Novel viruses associated with Amaryllidaceae species in South Africa

David A. Read[1*], Ronel Roberts[2], Dirk Swanevelder[1], Gerhard Pietersen[3] and Genevieve D. Thompson[1,4]

[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] Department of Genetics, Stellenbosch University, Stellenbosch, 7600, South Africa

[4] Gene Vantage, 34 Monte Carlo Crescent, Kyalami Business Park, Johannesburg, 1684, South Africa

* - Corresponding author

Email: david.read@fabi.up.ac.za

Telephone: +27833259077

Current address: Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

Abstract

Nineteen samples of *Agapanthus, Clivia, Hippeastrum* and *Scadoxus* were collected from gardens in the Gauteng and Western Cape provinces. Plants displayed highly variable symptoms of viral disease, including chlorosis, necrosis, streaking and ringspot. RNAtag-seq was used to characterize the associated viral populations. *Agapanthus* was shown to be associated with three novel viruses from the *Caulimo-, Clostero-* and *Betaflexiviridae* families, *Clivia* with novel members of the *Poty-* and *Betaflexiviridae* families and *Scadoxus* with a novel member of the *Tospoviridae* family. Nerine latent virus was associated with *Agapanthus, Clivia* and *Hippeastrum*, while hippeastrum mosaic virus was associated exclusively with a *Hippeastrum* cultivar.

Amaryllidaceae is a family of bulbous, plant species with several genera commonly grown as ornamentals [1,2]. Propagation typically occurs vegetatively, which promotes the spread and accumulation of viruses. A diverse set of viruses are known to infect *Agapanthus* [3], but data regarding *Clivia,* remains sparse [4]. *Amaryllis* and *Hippeastrum* have been associated with nerine yellow stripe virus [5], hippeastrum mosaic virus (HiMV) and nerine latent virus (NeLV) [6]. *Scadoxus multiflorus* is also a natural host of NeLV [7].

During 2019, 19 samples were collected from *Agapanthus, Clivia, Hippeastrum* and *Scadoxus*, from private and public collections in the Gauteng and Western Cape provinces, South Africa. Species, collection date and location, and descriptions of foliar symptoms are presented in Table 1, with images shown in (Supplementary Fig. 1, A-S). *Agapanthus* samples typically had mild to severe streaking and mottling, sometimes in the presence of necrotic lesions (A-L). *Clivia* showed mild to severe chlorotic streaks with two of the five samples also having additional necrotic lesions (M-Q). The *Hippeastrum* sample had mild mottling only (R), while the *Scadoxus* sample had severe ringspots, with both chlorotic and necrotic lesions (S).

Total RNA was isolated using the GeneJET Plant RNA kit (ThermoFischer Scientific, Vilnius, Lithuania). An RNAtag-seq library was prepared according to Shishkin *et al*. [8] and sequenced as paired-ends (2x125nt) using TruSeq V4 chemistry (Illumina, San Diego, USA) on a HiSeq 2500 instrument (Illumina, San Diego, USA) at the Agricultural Research Council Biotechnology Platform, Onderstepoort, Pretoria, South Africa.

Data was demultiplexed using *Je suite* [9] and trimmed with *CLC Genomics Workbench 9* (Qiagen Bioinformatics, Aarhus, Denmark) for length (reads <20bp discarded), quality (limit = 0.05) and adapter content (Illumina universal and RNAtag-seq adapters: 5'AGATCGGAAGAG and 5'TACACGACGCTCTTCC-GATCTNNNNNNNNT, respectively). The total number of reads was 50,164,374, of which 42,479,299 (85%) remained after quality and adapter trimming. The number of pre- and post-trim reads associated with each dataset is shown in Supplementary Table 1. Raw reads are available as an NCBI SRA, accession number PRJNA627819.

Trimmed reads were assembled using both *CLC Genomics Workbench 9* and *metaSPAdes* [10]. Plant virus contigs were identified using *blastn* [11] with the NCBI refseq database for viruses. The longest contig lengths between *CLC Genomics Workbench 9* and *metaSPAdes* were selected for further analysis. Genome organization, open reading frame (ORF) position and the amino acid sequences of putative genes were determined using *ORF finder* [12]. Average nucleotide identities (ANI) and average amino acid identities (AAI) were determined using the *Enveomics ANI* and *AAI* tools respectively [13].

