Molecular characterization of closteroviruses infecting Cordyline fruticosa L. in Hawaii

Michael J. Melzer*, Jari S. Sugano, Janice Y. Uchida, Wayne B. Borth, Michael K. Kawate and John S. Hu

Department of Plant and Environmental Protection Sciences, College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu, HI, USA

Keywords: pyrosequencing, Cloroviridae, Velarivirus, Cordyline, ti ringspot

INTRODUCTION
The family Cloroviridae represents a related group of mono- and multipartite, single-stranded, positive-sense RNA plant viruses with long, flexuous virions (Dolja et al., 1994). There are currently three genera in the family which are segregated largely based on vector species. Members of the genera Closterovirus, Crinivirus, and Ampelovirus are, in general, transmitted by aphids, whiteflies, and mealybugs, respectively (Karasev, 2000). Members of the genera Closterovirus and Ampelovirus have monopartite genomes, while members of the genus Crinivirus have multipartite genomes. Little cherry virus 1 (LCV-1) and Grapevine leafroll-associated virus 7 (GLRaV-7) are two members of the family that have not been assigned to a genus (Martelli et al., 2002). Molecular phylogenies and sequence similarity values associate them most closely to members of the genus Crinivirus, however, their monopartite genomes and their lack of a known insect vector have precluded their inclusion into this genus.

The ti plant, Cordyline fruticosa (L.), belongs to the plant family Agavaceae and has considerable cultural and economic importance in Hawaii and throughout most of Polynesia. In Hawaii, the common green variety was introduced by early Polynesian settlers and is a popular ornamental in residential settings that has also become naturalized in Hawaii’s forests. Vegetatively propagated due to sterility (Hinkle, 2007), it is also the most prominent variety grown commercially. In 2009, ti farmers on the island of Maui reported ringspot symptoms on common green ti plants. These ringspot symptoms were subsequently observed on commercial and residential ti plants on the islands of Maui and Hawaii. In search for a causal agent of the ringspot symptoms, it was previously described randomly primed complementary DNA (cDNA) library generated from double-stranded (ds) RNA isolated from a common green ti plant was used as the input material for multiplex pyrosequencing (Melzer et al., 2011). The most common cDNA length was estimated to be approximated 550 bp based on agarose gel electrophoresis. MID7 (5′-AGTGACACT-3′) was ligated to the cDNAs which then underwent pyrosequencing using a 454 GS FLX Titanium platform (Roche, Branford, CT, USA) at the University of Hawaii’s Advanced Studies in Genomics, Proteomics and Bioinformatics (ASGBP) laboratory. Following preliminary work has indicated that, aside from CoV-1, three additional closteroviruses may infect common green ti plants in Hawaii. In this study, pyrosequencing was used to characterize the genomes of closteroviruses infecting a single common green ti plant. The sequence data confirmed the presence of CoV-1 as well as three additional closteroviruses. Although all four viruses had the same general genome organization, the sequence divergence between the RNA-dependent RNA polymerase, heat shock protein 70 homolog, and coat protein ranged from 22 to 61%, indicating these represent four distinct closterovirus species. The names CoV-2, CoV-3, and CoV-4 are proposed for the three new viruses. Phylogenetic analyses placed CoV-2, CoV-3, and CoV-4 in the same clade as CoV-1, LChV-1, and GLRaV-7.

