Genomic characterization of two novel viruses, infecting *Barleria cristata* L. from the genera *Orthotospo-* and *Polerovirus*

David A. Read[1*], Ronel Roberts[2] and Genevieve 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

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

Barleria cristata L. has become naturalized in South Africa, where it commonly used as an ornamental. In 2019, plants of *B. cristata* showing putative viral symptoms were collected from two locations in Gauteng, South Africa. RNAtag-seq libraries were prepared and sequenced using an Illumina HiSeq 2500 platform. *De novo* assembly of the resulting data revealed the presence of a novel member of the *Tospoviridae* family, associated with the plants from both locations, with the tentative name of barleria chlorosis-associated virus. Segments L, M and S, have lengths of 8752, 4760 and 2906nt respectively. Additionally, one of the samples was associated with a novel polerovirus, provisionally named barleria polerovirus 1 and a complete genome length of 6096nt. This is the first study that shows the association of viruses with a species from the genus *Barleria.*

Barleria L. (family: Acanthaceae) is a diverse genus of herbaceous and shrub-like plants, with ~300 species distributed mainly within Africa and Asia [1]. *B. cristata* is a species native to large parts of southeast Asia and India but has become naturalized in many African countries, including South Africa [2]. While some species are used as ornamentals and sources of ethnomedicinal compounds [3], the economic importance of the genus is low. This has led to a corresponding lack of research of associated viruses and until now, no plant pathogenic virus has been reported for the genus.

In November 2019, two plants of *B. cristata,* one of which displayed large diffuse chlorotic spots (19-3031) and the other, severe foliar mosaic symptoms (19-3037) (Supplementary Fig. 1) were collected from two different locations in Gauteng province, South Africa. These samples are associated with accession numbers 19-3031 and 19-3037 with GPS collection coordinates of -25.751853; 28.229389 and -26.086949; 27.843592 respectively.

Samples were macerated using liquid nitrogen and total RNA was isolated using the GeneJET Plant RNA purification kit (Thermo Fischer Scientific, Vilnius, Lithuania), which was used as input for the generation of an RNAtag-seq library, according to Shishkin et al, (2015) [4]. A total of 32 samples were included in the library preparation, however 30 of these were unrelated to this study. Sequencing of the resulting library was performed using an Illumina HiSeq 2500 instrument (Illumina, San Diego, United States) at the Agricultural Research Council – Biotechnology Platform, Pretoria, South Africa. Demultiplexing of the raw reads output was performed using the Je suite [5]. Datasets associated with 19-3031 and 19-3037 were trimmed for read length, read quality and adapter content, using CLC Genomics Workbench 9 (Qiagen Bioinformatics, Aarhus, Denmark) with the following parameters. Minimum read length = 20bp; quality limit = 0.05 and trim adapters (Illumina universal adapter: 5’ AGATCGGAAGAG 3’; RNAtag-seq adapter: 5’ TACACGACGCTCTTCCGATCTNNNNNNNNT 3’). Trimming resulted in a total of 3,133,356 and 1,968,363 reads for 19-3031 and 19-3037 respectively, which are available as an NCBI SRA under BioProject number PRJNA658368.

Trimmed reads were assembled using metaSPAdes [6], with the identities of the resulting contigs being classified using blastn [7] with the virus refseq database. Only plant virus-associated contigs were selected for further analyses. Browser based blastx analyses (blast.ncbi.nlm.nih.gov) confirmed the presence of three RNA segments of a divergent member of *Orthotospovirus* associated with 19-3031 and 19-3037, as well as

a divergent polerovirus, present as a mixed infection, for 19-3037. The L, M and S orthotospovirus segments associated with 19-3031 had average coverage values of 1,243; 1,081 and 703x respectively and 163; 277 and 485x for 19-3037. The polerovirus contig had an average coverage of 248x.

The locations of putative open reading frames (ORFs) and associated amino acid sequences were determined using ORF finder [8]. Genome maps for both viruses are show in Fig. 1 with the genome organizations typical their respective families [9, 10]. The average amino acid identities (AAI) for the putative gene products of the viruses from this study and their closest relatives were determined using the AAI calculator from the enveomics collection [11]. The amino acid sequences of the N-protein of the orthotospovirus and the RNA-dependent RNA polymerase (RdRp) of the polerovirus were aligned against cognate sequences from related viruses, using BioEdit 7.2.5 [12]. Best-fit maximum likelihood phylogenies (LG and JTT models for orthotospo- and polerovirus respectively; 1000 bootstrap replicates) were generated using MEGA X [13].

