First report of orchid fleck virus and a novel strain of sweet potato chlorotic stunt virus on an ornamental cultivar of *Alcea rosea* L. in South Africa

David A. Read[1*†], Ronel Roberts[2] and Genevieve D. Thompson[1,3]

[1] Agricultural Research Council (ARC) - Biotechnology Platform, 100 Old Soutpan Road, Onderstepoort, Pretoria, 0110, South Africa.
[2] ARC - Plant Health and Protection, Private Bag X134, Queenswood, Pretoria 0121, South Africa
[3] Gene Vantage, 34 Monte Carlo Crescent, Kyalami Business Park, Johannesburg, 1684, South Africa

[*] Corresponding author – david.read@fabi.up.ac.za
[†] Current address - Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

ORCID – David A. Read
0000-0002-6614-9946

Keywords: Sweet potato chlorotic stunt virus; orchid fleck virus; *Alcea rosea*; South Africa

Abstract
Common hollyhock (*Alcea rosea*) is a ubiquitous ornamental in temperate climates but is highly adaptable and can be found growing in the tropics and subtropics. In 2019 an *A. rosea* plant showing symptoms of irregular chlorotic flecking on the basal leaves, with symptoms becoming gradually less severe toward the apex, was sampled in Pretoria, Gauteng province, South Africa. Total RNA was used to prepare an RNAtag-seq library, which was sequenced using an Illumina HiSeq 2500 instrument. Subsequent analysis of the data revealed the presence of two bipartite RNA viruses, namely orchid fleck virus (OFV) (segment 1: MW073772; segment 2: MW073773) and sweet potato chlorotic stunt virus (SPCSV) (segment 1: MW073774; segment 2: MW073775). OFV from this study was closely related to a strain from South Africa, associated with citrus leprosis disease, while SPCSV represented a novel strain. RT-PCR and bidirectional Sanger sequencing were used to confirm the presence of both viruses. Further samples were collected in 2020, which showed severe interveinal chlorosis, and were tested with RT-PCR; however only SPCSV was associated with these plants. This is the first time that both viruses have been associated with *A. rosea,* which should be considered a potential reservoir host of these agriculturally important viruses.

Alcea rosea L. (family: Malvaceae), also known as common hollyhock, is an annual, sometimes biennial plant native to China, but cultivars are cosmopolitan ornamentals in the tropics, subtropics and temperate climates. A number of viruses are known to affect *A. rosea* and include: hollyhock yellow vein mosaic virus (Srivastava et al. 2014), hollyhock leaf crumple virus (Abdel-Salam et al. 1998), tomato leaf curl virus (Ashwathappappa et al. 2020) (*Geminiviridae*); zucchini yellow mosaic virus (Choi et al. 2002), malva vein clearing virus (Menzel et al. 2010) (*Potyviridae*) and alcea rosea virus 1 (*Closteroviridae*) (Wang et al. 2020). Although the viruses known to infect *A. rosea* are diverse, observed symptoms have been mostly restricted to chlorotic mosaic and/or vein clearing of the leaves.

In November 2019, symptoms of interveinal chlorosis (*supplementary figure* 1) were observed on the mature leaves of an ornamental cultivar of *A. rosea* in Pretoria, Gauteng Province, South Africa (GPS coordinates: 25.649048 S; 28.186826 E). Newly emerged leaves appeared normal and symptomless. Leaf material was collected (sample number 19-3045) and total RNA was isolated, using a Thermo GeneJET Plant RNA purification kit (Thermo Fischer Scientific, Vilnius, Lithuania). Quality control of RNA was carried out using an Implen NanoPhotometer N60 instrument (Implen GmbH, Munich, Germany). An RNAtag-seq library was prepared according to Shishkin et al. (2015), with RiboZero plant leaf (Illumina, San Diego, CA, United States) for host ribosomal RNA depletion. Total RNA for 32 samples (31 of these were unrelated to this study) were used as input material for the preparation of the sequencing library, which was sequenced using an Illumina HiSeq 2500 instrument (Illumina, San Diego, United States) at Agricultural Research Council – Biotechnology Platform, Pretoria, South Africa.

