Characterization of viruses associated with Grapevine Leafroll disease in *Vitis* rootstocks in South Africa

Megan Harris

Submitted in partial fulfilment of the requirements for the degree of Master of Science (Microbiology)

In the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

South Africa

September 2017

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DECLARATION

I, Megan Harris, declare that the dissertation which I hereby submit for the degree of Master of Science at the University of Pretoria, is my own work and has not been previously submitted by me for a degree at this or any other tertiary institution.

_____

Signature

_____

Date

ACKNOWLEDGEMENTS

I would like to thank the following people and organisations for their valuable contributions made towards the completion of this study.

Winetech, University of Pretoria (UP), the National Research Foundation (NRF) and the Technology and Human Resources for Industry Program (THRIP) funded the study.

My supervisor, Prof. Gerhard Pietersen, for being an outstanding mentor and his comprehensive experience and knowledge of plant virology that guided this study to completion.

My parents for their strong and boundless support throughout my academic career, which has aided me to achieve this important achievement in my life.

Jennifer Wayland for her support and assistance, without which I would not have completed this achievement.

David Read, Kirsti Snyders, and Jennifer Wayland with assistance in collecting necessary material from the field.

Students in the Department of Microbiolgoy and Plant Pathology, particularly Jennifer Wayland, David Read, Gaby Carsten, Jackie Kleinhands, Erika Bruk, Kirsti Snyders and Ronel Roberts for their necessary friendship and support.

No greater thing is created suddenly, any more than a bunch of grapes or a fig. If you tell me you desire a fig, I answer you that there must be time. Let it first blossom, then bear fruit, then ripen.

- Epictetus

CONTENTS

LIST OF APPENDICES

LIST OF TABLES

Chapter 2

Chapter 3

Chapter 4

LIST OF FIGURES

Chapter 3

Figure number Page

Chapter 4

Abstract

One of the global and most economically significant viral disease of grape cultivation is grapevine-leafroll disease (GLD). Viruses associated with GLD are known as grapevine leafroll associated viruses (GLRaVs) and numbered in order of discovery, with the most important one being GLRaV-3. *Viti-* and *Foveaviruses* are often also often found in mixed infections with GLD, and speculation exists regarding a co-dependence of transmission of some Vitiviruses with GLRaV-3. As with most grapevine growing countries, many strains of GLRaVs, *Viti-* and *Foveavirus* are present in South Africa. A local certification scheme exists with the objective of providing grapevine scion and rootstock material free of these viruses to farmers. Rootstocks show no GLD symptoms thus making visual diagnostics impossible. Furthermore, it is generally accepted that GLRaV-3 is difficult to test for in rootstocks than in scions, possibly due to uneven distribution and low viral titre in rootstocks. It is however also possible that rootstocks select for specific strains of GLRaV-3 which may be less amenable to the detection technique. The difficulty in detection of GLRaV-3 could result in unwitting infection of healthy scion material by infected rootstock material during grafting. To gain greater insight into the detection of GLRaV-3 detection in rootstocks both rootstock and scion tissue were sampled separately from 95 grapevine individuals and tested using a broad spectrum PCR protocol directed at a highly conserved primer binding sequence flanking a highly variable region of the GLRaV-3 helicase gene. Amplicons of scion and rootstocks of 22 vines were subjected to Illumina next generation sequencing to determine the composition of the GLRaV-3 variant population. Poor detection (43% of samples) at low amplicon levels were obtained for GLRaV-3 in primarily R99 rootstock compared to the corresponding scions (GLRaV-3 detected in 93% of the scions samples) and corroborated the perceived poor detection of this virus in at least R99. We also determined that poor detection of GLRaV-3 is not due to the presence of a PCR-inhibitory substance in rootstocks as we obtained high levels of amplicons of *Viti-* and *Foveaviruses* associated with the rootstocks of GLD infected vines. This was achieved using universal degenerate primer nested RT-PCR combined with Illumina next generation sequencing of the resulting amplicons. Confirmation tests were performed using Vitivirus specific primers to determine the identity of the unidentified *Viti-* and *Foveaviruses* found. Direct Sanger sequencing was used on 37 rootstock

samples and 20 corresponding scion samples. *Viti-* and *Foveaviruses* were significantly less detected (61%) in rootstocks than corresponding scions (82%). The dominant component of the viral population was shown to differ between samples, with *Grapevine virus B* (GVB) and *Grapevine rupestris stem pitting associated virus* (GRSPaV) only found dominant in rootstocks. NGS data exhibited that the largest amount of the total reads belonged to GVB (44%), though no pattern could be differentiated between differing rootstocks. This is the most comprehensive study done on the viruses associated with GLD in rootstocks in South Africa, and will contribute to improve the certification scheme.

Chapter 1

Introduction

Grapevine leafroll disease (GLD) is one of the largest important global problem in grapevine growing regions and has a major economical impact on the grapevine industry (Golino*, et al.*, 2002, Charles*, et al.*, 2006, Beuve*, et al.*, 2007, Singh Brar*, et al.*, 2008, Bester*, et al.*, 2012, Tsai*, et al.*, 2012, Maree*, et al.*, 2013). The vine health and grape quality in both wine and table grapes are impacted by GLD (Tsai*, et al.*, 2012, Chooi*, et al.*, 2013, Maree*, et al.*, 2013). Symptoms can differ based on season, grape cultivar and climatic conditions (Maree*, et al.*, 2013, Naidu*, et al.*, 2015). Red cultivars show interveinal reddening of mature leaves and the veins stays green and the downward rolling of the edges of the leaves (Almeida*, et al.*, 2013). Symptoms in white cultivars are discontinuous chlorosis of mature leaves and are more difficult to note especially since only some of the white cultivars display them (Maree*, et al.*, 2013). Rootstocks like some white cultivars display no symptoms, which make visual diagnosis impossible (Martelli*, et al.*, 2012, Maree*, et al.*, 2013).

In South Africa, all commercial *Vitis* production is on grafted rootstocks. This is due to the presence of phylloxera in South Africa. *Daktulosphaira vitifoliae* commonly known as phylloxera is an aphid destructive to the roots and leaves of grapevines. It causes deformations on roots and affects water and nutrient uptake (Granett*, et al.*, 1985, Downie, 2002, Forneck & Huber, 2009). Phylloxera is indigenous to North America and was brought into the French wine industry in 1868, soon after into South Africa (Lounsbury, 1940). Many American *Vitis* species display natural resistance to phylloxera, and are therefore used as an instrument of controlling phylloxera globally by grafting *Vitis vinifera* cultivars onto rootstocks of these resistant species (Töpfer*, et al.*, 2011). Grafting healthy *V. vinifera* scion cultivars onto GLD infected rootstocks can infect the scion since GLD is graft and vector (*Pseudococcidae*) transmissible (Engelbrecht & Kasdorf, 1990, Tsai*, et al.*, 2010).

The main approach to controlling GLD is to decrease the introduction and spread in healthy vineyards. The production of virus-free plants by nurseries and screening of vineyards and removal of GLD infected vines aids in limiting the introduction and spread in vineyards. GLD control strategies therefore depend heavily on reliable

diagnostic tests, various diagnostic techniques are available for GLD diagnostics and include biological indexing, serological and molecular methods (Maree*, et al.*, 2013, Pietersen*, et al.*, 2013, Olmos*, et al.*, 2016). Absence of obvious symptoms (Alkowni*, et al.*, 2011, Al Rwahnih*, et al.*, 2012, Maree*, et al.*, 2013) and low viral titres (Chooi*, et al.*, 2016) presents problems for biological indexing and serological methods in turn inhibiting the diagnosis of GLD by the certification scheme. Misdiagnosis produces an opportunity for spread of virus to healthy material. The certification scheme in South Africa provides only the finest virus-free plant material to the wine industry (Carstens, 2002, Pietersen, 2017).

The viruses associated with GLD are known by the name “Grapevine leafroll associated viruses” followed by a number reflecting the order of discovery. The most prominent and important of these is *Grapevine leafroll associated virus 3* (GLRaV-3) (Fuchs*, et al.*, 2009, Gouveia*, et al.*, 2011, Sharma*, et al.*, 2011, Chooi*, et al.*, 2013). GLRaV-3 often occurs in mixed infections of variants worldwide (Martin*, et al.*, 2005, Turturo*, et al.*, 2005, Akbaş*, et al.*, 2007, Fiore*, et al.*, 2008, Fuchs*, et al.*, 2009, Sharma*, et al.*, 2011) including South Africa (Jooste*, et al.*, 2015). Representatives of each of these variant groups have been found in South Africa (Maree*, et al.*, 2008, Jooste*, et al.*, 2010, Bester*, et al.*, 2012, Maree*, et al.*, 2015).

GLD diseased vines are often not only infected with GLRaV-3 but are often co-infected with various *Vitiviruses*, including *Grapevine virus A* (GVA), GVB, GVE, GVF and *Grapevine rupestris stem pitting associated virus* (GRSPaV) (Martelli, 1993, Jooste*, et al.*, 2015). In addition to being closely associated with GLD both *Viti-* and *Foveaviruses* form part of the rugose wood (RW) complex. Co-infection of GLRaV-3 and *Vitiviruses* can have serious and destructive effects on some rootstocks and the severity of which depends on the rootstock genotype and the virus type (Golino*, et al.*, 2008).

The main objective of this study was gain more information regarding the infection of commercial rootstocks by GLRaVs and *Viti-* and *Foveaviruses* in South Africa, and to ascertain the reasons for poor virus detection including whether selection of poorly detected GLRaV-3 variants occurs in rootstocks. Various template preparations such as randomly amplified DNA can be sequenced using NGS (Roossinck*, et al.*, 2015), double-stranded (ds) RNA (Zablocki & Pietersen, 2014), and template amplified using gene-specific primers (Read & Pietersen, 2015). Scion and rootstock tissues of 95 vines were sampled separately from experimental and industrial farms in the

Western Cape. All 190 (95 scion and 95 rootstock) samples were tested for the presence of GLRaV-1, -2, -3, 4-like, -7, *Viti*- and *Foveaviruses*. The GLRaV-3 primers are highly conserved flanking a highly variable region in the helicase gene, allowing non-bias detection of a multitude of variants (Goszczynski, 2013). A subset of 22 amplicons of 11 vines that tested positive for GLRaV-3 in both rootstock and scion tissues were subjected to Illumina MiSeq technology. A subset of 37 rootstocks and 20 scions that tested positive for *Viti*- and *Foveaviruses* using universal primers (Dovas & Katis, 2003) were sequenced using direct Sanger sequencing to determine the dominant virus present. A subset of 19 amplicons from rootstocks only was submitted for sequencing using Illumina MiSeq technologies. To ensure the accuracy of the Illumina Miseq data analysis clones of the various amplicons were obtained and amplified. The amplicons were used as a positive control with known content allowing for optimization of data analysis parameters. In the following literature review we explore information and previously developed studies relevant to the objectives of the current study.

REFERENCES

Akbaş B., Kunter B. and Ilhan D. (2007). Occurrence and distribution of *Grapevine leafroll-associated viruses 1, 2, 3* and *7* in Turkey. Journal of Phytopathology vol.155 pp122-124

Al Rwahnih M., Osman F., Sudarshana M.*, et al.* (2012). Detection of *Grapevine leafroll-associated virus 7* using real time qRT-PCR and conventional RT-PCR. Journal of Virological Methods vol.179 pp383-389

Alkowni R., Zhang Y.-P., Rowhani A., Uyemoto J.K. and Minafra A. (2011). Biological, molecular, and serological studies of a novel strain of *Grapevine leafroll-associated virus 2*. Virus Genes vol.43 pp102-110

Almeida R.P., Daane K.M., Bell V.A., Blaisdell G.K., Cooper M.L., Herrbach E. and Pietersen G. (2013). Ecology and management of grapevine leafroll disease. Frontiers in Microbiology vol.4 pp1-13

Bester R., Maree H. and Burger J. (2012). Complete nucleotide sequence of a new strain of *Grapevine leafroll-associated virus 3* in South Africa. Archives of Virology vol.157 pp1815-1819

Bester R., Jooste A.E., Maree H.J. and Burger J.T. (2012). Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate *Grapevine leafroll-associated virus 3* variant groups I, II, III and VI. Virology Journal vol.9 pp1-12

Beuve M., Sempé L. and Lemaire O. (2007). A sensitive one-step real-time RT-PCR method for detecting *Grapevine leafroll-associated virus 2* variants in grapevine. Journal of Virological Methods vol.141 pp117-124

Carstens R. (2002). VINEYARD VIRUSSES: Leafroll Overview

Charles J., Cohen D., Walker J., Forgie S., Bell V. and Breen K. (2006). A review of *Grapevine Leafroll associated Virus type 3* (GLRaV-3) for the New Zealand wine industry. The Horticulture and Food Research Institute of New Zealand Ltd, Auckland

Chooi K.M., Cohen D. and Pearson M.N. (2013). Generic and sequence-variant specific molecular assays for the detection of the highly variable *Grapevine leafroll-associated virus 3*. Journal of Virological Methods vol.189 pp20-29

Chooi K.M., Cohen D. and Pearson M.N. (2016). Differential distribution and titre of selected *Grapevine leafroll-associated virus 3* genetic variants within grapevine rootstocks. Archives of Virology vol.161 pp1371-1375

Dovas C. and Katis N. (2003). A spot multiplex nested RT-PCR for the simultaneous and generic detection of viruses involved in the aetiology of grapevine leafroll and rugose wood of grapevine. Journal of Virological Methods vol.109 pp217-226

Downie D. (2002). Locating the sources of an invasive pest, grape phylloxera, using a mitochondrial DNA gene genealogy. Molecular Ecology vol.11 pp2013-2026

Engelbrecht D. and Kasdorf G. (1990). Transmission of grapevine leafroll disease and associated *closteroviruses* by the vine mealybug, Planococcus ficus. Phytophylactica vol.22 pp341-346

Fiore N., Prodan S., Montealegre J., Aballay E., Pino A. and Zamorano A. (2008). Survey of grapevine viruses in Chile. Journal of Plant Pathology pp125-130

Forneck A. and Huber L. (2009). (A) sexual reproduction–a review of life cycles of grape phylloxera, *Daktulosphaira vitifoliae*. Entomologia Experimentalis et Applicata vol.131 pp1-10

Fuchs M., Martinson T., Loeb G. and Hoch H. (2009). Survey for the three major leafroll disease-associated viruses in Finger Lakes vineyards in New York. Plant Disease vol.93 pp395-401

Golino D.A., Sim S.T., Gill R. and Rowhani A. (2002). California mealybugs can spread grapevine leafroll disease. California Agriculture vol.56 pp196-201

Golino D.A., Wolpert J., Sim B.J. and Aderson R.A. (2008). Virus effects on vine growth and fruit components of Cabernet Sauvignon on six rootstocks. In: *Proceedings of the second annual national viticulture research conference*

Goszczynski D.E. (2013). Brief Report of a New Highly Divergent Variant of *Grapevine leafroll-associated virus 3* (GLRaV-3). Journal of Phytopathology vol.161 pp874-879

Gouveia P., Santos M.T., Eiras-Dias J.E. and Nolasco G. (2011). Five phylogenetic groups identified in the coat protein gene of *Grapevine leafroll-associated virus 3* obtained from Portuguese grapevine varieties. Archives of Virology vol.156 pp413-420

Granett J., Timper P. and Lider L. (1985). Grape phylloxera (*Daktulosphaira vitifoliae*)(Homoptera: *Phylloxeridae*) biotypes in California. Journal of Economic Entomology vol.78 pp1463-1467

Jooste A., Maree H., Bellstedt D., Goszczynski D., Pietersen G. and Burger J. (2010). Three genetic *Grapevine leafroll-associated virus 3* variants identified from South African vineyards show high variability in their 5′ UTR. Archives of Virology vol.155 pp1997-2006

Jooste A.E., Molenaar N., Maree H.J., Bester R., Morey L., de Koker W.C. and Burger J.T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. European Journal of Plant Pathology vol.142 pp363-375

Lounsbury C. (1940). The pioneer period of economic entomology in South Africa. J Entomol Soc S Afr vol.3 pp9-29

Maree H., Freeborough M.-J. and Burger J. (2008). Complete nucleotide sequence of a South African isolate of *Grapevine leafroll-associated virus 3* reveals a 5′ UTR of 737 nucleotides. Archives of Virology vol.153 pp755-757

Maree H., Almeida R., Bester R.*, et al.* (2013). *Grapevine leafroll-associated virus 3*. Frontiers in Microbiology vol.4 pp1-12

Maree H.J., Pirie M.D., Oosthuizen K., Bester R., Rees D.J.G. and Burger J.T. (2015). Phylogenomic analysis reveals deep divergence and recombination in an economically important grapevine virus. PloS one vol.10 pp1-18

Martelli G. (1993). *Graft-transmissible diseases of grapevines: handbook for detection and diagnosis*. Food & Agriculture Org.

Martelli G., Abou Ghanem-Sabanadzovic N., Agranovsky A.*, et al.* (2012). Taxonomic revision of the family *Closteroviridae* with special reference to the grapevine leafroll-associated members of the genus *Ampelovirus* and the putative species unassigned to the family. Journal of Plant Pathology vol.94 pp7-19

Martin R., Eastwell K., Wagner A., Lamprecht S. and Tzanetakis I. (2005). Survey for viruses of grapevine in Oregon and Washington. Plant Disease vol.89 pp763-766

Naidu R.A., Maree H.J. and Burger J.T. (2015). Grapevine leafroll disease and associated viruses: a unique pathosystem. Annual Review of Phytopathology vol.53 pp613-634

Olmos A., Bertolini E., Ruiz-García A.B., Martínez C., Peiró R. and Vidal E. (2016). Modeling the Accuracy of Three Detection Methods of *Grapevine leafroll-associated virus 3* During the Dormant Period Using a Bayesian Approach. Phytopathology vol.106 pp510-518

Pietersen G. (2017). Creation of healthy *Vitis* planting material in the South African wine grape certification scheme. Institute for grape and wine sciences. Date acessed: http://igws.co.za/article/fact-sheets/leafroll/8-creation-of-healthy-vitis-planting-material-in-the-south-african-wine-grape-certification-scheme

Pietersen G., Spreeth N., Oosthuizen T.*, et al.* (2013). Control of grapevine leafroll disease spread at a commercial wine estate in South Africa: a case study. American Journal of Enology and Viticulture pp269*-305

Read D.A. and Pietersen G. (2015). Genotypic diversity of *Citrus tristeza virus* within red grapefruit, in a field trial site in South Africa. European Journal of Plant Pathology vol.142 pp531-545

Roossinck M.J., Martin D.P. and Roumagnac P. (2015). Plant virus metagenomics: Advances in virus discovery. Phytopathology vol.105 pp716-727

Sharma A.M., Wang J., Duffy S.*, et al.* (2011). Occurrence of grapevine leafroll-associated virus complex in Napa Valley. PLoS One vol.6 pp1-7

Singh Brar H., Singh Z., Swinny E. and Cameron I. (2008). Girdling and grapevine leafroll associated viruses affect berry weight, colour development and accumulation of anthocyanins in 'Crimson Seedless' grapes during maturation and ripening. Plant Science vol.175 pp885-897

Töpfer R., Hausmann L., Harst M., Maul E., Zyprian E. and Eibach R. (2011). New horizons for grapevine breeding. Methods In Temperate Fruit Breeding, Fruit, Vegetable and Cereal Science and Biotechnology vol.5 pp79-100

Tsai C.-W., Rowhani A., Golino D.A., Daane K.M. and Almeida R.P. (2010). Mealybug transmission of grapevine leafroll viruses: an analysis of virus–vector specificity. Phytopathology vol.100 pp830-834

Tsai C., Daugherty M. and Almeida R. (2012). Seasonal dynamics and virus translocation of *Grapevine leafroll-associated virus 3* in grapevine cultivars. Plant Pathology vol.61 pp977-985

Turturo C., Saldarelli P., Yafeng D., Digiaro M., Minafra A., Savino V. and Martelli G. (2005). Genetic variability and population structure of *Grapevine leafroll-associated virus 3* isolates. Journal of General Virology vol.86 pp217-224

Zablocki O. and Pietersen G. (2014). Characterization of a novel *Citrus tristeza virus* genotype within three cross-protecting source GFMS12 sub-isolates in South Africa by means of Illumina sequencing. Archives of Virology vol.159 pp2133-2139

Chapter 2

Literature Review

(2.1) Biology, pathology and epidemiology of Grapevine-leafroll disease

Grapevine-leafroll disease (GLD) has presented a major obstacle to grape farming, and is regarded as one of the most economically important viral diseases (Golino*, et al.*, 2002, Coetzee*, et al.*, 2010, Tsai*, et al.*, 2012). GLD is a global problem in grapevine growing regions (Golino*, et al.*, 2002, Charles*, et al.*, 2006, Beuve*, et al.*, 2007, Singh Brar*, et al.*, 2008, Bester*, et al.*, 2012, Tsai*, et al.*, 2012, Maree*, et al.*, 2013). It is known that GLD has a rapid rate of spreading due to vectors and grafting (Golino*, et al.*, 2002, Pietersen, 2004, Charles*, et al.*, 2006, Beuve*, et al.*, 2007). GLD not only impacts the vine health but also the grape quality of both wine and table grapes (Tsai*, et al.*, 2012, Chooi*, et al.*, 2013, Maree*, et al.*, 2013). Another unwanted effect of this disease is graft incompatibility and young vine failure (Beuve*, et al.*, 2007).

The disease is affiliated with a variety of viruses called Grapevine leafroll associated viruses (GLRaV); these include GLRaV 1-9 and a group viruses newly described GLRaV-De and GLRaV-Car. All of these belongs to the family of *Closteroviridae,* but are classified in different genera. The majority of viruses mentioned above all belong to *Ampelovirus* genus, with the exception of GLRaV-2, which is part of the *Closterovirus* genus, and also GLRaV-7 which is a putative member of *Velarivirus* (Ling*, et al.*, 2004, Martelli*, et al.*, 2012, Maree*, et al.*, 2013). The most prevalent of the leafroll associated viruses is GLRaV-3 (Almeida*, et al.*, 2013, Maree*, et al.*, 2013). A variety of viruses were tested for on commercial wine and rootstock cultivars in the Western Cape in 1970. A high composition of viruses in symptomatic vines was observed, 68.9% of the viruses detected were represented by leafroll (Jooste*, et al.*, 2011). It is still debatable whether a single infection of GLRaV or a mixture of GLRaVs is required to cause GLD (Naidu*, et al.*, 2015)

GLRaV-3 is the most prominent and severe strain of GLRaVs and is a part of the genus *Ampelovirus*, and is a flexious particle of approximately 1800nm long. Its

genome is 17,919 nucleotides of positive-sense single stranded RNA, and is organized into 13 open reading frames (Group VI variants only has 12 ORFs) (Golino*, et al.*, 2002, Jooste*, et al.*, 2015). This virus is known to occur in multiple variants that can be classified into super groups and further divided into groups as can be seen in *Table1* (Maree*, et al.*, 2015). These variants (*Table 1*) can occur in mixed and single infections. Complete nucleotide sequences have been identified in South Africa for variants for each group, these variants include 621 (Group I), 623 and GP18 (Group II), PL-20 (Group III), GH30 and GH11 (Group VI), and GH24 (VII).

Symptoms of GLD are known to vary with season, grape cultivar, and climatic conditions (Naidu*, et al.*, 2015). Red fruited cultivars show more dramatic symptoms than white fruited cultivars, in addition various GLRaVs induce different severities of symptoms in grapevines (Al Rwahnih*, et al.*, 2012, Martelli*, et al.*, 2012). Rootstocks are different *Vitis* species to that of the scion, and show no symptoms. (Maree*, et al.*, 2013, Teubes, 2014). This could present a problem since the rootstocks can function as a reservoir to infect other cultures of *V. vinifera*. Research has shown that using infected rootstocks results in an increase in occurrence of the virus, but also depends on the number of consecutive propagations that had been performed on the particular rootstock (Cowham, 2004). Infected vines could show delayed bud break and shoot development (Martelli*, et al.*, 2012). Symptoms are first observed during early – mid-summer and on water stressed plants (Maree*, et al.*, 2013), they increase and progress until late autumn. Autumn is therefore when the symptoms are most prominent.

In lighter red cultivars the GLD causes the discontinuous reddening of mature leaves with veins remaining green (Almeida*, et al.*, 2013). In the more deep pigmented red cultivars the reddening of the mature leaves is uniform devoid of green veins (Martelli*, et al.*, 2012, Almeida*, et al.*, 2013, Maree*, et al.*, 2013). White cultivars may present chlorosis at interveinal areas, but may be difficult to note (Maree*, et al.*, 2013). Later in autumn both red and white cultivars display downward rolling of the leaf borders, although the extent of this varies between diseased cultivars (Martelli*, et al.*, 2002). These symptoms increase with the progress of the growing season by increasing in titre from the base to the shoot and then to the tip of the vine's shoots.

Table1: Complete Grapevine leafroll associated virus 3 genomes

Isolate	GenBank accession	Country	*Vitis vinifera* cultivar	Genome size (nt)	Group	Super Group	reference
3138-07	JX559645	Canada	*Vitis vinifera*	18498			(Maree*, et al.*, 2015)
621	GQ352631	South Africa	Cabernet Sauvignon	18498			(Jooste*, et al.*, 2010)
CL-766	EU344893	Chile	Merlot	17919*	I		(Engel*, et al.*, 2008)
NY-1	NC_004667	USA	Pinot Noir	17919*			(Ling*, et al.*, 2004)
WA-MR	GU983863	USA	Merlot	18498		A	(Jarugula*, et al.*, 2010)
623	GQ352632	South Africa	Ruby Cabernet	18498	II		(Jooste*, et al.*, 2010)
GP18	EU259806	South Africa	Cabernet Sauvignon	18498			(Maree*, et al.*, 2008)
LN	JQ423939	China	Venus Seedless	18563	III		(Fei*, et al.*, 2013)
PL-20	GQ352633	South Africa	Cabernet Sauvignon	18433			(Jooste*, et al.*, 2010)
CA7246	JQ796828	USA	Merlot	18552			(Seah*, et al.*, 2012)
GH11	JQ655295	South Africa	Cabernet	18671	VI	B	(Bester*, et al.*, 2012)
GH30	JQ655296	South Africa	Cabernet	18576			(Bester*, et al.*, 2012)
GH24	KM58745	South Africa	Cabernet Sauvignon	18647	VII	C	(Maree*, et al.*, 2015)
139	JX266782	Australia	Sauvignon Blanc	18475	ND	ND	(Rast*, et al.*, 2012)

(*) = Genome not fully sequenced

The physiological effects of infected vines is the disruption of the phloem by the destruction of cells in leaves and petioles, this causes the accumulation of starch in the leaves. This may result in the halting of photosynthetic activities via a feedback mechanism (Namba*, et al.*, 1979, Charles*, et al.*, 2006, Mannini*, et al.*, 2012, Endeshaw*, et al.*, 2014, Hunter*, et al.*, 2017). Therefore similar symptoms can be observed when there is mechanical damage to the phloem, such as damage to the trunk or poor graft unions (Naidu*, et al.*, 2014, Cieniewicz*, et al.*, 2017). Effects that GLD has on the grapes include; reduced yield, cluster size, lower pH, lower brix (sugar content), and delayed ripening. These factors in turn affects the quality of wine (Over de Linden & Chamberlain, 1970, Lee & Martin, 2009, Mannini*, et al.*, 2012, Maree*, et al.*, 2013). The range of GLD symptoms could indicate distinct types of virus-host interactions.

Most plant viruses are vector-borne. A widespread means of transmission of GLRaVs is by means of contaminated plant material and mealybugs (*Pseudococcidae*) (Engelbrecht & Kasdorf, 1990, Tsai*, et al.*, 2010). There are strategies in place to limit virus spreading through contaminated plant material, founded on the manufacturing of virus-free propagative material using certification programs (Cowham, 2004, Rowhani*, et al.*, 2005). Pathogen transfer after establishment of new healthy vineyards is mainly due to vector transmission (Daane*, et al.*, 2012). There are three types of vector transmission which includes persistent (circulative) or non-persistent (stylet-borne) manner (Hohn, 2007), though there are fewer species capable of transmitting a virus through persistent vector transmission. This mechanism also requires prolonged acquisition and inoculation access periods, but will be able to inoculate the virus for extended periods (Hogenhout*, et al.*, 2008). With the non-persistent mechanism, virus-vector specificity is not high, and a larger variety of vectors species are often able to disseminate a specific the virus. Viruses that are transmitted non-persistently requires only a few minutes after acquisition to transmit the virus into plants, after which the virus is lost by the insect (Hogenhout*, et al.*, 2008)

There is also an intermediate transmission manner that has traits between persistent and non-persistent, called semi-persistent (Martelli*, et al.*, 2002, Charles*, et al.*, 2006, Hogenhout*, et al.*, 2008, Almeida*, et al.*, 2013, Maree*, et al.*, 2013). Tsai et al., (2008) found transmission of GLRaV-3 via vine mealybug to be semi-persistent transmission. Both non- and semi-persistent viruses do not have a latent

period (time between acquisition access period and inoculation access period) (Hogenhout*, et al.*, 2008).

Various vectors are able to transmit GLRaV-1 and GLRaV-3 in a semi-persistent manner including mealybug (*Pseudococidae*) and soft scale insect (*Coccidae*) vectors (Martelli*, et al.*, 2002, Tsai*, et al.*, 2010, Daane*, et al.*, 2012, Almeida*, et al.*, 2013, Krüger & Douglas-Smit, 2013). The vine mealybug *Planococcus ficus* (Signoret) has a high abundance vector of GLRaVs in South African vineyards, and is therefore considered the most significant vector of GLRaV-3 in this country (Walton & Pringle, 2004, Walton*, et al.*, 2009, Daane*, et al.*, 2012). Some of the GLRaVs like GLRaV-2 and GLRaV-7 have no known vector and are only transmissible by infected propagation material (Tsai*, et al.*, 2010, Maliogka*, et al.*, 2015, Wistrom*, et al.*, 2017).

Commonly known as the vine mealybug, *P. ficus* is about 4mm in length; oval shaped, clearly segmented and has a slate grey pinkish colour. Furthermore the mealybug is covered in a waxy layer, with waxy hair-like appendages on the edges of the body. There are thin dark stripes without wax along the back, leaving the midline darker in colour than the remainder of the body. *P. ficus* is found throughout summer and spring feeding on the root, trunk, cordon, cane, leaves and fruit of the vines (Walton & Pringle, 2004, Daane*, et al.*, 2012). The most important concern of *P. ficus* is the transmission of GLRaVs (Daane*, et al.*, 2012). During the season the feeding dynamics of the mealybugs change. In summer mealybugs focus their feeding on the grape bunches. During winter months there is no grapes or leaves to feed on, therefore they can be found on the stems and roots (Walton & Pringle, 2004, Bell*, et al.*, 2009).

(2.2) GLRaVs within hosts

It has been found that *Ampeloviruses* inhabit a great variety of plant taxa, but GLRaVs seems to be restricted to grapevines (*Vitis* spp.). The focus on commercial spread in *Vitis vinifera* could have caused the limited knowledge of the potential host range of GLRaVs (Almeida*, et al.*, 2013), though rootstocks have been observed to test positive for GLRaV-1, -2, and -3 (Alkowni*, et al.*, 1998, Kominek & Holleinova, 2003, Mslmanieh*, et al.*, 2006).

Leafroll is erratically distributed throughout the plant resulting in some parts of the plant having higher titre than other parts (Quinlan & Weaver, 1970, Cohen*, et al.*, 2003, Beccavin*, et al.*, 2009, Chooi*, et al.*, 2016). Grafted vines are constantly more infected than self-rooted vines, and the difference in infection are more prominent for viruses that are latent in rootstocks such as GLRaVs (Alkowni*, et al.*, 1998). GLRaV-3 is more challenging to detect in *V. rupestris* "St. George", 1103 Paulsen, 125AA (*V. berlandieri x V. riparia*), Kober 5BB (*V. berlandieri x V. riparia*), and Richter 110, that are a symptomatic, utilizing ELISA than in *V. vinifera* (Credi, 1997).

Tsai et al 2012 found also that the virus titre increased from the basal leaves to the apical leaves but the titre decreased approaching the dormant season using RT-qPCR (Tsai*, et al.*, 2012). The results found in Tsai et al 2012 study was only applicable to the *Vitis vinifera* samples that they used, furthermore they suggested that these results vary between methods used. Results demonstrated that the virus translocated rapidly from the trunks to new growing shoots and leaves in the early season, the highest population was reached within two months of the growing season and therefore decreased toward the end of the year. They made the observation that the virus translocation pattern could be linked to the seasonal abundance and distribution of mealybug species (Daane*, et al.*, 2008, Daane*, et al.*, 2012, Tsai*, et al.*, 2012). RT-PCR has been observed to be more robust in the detection of GLRaV-3 than ELISA (Cid*, et al.*, 2003), though contrasting observations do exist (Cohen*, et al.*, 2003, Fiore*, et al.*, 2009).

Chooi *et al* studied distribution and titre of Group I and VI GLRaV-3 variants in 3309C (*Vitis riparia x V. rupestris*) and Schwarzmann (*Vitis. riparia x V. rupestris*) rootstocks. The RT-qPCR results suggested that not only does erratic virus distribution add to the lower detection rates, but also low virus titres. Low viral titre and viral replication of GLRaV-3 infected rootstocks were also observed by Cid *et al*. using IC-RT-PCR (Cid*, et al.*, 2003), and suggested that an unclear resistance mechanism might be responsible. They also observed that poorer detection rates occur for group VI GLRaV-3 variants in comparison to the detection of group I variants. The viral titre declined from basal to the apex of the shoot. These findings corroborates observations made by Cohen *et al* when GLRaV-3 was only detected in trunks of Richter 110 using both ELISA and RT-PCR (Cohen*, et al.*, 2003). The observations of GLRaV-3 viral titre in rootstocks made by Chooi *et al*. is in contrast to what Tsai *et al.* observed in *Vitis vinifera* (Tsai*, et al.*, 2012, Chooi*, et al.*, 2016).

Localization of GLRaV-3 in grapevine micrografts was investigated by Hoa *et al.* utilizing virus immunolocalization with CP specific antibodies. They observed that the time it took for GLRaV-3 to infect micrograft conjunctions was shorter in virus-infected scion/healthy rootstock micrografts than in healthy scion/virus-infected rootstock micrografts (Hao*, et al.*, 2017). This suggests the possibility that scions infect rootstocks faster than rootstocks can infect scions. On a cellular level short distance translocation of virus occurs through the plamodesmata i.e. cell-to-cell movement, whereas long distance translocation of viruses are by means of sieve elements such as phloem (Harries & Ding, 2011). It was observed that viruses of the *Closterovirus* genus (including GLRaV-2) need a specialized movement device for translocation, and that this is possibly due to the large size of the virus, furthermore it could imply that that these viruses require additional energy for movement (Alzhanova*, et al.*, 2001). A cytopathology study was done by Cid *et al*. and observed that in Cabernet franc (*V. vinifera*) the viral particles were bundled together, but not in rootstocks.

(2.3) Rootstocks resistance/tolerance

Grapevines were always grown on their own roots, as is the case with most fruit crops, until 1860s. This is when phylloxera, an aphid that attacks roots of *Vitis* spp. with devastating effects, was unintentionally imported from Northern Americas to the vineyards of Western Europe. This introduction resulted in mass deaths of grapevines (Meng*, et al.*, 2006, Reisch*, et al.*, 2012). Rootstocks were found to be an effective and immediate way to control phylloxera. Wild vines form North America were used as rootstocks since they would be most likely to have a resistance to the insect, because of the evolution with phylloxera pressure. *Vitis riparia* and *V. rupestris* were the prevailing cuttings used to provide resistance to *Phylloxera*. Later entry of *V. cinerea* var. Helleri (*V.berlandieri*) vines took place because of the synergy between resistance to phylloxera and required adaption to calcareous soils (Walker & Stirling, 2008, Reisch*, et al.*, 2012). Breeding of rootstocks is centralized around improving resistance to soil-borne pests and diseases, but also for environmental adaption while keeping protection against *Phylloxera* and ease of propagation (Reisch*, et al.*, 2012).

The terms resistance and tolerance should not be used interchangeably, because each describes unique virus-host relationships they should be differentiated. Resistance refers to the ability of a plant to interfere with the disease cycle within the plant and thereby limiting virus replication (Fraile & García-Arenal, 2010, Oliver & Fuchs, 2011). Furthermore there are different degrees of resistance, the extremes exist of completely resistant and completely susceptible and those in between the extremes are known as more resistant and more susceptible. Completely resistant, also known as immune plants are incapable of sustaining virus replication, thus being unable to be infected and also show no symptoms. Plants that do not impair pathogen infection are referred to as completely susceptible (Oliver & Fuchs, 2011). More resistant plants are more able to impair virus replication and may produce larger crop yields in the presence of the virus than more vulnerable plants will be able.

Tolerance is defined as the capability of plants to decrease the damage caused by the virus infection to generate a good crop in the presence of a virus (Lecoq*, et al.*, 2004, Fraile & García-Arenal, 2010). As is the case with resistance, different degrees of tolerance can be differentiated as more tolerant, less tolerant, completely tolerant and completely intolerant or susceptible (Oliver & Fuchs, 2011). Plants that are able to produce a better crop when infected are said to be more tolerant than less tolerant plants. When the crop in unaffected in the presence of the pathogen it is completely tolerant to the virus. Completely intolerant or susceptible plants are not able to produce a crop in the presence of the virus (Oliver & Fuchs, 2011).

It should be noted that these definitions can also be utilized to describe the ability of the plant to hinder vector multiplication, or to evade the damaging influence on yield in the presence of the vector. This can complicate conclusions that are drawn, since the resistance/tolerance to a vector can appear to be resistance/tolerance to the virus (Oliver & Fuchs, 2011).

The rootstock interaction with the scion is determined by the viral status of the scion (Rowhani*, et al.*, 2005, Reisch*, et al.*, 2012). Death and vine decline can occur if a rootstock is sensitive to a scion infected with particular virus disease (Reisch*, et al.*, 2012). Rootstock responses to virus diseases can vary from highly tolerant to highly intolerant (Reisch*, et al.*, 2012). Meng et al 2006 found that in the case of grapevine *Rupestris stem pitting-associated virus* (GRSPaV) the scions showed infection with a mixture of variants, whereas corresponding rootstocks demonstrated

infection with homologous populations of the virus. It would therefore be enriching to study the resistance of rootstocks to other viruses as well.