Table 1: Provenance of samples used in this study. Viral genomes with terminal nucleotides confirmed with RACE are shown in bold and underlined. * Only 5' RACE performed.

Sample number	Genus/species	Collection date	General collection area	GPS coordinates	Description of symptoms	Plant viruses present
19-3001	*Agapanthus* sp.	01-2019	Pretoria, Gauteng	25.6491S 28.1868E	Mottling, chlorosis, streaks and necrosis	AgTV
19-3004	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.7802S 28.2919E	Mottling, chlorosis, streaks and stunting	AgTV **AgVV*** **AgVA**
19-3005	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.7807S 28.2920E	Mottling, chlorosis, streaks, stunting and necrosis	AgTV AgVV
19-3006	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.7813E 28.2933S	Mottling, streaks, chlorosis, stunting, necrosis and ring spotting	AgTV AgVV AgVA
19-3007	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.7811S 28.2930E	Mottling, streaks, chlorosis and ring spotting	AgTV AgVV
19-3008	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.6510S 28.1862E	Mottling, chlorosis, stunting, streaks	AgTV AgVV
19-3009	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.6505S 28.1858E	Mottling, streaks, chlorosis, stunting, wavy leaf deformation	AgVV NeLV
19-3010	*Agapanthus* sp.	08-2019	Pretoria, Gauteng	25.6505S 28.1861E	Mottling, chlorosis, stunting	AgTV AgVV
19-3013	*Agapanthus* sp.	10-2019	Pretoria, Gauteng	25.6151S 28.3656E	Mottling, chlorosis	AgTV AgVV NeLV
19-3017	*Agapanthus* sp.	10-2019	Pretoria, Gauteng	25.6150S 28.3652E	Mottling, chlorosis	AgTV **AgVV***
19-3021	*Agapanthus* sp.	10-2019	Pretoria, Gauteng	25.6151S 28.3645E	Mottling, streaks, chlorosis	AgTV
19-3035	*Agapanthus* sp.	11-2019	Stellenbosch, Western Cape	33.9279S 18.8799E	Mottling, streaks, chlorosis	**AgCVB**
19-3019	*Clivia* sp.	10-2019	Pretoria, Gauteng	25.6151S 28.3652E	Mild mottle	NeLV
19-3026	*Clivia* sp.	10-2019	Midrand, Gauteng	25.9558S 28.2174E	Severe chlorotic streaks	**ClYSV**
19-3027	*Clivia* sp.	10-2019	Midrand, Gauteng	25.9558S 28.2174E	Mild chlorotic streaks	ClYSV
19-3040	*Clivia* sp.	11-2019	Johannesburg, Gauteng	26.0900S 27.8430E	Chlorotic mottle, necrotic lesions	**ClCVA**
19-3044	*Clivia* sp.	11-2019	Johannesburg, Gauteng	26.0887S 27.8422E	Chlorotic mottle, necrotic lesions	NeLV
19-3029	*Hippeastrum* sp.	10-2019	Midrand, Gauteng	25.9558S, 28.2174E	Mild mottle	HiMV NeLV
19-3032	*Scadoxus puniceus*	11-2019	Pretoria, Gauteng	25.7526S 28.2284E	Chlorotic ringspot	ScCSV

Selections of gene regions for phylogenetic analysis, as well as cutoffs for species delineation, were made based on previous studies [14, 15, 16, 17, 18]. Nucleotide and amino acid alignments were performed using the *ClustalW* function within *BioEdit 7.2.5* [19]. Determination of best-fit DNA/protein models, as well as the phylogenetic analyses (maximum likelihood (ML) with 1,000 bootstrap replicates) were performed using *MEGA X* [20]. The GTR model was used for the agapanthus tungro virus (AgTV) nucleotide phylogeny and the LG model for all of the amino acid phylogenies with the gamma frequencies option (n=5).

