

Diversity of partial RNA-dependent RNA polymerase gene sequences of Soybean blotchy mosaic virus isolates from different host-, geographical- and temporal origins

Elrea Strydom^{a, b}

^aDepartment of Microbiology and Plant Pathology, University of Pretoria, Pretoria, 0002, South Africa;

^bForestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, 0002, South Africa.

elrea.appelgryn@fabi.up.ac.za

Gerhard Pietersen^{a, b, c}

^cGenetics Department, University of Stellenbosch, Stellenbosch, 7600, South Africa; ^aDepartment of Microbiology and Plant Pathology, University of Pretoria, Pretoria, 0002, South Africa; ^bForestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, 0002, South Africa.

Corresponding author at: Genetics Department, University of Stellenbosch, Stellenbosch, 7600, South Africa.

gerhard.pietersen@up.ac.za

+ (27) 12 420 3265

Abstract

Infection of soybean by the plant cytorhabdovirus Soybean blotchy mosaic virus results in significant yield losses in the temperate, lower-lying soybean production regions of South Africa. A 277 bp portion of the RNA-dependent RNA polymerase gene of 66 Soybean blotchy mosaic virus isolates from different hosts, geographical locations in South Africa, and time of collection spanning 16 years were amplified by RT-PCR and sequenced to investigate the genetic diversity of isolates. Phylogenetic reconstruction revealed three main lineages, designated Groups A, B and C, with isolates grouping primarily according to geographic origin. Pairwise nucleotide identities ranged between 85.7% and 100% among all isolates, with isolates in Group A exhibiting the highest degree of sequence identity, and isolates of Groups A and B being more closely related to each other than to those in Group C. This is the first study investigating the genetic diversity of SbBMV.

Keywords: Soybean blotchy mosaic virus; RNA-dependent RNA polymerase gene, maximum-likelihood analysis; pairwise nucleotide similarity

Soybean blotchy mosaic virus (SbBMV) is provisionally classified as a member of the genus *Cytorhabdovirus* within the family *Rhabdoviridae* based on viral shape, size, distribution of viral particles in plant cells and limited sequence information [1]. Members of the *Rhabdoviridae* have single-stranded or bipartite RNA genomes of negative polarity between 12 kb and 14.5 kb in size [2-4]. Rhabdovirus genomes encode six structural proteins, the nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA-dependent RNA polymerase (L). Eighteen genera are recognized within the family *Rhabdoviridae*, of which four, *Nucleorhabdovirus*, *Cytorhabdovirus*, *Dichorhavirus* and *Varicosavirus* infect plants [4]. The plant cyto- and nucleorhabdoviruses also encode one to four accessory proteins hypothesized to be involved in movement from cell to cell and other unknown functions.

Infection of soybean by SbBMV is characterized by blotchy mosaic symptoms on leaves early in the soybean production season, with symptom severity and incidence declining with time. The disease has only been detected in the lower lying soybean production areas of South Africa such as Limpopo, Mpumalanga, North West and KwaZulu-Natal [5, 6]. The virus has been shown to be mechanically transmissible to soybean and *Nicotiana benthamiana*, and the leafhopper *Peragallia carboverdensis* (*Cicadellidae*, *Agalliinae*) was identified as an insect vector of SbBMV [1], but the whole genome sequence of SbBMV still needs to be determined.

Despite the economic importance of diseases caused by plant rhabdoviruses, they remain relatively poorly studied when compared to their animal-infecting counterparts. The amount of sequence information (both partial gene sequences and whole genome sequences) generated for plant rhabdoviruses has increased in the last few years, but studies on the evolutionary relationships and variability of plant rhabdoviruses still largely rely on portions of a gene sequence or full gene sequences of a few gene regions [7-10]. The RNA-dependent RNA polymerase protein has conserved RNA-binding and polymerase domains, which are valuable for phylogenetic analysis, and consequently these are often used to infer evolutionary relationships and assess genetic diversity in plant rhabdoviruses [7, 8, 11-13].