The first rootstocks were named Jacquez, Aramon, *V. rupestris* du Lot and *V. riparia* Gloire de Montpellier (Saayman, 2009). The Aramon rootstock proved to have the opposite undesired effect in the Cape, and was speculated to be due to a more dangerous or virulent race of phylloxera that evolved in the Cape (Saayman, 2009). The rootstock usage in South Africa is ruled by 99 Richter, 110 Richter, Ramsey, and 101-14 Mgt (Southey & Jooste, 1991, Fourie & Halleen, 2004, Retief*, et al.*, 2006, Saayman, 2009). *Table 2* shows the most popular rootstocks used in South Africa, corresponding breeding, percentage distribution in South Africa in 2012, and the environmental characteristics. Jacquez used to be popular but is now very low on the list of demand due to poor resistance to phylloxera and nematodes, and problematic drought tolerance (*Table 2*). Ramsey is fourth on the list but is known to be unconducive to wine quality; this is likely due to its extreme vigour and high potassium uptake and is therefore more popular in the table grape industry (Teubes, 2014). What makes it popular is that it is broadly used in hot, irrigated areas where mass production is the fundamental objective (Saayman, 2009). In 2008 the demand shifted from 99 Richter to 110 Richter, making 99 Richter and Ramsey equally popular (Saayman, 2009).

The combinations of rootstock-scion combinations are chosen based on a number of factors, the vigor and length of vegetative cycle of the rootstocks will determine the time of harvest, quality and quantity of the harvest (*Table 2*). The area of cultivation, cultivar and their environmental factors need to be taken into consideration when considering the rootstock scion combination (Teubes, 2014).

(2.4) Certification scheme

In 1970 a survey indicated 99.7% incidence of harmful viruses in South Africa (le R Kriel, 1999). Since all GLRaV-3 are transmitted by infected materials control of GLD is achieved by provision of virus-free planting material. Re-infection can be caused by consecutive propagation on infected rootstocks, though the main cause is vector transmission by mealybugs (Krüger & Douglas-Smit, 2013). This issue of re-infection of healthy plants is not exclusive to South Africa (Charles*, et al.*, 2009).

Outstanding progress has been made in the improvement of techniques used to successfully eliminate notorious harmful viruses. Enhanced plant material is propagated by SA Plant Certification Scheme for Wine Grapes, known as the South African Vine Improvement Association (VIA) (Almeida*, et al.*, 2013), in mother blocks. It is impractical to ensure that the plant material in the mother blocks will stay virus-free, because of the sheer size. The VIA certification scheme in is one of the most advanced in the world. Constant improvement of plant material and certification of clones that meet the minimum requirements is provided by the scheme. The result is that only the finest accessible plant material be available to the wine industry. Available virus-free plant material is crucial since the only means of controlling for grapevine viruses is by permanent removal and replacement, also known as rouging.

Plants with desirable qualities like vigor, productivity and health are visually selected and propagated (Martelli, 1979, Pietersen, 2017). In order to propagate putative virus free material, plants are vegetatively grown under temperatures of 37-38ºC (Almeida*, et al.*, 2013, Pietersen, 2017). There are many advantages to heat treatment the plant grows at a higher rate than the virus, therefore reducing the particle movement to the apical regions, induced block on viral RNA synthesis, and inactivation of virus particles (Grout, 1999).

Table 2: Rating of most used rootstocks in South Africa (Malan & Meye, 1993, Saayman, 2009, Teubes, 2014)

Rootstock	Breeding	% distribution in SA 2012	Phylloxera	Nematodes	Phytophthora	Vigour	Lime	Drought
99 Richter	*V. Berlandieri* x *V. rupestris*	41.2	E-VG	G	VL	VH	G	G
101-14 Mgt	*V. Riparia* x *V. rupestris*	17.7	G	M=G	M	M	L	M-L
110 Richter	*V. Berlandieri* x *V. rupestris*	20	VG	G	M-L	H	G	VG
Ramsey	*V. Champinii*	12.6	G	E	VG-G	E	L-G	L
US 8-7	Jacquez x 99 Richter	3.9	M-G	M	VG	VH-E	M	M-G
140 Ruggeri	*V. Berlandieri* x *V. rupestris*	1.4	G	M	M	H-VH	E	E-VG
1103 Paulsen	*V. Berlandieri* x *V. rupestris*	1.7	VG	M-G	VL	VH	G	VG-E
Jacquez	*V. aestivalis* x *V. cinerea* x *V. vinifera*	0.5	VL	L	VG	M-G	M	L-M
143 B Mgt	*V. vinifera* x *V. riparia*	0.2	M	M	VG	VH-E	M	G
SO4	*V. Berlandieri* x *V. riparia*	0.3	G	VG	L	G	M-G	L

VL= very low; L = low; M = moderate; G = good; VG = very good; H = high; VH = very high; and E = exceptional

The apical meristem (0.2-0.3mm) is considered to be virus free since it is not yet differentiated into vascular tissue (Grout, 1999, Pietersen, 2017). These plants are subsequently grown up in an insect-free greenhouse as nuclear plants and tested as growth takes place for all viruses that can be tested for, therefore eliminating any plants that test positive. Nuclear plants can be defined as plant material of the maximum level of sanitation in the certification scheme (Almeida*, et al.*, 2013, Pietersen, 2017).

Tests are performed at a three to five year interval to ensure the nuclear material is virus free. Scion clones are tested by biological indexing, ELISA and ISEM, the ELISA assays test for the presence of Grapevine leafroll viruses 1, 2, 3, and Grapevine fanleaf. Grapevine fleck, Grapevine leafroll, Corky bark, Rupestris stem pitting, and Shiraz decline disease are all tested via biological indexing (Pietersen, 2004, Almeida*, et al.*, 2013, Vititec, 2014, Pietersen, 2017). Vines that are virus infected are rouged, as not to infect surrounding healthy vines. All scion cultivars are individually tested for GLRaV-1, -2 and 3 via ELISA, but the only rootstocks tested are US8-7 and 143 B Mgt, and are tested for the first season and then every three years thereafter (Pietersen, 2017). This is partly due to the fact that GLRaV-3 is erratically distributed within the rootstock, and is difficult to detect with ELISA (Cid*, et al.*, 2003, Cohen*, et al.*, 2003). Plants must test negative for all the above mentioned viruses in order to be certified as three star rating GLRaV free nuclear material (Almeida*, et al.*, 2013, Pietersen, 2017).

After grafted vines pass the greenhouse phase, they are established in foundation blocks. Foundation blocks must be on virgin soil (not previously used for grapevine propagation), and vector-free (Almeida*, et al.*, 2013). Foundation blocks need to be in isolated open field areas that are kept vector free (Pietersen, 2017). From here they are further cultivated in foundation nurseries that are also found isolated locations, this is in preparation of establishment of mother blocks Inspection for GLD symptoms of foundation blocks are performed annually (Pietersen, 2017) and vector control is practiced in foundation blocks by systemic insecticide application. Pre-selected isolated areas on farms of contracted collaborating producers are chosen for the establishment of said mother blocks (Vititec, 2014). Virgin soil is selected for the establishment of rootstock mother blocks.

(2.5) *Viti-* and *Foveaviruses*

The *Viti-* and *Foveaviruses* are a part of the recently established family of *Flexiviridae*. The viruses of this family all share a positive-strand RNA genome and the flexuous morphology of the elongated helical virions (Martelli*, et al.*, 2007). This same morphology is shared with two other plant virus families, namely *Potyviridae* and *Closteroviridae* (Martelli*, et al.*, 2002, Dolja*, et al.*, 2006). *Closteroviruses* are, as the name suggests, a genus of the *Closteroviridae* family.

Viti- and *Foveaviruses* are known to be agents of woody host diseases, and have the ability to induce modifications of the host xylem, otherwise known as stem pitting or grooving (Martelli*, et al.*, 2007, Uyemoto, 2009). An example of this modification or stem pitting is the disorder known as rugose wood complex of grapevine; this involves both the *Viti- and Foveaviruses*. *Grapevine virus A* (GVA) and *Grapevine virus B* (GVB) are the *Vitivirus* agents, and the *Foveavirus* agent involved is *Grapevine rupestris stem pitting-associated virus* (GRSPaV). *Grapevine virus D* (GVD), also a *Vitivirus*, is associated with a serious condition known as corky rugose wood that affect grapevine cultivars (Martelli*, et al.*, 2007, Uyemoto, 2009). *Vitivirus Grapevine virus E* (GVE) has no connection to the symptoms associated with stem pitting disease, though it has been isolated from a plant displaying stem pitting disease symptoms (Nakaune*, et al.*, 2008, Coetzee*, et al.*, 2010).

Four disorders contribute to the Rugose wood complex (RW), that is corky bark; Rupestris stem pitting; Kober stem grooving; and LN33 stem grooving (Bonavia*, et al.*, 1996). Rupestris stempitting and Kober stem grooving are the most regularly encountered components of RW complex (Garau*, et al.*, 1994). Symptoms of the complex include less vigour, swelling at bud of union, and a noticeable dissimilarity in diameter between scion and rootstock. The major characterizing feature is the pitting and/or grooving of stems on both or either rootstocks and scions, and can be identified by removal of cortex (Martelli, 1993, Bonavia*, et al.*, 1996). Kober stem grooving has been shown to be closely associated with GVA and Kober 5BB (*V. berladieri x V. riparia*) rootstocks are used for its biological indexing (Garau*, et al.*, 1994, Chevalier*, et al.*, 1995). *V. rupestris* St. George is used to index Rupestris stem pitting that is caused by RSPaVs (Zhang*, et al.*, 1998). Rootstocks have been observed test positive for GVA (Alkowni*, et al.*, 1998, Kominek & Holleinova, 2003), and GRSPaV (Mslmanieh*, et al.*, 2006, Giampetruzzi*, et al.*, 2015).

Corky bark disease is affiliated with a *Vitivirus* known as *Grapevine virus B* (GVB) (Bonavia*, et al.*, 1996). This disease is latent in *V. vinifera* cultivars, since it only shows symptoms after infected buds are grafted onto virus-susceptible phylloxera-resistant rootstocks. The graft union will gradually develop disorders, which includes pitting, grooving, lesions or necrosis. Vine death is a possibility, and sometimes only the scion dies, leaving the only the rootstock behind. Numerous rootstocks are symptomless; other rootstocks develop deep pits and grooves (Namba*, et al.*, 1991, Golino*, et al.*, 2013). LN33 (Couderc 1613 x *V. berlandieri*) rootstocks are utilized as indicators for corky bark (Habili*, et al.*, 1992, Martelli, 1993) that is associated with infection with GVB (Al Rwahnih*, et al.*, 2015), though other rootstocks have tested positive for GVB (Kominek & Holleinova, 2003). When GVB is present as a single infection, it causes a less severe disease, as opposed to when present with GLRaV-2. Transmission is usually facilitated by use of noncertified scion wood, and field spread (Golino*, et al.*, 2013).

Shiraz disease (SD) is a harmful disease that causes the canes of infected vines to never fully mature and lignify. This disease affects both own-rooted and grafted vines, and has been observed to affect Shiraz, Merlot, Malbec, Gamay, and Viognier cultivars in South Africa. A physiological effect includes delayed budburst, acutely influences crop production, and premature vine death. Other cultivars than the above mentioned could harbor the disease but remain symptomless (Goszczynski, 2007). SD infected plants are always observed to be infected with GVA (Goszczynski & Jooste, 2003, Habili & Randles, 2004). Shiraz disease is latent in rootstocks and can therefore unknowingly transmit the disease to healthy material by grafting (Goszczynski & Habili, 2012), though a study by Renault-Spilmont *et al.* observed Richter 99 and Richter 110 to have increased sensitivity to the disease (Renault-Spilmont*, et al.*, 2007).

REFERENCES

Al Rwahnih M., Dolja V.V., Daubert S., Koonin E.V. and Rowhani A. (2012). Genomic and biological analysis of *Grapevine leafroll-associated virus 7* reveals a possible new genus within the family *Closteroviridae*. Virus Research vol.163 pp302-309

Al Rwahnih M., Daubert S., Golino D., Islas C. and Rowhani A. (2015). Comparison of next-generation sequencing versus biological indexing for the optimal detection of viral pathogens in grapevine. Phytopathology vol.105 pp758-763

Alkowni R., Digiaro M. and Savino V. (1998). Viruses and virus diseases of grapevine in Palestine. EPPO Bulletin vol.28 pp189-195

Almeida R.P., Daane K.M., Bell V.A., Blaisdell G.K., Cooper M.L., Herrbach E. and Pietersen G. (2013). Ecology and management of grapevine leafroll disease. Frontiers in Microbiology vol.4 pp1-13

Alzhanova D.V., Napuli A.J., Creamer R. and Dolja V.V. (2001). Cell-to-cell movement and assembly of a plant closterovirus: roles for the capsid proteins and Hsp70 homolog. The EMBO Journal vol.20 pp6997-7007

Beccavin I., Beuve M., Lemaire O. and Grenan S. (2009). Detection of leafroll GLRaV-1, 2 and 3 on grapevine rootstock varieties. In: *Extended Abstracks 16th Meeting of ICVG*, Dijon, France

Bell V., Bonfiglioli R., Walker J., Lo P., Mackay J. and McGregor S. (2009). *Grapevine leafroll-associated virus 3* persistence in *Vitis vinifera* remnant roots. Journal of Plant Pathology pp527-533

Bester R., Maree H. and Burger J. (2012). Complete nucleotide sequence of a new strain of *Grapevine leafroll-associated virus 3* in South Africa. Archives of Virology vol.157 pp1815-1819

Bester R., Jooste A.E., Maree H.J. and Burger J.T. (2012). Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate *Grapevine leafroll-associated virus 3* variant groups I, II, III and VI. Virology Journal vol.9 pp1-12

Beuve M., Sempé L. and Lemaire O. (2007). A sensitive one-step real-time RT-PCR method for detecting *Grapevine leafroll-associated virus 2* variants in grapevine. Journal of Virological Methods vol.141 pp117-124

Bonavia M., Digiaro M., Boscia D., Boari A., Bottalico G., Savino V. and Martelli G. (1996). Studies on "corky rugose wood" of grapevine and on the diagnosis of *Grapevine virus B*. Vitis vol.35 pp53-58

Charles J., Froud K., van den Brink R. and Allan D. (2009). Mealybugs and the spread of *Grapevine leafroll-associated virus 3* (GLRaV-3) in a New Zealand vineyard. Australasian Plant Pathology vol.38 pp576-583

Charles J.G., Cohen D., Walker J.T.S., Forgie S.A., Bell V.A. and Breen K.C. (2006). A review of the ecology of the *Grapevine leafroll associated virus type 3* (GLRaV-3). New Zealand Plant Protection vol.59 pp330-337

Chevalier S., Greif C., Clauzel J.M., Walter B. and Fritsch C. (1995). Use of an Immunocapture-Polymerase Chain Reaction Procedure for the Detection of *Grapevine virus A* in Kober Stem Grooving-Infected Grapevines. Journal of Phytopathology vol.143 pp369-373

Chooi K.M., Cohen D. and Pearson M.N. (2013). Generic and sequence-variant specific molecular assays for the detection of the highly variable *Grapevine leafroll-associated virus 3*. Journal of Virological Methods vol.189 pp20-29

Chooi K.M., Cohen D. and Pearson M.N. (2016). Differential distribution and titre of selected *Grapevine leafroll-associated virus 3* genetic variants within grapevine rootstocks. Archives of Virology vol.161 pp1371-1375

Cid M., Cabaleiro C. and Segura A. (2003). Detection of *Grapevine leafroll-associated virus 3* in rootstocks. In: *Extended abstracts of the 14th Meeting of ICVG, Locorotondo*, Locorotondo

Cieniewicz E., Perry K. and Fuchs M. (2017). Grapevine Red Blotch: Molecular Biology of the Virus and Management of the Disease, pp303-314. In: *Grapevine Viruses: Molecular Biology, Diagnostics and Management*.Springer

Coetzee B., Maree H.J., Stephan D., Freeborough M.-J. and Burger J.T. (2010). The first complete nucleotide sequence of a *Grapevine virus E* variant. Archives of virology vol.155 pp1357-1360

Coetzee B., Freeborough M.-J., Maree H.J., Celton J.-M., Rees D.J.G. and Burger J.T. (2010). Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard. Virology vol.400 pp157-163

Cohen D., Van Den Brink R. and Habili N. (2003). Leafroll virus movement in newly infected grapevines. In: *Extended abstracts 14th Meeting ICVG, Locorotondo, Italy*

Cowham S. (2004). Growers' guide to top-working grapevines. Phylloxera and Grape Industry Board of South Australia. Date acessed: 2017/08/17. http://www.vinehealth.com.au/media/A-growers-guide-to-Topworking-Grapevines_V2008.pdf

Credi R. (1997). Characterization of grapevine rugose wood disease sources from Italy. Plant Disease vol.81 pp1288-1292

Daane K.M., Cooper M.L., Triapitsyn S.V.*, et al.* (2008). Vineyard managers and researchers seek sustainable solutions for mealybugs, a changing pest complex. California Agriculture vol.62

Daane K.M., Almeida R.P., Bell V.A.*, et al.* (2012). Biology and management of mealybugs in vineyards, pp271-307. In: *Arthropod Management in Vineyards:*.Springer,Netherlands

Dolja V.V., Kreuze J.F. and Valkonen J. (2006). Comparative and functional genomics of closteroviruses. Virus Research vol.117 pp38-51

Endeshaw S.T., Sabbatini P., Romanazzi G., Schilder A.C. and Neri D. (2014). Effects of *Grapevine leafroll associated virus 3* infection on growth, leaf gas exchange, yield and basic fruit chemistry of *Vitis vinifera* L. cv. Cabernet Franc. Scientia Horticulturae vol.170 pp228-236

Engel E.A., Girardi C., Escobar P.F., Arredondo V., Domínguez C., Pérez-Acle T. and Valenzuela P.D. (2008). Genome analysis and detection of a Chilean isolate of *Grapevine leafroll associated virus-3*. Virus Genes vol.37 pp110-118

Engelbrecht D. and Kasdorf G. (1990). Transmission of grapevine leafroll disease and associated closteroviruses by the vine mealybug, Planococcus ficus. Phytophylactica vol.22 pp341-346

Fei F., Lyu M.-D., Li J., Fan Z.-F. and Cheng Y.-Q. (2013). Complete nucleotide sequence of a Chinese isolate of *Grapevine leafroll-associated virus 3* reveals a 5′ UTR of 802 nucleotides. Virus Genes vol.46 pp182-185

Fiore N., Prodan S. and Pino A. (2009). Monitoring grapevine viruses by ELISA and RT-PCR throughout the year. Journal of Plant Pathology pp489-493

Fourie P. and Halleen F. (2004). Occurrence of grapevine trunk disease pathogens in rootstock mother plants in South Africa. Australasian Plant Pathology vol.33 pp313-315

Fraile A. and García-Arenal F. (2010). The coevolution of plants and viruses: resistance and pathogenicity. Advances in Virus Research vol.76 pp1-32

Garau R., Prota V.A., Piredda R., Boscia D. and Prota U. (1994). On the possible relationship between Kober stem grooving and *Grapevine virus A*. Vitis vol.33 pp161-163

Giampetruzzi A., Morelli M., Chiumenti M.*, et al.* (2015). Towards the definition of the absolute sanitary status of certified grapevine clones and rootstocks. In: *Proceedings of the 18th Congress of ICVG*, Ankara, Turkey

Golino D.A., Rowhani A. and Uyemoto J.K. (2013). *Grapevine virus B*iseases, pp157-173. In: *Grape Pest Management*. vol. 3343 (Bettiga LJ), University of California,Oakland, California

Golino D.A., Sim S.T., Gill R. and Rowhani A. (2002). California mealybugs can spread grapevine leafroll disease. California Agriculture vol.56 pp196-201

Goszczynski D. (2007). Single-strand conformation polymorphism (SSCP), cloning and sequencing reveal a close association between related molecular variants of *Grapevine virus A* (GVA) and Shiraz disease in South Africa. Plant Pathology vol.56 pp755-762

Goszczynski D. and Jooste A. (2003). Shiraz disease is transmitted by mealybug *Planococcus ficus* and associated with *Grapevine virus A*. In: *Extended abstracts 14th Meeting ICVG*, Locorotondo, vol.219

Goszczynski D. and Habili N. (2012). *Grapevine virus A* variants of group II associated with Shiraz disease in South Africa are present in plants affected by Australian Shiraz disease, and have also been detected in the USA. Plant Pathology vol.61 pp205-214

Grout B.W. (1999). Meristem-tip culture for propagation and virus elimination, pp115-125. In: *Plant Cell Culture Protocols*.Springer

Habili N. and Randles J. (2004). Descriptors for *Grapevine virus A*-associated syndrome in Shiraz, Merlot and Ruby Cabernet in Australia, and its similarity to Shiraz disease in South Africa (Nuredin Habili). Australian and New Zealand Grapegrower and Winemaker vol.488 pp71-74

Habili N., Krake L., Barlass M. and Rezaian M. (1992). Evaluation of biological indexing and dsRNA analysis in *Grapevine virus E*limination. Annals of Applied Biology vol.121 pp277-283

Hao X.Y., Bi W.L., Cui Z.H., Pan C., Xu Y. and Wang Q.C. (2017). Development, histological observations and *Grapevine leafroll-associated virus-3* localisation in *in vitro* grapevine micrografts. Annals of Applied Biology vol.170 pp379-390

Harries P. and Ding B. (2011). Cellular factors in plant virus movement: at the leading edge of macromolecular trafficking in plants. Virology vol.411 pp237-243

Hogenhout S.A., Ammar E.-D., Whitfield A.E. and Redinbaugh M.G. (2008). Insect vector interactions with persistently transmitted viruses. Annu. Rev. Phytopathol. vol.46 pp327-359

Hohn T. (2007). Plant virus transmission from the insect point of view. Proceedings of the National Academy of Sciences vol.104 pp17905-17906

Hunter J., Ruffner H. and Volschenk C. (2017). Starch concentrations in grapevine leaves, berries and roots and the effect of canopy management. South African Journal of Enology and Viticulture vol.16 pp35-40

Jarugula S., Gowda S., Dawson W.O. and Naidu R.A. (2010). Development of full-length infectious cDNA clone of *Grapevine leafroll-associated virus 3*. Genetic Diversity and Molecular Biology of Grapevine Leafroll-Associated Viruses pp112-154

Jooste A., Maree H., Bellstedt D., Goszczynski D., Pietersen G. and Burger J. (2010). Three genetic *Grapevine leafroll-associated virus 3* variants identified from South African vineyards show high variability in their 5′ UTR. Archives of Virology vol.155 pp1997-2006

Jooste A.E., Pietersen G. and Burger J.T. (2011). Distribution of *Grapevine leafroll associated virus-3* variants in South African vineyards. European Journal of Plant Pathology vol.131 pp371-381

Jooste A.E., Molenaar N., Maree H.J., Bester R., Morey L., de Koker W.C. and Burger J.T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. European Journal of Plant Pathology vol.142 pp363-375

Kominek P. and Holleinova V. (2003). Evaluation of sanitary status of grapevines in the Czech Republic. Plant Soil and Environment vol.49 pp63-66

Krüger K. and Douglas-Smit N. (2013). *Grapevine leafroll-associated virus 3* (GLRaV-3) transmission by three soft scale insect species (Hemiptera: Coccidae) with notes on their biology. African Entomology vol.21 pp1-8

le R Kriel G.J. (1999). Vine Improvement and Availability WineLand Media. Date acessed: 2017/08/17. http://www.wineland.co.za/vine-improvement-and-availability-of-plant-material/

Lecoq H., Moury B., Desbiez C., Palloix A. and Pitrat M. (2004). Durable virus resistance in plants through conventional approaches: a challenge. Virus Research vol.100 pp31-39

Lee J. and Martin R.R. (2009). Influence of grapevine leafroll associated viruses (GLRaV-2 and -3) on the fruit composition of Oregon Vitis vinifera L. cv. Pino noir: Phenolics. Food Chemistry vol.112 pp889-896

Ling K.-S., Zhu H.-Y. and Gonsalves D. (2004). Complete nucleotide sequence and genome organization of *Grapevine leafroll-associated virus 3*, type member of the genus *Ampelovirus*. Journal of General Virology vol.85 pp2099-2102

Malan A.P. and Meye A. (1993). Interaction between a South African population of *Xiphinema index* and different grapevine rootstocks. South African Journal of Enology and Viticulture vol.14 pp11-15

Maliogka V.I., Martelli G.P., Fuchs M. and Katis N.I. (2015). Chapter six-control of viruses infecting grapevine. Advances in Virus Research vol.91 pp175-227

Mannini F., Mollo A. and Credi R. (2012). Field performance and wine quality modification in a clone of *Vitis vinifera* cv. Dolcetto after GLRaV-3 elimination. American Journal of Enology and Viticulture vol.63 pp144-147

Maree H., Freeborough M.-J. and Burger J. (2008). Complete nucleotide sequence of a South African isolate of *Grapevine leafroll-associated virus 3* reveals a 5′ UTR of 737 nucleotides. Archives of Virology vol.153 pp755-757

Maree H., Almeida R., Bester R.*, et al.* (2013). *Grapevine leafroll-associated virus 3*. Frontiers in Microbiology vol.4 pp1-12

Maree H.J., Pirie M.D., Oosthuizen K., Bester R., Rees D.J.G. and Burger J.T. (2015). Phylogenomic analysis reveals deep divergence and recombination in an economically important grapevine virus. PloS one vol.10 pp1-18

Martelli G. (1979). Identification of virus diseases of grapevine and production of disease-free plants. Vitis vol.18 pp127-136

Martelli G. (1993). *Graft-transmissible diseases of grapevines: handbook for detection and diagnosis*. Food & Agriculture Org.

Martelli G., Agranovsky A., Bar-Joseph M.*, et al.* (2002). The family *Closteroviridae* revised. Archives of Virology vol.147 pp2039-2044

Martelli G., Abou Ghanem-Sabanadzovic N., Agranovsky A., *et al.* (2012). Taxonomic revision of the family *Closteroviridae* with special reference to the grapevine leafroll-associated members of the genus *Ampelovirus* and the putative species unassigned to the family. Journal of Plant Pathology vol.94 pp7-19

Martelli G.P., Adams M.J., Kreuze J.F. and Dolja V.V. (2007). Family *Flexiviridae*: a case study in virion and genome plasticity. Annu. Rev. Phytopathol. vol.45 pp73-100

Meng B., Rebelo A.R. and Fisher H. (2006). Genetic diversity analyses of grapevine *Rupestris stem pitting-associated virus* reveal distinct population structures in scion versus rootstock varieties. Journal of General Virology vol.87 pp1725-1733

Mslmanieh T., Digiaro M., Elbeaino T., Boscia D. and Martelli G. (2006). Viruses of grapevine in Syria. EPPO Bulletin vol.36 pp523-528

Naidu R., Rowhani A., Fuchs M., Golino D. and Martelli G.P. (2014). Grapevine leafroll: A complex viral disease affecting a high-value fruit crop. Plant Disease vol.98 pp1172-1185

Naidu R.A., Maree H.J. and Burger J.T. (2015). Grapevine leafroll disease and associated viruses: a unique pathosystem. Annual Review of Phytopathology vol.53 pp613-634

Nakaune R., Toda S., Mochizuki M. and Nakano M. (2008). Identification and characterization of a new *Vitivirus* from grapevine. Archives of Virology vol.153 pp1827-1832

Namba S., Yamashita S., Doi Y., Yora K., Terai Y. and Yano R. (1979). Grapevine leafroll virus, a possible member of closteroviruses. Japanese Journal of Phytopathology vol.45 pp497-502

Namba S., Boscia D., Azzam O., Maixner M., Hu J., Golino D. and Gonsalves D. (1991). Purification and properties of closteroviruslike particles associated with grapevine corky bark disease. Phytopathology vol.81 pp964-970

Oliver J.E. and Fuchs M. (2011). Tolerance and resistance to viruses and their vectors in *Vitis* sp.: a virologist's perspective of the literature. American Journal of Enology and Viticulture vol.62 pp438-451

Over de Linden A. and Chamberlain E. (1970). Effect of grapevine leafroll virus on vine growth and fruit yield and quality. New Zealand Journal of Agricultural Research vol.13 pp689-698

Pietersen G. (2004). Spread of grapevine leafroll disease in South Africa-a difficult but not insurmountable problem. Technical Yearbook 2004/5

Pietersen G. (2017). Creation of healthy *Vitis* planting material in the South African wine grape certification scheme. Institute for grape and wine sciences. Date acessed: http://igws.co.za/article/fact-sheets/leafroll/8-creation-of-healthy-vitis-planting-material-in-the-south-african-wine-grape-certification-scheme

Quinlan J.D. and Weaver R.J. (1970). Modification of Pattern of Photosynthate Movement within and between Shoots of *Vitis vinifera* L. Plant Physiology vol.46 pp527-530

Rast H., James D., Habili N. and Masri S. (2012). Genome organization and characterization of a novel variant of *Grapevine leafroll associated virus 3*. In: *Proceedings 17th Congress of ICVG 7–14 October*

Reisch B.I., Owens C.L. and Cousins P.S. (2012). Grape, pp225-262. In: *Fruit breeding*.Springer

Renault-Spilmont A.-S., Grenan S. and Boursiquot J.-M. (2007). Syrah decline in French vineyards: Rootstocks and Syrah clone impact, pathological and genetic studies. In: *Proceedings of the Syrah Vine Health Symposium*, University of California, Davis

Retief E., McLeod A. and Fourie P. (2006). Potential inoculum sources of *Phaeomoniella chlamydospora* in South African grapevine nurseries. European Journal of Plant Pathology vol.115 pp331-339

Rowhani A., Uyemoto J.K., Golino D.A. and Martelli G.P. (2005). Pathogen Testing and Certification of *Vitis* and *Prunus* Species. Annu. Rev. Phytopathol. vol.43 pp261-278

Saayman D. (2009). Rootstock choice: The South African experience. In: *Oral presentation at Rootstock Symposium, ASEV 60th Annual Meeting, Napa, California*, vol.23

Seah Y., Sharma A.M., Zhang S., Almeida R.P. and Duffy S. (2012). A divergent variant of *Grapevine leafroll-associated virus 3* is present in California. Virology Journal vol.9 pp235-240

Singh Brar H., Singh Z., Swinny E. and Cameron I. (2008). Girdling and grapevine leafroll associated viruses affect berry weight, colour development and accumulation of anthocyanins in 'Crimson Seedless' grapes during maturation and ripening. Plant Science vol.175 pp885-897

Southey J. and Jooste J. (1991). The effect of grapevine rootstock on the performance of *Vitis vinifera* L.(cv. Colombard) on a relatively saline soil. S. Afr. J. Enol. Vitic vol.12 pp32-40

Teubes A. (2014). History of roostocks in South Africa (Part 1). WineLand Media. Date acessed: 2017-05-04. http://www.wineland.co.za/the-history-of-rootstocks-in-south-africa-part-1/

Teubes A. (2014). History of rootstocks in South Africa (Part 3). WineLand. Date acessed: 2017-07-13. http://www.winetech.co.za/docs2014/AndrewTeubes-Part_3.pdf

Teubes A. (2014). History of rootstocks in South Africa (Part 8). WineLand Media. Date acessed: 2017-05-04. http://www.wineland.co.za/history-of-rootstocks-in-south-africa-part-8/

Tsai C.-W., Rowhani A., Golino D.A., Daane K.M. and Almeida R.P. (2010). Mealybug transmission of grapevine leafroll viruses: an analysis of virus–vector specificity. Phytopathology vol.100 pp830-834

Tsai C., Daugherty M. and Almeida R. (2012). Seasonal dynamics and virus translocation of *Grapevine leafroll-associated virus 3* in grapevine cultivars. Plant Pathology vol.61 pp977-985

Uyemoto J. (2009). Plant Virus Diseases: Fruit Trees and Grapevine. Desk Encyclopedia of Plant and Fungal Virology pp431

Vititec K. (2014). Grafted vines and scion. Date acessed: 11/08/2014. http://capevines.co.za/html/quicklinks/affiliates/kwv-vititecRootstock_and_scion.htm

Walker G. and Stirling G. (2008). Plant-parasitic nematodes in Australian viticulture: key pests, current management practices and opportunities for future improvements. Australasian Plant Pathology vol.37 pp268-278

Walton V. and Pringle K. (2004). Vine mealybug, *Planococcus ficus* (Signoret) (Hemiptera: Pseudococcidae), a key pest in South African vineyards. A review. S. Afr. J. Enol. Vitic vol.25 pp54-62

Walton V.M., Krüger K., Saccaggi D.L. and Millar I.M. (2009). A survey of scale insects (Sternorryncha: Coccoidea) occurring on table grapes in South Africa. Journal of Insect Science vol.9 pp47

Wistrom C., Blaisdell G., Wunderlich L., Botton M., Almeida R.P. and Daane K. (2017). No evidence of transmission of grapevine leafroll-associated viruses by

phylloxera (*Daktulosphaira vitifoliae*). European Journal of Plant Pathology vol.147 pp937-941

Zhang Y.-P., Uyemoto J.K., Golino D.A. and Rowhani A. (1998). Nucleotide sequence and RT-PCR detection of a virus associated with grapevine rupestris stem-pitting disease. Phytopathology vol.88 pp1231-1237

Chapter 3

Grapevine leafroll associated viruses of *Vitis* rootstocks used in South Africa

Harris, M.[1], and Pietersen, G.[1,2]

[1]FABI, Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria, 0002 South Africa.

[2]Agricultural Research Council-Plant Protection Research Institute, Pretoria, South Africa.

ABSTRACT

A number of viruses, known as grapevine leafroll associated viruses, have been associated with grapevine leafroll disease (GLD). The most predominant of these is *Grapevine leafroll associated virus 3* (GLRaV-3). A number of sequence variants of GLRaV-3 are known. In South Africa, to control phylloxera, grapevines (*Vitis vinifera*) are generally grafted onto rootstocks which generally are a different species of *Vitis*. The relative composition of GLRaV-3 variants infecting rootstocks versus that of the corresponding scions has not been investigated. In this study detection rate of GLRaVs was determined using RT-PCR, and the resulting amplicons of the vines that tested positive for GLRaV-3 in both rootstock and scion of the same vine submitted for Illumina MiSeq. Inconsistent detection of GLRaV-3 in the most predominant rootstock used in South Africa, Richter 99, was observed. Poorer detection occurred in rootstocks (43% of samples) compared to the scion (93% of samples) counterparts. Both rootstock and scion can harbour mixed infections, and the populations do not always correspond between rootstock and scion. No definitive pattern was observed concerning the difference in variants between rootstocks and scions. The dynamics of the interaction between rootstock and GLRaV-3 is not yet fully understood and more controlled studies are needed to investigate these interactions.

INTRODUCTION

Grapevine is possibly the oldest and most extensively cultivated fruit crop worldwide (Töpfer*, et al.*, 2011) and has been cultivated as early as 1688 in South Africa (Saayman, 2004). A destructive aphid-related insect, *Daktulosphaira vitifoliae* (Fitch 1855)(Forneck & Huber, 2009); family *Phylloxeridae* commonly known as phylloxera was introduced into South Africa, shortly after its 1868 introduction and destruction of the French wine industry in France (Lounsbury, 1940). Phylloxera feeds on the roots and leaves of grapevines, and in the case of *Vitis vinifera* results in deformations on roots which can girdle roots and affect nutrient and water uptake (Granett*, et al.*, 1985, Downie, 2002). A number of American *Vitis* species display natural resistance to phylloxera, and grafting of *Vitis vinifera* cultivars onto rootstocks of these species is an effective means of controlling phylloxera. This practice is now routinely performed worldwide wherever phylloxera occurs (Töpfer*, et al.*, 2011), including South Africa.

Grapevine leafroll is the most significant disease of grapevines in South Africa (Pietersen, 2004) and leads to reduced yield, sugar content, weaker growth, and reduced root growth (Over de Linden & Chamberlain, 1970, Almeida*, et al.*, 2013). GLD is caused by a complex of viruses named grapevine leafroll associated viruses (GLRaVs), and are numbered in order of discovery. GLRaV-3 is the most prevalent and important worldwide (Fuchs*, et al.*, 2009, Gouveia*, et al.*, 2011, Sharma*, et al.*, 2011, Chooi*, et al.*, 2013) also in South Africa (Pietersen, 2004, Jooste*, et al.*, 2011, Jooste*, et al.*, 2015). *Grapevine leafroll associated virus 3* (GLRaV-3) is a member of the genus *Ampelovirus*, family *Closteroviridae,* and is a positive-sense single-stranded RNA virus. GLRaV-3 is the core causal agent of the disease complex (Maree*, et al.*, 2013). Variants of GLRaV-3 are known to occur in mixed infections worldwide (Martin*, et al.*, 2005, Turturo*, et al.*, 2005, Akbaş*, et al.*, 2007, Fiore*, et al.*, 2008, Fuchs*, et al.*, 2009, Sharma*, et al.*, 2011, Jooste*, et al.*, 2015) and can be divided into groups based on phylogeny and each of these groups has representative variant South African isolates (Maree*, et al.*, 2008, Jooste*, et al.*, 2010, Bester*, et al.*, 2012, Maree*, et al.*, 2015).

Symptoms of grapevine leafroll disease (GLD) have high variability between grapevine cultivars, while rootstocks and white cultivars can display no symptoms (Rayapati*, et al.*, 2008, Maree*, et al.*, 2013, Pietersen*, et al.*, 2013). Red cultivars display red interveinal colouring with green venation, whereas a few white cultivars

exhibit subtle interveinal chlorosis, while both groups of cultivars may display downward rolling of leaf margins (Almeida*, et al.*, 2013, Maree*, et al.*, 2013). Some cultivars and species, like most white cultivars and American species rootstocks, are asymptomatic (Martelli*, et al.*, 2012, Maree*, et al.*, 2013). Grapevines have irregular distribution of GLRaV-3 along developing canes of both scions and rootstocks (Monis & Bestwick, 1996, Tsai*, et al.*, 2012, Chooi*, et al.*, 2016). In addition to irregular distribution, sequence variation and low viral titre could be a contributing factor to reduced detection rates, that in turn influences virus spread in vineyards (Pacifico*, et al.*, 2011, Tsai*, et al.*, 2012, Bester*, et al.*, 2014).

