Australasian Plant Pathology (2025) 54:131–135
https://doi.org/10.1007/s13313-025-01023-8

RESEARCH NOTE

The complete genome sequence of a novel legume closterovirus from Norfolk Island

John E. Thomas[1] · Mark Jackson[1] · Paul R. Campbell[2] · Nga T. Tran[1] · Andrew D.W. Geering[1]

Received: 7 November 2024 / Accepted: 10 February 2025 / Published online: 7 March 2025

Abstract

A novel species in the genus *Closterovirus*, family *Closteroviridae*, has been identified by high throughput sequencing in legume samples collected during plant pathogen surveys on Norfolk Island in 2014. The complete genome sequence of 16,815 nucleotides was obtained from a French bean (*Phaseolus vulgaris*) plant with symptoms of interveinal chlorosis in the older leaves. Complete or near complete coding sequences for the 10 open reading frames were also obtained from a second French bean plant, and batch samples of white clover (*Trifolium repens*) and pea (*Pisum sativum*). All isolates shared 99.7–99.9% nucleotide identity, indicating they are members of the same virus species.

Keywords Closterovirus · Closteroviridae · Survey, phaseolus vulgaris · Legume · Fabaceae

Norfolk Island is an isolated, sub-tropical island and is an external non-self-governing Australian territory. It is located in the Western Pacific Ocean, roughly 1400 km east of the Australian mainland, with New Caledonia 800 km to the north and New Zealand 700 km to the south. The island is largely self-sufficient in fresh fruit and vegetable production. General quarantine surveys, covering plant and animal pathogens, invertebrates and introduced plants, were conducted in 2012–2014 (Maynard et al. 2018). During these surveys, a wide range of cultivated and non-cultivated plants were sampled to test for the presence of plant viruses and archived for later analysis.

The leaf samples were initially desiccated over silica gel, and then for longer term storage they were lyophilised and stored under vacuum at −20 °C in the Queensland Department of Primary Industries Plant Virus Collection. Two individual French bean (*Phaseolus vulgaris*) plants from the same planting (isolates Q6302 and Q6303) displaying strong interveinal chlorosis of the lower leaves (Fig. 1) were selected for high throughput sequencing (HTS). Additionally, batch samples of white clover plants (*Trifolium repens*) and pea (*Pisum sativum*) were examined. The first batch (#1) contained leaves from thirteen white clover plants, three of which had leaf mosaic symptoms (isolates Q4974, Q6013, Q6142), and the remainder were asymptomatic (Q6300). The second batch (#2) contained leaves from four white clover plants, three of which had leaf mosaic symptoms (Q4847, Q4920, Q4938) and the final (Q4915), leaf rugosity. The last batch (#3) of leaves from five pea plants all had vein chlorosis and flecking symptoms (Q6356, Q6287, Q6279, Q6318, Q6320).

Total nucleic acids were extracted from the lyophilised tissue samples using the Trizol Plus RNA Purification System (Invitrogen cat. no. 12183555), diluted in nuclease-free water and stored at −80 °C until use. HTS was outsourced to the Australian Genome Research Facility (AGRF). Library preparation was done with the Illumina TruSeq® Stranded Total RNA Library Preparation Kit with Ribo-ZeroTM Plant (Illumina, San Diego CA, USA) with 500 ng of input RNA followed by sequencing of 150 bp pair-ended reads using a 300 cycle kit on an Illumina NovaSeq X sequencer. About 36–44 million reads were obtained per sample. Initial *de novo* contig assembly was conducted with CLC Workbench V12.0 using automatic bubble and word size and a 600-nt contig cut-off. The *de novo* assembly was repeated with 1% randomly sampled reads to yield longer contigs. Subsequent analysis was done using Geneious Prime version 2023.2.1 (Biomatters Ltd., Auckland). The contigs were annotated by doing a BLASTN search of a custom plant viral sequence

