' end of cDNA clones of CTV strain T30-CA. Only nt variants at frequencies greater than 10 percent among all the clones sequenced are shown. Numbers at the top of each bar represent the number of clones with the nt sequence shown at the base of that bar over the total number of clones sequenced. Consensus (conserved) or common nts are shown in black letters; variant nts are in red.

many other nt variants (data not shown) but the ones shown in **Figure 2** were the most common.

The 3' end of the T30-CA strain genome contained fewer nt variants than its 5' end. Data obtained from the sequencing of the (+)-strand DNA and the (-)-strand DNA revealed that
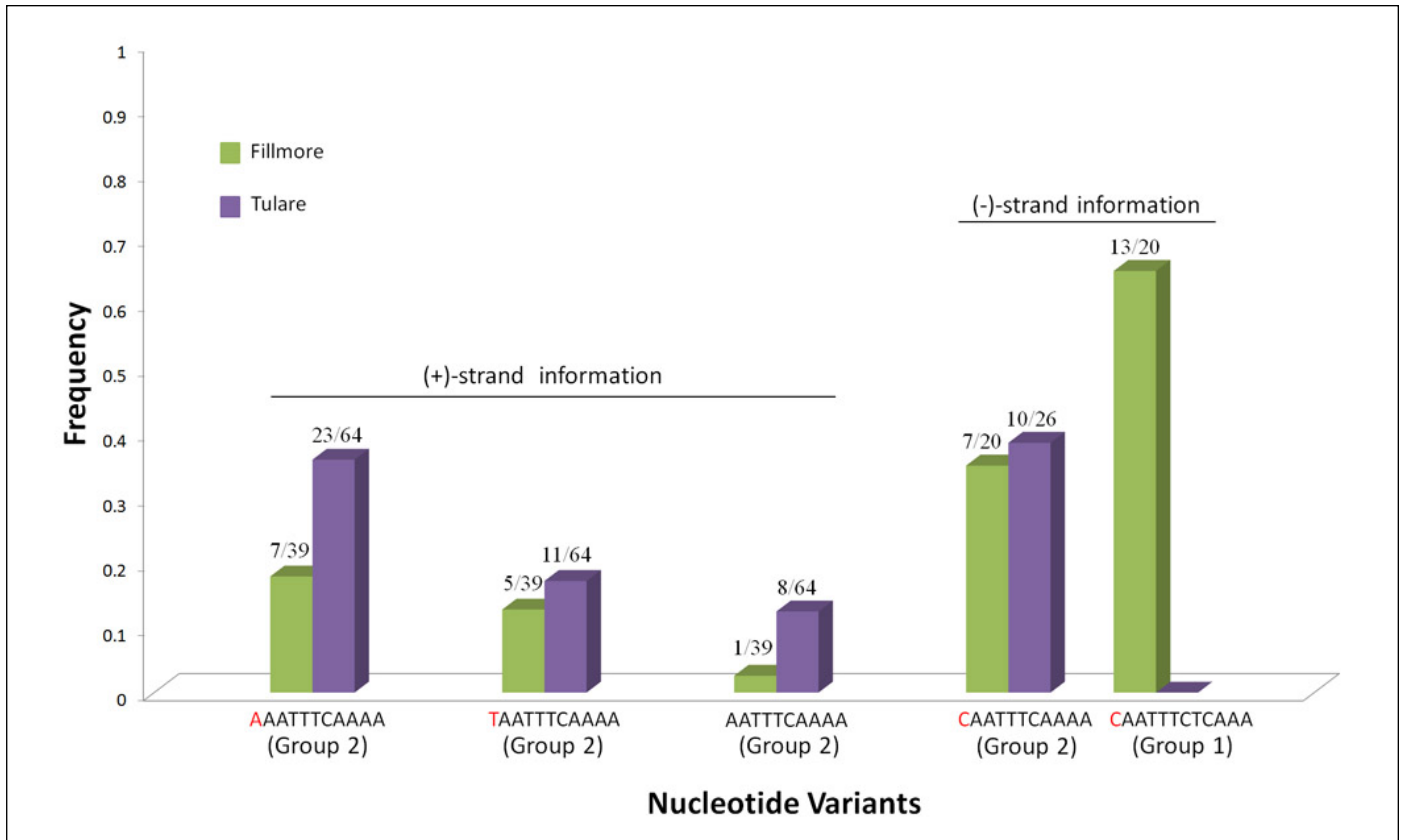
Figure 4. Nucleotide variations at 5' end of cDNA clones of CTV strain T36-CA. Only nt variants at frequencies greater than 10 percent among all the clones sequenced are shown. Numbers at the top of each bar represent the number of clones with the nt sequence shown at the base of that bar over the total number of clones sequenced. Consensus (conserved) or common nts are shown in black letters; variant nts are in red. Nucleotide sequences that are similar to the Group 1 consensus sequence, "AATTTCTCAAA," and the Group 2 consensus sequence, "AATTTCACAAA," are indicated.