The main control strategy available to reduce the impact of GLD is to decrease the introduction and dissemination of GLRaV-3 in healthy vineyards. Virus-free plants need to be made available by nurseries and vineyards need to be screened for virus-infected plants for removal of the plant, thus aiding in lessening introduction and spread in vineyards. Reliable diagnostic tests are an integral part of the control strategy, many techniques are available for diagnostics and includes biological indexing, serological and molecular methods (Maree*, et al.*, 2013, Pietersen*, et al.*, 2013, Olmos*, et al.*, 2016). Biological indexing and serological methods present problems for the certification scheme as absence of obvious symptoms (Alkowni*, et al.*, 2011, Al Rwahnih*, et al.*, 2012, Maree*, et al.*, 2013) and occurrence in low titres (Petersen & Charles, 1997, Chooi*, et al.*, 2016), this presents an opportunity for the virus to accidently spread to healthy material. Identifying a diseased plant in an already established vineyard, presents an opportunity of rouging the plant before it can spread to adjacent plants (Pietersen*, et al.*, 2013)

For improved detection protocols it is essential to investigate the dynamics of GLRaV-3 on scion-rootstock grafted grapevine plants. In this study the detection and identification of GLRaV-3 variants, as well as detection of other GLRaVs, are compared between the scion and its corresponding rootstock it is grafted on using RT-PCR and Illumina Miseq.

METHODS AND MATERIALS

Samples and RNA isolation

Fifty samples were collected from commercial and 45 from experimental red and white-cultivar wine grape vineyards from two regions (Wellington and Stellenbosch) of the Western Cape, South Africa during 2014, 2015, and 2016. Criteria for the sampling of vines were that scions display clear GLD symptoms, and that the rootstock of these vines had sizeable lignified suckers/canes growing from the stems. Cane material was collected separately from both the scion and rootstock of 95 individual vines and labelled accordingly. Samples were processed by removing the outer bark and preparing phloem shavings of the scion and rootstock material of each vine and stored at -80 ^{o}C until utilised for RNA extraction. Total RNA extraction was conducted on 200mg phloem scrapings from each sample using a modified cetyltrimethylammonium bromide (CTAB) (2% CTAB, 2.5% PVP-40, 100mM Tris-HCL pH8, 2M NaCl, 25mM EDTA pH8 and 3% β-mercaptoethanol) method (White*, et al.*, 2008).

RT-PCR

Samples collected were tested in RT-PCR for GLRaV-1; -2; -3; -4-like and GLRaV-7 using published primers (*Table 1*). First strand synthesis was accomplished by first performing a primer annealing step with a 5µl mixture containing 2µl of total RNA, 0.7µl of 10µM reverse primer (*Table 1*), and molecular grade water, and subsequently incubated at 70ºC for 5 minutes and then chilling it for 5 min. Reverse transcription was achieved by adding 5µl of the primer annealed RNA to 2.5µl of reaction buffer (5x), 1.25 mixed dNTPs (Roche, Basel, Switzerland), 0.05µl RiboLock™ (40U/µl) (Thermo Fisher Scientific, Waltham, MA, United States), 0.13µl Moloney murine leukemia virus reverse transcriptase (MMLV) (Promega, Madison, WI, United States), and molecular water to a final volume of 15µl, and afterwards incubated at 42ºC for 60 minutes. First strand cDNA was amplified with a 25µl reaction mixture containing 2.5µl reaction buffer (x10), 2µl $MgCl_2$ (50mM), 1µl dNTP mix (10mM), 0.5µl forward primer (10µM), 0.5µl reverse primer (10µM) (*Table 1*), 0.25µl BioTaq™ Polymerase (5U/µl) (Bioline, London, United Kingdom), and molecular grade water. The cycling conditions for the PCR were as follows: 95ºC for 1 min, 40 cycles at 95ºC, appropriate annealing temperature (*Table 1*) for 15

seconds, and 72ºC for 20 seconds, and 72ºC for 10 min. A 2% agarose electrophoresis was performed to determine the presence/absence of the various amplicons.

The GLRaV-3 was detected by firstly carrying out primer annealing as mentioned above using Hel2R (*Table 1*) (Goszczynski, 2013). First strand synthesis was performed using a 15 µl reaction mixture composed of 4µl reaction buffer (5x), 1.2µl $MgCl_2$ (25mM), 2µl dNTP mixture (Roche, Basel, Switzerland), 0.1µl RiboLock™, 1µl GoScript™ reverse transcriptase (160U/µl) (Promega, Madison, WI, United States), 5µl primer annealed RNA, and molecular grade water. Amplification of the first strand cDNAs was achieved using a 25µl reaction mixture consisting of 5µl reaction buffer (5x). 2µl $MgCl_2$ (25mM), 0.5µl dNTP mixture, 0.5µl forward primer (10µM), 0.5µl reverse primer (10µM) (*Table 1*) (Goszczynski, 2013), 0.25µl GoTaq® G2 Flexi DNA Polymerase (5U/µl) (Promega, Madison, WI, United States), 0.5µl cDNA, and molecular grade water to volume. Cycling conditions used were as follows: 95ºC for 5 minutes, 40 cycles of 95ºC for 30 seconds, 52ºC for 30 seconds, 72ºC for 1 minute, and 72ºC for 10 minute. The presence/absence of the 560bp product was analysed by performing a 2% agarose electrophoresis.

Table1: Primers for detecting Grapevine leafroll associated viruses

Viral target	Primer	Sequence (5' – 3')	Product size (bp)	Annealing temp (ºC)	Reference
GLRaV-1	HSP70-417 F HSP70-737 R	GAGCGACTTGCGACTTATCGA GGTAAACGGGTGTTCTTCAATTCT	320	61	**(Osman & Rowhani, 2006)**
GLRaV-2	V2dCPf2 V2CPrl	ACGGTGTGCTATAGTGCG GCAGCTAAGTACGAATCT	515	61	**(Bertazzon & Angelini, 2004)**
GLRaV-3	Hel2F Hel2R	GGCGAAGAGTATTCGCTC CCAGAAAAGGCCTTCGTC	560	52	**(Goszczynski, 2013)**
GLRaV-4	LRAmp-F LRAmp-R	ATTTAGGTAATGTWGTRGCTAC TATCCTCAGWGAGGAARCGG	485	46	**(Abou Ghanem-Sabanadzovic*, et al.*, 2012)**
GLRaV-7	LR7 G23metF LR7 G23metR	AATGACTGTGATGTCGCTTTTAC TACCACTACCAGGAGGTTTATTCA	190	61	**(Al Rwahnih*, et al.*, 2012)**

Illumina MiSeq positive control generation

Plasmid inserts were amplified using the GLRaV-3 assay previously described (*Table 1*). The products were purified using column purification (NucleoSpin® Gel and PCR cleanup, Macherey-Nagel), and the concentration of the purified products established using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, United States).

A T/A cloning approach was performed on the quantified products using pGEM-T Easy vector (Promega, Madison, WI, United States) following the manufacturer's instructions and used to transform competent *E. coli* JM109 cells. Fifty putative recombinants were chosen using blue/white selection and alkaline lysis used in performing plasmid extractions. All plasmids were amplified using the T7 (5'– TAA TAC GAC TCA CTA TAG GG -3') and SP6 (5'- ATT TAG GTG ACA CTA TAG AA -3') vector specific primers and the same PCR conditions as those used to amplify the original plasmid inserts. Non-recombinant plasmids were identified by subjecting resulting amplicons to 2% agarose gel electrophoresis and noting size difference in the amplicon size. Amplicons of recombinant plasmids were purified by addition of 2µl of FastAP and 0.5µl ExoI enzymes (Thermo Scientific, Vilnius, Lithuania) to 19µl of amplified PCR product.

Each purified product was subjected to Sanger sequencing in a single direction, using the vector specific primer T7. Sequencing reactions contained: 1µl BigDye® Terminator mix v3.1 (Applied Biosystems, Foster City, CA, United States), 0.75µl T7 primer (2µM), and molecular grade water to a final reaction volume of 10µl. The cycle conditions of the sequencing reaction were as follows: 94ºC for 1 minute, 25 cycles of 94ºC for 10 seconds, 50ºC for 5 seconds, and 60ºC for 4 minutes.

Precipitation of the sequencing reaction was performed by adding 1µl of 125mM EDTA, 1µl of 3M NaOAc, and 25µl of 100% non-denatured ethanol and incubating for 15 min. Whereafter it is centrifuged at maximum speed at 4ºC for 30 min, followed by removal of resulting supernatant, and addition of 100µl of 70% ethanol, and centrifugation at max speed for 15 min at 4ºC. The supernatant was subsequently removed and the samples air-dried for 20 min, after all ethanol was evaporated the sample was submitted to the African Centre for Gene Technologies (ACGT), Automated Sequencing Facility, Department of Genetics, University of

Pretoria, South Africa and sequenced using ABI Prism ® 3130XL Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

Data analyses of resulting chromatograms were viewed and corrected using Chromas Lite 2.1 (Technelysium, Brisbane, Australia), after which the sequences were NCBI BLASTed to identify the virus amplicon present in the specific clone. A mixture of selected identified plasmids was amplified using the GLRaV-3 assay with conditions as described above and the resulting amplicon used as an Illumina MiSeq positive control.

Illumina MiSeq Sequencing

GLRaV-3 amplicons of rootstocks and their respective scions plants where both tissues had tested positive by PCR were selected, based on the presence of sufficient concentration of amplicon obtained from the rootstock. The required concentrations for Illumina MiSeq submission was obtained by column purifying (NucleoSpin® Gel and PCR cleanup, Macherey-Nagel) a large volume of GLRaV-3 PCR product, and eluting it in a small volume of elution buffer. The concentration of the purified products determined using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, United States). Next generation sequencing, utilizing the MiSeq Platform was completed on selected purified amplicons and the positive control. The samples were submitted for sequencing at the Agricultural Research Council (ARC), Biotechnology Platform, Pretoria, South Africa.

MiSeq data analysis

CLC Genomics Workbench 6 (Aarhus) was used to carry out all trimming and analyses of the Illumina MiSeq data sets. After the data was imported as paired end reads, adapter and quality trimming was performed using the default program settings with Nextera V2 transposase adapter sequences
(Transposase1: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG;
Transposase2: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG). Quality control was performed, after which each sample was reference mapped to the cognate region that Hel2F/Hel2R amplified of GLRaV-3 variant group representatives. The following sequences were included in the reference mapping. Group I: AF037268.2 (NY-1); GQ352631.1 (621); EU344893.1 (Cl-766). Group II: GQ352632.1 (623); EU259806.1 (GP18). Group III: GQ 352633.1 (PL-20). Group VI: JQ 655296.1

(GH30). Group VII: KM 058745.1 (GH24). The data analysis pipeline was optimized for the specific cognate region, using a positive control consisting of amplicon obtained from clones. The optimized parameters for reference mapping were a 0.9 similarity fraction, 0.9 length fraction, and the use of the 'ignore' function where reads that were capable of multiple mappings are classified as unmapped. A cut off was determined to be ≤1% for an individual variant according optimization of the positive control and its known viral contents. These parameters were applied to the data received and the results are summarized in *Figure 3*. Unmapped reads were collected and subjected to *de novo* assembly at default parameters, thereafter continued to multiBLAST the generated contigs. A detailed flow chart of the data optimization can be found in appendix A *Figure 2,* and the reference mapping and *de novo* assembly results of the positive control in *Tables 3-21* under appendix A.

Statistical analysis

Differences in rootstock and scion results were statistically analysed using two proportion hypothesis z-test, at a 95% confidence interval.

RESULTS

Scion-rootstock combinations and number of each combination collected can be viewed in *Table 2*. The rootstocks sampled represents those most commonly used in the South African wine industry and include 101-14, Paulsen, Ramsey and Richter 99 (R99). Criteria for the sampling of vines were that scions display clear GLD symptoms, and that the rootstock of these vines had sizeable lignified suckers/canes growing from the stems. As rootstock suckers are usually pruned in practice, such plants were relatively rare.

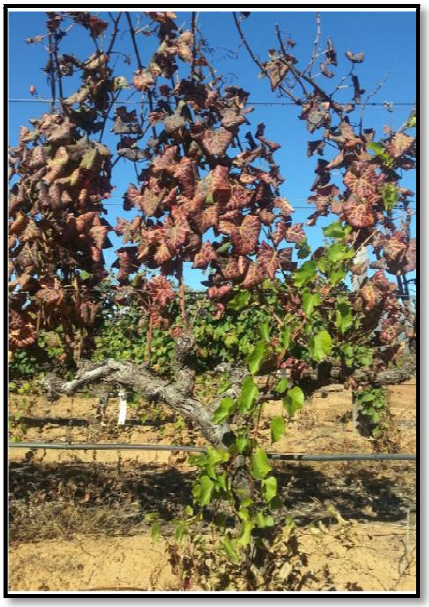


Figure 1: Image depicting a vine that meets all the criteria of sampling, which are clear scion symptoms and substantial lignified suckers/canes growing from stems

Detection of GLRaVs

Ninety five vines were tested for GLRaV-3, while only samples from the second season onwards i.e. 62 out of the 95 samples were tested for GLRaV-1, -2, -4-like, and -7. GLRaVs detected were compared between rootstocks and scions of the same plants. The majority the plants had rootstocks with a different GLRaV-3 status to that of its corresponding scion. Two out of 62 (3%) samples tested positive for GLRaV-1 in both rootstocks and scions both positive rootstock tissue samples were R99, 4 out of 62 (6%) tested positive for GLRaV-2 in rootstocks, and 8 out of 62 (13%) in scions. GLRaV-3 was detected in 41 out of 95 (43%) rootstocks in lower amplicon concentrations than the 88 out of 95 (93%) scions, *Figure 1* under appendix A illustrates typical image of a GLRaV-3 assay agarose electrophoresis gel and differences in amplicon band strengths. GLRaV-4-like and GLRaV-7 were not detected in any of the 62 rootstocks and scions tested. The small number of positive vines of GLRaV-1 or -2 do not provide enough evidence to suggest that the detection rate of either GLRaV-1 ($z=1$, $p>0.05$) or GVLRaV-2 ($z=1.215$, $p>0.05$) differ between scion and rootstock. However, the detection rate of GLRaV-3 differs significantly

(z=7.303, p<0.05) between scions and their respective rootstocks. In some cases only scion or rootstock tissue of a vine was found to be positive and in other cases both tissues of the vine was found to be positive for the respective viruses (*Figure 2*).

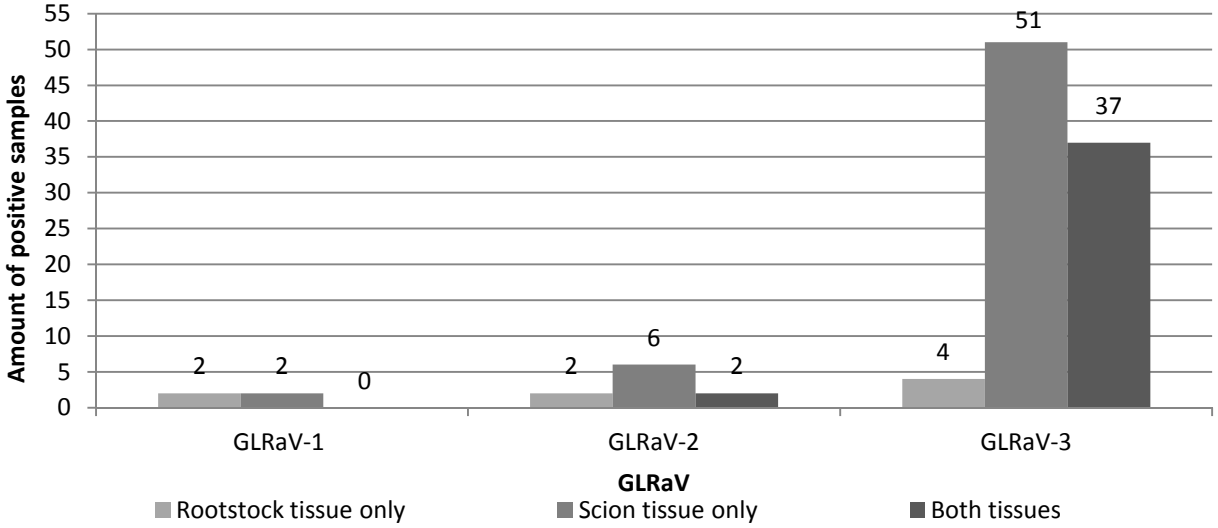


Figure 2: Graph indicating positive samples found in scion and rootstock tissues, and the respective tissue positive of the vine. The different colour bars represent the tissues in which the positives were found in each vine for the various GLRaVs

Table 2: GLRaV-3 positive samples per scion-rootstock combination

Rootstock-scion combination		Scion cultivar colour	Rootstocks GLRaV-3 Positive	Scions GLRaV-3 Positive	N	Selected for NGS
Rootstock	**Scion**					
101-14	**Cabernet Sauvignon**	Red	1	1	2	-
101-14	**Merlot**	Red	2	5	6	1
Paulsen	**La Rochelle**	Red	2	2	2	1
R 99	**Assyrtiko**	White	0	1	1	1
R 99	**Catarratto commune**	White	1	2	2	1
R 99	**Cabernet franc**	Red	2	3	4	2
R 99	**Cabernet Sauvignon**	Red	3	3	3	-
R 99	**CG 40318**	White	0	2	2	-
R 99	**Chardonnay**	White	1	8	8	-
R 99	**Cinsault**	Red	3	3	3	-
R 99	**Gamay hatif des vosges**	Red	0	2	2	-
R 99	**Lakemont seedless**	White	1	2	2	-
R 99	**L. Red**	Red	2	1	2	1
R 99	**Lumbrusco**	Red	0	2	2	-
R 99	**Malbec**	Red	3	5	5	-
R 99	**Merlot**	Red	3	14	16	1
R 99	**Planta nova**	Red	0	2	2	-
R 99	**Pinotage**	Red	2	6	7	-
R 99	**Pontac**	Red	1	3	3	-
R 99	**Ruby Cabernet**	Red	2	3	3	-
R 99	**Shiraz**	Red	2	5	5	1
R 99	**Tinta barrocca**	Red	2	2	2	-
R 99	**Zeni**	White	0	2	2	-
Ramsey	**Ruby Cabernet**	Red	5	6	6	2
Ramsey	**Shiraz**	Red	3	3	3	2
Total			41	88	95	11

N = Total amount of vines tested

The GLRaV-3 viral status for 101-14 was 3 out of 8 (38%), Paulsen had 2 out of 2 (100%) positives, Richter 99 had 28 out of 76 (39%), Ramsey had 8 out of 9 (89%), and Unknown rootstocks had 2 out of 8 (25%) (*Table 2*). Ramsey rootstocks present sufficient evidence to conclusively say that it had significantly higher GLRaV-3 viral status than average rootstock GLRaV-3 viral status. An expanded view of the GLRaV-3 presence, absence and amplicon concentration can be found in *Table 1* and *Table 2* under appendix A.

Illumina MiSeq sequencing data analysis

The 11 of the 37 vines that tested positive for both rootstock and scion tissue were selected (*Table 2*) and prepared for Illumina MiSeq sequencing. The samples submitted for Illumina MiSeq analysis yielded a maximum and minimum of 910 090 and 5 077 reads (average 478 499 reads) of 11 rootstock samples respectively. Eleven scions samples had a maximum and minimum of 1 103 838 and 3 444 reads (average 374 718 reads) respectively.

Analysis of the amount of positive samples for each GLRaV-3 variant group indicated that presence did not significantly differ between rootstocks and scions, 4 out of 11 (36%) of the individual rootstock-scion combinations differed in presence and absence of GLRaV-3 variants, and were 2 Cabernet franc-R99, 1 L. red-R99, 1 Ruby cabernet-Ramsey and 1 Merlot-101.14. The presence of the total amount of samples positive for the assorted groups of GLRaV-3 variants can be seen in *Figure 3*. No sufficient evidence was found to conclude that the total amount of positive samples for GLRaV-3 variant groups differ between rootstocks and scions.

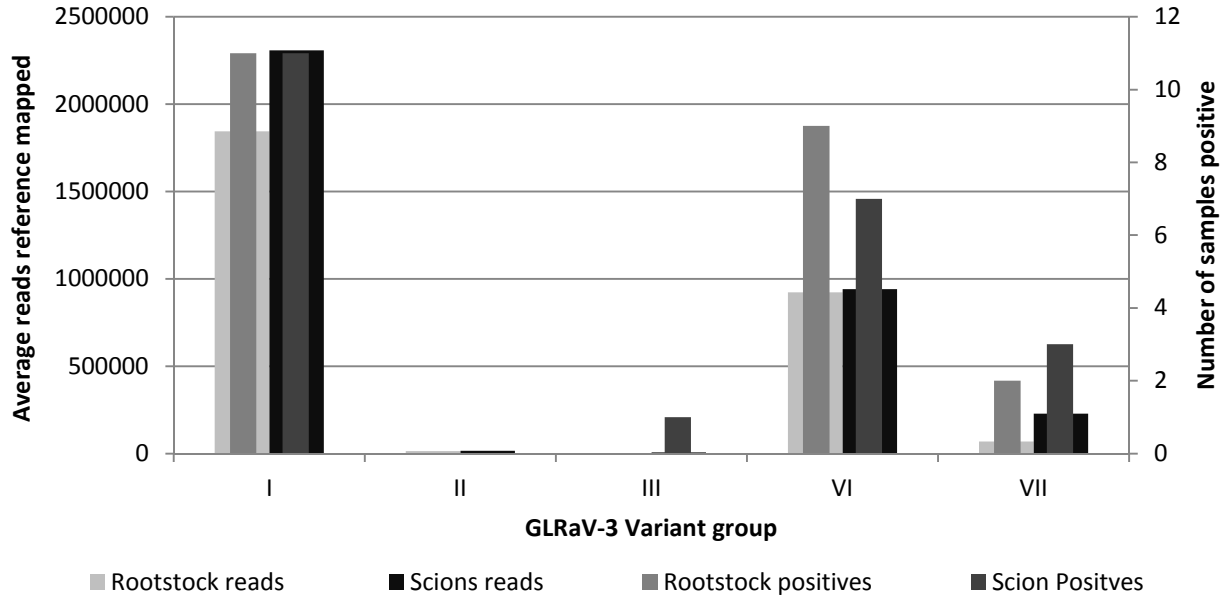


Figure 3: Graph illustrating the differences in number of positive samples and the average amount of reads mapped to GLRaV-3 variant groups between rootstocks versus scions. Rootstocks (n) = 11 and Scions (n) = 11

Illumina MiSeq read composition of each individual vine is listed in *Table 3* and both the rootstock and scion tissue results of each individual vine can be compared next to each other. Seven out of 11 vines differed in composition of reads mapped to various GLRaV-3 groups between rootstock and scion tissue. Detailed information on the Illumina MiSeq results of samples can be viewed in Appendix A *Table 22 – Table 43*, together with a graph (appendix A *Figure 3*) illustrating the read composition.

No new variants were discovered using *de novo* assembly, though additional mappings to existing variant were recovered implementing *de novo* assembly thus supplementing the reference mapping.

Table 3: Presence and absence and percentage reads mapped of various GLRaV-3 variant groups of individual scions and corresponding rootstocks samples

Accession	Scion-rootstock combination		Tissue	GLRaV-3 variant									
	Scion	Rootstock		Group I		Group II		Group III		Group VI		Group VII	
				-/+	%	-/+	%	-/+	%	-/+	%	-/+	%
14-9073	Merlot	**101-14**	Rootstock	+	96	-	1	-	0	+	3	-	0
14-9074			Scion	+	99	-	1	-	0	-	0	-	0
14-9001	La Rochelle	**Paulsen**	Rootstock	+	99	-	1	-	0	-	0	-	0
14-9002			Scion	+	99	-	1	-	0	-	0	-	0
14-9019	Cabernet franc	**R99**	Rootstock	+	88	-	1	-	0	+	10	-	0
14-9020			Scion	+	43	-	0	-	0	+	40	+	16
14-9021	Cabernet franc	**R99**	Rootstock	+	60	-	1	-	0	+	31	+	6
14-9022			Scion	+	44	-	0	-	0	+	37	+	18
14-9031	L. red	**R99**	Rootstock	+	84	-	1	-	0	+	5	+	10
14-9032			Scion	+	45	-	1	+	2	+	31	+	26
15-5015	Merlot	**R99**	Rootstock	+	80	-	0	-	0	+	20	-	0
15-5016			Scion	+	98	-	0	-	0	-	1	-	0
15-5067	Shiraz	**R99**	Rootstock	+	100	-	0	-	0	-	0	-	0
15-5068			Scion	+	100	-	0	-	0	-	0	-	0
15-5007	Ruby cabernet	**Ramsey**	Rootstock	+	4	-	0	-	0	+	96	-	0
15-5008			Scion	+	31	-	0	-	0	+	68	-	0
15-5009	Ruby cabernet	**Ramsey**	Rootstock	+	6	-	0	-	0	+	94	-	0
15-5010			Scion	+	37	-	0	-	0	+	63	-	0
15-5043	Shiraz	**Ramsey**	Rootstock	+	48	-	0	-	0	+	52	-	0
15-5044			Scion	+	66	-	0	-	0	+	34	-	0
15-5077	Shiraz	**Ramsey**	Rootstock	+	66	-	0	-	0	+	33	-	0
15-5078			Scion	+	67	-	0	-	0	+	33	-	0

+ = GLRaV-3 positive and - = GLRaV-3 negative

DISCUSSION AND CONCLUSIONS

The specific GLRaV-3 RT-PCR system was utilised as it contained primers directed at binding sites that are highly conserved amongst the GLRaV-3 variants known at initiation of this study, and amplifies a sequence that is variable amongst the GLRaV-3 variants. This allows identification of the GLRaV-3 variants present following Illumina targeted sequencing of the amplicon, an approach also utilised for *Citrus tristeza virus* viral population studies (Read & Pietersen, 2015). The RT-PCR with primers targeting all known GLRaV-3 variants revealed a major difference in the GLRaV-3 viral status in rootstocks to that of scions (*Figure 2*). The GLRaV-3 viral status of the scion tissue was higher than that of rootstock tissue, and vines that tested positive for both tissues. The low frequency of GLRaV-3 in rootstocks tested in this study supports the study of Chooi et al 2016 who found 15 out of 32 (47%) of 3309C (*Vitis riparia* x *V. rupestris*) and Schwarzmann (*V. riparia* x *V. rupestris)* rootstocks showed inconsistent GLRaV-3 detection and also that low viral titre and uneven distribution could contribute to lower detection rates using RT-qPCR. In the rootstocks basal material harboured higher viral titres than the apical material,

though the time during the growing season was not specified, nor was a V. *vinifera* control included to compare differences between the *Vitis* spp (Chooi*, et al.*, 2016). Tsai et al 2012 did seasonal and virus translocation studies on V. *vinifera* and observed that erratic viral titre occurs even in V. *vinifera* in growing canes and is season specific. The same was observed for V. *vinifera* material by Monis and Bestwick 1996 using serological methods. The difference in detection rates between scion and rootstock could thus be contributed by viral titre or the erratic distribution, a study comparing the viral titre and distribution between rootstock material and *V. vinifera* material is necessary to gain more insight into the dynamics responsible for the difference in detection rates.

Sufficient statistical evidence was found to conclude that GLRaV-3 frequency in assorted rootstocks differs. Ramsey rootstocks had a significantly higher detection than that of R99. Paulsen also exhibited a high detection rate than that of R99, though the sample size is too small and needs further assessment. The rate of GLRaV-3 detection in R99 is very similar to what Chooi et al found in rootstocks, though the rootstocks they used were of dissimilar parentage than that of R99 (*Table 4*). The rootstock sampled that shared most similar parentage to that of 3309C and Schwarzmann used in Chooi et al is 101-14 rootstock, the detection rate observed in 101-14 rootstocks did not resemble that found in Chooi et al. Richter 99 (R99) is one of the most predominant rootstocks used in the wine industry of South Africa (Teubes, 2014), whereas Ramsey/Saltcreek rootstocks are more frequently used in the table grape industry (Teubes, 2014). The parentage and *Vitis* spp. of Ramsey differs distinctively from other rootstocks (*Table 4*) (Malan & Meye, 1993), and could explain the putative higher detection of GLRaV-3 in Ramsey rootstocks compared to others. A higher incidence in Ramsey rootstocks could imply that it is more susceptible to GLRaV-3.

Table 4: Assorted tested rootstock varieties used in South Africa and their parentage (Malan & Meye, 1993)

Rootstock	**Cross source**
101-14	*V. Riparia* x *V. rupestris*
Pausen 1103	*V. Berlandieri* x *V. rupestris*
Ramsey	*V. Champinii*
Richter 99	*V. Belandieri* x *V. rupestris*

The other GRLaVs tested for in the same samples, were present in low frequencies or absent in all samples, even though samples were taken targeting GLD symptoms. The low presence of the other GLRaVs is in agreement with a survey conducted by Jooste et al 2015 in South Africa. When considering the tissues in which the positives were found (*Figure 2*) GLRaV-1 was never found in both tissues of a vine but rather in either the rootstock tissue or the scion tissue. The highest GLRaV-2 detection occurred in the scion tissue (*Figure 2*), the presence in both tissues of vines were low and also in rootstock tissue. GLRaV-2 has no insect vector and can only be transmitted between vines by graft inoculum.

The most abundant GLRaV-3 variants found in the samples were group I and group VI in both rootstocks and scions, most predominant to least is group I, group VI, group VII, group III (*Figure 3*). Contrary to what was found in a survey by Jooste et al 2015, where they found group II and group VI to be the most abundant of scions in South Africa. A possible explanation for this could be that this study focused analysing GLRaV-3 variants via Illumina sequencing on samples that tested positive for both rootstock and scion of the same plant; this could indicate that rootstocks become more frequently infected when group II is absent. Another explanation could be that our study was not survey orientated and therefore sampling was not as geographically distributed as in Jooste et al 2015. Furthermore it is worth noting that the vines selected for NGS was selected based on the highest concentration of GLRaV-3 amplicon obtained from the rootstock samples. This could possibly be responsible observing absence of any new variants amplified in low amounts.

Figure 3 exhibits the differences in the total number of samples positive for assorted GLRaV-3 variants and that they can differ between rootstocks and scions. All of the variants occur more frequently in scions than in rootstocks except for group VI, though this cannot be conclusively proven given this data. When considering presence and absence of individual samples comparing rootstock and scions, half of the plants tested differed in GLRaV-3 variants, but a distinctive pattern could not be discerned. This was not the case with Meng *et al* 2006 where they distinctly observed that the rootstocks could only harbour one *Grapevine rupestris stempitting associated virus* (GRSPaV) variant although the *V. vinifera* of the same plant contained mixed infections with up to four variants. They hypothesized that this could be due to co-evolution between the rootstocks and GRSPaV. Selection of an unusual, poorly detected variant is therefore not the cause of the observed poor

detection of GLRaV-3 in rootstock tissue. It is not clear whether GLRaV-3 variant selection does occur in rootstocks, only that distribution of GLRaV-3 variants in scion versus rootstocks is erratic.

Illumina MiSeq analysis did not reveal significant differences in presence of GLRaV-3 variant groups between the total positive scion and rootstock samples; though it cannot be conclusively stated. Some of the individual samples differed in GLRaV-3 presence in rootstocks versus scions, thus GLRaV-3 population can differ between rootstocks and scions or they can be the same. This indicates erratic behaviour in the infection of grapevines with GLRaV-3 variant groups in rootstock versus scions.

The large number sequences that Illumina MiSeq generates is expected to delivers a more realistic representation of the composition of sequences in a PCR amplification product than cloning would (Read & Pietersen, 2016). The average amount of reads mapping the various GLRaV-3 variant groups differs significantly (*Figure 3*), and indicates that in all cases except group VI more reads mapped to all GLRaV-3 in scion samples than in rootstocks. The amount of reads mapped to the various GLRaV-3 groups supports the presence of groups found in scion versus rootstocks. This supports findings by Chooi et al 2016 which focused their study on GLRaV-3 variant group I, group VI and NZ2, and observed that group VI had a higher viral titre in rootstocks than that of group I (Chooi*, et al.*, 2016). The primer pair used to detect the various GLRaV-3 variants utilizes highly conserved regions and contains no mismatches with GLRaV-3 variants considered in this study, and can thus be considered to introduce minimum PCR bias and to most accurately determine viral population in the samples using Illumina MiSeq sequencing (Read & Pietersen, 2016).

Characterization of the differences in GLRaV-3 variant populations in rootstocks versus scions was analysed (*Figure 4*). A greater variety of GLRaV-3 variants seem to infect R99 than the other rootstocks. When comparing read composition of Ramsey and R99 and their corresponding scions a trend was observed that both Ramsey and R99 favours the replication of one variant compared to corresponding scion. This possibly supports the difference seen in the spread of GLRaV-3 in micrografted grapevines, where infected scion-healthy rootstock combinations had a accelerated spread to healthy material than healthy scion-infected rootstock combinations (Hao*, et al.*, 2017).

This study confirms the inconsistent detection of GLRaV-3 in rootstocks (Chooi*, et al.*, 2016), and GLRaV-3 was undetectable in the vast majority of R99 samples. It is possible and important to test rootstocks for GLRaV-3 in a grapevine certification scheme, especially because the majority of rootstocks used in the wine industry are R99. Differences in GLRaV-3 populations between rootstock and scion occur, but no consistent pattern could be established regarding the differences in variants present. The observation that GLRaV-3 variant group VI constituted a larger component of the GLRaV-3 population in rootstocks than in scions supports previous studies done on group I and group VI (Chooi*, et al.*, 2016). The dynamics of GLRaV-3 and rootstocks are largely unknown and this study helps close the gap in the knowledge regarding this topic. More controlled studies are required to explore other possibilities for this phenomenon.