✉ John E. Thomas
j.thomas2@uq.edu.au

[1] The University of Queensland, QAAFI, Brisbane, Australia

[2] Queensland Department of Primary Industries, Brisbane, Australia

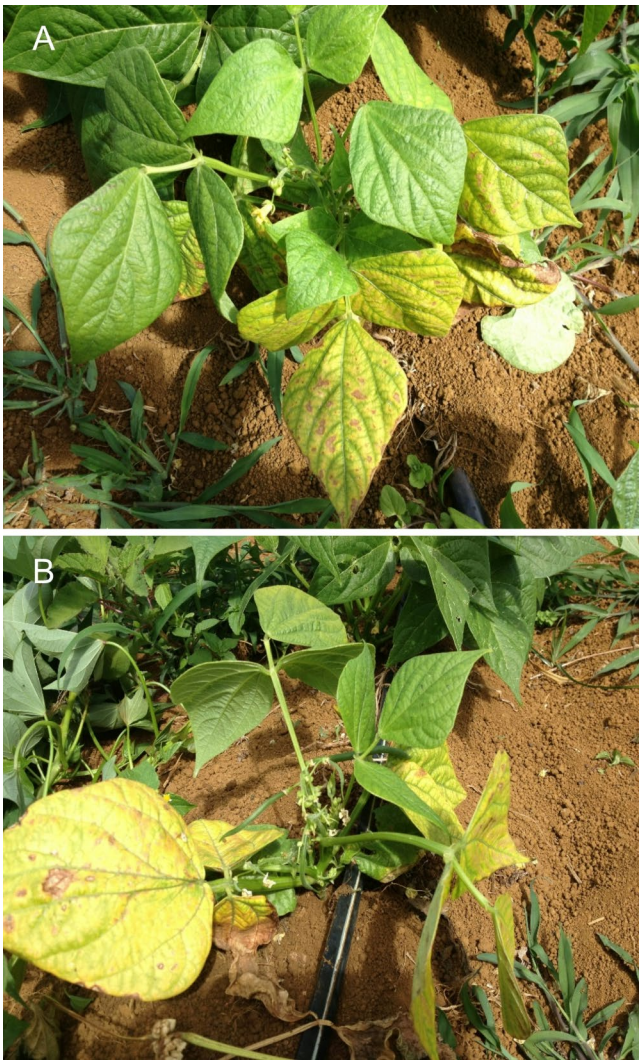

Fig. 1 Symptoms of lower leaf chlorosis in French bean (*Phaseolus vulgaris)* in plants infected with closterovirus isolates Q6302 (**A**) and Q6303 (**B**)

database. A contig from isolate Q6302 had a significant match to beet yellows virus (genus *Closterovirus*) sequence accession AF190581.1, with 62.3% pairwise nucleotide (nt) identity across ~3 kbp of aligned sequence. The sequence reads were then re-mapped to the contig using Geneious with the following parameter settings: disallow gaps, mismatch per read: 10%, minimum overlap: 25, and disallow word number filtering. The average read depth was 19,300, but for the 2 kbp of sequence at the 3′ end of the contig, the average read depth was 46,500. Other samples with contigs matching closteroviruses were mapped against the contig from sample Q6302.

To obtain the complete genome sequence of the exemplar virus isolate from French bean (Q6302), 5′ and 3′ rapid amplification of cDNA ends (RACE) was performed. For 5′ RACE, first strand cDNA was generated using Superscript III reverse transcriptase (cat. no. 18080093, ThermoFisher Scientific) and virus-specific oligo primer M96 (5′ AGAGA GCTACGCTCATAACC 3′). The cDNA was then purified using a Monarch DNA cleanup kit (cat. no. T1030L, New England Biolabs) and C-tailed using terminal transferase (cat. no. M0315S, New England Biolabs). Amplification of the tailed cDNA products was performed using Phusion High-Fidelity DNA polymerase (cat. no. M0530S, New England Biolabs) in a PCR containing the foward primer Tag1$(G)_7$HN (5′ CCACGCGTATCGATGTCGAC$(G)_7$HN 3′) and the virus-specific reverse primer M97 (5′ ATCTTC CAAGGATGTGTCGG 3′). A second PCR using the Tag1 primer (5′ CCACGCGTATCGATGTCGAC 3′) and the nested virus specific primer M98 (5′ GATGGCGTCCATA CTTCTC 3′) produced 5′ RACE amplicons that were gel-extracted and cloned into the the PCR 2.1 vector (cat. no. K202020, Invitrogen), enabling sequencing using M13F and M13R primers.