The presence of novel viral contigs was confirmed using PCR and two-step reverse-transcription PCR (RT-PCR). All amplification products were subjected to bidirectional Sanger sequencing (Inqaba Biotec, Pretoria, South Africa). The 5' and 3' terminal nucleotides were determined for a representative of each novel virus using the 5' RACE System for Rapid Amplification of cDNA Ends, version 2.0 (Invitrogen, Carlsbad, CA, USA) and a 3' RACE System for Rapid Amplification of cDNA Ends (Invitrogen, Carlsbad, CA, USA) respectively, with the exception of the velariviruses, whose 3' ends were recalcitrant to RACE, the 3' end of agapanthus carlavirus B (AgCVB) whose contig already had a poly-A-tract and the orthotospovirus, which displayed the terminal complementarity, typical of complete orthotospovirus genome segments. The PCR, RT-PCR and RACE primer sequences are listed in Supplementary Table 2.

Nine samples of *Agapanthus* were associated with a novel member of the *Caulimoviridae*, that has been putatively named agapanthus tungro virus (AgTV). PCR with the AgTV F/R primer pair was used to amplify the AgTV partial intergenic region. Four ORFs encoded for putative proteins, which share the greatest amino acid diversity (26.6 - 44.9%) with rice tungro bacilliform virus (RTBV) (NC001914) (Fig. 1). The putative minus strand binding site showed complete homology to the priming site of RTBV genomes [21] (TGGTATCAGAGC[G/A]ATG) with the exception of a single nucleotide deletion at position 13 (Supplementary Fig. 2). Despite low, shared AAI (<45% for the polyprotein) and the lack of a waikavirus helper [22] for AgTV, the shared genome organization (Table 2; Fig. 1) and phylogenetic relationships (Supplementary Fig. 3.1) suggest that the AgTV variants in this study could represent the second extant species of *Tungrovirus* to date. AgTV clusters into two distinct variants, designated as either AgTV 1 or 2. Complete genomes were generated for nine samples (Table 1) from two collection sites and varied between 7,935 and 8,061bp in length (GenBank accessions: MT501685 to MT501691, MT579870, MW080813, MW080814). Sample 19-3001 was infected with AgTV only and demonstrated severe chlorotic and necrotic foliar lesions and we putatively link these symptoms with AgTV.

Table 2: Characterization of novel viruses of Amaryllidaceae. Putative gene location, product size, product and function, average amino acid sequence identity shared with the most closely related viral species, molecular weight (MW) and isoelectric point (pI) for each open reading frame (ORF). **AgTV –** Agapanthus tungro virus; **AgVV** – Agapanthus velarivirus; **AgVA** – Agapanthus virus A; **AgCVB** – Agapanthus virus B; **ClYSV –** Clivia yellow stripe virus; **ClCVA** – Clivia carlavirus A; **ScCRV –** Scadoxus chlorotic ringspot virus. Reference viruses: [1] Rice tungro bacilliform virus (NC001914); [2] Grapevine leafroll-associated virus 7 (NC016436); [3] Cordyline virus 3 (NC043107); [4] Caucasus prunus virus (NC038325) [5] Peach mosaic virus (KY795997), [6] Grapevine berry inner necrosis virus (NC015220), [7] Nerine latent virus (NC_028111), [8] Birch carlavirus (MH536506), [9] Alfalfa latent virus (NC026616), [10] Garlic virus D (MN059387), [11] Turnip mosaic virus (MK241971), [12] Poplar mosaic virus (NC005343), [13] Rose virus A (MN053272) [14] Elderberry carlavirus B (NC029086) [15]Butterbur mosaic virus (NC_013527) [16] Daphne virus S (AB889483), [17] Hippeastrum chlorotic ringspot virus – Segment L (KY911358); [18] Hippeastrum chlorotic ringspot virus – Segment M (KY911356) [19] Alstroemeria yellow spot virus – Segment S (NC_040741) [20] Hippeastrum chlorotic ringspot virus – Segment S (KY911352)