REFERENCES

Abou Ghanem-Sabanadzovic N., Sabanadzovic S., Gugerli P. and Rowhani A. (2012). Genome organization, serology and phylogeny of *Grapevine leafroll-associated viruses 4* and *6*: taxonomic implications. Virus Research vol.163 pp120-128

Akbaş B., Kunter B. and Ilhan D. (2007). Occurrence and distribution of *Grapevine leafroll-associated viruses 1, 2, 3* and *7* in Turkey. Journal of Phytopathology vol.155 pp122-124

Al Rwahnih M., Dolja V.V., Daubert S., Koonin E.V. and Rowhani A. (2012). Genomic and biological analysis of *Grapevine leafroll-associated virus 7* reveals a possible new genus within the family *Closteroviridae*. Virus Research vol.163 pp302-309

Al Rwahnih M., Osman F., Sudarshana M.*, et al.* (2012). Detection of *Grapevine leafroll-associated virus 7* using real time qRT-PCR and conventional RT-PCR. Journal of Virological Methods vol.179 pp383-389

Alkowni R., Zhang Y.-P., Rowhani A., Uyemoto J.K. and Minafra A. (2011). Biological, molecular, and serological studies of a novel strain of *Grapevine leafroll-associated virus 2*. Virus Genes vol.43 pp102-110

Almeida R.P., Daane K.M., Bell V.A., Blaisdell G.K., Cooper M.L., Herrbach E. and Pietersen G. (2013). Ecology and management of grapevine leafroll disease. Frontiers in Microbiology vol.4 pp1-13

Bertazzon N. and Angelini E. (2004). Advances in the detection of *Grapevine leafroll-associated virus 2* variants. Journal of Plant Pathology pp283-290

Bester R., Jooste A.E., Maree H.J. and Burger J.T. (2012). Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate *Grapevine leafroll-associated virus 3* variant groups I, II, III and VI. Virology Journal vol.9 pp1-12

Bester R., Pepler P., Burger J. and Maree H. (2014). Relative quantitation goes viral: An RT-qPCR assay for a grapevine virus. Journal of Virological Methods vol.210 pp67-75

Chooi K.M., Cohen D. and Pearson M.N. (2013). Generic and sequence-variant specific molecular assays for the detection of the highly variable *Grapevine leafroll-associated virus 3*. Journal of Virological Methods vol.189 pp20-29

Chooi K.M., Cohen D. and Pearson M.N. (2016). Differential distribution and titre of selected *Grapevine leafroll-associated virus 3* genetic variants within grapevine rootstocks. Archives of Virology vol.161 pp1371-1375

Downie D. (2002). Locating the sources of an invasive pest, grape phylloxera, using a mitochondrial DNA gene genealogy. Molecular Ecology vol.11 pp2013-2026

Fiore N., Prodan S., Montealegre J., Aballay E., Pino A. and Zamorano A. (2008). Survey of grapevine viruses in Chile. Journal of Plant Pathology pp125-130

Forneck A. and Huber L. (2009). (A) sexual reproduction–a review of life cycles of grape phylloxera, *Daktulosphaira vitifoliae*. Entomologia Experimentalis et Applicata vol.131 pp1-10

Fuchs M., Martinson T., Loeb G. and Hoch H. (2009). Survey for the three major leafroll disease-associated viruses in Finger Lakes vineyards in New York. Plant Disease vol.93 pp395-401

Goszczynski D.E. (2013). Brief Report of a New Highly Divergent Variant of *Grapevine leafroll-associated virus 3* (GLRaV-3). Journal of Phytopathology vol.161 pp874-879

Gouveia P., Santos M.T., Eiras-Dias J.E. and Nolasco G. (2011). Five phylogenetic groups identified in the coat protein gene of *Grapevine leafroll-associated virus 3* obtained from Portuguese grapevine varieties. Archives of Virology vol.156 pp413-420

Granett J., Timper P. and Lider L. (1985). Grape phylloxera (*Daktulosphaira vitifoliae*)(Homoptera: *Phylloxeridae*) biotypes in California. Journal of Economic Entomology vol.78 pp1463-1467

Hao X.Y., Bi W.L., Cui Z.H., Pan C., Xu Y. and Wang Q.C. (2017). Development, histological observations and *Grapevine leafroll-associated virus-3* localisation in *in vitro* grapevine micrografts. Annals of Applied Biology vol.170 pp379-390

Jooste A., Maree H., Bellstedt D., Goszczynski D., Pietersen G. and Burger J. (2010). Three genetic *Grapevine leafroll-associated virus 3* variants identified from South African vineyards show high variability in their 5′ UTR. Archives of Virology vol.155 pp1997-2006

Jooste A.E., Pietersen G. and Burger J.T. (2011). Distribution of *Grapevine leafroll associated virus-3* variants in South African vineyards. European Journal of Plant Pathology vol.131 pp371-381

Jooste A.E., Molenaar N., Maree H.J., Bester R., Morey L., de Koker W.C. and Burger J.T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. European Journal of Plant Pathology vol.142 pp363-375

Lounsbury C. (1940). The pioneer period of economic entomology in South Africa. J Entomol Soc S Afr vol.3 pp9-29

Malan A.P. and Meye A. (1993). Interaction between a South African population of *Xiphinema index* and different grapevine rootstocks. South African Journal of Enology and Viticulture vol.14 pp11-15

Maree H., Freeborough M.-J. and Burger J. (2008). Complete nucleotide sequence of a South African isolate of *Grapevine leafroll-associated virus 3* reveals a 5′ UTR of 737 nucleotides. Archives of Virology vol.153 pp755-757

Maree H., Almeida R., Bester R.*, et al.* (2013). *Grapevine leafroll-associated virus 3*. Frontiers in Microbiology vol.4 pp1-12

Maree H.J., Pirie M.D., Oosthuizen K., Bester R., Rees D.J.G. and Burger J.T. (2015). Phylogenomic analysis reveals deep divergence and recombination in an economically important grapevine virus. PloS one vol.10 pp1-18

Martelli G., Abou Ghanem-Sabanadzovic N., Agranovsky A.*, et al.* (2012). Taxonomic revision of the family *Closteroviridae* with special reference to the grapevine leafroll-associated members of the genus *Ampelovirus* and the putative species unassigned to the family. Journal of Plant Pathology vol.94 pp7-19

Martin R., Eastwell K., Wagner A., Lamprecht S. and Tzanetakis I. (2005). Survey for viruses of grapevine in Oregon and Washington. Plant Disease vol.89 pp763-766

Meng B., Rebelo A.R. and Fisher H. (2006). Genetic diversity analyses of grapevine *Rupestris stem pitting-associated virus* reveal distinct population structures in scion versus rootstock varieties. Journal of General Virology vol.87 pp1725-1733

Monis J. and Bestwick R.K. (1996). Detection and localization of grapevine leafroll associated closteroviruses in greenhouse and tissue culture grown plants. American Journal of Enology and Viticulture vol.47 pp199-205

Olmos A., Bertolini E., Ruiz-García A.B., Martínez C., Peiró R. and Vidal E. (2016). Modeling the Accuracy of Three Detection Methods of *Grapevine leafroll-associated virus 3* During the Dormant Period Using a Bayesian Approach. Phytopathology vol.106 pp510-518

Osman F. and Rowhani A. (2006). Application of a spotting sample preparation technique for the detection of pathogens in woody plants by RT-PCR and real-time PCR (TaqMan). Journal of Virological Methods vol.133 pp130-136

Over de Linden A. and Chamberlain E. (1970). Effect of grapevine leafroll virus on vine growth and fruit yield and quality. New Zealand Journal of Agricultural Research vol.13 pp689-698

Pacifico D., Caciagli P., Palmano S., Mannini F. and Marzachì C. (2011). Quantitation of *Grapevine leafroll associated virus-1* and *-3, Grapevine virus A*, *Grapevine fanleaf virus* and *Grapevine fleck virus* in field-collected *Vitis vinifera* L. 'Nebbiolo' by real-time reverse transcription-PCR. Journal of Virological Methods vol.172 pp1-7

Petersen C. and Charles J. (1997). Transmission of grapevine leafroll-associated *closteroviruses* by *Pseudococcus longispinus* and *P. calceolariae*. Plant Pathology vol.46 pp509-515

Pietersen G. (2004). Spread of grapevine leafroll disease in South Africa-a difficult but not insurmountable problem. Technical Yearbook 2004/5

Pietersen G., Spreeth N., Oosthuizen T.*, et al.* (2013). Control of grapevine leafroll disease spread at a commercial wine estate in South Africa: a case study. American Journal of Enology and Viticulture pp269-305

Rayapati A., O'Neil S. and Walsh D. (2008). Grapevine leafroll disease. Washington State University Extension Bulletin EB2027E

Read D.A. and Pietersen G. (2015). Genotypic diversity of *Citrus tristeza virus* within red grapefruit, in a field trial site in South Africa. European Journal of Plant Pathology vol.142 pp531-545

Read D.A. and Pietersen G. (2016). PCR bias associated with conserved primer binding sites, used to determine genotype diversity within *Citrus tristeza virus* populations. Journal of Virological Methods vol.237 pp107-113

Saayman D. (2004). The South African vineyard and wine landscapes: Heritage and development. WineLand media. Date acessed: 28/03/2017. http://www.wineland.co.za/the-south-african-vineyard-and-wine-landscapes-heritage-and-development/

Sharma A.M., Wang J., Duffy S.*, et al.* (2011). Occurrence of grapevine leafroll-associated virus complex in Napa Valley. PLoS One vol.6 pp1-7

Teubes A. (2014). History of roostocks in South Africa (Part 1). WineLand Media. Date acessed: 2017-05-04. http://www.wineland.co.za/the-history-of-rootstocks-in-south-africa-part-1/

Teubes A. (2014). History of rootstocks in South Africa (Part 8). WineLand Media. Date acessed: 2017-05-04. http://www.wineland.co.za/history-of-rootstocks-in-south-africa-part-8/

Töpfer R., Hausmann L., Harst M., Maul E., Zyprian E. and Eibach R. (2011). New horizons for grapevine breeding. Methods In Temperate Fruit Breeding, Fruit, Vegetable and Cereal Science and Biotechnology vol.5 pp79-100

Tsai C., Daugherty M. and Almeida R. (2012). Seasonal dynamics and virus translocation of *Grapevine leafroll-associated virus 3* in grapevine cultivars. Plant Pathology vol.61 pp977-985

Turturo C., Saldarelli P., Yafeng D., Digiaro M., Minafra A., Savino V. and Martelli G. (2005). Genetic variability and population structure of *Grapevine leafroll-associated virus 3* isolates. Journal of General Virology vol.86 pp217-224

White E.J., Venter M., Hiten N.F. and Burger J.T. (2008). Modified Cetyltrimethylammonium bromide method improves robustness and versatility: the benchmark for plant RNA extraction. Biotechnology Journal vol.3 pp1424-1428

Chapter 4

Detection of *Viti-* and *Foveavirus*es of *Vitis* rootstocks used in South Africa

Harris, M.[1], and Pietersen, G.[1,2]

[1]FABI, Department of Microbiology and Plant Pathology, University of Pretoria, Pretoria, 0002 South Africa.

[2]Agricultural Research Council-Plant Protection Research Institute, Pretoria, South Africa.

ABSTRACT

Viti- and *Foveavirus*es are closely associated with the rugose wood (RW) complex of grapevine. This study aims to determine the relative detection of *Viti-* and *Foveaviruses* in Richter 99 (R99), Richter 110 (R110), 101-14, Paulsen and Ramsey rootstocks versus scions, and so also identify which specific *Vitiviruses* infect rootstocks frequently. Rootstocks and their corresponding scions were tested using universal degenerate primers detecting the RNA dependent RNA polymerase (RdRp) gene of both *Viti-* and *Foveavirus*es, and compared to each other. The detection rate is significantly lower in rootstocks, the majority of the rootstocks consisting of R99, than in scions. To identify the *Viti-* and *Foveavirus*es that infect the various grapevine rootstocks, Illumina MiSeq was conducted using *Viti-* and *Foveavirus* PCR amplicons as template, and the data analysed. PCR using specific *Vitivirus* primers was used to confirm the Illumina MiSeq data that indicated presence of GVA, GVB or GVE in the different rootstock tissues. In the samples analysed, GVB was found in rootstocks but not in scions, and this observation requires further investigation. Detection of GVE was poor using the universal *Viti-* and *Foveavirus* primers. The information generated in this study aids in the understanding of the dynamics of the RW complex and the role it plays in rootstocks versus scions, and provides evidence to warrant a change the approach in detection of *Viti-* and *Foveavirus*es of rootstocks in the certification scheme of South Africa.

INTRODUCTION

Plants afflicted with Grapevine leafroll disease are generally infected with Grapevine leafroll associated virus 3 (GLRaV-3) in South Africa, and are often co-infected with various *Vitiviruses*, including *Grapevine virus A* (GVA), *Grapevine virus B* (GVB), *Grapevine virus E* (GVE), and *Grapevine virus F* (GVF) (Martelli, 1993, Jooste*, et al.*, 2015). Both the *Foveavirus* known as *Grapevine rupestris stem pitting-associated viruses* and *Vitiviruses* are members of the *Betaflexiviridae* family (Martelli*, et al.*, 2007, ICTV, 2015), both genera are associated with a global destructive disease complex affecting grapevine known as rugose wood (RW) complex.

The diseases that form part of the RW complex are rupestris stem pitting, LN stem grooving, Kober stem grooving and corky Bark (Rosa & Rowhani, 2007). In addition to Kober stem grooving (Garau*, et al.*, 1994, Credi, 1997), GVA is also associated with Shiraz disease (Goszczynski, 2007). *Vitiviruses* associated with corky bark disease are GVB (Bonavia*, et al.*, 1996), GVD (Boscia*, et al.*, 2001, Martelli*, et al.*, 2007), GVE, and GVF (Al Rwahnih*, et al.*, 2012). GRSPaV has also been linked to Grapevine vein necrosis (Bouyahia*, et al.*, 2015).

The introduction of a destructive aphid-like insect, *Daktulosphaira vitifoliae* (Fitch 1855); family *Phylloxeridae* commonly known as phylloxera to South Africa in the 1860s created the necessity of grafting grapevine scions on *Phylloxera* resistant rootstocks (Lounsbury, 1940, Töpfer*, et al.*, 2011). Rootstocks often harbour GRSPaV and some rootstocks are indicators of Rupestris stem pitting disease (Meng*, et al.*, 2003, Habili*, et al.*, 2006, Meng*, et al.*, 2006). Co-infection of GLRaVs and *Vitiviruses* can have detrimental effects on some rootstocks, and the severity of the response of rootstocks to the virus infection depends on the rootstock genotype and virus type (Golino*, et al.*, 2008).

Limited information is available about the dynamics of *Viti*-and *Foveavirus*es in commercially used rootstocks in South Africa. The objective of this study was to determine which *Viti-* and *Foveavirus*es occur in rootstocks commercially utilised in South Africa, and compare the relative virus detection between the *Vitis* rootstocks and their corresponding scions in individual vines. Due to the relatively novel method of using amplicons in combination with Illumina MiSeq it was also necessary to optimize the parameters of the NGS data analysis.

METHODS AND MATERIALS

Viruses, sampling and RNA isolation

Phloem tissue were collected separately from bark scrapings of the scion and rootstock of 95 vine individuals from commercial vineyards in two wine growing regions (Wellington and Stellenbosch) of the Western Cape, South Africa between 2014 and 2016. The vines selected for sampling all displayed clear GLD symptoms on the scions and also possessed rootstocks with lignified suckers or canes. After being stored at -80ºC total RNA isolation of 200mg phloem was done using a modified cetyltrimethylammonium bromide (CTAB) (2% CTAB, 2.5% PVP-40, 100mM Tris-HCL pH8, 2M NaCl, 25mM EDTA pH8 and 3% β-mercaptoethanol) method (White*, et al.*, 2008). Quantity and quality of total RNA was determined using both gel electrophoresis (2% Agarose-TAE), and spectrophotometry (Nanodrop 1000).

RT-PCR

Detection of *Viti-* and *Foveavirus*es was achieved by using a modified universal nested RT-PCR system (Dovas & Katis, 2003) (*Table 1*). The one step RT-PCR was performed using a 25µl reaction mixture containing 2.5µl reaction buffer (10x), 1.25µl dNTP mixture (10mM) (Roche, Basel, Switzerland), 1.25µl forward primer dRW_up1 (10µM), 1.25µl reverse primer dRW_do2 (10µM), 0.05µl RiboLock (40U/µl) (Thermo Fisher Scientific, Waltham, MA, United States), 1.25µl dithiothreitol (DTT) (0.2M), 0.1µl Moloney murine leukemia virus reverse transcriptase (MMLV) (200U/µl) (Promega, Madison, WI, United States), 0.75µl $MgCl_2$ (50mM), 0.05µl BioTaq™ Polymerase (Bioline, London, United Kingdom), 1µl of total RNA, and molecular grade water. Cycle conditions used for the first stage, which is a one-step PCR, are as follows: 1 cycle of 37ºC for 45 min, 50ºC for 2 min, 94ºC for 4 min. 5 cycles of 95ºC for 30 s, 43ºC for 10 s, and 72ºC. 35 cycles of 95ºC for 30 s, 43ºC for 30 s, and 72ºC for 20 s, 1 cycle at 72ºC for 2 min. The second phase is the nested RT-PCR and consists of amplification of the first stage product, which includes both cDNA and amplicon molecules yielding a 200bp product. The nested PCR reaction mixture contained 2.5µl of reaction buffer (10x), 1.25µl of dNTP mixture, 1.25µl of forward

primer dRW_nest1 (10µM), 1.25µl of reverse primer dRW_nest2 (10µM), 0.75 of $MgCl_2$ (50mM), 0.05µl of BioTaq™ polymerase, 0.5µl of the one-step product, and molecular grade water to a final volume of 25µl. Cycle conditions used for the second stage are as follows: 1 Cycle of 95ºC for 3 min, 48ºC for 15 s, and 72ºC for 15 s. 39 Cycles of 95ºC for 30 s, 54ºC for 30 s, and 72ºC 1 Cycle at 72ºC for 2 min. The presence/absence of the product was analysed by performing gel electrophoresis (2% Agarose-TAE).

Direct Sanger sequencing

Amplicons of the correct size was purified by adding 2 Units of FastAP (1U/µl) and 10 Units of ExoI (20U/µl) enzymes (Thermo Scientific, Vilnius, Lithuania) to 19µl of amplified PCR product and sequenced in both directions using dRW_nest1 as forward primer and dRW_nest2 as reverse primer and held at constant temperature of 37ºC for 15 min followed by 85ºC for 15 min. Sequencing reactions consisted of: 1ul of BigDye® Terminator mix v3.1 (Applied Biosystems, Foster City, CA, United States), 2.25µl 5x BigDye ® v3.1 sequencing buffer (Applied Biosystems, Foster City, CA, United States), 0.75µl of the appropriate primer (2µM), and molecular grade water to a final reaction volume of 10µl. Sequencing reaction cycling conditions were: 1 Cycle at 94ºC for 1 min, 25 cycles of 94ºC for 10 s, 50ºC for 5 s, 60ºC for 4 min.

Precipitation of the sequencing reaction was performed by adding 1µl of 125mM EDTA, 1µl of 3M NaOAc, and 25µl of 100% non-denatured ethanol to the sequencing reaction and incubating for 15 min. Followed by centrifugation at maximum speed at 4ºC for 30 min, removal of resulting supernatant, and addition of 100µl of 70% ethanol, where after was centrifuged at max speed for 15 min at 4ºC. The supernatant is subsequently removed and the samples air-dried for 20 min, after all ethanol is evaporated the sample was submitted to the African Centre for Gene Technologies (ACGT), Automated Sequencing Facility, Department of Genetics, University of Pretoria, South Africa and sequenced using ABI Prism ® 3130XL Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

Data analyses of resulting chromatograms were viewed and corrected using Chromas Lite 2.1 (Technelysium, Brisbane, Australia). Alignments were done of the forward and reverse compliment of the reverse sequence in BioEdit Sequence

alignment editor 7.1.3 program employing the CLUSTAL W alignment software (EBI, Cambridgeshire, England) (Hall, 1999), where after consensus sequences were obtained. Forward, reverse and consensus sequences were NCBI BLASTed to identify the dominant virus present in the amplicon sample.

Illumina MiSeq Sequencing

The required concentrations (50ng/µl) for Illumina MiSeq submission was obtained by producing multiple 50µl PCR *Viti-* and *Foveavirus* products, pooling, and column purifying (NucleoSpin® Gel and PCR clean-up, Macherey-Nagel) the products, thereafter eluting in 50µl of elution buffer.

Next generation sequencing, utilizing the Illumina MiSeq (Illumina, San Diego, CA, United States) platform was performed on amplicons obtained from samples and a positive control, amplicon from identified clones, in parallel. The positive control consisted of *Viti-* and *Foveavirus* nested RT-PCR amplicon products amplified from clones containing identified viral amplicons from the same RdRp region, thus allowing certainty and knowledge of the exact constituents of the control. Nextera V2 sample kit (Epicentre, Madison, WI, United States) was used to prepare paired-end DNA libraries and run on $1/8^{th}$ of a lane on an Illumina MiSeq flow cell. The samples were submitted and sequenced at the Agricultural Research Council (ARC), Biotechnology Platform, Pretoria, South Africa.

MiSeq data analysis

CLC Genomics Workbench 6 (CLC bio, Aarhus, Denmark) was used to carry out all trimming and analyses of the Illumina MiSeq data sets. After the data was imported as paired end reads with a distance rage of 180-250, adapter and quality trimming was performed using the Fast QC function on default settings with Nextera V2 transposase adapter sequences
(Transposase1: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG;
Transposase2: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG). After quality control was performed on the data set, each sample was reference mapped using optimized parameters to the cognate amplified region represented by 69 GVA GenBank sequences, 16 GVB reference sequences, the four known GVE sequences, 2 GVF sequences, and 8 RSPaVs. Above mentioned sequence description and GenBank accession numbers can be viewed in Appendix B *Table 30*

Optimization of reference mapping parameters included the use of a positive control with known viral amplicon content. A combination of a variety of similarity fractions and length fraction parameters were used to analyse the % reads that mapped on the positive control, the most optimal combination was applied to the samples. The optimal option for non-specific match handling was determined and implemented. In addition to the optimized settings global alignment was switched off. The cut off of % reads mapped was determined based on positive control and implemented on all samples.

De novo assembly was carried out on the collected unmapped reads at default parameters, thereafter continued to multiBLAST the generated contigs. The number of reads that matched to contigs was added to the number reads reference mapped to form a complete profile of the *Viti-* and *Foveavirus*es present in amplicon sequenced.

Confirmation tests

A number of published and newly designed virus specific primers (*Table 1*) were used to detect specific *Vitiviruses*. Reverse transcription was achieved by first performing a primer annealing step in a 5µl reaction mixture containing 2µl of total RNA, 0.7µl appropriate reverse primer (*Table 1*), and molecular grade water, and subsequently incubated at 70ºC for 5 minutes and thereafter 4ºC for 5 minutes. First strand synthesis was performed using a reaction mixture consisting of 0.5µl reaction buffer (5x), 1.25µl dNTP mixture (10mM) (Roche, Basel, Switzerland), 0.05µl RiboLock™ (40U/µl), Moloney murine leukemia virus reverse transcriptase (MMLV) (Promega, Madison, WI, United States), 5µl cDNA, and molecular grade water to a volume of 15µl, and incubated at 42ºC for 60 minutes. Second strand synthesis was achieved by using a 25µl reaction mixture constituting of 2.5µl NH_4 BioTaq™ reaction buffer, 1.5µl $MgCl_2$ (50mM), 0.5µl dNTP mixture (10mM), 0.5µl appropriate forward primer (10µM), 0.05µl appropriate reverse primer (10µM) (*Table1*), 0.25µl BioTaq™ Polymerase (Bioline, London, United Kingdom), 0.5µl cDNA, and molecular grade water. The cycling conditions used were as follows: 1 Cycle at 95ºC for 1 min. 40 Cycles at 95ºC for 15 s, appropriate annealing temperature (*Table 1*) for 15 s, and 72ºC for 20 s. 1 Cycle at 72ºC for 10 min. The PCR product was analysed using gel electrophoresis (2% Agarose-TAE). Direct Sanger sequencing

was used to confirm identity of the products. Identities of the amplicons were verified using direct Sanger sequencing in both directions using appropriate primers (*Table 1*) and conditions described earlier.

Statistical analysis

A two proportion hypothesis test (z-test) was used to compare two proportions of results to each other with a significance level of p=0.05 and a two tailed approach.

.

Table 1: Primers used in Viti- and Foveavirus universal detection and Vitivirus specific primers

Virus	F/R	Primer name	Primer Sequence (5' to 3')	Annealing temperature	Amplicon length (nt)	Gene region amplified	Reference
***Viti*- and *Foveavirus*es**	F	dRW_up1	WGC IAA RDC IGG ICA RAC	Nested PCR	199bp	RNA dependent RNA polymerase	(Dovas & Katis, 2003)
	R	dRW_do2	RMY TCI CCI SWR AAI CKC AT				
	F	dRW_nest1	GGG GCA RAC IHT IGC ITG YTT				
	R	dRW_nest2	AAI GCY TCR TAR TCI GAI TCN GT				
GVA	F	MP	TGCCAGAGGTGTTTGAGACAAT	61	986	ORF 3,4,5 +3'UTR (MP, CP, and nucleic acid-binding protein	De Meyer 2000, Goszczynski and Jooste 2002
	R	CPdt	TTTTTGTCTTCGTGTGACAACCT				
GVB	F	ACPF	CAATAAGCAAGCARTTCCC	58	751	ORF 4 + IR	Megan (unpublished)
	R	ACPR	CACTCTAMTCTACCACAACA				
GVD	F	GD1	GTACCTTAGGACGCTCTTCGGG	48	700	CP	Abou-Ghanem 1997
	R	GD2	CGTT GGGT CG AGT GT G AGTACG				
GVE	F	EF1	CGTGCGGARGGCAAT	63	804	Mp and Nucleic acid-binding protein	Megan (unpublished)
	R	ER1	CGCCGGGGTTCTTATG				
	F	EF2	AACTTCACCTACCCACCA	63	822	ORF1	Megan (unpublished)
	R	ER2	TTTCATCTCMAGCCTATCC				
	F	EF3	GGATAGGCTKGAGATGAAA	63	819	ORF1 + hypothetical protein	Megan (unpublished)
	R	ER3	CCAAAGGGTAAAGGAGGT				

RESULTS

Detection of Viti- and *Foveavirus*es

The detection of the *Viti-* and *Foveavirus*es was compared between rootstocks and scions of the same plants. Rootstocks had a detection rate of 61% (58/95) and the corresponding scions' detection rate was 86% (82/95), therefore exhibited a significant (z=3.958, p>0.05) (*Table 2*) difference of 33% in detection rate between rootstock and scion. The number of positive rootstocks and scions for the different scion-rootstock combination vines are summarized in *Table 2*. A more detailed table of the results can be viewed in appendix B *Table 1.*

Table 2: Number of Viti- and Foveavirus positive scion and rootstock samples per scion-rootstock combination

Scion-rootstock combination		Red/white	Rootstock	Scion	Total	Selected
Rootstock	Scion	cultivar scion	positives	positives	plants	for MiSeq
101-14	**Cabernet Sauvignon**	Red	0	1	2	-
101-14	**Merlot**	Red	3	4	6	1
Paulsen	**La Rochelle**	Red	1	1	2	-
R99	**Assyrtiko**	White	1	1	1	1
R99	**Cabernet Franc**	Red	3	4	4	2
R99	**Cabernet Sauvignon**	Red	3	3	3	-
R99	**Catarrato commune**	White	1	2	2	1
R99	**CG40318**	White	2	2	2	-
R99	**Chardonnay**	White	6	7	8	2
R99	**Cincault**	Red	3	3	3	-
R99	**Gamay hatif des Vosges**	Red	1	1	2	-
R99	**L rooi**	Red	1	2	2	-
R99	**Lumbrusco**	Red	1	2	2	1
R99	**Lakemont seedless**	White	2	2	2	2
R99	**Malbec**	Red	2	5	5	-
R99	**Merlot**	Red	8	13	16	1
R99	**Planta Nova**	Red	2	1	2	1
R99	**Pinotage**	Red	2	7	7	2
R99	**Pontac**	Red	0	3	3	-
R99	**Ruby Cabernet**	Red	1	3	3	0
R99	**Shiraz**	Red	4	5	5	1
R99	**Tinta barrocca**	Red	2	2	2	-
R99	**Zeni**	White	2	2	2	2
Ramsey	**Ruby Cabernet**	Red	5	3	6	1
Ramsey	**Shiraz**	Red	2	3	3	1
	Total		58	82	95	19

Direct Sanger sequencing

Amplicons of 36 rootstock and 20 scions were submitted for sequencing, the dominant virus identified and listed in *Table 3*. Viruses that had a percentage nucleotide identity below 75% were considered to be an unknown virus according to

the suggested % nucleotide identity separating different *Vitiviruses* in the RdRP gene (Adams*, et al.*, 2004). The top BLAST hits for all the unknown viruses were *Vitiviruses*.

The sample sizes of the various scion-rootstock combinations were too small to make statistically significant inferences (*Table 4*). Corresponding rootstocks of 11 of the 20 scions were sequenced and observed that 5 out of 11 of the rootstocks and scions differed from each other (*Table 5*). The dominant strains of the other scion-rootstock combinations that did not differ between scion and rootstock results are listed in *Table 3*. More detailed results of the direct Sanger sequencing can be viewed in Appendix B, *Table 2*.

Table 3: Table of Vitivirus strains dominant in the population not differing amongst rootstock and scions. Determined by Sanger sequencing

Scion-rootstock combination		***Viti*- and *Foveavirus*es**							
Scion	Rootstock	GVA	GVB	GVD	GVE	GVF	GRSPaV	Unknown	N
Cincault	R99	2	0	0	0	0	0	1	3
Cabernet Sauvignon	R99	0	0	0	0	0	0	2	2
Malbec	R99	0	0	0	1	0	0	0	1
Total		**2**	0	0	1	0	0	3	6

N = Number of vines that had the same dominant virus in both rootstock and scion

Table 4: Number of positive samples for each Viti- and Foveavirus per scion-rootstock combination according to direct Sanger sequencing

Scion-rootstock combination		**Rootstock**								**Scion**							
Scion	Rootstock	GVA	GVB	GVD	GVE	GVF	RSPaV	Unknown	N	GVA	GVB	GVD	GVE	GVF	RSPaV	Unknown	N
Merlot	101-14	0	1	0	0	0	1	1	3	0	0	0	0	0	0	0	0
Cabernet Sauvignon	R 99	1	0	0	0	0	0	2	3	0	0	0	0	0	0	3	3
Chardonnay	R 99	1	0	0	0	0	0	4	5	0	0	0	0	0	0	0	0
Cinsault	R 99	2	0	0	0	0	0	1	3	2	0	0	0	0	0	1	3
Malbec	R 99	0	1	0	1	0	0	0	2	1	0	0	1	1	0	0	3
Merlot	R 99	0	1	0	0	0	5	2	8	0	0	0	0	0	0	0	0
Pinotage	R 99	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	3
Pontac	R 99	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	3
Ruby Cabernet	R 99	0	0	0	3	0	1	0	4	3	0	0	0	0	0	0	3
Shiraz	R 99	1	3	0	0	0	0	1	5	0	0	0	0	0	0	0	0
Tinta barrocca	R 99	0	0	0	0	0	2	0	2	1	0	0	0	0	0	1	2
Shiraz	Ramsey	1	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Total		6	7	0	4	0	9	11	37	10	0	0	3	1	0	6	20

N = Total number of samples tested for each scion-rootstock combination.

Table 5: Number of dominant Viti- and *Foveavirus*es *in rootstocks and scions per scion-rootstock combination*

Scion-rootstock combination		**Rootstock**								**Scion**							
Scion	Rootstock	GVA	GVB	GVD	GVE	GVF	RSPaV	Unknown	N	GVA	GVB	GVD	GVE	GVF	RSPaV	Unknown	N
Cabernet Sauvignon	R 99	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1
Malbec	R 99	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1
Ruby Cabernet	R 99	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	1
Tinta barrocca	R 99	0	0	0	0	0	2	0	2	1	0	0	0	0	0	1	2
Total		1	1	0	1	0	2	0	5	3	0	0	0	0	0	2	5

N = Total number of samples tested for each scion-rootstock combination.

Illumina MiSeq sequencing data analysis optimization

The data analysis pipeline was optimized for use of CLC genomics workbench 6 and the specific cognate region to be used, using a positive control consisting of amplicon obtained from clones. Reference mapping was done using different parameters for length fraction and similarity fraction including default settings.

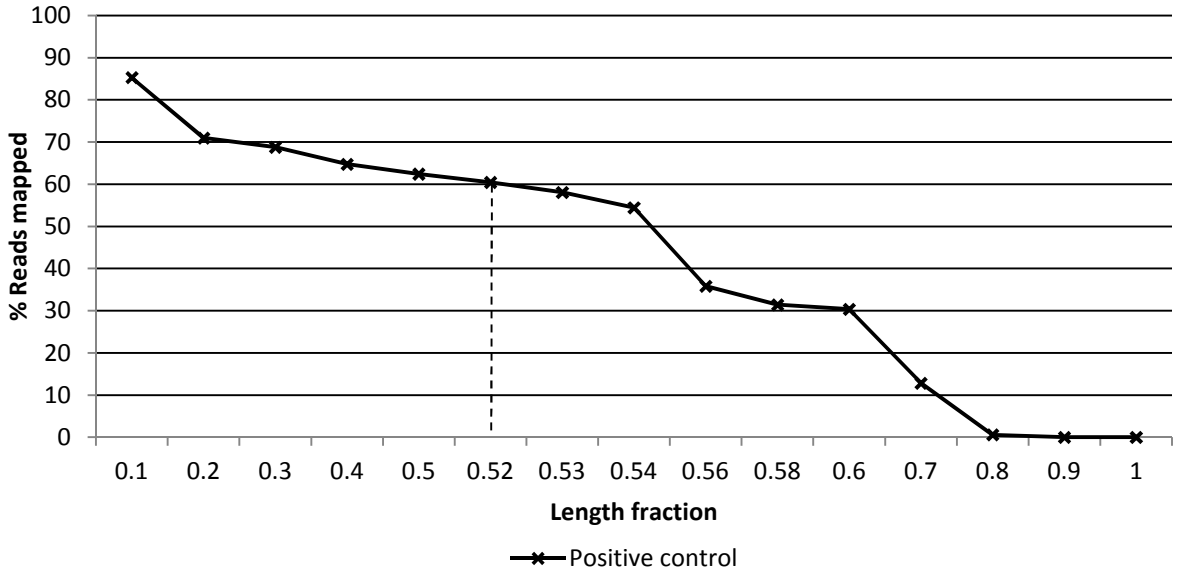


Figure 1: Graph depicting reference mapping optimization of length fraction at a stringent similarity fraction of 0.95

The stability of the parameter was determined by the difference in percentage reads mapped of the parameters preceding and following the parameter, thus the lower the difference the less fluctuation occurs around that parameter. The most stable and stringent parameters for reference mapping was 0.52 length fraction and 0.95 similarity fraction using the 'ignore' function as can be seen in *Figure 1*.

The reads that didn't map during reference mapping were collected and used in de novo assembly to ensure no possibly new viruses would be overlooked. Samples that had a low percentage of total reads mapping during reference mapping produced large (>2000) amounts of contigs. Processing the contigs individually for each sample is impractical, and therefore an arbitrary cut off of 0.20%, based on the amount of reads that mapped back to the contig, was implemented. The de novo parameters were optimised for the 0.20% cut off implemented, testing variations of similarity fraction and length fraction as can be seen in *Figure 2*.

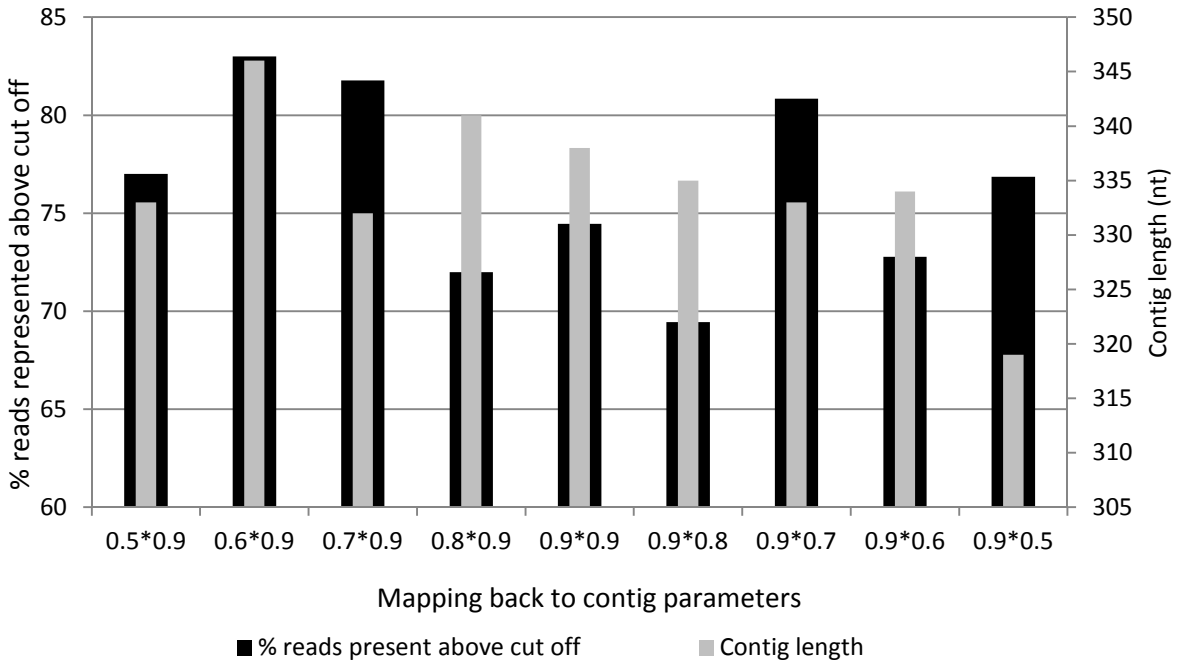


Figure 2: Graph optimizing the best % reads represented above the arbitrary 0.20% cut off of de novo results

The resulting list of contigs, were subsequently identified using the lowest e-value identity, and the corresponding % nucleotide identity and greatest hit length noted. It was observed that each of the contigs shared a signature of highly similar combination of e-value, % nucleotide identity and hit length. Due to this observation the percentage reads that mapped back to each of the contigs exhibiting the same signature were combined to thus yield a clear profile of the viruses present and the amount of reads that mapped back to that virus.

Since some of the reference mappings had as low as 7% of total reads being used, these reference mappings could not be recognised as reliable by itself. The reference mapping was therefore amended with the % reads mapped back to contigs during de novo to produce a whole profile of the viruses present. In addition some classification limits were implemented to the complete profile where if x > 9.15% it is considered present in the sample, 9.15% > x > 0.38% further testing is necessary, and x < 0.38% not present (background noise). A flow chart describing the Illumina MiSeq data analysis of the *Viti-* and *Foveavirus* amplicon and the thought process behind it can be viewed in appendix B *Figure 1.* The optimization based on the positive control can be observed in appendix B *Tables 3 – 9.*

MiSeq Illumina data results

To acquire better resolution into co-infections that possibly occurred in 19 rootstocks (*Table 2*) the *Viti-* and *Foveavirus* amplicons of rootstock samples were subjected to MiSeq Illumina sequencing. An average of 531 978 reads were obtained from Illumina MiSeq sequencing and after trimming the average read length was 242,4bp. The average amount of reads that reference mapped was 112 701 reads (21%) with an average coverage of 102 928 reads. An average of 273 101 (51%) reads were utilized by *de novo* assembly and mapping back, thus 28% of the reads were not utilized.

GVB was detected in 8 out of 19 (53%) samples, GVA in 7 out of 19 (42%), Unknown *Vitivirus* in 6 out of 19 (32%), GRSPaV in 4 out of 19 (26%), GVD in 1 out of 19 (5%), and GVE in 1 out of 19 (5%). The read composition observed for the individual samples can be viewed in *Table 6*. The prevalence of *Viti-* and *Foveavirus* mixed infections was investigated and found that 11 out of 19 (47%) samples had single infections, and 8 out of 19 (53%) samples contained more than one virus. The viruses in single infections from most prevalent to least were GVB (4 out of 11), Unknown (3 out of 11), GVA (2 out of 11), GVE (1 out 11), and GRSPaV (1 out of 11). The most prevalent combination was GVA co-infected with GVB (2 out of 8), then GVA co-infected with unknown virus (2 out of 8), and only one of each of the following combinations: GVB and Unknown; GVA, GVB, and GRSPaV; GVB and GRSPaV; and GVD and unknown.

Table 6: Percentage read composition of individual samples

Rootstock-scion combination		***Viti-* or *Foveavirus***						
Rootstock	Scion	GVA	GVB	GVD	GVE	GVF	GRSPaV	Unknown
101-14	Merlot	1%	11%	0%	0%	0%	4%	84%
R 99	Assyrtiko	0%	0%	0%	0%	0%	100%	0%
R 99	Cabernet franc	0%	0%	0%	0%	0%	0%	100%
R 99	Cabernet franc	0%	100%	0%	0%	0%	0%	0%
R 99	Cabernet franc	18%	81%	0%	0%	0%	1%	0%
R 99	Catarrato commune	0%	0%	61%	0%	0%	0%	38%
R 99	Chardonnay	20%	0%	0%	0%	0%	79%	1%
R 99	Chardonnay	0%	0%	0%	0%	0%	0%	100%
R 99	Lakemont seedless	0%	99%	0%	0%	0%	1%	0%
R 99	Lakemont seedless	0%	4%	0%	0%	0%	0%	96%
R 99	Merlot	0%	65%	0%	0%	0%	35%	0%
R 99	Pinotage	50%	21%	0%	0%	0%	29%	0%
R 99	Pinotage	22%	1%	0%	0%	0%	0%	77%
R 99	Planta nova	50%	50%	0%	0%	0%	0%	0%
R 99	Shiraz	100%	0%	0%	0%	0%	0%	0%
R 99	Zeni	4%	96%	0%	0%	0%	0%	0%
R 99	Zeni	0%	98%	1%	0%	0%	0%	0%
Ramsey	Ruby cabernet	0%	0%	0%	100%	0%	0%	0%
Ramsey	Shiraz	100%	0%	0%	0%	0%	0%	0%

The percentage of the total amount of reads used for each *Viti-* and *Foveavirus* was as follows: GVB (40%), Unknown (34%), GVA (10%), GRSPaV (7%), GVD (5%), and GVE (4%). Detailed information on the read composition calculation of virus in the rootstock tissues can be viewed in Appendix B *Table 10 – 28* and a graphical demonstration of the read composition (*Table 5)* can be vied in *Figure 2* of Appendix B.