"AGGTCC" was the consensus **(Figure 3)**. All of the T30 genotype sequences retrieved from the NCBI GenBank database have highly conserved nts at their genomic 3' ends, and the consensus sequence is "AGGUCCA." Sequencing of the (-)-strand DNA indicated that about 21 percent (five out of 24) of the sequences contained the last "A" while another 21 percent (five out of 24) did not **(Figure 3)**. This suggests that T30-CA strain genome contains a subpopulation with the "AGGUCC" sequence and another subpopulation with the "AGGUCCA" sequence. There were other nt variants, where one or more extra nucleotides were observed after the last "C" in the consensus "AGGTCC" (an example of which is "AGGTCCG"), but these were detected infrequently (at frequencies < 20 percent as shown in **Figure 3**).

These results will be used to create the initial nt sequence to be incorporated in the cDNA clone of a T30 CTV genome from California.

VARIABILITY AT GENOME ENDS OF T36-CA

A comparison of selected T36 genotype sequences retrieved from the NCBI GenBank database indicated that the consensus genomic 5' end sequence can be placed in two groups: group 1 consensus sequence is "AAUUUCUCAAA;" group 2 consensus

sequence is "AAUUUCACAAA.." In this study, T36-CA sequences were determined using infected tissues collected from two California locations – Fillmore County and Tulare County. Sequencing of the (+)- and (-)-strands DNA revealed that the group 1 consensus sequence (i.e. "AATTTCTCAAA") was found in more than 15 nt variants from the Fillmore County sample (only one is shown in **Figure 4**), whereas none was present in the nt variants from the Tulare County sample **(Figure 4)**. The nt variants from both the Fillmore and the Tulare samples have sequences similar (but not identical) to the group 2 consensus sequence (i.e. AATTTCACAAA), except that none of them has the cytosine ("C") at position 8 of the consensus **(Figure 4)**. The "AATTTCAAAA" variant (seen more frequently in the Tulare sample than the Fillmore sample) is most similar to the group 2 consensus sequence **(Figure 4)**. In many of the clones sequenced, one or more extra (variant) nts were observed upstream of this consensus sequence, but variants shown in **Figure 4** were the most common.

As with the T30 genomic sequences, T36 genomic sequences retrieved from the NCBI GenBank database have highly conserved nts at their genomic 3' ends, and the consensus sequence is "AGGUCCA."