Sample	(bp/nt)		location (bp/nt)	size (aa)	gene product	function/s	(amino acid)	(kDa)	
AgTV 19-3001	7942	ORF1	90-653	187	p24	Unknown	34.1[1]	21.9	7.8
		ORF2	650-973	107	p12	Virion-associated protein	41.1[1]	12.1	9.7
		ORF3	970-5868	1632	p194	Polyprotein	44.9[1]	188.0	7.7
		ORF4	5822-7126	434	p46	Unknown	26.6[1]	49.0	6.3
AgVV 19-3004	17479	ORF1a	59-7750	2563	Pol	Polyprotein	35.4[2]	291.0	9.2
		ORF1b	7839-9266	475	RdRp	RNA dependent RNA polymerase	60.2[2]	55.4	8.3
		ORF2	9373-11031	552	HSP70	Heat shock protein 70 homolog	56.2[2]	61.7	5.9
		ORF3	10958-11218	86	p10	Unknown	30.4[3]	10.3	4.9
		ORF4	11209-12771	520	p60	Heat shock protein 90 homolog	33.3[2]	60.0	8.1
		ORF5	12779-13681	300	CP	Coat protein	36.6[2]	33.3	9.0
		ORF6	13701-15542	613	mCP	Minor coat protein	34.8[2]	70.2	7.6
		ORF7	15529-16203	224	p25	Unknown	22.7[2]	25.6	6.8
		ORF8	16254-16985	243	p27	Unknown	36.8[2]	28.3	5.4
		ORF 9	17040-17291	83	p9	Unknown	n/a	9.4	6.6
AgVA 19-3004	7116	ORF1	45-5429	1794	Rep	Replicase	34.7[4]	206.8	6.2
		ORF2	5338-6354	338	MP	Movement protein	26.0[5]	37.6	8.9
		ORF3	6432-7040	202	CP	Coat protein	32.1[6]	22.6	8.3
AgCVB 19-3035	8727	ORF1	61-6087	2008	Rep	Replicase	42.6[7]	229.6	7.3
		ORF2	6100-6810	236	TGB1	Viral movement	43.4[7]	26.8	5.7
		ORF3	6788-7123	111	TGB2		55.8[8]	11.9	9.4
		ORF4	7102-7311	69	TGB3		45.7[9]	8.0	6.1
		ORF5	7154-8053	286	CP	Coat protein	62.6[7]	32.3	6.5
		ORF6	8055-8357	140	NABP	Nucleic acid binding	29.1[10]	16.4	9.9
ClYSV 19-3027	9800	ORF1	134-9625	3163	Pol	Polyprotein	48.1[11]	359.2	8.4
ClCVA 19-3040	8411	ORF1	96-5900	1934	Rep	Replicase	40.9[12]	221.0	7.9
		ORF2	5923-6630	235	TGB1	Viral movement	45.6[13]	26.7	6.0
		ORF3	6605-6940	111	TGB2		55.6[14]	12.2	9.5
		ORF4	6904-7137	77	TGB3		48.8[14]	8.0	4.4
		ORF5	7154-8053	299	CP	Coat protein	56.6[15]	33.2	8.6
		ORF6	8055-8357	100	NABP	Nucleic acid binding protein	43.6[16]	11.8	9.6
ScCRV 19-3032	8915-L segment	ORF1	8876-267	2869	RdRp	RNA-dependent-RNA-polymerase	73.3[17]	331.4	6.4
	4885-M segment	ORF1	65-991	308	NSm	Movement protein	72.0[18]	33.7	6.9
		ORF2	4836-1453	1127	GP	Glycoprotein	62.0[18]	127.4	5.9
	3406-S segment	ORF1	72-1397	441	NSs	Non-structural protein	57.0[19]	50.2	6.7
		ORF2	3337-2525	270	N	Nucleocapsid	50.7[20]	30.0	8.9