The raw reads of the resulting datasets were demultiplexed using Je (Girardot et al. 2016), which resulted in 3,745,216 reads (average read length = 121bp) being associated with the *A. rosea* sample. Trimming for quality and adapter content was performed using CLC Genomics Workbench 9 (Qiagen Bioinformatics, Aarhus, Denmark), with the following trim parameters: Minimum read length = 20bp;

quality limit = 0.05 and trim adapters (Illumina universal adapter: 5' AGATCGGAAGAG 3'; RNAtag-seq adapter: 5' TACACGACGCTCTTCCGATCTNNNNNNNNT 3') and led to 3,552,785 post-trim reads (average reads length = 114.5bp). Raw trimmed reads are available as an NCBI sequence read archive (PRJNA664778). Assembly was performed using metaSPAdes 3.14.0 in paired-end mode with read error correction (Nurk et al. 2017). Virus-specific contigs were identified using blastn (Altschul et al. 1990) and the viral refseq database. Amino acid sequences for putative open reading frames were determined using the ORF finder (Wheeler et al. 2003). Average amino acid identities (AAI) between the putative amino acid sequences for the viruses from this study and the most closely related extant viruses, were determined using the AAI calculators from the enveomics collection (Rodriguez-R and Konstantinidis 2016). Deduced amino acid sequences (RNA-dependent RNA polymerase (RdRp)) for OFV; ORF1a polyprotein product for SPCSV) were aligned against reference sequences from NCBI, using ClustalW within BioEdit 7.2.5 (Hall 1999). Determination of best-fit maximum likelihood amino acid models, as well as the generation of phylogenies was performed using MEGA X (Stecher et al. 2020).

Contigs corresponding to the complete genomes of sweet potato chlorotic stunt virus (SPCSV) (family: *Closteroviridae*; genus: *Crinivirus*) and orchid fleck virus (OFV) (family: *Rhabdoviridae*; genus: *Dichorhavirus*) were associated with *A. alcea* sample 19-3045. The genomes of both viruses are single-stranded and bipartite, with the SPCSV genome being positive sense (Kiss et al. 2013) and OFV being negative sense (Dietzgen et al. 2018). OFV genome segments 1 and 2 were 6,431 (GenBank accession: MW073772) and 6,001nt (MW073773), with 206 and 300x average coverage respectively, while SPCSV segments 1 and 2 were 8,680 (MW073774) and 8,094nt (MW073775), with 23 and 28x coverage respectively. The locations of open reading frames and putative functions of gene products are listed in *table 1*. Deduced amino acid sequences for the OFV genes showed the greatest homology (98-100%) to MK522804/5 isolate CL, which has recently been shown to be associated with citrus leprosis-N disease in South Africa (Cook et al. 2019). The leader and trailer sequences show inverted complementarity for the terminal 6 and 10nt of segments 1 and 2 respectively. This type of complementarity is characteristic for members of the *Rhabdoviridae* family (Heim et al. 2018), which is thought to function in genome replication (Jackson et al. 2005). Comparison of deduced amino acid sequences for SPCSV showed that the strain associated with *Alcea* is most closely related to the m2-47 strain from Peru (HQ291259/60) (Cuellar et al. 2011). AAI values vary between 55.1% for the RNaseIII gene product and 92.4% for the heat shock protein 70 homolog (*table 1*). Both SPCSV and OFV phylogenies were generated using the Le Gascuel model with frequencies and gamma distribution (n = 5) with invariant sites, as well as 1000 bootstrap replicates. The SPCSV phylogeny, based on the ORF1a polyprotein (*figure 1*), suggested that the virus from this study represents a novel strain, since it is formed a branch that was isolated from other extant SPCSV strains. The high levels if homogeneity observed between the RdRp of the OFV strain from this study and those recently associated with citrus and orchid species in South Africa is confirmed by the close grouping in the respective phylogeny (*figure 2*).

Table 1: Genome segment lengths (in nucleotides (nt)), average nucleotide identity (ANI) shared with the most closely related viral strain, genes, gene locations, putative gene product sizes (in amino acids (aa)), product functions and average amino acid sequence identity (AAI) shared between the viruses in this study and cognate gene products the most closely related viral strains. ***SPCSV*** *– Sweet potato chlorotic stunt virus.* ***OFV*** *– Orchid fleck dichorhavirus.* [a] *HQ291259 SPCSV m2-47 RNA1;* [b] *HQ291260 SPCSV m2-47 RNA2;* [c] *MK522804 OFV CL RNA1* [d] *MK522805 OFV CL RNA2*