The phylogenetic relationships and genetic diversity of SbBMV isolates have not been evaluated, and have important implications for viral diagnostics and the effective detection of all variants in a population, as well as understanding processes such as evolution and interactions between viruses and their hosts and vectors [14]. In this paper, partial sequences of the L gene of SbBMV isolates were generated by RT-PCR and Sanger sequencing and analyzed to assess the genetic diversity of SbBMV isolates from different plant and insect hosts, geographical

origins in South Africa and three general collection periods spanning 16 years. Maximum-likelihood phylogenetic reconstruction was used to infer phylogenetic relationships between isolates, and *p*-distances were calculated to assess genetic variability.

Isolates originating from plant hosts were obtained either as dried leaf material from the Virus, Antiserum and Seroreagent (PVAS) collection at the Agricultural Research Council-Plant Protection Research Institute (ARC-PPRI), in Pretoria, South Africa, or fresh leaf material displaying the typical blotchy mosaic symptoms associated with SbBMV. Samples were collected in the Brits (North West), Thabazimbi (Limpopo), Pretoria (Gauteng), Loskop Irrigation Scheme (Mpumalanga), Schoemanskloof (Mpumalanga), Dundee (KwaZulu-Natal) and Lusikisiki (Eastern Cape) areas of South Africa. Isolates obtained from asymptomatic alternative hosts of SbBMV, *Gymnosporia buxifolia*, and *Lamium amplexicaule* were also included. Each plant was assigned a unique accession number and fresh plant material samples were stored at 4°C prior to homogenization with liquid nitrogen. Plant total RNA was extracted using 1.8 ml of CTAB buffer per sample according to the method described by White and co-workers [15] modified by the omission of spermidine in the CTAB buffer, and centrifugation for 15 min. Individuals of *P. caboverdensis* were collected from soybean cultivar trials in Brits, North West in which high incidences of the disease was observed, and preserved in 100% ethanol until processed. RNA was extracted from individuals of *P. caboverdensis* using the method described by Mallory and co-workers [16] prior to RT-PCR and sequencing.

Partial L gene sequences of 66 isolates (**Table 1**) were determined. The Soyblotch F (5' CTTTGCCCAACTGGACTCCC 3') and Soyblotch R (5' TCCAAACAGTCTTCCCAGGC 3') primer pair were designed to amplify a 354 bp portion of 522 nt of the SbBMV L gene which is available in the National Centre for Biotechnology Information (NCBI) database (EU877231) [1, 17]. RT-PCR using the Soyblotch primer pair was performed as previously described [17]. Samples were submitted for Sanger sequencing using an ABI 3500xL automated sequencer at the University of Pretoria, South Africa, and sequences subjected to a Basic Local Alignment Search Tool (BLASTn) search in the NCBI database in order to confirm the identity of isolates as SbBMV. Forward and reverse raw sequence reads were assembled using BioEdit software [18], and consensus sequences were manually curated and aligned online using Mafft [19] with default parameters. Alignments were trimmed from both the 5' and 3' ends in BioEdit to minimise the influence of gaps present in alignments. Phylogenetic trees were constructed using the maximum-likelihood method based on the Tamura-Nei model [22] in MEGA software version 6 [20] with 1000 replicates of bootstrap analysis, and Northern cereal mosaic virus

Table I. Host, geographical origin, year of collection and accession information of SbBMV isolates used in this study.