Figure 1: Genome maps of the genomic segments of barleria chlorosis virus (BCaV) and barleria polerovirus 1 (BPoV1). Each genome or genome segment is shown in the positive sense and each open reading frame (ORF) is represented by a grey block with the direction of each ORF indicated. The putative function of the gene products expressed by each ORF is shown. RdRp – RNA-dependent RNA polymerase; NSm – non-structural protein (M segment); GP – glycoprotein; NSs - non-structural protein (S segment); N – nucleocapsid; P0 - suppressor of RNA silencing; P1 – protease; P1-P2 - RNA-dependent RNA polymerase; P3 - Major coat protein; P4 – movement protein; P3-P5 – minor coat protein.

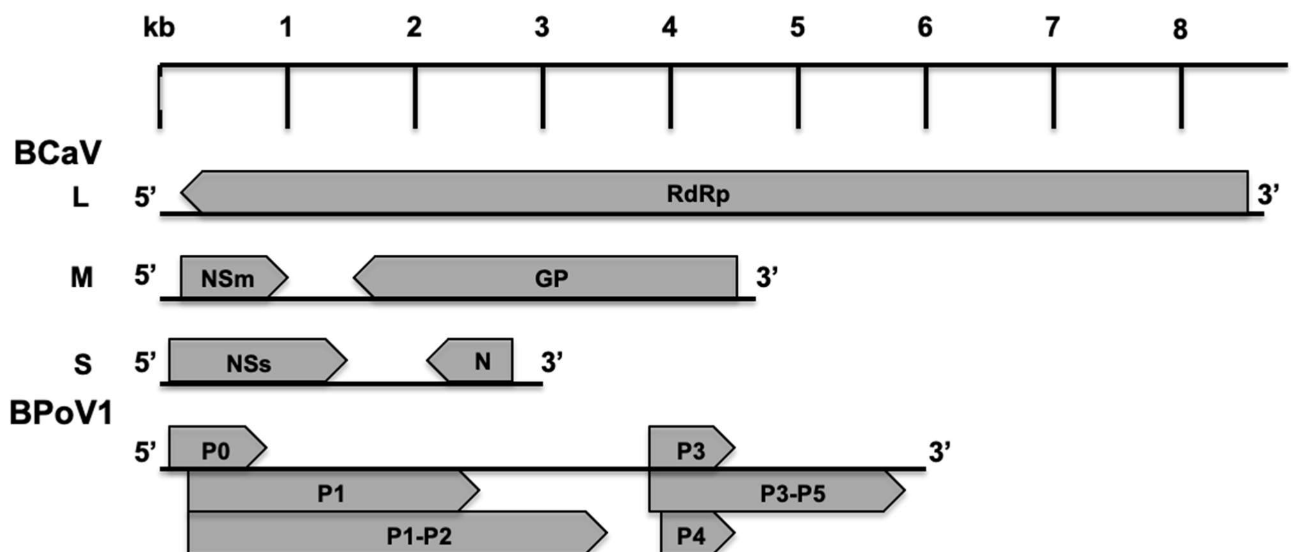

Table 1: Genome segment lengths (in nucleotides (nt)), genes, gene locations, putative gene product sizes (in amino acids (aa)), product functions and average amino acid sequence identity (AAI) shared between the viruses in this study and cognate gene products the most closely related viral strains. BCaV – Barleria chlorosis-associated virus. BPoV1 – Barleria polerovirus 1. [a] KX468767 - Polygonum ringspot virus; [b] GU591752 – Impatiens necrotic spot virus; [c] KT984753_Tomato_necrotic_spot_associated_virus; [d] NC 018071 Bean necrotic mosaic virus; [e] NC_027719_Chrysanthemum_stem_necrosis_virus; [f] KR231963 - Cucurbit aphid-borne yellows virus [g] MK050790_Luffa_aphid-borne_yellows_virus; [h] HQ388351 Brassica yellows virus; [i] NC_034247 Cowpea polerovirus 2 [j] NC_033229 Hubei polero-like virus 2 [k] HF969027 Pepper yellows virus.[*] P3a is a putative gene product that is potentially expressed from a non-AUG initiation site through leaky scanning.