MATERIALS AND METHODS

LIBRARY SEQUENCING AND ASSEMBLY
A previously described randomly primed complementary DNA (cDNA) library generated from double-stranded (ds) RNA isolated from a common green ti plant was used as the input material for multiplex pyrosequencing (Melzer et al., 2011). The most common cDNA length was estimated to be approximated 550 bp based on agarose gel electrophoresis. MID7 (5′-AGTGACACT-3′) was ligated to the cDNAs which then underwent pyrosequencing using a 454 GS FLX Titanium platform (Roche, Branford, CT, USA) at the University of Hawaii’s Advanced Studies in Genomics, Proteomics and Bioinformatics (ASGBP) laboratory. Following
pyrosequencing, the MID7 and random primer sequences as well as low quality basecalls at the end of reads were trimmed. Short length (<65 nt) and low quality reads as well as those that mapped to the CoV-1 genome were purged from the dataset. The remaining reads underwent de novo assembly using Geneious® Pro 5.6.5 (Biomatters Ltd., Auckland, New Zealand). To validate low coverage regions or where unexpected stop codons or frameshifts occurred, and to bridge sequence gaps between contiguous sequences (contigs), primers were designed flanking the region in question and PCR was performed using the cDNA library as template. PCR products were either directly sequenced following treatment with ExoAP-T® (USB/Affymetrix, Santa Clara, CA, USA) or ligated into pGEM®-T Easy (Promega, Madison, WI, USA) followed by Sanger-based sequencing at the ASGPB.

The 3′-terminal sequences were obtained by polyadenylating heat-denatured dsRNA using yeast poly(A) polymerase following the manufacturer’s instructions. An oligo d(T) primer (5′-CAGCTGCTATATTCCAGG(T)3′) was used to initiate cDNA synthesis and also used in the subsequent PCR reaction along with a virus-specific primer designed to anneal near the 3′-end of the available virus sequence. Amplification products were cloned using pGEM®-T Easy and at least five clones underwent Sanger-based sequencing at the ASGPB.

PHYLOGENETIC ANALYSES
The combined helicase domain (HEL), RNA-dependent RNA polymerase (RdRp), heat shock protein 70 homolog (HSP70h), p61 (FP03235), and coat protein (CP) amino acid sequences of members and tentative members of the family Closteroviridae were aligned using ClustalX 2.0.12 (Larkin et al., 2007). With this alignment, the phylogenetic relationships of the sequences were inferred using neighbor-joining (NJ) and maximum likelihood (ML) algorithms. The NJ algorithm was performed using ClustalX 2.0.12 and bootstrapped with 1000 replications. The ML algorithm was performed with PhyML 3.0 (Guindon et al., 2010) using the WAG model and bootstrapped with 1000 replications.

RESULTS
LIBRARY SEQUENCING AND ASSEMBLY
A total of 107,655 high quality reads > 64 nt were generated from the cDNA library, with maximum, minimum, and mean lengths of 772, 65, and 392.1 nt, respectively. Of those reads, 4,424 mapped to the CoV-1 genome. The majority of the remaining reads assembled into three contigs, each in excess of 10 kb in length (Table 1). Based on comparisons with accessions in GenBank, all three contigs represented closterovirus-based genomes. The first contig was found to be essentially identical to the previously identified Contig5 (Melzer et al., 2011) where the two sequences overlapped, and was thus designated Contig5. Similarly, the second and third contigs were found to be essentially identical in overlapping regions with Contig5 and CloneH11, respectively, and were designated as such. The 15,031 nt Contig5 was extended to 15,107 nt following the addition of the 3′-terminal sequence. The 14,941 nt Contig8 was extended to 16,274 nt following the addition of the 3′-terminal sequence and a contig in the 5′-region of the genome. The 10,684 nt CloneH11 was extended to 14,620 nt with the addition of the 3′-terminal sequence and a contig in the 5′-region of the genome. The 5′-terminal sequences were not obtained for any of the contigs.

| Table 1 | Summary of pyrosequencing results and assembly of reads into contiguous sequences (contigs). |
| --- | --- |
| # of reads (% of total) | Length of contig | Mean coverage |
| CoV-1 | 4,424 (4.1) | n/a | n/a |
| CoV-2 (Contig8) | 11,395 (10.8) | 15,031 | 342.8 |
| CoV-3 (Contig8) | 79,593 (73.9) | 14,941 | 2410.8 |
| CoV-4 (CloneH11) | 3,537 (3.2) | 10,684 | 186.7 |
| Uncorporated | 6,728 (6.1) | n/a | n/a |

1 Post-editing.
CloneH11 were identical to that of CoV-1, having the sequence 5′-AAAGGUGCG-3′. Contig8 also ended with this sequence, but appeared to lack the terminal guanine residue.