Isolate	Genbank Accession number	Host	Origin	Year	Group
91/0084	MF964900	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1991	A
91/0085	MF964901	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1991	A
91/0088	MF964902	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1991	A
91/0090	MF964903	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1991	A
91/0091	MF964871	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1991	A
91/0095	MF964872	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1991	A
92/0010	MF964904	<i>Glycine max</i>	Dundee, KwaZulu-Natal	1992	A
92/0026	MF964905	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1992	A
92/0030	MF964907	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1992	A
92/0033	MF964908	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1992	A
92/0034	MF964906	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1992	C
92/0404	MF964909	<i>Glycine max</i>	Thabazimbi, Limpopo, South Africa	1992	C
94/0398	MF964873	<i>Glycine max</i>	Brits, North West, South Africa	1994	A
94/2068	MF964912	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1994	A
94/2074	MF964911	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1994	A
94/2089	MF964870	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1994	A
95/0011	MF964913	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1995	A
95/0013	MF964914	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1995	A
95/0014	MF964915	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1995	A
95/0015	MF964916	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1995	B
95/0017	MF964917	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	1995	A
95/0038	MF964918	<i>Glycine max</i>	Thabazimbi, Limpopo, South Africa	1995	C
95/0073	MF964919	<i>Glycine max</i>	Pretoria, Gauteng, South Africa	1995	C
03/4013	MF964910	<i>Glycine max</i>	Lusikisiki, Eastern Cape, South Africa	2003	B
03/4025	MF964896	<i>Glycine max</i>	Lusikisiki, Eastern Cape, South Africa	2003	B
03/4029	MF964897	<i>Glycine max</i>	Lusikisiki, Eastern Cape, South Africa	2003	B
03/4030	MF964898	<i>Glycine max</i>	Lusikisiki, Eastern Cape, South Africa	2003	B
03/4033	MF964899	<i>Glycine max</i>	Lusikisiki, Eastern Cape, South Africa	2003	B
14/3001	MF964874	<i>Glycine max</i>	Brits, North West, South Africa	2014	C
14/3002	MF964875	<i>Glycine max</i>	Brits, North West, South Africa	2014	C
14/3005	MF964876	<i>Glycine max</i>	Brits, North West, South Africa	2014	C
15/3002	MF964877	<i>Glycine max</i>	Brits, North West, South Africa	2015	C
15/3006	MF964878	<i>Glycine max</i>	Brits, North West, South Africa	2015	C
15/3011	MF964879	<i>Glycine max</i>	Brits, North West, South Africa	2015	C
15/3067	MF964880	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	2015	B
15/3069	MF964881	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	2015	C
15/3086	MF964882	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	2015	A
15/3102	MF964883	<i>Glycine max</i>	Loskop Irrigation Scheme, Mpumalanga, South Africa	2015	C
15/3137	MF964884	<i>Glycine max</i>	Schoemanskloof, Mpumalanga, South Africa	2015	A
15/3142	MF964885	<i>Glycine max</i>	Schoemanskloof, Mpumalanga, South Africa	2015	B
15/3146	MF964886	<i>Glycine max</i>	Schoemanskloof, Mpumalanga, South Africa	2015	B
15/3147	MF964887	<i>Glycine max</i>	Schoemanskloof, Mpumalanga, South Africa	2015	B
15/3155	MF964888	<i>Glycine max</i>	Schoemanskloof, Mpumalanga, South Africa	2015	B
15/3156	MF964889	<i>Glycine max</i>	Schoemanskloof, Mpumalanga, South Africa	2015	B
15/3250	MF964890	<i>Glycine max</i>	Brits, North West, South Africa	2015	C
16/4129	MF964893	<i>Glycine max</i>	Brits, North West, South Africa	2016	C
16/4131	MF964894	<i>Glycine max</i>	Brits, North West, South Africa	2016	A
16/4132	MF964920	<i>Glycine max</i>	Brits, North West, South Africa	2016	C
16/4134	MF964921	<i>Glycine max</i>	Brits, North West, South Africa	2016	C
16/4427	MF964892	<i>Gymnosporia buxifolia</i>	Brits, North West, South Africa	2016	C
16/4712	MF964891	<i>Lamium amplexicaule</i>	Brits, North West, South Africa	2016	C
17/5000	MF964922	<i>Glycine max</i>	Brits, North West, South Africa	2017	A
17/5002	MF964923	<i>Glycine max</i>	Brits, North West, South Africa	2017	C
17/5003	MF964924	<i>Glycine max</i>	Brits, North West, South Africa	2017	C
17/5004	MF964925	<i>Glycine max</i>	Brits, North West, South Africa	2017	C
17/5005	MF964926	<i>Glycine max</i>	Brits, North West, South Africa	2017	C
17/5006	MF964927	<i>Glycine max</i>	Brits, North West, South Africa	2017	C
17/5008	MF964928	<i>Glycine max</i>	Brits, North West, South Africa	2017	C
17/5501	MF964929	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5503	MF964930	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5505	MF964931	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5506	MF964932	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5507	MF964933	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5508	MF964934	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5539	MF964935	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C
17/5633	MF964895	<i>Peragallia caboverdensis</i>	Brits, North West, South Africa	2017	C