For 3′ RACE, total RNA was first labelled using either *E. coli* Poly(A) or Poly(U) Polymerase to introduce 3′ A and U tails, respectively. First strand cDNA was then generated using Superscript III reverse transcriptase and primers Tag1-oligo(dT) (5′ CCACGCGTATCGATGTCGA $C(T)_{40}$VN 3′) or Tag1-oligo(dA) (5′ CCACGCGTATCGAT GTCGAC$(A)_{40}$BN 3′). Amplification and sequence analysis of tailed cDNA products was esentially performed as per 5′ RACE with initial PCR using the virus-specific forward primer M99 (5′ TCGACTCGCGACATCTGC 3′), with

Table 1 Position, size and predicted functions of putative ORFs in closterovirus isolate Q6302

ORF	Nucleotide position	Amino acids	size (kDa)	Translation frame	Predicted function[a]
1 A	166–8,007	2613	294.65	1	PRO/MTR/HEL
1B	8,060–9, 394	444	51.47	2	RdRp
2	9,484–10,485	333	37.76	1	no matches
3	10,482–10,676	64	7.41	3	movement protein?
4	10,677–12,479	600	65.47	3	HSP70
5	12,385–14,067	560	64.89	1	HSP90
6	14,039–14,737	232	26.62	2	minor CP
7	14,800–15,420	206	22.48	1	major CP
8	15,329–15,994	221	25.23	2	no matches
9	16,042–16,626	194	22.90	1	RNA silencing suppressor?

[a] PRO, protease; MTR, methyltransferase; HEL, helicase; RdRp, RNA-dependent RNA polymerase; HSP70, heat shock protein 70 analogue; HSP90, heatshock protein 90 analogue; CP, coat protein

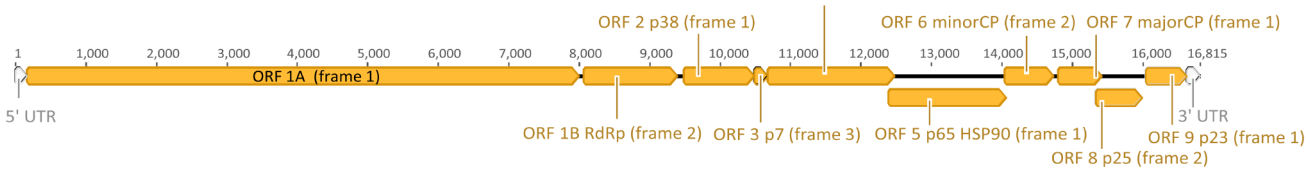

Fig. 2 Predicted genome organization of closterovirus isolate Q6302

subsequent nested PCR using virus-specific primer M101 (5′ CGTTACAGAGTAACGAGACG 3′).

From French bean isolate Q6302, a contig of 16,841 nt was initially obtained. After 3′ and 5′ RACE, the genome length was corrected to 16,815 nt. The viral genome contained ten putative ORFs, determined by alignments with recognised closteroviruses (Table 1), and a 189-nt 3′ untranslated region (UTR) and a 165-nt 5′ UTR (Fig. 2). The genome organization resembles that of citrus tristeza virus (Folimonova 2020), with a potential movement protein (ORF 2, p38) following ORF 1B (RNA-dependent RNA polymerase; RdRp).