Vitivirus confirmation testing

The samples containing unknown viruses were tested using primers targeting coat protein of GVA and GVB initially, and subsequently GVE to eliminate the possibility of other *Vitiviruses*. Of the 19 samples that was MiSeq Illumina sequenced 17 samples were tested for GVA and GVB, and 16 samples tested for GVE according to RNA availability.

Samples were tested with *Vitivirus* specific primers (*Table 1*) to determine the identity of unknown virus. Detection rates using *Vitivirus* specific primers for GVB

was 35% (6 out of 17), GVE 33% (6 out of 17) and. GVA 24% (4 out of 17). The presence of *Vitiviruses* using confirmation testing were compared to that found using Illumina MiSeq sequencing using the same sample set (Figure 6). Sufficient evidence was that GVE had a significantly higher detection rate using specific primer testing that using universal primers combined with Illumina MiSeq ($z = 2.138$; $p<0.05$). The confirmation testing of all the samples can be viewed in Appendix B *Table 29.*

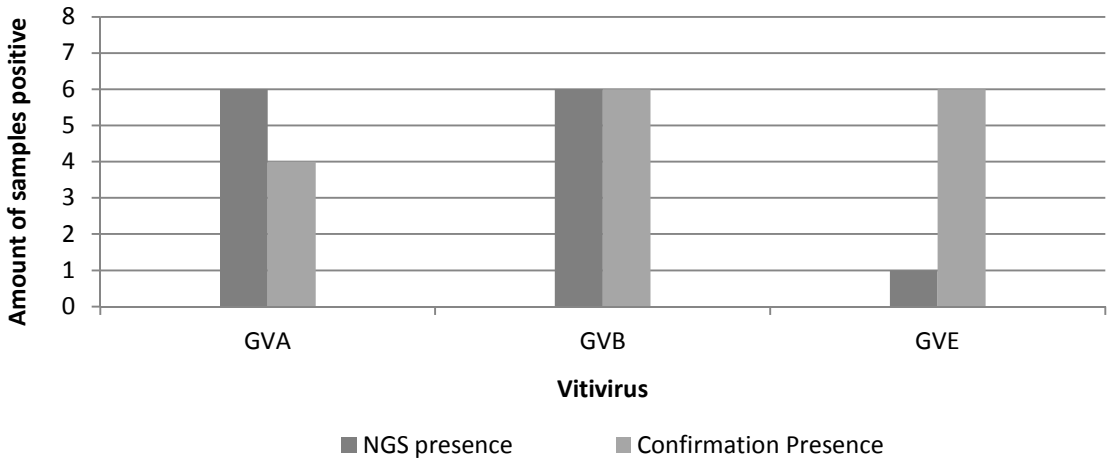


Figure 3: The presence determined using Illumina MiSeq compared to confirmation PCR testing

DISCUSSION

The present study involved determining detection of *Viti-* and *Foveavirus*es in rootstocks versus that of scions of the same plants, and the general population found in rootstocks. A statistically significant difference was found in the detection of *Viti-* and *Foveavirus*es using universal degenerate nested primers with scions having much higher detection than that of corresponding rootstocks. This could be explained by the fact that more GRSPaVs variants occur in scions than in rootstocks (Meng*, et al.*, 2006), though it was observed in this study that GRSPaV are more often the dominant virus in rootstocks than in scions (*Table 4*).

Scions have frequently tested positive for GRSPaVs (Habili*, et al.*, 2006, Alabi*, et al.*, 2010, Coetzee*, et al.*, 2010, Terlizzi*, et al.*, 2010, Jooste*, et al.*, 2015), but this was not observed in this study on samples subjected to direct Sanger sequencing. Unexpectedly the direct Sanger sequencing results intended to identify the dominant

virus present, indicated that GRSPaVs detected was present in rootstocks only. The universal *Viti-* and *Foveavirus* direct Sanger sequencing is limited in its resolution capability because it only indicates the dominant virus in a mixed infection (Petrosino*, et al.*, 2009). Grapevine scions hosts mixed infections of *Viti-* and *Foveavirus*es (Jooste*, et al.*, 2015) and GRSPaV is present in both scion and rootstock but is possibly more dominant in corresponding rootstocks, therefore not detectable using direct Sanger sequencing. The dominance of GRSPaV in rootstocks possibly correlates to the use of rootstocks as indicators of Rupestris stem pitting disease (Meng*, et al.*, 2006).

Scion sample were not analysed using Illumina MiSeq, thus an analysis on the possibility of greater population dominance in rootstocks than in scions could not be further pursued. The rootstock that showed the highest read composition of GRSPaV was Richter 99, this is further highlighted by observing Merlot on two different rootstocks and both infected with GRSPaV (Figure 5). GRSPaV read composition was observed to be higher in Richter 99 than in 101-14 both grafted with merlot.

The identification rate using all three methods, which includes universal primer RT-PCR, amplicon Illumina MiSeq sequencing, and *Vitivirus* specific RT-PCR, indicates that all three detection methods are equally effective in detecting GVA, GVB and GRSPaV in rootstocks. Utilizing direct Sanger sequencing revealed a significant difference in detection between scions and rootstocks with rootstocks having a higher detection of GVB and GRSPaV. As previously mentioned this could be due to the lacking resolution direct Sanger sequencing provides using universal primers. It could therefore be possible that GVB occurs in more dominant titres in rootstocks than in scions.

Jooste et al 2015 found that GVA occurred in a frequency of 19.3% in scions this frequency strongly correlates with observed detection of this study in rootstocks utilizing direct Sanger sequencing ($z=0.065$, $p>0.05$) and MiSeq ($z=1.54$, $p>0.05$). On the contrary the frequency of GVA found in scions that Jooste et al tested is significantly lower than scion samples observed using direct Sanger sequencing in this study ($z=3.251$, $p<0.05$). This could be explained by the nature of the sampling strategy of this study that was not designed to be for surveying purposes.

Direct Sanger Sequencing showed that GVB identification differed significantly between rootstocks and scions ($z=2.3$, $p<0.05$) and is possibly due to mixed infections occurring in the samples, and that GVB is more dominant in rootstocks

than in scion infections. Jooste et al only tested scions and none of them were positive, therefore strongly correlates with the direct Sanger sequencing results ($z=0$, $p>0.05$). When comparing all GVB tests done on rootstock samples to what was found in Jooste et al scion samples there exists a strong possibility that rootstocks in South Africa may be harbouring GVB but not the corresponding scions ($z=8.35$, $p<0.05$). GVB was observed to have the highest read composition overall and especially in Richter 99, this further indicate that GVB occurs in high titres in rootstocks.

There is no difference in detecting GVE using either direct Sanger sequencing or Illumina MiSeq in combination with universal *Viti-* and *Foveavirus* primers. (Dovas & Katis, 2003). Evaluation using selected specific *Vitivirus* primers revealed that the universal *Viti-* and *Foveavirus* primers are lacking in detection of GVE ($z=2.36$, $p<0.05$) (Figure 6). This suggests bias of the universal primers against GVE but can be explained by the date of which these primers were designed. The Dovas et al primers were designed before the discovery of GVE, though the nested and degenerate nature of the primers does allow detection of GVE it is possible that when present in mixed infection GVE is not detected. The degeneracy of the primers used in this study could introduce a large bias towards population estimation. Read *et al* suggest that polyspecific primes targeting heterogeneous binding sites should be revised constantly as new sequences of viruses and variants emerge to prevent bias during PCR amplification of sequences (Read & Pietersen, 2016).

The theory of occurrence of bias can be further corroborated by findings of Jooste et al survey in South Africa where GVE occurred in a frequency of 57.4% in scion samples that differs from what was seen in the direct Sanger sequencing using universal primers of scion samples ($z=3.7$, $p<0.05$). The difference can also be explained by the low resolution of the combination that the universal primers have in combination with Sanger sequencing, but unfortunately no scion material were analysed using Illumina MiSeq. The frequency which GVE was detected in rootstocks using specific primers differed considerably to the frequency Jooste et al found to occur in scions in South Africa ($z=2.16$, $p<0.05$), and suggests the possibility that rootstocks harbour GVE less frequently than scions do.

This study confirms that rootstocks can be infected with *Viti-* and *Foveavirus*es but in a lesser frequency than in scions. The lesser frequent infection of rootstocks could be due to one virus or more. Difference in dominance of *Viti-* and *Foveavirus*es in

rootstocks and scions was observed. *Viti-* and *Foveavirus*es observed in rootstocks using Illumina MiSeq do not correlate with what one would expect to find in South Africa. This suggests that rootstocks should be given the same consideration in diagnostics as scions. Further studies can pursue the possible differences found in some *Vitivirus* infection in rootstocks to further resolve the dynamics involved in rootstock-virus interaction, and further better diagnostic strategies in rootstocks.

REFERENCES

Adams M., Antoniw J., Bar-Joseph M.*, et al.* (2004). Virology Division News: The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. Archives of Virology vol.149 pp1045-1060

Al Rwahnih M., Sudarshana M.R., Uyemoto J.K. and Rowhani A. (2012). Complete genome sequence of a novel *Vitivirus* isolated from grapevine. Journal of Virology vol.86 pp9545-9545

Alabi O.J., Martin R.R. and Naidu R.A. (2010). Sequence diversity, population genetics and potential recombination events in *Grapevine rupestris stem pitting-associated virus* in Pacific North-West vineyards. Journal of General Virology vol.91 pp265-276

Bonavia M., Digiaro M., Boscia D., Boari A., Bottalico G., Savino V. and Martelli G. (1996). Studies on "corky rugose wood" of grapevine and on the diagnosis of *Grapevine virus B*. Vitis vol.35 pp53-58

Boscia D., Digiaro M., Safi M., Garau R., Zhou Z., Minafra A. and Ghanem-Sabanadzovic N.A. (2001). Production of monoclonal antibodies to *Grapevine virus B* and contribution to the study of its aetiological role in grapevine diseases. VITIS-Journal of Grapevine Research vol.40 pp69-74

Bouyahia H., Boscia D., Savino V.*, et al.* (2015). *Grapevine rupestris stem pitting-associated virus* is linked with grapevine vein necrosis. VITIS-Journal of Grapevine Research vol.44 pp133-137

Coetzee B., Freeborough M.-J., Maree H.J., Celton J.-M., Rees D.J.G. and Burger J.T. (2010). Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard. Virology vol.400 pp157-163

Credi R. (1997). Characterization of grapevine rugose wood disease sources from Italy. Plant Disease vol.81 pp1288-1292

Dovas C. and Katis N. (2003). A spot multiplex nested RT-PCR for the simultaneous and generic detection of viruses involved in the aetiology of grapevine leafroll and rugose wood of grapevine. Journal of Virological Methods vol.109 pp217-226

Garau R., Prota V.A., Piredda R., Boscia D. and Prota U. (1994). On the possible relationship between Kober stem grooving and *Grapevine virus A*. Vitis vol.33 pp161-163

Golino D.A., Wolpert J., Sim B.J. and Aderson R.A. (2008). Virus effects on vine growth and fruit components of Cabernet Sauvignon on six rootstocks. In: *Proceedings of the second annual national viticulture research conference*,

Goszczynski D. (2007). Single-strand conformation polymorphism (SSCP), cloning and sequencing reveal a close association between related molecular variants of *Grapevine virus A* (GVA) and Shiraz disease in South Africa. Plant Pathology vol.56 pp755-762

Habili N., Farrokhi N., Lima M., Nicholas P. and Randles J. (2006). Distribution of *Rupestris stem-pitting-associated virus* variants in two Australian vineyards showing different symptoms. Annals of Applied Biology vol.148 pp91-96

Hall T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: *Nucleic acids symposium series*, vol.41

ICTV (2015). Virus Taxonomy: 2015 Release. Date acessed: http://www.ictvonline.org/virusTaxonomy.asp?version2015

Jooste A.E., Molenaar N., Maree H.J., Bester R., Morey L., de Koker W.C. and Burger J.T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. European Journal of Plant Pathology vol.142 pp363-375

Lounsbury C. (1940). The pioneer period of economic entomology in South Africa. J Entomol Soc S Afr vol.3 pp9-29

Martelli G. (1993). *Graft-transmissible diseases of grapevines: handbook for detection and diagnosis*. Food & Agriculture Org.

Martelli G.P., Adams M.J., Kreuze J.F. and Dolja V.V. (2007). Family *Flexiviridae*: a case study in virion and genome plasticity. Annu. Rev. Phytopathol. vol.45 pp73-100

Meng B., Rebelo A.R. and Fisher H. (2006). Genetic diversity analyses of grapevine *Rupestris stem pitting-associated virus* reveal distinct population structures in scion versus rootstock varieties. Journal of General Virology vol.87 pp1725-1733

Meng B., Credi R., Petrovic N., Tomazic I. and Gonsalves D. (2003). Antiserum to recombinant virus coat protein detects *Rupestris stem pitting associated virus* in grapevines. Plant Disease vol.87 pp515-522

Petrosino J.F., Highlander S., Luna R.A., Gibbs R.A. and Versalovic J. (2009). Metagenomic pyrosequencing and microbial identification. Clinical Chemistry vol.55 pp856-866

Read D.A. and Pietersen G. (2016). PCR bias associated with conserved primer binding sites, used to determine genotype diversity within *Citrus tristeza virus* populations. Journal of Virological Methods vol.237 pp107-113

Rosa C. and Rowhani A. (2007). Etiology of “Rugose Wood Complex”. University of California, California

Terlizzi F., Ratti C., Filippini G., Pisi A. and Credi R. (2010). Detection and molecular characterization of Italian *Grapevine rupestris stem pitting-associated virus* isolates. Plant Pathology vol.59 pp48-58

Töpfer R., Hausmann L., Harst M., Maul E., Zyprian E. and Eibach R. (2011). New horizons for grapevine breeding. Methods In Temperate Fruit Breeding, Fruit, Vegetable and Cereal Science and Biotechnology vol.5 pp79-100

White E.J., Venter M., Hiten N.F. and Burger J.T. (2008). Modified Cetyltrimethylammonium bromide method improves robustness and versatility: the benchmark for plant RNA extraction. Biotechnology Journal vol.3 pp1424-1428

Chapter 5

Concluding Remarks

An aim of this study was to determine if GLRaV-3 can infect rootstocks commonly used in South Africa, if so whether selection of GLRaV-3 variants occur in rootstocks. Combining the approaches of Goszczynski, (2013) and David and Pietersen, (2016) allowed for unbiased resolution of the GLRaV-3 variant populations. Scion and Rootstock tissues were sampled from 95 vines in the Western Cape Province, allowing for comparison of GLRaV-3 detection and populations between the two tissues of a single vine. Though GLRaV-3 was detected in rootstocks, the detection was significantly lower in Richter 99 rootstock tissues and in low amplicon levels than that of the corresponding scion tissues of grapevines collected locally. PCR inhibitory substances were not responsible for the poor detection in the rootstocks, since the same rootstocks more often tested positive for *Viti-* and *Foveaviruses*. Populations of GLRaV-3 variants differed between scion tissue and rootstock tissue that were both positive for GLRaV-3 in the same vines. No definitive pattern could be discerned between populations in rootstock and scion tissues, though GLRaV-3 variant group VI did appear to have slightly higher incidence in rootstocks than in scions. The majority of the samples differed in GLRaV-3 read composition therefore also population between rootstocks and scion tissues of the same vine, though some populations were the same, thus corroborating previous observations of the erratic behaviour of GLRaV-3 in rootstocks (Cid*, et al.*, 2003, Cohen*, et al.*, 2003, Chooi*, et al.*, 2016). The observations made in this study suggest that rootstocks more commonly used in South Africa by the South African grapevine certification scheme should be screened for GLRaV-3. The dynamics between GLRaV-3 and rootstocks remain largely unknown therefore this study aids in closing a gap in the knowledge regarding this issue.

GLD diseased vines in South Africa are generally not only infected with GLRaV-3 but often co-infected with various *Vitiviruses* (Jooste*, et al.*, 2015). A method for simultaneous detection of *Vitivirus* and *Foveavirus* species was developed by Dovas and Katis, (2003) that included a one step RT-PCR followed by a nested-PCR in conjunction with universal degenerate primers (Dovas & Katis, 2003). Combining this

assay with NGS technologies provides a high throughput method of detection and identification of *Viti-* and *Foveaviruses* in grapevines and the capacity for detection of novel viruses. The second aim of this study was to characterize the populations of *Viti-* and *Foveaviruses* found in rootstocks commonly used in South Africa, in addition to comparison of detection between rootstock and scion tissues of the same vines. The same 190 samples used for the detection of GLRaVs were used for the detection of *Viti-* and *Foveaviruses* using the primers designed by Dovas and Katis, 2003. The dominant viruses present in the amplicons were determined by direct Sanger sequencing of amplicons obtained from 35 rootstocks and 20 scions that tested positive for *Viti-* and *Foveaviruses*. Amplicons of 19 positive rootstocks were sequenced using Illumina MiSeq technologies to fully characterize populations present in various rootstocks.

Rootstocks exhibited significantly lower detection of *Viti-* and *Foveaviruses* than that of corresponding scion tissues. Direct Sanger sequencing revealed that dominant viruses in scions and rootstocks differed, with GVB identified to be more dominant in rootstocks than in scions. Illumina MiSeq sequencing suggested that GVB had the highest incidence in the rootstocks, in addition to the presence of an unknown virus. GVB being the dominant *Vitivirus* in rootstocks was an unexpected observation since South African survey studies have shown GVE to be the dominant Vitivirus in scion cultivars (Jooste*, et al.*, 2015). The dominance of GVB in rootstocks could be explained by its involvement with Corky bark disease for which some rootstocks are used as indicators (Habili*, et al.*, 1992, Martelli, 1993), further suggesting a high incidence of Corky bark disease in the vineyards sampled. Therefore it is necessary to do more rigorous GVB testing of rootstocks in the certification scheme.

The unknown *Vitivirus* could not be identified using *Vitivirus* specific primers, though it did reveal a bias in the universal degenerate primers used for the simultaneous detection of *Viti-* and *Foveaviruses*. The *Vitivirus* specific testing suggested a bias of the universal primers toward GVE even though the universal primers are able to detect GVE, which could explain the unexpected lack of GVE presence. The primers designed by Dovas and Katis, 2003 were designed using the limited sequences available at that time, since then many more *Vitivirus* species have been identified. Furthermore the degeneracy of the primers could have a large impact on the viral composition of the amplicons by introducing primer-associated

bias (Dieffenbach*, et al.*, 1993). Future studies would better benefit from using more specific primers in determining *Viti-* and *Foveavirus* populations in grapevines. The certification scheme would gain from identification of novel virus identification by adjusting variety of diagnostics used.

This study yielded one of the most comprehensive characterizations of GLD associated viruses in rootstocks popular in the South African grapevine industry to date. Comparison of GLRaV-3 variant populations in rootstock versus that of scion tissue of the same vine in combination with the in depth identification that a NGS platform can provide, led to detailed observations of the dynamics between GLRaV-3 and rootstocks commonly used in South Africa. Much of the dynamics between GLD and rootstocks remain unexplored but observations made in this study highlight the necessity of investigating this subject.

REFERENCES

Chooi K.M., Cohen D. and Pearson M.N. (2016). Differential distribution and titre of selected *Grapevine leafroll-associated virus 3* genetic variants within grapevine rootstocks. Archives of Virology vol.161 pp1371-1375

Cid M., Cabaleiro C. and Segura A. (2003). Detection of *Grapevine leafroll-associated virus 3* in rootstocks. In: *Extended abstracts of the 14th Meeting of ICVG, Locorotondo*, Locorotondo,

Cohen D., Van Den Brink R. and Habili N. (2003). Leafroll virus movement in newly infected grapevines. In: *Extended abstracts 14th Meeting ICVG, Locorotondo, Italy*,
Dieffenbach C., Lowe T. and Dveksler G. (1993). General concepts for PCR primer design. PCR Methods Appl vol.3 pp30-37

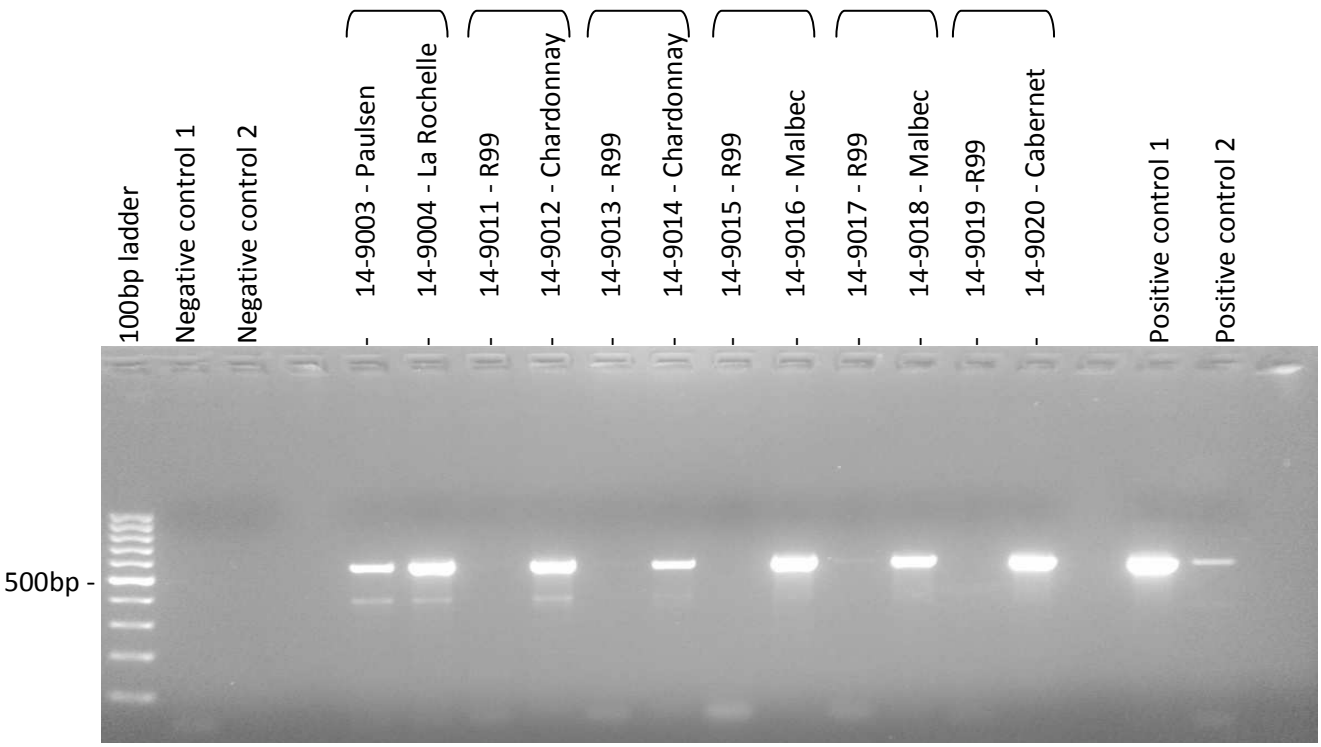
Dovas C. and Katis N. (2003). A spot nested RT-PCR method for the simultaneous detection of members of the *Vitivirus* and *Foveavirus* genera in grapevine. Journal of Virological Methods vol.107 pp99-106

Habili N., Krake L., Barlass M. and Rezaian M. (1992). Evaluation of biological indexing and dsRNA analysis in *Grapevine virus E*limination. Annals of Applied Biology vol.121 pp277-283

Jooste A.E., Molenaar N., Maree H.J., Bester R., Morey L., de Koker W.C. and Burger J.T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. European Journal of Plant Pathology vol.142 pp363-375

Martelli G. (1993). *Graft-transmissible diseases of grapevines: handbook for detection and diagnosis*. Food & Agriculture Org.

APPENDIX A



Negative control 1 = cDNA synthesis control, Negative control 2 = PCR control, Positive 1 = RNA cDNA synthesis control, and Positive 2 = GLRaV-3 amplicon PCR control.

Figure 1: Image of an example of a 2% agarose gel electrophoresis performed of grapevine leafroll associated virus -3 (GLRaV-3) amplicons obtained from grapevine rootstock and scion tissues of the same vines

Table 1: GLRaV-3 status of rootstock and scion tissue combinations of individual grapevines with PCR band strength observed

Accession # (Rootstock/Scion)	Rootstock-scion combination		Scion cultivar colour	Rootstocks GLRaV-3 status	Scions GLRaV-3 status	Selected for NGS
	Rootstock	Scion				
14-9071/9072	**101-14**	**Cabernet Sauvignon**	Red	+	++	-
15-9041/9042	**101-14**	**Cabernet Sauvignon**	Red	-	-	-
14-9073/9074	**101-14**	**Merlot**	Red	++	++	*
15-5021/5022	**101-14**	**Merlot**	Red	-	+++	-
15-5023/5024	**101-14**	**Merlot**	Red	++	-	-
15-5025/5026	**101-14**	**Merlot**	Red	-	+++	-
15-5027/5028	**101-14**	**Merlot**	Red	-	+++	-
15-5029/5030	**101-14**	**Merlot**	Red	-	+++	-
14-9001/9002	**Paulsen**	**La Rochelle**	Red	+	++	*
14/9003/9004	**Paulsen**	**La Rochelle**	Red	++	++	-
14-9005/9006	**R 99**	**Assyrtiko**	White	-	++	-
14-9007/0908	**R 99**	**Catarratto commune**	White	+	++	-
14-9009/9010	**R 99**	**Catarratto commune**	White	-	++	-
14-9019/9020	**R 99**	**Cabernet franc**	Red	-	++	*
14-9021/9022	**R 99**	**Cabernet franc**	Red	+	++	*
14-9057/9058	**R 99**	**Cabernet franc**	Red	-	++	-
14-9059/9060	**R 99**	**Cabernet franc**	Red	++	-	-
16-0045/0046	**R 99**	**Cabernet Sauvignon**	Red	++	+++	-
16-0047/0048	**R 99**	**Cabernet Sauvignon**	Red	+++	+++	-
16-0049/0050	**R 99**	**Cabernet Sauvignon**	Red	+++	+++	-
14-9049/9050	**R 99**	**CG 40318**	White	-	++	-
14-9051/9052	**R 99**	**CG 40318**	White	-	++	-
14-9011/9012	**R 99**	**Chardonnay**	White	-	++	-
14-9013/9014	**R 99**	**Chardonnay**	White	+	++	-
15-5031/5032	**R 99**	**Chardonnay**	White	-	+++	-
15-5033/5034	**R 99**	**Chardonnay**	White	-	+++	-
15-5035/5036	**R 99**	**Chardonnay**	White	-	+++	-
15-5037/5038	**R 99**	**Chardonnay**	White	-	+++	-
15-5039/5040	**R 99**	**Chardonnay**	White	-	+++	-
15-5083/5084	**R 99**	**Chardonnay**	White	-	+++	-
16-0039/0040	**R 99**	**Cinsault**	Red	++	+++	-
16-0041/0042	**R 99**	**Cinsault**	Red	+	+++	-
16-0043/0044	**R 99**	**Cinsault**	Red	++	+++	-
14-9045/9046	**R 99**	**Gamay hatif des vosges**	Red	-	++	-
14-9047/9048	**R 99**	**Gamay hatif des vosges**	Red	-	++	-
14-9037/9038	**R 99**	**Lakemont seedless**	White	+	++	-
14-9039/9040	**R 99**	**Lakemont seedless**	White	-	++	-
14-9029/9030	**R 99**	**L. Red**	Red	+	-	-
14-9031/9032	**R 99**	**L. Red**	Red	+	+	*
14-9033/9034	**R 99**	**Lambrusco**	Red	-	++	-
14-9035/9036	**R 99**	**Lambrusco**	Red	-	++	-
14-9015/9016	**R 99**	**Malbec**	Red	-	++	-
14-9017/9018	**R 99**	**Malbec**	Red	+	++	-
16-0017/0018	**R 99**	**Malbec**	Red	+	+++	-
16-0019/0020	**R 99**	**Malbec**	Red	+	+++	-
16-0021/0022	**R 99**	**Malbec**	Red	-	+++	-
15-5011/5012	**R 99**	**Merlot**	Red	-	++	-
15-5013/5014	**R 99**	**Merlot**	Red	-	-	-
15-5015/5016	**R 99**	**Merlot**	Red	++	+++	*
15-5017/5018	**R 99**	**Merlot**	Red	++	+++	*
15-5019/5020	**R 99**	**Merlot**	Red	-	+++	-
15-5045/5046	**R 99**	**Merlot**	Red	-	+++	-
15-5047/5048	**R 99**	**Merlot**	Red	-	+++	-
15-5049/5050	**R 99**	**Merlot**	Red	-	+++	-
15-5051/5052	**R 99**	**Merlot**	Red	-	+++	-
15-5053/5054	**R 99**	**Merlot**	Red	+++	-	-

Accession # (Rootstock/Scion)	Scion-rootstock combination		Scion cultivar colour	Rootstocks GLRaV-3 status	Scions GLRaV-3 status	Selected for NGS
	Rootstock	Scion				
15-5055/5056	**R 99**	**Merlot**	Red	-	++	-
15-5057/5058	**R 99**	**Merlot**	Red	-	+++	-
15-5059/5060	**R 99**	**Merlot**	Red	-	+++	-
15-5061/5062	**R 99**	**Merlot**	Red	-	+++	-
15-5063/5064	**R 99**	**Merlot**	Red	-	+++	-
15-5085/5086	**R 99**	**Merlot**	Red	-	+++	-
14-9041/9042	**R 99**	**Planta nova**	Red	-	++	-
14-9043/9044	**R 99**	**Planta nova**	Red	-	+	-
14-9077/9078	**R 99**	**Pinotage**	Red	-	++	-
14-9079/9080	**R 99**	**Pinotage**	Red	-	++	-
14-9081/9082	**R 99**	**Pinotage**	Red	-	-	-
14-9083/9084	**R 99**	**Pinotage**	Red	-	++	-
16-0023/0024	**R 99**	**Pinotage**	Red	+	+++	-
16-0025/0026	**R 99**	**Pinotage**	Red	+	+++	-
16-0027/0028	**R 99**	**Pinotage**	Red	-	+++	-
16-0033/0034	**R 99**	**Pontac**	Red	-	+++	-
16-0035/0036	**R 99**	**Pontac**	Red	-	+++	-
16-0037/0038	**R 99**	**Pontac**	Red	++	+++	-
16-0011/0012	**R 99**	**Ruby Cabernet**	Red	-	+++	-
16-0013/0014	**R 99**	**Ruby Cabernet**	Red	+++	+++	-
16-0015/0016	**R 99**	**Ruby Cabernet**	Red	++	+++	-
15-5065/5066	**R 99**	**Shiraz**	Red	++	+++	-
15-5067/5068	**R 99**	**Shiraz**	Red	+++	+++	*
15-5069/5070	**R 99**	**Shiraz**	Red	-	+++	-
15-5073/5074	**R 99**	**Shiraz**	Red	-	+++	-
15-5075/5076	**R 99**	**Shiraz**	Red	-	+++	-
16-0029/0030	**R 99**	**Tinta barrocca**	Red	+	+++	-
16-0031/0032	**R 99**	**Tinta barrocca**	Red	++	+++	-
14-9053/9054	**R 99**	**Zeni**	White	-	++	-
14-9055/9056	**R 99**	**Zeni**	White	-	++	-
15-5001/5002	**Ramsey**	**Ruby Cabernet**	Red	+++	++	-
15-5003/5004	**Ramsey**	**Ruby Cabernet**	Red	++	+++	-
15-5005/5006	**Ramsey**	**Ruby Cabernet**	Red	+++	+++	-
15-5007/5008	**Ramsey**	**Ruby Cabernet**	Red	+++	+++	*
15-5009/5010	**Ramsey**	**Ruby Cabernet**	Red	+++	+++	*
15-5081/5082	**Ramsey**	**Ruby Cabernet**	Red	-	+++	-
15-5043/5044	**Ramsey**	**Shiraz**	Red	+++	+++	*
15-5077/5078	**Ramsey**	**Shiraz**	Red	+++	+++	-
15-5079/5080	**Ramsey**	**Shiraz**	Red	+++	+++	-
		Total		41	88	11

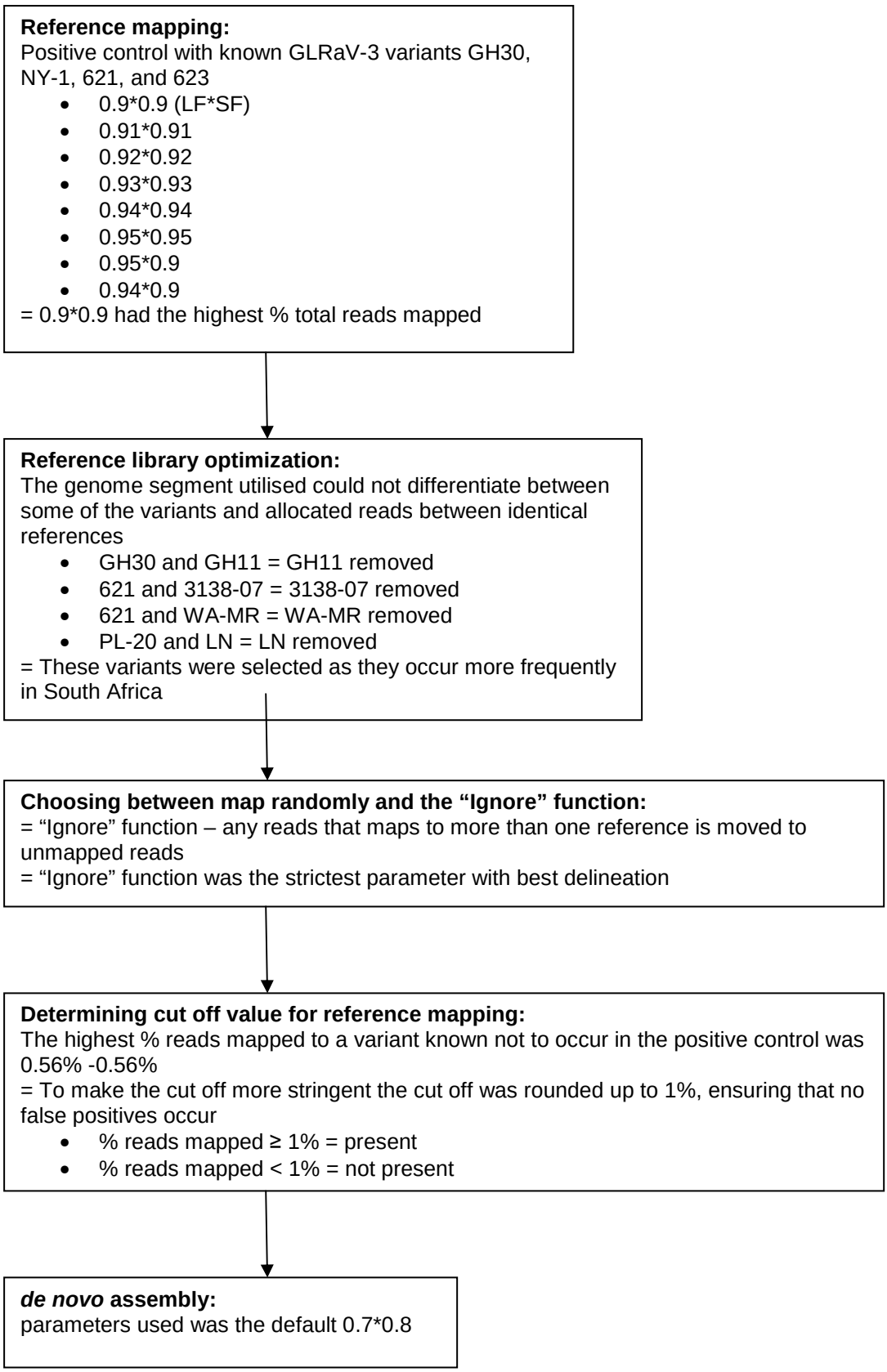
(-) = Negative for GLRaV-3; (+) = Weak PCR band, (++) = Medium PCR band; and (+++) = Strong PCR band. (*) = Selected for further, NGS analysis and (-) = Not selected for NGS analysis.