(NCMV) (NC_002251.1) as outgroup. Pairwise nucleotide distances (p -distance) for the same dataset was determined using SDT v.1.2 software [21], and is presented as colour-coded blocks.

Analysis of 277 bp of the L gene of 66 SbBMV isolates using the maximum-likelihood method is presented in **Fig 1**. Phylogenetic analysis showed segregation of isolates into three main lineages, referred to here as groups A, B and C. Group A was dominated by isolates collected in the 1990s from the Loskop Irrigation Scheme, but also contains isolates from two other geographical locations (Schoemanskloof and Dundee) and the more recently isolated accessions 15/3086 (Loskop Irrigation Scheme), 15/3137 (Schoemanskloof), 16/4131 (Brits) and 17/5000 (Brits). Clade B consists of two sub-clades mostly of isolates collected from Lusikisiki in 2003 and Schoemanskloof during 2015, with the exception of isolates 15/3067 and 95/0015, which were collected in the Loskop Irrigation Scheme. Group C is mostly composed of isolates collected from soybean in Brits, but isolates from Thabazimbi (2), Pretoria (1) and Loskop Irrigation Scheme (3), the leafhopper vector *P. caboverdensis* and the alternative plant hosts *G. buxifolia* and *L. amplexicaule* represented the rest of the clade.

A clear correlation by grouping of isolates and geographic origin was observed. Group A, the two sub-clades within Group B and Group C were each dominated by isolates from a single geographic location, which were the Loskop Irrigation Scheme, KwaZulu-Natal and Schoemanskloof, and Brits respectively. The presence of isolates from other locations within each of these clades indicated that these isolates were not geographically isolated with respect to gene flow, but rather that each region might select for specific strains or genotypes based on specific selection pressures present. As an example, the high variability and absence of isolates from Lusikisiki in Groups A and C and the limited presence of isolates from Schoemanskloof in the other groups might be the result of competitive exclusion. Enhanced interactions of isolates in Groups A and C with plant or insect hosts may lead to their apparent dominance, as also described in Lettuce necrotic yellows virus (LNYV) [23]. Soybean cultivar selection varies greatly in different regions, as does the natural vegetation, which may serve as alternative hosts, and these may act in strain selection, furthermore insect vectors often select for specific genotypes adapted to plant-vector systems [24]. Multiple genetically distinct strains of SbBMV thus exist in each location, in which one genotype is favoured, increasing their prevalence.