In this study, data were consistent with "AGGUCCA" as being the consensus. For example, of the 21 sequences detected in the Fillmore and Tulare samples, nine contained "AGGTCC" as determined by sequencing the (+)-strand DNA **(Figure 5)**.
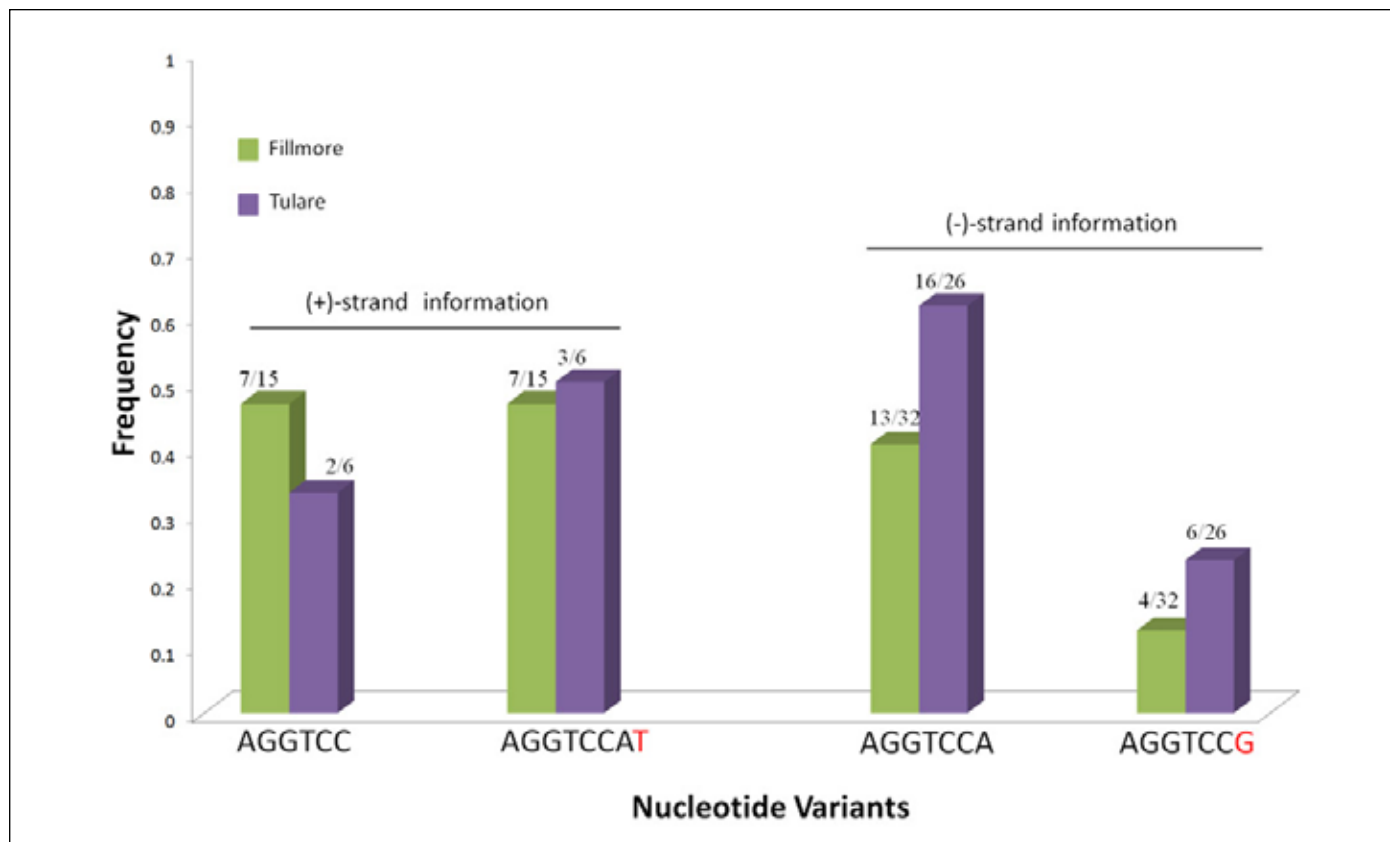
Figure 5. Nucleotide variations at 3' end of cDNA clones of CTV strain T36-CA. Numbers at the top of each bar represent the number of clones with the nt sequence shown at the base of that bar over the total number of clones sequenced. Consensus (conserved) nts are written in black letters; variant nts are in red.