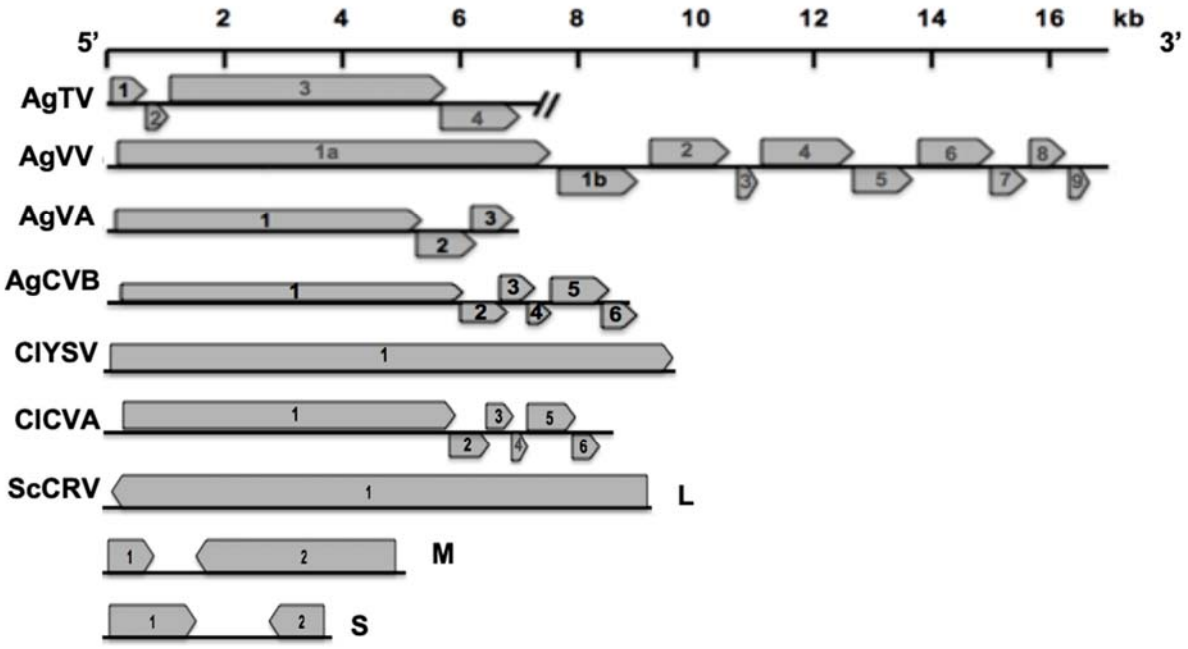

Fig 1: Genome maps of the novel viruses of Amaryllidaceae **AgTV –** Agapanthus tungro virus. 1 – p24; 2 – p12; 3 – p194; 4 – p46. **AgVV** – Agapanthus velarivirus. 1 – Polyprotein; 2 – RNA-dependent RNA polymerase (RdRp); 3 – p10; 4 – p60; 5 – coat protein (CP); 6 – minor CP; 7 – p25; 8 – p27; 9 - p9. **AgVA** – Agapanthus virus A. 1 – Replicase (Rep); 2 – movement protein; 3 – CP. **AgCVB** – Agapanthus virus B. 1 – Rep; 2 – Triple gene block (TGB) 1; 3 – TGB 2; 4 – TGB 3; 5 – CP; 6 – Nucleic acid binding protein. **ClYSV –** Clivia yellow stripe virus. 1 - **ClCVA** – Clivia carlavirus A. 1 – Rep; 2 – Triple gene block (TGB) 1; 3 – TGB 2; 4 – TGB 3; 5 – CP; 6 – Nucleic acid binding protein. **ScCRV –** Scadoxus chlorotic ringspot virus. L-segment: 1 – RdRp. M-segment: 1 – NSm; 2 – Glycoprotein. S-segment: 1 – NSs; 2 – Nucleocapsid.

A member of the *Velarivirus* genus (*Closteroviridae*) was found to be associated with *Agapanthus* (Table 1), with the proposed name of agapanthus velarivirus (AgVV) and having the greatest level of sequence homology to grapevine leafroll-associated virus 7 (GLRaV-7). This was confirmed by a ML phylogeny (Supplementary Fig. 3.2), which also indicated three variants of AgVV (variants 1 to 3). The viral genomes varied in length from 17,479 to 17,656nt and encode for ten putative ORFs (MT533600 - MT533607 and MW080811 - MW080812), however with the lack of 3' terminal confirmation, these genomes may be incomplete. The position, product size and putative gene function for each ORF, are listed in Table 2, as well as AAI with respect to the cognate gene product of the most closely related virus. AgVV genome organization and gene lengths are most analogous to GRLaV-7, with the exception of a putative p9 gene, which appears to be AgVV specific (Figure 1). The AgVV variants from this study were always components of mixed viral populations (Table 1), preventing the determination of their contribution to symptomatology. Assigning definitive symptomology to other velariviruses has been also been difficult, with cordyline virus variants being associated with both symptomatic and asymptomatic plants [23], however, a strong association between areca palm velarivirus 1 and yellow leaf disease of *Areca catechu* has been observed [24].