**DIVERSITY AND PHYLOGENETIC PLACEMENT**

The amino acid identity between Contig5, Contig8, CloneH11, and CoV-1 for their respective RdRp, HSP70h, and CP sequences was under 70% in all cases except for the RdRp sequences of Contig5 and CoV-1 which were 78% identical (Table 2). Using the current criteria for closterovirus species demarcation recently revised by the International Committee on Taxonomy of Viruses (Martelli et al., 2011), these contigs would represent distinct closterovirus species. As such, Contig5, Contig8, and CloneH11 were tentatively designated CoV-2, CoV-3, and CoV-4, respectively.

Cordyline virus 1 and CoV-2 appear to be the most closely related of the CoVs characterized in this study with an overall nucleotide identity of 63.7%. Toward the 3′-termini of their genomes, however, the similarity gradually increased. This similarity peaked in the 3′UTR of CoV-1 and CoV-2 which shared a 90.8% nucleotide identity.

Over 8% of the total sequence reads did not map to the genomes of CoV-1, CoV-2, CoV-3, or CoV-4 (Table 1). The majority of these appeared to be either of plant or prokaryotic origin, or did not share significant similarity to any of the sequence accessions in the current databases. Approximately 11% of these reads, however, represented closterovirus sequences. One thousand and ninety-two of these reads were selected for further investigation. One hundred and seventy-four of these reads, when translated to amino acid sequence, were similar to the N-terminal region of a closterovirus ORF1a. Reverse-transcription PCR revealed this region was part of CoV-4. The remaining reads, when translated to amino acid sequence, had high similarity (between 80 and 92% identity) to proteins encoded by CoV-1.

**Phylogenetic analyses of the combined HEL domain, RdRp, HSP70h, CP, and p60/61 amino acid sequences using distance-based (NJ) and character-based (ML) algorithms inferred almost identical relationships between CoV-2, CoV-3, and CoV-4.**

**Table 2 | RNA-dependent RNA polymerase/heat shock protein 70 homolog/coat protein amino acid percent identities between the CoVs infecting common green ti plants.**

| CoV-1 | CoV-2 (Contig5) | CoV-3 (Contig8) | CoV-4 (CloneH11) |
|---|---|---|---|
| CoV-2 (Contig5) | 78/68/67 | 78/68/67 | 78/68/67 |
| CoV-3 (Contig8) | 62/50/44 | 62/50/44 | 62/50/44 |
| CoV-4 (CloneH11) | 59/53/39 | 57/51/40 | 55/55/39 |
DISCUSSION

We have previously reported the presence of one closterovirus, CoV-1, infecting common green ti plants in Hawaii, and provided preliminary evidence for the presence of additional closterovirus species using a Sanger-based sequencing approach (Melzer et al., 2011). In this study we used a massively parallel sequencing approach to identify, in addition to CoV-1, three new closterovirus species which we have tentatively designated CoV-2, CoV-3, and CoV-4.

Cordyline virus 2 and CoV-3 share the same general genome organization as CoV-1, and differ from each other only in the molecular weight of their small hydrophobic proteins and the small protein encoded by the ORF located between their HSP70h and p61 ORFs. CoV-4, however, is unusual by lacking an ORF encoding a small transmembrane protein between the ORFs of the RdRp and HSP70h. The small transmembrane protein of Beet yellow virus (BYV) associates with the host endoplasmic reticulum and is involved in the cell-to-cell movement of BYV and presumably other closteroviruses as well (Peremyslov et al., 2004). An ORF which could encode a small protein possessing a transmembrane domain does exist in CoV-4, although it is located within the HSP70h ORF in a +1 reading frame relative to the HSP70h ORF. If this ORF is not expressed, it is possible that CoV-4 requires co-infection with another closterovirus for cell-to-cell movement. Since all four CoVs were present in a single host plant, multiple infections in a single host plant may not be uncommon.