Table 2: Subset of samples tested for GLRaV-1; -2; -4-like; and -7 in rootstock and scion tissue combinations and associated PCR band strength

Accession # (Rootstock/Scion)	Rootstock-scion combination		GLRaV-1 status		GLRaV-2 status		GLRaV-4-like status		GLRaV-7 status	
	Rootstock	Scion	Rootstocks	Scions	Rootstocks	Scions	Rootstocks	Scions	Rootstocks	Scions
15-9041/9042	**101-14**	**Cabernet Sauvignon**	-	-	-	-	-	-	-	-
15-5021/5022	**101-14**	**Merlot**	-	-	-	-	-	-	-	-
15-5023/5024	**101-14**	**Merlot**	-	-	-	-	-	-	-	-
15-5025/5026	**101-14**	**Merlot**	-	-	-	-	-	-	-	-
15-5027/5028	**101-14**	**Merlot**	-	-	-	-	-	-	-	-
15-5029/5030	**101-14**	**Merlot**	-	-	-	-	-	-	-	-
16-0045/0046	**R 99**	**Cabernet Sauvignon**	-	-	-	-	-	-	-	-
16-0047/0048	**R 99**	**Cabernet Sauvignon**	-	-	-	-	-	-	-	-
16-0049/0050	**R 99**	**Cabernet Sauvignon**	-	-	-	-	-	-	-	-
15-5031/5032	**R 99**	**Chardonnay**	-	-	-	-	-	-	-	-
15-5033/5034	**R 99**	**Chardonnay**	-	-	-	-	-	-	-	-
15-5035/5036	**R 99**	**Chardonnay**	-	-	-	-	-	-	-	-
15-5037/5038	**R 99**	**Chardonnay**	-	-	-	-	-	-	-	-
15-5039/5040	**R 99**	**Chardonnay**	-	-	-	-	-	-	-	-
15-5083/5084	**R 99**	**Chardonnay**	-	-	-	-	-	-	-	-
16-0039/0040	**R 99**	**Cinsault**	-	-	-	-	-	-	-	-
16-0041/0042	**R 99**	**Cinsault**	-	-	-	-	-	-	-	-
16-0043/0044	**R 99**	**Cinsault**	-	-	-	-	-	-	-	-
16-0017/0018	**R 99**	**Malbec**	-	-	-	-	-	-	-	-
16-0019/0020	**R 99**	**Malbec**	-	-	-	-	-	-	-	-
16-0021/0022	**R 99**	**Malbec**	-	-	-	-	-	-	-	-
15-5011/5012	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
15-5013/5014	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
15-5015/5016	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
15-5017/5018	**R 99**	**Merlot**	+	-	-	-	-	-	-	-
15-5019/5020	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
15-5045/5046	**R 99**	**Merlot**	-	-	++	++	-	-	-	-
15-5047/5048	**R 99**	**Merlot**	-	+	-	-	-	-	-	-
15-5049/5050	**R 99**	**Merlot**	-	-	-	++	-	-	-	-
15-5051/5052	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
15-5053/5054	**R 99**	**Merlot**	+	-	++	-	-	-	-	-
15-5055/5056	**R 99**	**Merlot**	-	-	++	-	-	-	-	-
15-5057/5058	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
15-5059/5060	**R 99**	**Merlot**	-	-	-	+++	-	-	-	-
15-5061/5062	**R 99**	**Merlot**	-	-	-	-	-	-	-	-

Accession # (Rootstock/Scion)	Rootstock-scion combination		GLRaV-1 status		GLRaV-2 status		GLRaV-4-like status		GLRaV-7 status	
	Rootstock	Scion	Rootstocks	Scions	Rootstocks	Scions	Rootstocks	Scions	Rootstocks	Scions
15-5063/5064	**R 99**	**Merlot**	-	-	-	++	-	-	-	-
15-5085/5086	**R 99**	**Merlot**	-	-	-	-	-	-	-	-
16-0023/0024	**R 99**	**Pinotage**	-	-	-	-	-	-	-	-
16-0025/0026	**R 99**	**Pinotage**	-	-	-	-	-	-	-	-
16-0027/0028	**R 99**	**Pinotage**	-	-	-	-	-	-	-	-
16-0033/0034	**R 99**	**Pontac**	-	-	-	-	-	-	-	-
16-0035/0036	**R 99**	**Pontac**	-	-	-	-	-	-	-	-
16-0037/0038	**R 99**	**Pontac**	-	-	-	-	-	-	-	-
16-0011/0012	**R 99**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
16-0013/0014	**R 99**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
16-0015/0016	**R 99**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5065/5066	**R 99**	**Shiraz**	-	-	-	+++	-	-	-	-
15-5067/5068	**R 99**	**Shiraz**	-	-	+	+++	-	-	-	-
15-5069/5070	**R 99**	**Shiraz**	-	-	-	-	-	-	-	-
15-5073/5074	**R 99**	**Shiraz**	-	-	-	++	-	-	-	-
15-5075/5076	**R 99**	**Shiraz**	-	-	-	++	-	-	-	-
16-0029/0030	**R 99**	**Tinta barrocca**	-	-	-	-	-	-	-	-
16-0031/0032	**R 99**	**Tinta barrocca**	-	-	-	-	-	-	-	-
15-5001/5002	**Ramsey**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5003/5004	**Ramsey**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5005/5006	**Ramsey**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5007/5008	**Ramsey**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5009/5010	**Ramsey**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5081/5082	**Ramsey**	**Ruby Cabernet**	-	-	-	-	-	-	-	-
15-5043/5044	**Ramsey**	**Shiraz**	-	-	-	-	-	-	-	-
15-5077/5078	**Ramsey**	**Shiraz**	-	-	-	-	-	-	-	-
15-5079/5080	**Ramsey**	**Shiraz**	-	+	-	-	-	-	-	-
Total										

(-) = Negative; (+) = Weak PCR band; (++) = Medium PCR band; and (+++) = Strong PCR band


Reference mapping:
Positive control with known GLRaV-3 variants GH30, NY-1, 621, and 623
0.9*0.9 (LF*SF)
0.91*0.91
0.92*0.92
0.93*0.93
0.94*0.94
0.95*0.95
0.95*0.9
0.94*0.9
= 0.9*0.9 had the highest % total reads mapped
Reference library optimization:
The genome segment utilised could not differentiate between some of the variants and allocated reads between identical references
GH30 and GH11 = GH11 removed
621 and 3138-07 = 3138-07 removed
621 and WA-MR = WA-MR removed
PL-20 and LN = LN removed
= These variants were selected as they occur more frequently in South Africa
Choosing between map randomly and the "Ignore" function:
= "Ignore" function – any reads that maps to more than one reference is moved to unmapped reads
= "Ignore" function was the strictest parameter with best delineation
Determining cut off value for reference mapping:
The highest % reads mapped to a variant known not to occur in the positive control was 0.56% -0.56%
= To make the cut off more stringent the cut off was rounded up to 1%, ensuring that no false positives occur
% reads mapped ≥ 1% = present
% reads mapped < 1% = not present
***de novo* assembly:**
parameters used was the default 0.7*0.8

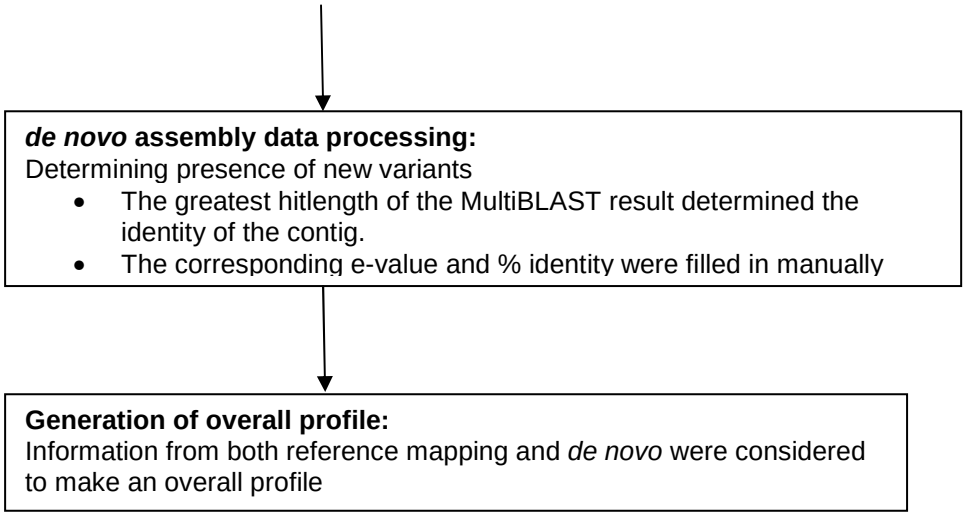


Figure 2: Flow chart detailing the optimization of Illumina MiSeq data using CLC genomic workbench 6.0

*Table 3: Number of reads of **positive control** mapped to references at **0.9*0.9** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total mapped reads	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	170250	53014	31.14%	23124	29890	26171.17	NY-1
GH30 mapping	513	170250	41974	24.65%	2218	39756	21374	GH30
3138-07 mapping	513	170250	19110	11.22%	5826	13284	9483.396	3138-07
621 mapping	513	170250	19020	11.17%	5864	13156	9423.201	621
623 mapping	513	170250	17062	10.02%	16942	120	8429.396	623
GH11 mapping	513	170250	10032	5.89%	1310	8722	4871.585	GH11
WA-MR mapping	513	170250	4992	2.93%	3858	1134	2424.737	WA-MR
GP18 mapping	513	170250	3835	2.25%	3763	72	1595.413	GP18
Cl-766 mapping	513	170250	1210	0.71%	1210	0	570.7466	Cl-766
PL-20 mapping	257	170250	1	0.00%	1	0	0.500975	PL-20
139	0	170250	0	0.00%	0	0	0	139
LN_CHINA	0	170250	0	0.00%	0	0	0	LN_CHINA

*Table 4: Number of reads of **positive control** mapped to references at **0.91*0.91** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	164790	51202	31.07%	23140	28062	25468.71	NY-1
GH30 mapping	513	164790	41443	25.15%	2805	38638	21179.09	GH30
3138-07 mapping	513	164790	18577	11.27%	5949	12628	9280.591	3138-07
621 mapping	513	164790	18266	11.08%	6090	12176	9118.899	621
623 mapping	513	164790	16588	10.07%	16500	88	8240.261	623
GH11 mapping	513	164790	9476	5.75%	1578	7898	4637.76	GH11
WA-MR mapping	513	164790	4728	2.87%	3738	990	2320.977	WA-MR
GP18 mapping	513	164790	3413	2.07%	3359	54	1446.924	GP18
Cl-766 mapping	512	164790	1096	0.67%	1096	0	526.154	Cl-766
PL-20 mapping	257	164790	1	0.00%	1	0	0.500975	PL-20
139	0	164790	0	0.00%	0	0	0	139
LN_CHINA	0	164790	0	0.00%	0	0	0	LN_CHINA

*Table 5: Number of reads of **positive control** mapped to references at **0.92*0.92** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	145975	45413	31.11%	24391	21022	23146.64	NY-1
GH30 mapping	513	145975	38544	26.40%	5734	32810	20066.97	GH30
621 mapping	513	145975	16330	11.19%	6542	9788	8343.865	621
3138-07 mapping	513	145975	16202	11.10%	6440	9762	8261.501	3138-07
623 mapping	513	145975	14037	9.62%	13979	58	7230.901	623
GH11 mapping	513	145975	7920	5.43%	2366	5554	3987.764	GH11
WA-MR mapping	513	145975	4240	2.90%	3544	696	2169.023	WA-MR
GP18 mapping	513	145975	2395	1.64%	2371	24	1058.52	GP18
Cl-766 mapping	288	145975	894	0.61%	894	0	443.729	Cl-766
139	0	145975	0	0.00%	0	0	0	139
LN_CHINA	0	145975	0	0.00%	0	0	0	LN_CHINA
PL-20	0	145975	0	0.00%	0	0	0	PL-20

*Table 6: Number of reads of **positive control** mapped to references at **0.93*0.93** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	113950	34883	30.61%	24237	10646	18505.19	NY-1
GH30 mapping	513	113950	31817	27.92%	10387	21430	17063.92	GH30
3138-07 mapping	513	113950	12246	10.75%	7388	4858	6503.737	3138-07
621 mapping	513	113950	12224	10.73%	7428	4796	6492.074	621
623 mapping	513	113950	11654	10.23%	11634	20	6169.55	623
GH11 mapping	513	113950	6140	5.39%	4306	1834	3282.448	GH11
WA-MR mapping	513	113950	3750	3.29%	3436	314	1996.033	WA-MR
GP18 mapping	513	113950	705	0.62%	703	2	356.6472	GP18
Cl-766 mapping	288	113950	531	0.47%	531	0	281.6023	Cl-766
139	0	113950	0	0.00%	0	0	0	139
LN_CHINA	0	113950	0	0.00%	0	0	0	LN_CHINA
PL-20	0	113950	0	0.00%	0	0	0	PL-20

*Table 7: Number of reads of **positive control** mapped to references at **0.94*0.94** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
GH30 mapping	513	45261	13444	29.70%	10426	3018	7367.029	GH30
NY-1 mapping	513	45261	13441	29.70%	12407	1034	7319.234	NY-1
GH11 mapping	513	45261	4636	10.24%	4490	146	2542.24	GH11
3138-07 mapping	513	45261	4008	8.86%	3600	408	2176.452	3138-07
621 mapping	513	45261	3959	8.75%	3515	444	2149.285	621
623 mapping	513	45261	3768	8.33%	3760	8	2026.146	623
WA-MR mapping	513	45261	1617	3.57%	1587	30	879.6706	WA-MR
GP18 mapping	513	45261	227	0.50%	225	2	117.9688	GP18
Cl-766 mapping	287	45261	161	0.36%	161	0	88.00975	Cl-766
139	0	45261	0	0.00%	0	0	0	139
LN_CHINA	0	45261	0	0.00%	0	0	0	LN_CHINA
PL-20	0	45261	0	0.00%	0	0	0	PL-20

*Table 8: Number of reads of **positive control** mapped to references at **0.95*0.95** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	4557	1296	28.44%	1248	48	667.7836	NY-1
GH30 mapping	513	4557	787	17.27%	745	42	415.1462	GH30
623 mapping	444	4557	756	16.59%	746	10	382.8889	623
3138-07 mapping	513	4557	579	12.71%	569	10	295.8265	3138-07
621 mapping	513	4557	524	11.50%	508	16	271.1774	621
GH11 mapping	513	4557	418	9.17%	416	2	212.2456	GH11
WA-MR mapping	513	4557	150	3.29%	148	2	76.89084	WA-MR
GP18 mapping	513	4557	42	0.92%	42	0	17.90838	GP18
Cl-766 mapping	287	4557	5	0.11%	5	0	2.393762	Cl-766
139	0	4557	0	0.00%	0	0	0	139
LN_CHINA	0	4557	0	0.00%	0	0	0	LN_CHINA
PL-20	0	4557	0	0.00%	0	0	0	PL-20

*Table 9: Number of reads of **positive control** mapped to references at **0.95*0.9** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	5315	1491	28.05%	1443	48	758.1676	NY-1
623 mapping	513	5315	937	17.63%	925	12	464.4737	623
GH30 mapping	513	5315	859	16.16%	811	48	447.4639	GH30
3138-07 mapping	513	5315	660	12.42%	640	20	334.575	3138-07
621 mapping	513	5315	622	11.70%	614	8	316.0682	621
GH11 mapping	513	5315	459	8.64%	457	2	229.3762	GH11
WA-MR mapping	513	5315	218	4.10%	216	2	108.8187	WA-MR
GP18 mapping	513	5315	59	1.11%	59	0	24.59064	GP18
Cl-766 mapping	288	5315	10	0.19%	10	0	4.988304	Cl-766
139	0	5315	0	0.00%	0	0	0	139
LN_CHINA	0	5315	0	0.00%	0	0	0	LN_CHINA
PL-20	0	5315	0	0.00%	0	0	0	PL-20

*Table 10: Number of reads of **positive control** mapped to references at **0.94*0.95** (Length fraction*Similarity fraction) used for parameter optimization, with the positive control consisting of GLRaV-3 helicase gene amplicons of identified clones. The GLRaV-3 amplicons that are known to be present in the positive control sample are the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. Consensus length in column 2 refers to the length of the reference the read had matched to.*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	46461	13840	29.79%	12764	1076	7526.682	NY-1
GH30 mapping	513	46461	13495	29.05%	10415	3080	7389.571	GH30
GH11 mapping	513	46461	4767	10.26%	4593	174	2609.706	GH11
621 mapping	513	46461	4131	8.89%	3655	476	2239.191	621
3138-07 mapping	513	46461	4102	8.83%	3698	404	2221.142	3138-07
623 mapping	513	46461	3926	8.45%	3918	8	2101.883	623
WA-MR mapping	513	46461	1758	3.78%	1724	34	952.0253	WA-MR
GP18 mapping	513	46461	240	0.52%	238	2	121.3314	GP18
Cl-766 mapping	288	46461	202	0.43%	202	0	110.3255	Cl-766
139	0	46461	0	0.00%	0	0	0	139
LN_CHINA	0	46461	0	0.00%	0	0	0	LN_CHINA
PL-20	0	46461	0	0.00%	0	0	0	PL-20

*Table 11: Number of reads of the **positive control** mapped to references at **0.9*0.9** and GLRaV-3 **3138-07 removed** from the reference library, with the positive control consisting of identified amplicons obtained from clones. The GLRaV-3 amplicons known to occur in the positive control can be viewed as the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	170250	52959	31.11%	23069	29890	26140.46	NY-1
GH30 mapping	513	170250	42012	24.68%	2226	39786	21393.57	GH30
621 mapping	513	170250	37080	21.78%	10672	26408	18390.11	621
623 mapping	513	170250	17095	10.04%	16975	120	8446.86	623
GH11 mapping	513	170250	9994	5.87%	1302	8692	4852.033	GH11
WA-MR mapping	513	170250	6014	3.53%	4848	1166	2928.809	WA-MR
GP18 mapping	513	170250	3861	2.27%	3789	72	1610.029	GP18
Cl-766 mapping	512	170250	1234	0.72%	1234	0	581.7622	Cl-766
PL-20 mapping	257	170250	1	0.00%	1	0	0.500975	PL-20
139	0	170250	0	0.00%	0	0	0	139
LN_CHINA	0	170250	0	0.00%	0	0	0	LN_CHINA

*Table 12: Number of reads of the **positive control** mapped to references at **0.9*0.9** and GLRaV-3 3138-07 and **WA-MR removed** from the reference library, with the positive control consisting of identified amplicons obtained from clones. The GLRaV-3 amplicons known to occur in the positive control can be viewed as the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total Reads Mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	170224	53044	31.16%	23044	30000	26181.47	NY-1
621 mapping	513	170224	42896	25.20%	15398	27498	21225.16	621
GH30 mapping	513	170224	41934	24.63%	2218	39716	21359.42	GH30
623 mapping	513	170224	17114	10.05%	16994	120	8457.285	623
GH11 mapping	513	170224	10072	5.92%	1310	8762	4886.197	GH11
GP18 mapping	513	170224	3885	2.28%	3805	80	1622.25	GP18
Cl-766 mapping	512	170224	1278	0.75%	1278	0	601.1559	Cl-766
PL-20 mapping	257	170224	1	0.00%	1	0	0.500975	PL-20
139	0	170224	0	0.00%	0	0	0	139
LN_CHINA	0	170224	0	0.00%	0	0	0	LN_CHINA

*Table 13: Number of reads of the **positive control** mapped to references at **0.9*0.9** and GLRaV-3 3138-07, WA-MR, **LN, and GH11 removed** from the reference library, with the positive control consisting of identified amplicons obtained from clones. The GLRaV-3 amplicons known to occur in the positive control can be viewed as the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total reads mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	170248	53009	31.14%	22995	30014	26171.05	NY-1
GH30 mapping	513	170248	52030	30.56%	3496	48534	26256.98	GH30
621 mapping	513	170248	42906	25.20%	15408	27498	21230.36	621
623 mapping	513	170248	17127	10.06%	17007	120	8462.719	623
GP18 mapping	513	170248	3902	2.29%	3836	66	1624.22	GP18
Cl-766 mapping	513	170248	1273	0.75%	1273	0	598.9844	Cl-766
PL-20 mapping	257	170248	1	0.00%	1	0	0.500975	PL-20
139	0	170248	0	0.00%	0	0	0	139

*Table 14: Number of reads of the **positive control** mapped to references at **0.9*0.9** and GLRaV-3 3138-07, WA-MR, LN, and GH11 removed from the reference library and **GH24 added**, with the positive control consisting of identified amplicons obtained from clones. The GLRaV-3 amplicons known to occur in the positive control can be viewed as the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total reads mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
NY-1 mapping	513	153929	49140	31.92%	24730	24410	27029.97	NY-1
GH30 mapping	513	153929	48541	31.53%	6149	42392	26702.02	GH30
621 mapping	513	153929	38857	25.24%	15129	23728	21359.07	621
623 mapping	513	153929	14939	9.71%	14883	56	8213.236	623
GP18 mapping	513	153929	1478	0.96%	1472	6	811.3938	GP18
Cl-766 mapping	513	153929	970	0.63%	970	0	533.2904	Cl-766
PL-20 mapping	514	153929	2	0.00%	2	0	1.093567	PL-20
GH24 mapping	276	153929	2	0.00%	2	0	1.074219	GH24
139	0	153929	0	0.00%	0	0	0	139

*Table 15: Number of reads **positive control** mapped to references at 0.9*0.9 using **amended reference library** and the **"Ignore" function** instead of the default "map randomly" function. The positive control consists of known and identified amplicons obtained from clones, and can be seen in the shaded rows in the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8. This is considered the most optimal set of conditions for reference mapping.*

Name	Consensus length	Total reads mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
GH30 mapping	513	151516	48541	32.04%	6149	42392	26702.02	GH30
NY-1 mapping	513	151516	48078	31.73%	23674	24404	26443.09	NY-1
621 mapping	513	151516	38746	25.57%	15018	23728	21297.95	621
623 mapping	513	151516	14753	9.74%	14697	56	8106.982	623
Cl-766 mapping	302	151516	883	0.58%	883	0	485.3665	Cl-766
GP18 mapping	513	151516	511	0.34%	505	6	280.6803	GP18
PL-20 mapping	514	151516	2	0.00%	2	0	1.093567	PL-20
GH24 mapping	276	151516	2	0.00%	2	0	1.074219	GH24
139	0	151516	0	0.00%	0	0	0	139

*Table 16: Identity of contigs obtained during the de novo of unmapped reads discarded by **positive control** sequencing run replicate 1 reference mapping (Table 15). Default parameters of 0.7*0.8 and mapping back to contigs was used. Greatest hit length of MultiBLAST results determined identity of contig that can be found in the last column, corresponding information such as e-value and % identity can be found in column 6 and 7*

Query	Total reads matched	Reads matched	% Reads matched	Contig length	E-value	identity %	Greatest hit length	Accession (hit length)	Description (hit length)
Contig 1	7690	7682	99.90%	622	0	99.82	545	JQ655296	Grapevine leafroll-associated virus 3 isolate GH30,

*Table 17: The overall amount of reads of the **positive control** sequencing run replicate 1mapping during reference mapping and de novo, with the amplicons in positive control known. The GLRaV-3 helicase amplicons known to be present in the positive control can be seen in the shaded rows of the table. This represents the overall population present in the amplicon. The reference mapping and de novo assembly results are used to generate overall profile. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	151516	48541	32.04%	7690	7682	99.90%	159206	56223	35.31%
NY-1	151516	48078	31.73%	7690	0	0.00%	159206	48078	30.20%
621	151516	38746	25.57%	7690	0	0.00%	159206	38746	24.34%
623	151516	14753	9.74%	7690	0	0.00%	159206	14753	9.27%
Cl-766	151516	883	0.58%	7690	0	0.00%	159206	883	0.55%
GP18	151516	511	0.34%	7690	0	0.00%	159206	511	0.32%
PL-20	151516	2	0.00%	7690	0	0.00%	159206	2	0.00%
GH24	151516	2	0.00%	7690	0	0.00%	159206	2	0.00%
139	151516	0	0.00%	7690	0	0.00%	159206	0	0.00%

*Table 18: Number of reads of **positive control** Illumina MiSeq sequencing run replicate 2 mapped to references at 0.9*0.9 and other optimal conditions. The same positive control was Illumina MiSeq sequenced twice, thus a duplicate, to control for variance between sequencing runs, and to acquire more reads. The GLRaV-3 helicase amplicons known to occur in the positive control can be seen in the shaded rows of the table. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total reads mapped	Total read count	Percentage (%) Reads Mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
GH30 mapping	513	29802	9620	32.28%	3208	6412	5268.719	GH30
NY-1 mapping	513	29802	9613	32.26%	6131	3482	5275.53	NY-1
621 mapping	513	29802	7517	25.22%	4223	3294	4122.4	621
623 mapping	513	29802	2809	9.43%	2807	2	1539.75	623
CI-766 mapping	294	29802	166	0.56%	166	0	90.98246	CI-766
GP18 mapping	513	29802	77	0.26%	77	0	42.15205	GP18
139	0	29802	0	0.00%	0	0	0	139
PL-20	0	29802	0	0.00%	0	0	0	PL-20
GH24	0	29802	0	0.00%	0	0	0	GH24

*Table 19: Identity of contigs obtained during the de novo of unmapped reads discarded by **positive control** sequencing replicate 2 run reference mapping (Table 18). The same positive control was Illumina MiSeq sequenced twice, thus a duplicate, to control for variance between sequencing runs, and to acquire more reads. Default parameters of 0.7*0.8 and mapping back to contigs was used. Greatest hit length of MultiBLAST results determined identity of contig that can be found in the last column, corresponding information such as e-value and % identity can be found in column 6 and 7*

Query	Total reads matched	Reads matched	% Reads matched	Contig length	E-value	identity %	Greatest hit length	Accession (hit length)	Description (hit length)
Contig 1	2176	2157	99.13%	307	5.8E-156	100	307	JQ655296	Grapevine leafroll-associated virus 3 isolate GH30

*Table 20: Overall amount of reads mapping of the **positive control** Illumina MiSeq sequencing run replicate 2. The same positive control was Illumina MiSeq sequenced twice, thus being a duplicate, to control for variance between sequencing runs, and to acquire more reads. The GLRaV-3 helicase amplicons known to be present in the positive control can be seen in the shaded rows of the table. This represents the overall population present in the amplicon. This represents the overall population present in the amplicon. The reference mapping and de novo assembly results are used to generate overall profile. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	29802	9620	32.28%	2176	2157	99.13%	31978	11777	36.83%
NY-1	29802	9613	32.26%	2176	0	0.00%	31978	9613	30.06%
621	29802	7517	25.22%	2176	0	0.00%	31978	7517	23.51%
623	29802	2809	9.43%	2176	0	0.00%	31978	2809	8.78%
Cl-766	29802	166	0.56%	2176	0	0.00%	31978	166	0.52%
GP18	29802	77	0.26%	2176	0	0.00%	31978	77	0.24%
139	29802	0	0.00%	2176	0	0.00%	31978	0	0.00%
PL-20	29802	0	0.00%	2176	0	0.00%	31978	0	0.00%
GH24	29802	0	0.00%	2176	0	0.00%	31978	0	0.00%

*Table 21: Number reads mapped and matched of both the **positive control replicate 1**and **positive control replicate 2**Illumina MiSeq run and their total. This increases the amount of data for the positive control giving a more accurate representation of the sample*

GLRaV-3 variant	Positive control sequencing run replicate 1 overall			Positive control sequencing run replicate 2 overall			Positive control replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GH30	159206	56223	35.31%	31978	11777	36.83%	191184	68000	35.57%
NY-1	159206	48078	30.20%	31978	9613	30.06%	191184	57691	30.18%
621	159206	38746	24.34%	31978	7517	23.51%	191184	46263	24.20%
623	159206	14753	9.27%	31978	2809	8.78%	191184	17562	9.19%
Cl-766	159206	883	0.55%	31978	166	0.52%	191184	1049	0.55%
GP18	159206	511	0.32%	31978	77	0.24%	191184	588	0.31%
139	159206	2	0.00%	31978	0	0.00%	191184	2	0.00%
PL-20	159206	2	0.00%	31978	0	0.00%	191184	2	0.00%
GH24	159206	0	0.00%	31978	0	0.00%	191184	0	0.00%

*Table 22: Number reads mapped and matched of both **14-9001 sequencing run replicate 1** and **14-9001 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9001, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9001 sequencing run replicate 1 overall			14-9001 sequencing run replicate 2 overall			14-9001 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	359651	356456	99.11%	44294	43918	99.15%	403945	400374	99.12%
GP18	359651	2425	0.67%	44294	291	0.66%	403945	2716	0.67%
623	359651	531	0.15%	44294	73	0.16%	403945	604	0.15%
621	359651	66	0.02%	44294	5	0.01%	403945	71	0.02%
GH30	359651	33	0.01%	44294	5	0.01%	403945	38	0.01%
GH24	359651	14	0.00%	44294	2	0.00%	403945	16	0.00%
139	359651	2	0.00%	44294	0	0.00%	403945	2	0.00%
CI-766	359651	0	0.00%	44294	0	0.00%	403945	0	0.00%
PL-20	359651	0	0.00%	44294	0	0.00%	403945	0	0.00%

*Table 23: Number reads mapped and matched of both **14-9002 sequencing run replicate 1** and **14-9002 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9002, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9002 sequencing run replicate 1 overall			14-9002 sequencing run replicate 2 overall			14-9002 replicate 1 + replicate2 overall		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	865263	851973	98.46%	84155	82923	98.54%	949418	934896	98.47%
GP18	865263	5948	0.69%	84155	388	0.46%	949418	6336	0.67%
623	865263	1336	0.15%	84155	100	0.12%	949418	1436	0.15%
GH30	865263	710	0.08%	84155	44	0.05%	949418	754	0.08%
GH24	865263	104	0.01%	84155	5	0.01%	949418	109	0.01%
621	865263	52	0.01%	84155	2	0.00%	949418	54	0.01%
Cl-766	865263	18	0.00%	84155	1	0.00%	949418	19	0.00%
139	865263	2	0.00%	84155	0	0.00%	949418	2	0.00%
PL-20	865263	0	0.00%	84155	0	0.00%	949418	0	0.00%

*Table 24: Number reads mapped and matched of both **14-9019 sequencing run replicate 1** and **14-9019 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9019, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9019 sequencing run replicate 1 overall			14-9019 sequencing run replicate 2 overall			14-9019 replicate 1 + replicate 2 run		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	296994	225192	75.82%	114872	61863	53.85%	411866	287055	69.70%
GH30	296994	23458	7.90%	114872	10487	9.13%	411866	33945	8.24%
GP18	296994	1799	0.61%	114872	470	0.41%	411866	2269	0.55%
623	296994	414	0.14%	114872	148	0.13%	411866	562	0.14%
GH24	296994	413	0.14%	114872	133	0.12%	411866	546	0.13%
621	296994	118	0.04%	114872	40	0.03%	411866	158	0.04%
PL-20	296994	53	0.02%	114872	18	0.02%	411866	71	0.02%
Cl-766	296994	5	0.00%	114872	2	0.00%	411866	7	0.00%
139	296994	0	0.00%	114872	0	0.00%	411866	0	0.00%

*Table 25: Number reads mapped and matched of both **14-9020 sequencing run replicate 1** and **14-9020 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9020, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9020 sequencing run replicate 1 overall			14-9020 sequencing run replicate 2 overall			14-9020 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	396984	165654	41.73%	129290	58931	45.58%	526274	224585	42.67%
GH30	396984	158872	40.02%	129290	53855	41.65%	526274	212727	40.42%
GH24	396984	70665	17.80%	129290	16085	12.44%	526274	86750	16.48%
GP18	396984	1297	0.33%	129290	262	0.20%	526274	1559	0.30%
623	396984	244	0.06%	129290	70	0.05%	526274	314	0.06%
PL-20	396984	34	0.01%	129290	18	0.01%	526274	52	0.01%
621	396984	2	0.00%	129290	0	0.00%	526274	2	0.00%
Cl-766	396984	4	0.00%	129290	0	0.00%	526274	4	0.00%
139	396984	0	0.00%	129290	0	0.00%	526274	0	0.00%

*Table 26: Number reads mapped and matched of both **14-9021 sequencing run replicate 1** and **14-9021 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9021, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9021 sequencing run replicate 1 overall			14-9021 sequencing run replicate 2 overall			14-9021 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	369089	214120	58.01%	66664	40288	60.43%	435753	254408	58.38%
GH30	369089	112487	30.48%	66664	20694	31.04%	435753	133181	30.56%
GH24	369089	33432	9.06%	66664	4227	6.34%	435753	37659	8.64%
621	369089	4403	1.19%	66664	895	1.34%	435753	5298	1.22%
GP18	369089	2053	0.56%	66664	310	0.47%	435753	2363	0.54%
Cl-766	369089	858	0.23%	66664	159	0.24%	435753	1017	0.23%
623	369089	504	0.14%	66664	91	0.14%	435753	595	0.14%
139	369089	0	0.00%	66664	0	0.00%	435753	0	0.00%
PL-20	369089	0	0.00%	66664	0	0.00%	435753	0	0.00%

*Table 27: Number reads mapped and matched of both **14-9022 sequencing run replicate 1** and **14-9022 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9022, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9022 sequencing run replicate 1 overall			14-9022 sequencing run replicate 2 overall			14-9022 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	368919	137319	37.22%	71529	30062	42.03%	440448	167381	38.00%
NY-1	368919	129431	35.08%	71529	17197	24.04%	440448	146628	33.29%
GH24	368919	53110	14.40%	71529	9505	13.29%	440448	62615	14.22%
621	368919	27440	7.44%	71529	4803	6.71%	440448	32243	7.32%
GP18	368919	1010	0.27%	71529	122	0.17%	440448	1132	0.26%
CI-766	368919	824	0.22%	71529	99	0.14%	440448	923	0.21%
623	368919	196	0.05%	71529	33	0.05%	440448	229	0.05%
PL-20	368919	4	0.00%	71529	0	0.00%	440448	4	0.00%
139	368919	0	0.00%	71529	0	0.00%	440448	0	0.00%

*Table 28: Number reads mapped and matched of both **14-9031 sequencing run replicate 1** and **14-9031 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9031, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9031 sequencing run replicate 1 overall			14-9031 sequencing run replicate 2 overall			14-9031 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	202710	142019	70.06%	158085	107446	67.97%	360795	249465	69.14%
GH24	202710	20289	10.01%	158085	10688	6.76%	360795	30977	8.59%
621	202710	11898	5.87%	158085	7280	4.61%	360795	19178	5.32%
GH30	202710	11202	5.53%	158085	5426	3.43%	360795	16628	4.61%
GP18	202710	989	0.49%	158085	501	0.32%	360795	1490	0.41%
Cl-766	202710	350	0.17%	158085	200	0.13%	360795	550	0.15%
623	202710	229	0.11%	158085	146	0.09%	360795	375	0.10%
PL-20	202710	42	0.02%	158085	9	0.01%	360795	51	0.01%
139	202710	0	0.00%	158085	0	0.00%	360795	0	0.00%

*Table 29: Number reads mapped and matched of both **14-9032 sequencing run replicate 1** and **14-9032 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9032, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9032 sequencing run replicate 1 overall			14-9032 sequencing run replicate 2 overall			14-9032 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	273820	124500	45.47%	81840	34558	42.23%	355660	159058	44.72%
GH30	273820	84538	30.87%	81840	25258	30.86%	355660	109796	30.87%
GH24	273820	58543	21.38%	81840	20768	25.38%	355660	79311	22.30%
PL-20	273820	4756	1.74%	81840	1051	1.28%	355660	5807	1.63%
GP18	273820	999	0.36%	81840	193	0.24%	355660	1192	0.34%
623	273820	200	0.07%	81840	28	0.03%	355660	228	0.06%
621	273820	78	0.03%	81840	14	0.02%	355660	92	0.03%
Cl-766	273820	4	0.00%	81840	0	0.00%	355660	4	0.00%
139	273820	0	0.00%	81840	0	0.00%	355660	0	0.00%

*Table 30: Number reads mapped and matched of both **14-9073 sequencing run replicate 1** and **14-9073 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9073, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9073 sequencing run replicate 1 overall			14-9073 sequencing run replicate 2 overall			14-9073 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	296868	277878	93.60%	53311	51310	96.25%	350179	329188	94.01%
GH30	296868	8571	2.89%	53311	1481	2.78%	350179	10052	2.87%
GP18	296868	2259	0.76%	53311	409	0.77%	350179	2668	0.76%
623	296868	630	0.21%	53311	107	0.20%	350179	737	0.21%
GH24	296868	12	0.00%	53311	3	0.01%	350179	15	0.00%
621	296868	4	0.00%	53311	1	0.00%	350179	5	0.00%
CI-766	296868	2	0.00%	53311	0	0.00%	350179	2	0.00%
139	296868	0	0.00%	53311	0	0.00%	350179	0	0.00%
PL-20	296868	0	0.00%	53311	0	0.00%	350179	0	0.00%

*Table 31: Number reads mapped and matched of both **14-9074 sequencing run replicate 1** and **14-9074 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9074, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9074 sequencing run replicate 1 overall			14-9074 sequencing run replicate 2 overall			14-9074 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	251290	242526	96.51%	138351	136582	98.72%	389641	379108	97.30%
621	251290	5570	2.22%	138351	1147	0.83%	389641	6717	1.72%
GP18	251290	2353	0.94%	138351	405	0.29%	389641	2758	0.71%
623	251290	447	0.18%	138351	96	0.07%	389641	543	0.14%
Cl-766	251290	231	0.09%	138351	51	0.04%	389641	282	0.07%
GH30	251290	129	0.05%	138351	22	0.02%	389641	151	0.04%
GH24	251290	34	0.01%	138351	2	0.00%	389641	36	0.01%
139	251290	0	0.00%	138351	0	0.00%	389641	0	0.00%
PL-20	251290	0	0.00%	138351	0	0.00%	389641	0	0.00%

*Table 32: Overall amount reads of accession number **15-5007** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	223856	221819	99.09%	226885	208321	91.82%	450741	430140	95.43%
NY-1	223856	1807	0.81%	226885	18309	8.07%	450741	20116	4.46%
PL-20	223856	196	0.09%	226885	0	0.00%	450741	196	0.04%
621	223856	18	0.01%	226885	0	0.00%	450741	18	0.00%
GP18	223856	6	0.00%	226885	0	0.00%	450741	6	0.00%
623	223856	5	0.00%	226885	0	0.00%	450741	5	0.00%
GH24	223856	5	0.00%	226885	0	0.00%	450741	5	0.00%
139	223856	0	0.00%	226885	0	0.00%	450741	0	0.00%
Cl-766	223856	0	0.00%	226885	0	0.00%	450741	0	0.00%

*Table 33: Overall amount reads of accession number **15-5008** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	396 724	261806	65.99%	97532	75807	77.73%	494256	337613	68.31%
621	396 724	128605	32.42%	97532	0	0.00%	494256	128605	26.02%
3138-07	396 724	0	0.00%	97532	21423	21.97%	494256	21423	4.33%
NY-1	396 724	5999	1.51%	97532	0	0.00%	494256	5999	1.21%
GH24	396 724	213	0.05%	97532	0	0.00%	494256	213	0.04%
Cl-766	396 724	73	0.02%	97532	0	0.00%	494256	73	0.01%
GP18	396 724	19	0.00%	97532	0	0.00%	494256	19	0.00%
623	396 724	8	0.00%	97532	0	0.00%	494256	8	0.00%
PL-20	396 724	1	0.00%	97532	0	0.00%	494256	1	0.00%
139	396 724	0	0.00%	97532	0	0.00%	494256	0	0.00%