ORF 1 A contains an N-terminal papain-like endopeptidase domain (aa 552–635), an RNA methyltransferase domain (aa 693–1024) and at the C-terminal end a viral (Superfamily 1) RNA helicase (aa 2237–2498). The recently recognised "Zemlya region" (Gushchin et al. 2017), thought to be involved in the formation of closterovirus replication platforms, is also present in the central region of the protein (aa 1353–1454). ORF 1B encodes an RNA dependent RNA polymerase, ORF 3 a putative movement protein, ORF 4 an

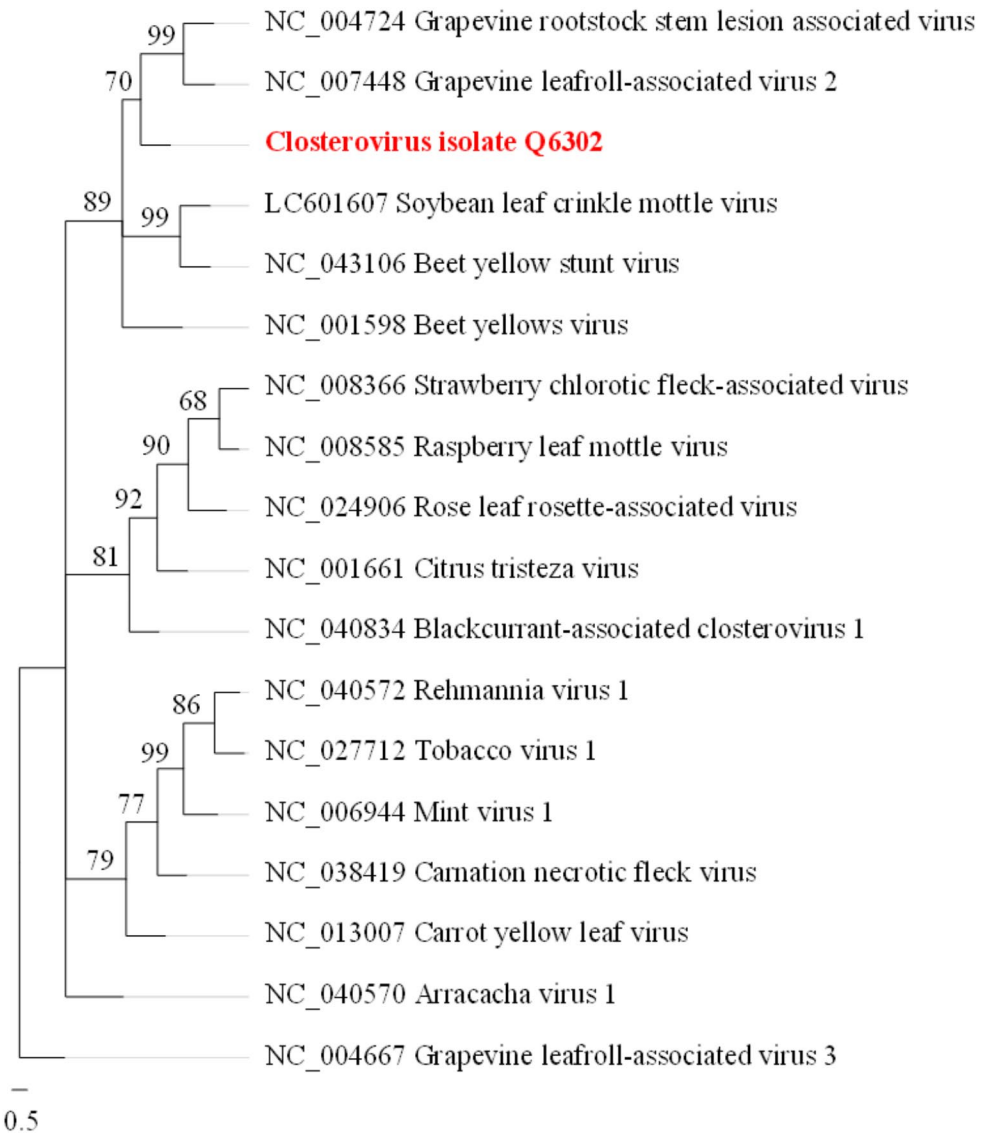

Fig. 3 Maximum-likelihood phylogenetic tree showing the relationships between closterovirus isolate Q6302 and members of the genus *Closterovirus* based on amino acid sequence alignment of the HSP70 region. Grapevine leafroll-associated virus 3 (genus *Ampelovirus*) is used as an outgroup. Values at the nodes denote consensus bootstrap support with values less than 50% collapsed into polytomies. Tip labels include RefSeq numbers or GenBank accessions, followed by virus names. The scale bar represents numbers of amino acid substitutions per position in the alignment

Table 2 Nearest BLASTP matches for putative protein products of closterovirus isolate Q6302

Protein	Nearest GenBank match		% amino acid identity
	Virus	GenBank accession	
1 A	Grapevine leafroll associated virus −2	ON645912.1	35.2
1B	Beet yellows virus	ON738345.1	66.3
2	No matches	-	-
3	No matches	-	-
4 S	Beet yellows virus	ON738341.1	54.4
4 L	Beet yellows virus	MT815988.1	54.3
5	Soybean leaf crinkle mottle virus	LC601607.1	43.1
6	Grapevine leafroll associated virus −2	MW715829.1	41.1
7	Grapevine leafroll associated virus −2	MT899927.1	36.3
8	No matches	-	-
9	No matches	-	-

HSP70 homologue and ORF 5 an HSP90 homologue. ORFs 6 and 7 encode the minor and major coat proteins, respectively, and ORF 9 a putative suppressor of RNA silencing. The possible functions of the proteins encoded by ORFs 2 and 8 are not known. Interestingly, ORFs IA, 1B, 3, 5 and 6 have TTG start codons, uncommon in Eukaryota (Touriol et al. 2003) and in contrast to all other published closterovirus reference isolate coding sequences.