Virus	Genome segment	Length (nt)	ANI	Gene	Gene location (nt)	Product size (aa)	Putative gene function/s	AAI
SPCSV	RNA1 (MW073774)	8680	70.0 [a]	p227	88-6051	1987	ORF1a (Protease, methyltransferase and helicase domains)	73.2[a]
				RdRp	6119-7567	482	RNA-dependent RNA-polymerase	88.8[a]
				RNase3	7571-8281	236	Suppressor of RNA silencing	55.1[a]
				p7	8281-8451	56	Hypothetical	60[a]
	RNA2 (MW073775)	8094	75.0 [b]	HSP70h	874-2538	554	Heat shock protein 70 homolog	92.4[b]
				p60	2560-4116	518	Movement protein	81.1[b]
				p8	4098-4319	73	Hypothetical	71.2[b]
				CP	4352-5125	257	Coat protein	82.4[b]
				CPm	5128-7182	684	Minor coat protein	65.4[b]
				p28	7187-7915	242	Hypothetical	80.6[b]
OFV	RNA1 (MW073772)	6431	98.5 [c]	N	228-1580	450	Nucleocapsid	99[c]
				P	1730-2443	237	Phosphoprotein	99[c]
				ORF3	2579-3691	370	Movement protein	99[c]
				M	3832-4383	183	Matrix protein	100[c]
				G	4484-6226	580	Glycoprotein	98[c]
	RNA2 (MW073773)	6001	98.3 [d]	L	183-5816	1877	RNA-dependent RNA-polymerase	99[d]

In October 2020, leaf samples were collected from symptomatic and asymptomatic plants that were presumed to be self-seeded volunteers, from the same collection site as 2019. Samples 19-3045 and 20-3001 to 20-3003 were symptomatic (*supplementary figure 1*), while samples 20-3004 to 20-3007 were asymptomatic. An unidentified species of *Brevipalpus* mite (Tenuipalpidae), as well as indivduals *Bemisia tabaci* were observed simultaneously on the plants sampled in 2020. Plants of *Ipomea batatas* (sweet potato) and *Cymbidium* sp. were also both observed close to the collection site. Total RNA was isolated, as described previously, and the presence of SPCSV and OFV determined using two-step RT-PCR. SPCSV primers were designed to amplify part of the coat protein gene (CP) with the following sequences: SPCSV-CP-F (5' CTC AGA TTC AGA CAT TGG AGC TGG A 3') / SPCSV-CP-R (5' AGG TAG TGA GTA GTC CTG GTT GA 3'). Primers targeting OFV were designed to amplify part of the N-gene and had the following sequences: OFV-N-F (5' ACC TGG TGT TTA TGT CGG CTA TGA C 3') / OFV-N-R (5' CTG GAA ACC ATG ACC ATG CCT GAT 3'). Initial denaturation of 30ng of RNA was performed at 70° C for 5 min in the presence of 2.5 μM of each RT primer (forward for OFV, reverse for SPCSV), with RNase-free water to a final volume of 5μl. A further incubation step at 4° C for 5min, completed RT primer annealing. RT reactions were setup using Promega GoScript™ Reverse Transcriptase in accordance with manufacturer instructions, with the following reaction conditions: 25° C for 5min; 42° C for 60min; 70° C for 15min. PCR reactions

were setup using using GoTaq® Taq polymerase (Promega, Madison, WI, United States), in accordance with manufacturer instructions, with 1µl of cDNA as template and 5µM of each respective primer. Reaction conditions were: 5 min at 94° C followed by 35 cycles of 94° C for 30s; 55° C for 30s; 1 min at 72° C before a final extension step at 72° C for 10 min. PCR amplicons were visualized on an agarose gel (*supplementary figure 2*) and positive reactions products for sample 19-3045 were subjected to bidirectional Sanger sequencing to confirm their identity (Inqaba Biotechnical Industries, Pretoria, South Africa). The sequences of the amplicons were >99% similar to the cognate regions of the complete genomes from 19-3045, with GenBank accession numbers MW489857 and MW489858.

The presence of SPCSV was confirmed in all symptomatic samples, while the presence of OFV was only confirmed in the sample 19-3045. The presence of only SPCSV in samples collected in 2020, suggests that it responsible for the observed symptoms, which were highly characteristic of criniviruses infections (Tzanetakis et al. 2013), which included interveinal chlorosis on basal leaves with a decline in symptom severity toward the apex of the plants. Although SPCSV from this study appears to be highly divergent from other extant SPCSV strains, it does however not represent a novel species since shared AAI among key genes (polymerase, CP and HSP70h) are above the lower limit for crinivirus species demarcation (Martelli et al. 2012).