*Table 34: Overall amount reads of accession number **15-5009** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	163522	149722	91.56%	83737	83563	99.79%	247259	233285	94.35%
NY-1	163522	13744	8.40%	83737	0	0.00%	247259	13744	5.56%
GP18	163522	33	0.02%	83737	0	0.00%	247259	33	0.01%
623	163522	17	0.01%	83737	0	0.00%	247259	17	0.01%
621	163522	6	0.00%	83737	0	0.00%	247259	6	0.00%
139	163522	0	0.00%	83737	0	0.00%	247259	0	0.00%
Cl-766	163522	0	0.00%	83737	0	0.00%	247259	0	0.00%
GH24	163522	0	0.00%	83737	0	0.00%	247259	0	0.00%
PL-20	163522	0	0.00%	83737	0	0.00%	247259	0	0.00%

*Table 35: Overall amount reads of accession number **15-5010** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	115655	68391	59.13%	12727	12667	99.53%	128382	81058	63.14%
NY-1	115655	46973	40.61%	12727	0	0.00%	128382	46973	36.59%
GP18	115655	185	0.16%	12727	0	0.00%	128382	185	0.14%
623	115655	88	0.08%	12727	0	0.00%	128382	88	0.07%
621	115655	17	0.01%	12727	0	0.00%	128382	17	0.01%
Cl-766	115655	1	0.00%	12727	0	0.00%	128382	1	0.00%
139	115655	0	0.00%	12727	0	0.00%	128382	0	0.00%
GH24	115655	0	0.00%	12727	0	0.00%	128382	0	0.00%
PL-20	115655	0	0.00%	12727	0	0.00%	128382	0	0.00%

*Table 36: Overall amount reads of accession number **15-5015** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
621	36908	24494	66.37%	181143	0	0.00%	218051	24494	11.23%
GH30	36908	11608	31.45%	181143	0	0.00%	218051	11608	5.32%
3138-07	36908	0	0.00%	181143	21327	11.77%	218051	21327	9.78%
NY-1	36908	784	2.12%	181143	0	0.00%	218051	784	0.36%
Cl-766	36908	17	0.05%	181143	0	0.00%	218051	17	0.01%
623	36908	4	0.01%	181143	0	0.00%	218051	4	0.00%
GP18	36908	1	0.00%	181143	0	0.00%	218051	1	0.00%
139	36908	0	0.00%	181143	0	0.00%	218051	0	0.00%
GH24	36908	0	0.00%	181143	0	0.00%	218051	0	0.00%
PL-20	36908	0	0.00%	181143	0	0.00%	218051	0	0.00%

*Table 37: Number reads of accession number **15-5016** that mapped to references using 0.9*0.9 and other optimal conditions. No replicate sequencing runs were performed with this sample. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. The reads above the line represents what is considered to be present in the sample. No reads yielded any results during de novo assembly (Table 38), thus the reference mapping represents the overall profile. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total reads mapped	Total read count	% reads mapped	Single reads	Reads in pairs	Average coverage	Reference sequence	Reference length
621 mapping	513	824	810	98.30%	232	578	411.8265	621	513
GH30 mapping	513	824	12	1.46%	2	10	6.413255	GH30	513
Cl-766 mapping	282	824	1	0.12%	1	0	0.549708	Cl-766	513
NY-1 mapping	269	824	1	0.12%	1	0	0.524366	NY-1	513
139	0	824	0	0.00%	0	0	0	139	513
623	0	824	0	0.00%	0	0	0	623	513
GH24	0	824	0	0.00%	0	0	0	GH24	512
GP18	0	824	0	0.00%	0	0	0	GP18	513
PL-20	0	824	0	0.00%	0	0	0	PL-20	513

*Table 36: Identity of contigs obtained during the de novo of unmapped reads discarded by accession number **15-5016** reference mapping (Table 37). Default parameters of 0.7*0.8 and mapping back to contigs was used. Greatest hit length of MultiBLAST results determined identity of contig that can be found in the last column, corresponding information such as e-value and % identity can be found in column 6 and 7*

Query	Total reads matched	Reads matched	% Reads matched	E-value	Identity %	Greatest hit length	Accession (hit length)	Description (hit length)
No Results								

*Table 37: Overall amount reads of accession number **15-5043** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	17157	7619	44.41%	2707	2673	98.74%	19864	10292	51.81%
NY-1	17157	5010	29.20%	2707	0	0.00%	19864	5010	25.22%
621	17157	4448	25.93%	2707	0	0.00%	19864	4448	22.39%
Cl-766	17157	35	0.20%	2707	0	0.00%	19864	35	0.18%
GP18	17157	34	0.20%	2707	0	0.00%	19864	34	0.17%
623	17157	11	0.06%	2707	0	0.00%	19864	11	0.06%
139	17157	0	0.00%	2707	0	0.00%	19864	0	0.00%
GH24	17157	0	0.00%	2707	0	0.00%	19864	0	0.00%
PL-20	17157	0	0.00%	2707	0	0.00%	19864	0	0.00%

*Table 38: Overall amount reads of accession number **15-5044** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	30118	11160	37.05%	2752	0	0.00%	32870	11160	33.95%
NY-1	30118	9847	32.69%	2752	1121	40.73%	32870	10968	33.37%
621	30118	8981	29.82%	2752	0	0.00%	32870	8981	27.32%
3138-07	30118	0	0.00%	2752	1146	41.64%	32870	1146	3.49%
Cl-766	30118	90	0.30%	2752	0	0.00%	32870	90	0.27%
GP18	30118	25	0.08%	2752	0	0.00%	32870	25	0.08%
623	30118	15	0.05%	2752	0	0.00%	32870	15	0.05%
139	30118	0	0.00%	2752	0	0.00%	32870	0	0.00%
GH24	30118	0	0.00%	2752	0	0.00%	32870	0	0.00%
PL-20	30118	0	0.00%	2752	0	0.00%	32870	0	0.00%

*Table 39: Number reads of accession number **15-5067** that mapped to references using 0.9*0.9 and other optimal conditions. No replicate sequencing runs were performed with this sample. Percentage of reads mapped is listed in column 5, indicating relative abundance of specific strains in the population. The reads above the line represents what is considered to be present in the sample. No reads yielded any results during de novo assembly (Table 40), thus the reference mapping represents the overall profile. Additional information relating to the mapping can be found in column 6, 7, and 8*

Name	Consensus length	Total reads mapped	Total read count	% Reads mapped	Single reads	Reads in pairs	Average coverage	Reference sequence	Reference length
621 mapping	513	122396	120621	98.55%	15895	104726	61865.68	621	513
NY-1 mapping	513	122396	1505	1.23%	457	1048	774.0604	NY-1	513
GH30 mapping	513	122396	169	0.14%	57	112	88.09942	GH30	513
Cl-766 mapping	340	122396	63	0.05%	63	0	31.13645	Cl-766	513
GP18 mapping	513	122396	36	0.03%	28	8	18.14425	GP18	513
623 mapping	512	122396	2	0.00%	2	0	1.019493	623	513
139	0	122396	0	0.00%	0	0	0	139	513
GH24	0	122396	0	0.00%	0	0	0	GH24	512
PL-20	0	122396	0	0.00%	0	0	0	PL-20	513

*Table 40: Identity of contigs obtained during the de novo of unmapped reads discarded by accession number **15-5067** reference mapping (Table 37). Default parameters of 0.7*0.8 and mapping back to contigs was used. Greatest hit length of MultiBLAST results determined identity of contig that can be found in the last column, corresponding information such as e-value and % identity can be found in column 6 and 7*

Query	Total reads matched	Reads matched	% Reads matched	E-value	Identity %	Greatest hit length	Accession (hit length)	Description (hit length)

No Results

*Table 41: Overall amount reads of accession number **15-5068** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	99569	51666	51.89%	59022	16513	27.98%	158591	68179	42.99%
621	99569	47421	47.63%	59022	0	0.00%	158591	47421	29.90%
3138-07	99569	0	0.00%	59022	41811	70.84%	158591	41811	26.36%
Cl-766	99569	172	0.17%	59022	0	0.00%	158591	172	0.11%
GP18	99569	156	0.16%	59022	0	0.00%	158591	156	0.10%
623	99569	120	0.12%	59022	0	0.00%	158591	120	0.08%
GH30	99569	34	0.03%	59022	0	0.00%	158591	34	0.02%
139	99569	0	0.00%	59022	0	0.00%	158591	0	0.00%
GH24	99569	0	0.00%	59022	0	0.00%	158591	0	0.00%
PL-20	99569	0	0.00%	59022	0	0.00%	158591	0	0.00%

*Table 42: Overall amount reads of accession number **15-5077** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
NY-1	128576	51321	39.91%	203	0	0.00%	128779	51321	39.85%
GH30	128576	42808	33.29%	203	0	0.00%	128779	42808	33.24%
621	128576	34064	26.49%	203	0	0.00%	128779	34064	26.45%
Cl-766	128576	181	0.14%	203	0	0.00%	128779	181	0.14%
GP18	128576	137	0.11%	203	0	0.00%	128779	137	0.11%
623	128576	65	0.05%	203	0	0.00%	128779	65	0.05%
GH11	128576	0	0.00%	203	20	9.85%	128779	20	0.02%
139	128576	0	0.00%	203	0	0.00%	128779	0	0.00%
GH24	128576	0	0.00%	203	0	0.00%	128779	0	0.00%
PL-20	128576	0	0.00%	203	0	0.00%	128779	0	0.00%

*Table 43: Overall amount reads of accession number **15-5078** mapping. No replicate sequencing runs were performed with this sample. The reference mapping and de novo assembly results are used to generate overall profile, thus representing the overall population present in the amplicon. The reads above the line represents what is considered to be present in the sample, and the overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GH30	59114	20899	35.35%	3879	0	0.00%	62993	20899	33.18%
NY-1	59114	19211	32.50%	3879	0	0.00%	62993	19211	30.50%
621	59114	18696	31.63%	3879	3631	93.61%	62993	22327	35.44%
Cl-766	59114	205	0.35%	3879	0	0.00%	62993	205	0.33%
GP18	59114	81	0.14%	3879	0	0.00%	62993	81	0.13%
623	59114	22	0.04%	3879	0	0.00%	62993	22	0.03%
139	59114	0	0.00%	3879	0	0.00%	62993	0	0.00%
GH24	59114	0	0.00%	3879	0	0.00%	62993	0	0.00%
PL-20	59114	0	0.00%	3879	0	0.00%	62993	0	0.00%

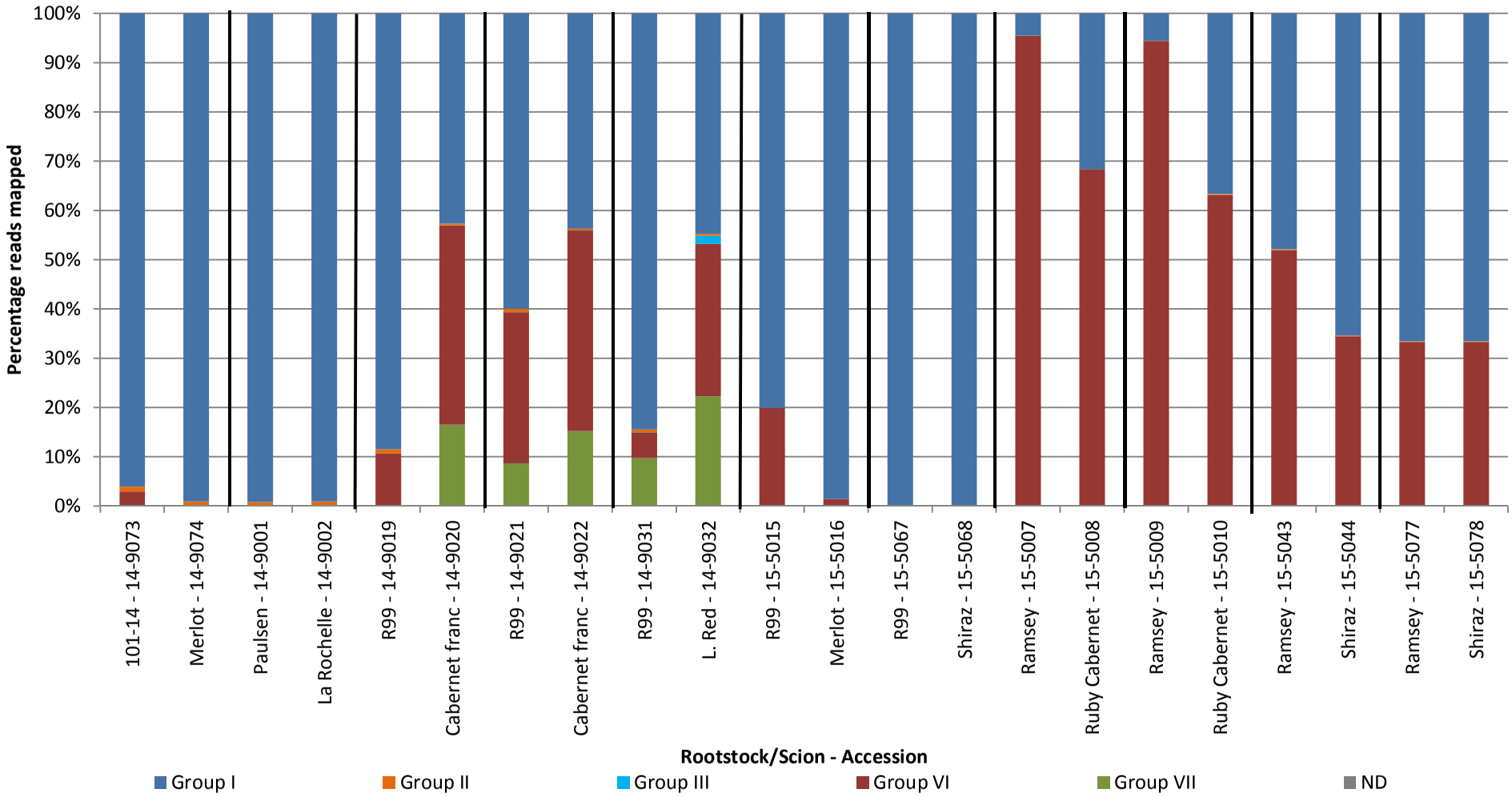


Figure 3: Graphical illustration of read composition in rootstock and scion tissues of individual samples

APPENDIX B

Table 1: Viti- and Foveavirus status of rootstock and scion tissue combinations of individual grapevines with PCR band strength observed

Accession # (Rootstock/Scion)	Rootstock-scion combination		Scion cultivar colour	Rootstocks *Viti-* and *Foveavirus* status	Scions *Viti-* and *Foveavirus* status	Rootstock selected for NGS
	Rootstock	Scion				
14-9071/9072	**101-14**	**Cabernet Sauvignon**	Red	-	+	-
15-9041/9042	**101-14**	**Cabernet Sauvignon**	Red	-	-	-
14-9073/9074	**101-14**	**Merlot**	Red	-	-	-
15-5021/5022	**101-14**	**Merlot**	Red	-	+	-
15-5023/5024	**101-14**	**Merlot**	Red	+	-	-
15-5025/5026	**101-14**	**Merlot**	Red	+++	+++	-
15-5027/5028	**101-14**	**Merlot**	Red	-	+++	-
15-5029/5030	**101-14**	**Merlot**	Red	++	+++	*
14-9001/9002	**Paulsen**	**La Rochelle**	Red	-	-	-
14/9003/9004	**Paulsen**	**La Rochelle**	Red	+	+	-
14-9005/9006	**R 99**	**Assyrtiko**	White	+	+	*
14-9007/0908	**R 99**	**Catarratto commune**	White	+	+	*
14-9009/9010	**R 99**	**Catarratto commune**	White	-	+	-
14-9019/9020	**R 99**	**Cabernet franc**	Red	-	+	-
14-9021/9022	**R 99**	**Cabernet franc**	Red	+	+	*
14-9057/9058	**R 99**	**Cabernet franc**	Red	+	+	*
14-9059/9060	**R 99**	**Cabernet franc**	Red	+	+	-
16-0045/0046	**R 99**	**Cabernet Sauvignon**	Red	+++	+++	-
16-0047/0048	**R 99**	**Cabernet Sauvignon**	Red	+++	+++	-
16-0049/0050	**R 99**	**Cabernet Sauvignon**	Red	+++	+++	-
14-9049/9050	**R 99**	**CG 40318**	White	+	+	-
14-9051/9052	**R 99**	**CG 40318**	White	+	+	-
14-9011/9012	**R 99**	**Chardonnay**	White	-	-	-
14-9013/9014	**R 99**	**Chardonnay**	White	+	+	*
15-5031/5032	**R 99**	**Chardonnay**	White	++	+++	*
15-5033/5034	**R 99**	**Chardonnay**	White	-	+++	-
15-5035/5036	**R 99**	**Chardonnay**	White	++	+++	-
15-5037/5038	**R 99**	**Chardonnay**	White	+++	+	-
15-5039/5040	**R 99**	**Chardonnay**	White	++	++	-
15-5083/5084	**R 99**	**Chardonnay**	White	+	+	-
16-0039/0040	**R 99**	**Cinsault**	Red	++	+++	-
16-0041/0042	**R 99**	**Cinsault**	Red	+++	+++	-
16-0043/0044	**R 99**	**Cinsault**	Red	+++	+++	-
14-9045/9046	**R 99**	**Gamay hatif des vosges**	Red	-	-	-
14-9047/9048	**R 99**	**Gamay hatif des vosges**	Red	+	+	-
14-9037/9038	**R 99**	**Lakemont seedless**	White	+	+	*
14-9039/9040	**R 99**	**Lakemont seedless**	White	+	+	*
14-9029/9030	**R 99**	**L. Red**	Red	+	+	-
14-9031/9032	**R 99**	**L. Red**	Red	-	+	-
14-9033/9034	**R 99**	**Lumbrusco**	Red	-	+	-
14-9035/9036	**R 99**	**Lumbrusco**	Red	+	+	*
14-9015/9016	**R 99**	**Malbec**	Red	-	+	-
14-9017/9018	**R 99**	**Malbec**	Red	-	+	-
16-0017/0018	**R 99**	**Malbec**	Red	+	+++	-
16-0019/0020	**R 99**	**Malbec**	Red	++	+++	-
16-0021/0022	**R 99**	**Malbec**	Red	-	+++	-
15-5011/5012	**R 99**	**Merlot**	Red	+++	+	-
15-5013/5014	**R 99**	**Merlot**	Red	+++	-	-
15-5015/5016	**R 99**	**Merlot**	Red	-	+++	-
15-5017/5018	**R 99**	**Merlot**	Red	-	++	-
15-5019/5020	**R 99**	**Merlot**	Red	++	+++	-
15-5045/5046	**R 99**	**Merlot**	Red	+++	++	*
15-5047/5048	**R 99**	**Merlot**	Red	-	+	-
15-5049/5050	**R 99**	**Merlot**	Red	-	+	-
15-5051/5052	**R 99**	**Merlot**	Red	+	++	-

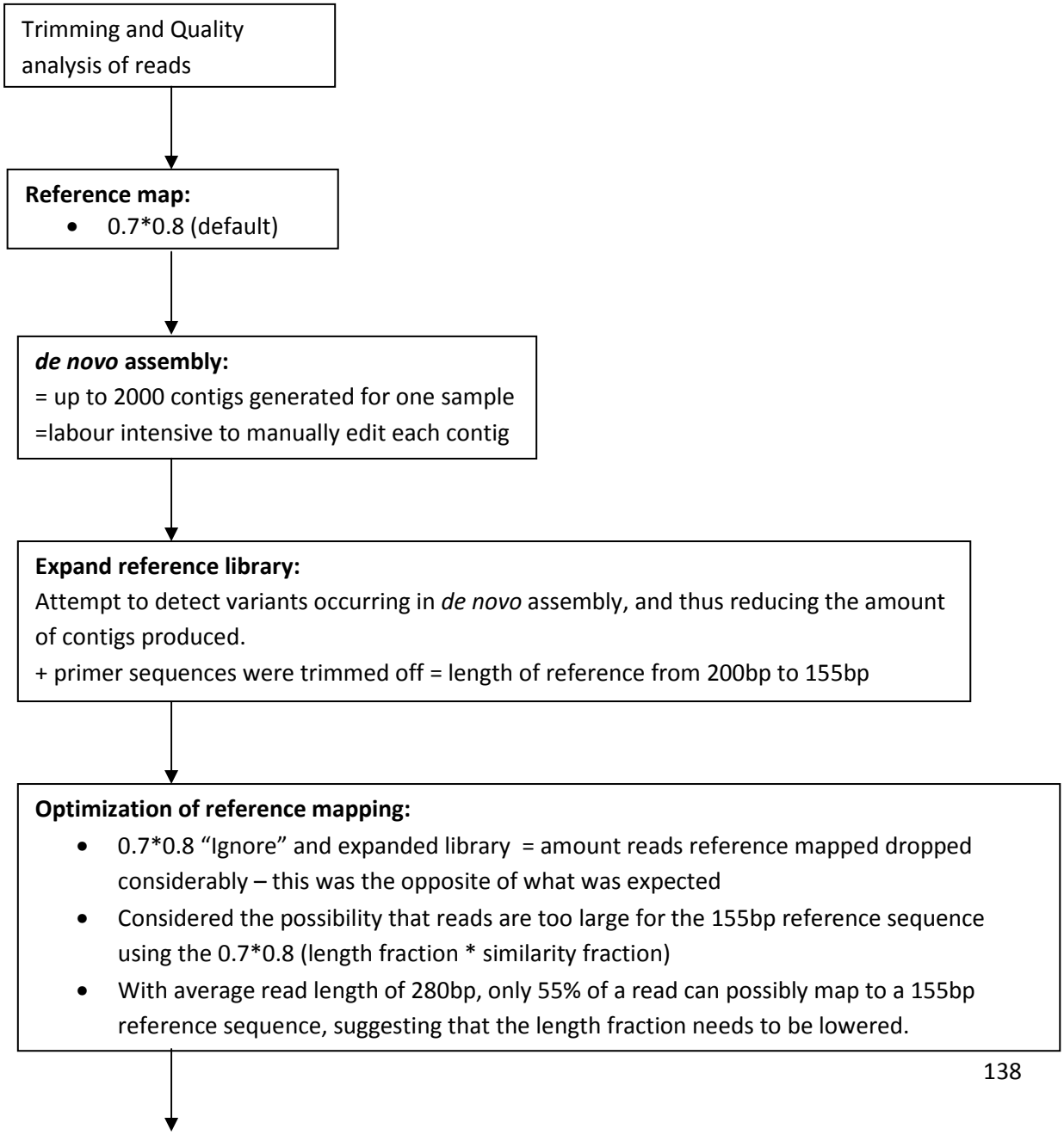
Accession # (Rootstock/Scion)	Scion-rootstock combination		Scion cultivar colour	Rootstocks GLRaV-3 status	Scions GLRaV-3 status	Selected for NGS
	Rootstock	Scion				
15-5053/5054	**R 99**	**Merlot**	Red	+++	-	-
15-5055/5056	**R 99**	**Merlot**	Red	-	+++	-
15-5057/5058	**R 99**	**Merlot**	Red	-	+	-
15-5059/5060	**R 99**	**Merlot**	Red	-	+	-
15-5061/5062	**R 99**	**Merlot**	Red	-	-	-
15-5063/5064	**R 99**	**Merlot**	Red	+	+	-
15-5085/5086	**R 99**	**Merlot**	Red	+	++	-
14-9041/9042	**R 99**	**Planta nova**	Red	+	+	*
14-9043/9044	**R 99**	**Planta nova**	Red	+	-	-
14-9077/9078	**R 99**	**Pinotage**	Red	-	+	-
14-9079/9080	**R 99**	**Pinotage**	Red	-	+	-
14-9081/9082	**R 99**	**Pinotage**	Red	+	+	*
14-9083/9084	**R 99**	**Pinotage**	Red	+	+	*
16-0023/0024	**R 99**	**Pinotage**	Red	-	+	-
16-0025/0026	**R 99**	**Pinotage**	Red	-	+++	-
16-0027/0028	**R 99**	**Pinotage**	Red	-	+	-
16-0033/0034	**R 99**	**Pontac**	Red	-	+++	-
16-0035/0036	**R 99**	**Pontac**	Red	-	+++	-
16-0037/0038	**R 99**	**Pontac**	Red	-	+++	-
16-0011/0012	**R 99**	**Ruby Cabernet**	Red	-	+++	-
16-0013/0014	**R 99**	**Ruby Cabernet**	Red	+++	+++	-
16-0015/0016	**R 99**	**Ruby Cabernet**	Red	-	+++	-
15-5065/5066	**R 99**	**Shiraz**	Red	+	+	-
15-5067/5068	**R 99**	**Shiraz**	Red	+	+	-
15-5069/5070	**R 99**	**Shiraz**	Red	+	+	*
15-5073/5074	**R 99**	**Shiraz**	Red	+	++	-
15-5075/5076	**R 99**	**Shiraz**	Red	-	++	-
16-0029/0030	**R 99**	**Tinta barrocca**	Red	+++	+++	-
16-0031/0032	**R 99**	**Tinta barrocca**	Red	++	+++	-
14-9053/9054	**R 99**	**Zeni**	White	+	+	*
14-9055/9056	**R 99**	**Zeni**	White	+	+	*
15-5001/5002	**Ramsey**	**Ruby Cabernet**	Red	+++	++	-
15-5003/5004	**Ramsey**	**Ruby Cabernet**	Red	+	+++	-
15-5005/5006	**Ramsey**	**Ruby Cabernet**	Red	+	-	-
15-5007/5008	**Ramsey**	**Ruby Cabernet**	Red	+	-	-
15-5009/5010	**Ramsey**	**Ruby Cabernet**	Red	+++	+++	*
15-5081/5082	**Ramsey**	**Ruby Cabernet**	Red	-	-	-
15-5043/5044	**Ramsey**	**Shiraz**	Red	+++	++	-
15-5077/5078	**Ramsey**	**Shiraz**	Red	-	+	-
15-5079/5080	**Ramsey**	**Shiraz**	Red	++	++	*
		Total		58	82	19

(-) = Negative for GLRaV-3; (+) = Weak PCR band strength, (++) = Medium PCR band strength; and (+++) = Strong PCR band strength. (*) = Selected for further NGS analysis and (-) = Not selected for NGS analysis

Table 2: *Viti*- and *Foveavirus* amplicon Direct Sanger sequencing BLAST results

Accession	Sample	E-	%	GenBank	Description
14-9008	R99	1E-18	70%	EU247957.1	GVB clone 99B.SdP2.10
14-9021	R99	2E-21	72%	EU247957.1	GVB clone 99B.SdP2.10
14-9022	R99	2E-14	80%	EU247957.1	GVB clone 99B.SdP2.10
15-5001	Ramsey	8E-77	91%	JX402759.1	GVE isolate WAHH2
15-5003	Ramsey	9E-38	87%	JN683371.1	GRSPaV isolate GR1
15-5009	Ramsey	2E-46	90%	JX402759.1	GVE isolate WAHH2
15-5011	R99	2E-54	92%	JN683371.1	GRSPaV isolate GR1
15-5013	R99	4E-68	89%	FR691076.1	GRSPaV isolate MG
15-5019	R99	2E-66	88%	FR691076.1	GRSPaV isolate MG
15-5023	101-14	4E-55	93%	JN683371.1	GRSPaV isolate GR1
15-5025	101-14	5E-16	80%	FJ884330.1	GVB clone
15-5029	101-14	1E-17	72%	EU247957.1	GVB clone 99B.SdP2.10
15-5031	R99	2E-19	72%	EU247957.1	GVB clone 99B.SdP2.10
15-5035	R99	2E-20	73%	EU247957.1	GVB clone 99B.SdP2.10
15-5037	R99	1E-22	74%	EU247957.1	GVB clone 99B.SdP2.10
15-5039	R99	7E-19	82%	EU247957.1	GVB clone 99B.SdP2.10
15-5043	Ramsey	7E-14	91%	EU247957.1	GVB clone 99B.SdP2.10
15-5045	R99	5E-54	93%	FJ884329.1	GVB clone SH470.N.5A34
15-5051	R99	1E-16	81%	EU247957.1	GVB clone 99B.SdP2.10
15-5053	R99	1E-34	86%	KF013741.1	GVA clone LVCH 92-072
15-5063	R99	2E-71	90%	KC427107.1	GRSPaV isolate WA
15-5065	R99	7E-71	90%	X75448.1	GVB genomic RNA
15-5067	R99	2E-66	87%	X75448.1	GVB genomic RNA
15-5069	R99	5E-42	86%	DQ864490.1	GVB isolate Sd7
15-5073	R99	4E-61	86%	DQ855088.1	GVA isolate P163-1
15-5079	Ramsey	7E-58	85%	KC962564.1	GVA isolate I327-5
15-5083	R99	2E-70	89%	DQ855084.2	GVA isolate GTG11-1
15-5085	R99	2E-58	93%	JN683371.1	GRSPaV isolate GR1
16-0012	Ruby cabernet	8E-25	78%	KF013715.1	GVA clone H6TM 2-3-2
16-0013	R99	3E-71	94%	JX402759.1	GVE isolate WAHH2
16-0014	Ruby cabernet	5E-63	88%	KF013115.1	GVA clone H6TM2-3-2
16-0016	Ruby cabernet	1E-50	92%	DQ855082.2	GVA isolate P163-M5
16-0017	R99	3E-50	92%	JX402759.1	GVE isolate WAHH2
16-0018	Malbec	2E-16	94%	EU247957.1	GVB clone 99B.SdP2.10
16-0019	R99	8E-23	73%	EU247957.1	GVB clone 99B.SdP2.10
16-0020	Malbec	1E-19	78%	DQ855082.2	GVA isolate P163-M5
16-0022	Malbec	7E-59	97%	KP114220.1	GVF isolate V5
16-0024	Pinotage	6E-58	97%	JX402759.1	GVE isolate WAHH2
16-0026	Pinotage	3E-27	75%	EU247957.1	GVB clone 99B.SdP2.10
16-0028	Pinotage	4E-60	97%	JX402759.1	GVE isolate WAHH2
16-0029	R99	1E-76	93%	FR691076.1	GRSPaV isolate MG
16-0030	Tinta barrocca	2E-09	76%	EU247957.1	GVB clone 99B.SdP2.10
16-0031	R99	2E-67	87%	FR691076.1	GRSPaV isolate MG
16-0032	Tinta barrocca	2E-51	90%	KF013715.1	GVA clone H6TM 2-32

16-0034	Pontac	5E-68	89%	KF013715.1	GVA clone H6TM 2-32
16-0036	Pontac	2E-66	89%	KF013715.1	GVA clone H6TM 2-32
16-0038	Pontac	2E-66	85%	KF013715.1	GVA clone H6TM 2-32
16-0039	R99	4E-57	82%	KF013715.1	GVA clone H6TM 2-32
16-0040	Cinsault	4E-43	88%	KF013715.1	GVA clone H6TM 2-32
16-0041	R99	2E-23	73%	EU247957.1	GVB clone 99B.SdP2.10
16-0042	Cinsault	4E-41	69%	KF013715.1	GVA clone H6TM 2-32
16-0043	R99	9E-64	87%	DQ787959.1	GVA isolate GTR1-1
16-0044	Cincault	2E-61	86%	DQ787959.1	GVA isolate GTR1-1
16-0045	R99	6E-60	86%	DQ787959.1	GVA isolate GTR1-1
16-0046	Cabernet	2E-14	76%	KF013686.1	GVA clone HHC815c
16-0047	R99	2E-23	73%	EU247957.1	GVB clone 99B.SdP2.10
16-0048	Cabernet	7E-23	73%	EU247957.1	GVB clone 99B.SdP2.10
16-0049	R99	3E-19	77%	KX522545.1	GVB isolate GVB 248
16-0050	Cabernet	3E-19	77%	KX522545.1	GVB isolate GVB 248



Optimization of length fraction and similarity fraction:
The more reads that map during reference mapping, the less reads are available for *de novo* assembly, possibly decreasing amount of contigs generated.

- Various length fraction*similarity fraction combinations were tested on the positive control and a graph generated
- A point, with the highest reads that was surrounded by a smaller slope of change in the reads that mapped, was chosen and was determined to be 0.52*0.95
- In retrospect this makes sense when considering the reference size divided by the average read length (155bp/280bp = 0.55)

↓

Reference mapping all samples using optimized parameters:
Length fraction*Similarity fraction combination used = 0.52*0.95
= Even thought the amount of reads of reference mapping was optimized, sequences previously found in *de novo* assembly and added to reference library was not detected in the reference mapping.

↓

***de novo* assembly of unmapped reads:**
Unmapped reads of the 0.52*0.95 "ignore" reference mapping was used in *de novo* assembly, using default *de novo* assembly parameters.
= Found that amount of contigs are still impractically high

- Expanding the reference library was unsuccessful in reducing the amount of contigs generated.

↓

Implementation of arbitrary reads mapping back cut-off:
The next attempt in reducing the processing of contigs is to implement a cut off based on the amount of reads that mapped back to a contig. Doing this will make sure that the highest represented contigs are given priority to enable best representation of presence.

- Arbitrary cut-off of 0.20% was implemented
- To express the amount of reads mapped back to the contig in a percentage it was divided by the total reads matched during *de novo*.
- The cut off was not tested on the positive control, since it had a high % reads reference mapped and not many contigs resulting of *de novo* assembly. Instead the cut-off was tested on 14-9007 because it had a % reads reference mapped of 0.12%, and generated up to 700 contigs.

= The result was a change of 752 total contigs generated to 125 contigs remaining after 0.20% cut-off, and in the process only discarding 1.5% of the de novo matched reads

↓

Implementation of arbitrary reads mapping back cut-off:

The next attempt in reducing the processing of contigs is to implement a cut off based on the amount of reads that mapped back to a contig. Doing this will make sure that the highest represented contigs are given priority to enable best representation of presence.

- Arbitrary cut-off of 0.20% was implemented
- To express the amount of reads mapped back to the contig in a percentage it was divided by the total reads matched during *de novo*.
- The cut off was not tested on the positive control, since it had a high % reads reference mapped and not many contigs resulting of *de novo* assembly. Instead the cut-off was tested on 14-9007 because it had a % reads reference mapped of 0.12%, and generated up to 700 contigs.

= The result was a change of 752 total contigs generated to 125 contigs remaining after 0.20% cut-off, and in the process only discarding 1.5% of the de novo matched reads

How *de novo* mapping back parameters affect amount of contigs:

de novo also has length fraction*similarity fraction parameters regarding the mapping back to contig feature. Optimization of this parameter was done by varying the length fraction*similarity fraction combination of the *de novo* process, using s=14-9007 sample mentioned earlier.

- A bar graph generated showed the % of reads represented above the 0.20% arbitrary cut-off of every parameter tested. The optimal combination would be one that gets the most reads to be represented above the 0.20% arbitrary cut-off.

=The parameter that gave the most optimal representation (84.52%) of reads above the cut-off was 0.6*0.9.

***de novo* of unmapped reads:**

The optimal *de novo* mapping back parameter was used on the rest of the samples' unmapped reads obtained from the optimized reference mapping.