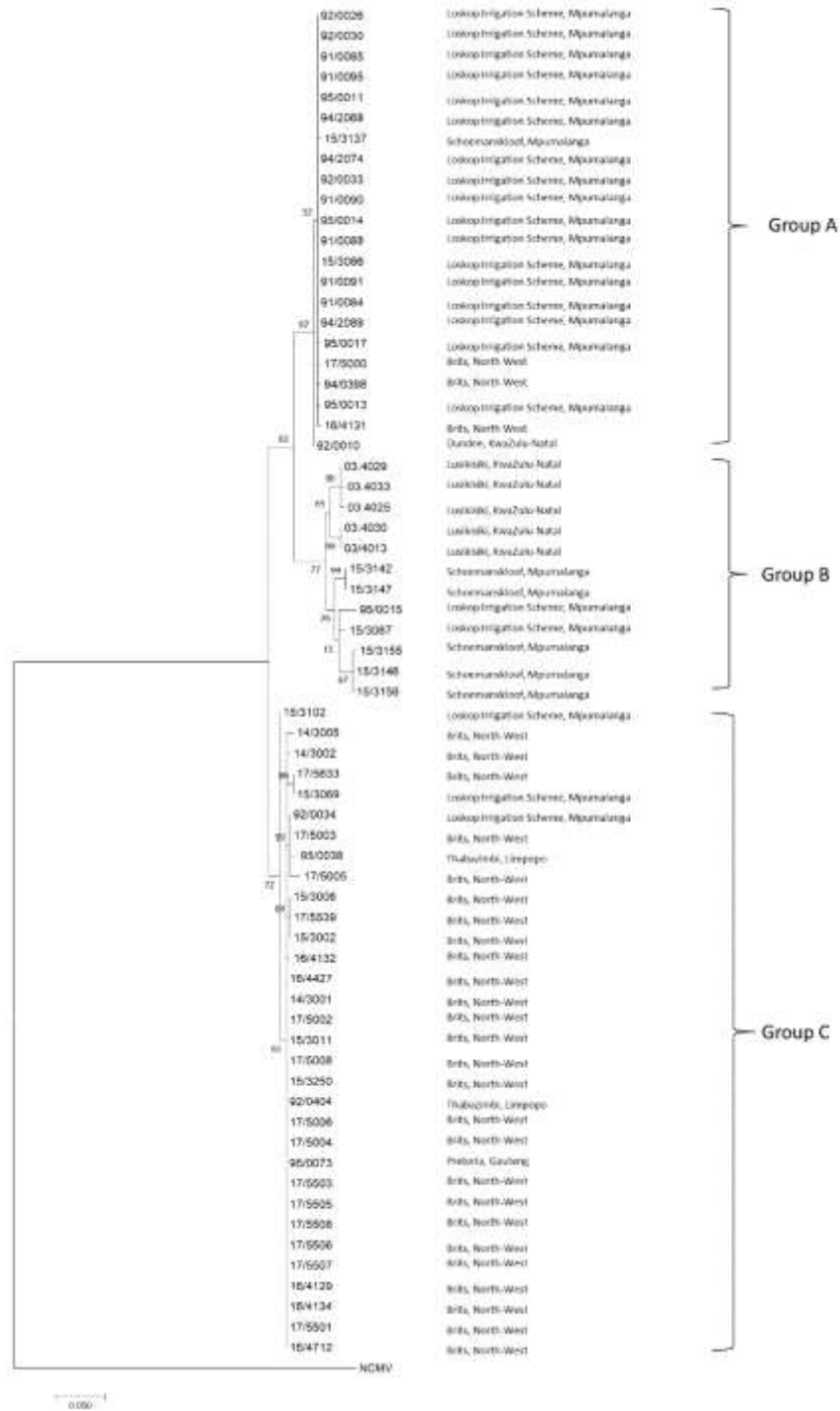


Fig 1. Maximum-likelihood phylogenetic analysis of partial L genes of 66 Soybean blotchy mosaic virus isolates from different geographical origins and hosts. Phylogenetic tree is mid-point rooted, and bootstrap percentages (1000 replicates) are indicated at nodes of each clade. Northern cereal mosaic virus (NCMV) was used as outgroup. Evolutionary analysis was conducted in MEGA 6.

Little is known of the epidemiology and variation of *P. caboverdensis*. Accession 15/3102 appears to be the founder sequence of Group C, and its association with mainly accessions from Brits suggests an introduction of SbBMV from the Loskop Irrigation Scheme into Brits in the early 1990s. The manner in which this may have occurred is unclear. Irrigation schemes are used at both Loskop and Brits, and similar crops are grown in both areas. Movement of plant material between these regions and/or enhanced vector breeding associated with high rainfall (irrigation) might have facilitated the introduction of *P. caboverdensis* into Brits. Lastly, rising temperatures can also increase the range, densities and migration potential of host plants and insect vectors [25]. Lengthened breeding seasons for *P. caboverdensis* or expansion of the geographic ranges of alternative hosts, acting as a corridor for the spread of the vector, could also be plausible.