Of the 58 sequences originating from the Fillmore and the Tulare samples, 29 contained the consensus "AGGTCCA" as determined by sequencing the (-)-strand DNA **(Figure 5)**. In a number of cloned sequences, one or more extra nucleotides were observed after the last "A" in the consensus "AGGTCCA" (an example of which is "AGGTCCAT") or after the last "C" in "AGGTCC" (an example of which is "AGGTCCG" **(Figure 5)**.

These results will be used to create the initial nt sequence to be incorporated in the cDNA clone of a T36 CTV genome from California.

Taken together, our analysis of the genomic 5' and 3' ends of the T30-CA and T36-CA viral dsRNA consistently showed predominant consensus sequences and a population of nucleotide variants at the extreme ends. This information is critical to the success of constructing the full-length cDNA clones of both the T30 and the T36 CTV strains from California.

Acknowledgements

James Ng, Ph.D., associate professor of Plant Pathology, and Angel Chen, Ph.D., assistant specialist, are from the Department of Plant Pathology and Microbiology at the University of California, Riverside. Raymond Yokomi, Ph.D., is a plant pathologist at USDA-Agricultural Research Service, Parlier, California.

References

1. Lee, R.F. and Bar-Joseph, M. 2003. Graft-transmissible diseases of citrus. p. 607-639. In: G. Loebenstein and G. Thottappilly (eds.), *Virus and Virus-like Diseases of Major Crops in Developing Countries.* Kluwer Academic Publishers, the Netherlands.

2. Ng, J.C.K. and Zhou, J.S. 2015. Insect vector-plant virus interactions associated with non-circulative, semi-persistent transmission: current perspectives and future challenges. *Current Opinion in Virology* 15:48-55.

3. 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* 24:1119-1131.

4. Tatineni, S., Gowda, S., Boyko, V.P., Albiach-Marti, M.R., Mawassi, M., Navas-Castillo, J., Karasev, A.V., Dolja, V., Hilf, M.E., Lewandowski, D.J., Moreno, P., Bar-Joseph, M., Garnsey, S.M. and W.O. Dawson. 1999. An engineered closterovirus RNA replicon and analysis of heterologous terminal sequences for replication. *Proceedings of the National Academy of Sciences USA* 96:7433-7438.

5. 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* USA 108:17366-17371.

6. El-Mohtar, C. and Dawson, W.O. 2014. Exploring the limits of vector construction based on *Citrus tristeza virus. Virology* 448:274-283.

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

8. Hajeri, S., Killiny, N., El-Mohtar, C., Dawson, W.O. and Gowda, S. 2014. *Citrus tristeza virus*-based RNAi in citrus plants induces gene silencing in *Diaphorina citri*, a phloem-sap sucking insect vector of citrus greening disease (Huanglongbing). *Journal of Biotechnology* 176:42-49.

9. Albiach-Marti, M.R., Mawassi, M., Gowda, S., Tatineni, S., 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. *Journal of Virology* 74(15):6856-6865.

10. Rubio, L., Ayllon, 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.

11. Dawson, W.O., Garnsey, S.M., Tatineni, S., Folimonova, S.Y., Harper, S.J., Gowda, S. 2013. *Citrus tristeza virus*-host interactions. *Frontiers in Microbiology* 4:1-10.

Glossary

Infectious complementary (c)DNA clone – Many plant viruses, including CTV, use RNA (as opposed to DNA) as their genetic material. This RNA can be converted into DNA, which is complementary to the RNA (hence the name "complementary DNA"), and inserted (cloned) into a piece of autonomously replicating DNA (called a cloning plasmid) that exists outside a bacterium's DNA chromosome to produce many copies of the inserted cDNA. When it is introduced into a suitable plant, this cloned cDNA is deemed infectious when it replicates and moves throughout the plant.

RNA silencing/interference – This is an all-natural gene-inactivation system that allows organisms, including plants and animals, to regulate their biological processes, such those involved in development and in defense against viruses. With biotechnology, RNAi can be used to inactivate the genes of virtually any organism that has the RNAi mechanism.