A phylogenetic analysis was performed in Geneious Prime version 2024.0.3 (Biomatters Ltd., Auckland) using the conserved HSP70 protein, a distinguishing taxonomic feature of the *Closteroviridae* (Agranovsky 2016). The sequences were aligned using MUSCLE and a maximum likelihood (ML) tree generated using the RA×ML Geneious Plugin version 8, selecting the GAMMA JTT model of protein evolution, rapid bootstrapping of 1,000 replicates, and starting with a complete random tree topology. The ML bootstrapping trees were then used to generate a consensus tree, as illustrated in Fig. 3. Closterovirus isolate Q6302 formed a well-supported clade with grapevine leafroll-associated virus 2 and grapevine rootstock stem lesion associated virus as sister taxa.

French bean isolate Q6303 and leaf batches #1 and #2 from white clover also contained contigs > 16,800 nt, which contained all ten putative ORFs that were identified in Q6302. The contig obtained from pea was truncated at the 3′ end on the genome and was 16,216-nt long, lacking part of ORF 9 and the 3′ UTR. These complete or near complete coding sequences had 98.8–99.8% nt identity to the genome sequence from isolate Q6302. The novel closterovirus was distinct from other reported closteroviruses across all putative protein products (Table 2). It was phylogenetically distinct in the HSP70 homologue (Fig. 3), in the coat protein (aa) and across the complete genome (nt) (data not shown) with a similar pattern of clustering in each case. The aa sequences of all gene products differ from those of other recognised closteroviruses by more than 25%, thus satisfying the current ICTV species demarcation guidelines (Fuchs et al. 2020). All new sequences have been deposited in GenBank under accessions PQ661949 (batch #3), PQ661950 (batch #2), PQ661951 (batch #1), PQ661952 (isolate Q6302) and PQ661953 (isolate Q6303).