Virus	Genome segment/ GenBank acc.	Length (nt)	Gene	Gene location (nt)	Product size (aa)	Putative gene function/s	MW (kDa)	pI	AAI
BSMoV	L segment MW251496	8752	RdRp	8711-111	2866	RNA-dependant-RNA-polymerase	331.7	6.5	42.5[a]
	M segment MW251497	4760	NSm	80-1027	315	Movement protein	35.5	8.7	43.9[b]
			GP	4669-1364	1101	Glycoprotein	125.9	8.3	41.5[c]
	S segment/ MW251498	2906	NSs	60-1445	461	Non-structural protein	52.4	6.6	19.1[d]
			N	2837-2079	252	Nucleocapsid	28.9	8.7	33.5[e]
BPoV1	MW251502	6096	P0	28-768	246	Suppressor of RNA silencing	27.9	9.2	23.6[f]
			P1	167-2317	716	Protease	78.7	9.1	32.8[g]
			P1-P2	167-1879:1879-3594	1063	RNA-dependent RNA polymerase	119.5	6.5	49.6[h]
			P3a[*]	3661-3795	45	Protein 3a	4.9	9.4	60[i]
			P3	3779-4384	201	Major coat protein	22.3	11.4	70.2[j]
			P3-P5	3779-5803	674	Minor coat protein (P3-P5 fusion)	74.4	8.9	54.3[f]
			P4	3813-4388	191	Putative movement protein	21.1	6.7	41.3[k]

AAI figures and phylogenetic dendrograms (Table 1; Supplementary Fig. 2 and 3) suggested that the plant viruses associated with *B. cristata* in this study are novel and the tentative names of barleria chlorosis-associated virus (BCaV) and barleria polerovirus 1 (BPoV1) have been proposed for the orthotospo- and polerovirus respectively. GenBank accessions for BCaV 19-3031 and 19-3037 segments are MW251496 – MW251498 and MW251499 – MW251501 respectively, while BPoV1 is associated with MW251502. The BCaV N-gene encoded product shares a 34% AAI with the cognate gene of chrysanthemum stem necrosis virus, the most closely related extant orthotospovirus (Supplementary Fig. 2). This is well below the demarcation threshold of 90% AAI to be considered a novel species [13]. The AAI shared between

the putative gene products for BCaV from 19-3031 and 19-3037 were all >99% with the exception of NSs (96.3%).

BPoV1 clusters with other extant members of the *Polerovirus* genus (Supplementary Fig. 3) grouping most closely to chickpea chlorotic stunt virus (CpCSV). The AAI shared between the putative gene products of BPoV1 and the cognate gene products of the most closely related viruses ranged between 23.6% (P0) – 70.2% (P3) (Table 1). BPoV1 has a putative ORF3a, such as that described for *Turnip yellows virus* [15], with a non-AUG (ACG) at 3661..3663nt and a UAG stop codon at 3796..3798nt. All AAI values were below 90%, which suggested that BPoV1 is a novel member of the *Polerovirus* genus [16]. BPoV1 has a slippery site between 1873-1879nt (GGGAAAC), which facilitates the production of the P1-P2 fusion protein (RdRp), through a -1 ribosomal frameshift [17].

The presence of both BCaV (19-3031 and 19-3037) and BPoV1 (19-3037) were confirmed using RT-PCR. Primers were designed to target a partial segment of the coat protein gene of BPoV1 with forward and reverse sequences of BPoV1-CP-F: 5' AGA CTC CCT ATA CCA GTG GTT GGT 3' and BPoV1-CP-R: 5' CAA GGA TAC CCG ATA AGT GAT CTG GA 3'. The N-gene of BCaV was targeted using BCaV -N-F: 5' TCA TTC TTG GAC CTT TAG AGC CAG T 3' and BCaV-N-R: CTA GAA TGG CTC CAC CAG TGT AC 3'. Two-step RT-PCR reactions were carried using Promega GoScript™ Reverse Transcriptase and GoTaq® Taq polymerase (Promega, Madison, WI, United States), according to manufacturer instructions. Amplicons of the expected size were visualized on an agarose gel (758 and 826bp for BCaV and BPoV1 respectively) and the identities confirmed using bidirectional Sanger sequencing (Inqaba Biotechnical Industries, Pretoria, South Africa).

The 5' terminal nucleotides of barleria polerovirus 1 was confirmed using the 5' RACE System for Rapid Amplification of cDNA Ends (RACE), version 2.0 (Invitrogen, Carlsbad, CA, United States), according the to manufacturer's specifications, using the following gene specific primers: GSP1: 5' AGA GAA TCG TCC GCA CTA 3'; GSP2: 5 'ATA CAC CTG AGA GTA CGC CCA ACA T 3'. RACE amplification was not performed for either of the two BCaV genomes, since all three segments from both 19-3031 and 19-3037 showed the expected inverted complementary [18] in the terminal nine nucleotides (5' terminus '5 AGA GCA TTC 3'). One exception was a single nucleotide mismatch within the 3' terminus of the M

segment. Segments L, M and S, of BCaV have lengths of 8752, 4760 and 2906nt respectively and the genome of BPoV1 has a length of 6096nt.