Clustering of isolates according to geographic origin has also been reported in other plant rhabdoviruses. Investigation of more than 1kB of sequence data for each of 14 N and 19 L genes of Taro vein chlorosis virus isolates showed distinctive relationships between viral subgroups and geographical origin in the Pacific Islands [8]. This was also observed in a study which focussed on conserved regions of the L gene of 20 Eggplant mottled dwarf virus (EMDV) isolates, where phylogenetic analysis indicated two distinct genetic groups based on geographical origin [12]. Furthermore, analysis of sequences corresponding to the N, X, P, Y, G and L genes, as well as the untranslated regions of EMDV from different alternative hosts generally clustered isolates into three subgroups based on geographical origin, and not host [14].

The lack of a correlation between the clustering of isolates and host in EMDV isolates [14] was also observed in this study. Accessions isolated from different hosts such as soybean, other alternative plant hosts and the insect vector *P. caboverdensis* all grouped together in clade C, illustrating an observed absence of divergence of SbBMV in different host backgrounds. A correlation by grouping of isolates and date of collection was also absent. Although the majority of isolates originating from the early 1990's clustered together in Group A, older collection accessions were also present in Groups B and C, and similarly, isolates collected more recently present in clade A. The isolates in the two sub-clades which represent Group B were also collected more than 10 years apart. Similarly, sequencing and phylogenetic analysis of the complete N gene of eight LNYV isolates from Australia also showed no temporal separation in subgroups [10]. However, this is in contrast to later studies of LNYV which reported the

clustering of more recently collected isolates in a subgroup, with older isolates forming a distinct clade [23].