Overall, the genomes of CoV-1 and CoV-2 had a moderate sequence similarity. Near the 3′-terminus, however, this similarity gradually increased to the point where the 3′ UTRs of these viruses were nearly identical, indicative of a potential recombination.
event. Putative examples of closterovirus recombination are abundant (Cuellar et al., 2008; Melzer et al., 2010; Farooq et al., 2013). The gradual increase in genome similarity is comparable to that proposed for Citrus tristeza virus (CTV) strain T36, and suggests the potential recombination event was not recent (Mawassi et al., 1996). The presence of multiple CoVs in a single host plant would provide an environment conducive for such recombination events.

The family Clustoroviridae Subcommittee to the International Committee on Taxonomy of Viruses (ICTV) has recently amended a commonly used criterion for species demarcation of closteroviruses. To be considered a distinct species, the level of sequence divergence in a phylogenetically informative protein (RdRp, HSP70h, or CP) was raised from 10 to 25% (Martelli et al., 2011). This increase in stringency was undertaken to address the proliferation of closteroviruses that had a similar genome organization, host range, and biological properties but, in some cases, only marginally exceeded the previous 10% sequence divergence criterion, thereby elevating them to species status (Martelli et al., 2012). Following the implementation of this more stringent criterion, a group of seven GLRaVs species and their “variants” (GLRaV-4, GLRaV-5, GLRaV-6, GLRaV-6DE, GLRaV-9, GLRaV-Gar, and GLRaV-P) were condensed into a single species, GLRaV-4 (Martelli et al., 2012). Based on the amino acid identities of the RdRp, HSP70h, and CP sequences it is clear that CoV-1, CoV-3, and CoV-4 are distinct species under the new criterion. The two most closely related CoVs, CoV-1 and CoV-2, have amino acid divergence values for the RdRp, HSP70h, and CP of 22, 31, and 33%, respectively. Although the sequence divergence between the CoV-1 and CoV-2 RdRp does not exceed the 25% threshold, the average sequence divergence for these three phylogenetically informative proteins is 29%, and we therefore contend that CoV-1 and CoV-2 should represent two distinct species. Additional closterovirus-like sequences were also identified in the library. The majority of these, when translated into amino acid sequences, were only 10–20% divergent from CoV-1 and are likely to have come from a second strain of CoV-1 that also infects common green ti.

The discovery of four related closterovirus species co-infecting the same host which share a similar genome organization and perhaps similar biological properties presents a situation reminiscent to the GLRaV-4 group. Since these four CoVs were discovered through the intense study of a single ti plant, it is also plausible that additional CoV species exist. Some members of the GLRaV-4 group, however, were found to be serologically related (Ghanem-Sabanadzovic et al., 2012). There are currently no antiserum raised against any of the CoVs that would allow experiments to be conducted to determine their serological relationships. Given the amount of sequence divergence between the currently known CoVs, particularly in their respective structural proteins, it seems unlikely that they will be serologically related.

Within the family Clustoroviridae, LCV-C, GLRaV-7, and CoV-1 form a distinct monophyletic clade for which the genus "Velarivirus" has been proposed (Al Rwahnih et al., 2012; Martelli et al., 2012). Phylogenetic analyses placed CoV-2, CoV-3, and CoV-4 within this clade. We therefore propose that CoV-2, CoV-3, and CoV-4 also be members of the proposed genus "Velarivirus", should it be ratified by the ICTV.

ACKNOWLEDGMENTS

The authors would like to thank Julie Coughlin, Dr. Shaobin Hou, Kenneth Hu, Chris Kadosaka, James Kam, and Christina Liu for their assistance with this project. This project was supported, in part, by Award 595955 from the Hawaii Department of Agriculture.