Two *Agapanthus* samples contained a novel virus within the *Betaflexiviridae* family, and agapanthus virus A (AgVA) is the proposed name. The two AgVA genomes vary by 7 nucleotides in length, i.e., 7,109 (MT533608) vs 7,116nt (MT533609), and encode three putative ORFs – a replicase, movement and coat proteins (Table 2), which is consistent with the *Citri-, Capillo-* and *Trichovirus* genera. All gene products share <35% AAI with the cognate gene of the most closely related viruses, i.e., the replicase of caucasus prunus virus, movement protein of peach mosaic virus and coat protein of grapevine berry inner necrosis virus (Fig. 1). An ML phylogeny (Supplementary Fig. 3.3) shows that AgVA genomes group with other members of the *Betaflexiviridae* family, but as an isolated clade between those of the *Citri-* and *Capillovirus* genera, which suggests that it belongs to a yet unclassified genus of the *Betaflexiviridae* family. A second novel betaflexivirus was detected in *Agapanthus* from the Western Cape with proposed name AgCVB. The genome is 8,727nt long, (MW303999), and encodes for six putative ORFs (Table 2), consistent with the *Carlavirus* genus [16]. Shared AAI with cognate gene products of the most closely related viruses varies from 62.6% for the coat protein to 29.1% for the nucleic acid binding protein (Fig. 1). The associated phylogeny of the putative replicase product suggests that this virus is most closely related to nerine latent virus (NeLV). The AAI of <80% shared between replicase of AgCVB and the cognate sequence of the most closely related virus, provides further evidence that AgCVB us a putative novel species.

Two *Clivia* cultivars showing symptoms of chlorotic streaking (Supplementary Fig. 1N and O) appeared to be infected with a novel member of the *Potyviridae* family, with proposed name clivia yellow stripe virus (ClYSV). The genome sizes were both 9,800nt (MT533610 and MT533611). The polyprotein amino acid sequence is most homologous (48.1%) to the cognate sequence of turnip mosaic virus (TuMV) (Table 2), suggesting that it is a novel species of the *Potyvirus* genus [17]. The ML phylogeny shows that ClYSV clusters with other extant potyviruses (Supplementary Fig. 3.2). A second virus associated with *Clivia* appears to be a novel species of the *Carlavirus* genus with proposed name, clivia carlavirus A (ClCVA). This virus has a genome length of 8,412nt (MT533598) that encodes for six ORFs with amino acid identities that varied between 40.9% (replicase) and 56.6% (coat protein) (Table 2). ClCVA clusters with other carlavirus members. Being the only plant virus associated with the sample, it is likely that ClCVA is responsible for the interveinal chlorotic symptoms observed.

Despite having a high diversity and wide geographic distribution [25], no viruses have previously been reported for any African species of *Scadoxus*. Three contigs of a putatively novel virus from the *Tospoviridae* family, provisionally named scadoxus chlorotic ringspot

virus (ScCRV), was detected in *S. puniceus*, with symptoms of severe foliar chlorotic ringspotting observed (Supplementary Fig. 1S). The genome consisted of three contigs of 8,958, 4,937 and 3,495nt in length and correspond to the L, M and S genomic segments (MW080808, MW080809 and MW080810) respectively. The highest AAI was shared with hippeastrum chlorotic ringspot virus (HCRV) replicase (73.3%) and the nucleocapsid gene (50.7%) (Fig. 4) and the NSs of alstroemeria yellow spot virus (AYSV) (57%). ScCRV exhibited typical genome organization of the *Tospoviridae* family, except for an unusually large intergenic region associated with the S segment. This has previously been observed for variants of other orthotospoviruses, such as capsicum chlorosis virus [26]. The ML phylogeny of the nucleocapsid gene product suggests a relationship between ScCRV and AYSV (Supplementary Fig. 3.5). The inverted complementarity observed in the terminal nucleotides (5' AGAGCAATC 3') is typical for members of the *Tospoviridae* family [27]. ScCRV is the first virus to be associated with *Scadoxus* in South Africa. Given that no other additional plant viruses were detected, it is probable that the severe ringspot symptoms observed were the result of ScCRV infection.