Synergistic co-infections of SPCSV and sweet potato feathery mottle virus in *I. batatas* lead to a severe, yield-limiting disease of sweet potato (sweet potato virus disease), particularly in the tropics of Africa (Gibson et al. 1998). SPCSV has been shown to infect wild species of *Ipomea* (Tugume et al. 2012) but has up to now been considered to be limited to members of the *Convolvulaceae* family (Maliogka et al. 2019). OFV host range has been limited to species of *Orchidaceae, Asparagaceae* and *Citrus* (Dietzgen et al. 2018). *Alcea* is a known host of both *Brevipalpus* species and *Bemisia tabaci,* which are also confirmed vectors of OFV and SPCSV respectively (Childers et al. 2003; Srivastava et al. 2014; Tzanetakis et al. 2013). An active putative vector population, along with the presence of known OFV and SPCSV hosts close to the sampling site, suggests an existing mechanism for the transfer of these viruses to *Alcea.* This study represents the first report of SPCSV and OFV infecting *A. rosea* and the first association of these viruses with a member of the *Malvaceae* family. *A. rosea* has naturalized in temperate regions and is adaptable to a wide range of climates including the tropics (Lim 2014). Both sweet potato and citrus are important crops in both tropical and subtropical climates, which presents a significant geographic overlap with *A. rosea* as an ornamental. Future research should include determining whether the infection of *A. rosea* by SPCSV and OFV is widespread or localized, fulfilling Koch's postulates*,* determining whether any synergy exists for multiple infections of these viruses, as well as confirming the identity of their respective insect vectors. However, *A. rosea* should be considered a potentially important reservoir host for these two agriculturally important viruses.

Declarations

Funding: David Read is grateful for the financial support provided by the National Research Foundation of South Africa, under grant UID 104901.

Conflict of interest: All authors declare that they have no conflict of interest.

Availability of data and material: The data that support the findings of this study, are openly available in NCBI public databases.

Code availability: Not applicable

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abdel-Salam, AM., El-Shazly, MA., & Thouvenel, JC. (1998). Biological. biochemical and serological studies on hollyhock leaf crumple virus (HLCrV): A newly discovered whitefly transmitted Geminivirus. *Arab Journal of Biotechnology, 1*, 41-58

Altschul, SF., Gish, W., Miller, W., Myers, EW., & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*, 403-410

Ashwathappappa, K.V., Venkataravanappa, V., Lakshminarayana Reddy, C.N., & Krishna Reddy, M. (2020). Molecular characterization of Tomato leaf curl virus infecting hollyhock (Alcea rosea L.) in India. *Indian Phytopathology, 73*, 339–347

Childers, C.C., Rodrigues, J.C.V., & Welbourn, W.C. (2003). Host plants of *Brevipalpus californicus*, *B. obovatus*, and *B. phoenicis* (Acari: Tenuipalpidae) and their potential involvement in the spread of viral diseases vectored by these mites. *Experimental and Applied Acarology*, *30,* 29–105

Choi, S.K., Yoon, J.Y., Ryu, K.H., Choi, JK., & Park, WM. (2002). First report of *Zucchini yellow mosaic virus* on hollyhock (*Althaea rosea*). *The Plant Pathology Journal, 18*, 121-125

Cook, G., Kirkman, W., Clase, R., Steyn, C., Basson, E., Fourie, P.H., Moore, S.D., Grout, T.G., & Hattingh, V. (2019). Orchid fleck virus associated with the first case of citrus leprosis-N in South Africa. *European Journal of Plant Pathology, 155*, 1373–1379

Cuellar, W.J., Cruzado, R.K., Fuentes, S., Untiveros, S., Soto, M., & Kreuze, J.F. (2011). Sequence characterization of a Peruvian isolate of Sweet potato chlorotic stunt virus: Further variability and a model for p22 acquisition. *Virus Research, 157*, 111-115

Dietzgen, R.G., Freitas-Astúa, J., Chabi-Jesus, C., Ramos-González, P.L., Goodin, M.M., Kondo, H., Tassi, A.D., & Kitajima, E.W. (2018). Dichorhaviruses in their host plants and mite vectors. *Advances in Virus Research, 102*, 119-148

Gibson, R.W., Mpembe, I., Alicai, T., Carey, E.E., Mwanga, R.O.M., Seal, S.E., & Vetten, H.J. (1998). Symptoms, aetiology and serological analysis of sweetpotato virus disease in Uganda. *Plant Pathology, 47*, 95–102