Orthotospoviruses are known for having highly variable host ranges from extremely wide, such as that of *Tomato spotted wilt tospovirus* [19], to very narrow, such as *Iris yellow spot tospovirus* [20]. Members of the *Polerovirus* genus are also known to have highly variable host ranges [21], although CpCSV, to which BPoV1 is most closely related to in a phylogenetic sense, has a narrow host range [22]. With no additional plant viruses being detected in either of the two plants, it is likely that the symptoms observed are the result of the infection of the orthotospo- and poleroviruses described here. The potential for host jumps from native or naturalized plant species to agriculturally important crops has become somewhat more predictable in recent years, when taking virus taxonomy into consideration [23]. However, the potential for broad host ranges with both genera associated with this study, warrants further research on determining whether these viruses are able to infect agriculturally important species.

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: All sequencing data related to this study has been deposited into the relevant NCBI public databases, with the accession number listed in the manuscript.

Code availability: Not applicable.

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

Consent for publication: This manuscript constitutes original research and has not been previously published nor submitted for consideration at any other journal. All authors have agreed to the submission of the manuscript.

References

[1] Balkwill MJ, Balkwill K (1998) A preliminary analysis of distribution patterns in a large, pantropical genus, Barleria L. (Acanthaceae). J Biogeogr 25:95–110

[2] Kumar H, Agrawal R, Kumar V (2018) Barleria cristata: perspective towards phytopharmacological aspects. J Pharm Pharmacol 70:475-487

[3] Kumari R, Kumar S, Kumar A, Goel KK, Dubey RC (2017) Antibacterial, antioxidant and immuno-modulatory properties in extracts of Barleria lupulina Lindl. BMC Complement Altern Med 17:484

[4] Shishkin AA, Giannoukos G, Kucukural A, Ciulla D, Busby M, Surka C, Chen J (2015) Simultaneous generation of many RNA-seq libraries in a single reaction. Nat Methods 12:323–325.

[5] Girardot C, Scholtalbers J, Sauer S, Su SY, Furlong EE (2016) Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. BMC Bioinform 17(1):419

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

[7] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410

[8] Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA, Wagner L (2003) Database resources of the National Center for Biotechnology. Nucleic Acids Res 31:28-33

[9] Nigam D, Garcia-Ruiz H (2020) Variation profile of the orthotospovirus genome. Pathogens 9:521

[10] Ghosh S, Kanakala S, Lebedev G, Kontsedalov S, Silverman D, Alon T, Mor N, Sela N, Luria N, Dombrovsky A, Mawassi M, Haviv S, Czosnek H, Ghanima M (2019) Transmission of a new polerovirus infecting pepper by the whitefly *Bemisia tabaci.* J Virol 93(15): e00488-19

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

[12] 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.

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

[14] Goldbach R, Kuo G (1996) Introduction. In: International symposium on tospoviruses and thrips of floral and vegetable crops. Acta Hortic 431:21–26

[15] Smirnova E, Firth AE, Miller WA, Scheidecker D, Brault V, Reinbold C, Rakotondrafara MA, Chung BYW, Ziegler-Graff V (2015) Discovery of a small non-AUG-initiated ORF in poleroviruses and luteoviruses that is required for long-distance movement. PLoS Pathog 11(5): e1004868

[16] Zhang P, Liu Y, Liu W, Cao M, Massart S, Wang X (2017) Identification, characterization and full-length sequence analysis of a novel polerovirus associated with wheat leaf yellowing disease. Front Microbiol 8:1689

[17] Kim YG, Maas S, Wang SC, Rich A (2000) Mutational study reveals that tertiary interactions are conserved in ribosomal frameshifting pseudoknots of two luteoviruses. RNA 6:1157–1165

[18] Oliver JE, Whitfield AE (2016) The genus *Tospovirus*: Emerging bunyaviruses that threaten food security. Annu Rev Virol 3:101-124

[19] Parrella G, Gognalons P, Gebre-Selassiè K, Vovlas C, Marchoux G (2003) An update of the host range of Tomato spotted wilt virus. J Plant Pathol 85:227-264

[20] Tsompana M, Moyer JW (2008) Tospovirus. In: BWJ Mahy and MHV Van Regenmortel. (Eds) *Encyclopedia of Virology, Third Edition.* (pp157-163). Elsevier:Amsterdam

[21] Yoshida N, Tamada T (2019) Host range and molecular analysis of Beet leaf yellowing virus, Beet western yellows virus ‐ JP and Brassica yellows virus in Japan. Plant Pathol 68, 1045–1058

[22] Abraham AD, Menzel W, Lesemann DE, Varrelmann M, Vetten HJ (2006) Chickpea chlorotic stunt virus: A new polerovirus infecting cool-season food legumes in Ethiopia. Phytopathol 96:437-446

[23] Moury B, Fabre F, Hébrard E, Froissart R (2017) Determinants of host species range in plant viruses. J Gen Virol 98:862–873