Amongst known viruses, HiMV was associated in a single sample of a *Hippeastrum* cultivar, with a length of 9,722nt (MT533599). HiMV has a wide geographic range, being detected in Australasia [6], South America [28], Asia [29] and Europe [30]. This is the first genomic characterization of an HiMV from South Africa. The polyprotein of HiMV from this study shares a >80% AAI with the most closely related variants from Australia, which is above the threshold for species demarcation among potyviruses [17]. Therefore, while the South African variant appears phylogenetically distinct, it should be designated as an HiMV genetic variant.

NeLV was observed in two samples of *Agapanthus*, two *Clivia* samples and *Hippeastrum*. The *Hippeastrum* sample was also associated with two contigs that shared 95% and 98% similarity with NeLV isolates from *Narcissus* (NC008552 and NC028111), with an ML phylogeny (replicase) confirming the separation of these two variants. Contigs associated with *Agapanthus* had sizes of 8,282nt (MT536155) and 8,409nt (MT536160) (Supplementary Table 3), which grouped with NeLV isolates from *Crinum* (MG012804) and *Narcissus* (NC008552 and JQ395044) (Supplementary Fig. 3.3). NeLV was detected also associated with *Clivia*, with genome lengths of 8,282 and 8,275nt (MT536157 and MT536160 respectively) and grouping with NeLV isolates from *Narcissus*. Sample 19-3019 showed mild chlorotic mottle on the leaves, while 19-3044 showed more severe chlorotic mottle with necrotic lesions (Supplementary Fig. 1M and Q). NeLV appears to show a wide host range within the

Amaryllidaceae and already is known to infect *Hippeastrum*. This is the first report of NeLV in South Africa and the first time the virus has been shown to infect *Agapanthus* and *Clivia.*

The last decade has seen the mainstream use of next-generation sequencing for the characterization of plant virus populations and the discovery of novel viruses. The focus, however, was economically important crop species with studies on wild and ornamental species comparatively lagging. Additionally, the ability to diagnose plant viruses based on sequencing alone has continued to outstrip capacity to associate these viruses with specific diseases, often showing the presence of mixed viral populations [31, 32]. This is true of the current study, especially for *Agapanthus*, where most novel viruses were present as mixed infections. This necessitates additional work to determine the biological role that each component in these populations play. Despite this limitation, the work has uncovered an unprecedented diversity of viruses among members of the Amaryllidaceae and is likely to be the largest metaviromic study of the family to date. Additionally, it represents the first report of members of the *Clostero-* and *Caulimoviridae* families associated with *Agapanthus,* as well as the first molecular descriptions of AgTV, AgVV, AgVA, AgCVB, ClYSV, ClCVA and ScCRV. Many of the observed symptoms were severe and are possibly a threat to natural plant populations and the horticultural industry. The presence of viruses within horticulturally important families has the potential to reduce profitability through yield reduction and exports limited by phytosanitary restrictions.

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. The authors have notified the relevant local agricultural authority (DALRRD) of these findings.

References

1 Reinten EY, Coetzee JH, van Wyk BE (2011) The potential of South African indigenous plants for the international cutflower trade. S Afr J Bot 77:934-946

2 Baskaran P, van Staden J (2013) Rapid *in vitro* micropropagation of *Agapanthus praecox*. S Afr J Bot 86:46-50

3 Wilson CR, Wilson AJ, Pethybridge SJ (2000) First report of tomato spotted wilt virus in common *Agapanthus*. Plant Dis 84:491

4 Jones DR (2005) Plant viruses transmitted by thrips. Eur J Plant Pathol 113:119-157

5 Pearson MN, Cohen AD, Cowell BSJ et al (2009). A survey of viruses of flower bulbs in New Zealand. Australas Plant Pathol 38:305–309

6 Wylie SJ, Jones MGK (2012) Complete genome sequences of seven carlavirus and potyvirus isolates from *Narcissus* and *Hippeastrum* plants in Australia, and proposals to clarify their naming. Arch Virol, 157:1471–1480

7 Chen CC, Chang CA, Chiang FL (2016). Serological and molecular identification of Nerine latent virus infecting blood lily (*Haemanthus multiflorus* Martyn). J Taiwan Agric Res 65:194-206.