Girardot, C., Scholtalbers, J., Sauer, S., Su, S.Y., & Furlong, E.E. (2016). Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics, 17(1),* 419

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

Heim, F., Lot, H., Delecolle, B., Bassler, A., Krczal, G., & Wetzel, T. (2008). Complete nucleotide sequence of a putative new cytorhabdovirus infecting lettuce. *Archives of Virology, 153*, 81–92

Jackson, A.O., Dietzgen, R.G., Goodin, M.M., Bragg, J.N., & Deng, M. (2005). Biology of plant rhabdoviruses. *Annual Review of Phytopathology, 43*, 623 - 660

Kiss, Z.A., Medina, V., & Falk, B.W. (2013). Crinivirus replication and host interactions. *Frontiers in Microbiology, 4*, 99

Lim, T.K. (2014). Alcea rosea. In: *Edible Medicinal and Non Medicinal Plants.* (pp 292-299) Dordrecht: Springer

Maliogka, V.I., Wintermantel, W.M., Orfanidou, C.G., & Katis, N.I. (2019). Criniviruses infecting vegetable crops. In: P. Poltronieri, Y. Hong (Eds) *Applied plant biotechnology for improving resistance to biotic stress.* (pp 251-289) Amsterdam: Elsevier

Martelli, G.P., Agranovsky, A.A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R.H.A., Dolja V.V., et al. (2012). Closteroviridae. In: A. M. Q. King, M. J. Adams, E. B. Carstens, E. Lefkowitz (Eds) *Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses.* (pp 987-1001) San Diego: Elsevier

Menzel, W., Winter, S., & Richert-Pöggeler, K.R. (2010). First report of Malva vein clearing virus naturally occurring in hollyhock in Germany. *Plant Disease, 94*, 276

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P.A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Research, 27*, 824-834

Rodriguez-R, L.M., & Konstantinidis, K.T. (2016). The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints, 4*, e1900v1.

Shishkin, A.A., Giannoukos, G., Kucukural, A., Ciulla, D., Busby, M., Surka, C., & Chen, J. (2015). Simultaneous generation of many RNA-seq libraries in a single reaction. *Nature Methods, 12*, 323–325

Srivastava, A., Kumar, S., Raj, S.K., & Pande, S.S. (2014). Association of a distinct strain of hollyhock yellow vein mosaic virus and Ludwigia leaf distortion betasatellite with yellow vein mosaic disease of hollyhock (Alcea rosea) in India. *Archives of Virology, 159*, 2711–2715

Stecher, G., Tamura, K., & Kumar, S. (2020). Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular Biology and Evolution, 37*, 1237–1239

Tugume, A.K., Amayo, R., Weinheimer, I., Mukasa, S.B., Rubaihayo, P.R., & Valkonen, J.P.T. (2012). Genetic variability and evolutionary implications of RNA silencing suppressor genes in RNA1 of sweet potato chlorotic stunt virus isolates infecting sweetpotato and related wild species. *PLoS ONE, 8(11)*, e81479

Tzanetakis, I.E., Martin, R.R., & Wintermantel, W.M. (2013). Epidemiology of criniviruses: an emerging problem in world agriculture. *Frontiers in Microbiology, 4,* 119

Wang, D., Cui, L., Yu, P., Qi, Y., Wang, B., & Niu, Y. (2020). A new closterovirus in Alcea rosea identified by small RNA deep sequencing. *Journal of Plant Pathology, 102*, 29–37

Wheeler, D.L., Church, D.M., Federhen, S., Lash, A.E., Madden, T.L., Pontius, J.U., Schuler, G.D., et al. (2003). Database resources of the National Center for Biotechnology. *Nucleic Acids Research, 31*, 28-33