The presence of pea seedborne mosaic virus in pea, and white clover mosaic virus and clover yellow mosaic virus in white clover (J. E. Thomas et al., unpublished results) precludes identifying symptoms caused by the closterovirus in these hosts. Furthermore, individual plants in these leaf batches were not tested to determine which plants were infected. However, both French bean plants displayed the same symptoms of interveinal chlorosis in the lower leaves. No additional viruses were detected in Q6302 and only phaseolus vulgaris alphaendornavirus 1, not known to cause symptoms (Mrkvová et al. 2023), was additionally detected in Q6303. Thus, we propose the common name ‘bean yellows virus’ for members of this novel closterovirus species. This appears to be the first record of this closterovirus anywhere in the world, and our results suggest it may be widespread in legumes on Norfolk Island. We hypothesize that the virus has been introduced onto the island since European settlement, and it will likely be present elsewhere in the world.

Acknowledgements We thank the Norfolk Island Government for the provision of resources on Norfolk Island and local producers for their generous cooperation and access to their crops.

Author contributions A.D.W.G. and J.E.T. contributed to the study conception and design, field surveys and data collection. Material preparation, and analysis were performed by N.T.T., M.J., P.R.C., J.E.T. and A.D.W.G. The first draft of the manuscript was written by J.E.T. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions
This work was supported in part by funds from the federal Department of Infrastructure and Regional Development to A.D.W.G. We would also like to acknowledge the contribution of the Plant Pathogen “Omics” consortium in the generation of data used in this publication. The Initiative is supported by funding to J.E.T. from Bioplatforms Australia (enabled by NCRIS).

Data availability All sequence data analysed in this work are available from the NCBI GenBank database.

Declarations

Ethics approval This article contains no studies with human participants or animals performed by any authors. Therefore, informed consent was not required.

Competing interests The authors declare no competing interests.

References

Agranovsky A (2016) Closteroviruses: Molecular Biology, Evolution and Interactions with Cells. In: Rajarshi Kumar Gaur NMP, Basavaprabhu L. Patil, Mariya Ivanova Stoyanova (ed) Plant Viruses: Evolution and Management. Springer Singapore, pp 231–252. https://doi.org/10.1007/978-981-10-1406-2

Folimonova SY (2020) Citrus tristeza virus: a large RNA virus with complex biology turned into a valuable tool for crop protection. PLoS Pathog 16(4):e1008416. https://doi.org/10.1371/journal.ppat.1008416

Fuchs M, Bar-Joseph M, Candresse T, Maree HJ, Martelli GP, Melzer MJ, Menzel W, Minafra A, Sabanadzovic S (2020) ICTV Virus Taxonomy Profile: *Closteroviridae*. J Gen Virol 101:364–365. https://doi.org/10.1099/jgv.0.001397

Gushchin VA, Karlin DG, Makhotenko AV, Khromov AV, Erokhina TN, Solovyev AG, Morozov SY, Agranovsky AA (2017) A conserved region in the Closterovirus 1a polyprotein drives extensive remodeling of endoplasmic reticulum membranes and induces motile globules in *Nicotiana benthamiana* cells. Virology 502:106–113. https://doi.org/10.1016/j.virol.2016.12.006

Maynard GV, Lepschi BJ, Malfroy SF (2018) Norfolk island quarantine survey 2012–2014 - a comprehensive assessment of an isolated subtropical island. ProcLinnean Soc NSW 140:7–243

Mrkvová M, Achs A, Alaxin P, Šubr Z, Predajňa L, Zetochová E, Hauptvogel P, Šoltys K, Candresse T, Glasa M (2023) Phaseolus vulgaris alphaendornavirus-1 is frequent in bean germplasm in Slovakia and shows low molecular variability. Acta Virol 67. https://doi.org/10.3389/av.2023.11484

Touriol C, Bornes S, Bonnal S, Audigier S, Prats H, Prats A-C, Vagner S (2003) Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. BiolCell 95(3–4):169–178. https://doi.org/10.1016/S0248-4900(03)00033-9