8 Shishkin AA, Giannoukos G, Kucukural A et al (2015) Simultaneous generation of many RNA-seq libraries in a single reaction. Nat Meth 12:323–325.

9 Girardot C, Scholtalbers J, Sauer S et al (2016) Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. BMC Bioinformatics 17(1):419

10 Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. Genome Res 27(5):824-834

11 Altschul SF, Gish W, Miller W et al (1990) Basic local alignment search tool. J Mol Biol 215:403–410

12 Wheeler DL, Church DM, Federhen S et al (2003) Database resources of the National Center for biotechnology. Nucleic Acids Res 31:28-33

13 Rodriguez-R LM, Konstantinidis KT (2016) The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. PeerJ Prepr 4:e1900v1.

14 Bousalem M, Douzery EJP, Seal SE (2008) Taxonomy, molecular phylogeny and evolution of plant reverse transcribing viruses (family Caulimoviridae) inferred from full-length genome and reverse transcriptase sequences. Arch Virol 153:1085-1102

15 Klaassen VA, Boeshore ML, Koonin EV et al (1995) Genome structure and phylogenetic analysis of lettuce infectious yellows virus, a whitefly-transmitted, bipartite closterovirus. Virology 208:99-110

16 Adams MJ, Antoniw JF, Bar-Joseph M et al (2004) The new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation. Arch Virol 149:1045–1060

17 Adams MJ, Antoniw JF, Fauquet CM (2005) Molecular criteria for genus and species discrimination within the family *Potyviridae.* Arch Virol 150:459–479

18 Dong JH, Cheng XF, Yin YY et al (2008) Characterization of tomato zonate spot virus, a new tospovirus in China. Arch Virol 153:855–864

19 Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

20 Stecher G, Tamura K, Kumar S (2020) Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Mol Biol Evol 37(4):1237–1239

21 Hay JM, Jones MC, Blakebrough ML et al (1991) An analysis of the sequence of an infectious clone of rice tungro bacilliform virus, a plant pararetrovirus. Nucleic Acids Res 19:2615-2621

22 Hull R (1996) Molecular biology of rice tungro viruses. Annu Rev Phytopathol 34:275–97

23 Melzer M, Ayin C, Sugano J et al (2013) Differentiation and distribution of cordyline viruses 1–4 in Hawaiian ti plants (*Cordyline fruticosa* L.). Viruses 5:1655-1663

24 Wang H, Zhao R, Zhang H et al (2020) Prevalence of yellow leaf disease (YLD) and its associated areca palm velarivirus 1 (APV1) in betel palm (*Areca catechu*) plantations in Hainan, China. Plant Dis 104(10):2556-2562

25 Nordal I, Duncan T (1984) A cladistic analysis of *Haemanthus* and *Scadoxus*. Nord J Bot 4:145-153

26 Gamage SW, Persley DM, Higgins CM, Dietzgen RG (2015) First complete genome sequence of a capsicum chlorosis tospovirus isolate from Australia with an unusually large S RNA intergenic region. Arch Virol 160:869–872

27 Dehaan P, Wagemakers L, Peters D, Goldbach R (1989) Molecular-cloning and terminal sequence determination of the S RNA and M RNA of Tomato spotted wilt virus. J Gen Virol 70:3469–3473

28 Alexandre MAV, Duarte LML, Rivas EB et al (2011) Hippeastrum mosaic virus diagnosed in *Hippeastrum* and *Eucharis* in Brazil. J Plant Pathol 93:643-649

29 Xu XH, Tang W, Gao R et al (2017) First Report of hippeastrum mosaic virus in *Hippeastrum* spp. in Mainland China. Plant Dis 101:1064

30 Ágoston J, Almási A, Nemes K et al (2020) First report of hippeastrum mosaic virus, narcissus late season yellows virus, narcissus latent virus and narcissus mosaic virus in daffodils from Hungary. J Plant Pathol 102:1275–1276

31 Wright AA, Cross AR, Harper SJ (2020) A bushel of viruses: Identification of seventeen novel putative viruses by RNA-seq in six apple trees. PLoS ONE 15: e0227669

32 Hull R, Rima B (2020) Virus taxonomy and classification: naming of virus species. Arch Virol 165:2733–2736