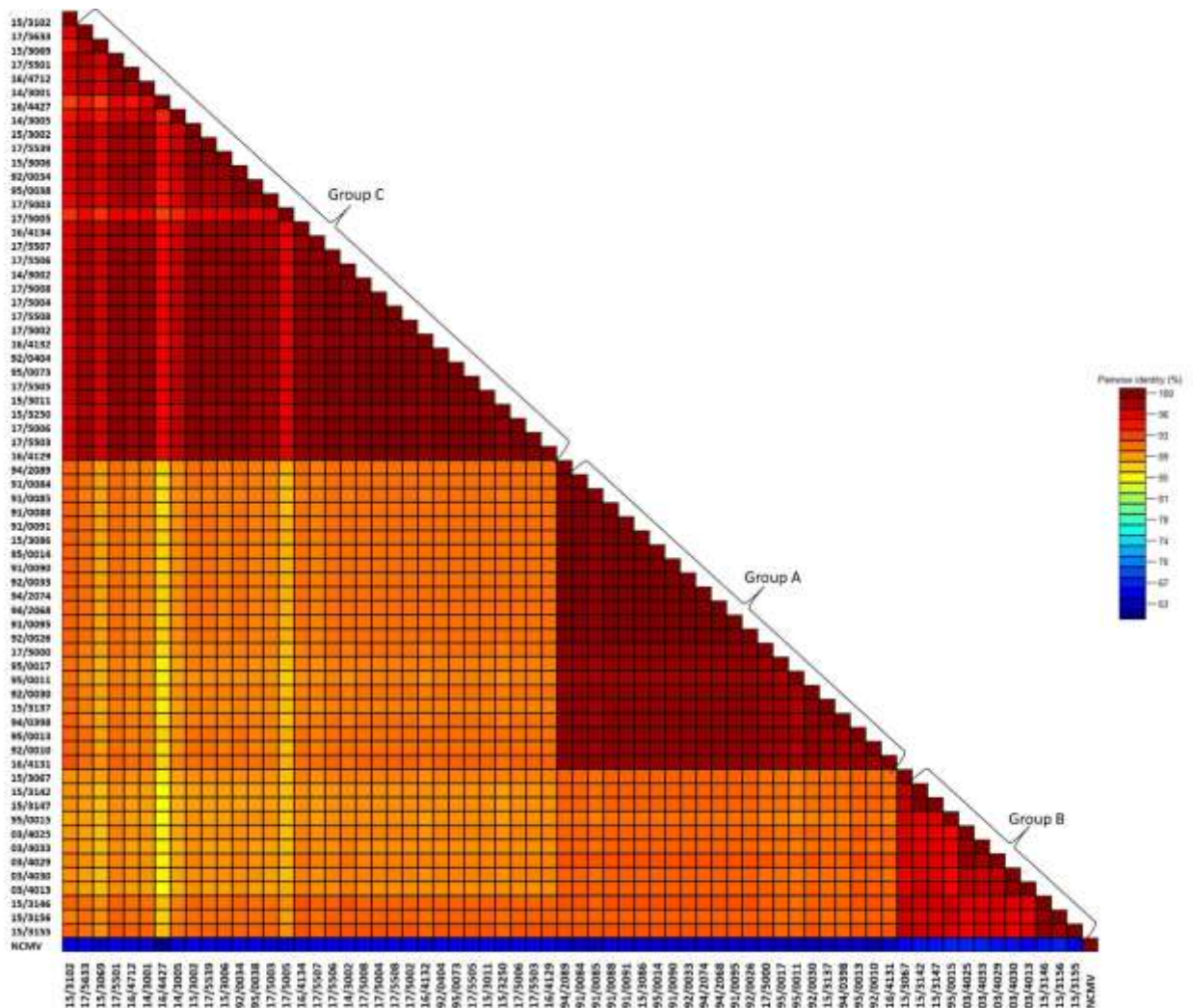


Fig 2. Graphical representation of pairwise nucleotide identities of 66 Soybean blotchy mosaic virus isolates used in this study with percentage identity scale. Pairwise nucleotide identities were obtained with SDT v1.2 software. Northern cereal mosaic virus (NC_002251.1) was used as outgroup in phylogenetic analysis, and is included here. Origin and hosts of isolates are listed in Table 1.

Pairwise nucleotide identities ranged between 98.5%-100%, 95.1%-100% and 92.7%-100% between SbBMV isolates in Groups A, B and C respectively (**Fig 2**). Between isolates of Groups A and B, the highest nucleotide similarity was 92.7%, and the lowest 90%. The minimum nucleotide identity between clade A and clade C was 86.3, and the maximum 92.1. The

minimum sequence identity between clades B and C was 85.7%, and the maximum 92.4%. Isolates within Group A showed the least amount of genetic diversity, while those in Group C showed the highest amount of sequence diversity. Clades A and B appear to more closely related to each other than to Group C, as indicated by the larger minimum and maximum nucleotide identities. It will be important to determine the whole genome of representative samples from each clade.

A previous study reported that NCMV had the highest nucleotide similarity to SbBMV (60.7%) among the cytorhabdoviruses [1], and in this study, nucleotide similarities between SbBMV isolates and NCMV varied between 57.8% and 62.8 %. Despite the use of the closest member in the genus *Cytorhabdovirus* as outgroup, a long root was still obtained, which is attributed to the diversity present in different species of the cyto- and nucleorhabdoviruses. Cyto- and nucleorhabdoviruses show higher levels of genetic diversity than other genera in the *Rhabdoviridae* such as the lyssaviruses [11], which could be the result of purifying selection in the lyssaviruses or a more recent evolution.

In this study, we report an increase in the known distribution of SbBMV, now confirmed to also be present in the Eastern Cape, furthermore we demonstrate the presence of diverse populations of SbBMV isolates for the first time. In future, additional, longer gene regions or whole genome sequences should be used to confirm evolutionary relationships inferred here, and can be used to determine whether the genetic variation observed in a portion of the L gene also translates to the full genome sequence.

Acknowledgements

Funding was provided by the Association of African Universities (AAU), the Genomics Research Institute (GRI) at the University of Pretoria, South Africa and through the National Research Foundation Incentive Grant for Rated Scientists.