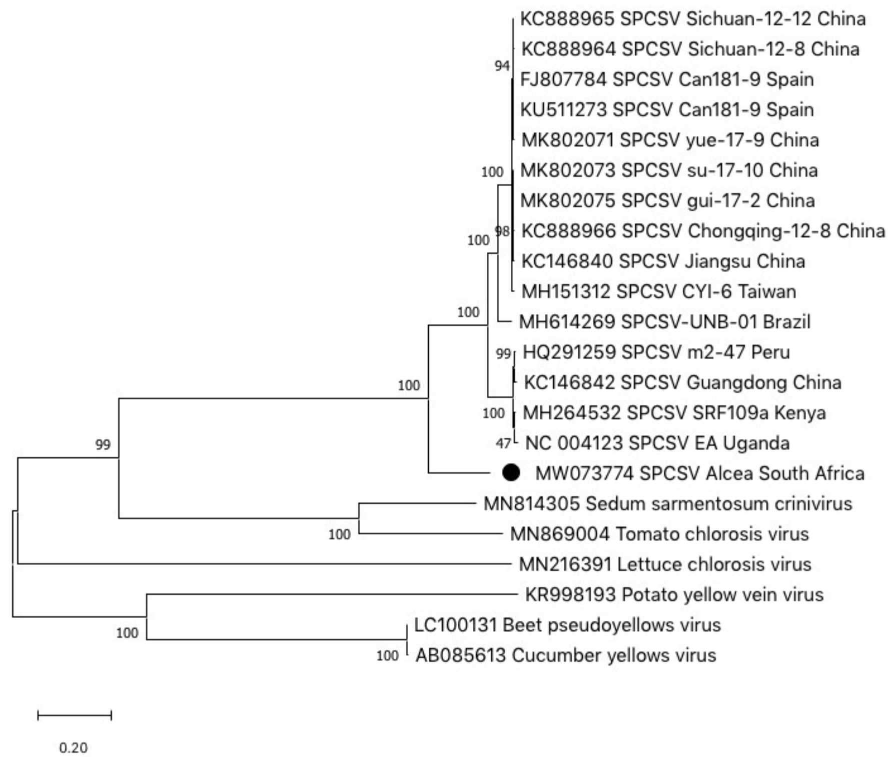

***Figure 1**: Maximum likelihood phylogeny based on the amino acid sequences of the ORF1a polyprotein of the SPCSV strain from Alcea (indicated by a solid circle marker) and the cognate amino acid sequences from selected strains of SPCSV as well as additional criniviruses, as references. The phylogeny represents the tree with the highest log likelihood and was generated in MEGA X using the Le Gascuel model with frequencies, gamma distribution (n = 5) with invariant sites, as well as bootstrapping (1000 replicates). The percentage of trees in which the associated taxa clustered together is shown next to the branches (Percentages lower than 50 are not shown)*

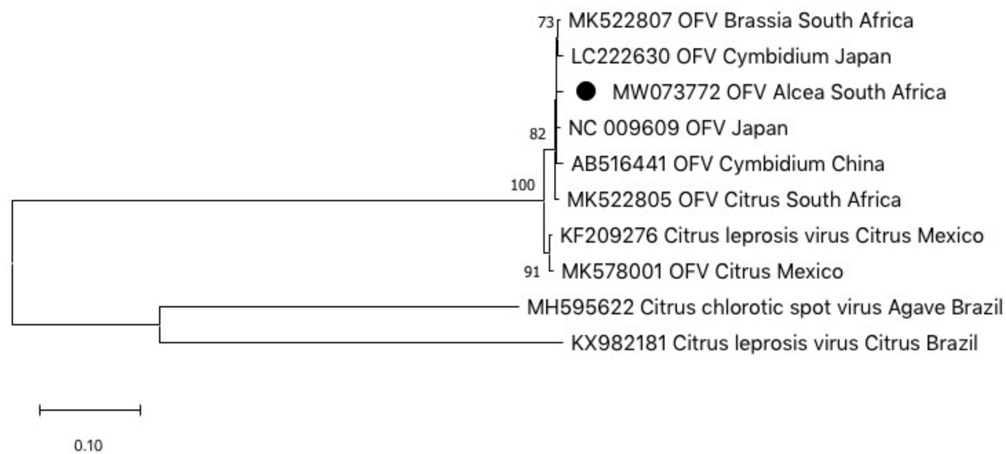

***Figure 2**: Maximum likelihood phylogeny based on the amino acid sequences of the RNA-dependent RNA polymerase of the OFV strain from Alcea (indicated by a solid circle marker) and the cognate amino acid sequences from selected strains of OFV as well as additional dichorhaviruses, as references. The phylogeny represents the tree with the highest log likelihood and was generated in MEGA X using the Le Gascuel model with frequencies, gamma distribution (n = 5) with invariant sites, as well as bootstrapping (1000 replicates). The percentage of trees in which the associated taxa clustered together is shown next to the branches (Percentages lower than 50 are not shown)*