REFERENCES

Al Rwahnih, M., Dolja, V. V., Daubert, S., Koontz, E. V., and Rundh, A. (2012). Genomic and biological analysis of grapevine leafroll-associated virus 7 reveals a possible new genus within the family Clustoroviridae Virus Res. 163, 592–599.

Cuellar, W. J., Taipo, F., Kreuze, J. F., and Vest, K. E. (2005). Analysis of genomic content in sweet potato chlorotic stunt virus RNA1 reveals the presence of the p22 RNA silencing suppressor in only a few isolates: implication for viral evolution and pathogenicity. J. Gen. Virol. 86, 573–582.

Dolja, V. V., Kasera, A. V., and Kurapov E. V. (1994). Molecular biology and evolution of closteroviruses: sophisticated build-up of large RNA genomes. Ann. Rev. Phytopathol. 32, 262–287.

Fausing, A. B. R., Ma, Y. X., Wang, Z., Zhou, N., Wenig, X., Wang, G. P., et al. (2013). Genomic diversity analysis reveals novel recombination events in Grapevine leafroll-associated virus 3 in China. Virus Res. 171, 15–21.

Ghanem-Sabanadzovic, N. A., Sabanadzovic, S., Gagarin, P., and Rowhani, A. (2012). Genome organization, serology and phylogeny of Grapevine leafroll-associated viruses 4 and 6: taxonomic implications. Virus Res. 163, 120–128.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Hinkle, A. E. (2007). Population structure of Pacific Codline fly (Laminauminaceae) with implications for human settlement of Polynesia. Am. J. Hum. 82, 828–839.

Jilka, F., Faucherre, R., and Agra- novsky, A. A. (1977). Complete genome structure and phylogenetic analysis of little cherry virus, a mealybug-transmissible closterovirus. J. Gen. Virol. 79, 2007–2017.

Kasera, A. V. (2000). Genetic diversity and evolution of closteroviruses. Annu. Rev. Phytopathol. 38, 293–324.

Koontz, E. V. (1991). The phylogeny of RNA1-dependent RNA polymerase of positive-strand RNA viruses. J. Gen. Virol. 72, 2197–2206.

Kulkin, M. A., Blackshields, G., Brown, N. P., Chetan, R., McGregor, P. A., McWilliam, H., et al. (2004). Clustal W and Clustal X version 2.0: Bioinformatics 23, 2947–2948.

Martelli, G. P., Agranovsky, A. A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R. H. A., et al. (2002). The family Clustoroviridae revised. Arch. Virol. 147, 209–2064.

Martelli, G. P., Agranovsky, A. A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R. H. A., et al. (2011). "Family Clustoroviridae" in Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. vi 30 A. King, M. J. Adams, E. R. Carreira, and E. Lefkowitz (Amsterdam Elsevier Academic Press), 907–1062.

Mawassi, M., Mlikotowska, E., Gofman, R., Yang, G., and Bar-Joseph, M. (1998). Unusual sequence relationships between two isolates of Citrus tristeza virus. J. Gen. Virol. 77, 2359–2364.
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.

Melzer, M. J., Sether, D. M., Borth, W. B., Mersino, E. F., and Hu, J. S. (2011). An assemblage of closteroviruses infecting Hawaiian ti (Cordyline fruticosa L.), Virus Genes 42, 254–260.

Peremyslov, V. V., Pan, Y. W., and Dolja, V. V. (2004). Movement protein of a closterovirus is a type III integral transmembrane protein localized to the endoplasmic reticulum. J. Virol. 78, 3704–3709.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 December 2012; paper pending publication: 27 January 2013; accepted: 05 March 2013.

Citation: Melzer MJ, Sugano JS, Uchida JY, Borth WB, Kawate MK and Hu JS (2013) Molecular characterization of closteroviruses infecting Cordyline fruticosa L. in Hawaii. Front. Microbiol. 4:39. doi: 10.3389/fmicb.2013.00039

This article was submitted to Frontiers in Virology, a specialty of Frontiers in Microbiology.

Copyright © 2013 Melzer, Sugano, Uchida, Borth, Kawate and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.