Multi-BLAST

Arbitrary 0.20% cut-off implementation

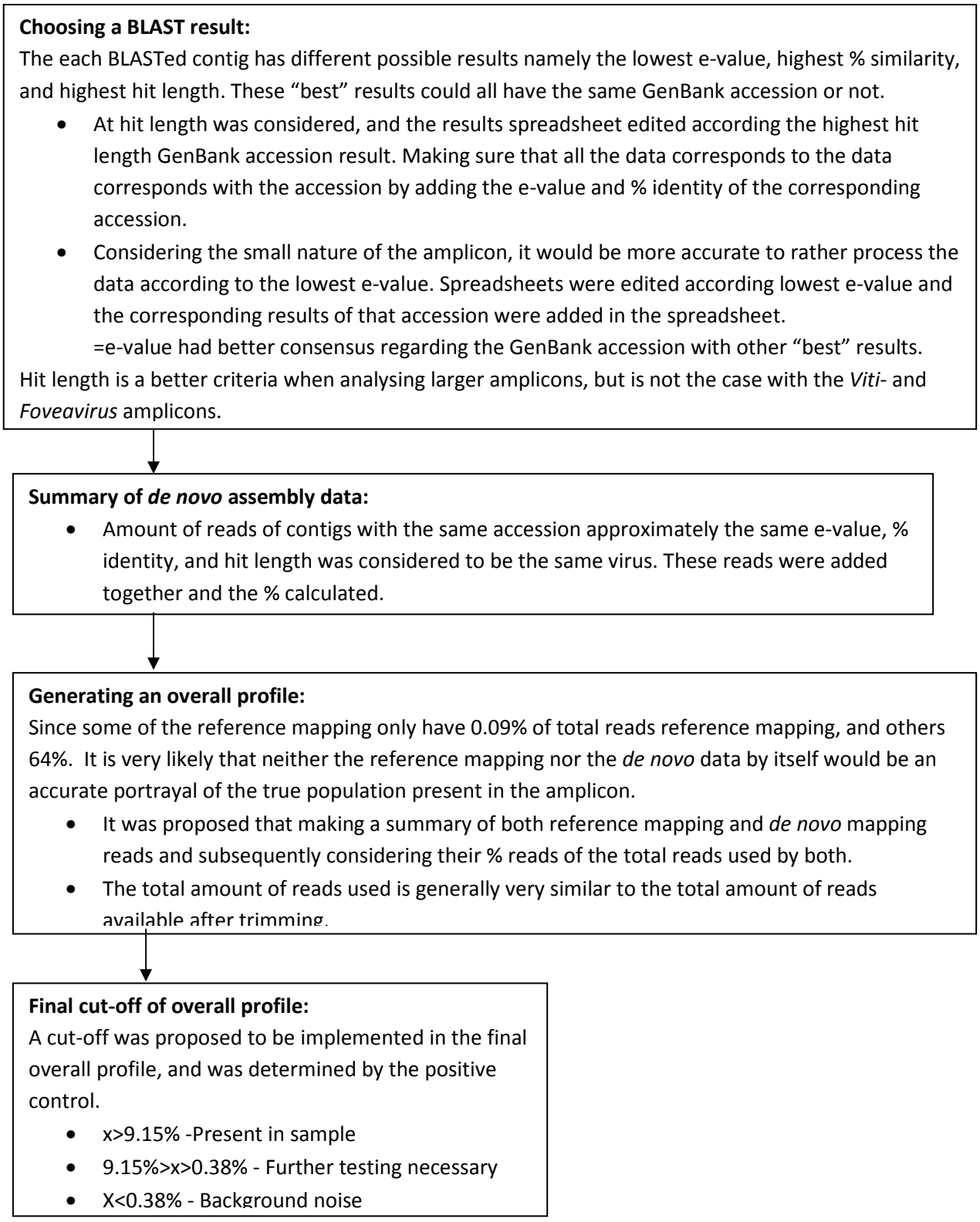


Figure 1: Flow chart illustrating both the thought process and actual process in Viti- and Foveavirus amplicon Illumina MiSeq data analysis pipeline and the thought process thereof

*Table 3: The top results of number of reads of **positive control** sequencing run **replicate 1** that mapped to references at 0.52*0.95 (Length fraction*Similarity fraction) used for parameter optimization. The positive control consists of Viti- and Foveavirus nested PCR amplicons of the RNA dependent RNA polymerase (RdRp) obtained from identified clones. Viruses in shaded rows of the table are known to occur in the positive control sample. Percentage of reads mapped is listed in column 4, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 5, 6, and 7. Reads above the line are known constituents of the positive control*

Name	Total reads mapped	Reads mapped	% Reads mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
GVA-92-063-GVA-B	660040	190487	28.86%	4615	185872	190181	GVA-92-063-GVA-B
GRSPaV VA	660040	174421	26.43%	4231	170190	174288.2	GRSPaV VA
GVA-GTG11-1	660040	134292	20.35%	6768	127524	133242.2	GVA-GTG11-1
GVA-92-063-GVA-A	660040	98132	14.87%	2726	95406	97908.04	GVA-92-063-GVA-A
GVB clone	660040	62222	9.43%	1720	60502	61723.4	GVB clone
GVA-P163-1	660040	202	0.03%	92	110	201.2922	GVA-P163-1
GRSPaV-GR3	660040	93	0.01%	25	68	92.68831	GRSPaV-GR3
GVB-clone-99B.SdZ5.2	660040	55	0.01%	37	18	54.46753	GVB-clone-99B.SdZ5.2
GRSPaV-GG	660040	52	0.01%	8	44	51.94156	GRSPaV-GG
GVB-H1	660040	49	0.01%	11	38	48.81169	GVB-H1
RSPaV-RSP47-4	660040	12	0.00%	0	12	12	RSPaV-RSP47-4

*Table 4: Summarized table of the identity of contigs obtained during the de novo of unmapped reads discarded by **positive control** sequencing run **replicate 1** reference mapping (Table 3), while using 0.6*0.9 mapping back parameters. Lowest e-value of MultiBLAST results determined identity of contig that can be found in the last column, corresponding information such as % identity and hit length can be found in column 5 and 6*

Reads matched	Total reads matched	% Reads matched	Average Lowest E-value	Average identity %	Average hit length	Accession (E-value)	Description (E-value)
10632	13070	4.99%	5E-64	91.5	200	DQ855084	*Grapevine virus A* isolate GTG11-1, complete genome
1742	13070	13.33%	2E-86	95.98	199	EU247952	Rupestris stem pitting-associated virus clone 31.99B.SdP2.SY.SA RNA-dependent RNA polymerase gene, partial cds
696	13070	5.33%	7E-93	98.4925	199	EU247956	*Grapevine virus B* clone 99B.SdZ5.2 RNA-dependent RNA polymerase gene, partial cds

*Table 5: Overall amount of **positive control** sequencing run **replicate 1** reads mapping. The positive control consists of Viti- and Foveavirus nested PCR amplicons of the RNA dependent RNA polymerase (RdRp) obtained from identified clones. Reference mapping and de novo assembly results are combined to generate overall profile to better represent amplicon contents of the sample. Viruses above the line are considered to be present in the sample, and viruses in shaded rows of the table are known to occur in the positive control sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GVA-92-063-GVA-B	660040	190487	28.86%	13070	0	0.00%	673110	190487	28.30%
GRSPaV VA	660040	174421	26.43%	13070	0	0.00%	673110	174421	25.91%
GVA-GTG11-1	660040	134292	20.35%	13070	10632	81.35%	673110	144924	21.53%
GVA-92-063-GVA-A	660040	98132	14.87%	13070	0	0.00%	673110	98132	14.58%
GVB clone	660040	62222	9.43%	13070	696	5.33%	673110	62918	9.35%
Rupestris stem pitting-associated virus clone 31.99B.SdP2.SY.SA	660040	0	0.00%	13070	1742	13.33%	673110	1742	0.26%
GVA-P163-1	660040	202	0.03%	13070	0	0.00%	673110	202	0.03%

*Table 6: The top results of number of reads of **positive control** sequencing run **replicate 2** that mapped to references at 0.52*0.95 (Length fraction*Similarity fraction) used for parameter optimization. The positive control consists of Viti- and Foveavirus nested PCR amplicons of the RNA dependent RNA polymerase (RdRp) obtained from identified clones. Viruses in shaded rows of the table are known to occur in the positive control sample. Percentage of reads mapped is listed in column 4, indicating relative abundance of specific strains in the population. Additional information relating to the mapping can be found in column 5, 6, and 7. Reads above the line are known constituents of the positive control*

Name	Total reads mapped	Reads mapped	% Reads mapped	Single reads	Reads in pairs	Average coverage	Reference sequence
GVA-92-063-GVA-B	164168	46962	28.61%	1406	45556	46881.03	GVA-92-063-GVA-B
GRSPaV VA	164168	42977	26.18%	1665	41312	42937.51	GRSPaV VA
GVA-GTG11-1	164168	34275	20.88%	1883	32392	33997.49	GVA-GTG11-1
GVA-92-063-GVA-A	164168	24678	15.03%	816	23862	24626.46	GVA-92-063-GVA-A
GVB clone	164168	15197	9.26%	647	14550	15073.26	GVB clone
GVA-P163-1	164168	30	0.02%	12	18	29.77273	GVA-P163-1
GRSPaV-GR3	164168	18	0.01%	2	16	17.81818	GRSPaV-GR3
GVB-clone-99B.SdZ5.2	164168	15	0.01%	5	10	14.98701	GVB-clone-99B.SdZ5.2
GRSPaV-GG	164168	9	0.01%	1	8	9	GRSPaV-GG
RSPaV-clone-31.99B.SdP2.SY.SA	164168	3	0.00%	3	0	3	RSPaV-clone-31.99B.SdP2.SY.SA

*Table 7: Summarized table of the identity of contigs obtained during the de novo of unmapped reads discarded by **positive control** sequencing run **replicate 2** reference mapping (Table 3), while using 0.6*0.9 mapping back parameters. Lowest e-value of MultiBLAST results determined identity of contig that can be found in the last column, corresponding information such as % identity and hit length can be found in column 5 and 6*

Reads matched	Total reads matched	% Reads matched	Average Lowest E-value	Average identity %	Average hit length	Accession (E-value)	Description (E-value)
2653	5415	48.99%	5E-78	94.0904	194	DQ855084	*Grapevine virus A* isolate GTG11-1, complete genome
678	5415	12.52%	2E-80	95.7447	188	DQ855088	*Grapevine virus A* isolate P163-1 putative replicase, movement protein, capsid protein, and RNA-binding protein genes, complete cds; and unknown gene
729	5415	13.46%	3E-85	96.4103	195	EU247952	Rupestris stem pitting-associated virus clone 31.99B.SdP2.SY.SA

*Table 8: Overall amount of **positive control** sequencing run **replicate 2** reads mapping. Positive control consists of Viti- and Foveavirus nested PCR amplicons of the RNA dependent RNA polymerase (RdRp) obtained from identified clones. Reference mapping and de novo assembly results are combined to generate overall profile to better represent amplicon contents of the sample. Viruses above the line are considered to be present in the sample, and viruses in shaded rows of the table are known to occur in the positive control sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GVA-92-063-GVA-B	164168	46962	28.61%	5415	0	0.00%	169583	46962	27.69%
GRSPaV VA	164168	42977	26.18%	5415	0	0.00%	169583	42977	25.34%
GVA-GTG11-1	164168	34275	20.88%	5415	2653	48.99%	169583	36928	21.78%
GVA-92-063-GVA-A	164168	24678	15.03%	5415	0	0.00%	169583	24678	14.55%
GVB clone	164168	15197	9.26%	5415	0	0.00%	169583	15197	8.96%
Rupestris stem pitting-associated virus clone 31.99B.SdP2.SY.SA	164168	0	0.00%	5415	729	13.46%	169583	729	0.43%
GVA-P163-1	164168	30	0.02%	5415	678	12.52%	169583	708	0.42%

*Table 9: Number reads mapped and matched of both the **positive control replicate 1** and **positive control replicate 2** Illumina MiSeq run and their total. Positive control consists of Viti- and Foveavirus nested PCR amplicons of the RNA dependent RNA polymerase (RdRp) obtained from identified clones. Reference mapping and de novo assembly results are combined to generate overall profile to better represent amplicon contents of the sample. The two replicate's information is added together to maximize the depth and control for variance in the Illumina MiSeq sequencing itself. Viruses above the line are considered to be present in the sample, and viruses in shaded rows of the table are known to occur in the positive control sample.*

GLRaV-3 variant	Positive control sequencing run replicate 1 overall			Positive control sequencing run replicate 2 overall			Positive control replicate 1 + Replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVA-92-063-GVA-B	673110	190487	28.30%	169583	46962	27.69%	842693	237449	28.18%
GRSPaV VA	673110	174421	25.91%	169583	42977	25.34%	842693	217398	25.80%
GVA-GTG11-1	673110	144924	21.53%	169583	36928	21.78%	842693	181852	21.58%
GVA-92-063-GVA-A	673110	98132	14.58%	169583	24678	14.55%	842693	122810	14.57%
GVB clone	673110	62918	9.35%	169583	15197	8.96%	842693	78115	9.27%
Rupestris stem pitting-associated virus clone 31.99B.SdP2.SY.SA	673110	1742	0.26%	169583	729	0.43%	842693	2471	0.29%
GVA-P163-1	673110	202	0.03%	169583	708	0.42%	842693	910	0.11%

*Table 10: Number reads mapped and matched of both **14-9005 sequencing run replicate 1** and **14-9005 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9005, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9005 sequencing run replicate 1 overall			14-9005 sequencing run replicate 2 overall			14-9005 replicate 1 + Duplicate replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GRSPaV-GG	239286	170166	71.11%	61956	41508	67.00%	301242	211674	70.27%
RSPaV-RSP47-4	239286	40053	16.74%	61956	10893	17.58%	301242	50946	16.91%
RSPaV-1	239286	26685	11.15%	61956	7454	12.03%	301242	34139	11.33%
GRSPaV (AF026279)	239286	0	0.00%	61956	1804	2.91%	301242	1804	0.60%
GVB-953-1	239286	708	0.30%	61956	4	0.01%	301242	712	0.24%
GVB isolate Sd7	239286	611	0.26%	61956	0	0.00%	301242	611	0.20%
Unknown	239286	403	0.17%	61956	128	0.21%	301242	531	0.18%
GVD	239286	201	0.08%	61956	0	0.00%	301242	201	0.07%
GVA-P163-1	239286	0	0.00%	61956	129	0.21%	301242	129	0.04%
GVB-H1	239286	108	0.05%	61956	0	0.00%	301242	108	0.04%

*Table 11: Number reads mapped and matched of both **14-9007 sequencing run replicate 1** and **14-9007 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9007, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9007 sequencing run replicate 1 overall			14-9007 sequencing run replicate 2 overall			14-9007 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVD (Y15892)	487232	251888	51.70%	130355	81126	62.23%	617587	333014	53.92%
Unknown	487232	160084	32.86%	130355	48674	37.34%	617587	208758	33.80%
GRSPaV-GG	487232	330	0.07%	130355	78	0.06%	617587	408	0.07%
RSPaV-1	487232	82	0.02%	130355	22	0.02%	617587	104	0.02%
GVA-GTG11-1	487232	71	0.01%	130355	0	0.00%	617587	71	0.01%
GRSPaV-GR3	487232	61	0.01%	130355	0	0.00%	617587	61	0.01%
RSPaV-RSP47-4	487232	13	0.00%	130355	6	0.00%	617587	19	0.00%

*Table 12: Number reads mapped and matched of both **14-9013 sequencing run replicate 1** and **14-9013 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9013, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

	14-9013 sequencing run replicate 1 overall			14-9013 sequencing run replicate 2 overall			14-9013 replicate 1 + replicate 2		
GLRaV-3 variant	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GRSPaV clone 99B.SdP2.8	402861	318349	79.02%	73909	53653	72.59%	476770	372002	78.03%
GVA clone LVCH92-07-2	402861	82939	20.59%	73909	7404	10.02%	476770	90343	18.95%
Unknown	402861	0	0.00%	73909	5483	7.42%	476770	5483	1.15%
GVA clone SLWZF1-3c1	402861	0	0.00%	73909	4702	6.36%	476770	4702	0.99%
GVA clone LVZT93-09-19	402861	0	0.00%	73909	708	0.96%	476770	708	0.15%
GVA clone LREP100 4c	402861	0	0.00%	73909	232	0.31%	476770	232	0.05%

.

Table 13: *Number reads mapped and matched of both **14-9021 sequencing run replicate 1** and **14-9021 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9021, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9021 sequencing run replicate 1 overall			14-9021 sequencing run replicate 2 overall			14-9021 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
Unknown	822181	820169	99.76%	192840	192550	99.85%	1015021	1012719	99.77%
GRSPaV-GG	822181	202	0.02%	192840	49	0.03%	1015021	251	0.02%
GVB-clone-SH470.N.5A34	822181	115	0.01%	192840	30	0.02%	1015021	145	0.01%
GVB-H1	822181	54	0.01%	192840	11	0.01%	1015021	65	0.01%
GVA-GTG11-1	822181	31	0.00%	192840	0	0.00%	1015021	31	0.00%

*Table 14: Number reads mapped and matched of both **14-9035 sequencing run replicate 1** and **14-9035 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9035, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9035 sequencing run replicate 1 overall			14-9035 sequencing run replicate 2 overall			14-9035 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVB-H1	356841	335712	94.08%	112286	104801	93.33%	469127	440513	93.90%
GVB clone 99B.SdP2.10	356841	20212	5.66%	112286	7401	6.59%	469127	27613	5.89%
GVA P163-1	356841	436	0.12%	112286	3	0.00%	469127	439	0.09%
GRSPaV-GG	356841	93	0.03%	112286	23	0.02%	469127	116	0.02%
RSPaV-1	356841	9	0.00%	112286	10	0.01%	469127	19	0.00%

*Table 15: Number reads mapped and matched of both **14-9037 sequencing run replicate 1** and **14-9037 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9037, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9037 sequencing run replicate 1 overall			14-9037 sequencing run replicate 2 overall			14-9037 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVB-clone-SH470.N.5A34	442498	435251	98.36%	90126	87682	97.29%	532624	522933	98.18%
GRSPaV-GG	442498	3708	0.84%	90126	847	0.94%	532624	4555	0.86%
RSPaV-1	442498	714	0.16%	90126	184	0.20%	532624	898	0.17%
GVB-953-1	442498	139	0.03%	90126	36	0.04%	532624	175	0.03%
GVA-P163-1	442498	120	0.03%	90126	10	0.01%	532624	130	0.02%
GRSPaV-GR3	442498	65	0.01%	90126	0	0.00%	532624	65	0.01%
GVB-H1	442498	54	0.01%	90126	15	0.02%	532624	69	0.01%

*Table 16: Number reads mapped and matched of both **14-9039 sequencing run replicate 1** and **14-9039 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9039, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9039 sequencing run replicate 1 overall			14-9039 sequencing run replicate 2 overall			14-9039 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
Unknown	168792	161852	95.89%	64810	62604	96.60%	233602	224456	96.08%
GVB-H1	168792	6202	3.67%	64810	2082	3.21%	233602	8284	3.55%
GVA-P163-1	168792	52	0.03%	64810	70	0.11%	233602	122	0.05%
GRSPaV-GG	168792	14	0.01%	64810	0	0.00%	233602	14	0.01%
GVA-GTG11-1	168792	9	0.01%	64810	6	0.01%	233602	15	0.01%
GVB 953-1	168792	0	0.00%	64810	6	0.01%	233602	6	0.00%

*Table 17: Number reads mapped and matched of both **14-9041 sequencing run replicate 1** and **14-9041 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9041, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9041 sequencing run replicate 1 overall			14-9041 sequencing run replicate 2 overall			14-9041 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVA-Bio	19622	10091	51.43%	2846	0	0.00%	22468	10091	44.91%
GVB clone 99B.SdP2.10	19622	6279	32.00%	2846	587	20.63%	22468	6866	30.56%
GVB isolate Sd7	19622	2100	10.70%	2846	248	8.71%	22468	2348	10.45%
GVB H1	19622	0	0.00%	2846	1890	66.41%	22468	1890	8.41%
GVA-BMo32-1	19622	275	1.40%	2846	0	0.00%	22468	275	1.22%
GVA-clone-BVPN1_6c	19622	272	1.39%	2846	0	0.00%	22468	272	1.21%
GVA clone LVCH92-07_3c2	19622	199	1.01%	2846	0	0.00%	22468	199	0.89%
GVA-clone-CBSM119-5	19622	197	1.00%	2846	0	0.00%	22468	197	0.88%
GVA-clone-CBSM119-10	19622	118	0.60%	2846	0	0.00%	22468	118	0.53%
GVA P163-1	19622	0	0.00%	2846	38	1.34%	22468	38	0.17%

*Table 18: Number reads mapped and matched of both **14-9053 sequencing run replicate 1** and **14-9053 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9053, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9053 sequencing run replicate 1 overall			14-9053 sequencing run replicate 2 overall			14-9053 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVB isolate Sd7	1977040	721230	36.48%	227765	155589	68.31%	2204805	876819	39.77%
GVB-H1	1977040	307609	15.56%	227765	36519	16.03%	2204805	344128	15.61%
GVB-clone-99B.SdP2.10	1977040	124891	6.32%	227765	23215	10.19%	2204805	148106	6.72%
GVB-953-1	1977040	33029	1.67%	227765	3832	1.68%	2204805	36861	1.67%
GVB clone SH470.N.5A34	1977040	36791	1.86%	227765	0	0.00%	2204805	36791	1.67%
GVA-P163-1	1977040	20033	1.01%	227765	2277	1.00%	2204805	22310	1.01%
GVA-GTG11-1	1977040	14559	0.74%	227765	1496	0.66%	2204805	16055	0.73%
GVA clone LVCH92-07_3c2	1977040	6902	0.35%	227765	940	0.41%	2204805	7842	0.36%
GRSPaV-GR3	1977040	5816	0.29%	227765	630	0.28%	2204805	6446	0.29%
GVA-P163-M5	1977040	4626	0.23%	227765	1544	0.68%	2204805	6170	0.28%
GVB isolate 94/971	1977040	3842	0.19%	227765	0	0.00%	2204805	3842	0.17%
GRSPaV-GG	1977040	329	0.02%	227765	23	0.01%	2204805	352	0.02%

*Table 19: Number reads mapped and matched of both **14-9055 sequencing run replicate 1** and **14-9055 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9055, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9055 sequencing run replicate 1 overall			14-9055 sequencing run replicate 2 overall			14-9055 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVB isolate Sd7	9712	9130	94.01%	2879	2863	99.44%	12591	11993	95.25%
GVD (Y15892)	9712	92	0.95%	2879	0	0.00%	12591	92	0.73%
GVA-GTG11-1	9712	58	0.60%	2879	0	0.00%	12591	58	0.46%
GRSPaV-GR3	9712	39	0.40%	2879	0	0.00%	12591	39	0.31%
GVB-H1	9712	30	0.31%	2879	0	0.00%	12591	30	0.24%
GVB-clone-SH470.N.5A34	9712	12	0.12%	2879	0	0.00%	12591	12	0.10%
GRSPaV-GG	9712	2	0.02%	2879	0	0.00%	12591	2	0.02%
GVA-Bio	9712	0	0.00%	2879	2	0.07%	12591	2	0.02%

*Table 20: Number reads mapped and matched of both **14-9057 sequencing run replicate 1** and **14-9057 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9057, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9057 sequencing run replicate 1 overall			14-9057 sequencing run replicate 2 overall			14-9057 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVB isolate Sd7	605	3	0.50%	1959	1594	81.37%	2564	1597	62.29%
GVA-BMo32-1	605	299	49.42%	1959	100	5.10%	2564	399	15.56%
GVB-clone-SH470.N.5A34	605	13	2.15%	1959	243	12.40%	2564	256	9.98%
GVA-GTG11-1	605	36	5.95%	1959	0	0.00%	2564	36	1.40%
GRSPaV-GR3	605	30	4.96%	1959	0	0.00%	2564	30	1.17%
GVB-H1	605	24	3.97%	1959	2	0.10%	2564	26	1.01%
GVB-953-1	605	8	1.32%	1959	0	0.00%	2564	8	0.31%
Unknown	605	0	0.00%	1959	4	0.20%	2564	4	0.16%
GVA-P163-1	605	2	0.33%	1959	0	0.00%	2564	2	0.08%
GVB isolate 953-1	605	0	0.00%	1959	2	0.10%	2564	2	0.08%

*Table 21: Number reads mapped and matched of both **14-9081 sequencing run replicate 1** and **14-9081 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9081, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9081 sequencing run replicate 1 overall			14-9081 sequencing run replicate 2 overall			14-9081 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
GVA-GTG11-1	33	12	36.36%	11	0	0.00%	44	12	27.27%
GRSPaV-GR3	33	12	36.36%	11	0	0.00%	44	12	27.27%
GVB-clone-SH470.N.5A34	33	8	24.24%	11	1	9.09%	44	9	20.45%
GVA-P163-1	33	1	3.03%	11	6	54.55%	44	7	15.91%
GVA isolate GTR1-1	33	0	0.00%	11	2	18.18%	44	2	4.55%
GVA-Bio	33	0	0.00%	11	0	0.00%	44	0	0.00%

*Table 22: Number reads mapped and matched of both **14-9083 sequencing run replicate 1** and **14-9083 sequencing run replicate 2** Illumina MiSeq run and their total. By adding the two sequencing run replicates together it maximizes the amount of data for accession 14-9083, giving a more accurate representation of the amplicon presence in the sample. The reads above the line represents what is considered to be present in the sample*

GLRaV-3 variant	14-9083 sequencing run replicate 1 overall			14-9083 sequencing run replicate 2 overall			14-9083 replicate 1 + replicate 2		
	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence	Total reads used (Positive control + duplicate)	Total reads mapped	% Overall presence
Unknown	289	142	49.13%	134	117	87.31%	423	259	61.23%
GVA-P163-1	289	51	17.65%	134	2	1.49%	423	53	12.53%
GVA isolate GTR1-1	289	13	4.50%	134	6	4.48%	423	19	4.49%
GVB-H1	289	2	0.69%	134	2	1.49%	423	4	0.95%
GVA clone LVCH92-07_3c2	289	2	0.69%	134	0	0.00%	423	2	0.47%
GVA-Bio	289	0	0.00%	134	0	0.00%	423	0	0.00%

*Table 23: Number reads mapped and matched of accession number **15-5009** Illumina MiSeq run. Adding reference mapping and de novo assembly results together represents the overall 15-5009, giving a more accurate representation of the amplicon presence in the sample. Reads above the line represents what is considered to be present in the sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GVE-WAHH2	239835	239576	99.89%	94895	22	0.02%	334730	239598	71.58%
GVE-SA94	239835	161	0.07%	94895	0	0.00%	334730	161	0.05%
GVB-clone-SH470.N.5A34	239835	67	0.03%	94895	0	0.00%	334730	67	0.02%
RSPaV-2	239835	12	0.01%	94895	0	0.00%	334730	12	0.00%
GVA-GTG11-1	239835	11	0.00%	94895	0	0.00%	334730	11	0.00%

*Table 24: Number reads mapped and matched of accession number **15-5029** Illumina MiSeq run. Adding reference mapping and de novo assembly results together represents the overall 15-5029, giving a more accurate representation of the amplicon presence in the sample. Reads above the line represents what is considered to be present in the sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
Unknown	2473	0	0.00%	619207	502485	81.15%	621680	502485	80.83%
GVB clone 99B.SD8.32.7.12	2473	0	0.00%	619207	36522	5.90%	621680	36522	5.87%
GRSPaV isolate GR1	2473	0	0.00%	619207	23079	3.73%	621680	23079	3.71%
GVB-clone-SH470.N.5A34	2473	67	2.71%	619207	16703	2.70%	621680	16770	2.70%
GVB isolate Sd7	2473	0	0.00%	619207	7890	1.27%	621680	7890	1.27%
GVA clone CBSM119-8	2473	0	0.00%	619208	3646	0.59%	621681	3646	0.59%
GRSPaV-MG	2473	1733	70.08%	619207	0	0.00%	621680	1733	0.28%
GVB clone 56.AA219G.N	2473	0	0.00%	619207	1257	0.20%	621680	1257	0.20%
RSPaV-1	2473	516	20.87%	619207	0	0.00%	621680	516	0.08%
GRSPaV-WA	2473	97	3.92%	619207	0	0.00%	621680	97	0.02%
GVE-WAHH2	2473	26	1.05%	619207	0	0.00%	621680	26	0.00%

*Table 25: Number reads mapped and matched of accession number **15-5031** Illumina MiSeq run. Adding reference mapping and de novo assembly results together represents the overall 15-5031, giving a more accurate representation of the amplicon presence in the sample. Reads above the line represents what is considered to be present in the sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
Unknown	139	0	0.00%	270869	263477	97.27%	271008	263477	97.22%
GVB isolate H1	139	0	0.00%	270869	532	0.20%	271008	532	0.20%
GVB-clone-SH470.N.5A34	139	81	58.27%	270869	23	0.01%	271008	104	0.04%
GVA isolate KWVMo4-1	139	0	0.00%	270869	81	0.03%	271008	81	0.03%
GVA clone CBSM119-8	139	0	0.00%	270869	48	0.02%	271008	48	0.02%
GVE-WAHH2	139	44	31.65%	270869	0	0.00%	271008	44	0.02%
GVA clone VHLM16-17	139	0	0.00%	270869	13	0.00%	271008	13	0.00%

*Table 26: Number reads mapped and matched of accession number **15-5045** Illumina MiSeq run. Adding reference mapping and de novo assembly results together represents the overall 15-5045, giving a more accurate representation of the amplicon presence in the sample. Reads above the line represents what is considered to be present in the sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GVB-clone-SH470.N.5A34	15960	15248	95.54%	11273	1990	17.65%	27233	17238	63.30%
GRSPaV clone 61.22F.N.SY.SA	15960	0	0.00%	11273	4953	43.94%	27233	4953	18.19%
GRSPaV clone SH470.N.5A38	15960	0	0.00%	11273	3711	32.92%	27233	3711	13.63%
RSPaV-2	15960	653	4.09%	11273	0	0.00%	27233	653	2.40%
GVB clone 22F.90	15960	0	0.00%	11273	143	1.27%	27233	143	0.53%
GVA isolate I327-5	15960	0	0.00%	11273	110	0.98%	27233	110	0.40%
GVB clone 59.22F.N	15960	0	0.00%	11273	64	0.57%	27233	64	0.24%
GVE-WAHH2	15960	42	0.26%	11273	0	0.00%	27233	42	0.15%
GRSPaV-MG	15960	9	0.06%	11273	0	0.00%	27233	9	0.03%
GVA-GTG11-1	15960	7	0.04%	11273	0	0.00%	27233	7	0.03%

*Table 27: Number reads mapped and matched of accession number **15-5071** Illumina MiSeq run. Adding reference mapping and de novo assembly results together represents the overall 15-5071, giving a more accurate representation of the amplicon presence in the sample. Reads above the line represents what is considered to be present in the sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference Mapping			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference mapping + *de novo*)	Total reads mapped	% Overall presence
GVA-P163-M5	43653	5451	12.49%	395941	355900	89.89%	439594	361351	82.20%
GVA-GTG11-1	43653	38076	87.22%	395941	3527	0.89%	439594	41603	9.46%
GVA isolate KWVMo--1	43653	0	0.00%	395942	28496	7.20%	439595	28496	6.48%
GVA clone SLWZF30-20	43653	0	0.00%	395941	1766	0.45%	439594	1766	0.40%
GVA isolate IR-S7	43653	0	0.00%	395941	475	0.12%	439594	475	0.11%
GVA clone H6TM2-3-2	43653	0	0.00%	395941	470	0.12%	439594	470	0.11%
GVA isolate I327-5	43653	0	0.00%	395941	235	0.06%	439594	235	0.05%
GVB-clone-SH470.N.5A34	43653	26	0.06%	395941	0	0.00%	439594	26	0.01%
GVE-WAHH2	43653	20	0.05%	395941	0	0.00%	439594	20	0.00%

*Table 28: Number reads mapped and matched of accession number **15-5073** Illumina MiSeq run. Adding reference mapping and de novo assembly results together represents the overall 15-5073, giving a more accurate representation of the amplicon presence in the sample. Reads above the line represents what is considered to be present in the sample. The overall profile results can be seen in the last three columns*

GLRaV-3 variant	Reference			*de novo* assembly			Overall		
	Total reads reference mapped	Reads reference mapped	% Reads reference mapped	Total reads matched *de novo*	Reads matched *de novo*	% Reads matched *de novo*	Total reads used (Reference + de novo)	Total reads mapped	% Overall presence
GVA-I327-5-1	145	28	19.31%	34490	34382	99.69%	34635	34410	99.35%
GVB-clone-SH470.N.5A34	145	37	25.52%	34490	0	0.00%	34635	37	0.11%
GVE-WAHH2	145	27	18.62%	34490	0	0.00%	34635	27	0.08%
GVA-GTG11-1	145	17	11.72%	34490	0	0.00%	34635	17	0.05%
GVA-isolate-IR-S7-1	145	12	8.28%	34490	0	0.00%	34635	12	0.03%
RSPaV-2	145	11	7.59%	34490	0	0.00%	34635	11	0.03%
GRSPaV-MG	145	6	4.14%	34490	0	0.00%	34635	6	0.02%

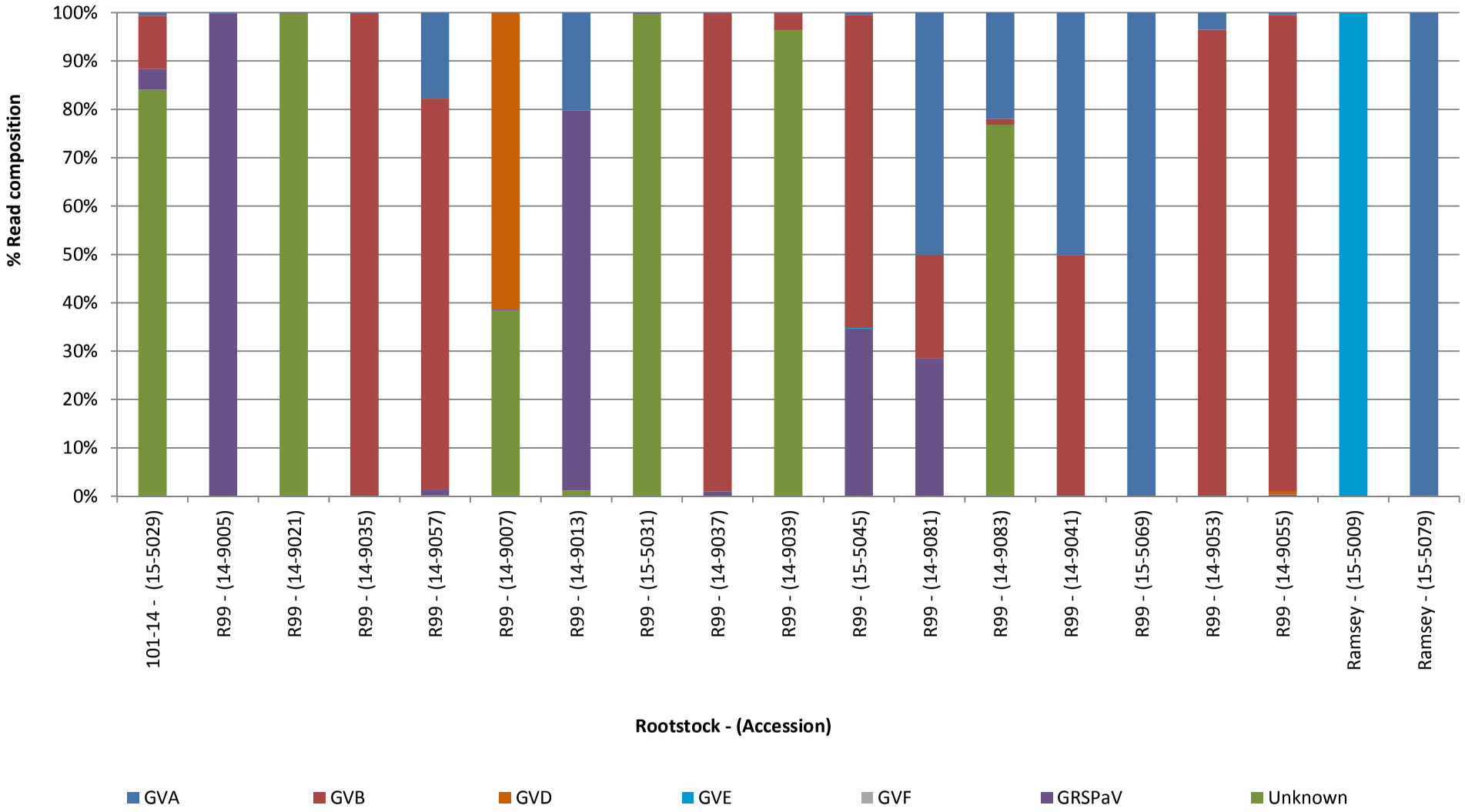


Figure 2: Graphical demonstration of Viti- and Foveavirus Illumina MiSeq read composition of the various rootstock tissue samples

Table 29: Confirmation testing results of rootstock and scion tissue samples

Sample	Rootstock- scion combination		Test			Illumina Sequenced
	Rootstock	Scion	GVA	GVB	GVE	
14/9013	R99	Assyrtiko	-	-	-	*
14/9021	R99	Cabernet franc	-	-	+	*
14/9035	R99	Lumbrusco	-	+	-	*
14/9037	R99	Lakemont seedless	-	-	-	*
14/9039	R99	Lakemont seedless	-	+	N/A	*
14/9041	R99	Planta nova	-	+	+	*
14/9053	R99	Zeni	-	+	+	*
14/9055	R99	Zeni	-	+	-	*
14/9057	R99	Cabernet franc	+	-	+	*
14/9081	R99	Pinotage	-	-	+	*
14/9083	R99	Pinotage	+	-	-	*
15/5007	Ramsey	Ruby cabernet	-	-	+	N/A
15/5009	Ramsey	Ruby cabernet	-	-	+	*
15/5029	101-14	Merlot	-	-	-	*
15/5031	R99	Chardonnay	-	-	-	*
15/5045	R99	Merlot	-	+	-	*
15/5069	R99	Shiraz	+	-	-	*
15/5079	Ramsey	Shiraz	+	-	-	*
16/0013	R99	Ruby cabernet	-	-	+	N/A
16/0014	R99	Ruby cabernet	+	-	+	N/A
16/0049	R99	Cabernet sauvignon	-	-	+	N/A

Table 30: Virus description and GenBank accession number of references sequences used in identification of Viti- and Foveaviruses via CLC genomic workbench analysis of the Illumina MiSeq data of amplicons

Virus	Description	GenBank accession number
GVA	Grapevine virus A isolate Bio	AY676325
	Grapevine virus A isolate BMo21	DQ855087
	Grapevine virus A isolate GTG 11-1	DQ855084
	Grapevine virus A isolate GTR 1-1	DQ787959
	Grapevine virus A isolate GTR 1-2	DQ855086
	Grapevine virus A isolate GTR 1SD-1	DQ855081
	Grapevine virus A isolate I327-5	KC962564
	Grapevine virus A isolate IR-S7	GU084166
	Grapevine virus A isolate Is151	X75433
	Grapevine virus A isolate KWVMo4-1	DQ855083
	Grapevine virus A isolate P163-M5	DQ855082
	Grapevine virus A isolate P163-1	DQ855088
	Grapevine virus A isolate PA3	AF007415
GVB	Grapevine virus B isolate 94_971	EF583906
	Grapevine virus B isolate 953-1	KJ524452
	Grapevine virus B Italy	X75448
	Grapevine virus B isolate H1	GU733707
	Grapevine virus B clone 99B.SdP2.10	EU247957
	Grapevine virus B isolate 3138-01	JX513897
	Grapevine virus B clone 99B.Sd8.32.7.12	FJ884330
	Grapevine virus B isolate Sd7	DQ864490
GVE	Grapevine virus E isolate GFMG-1	KF88015
	Grapevine virus E isolate SA94	GU903012
	Grapevine virus E isolate TvAQ7	AB432910
	Grapevine virus E isolate WAHH2	JX402759
GVF	Grapevine virus F isolate AUD46129	JX105428
	Grapevine virus F isolate V5	KP114220
GRSPaV	Grapevine rupestris stempitting associated virus isolate BS	AY881627
	Grapevine rupestris stempitting associated virus -1	AF057136
	Grapevine rupestris stempitting associated virus isolate SG1	AY881626
	Grapevine rupestris stempitting associated virus isolate GG	JQ922417
	Grapevine rupestris stempitting associated virus isolate MG	FR691076
	Grapevine rupestris stempitting associated virus isolate PN	AY368172
	Grapevine rupestris stempitting associated virus isolate Syrah	AY368590
	Grapevine rupestris stempitting associated virus isolate WA	KC427107