References

1. Lamprecht, R.L., et al., Soybean blotchy mosaic virus, a new cytorhabdovirus found in South Africa. *Plant Dis*, 2010. **94**(11): p. 1348-1354.
2. Jackson, A.O., et al., Biology of plant rhabdoviruses. *Annu Rev Phytopathol*, 2005. **43**: p. 623-660.
3. Dietzgen, R.G., et al., The family *Rhabdoviridae*: mono- and bipartite negative-sense RNA viruses with diverse genome organization and common evolutionary origins. *Virus Res*, 2017. **227**: p. 158-170.
4. Amarasinghe, G.K., et al., Taxonomy of the order *Mononegavirales*: update 2017. *Arch Virol*, 2017. **162**(8): p. 2493-2504.
5. Pietersen, G. and H.M. Garnett, A survey for the viruses of soybeans (*Glycine max*) in the Transvaal, South Africa. *Phytophylactica*, 1990. **22**(1): p. 35-40.
6. Pietersen, G., et al., Relative abundance of soybean viruses in South Africa. *Afr Plant Prot*, 1998. **4**(2): p. 65-70.
7. Klerks, M.M.J., et al., Detection and tentative grouping of Strawberry crinkle virus isolates. *Eur J Plant Pathol* 2004. **110**: p. 45-52.
8. Revill, P., et al., Taro vein chlorosis virus: characterization and variability of a new nucleorhabdovirus. *J Gen Virol*, 2005. **86**(2): p. 491-499.
9. Talbi, C., et al., Genetic diversity of perch rhabdovirus isolates based on the nucleoprotein and glycoprotein genes. *Arch Virol*, 2011. **156**: p. 2133-2144.
10. Callaghan, B. and R.G. Dietzgen, Nucleocapsid gene variability reveals two subgroups of Lettuce necrotic yellows virus. *Arch Virol* 2005. **150**: p. 1661–1667.
11. Bourhy, H., et al., Phylogenetic relationships among rhabdoviruses inferred using the L polymerase gene. *J Gen Virol*, 2005. **86**(10): p. 2849-2858.
12. Parrella, G. and B. Greco, Sequence variation of block III segment identifies three distinct lineages within Eggplant mottled dwarf virus isolates from Italy, Spain and Greece. *Acta Virol*, 2016. **60**: p. 100-105.
13. Petrzik, K., Bioinformatic analysis of the L polymerase gene leads to discrimination of new rhabdoviruses. *J Phytopathology*, 2012. **160**(7-8): p. 377-381.
14. Pappi, P.G., et al., Genetic variation of Eggplant mottled dwarf virus from annual and perennial plant hosts. *Arch Virol*, 2016. **161**(3): p. 631-639.

15. White, E.J., et al., Modified cetyltrimethylammonium bromide method improves robustness and versatility: The benchmark for plant RNA extraction. *Biotechnol J*, 2008. **3**: p. 1424–1428.
16. Mallory, A., et al., HC-Pro suppression of transgene silencing eliminates the small RNAs but not transgene methylation or the mobile signal. *Plant Cell*, 2001. **13**(3): p. 571-583.
17. Strydom, E. and G. Pietersen, Alternative hosts and seed transmissibility of Soybean blotchy mosaic virus. *Eur J Plant Pathol*, 2017 doi:10.1007/s10658-017-1361-z.
18. Hall, T.A., BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98NT. *Nucleic Acids Symp Ser*, 1999. **41**: p. 95-98.
19. Katoh, K., et al., A novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res*, 2002. **30**(14): p. 3059-3066.
20. Tamura, K., et al., MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011. **28**: p. 2371-2739.
21. Muhire, B., et al., A genome-wide pairwise-identity-based proposal for the classification of viruses in the genus *Mastrevirus* (family *Geminiviridae*). *Arch Virol*, 2013. **158**(6): p. 1411-1424.
22. Tamura, K. and M. Nei, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol*, 1993. **10**(3): p. 512-526.
23. Higgins, C.M., et al., Diversity and evolutionary history of Lettuce necrotic yellows virus in Australia and New Zealand. *Arch Virol*, 2016. **161**(2): p. 269-277.
24. Ali, A., et al., Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus. *J Virol*, 2006. **80**(17): p. 8345-8350.
25. Canto, T., M.A. Aranda, and A. Fereres, Climate change effects on physiology and population processes of hosts and vectors that influence the spread of hemipteran-borne plant viruses. *Global Change Biol*, 2009. **15**(8): p. 1884-1894.