TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Glossary

aa	amino acid
ABI	Applied Biosystems, Inc.
AGVd	Australian grapevine viroid
ArMV	Arabis mosaic virus
ATLV	Agave tequilana leaf virus
AVV	Arracacha virus V
CCD	Charge-Coupled Device
CLRV	Cherry leaf roll virus
CMV	Cucumber mosaic virus
CP	Coat Protein
CSIRO	Commonwealth Scientific and Industrial Research Organisation ,
CTAB	Cetyltrimethylammonium Bromide
ddNTP	Dideoxynucleoside triphosphate
dNTPs	deoxynucleoside triphosphates
DR	Samples from the disease response grou
dsRNA	double-stranded RNA
DW	Dry Weight
ELISA	Enzyme-Linked Immunosorbent Assay
emPCR	emulsion PCR or
EPPO	European and Mediterranean Plant Protection Organisation
FW	Fresh Weight
GAMaV	Grapevine asteroid mosaic-associated virus
GC-HH	Samples from the New Zealand germplasm ‘high-health’ block
GC-LH	Samples from the New Zealand germplasm ‘low-health’ block
GCSV	Grapevine Cabernet Sauvignon reovirus
GFkV	Grapevine fleck virus
GFLV	Grapevine fanleaf virus
GGS	Grafted Grapevine Standards
GGVA	Grapevine geminivirus A
GLRaV-1	Grapevine leafroll-associated virus 1

GLRaV-2	Grapevine leafroll-associated virus 2
GLRaV-3	Grapevine leafroll-associated virus 3
GLRaV-4	Grapevine leafroll-associated virus 4
GLRaV-7	Grapevine leafroll-associated virus 7
GPGV	Grapevine Pinot gris virus
GRBV	Grapevine red blotch virus
GRGV	Grapevine red globe virus
GRLDaV	Grapevine roditis leaf discoloration-associated virus
GRSPaV	Grapevine rupestris stem pitting-associated virus
GRSPaV-1	Grapevine rupestris stem pitting-associated virus 1
GRVFV	Grapevine rupestris vein feathering virus
GSyV-1	Grapevine syrah virus-1
GVA	Grapevine virus A
GVB	Grapevine virus B
GVCV	Grapevine vein clearing virus
GVD	Grapevine virus D
GVE	Grapevine virus E
GVE-like	a divergent isolate related to Grapevine virus E
GVF	Grapevine virus F
GVG	Grapevine virus G
GVH	Grapevine virus H
GVI	Grapevine virus I
GVJ	Grapevine virus J
GYSVd-1	Grapevine yellow speckle viroid 1
HEL	Helicase domain
HLV	Heracleum latent virus
HMMs	Hidden Markov Models
HSP70h	Heat shock 70 homologue
HSP90h	Heat shock 90 homologue
HSVd	Hop stunt viroid
HTS	High Throughput Sequencing
IC-RT-PCR	Immunocapture Reverse Transcriptase Polymerase Chain Reaction
ICTV	International Committee on Taxonomy of Viruses

ICVG	International Council for the Study of Viruses and Virus Diseases of the Grapevine
KAc	Potassium Acetate
kb	kilobase
mAbs	Monoclonal antibodie s
MAVRIC	Monoclonal Antibodies against Viral RNA Intermediates in Cells
MB	Million Bases
MP	Movement Protein
MPI	Ministry for Primary Industries
MTR	Methyl transferase domain
NABP	Nucleic acid binding protein
NGS	Next Generation Sequencing
NZW	New Zealand Winegrowers
ORFs	Open Reading Frames
PAbs	Polyclonal antibodies
PacBio	Pacific Biosciences
PAMV	Potato aucuba mosaic virus
PVS	Potato virus S
PVY	Potato virus Y
QC	Quality Control
RCA	Rolling Circle Amplification
RdRp	RNA-dependant RNA-polymerase
RISC	RNA-induced silencing complex
RMPM	Read Mapped per Million
rRNA	ribosomal RNA
RSP	Rupestris Stem Pitting
RT-PCR	Reverse-Transcription Polymerase Chain reaction
RT-qPCR	Reverse Transcriptase quantitative PCR
RVY	Rumex virus Y renamed Broad-leafed dock virus A by the ICTV committee
SB	Sample from the Sauvignon blanc group
scFv	single-chain Fragment variable
SDS	Sodium Dodecyl Sulfate
sgRNAs	subgenomic RNAs
siRNA	small interfering RNAs

SMRT	Single Molecule Real Time sequencing
SOLiD	Sequencing by Oligo Ligation Detection.
sRNA	small RNA
ssRNA	single-stranded RNA
TA	Titratable Acidity
TFTDaV	Temperate fruit decay-associated virus
ToRSV	Tomato ringspot virus
TRSV	Tobacco ringspot virus
USDA-ARS	United States Department of Agriculture, Agriculture Research Service
UTR	Untranslated region
VANA	Virion-Associated Nucleic Acids
VSD	Virus Surveillance and Diagnosis
VSR	viral suppressor of RNA silencing
WVV-1	Wild vitis virus 1
ZMW	Zero Mode Waveguide

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. Serological methods for the detection of major grapevine viruse Section 1) Introdution; 8) Advantages and limitations of immuno-assays; and 9) Conclusions and Future Directions	
Nature of contribution by PhD candidate	Arnaud Blouin: Contact with the editors manucript direction. redaction of the sections used in the thesis (listed above)
Extent of contribution by PhD candidate (%)	90%

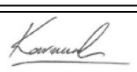
CO-AUTHORS

Name	Nature of Contribution
Kar Mun Chooi	Contributed to other sections of the manuscript and review the completed version
Daniel Cohen	Contributed to other sections of the manuscript and review the completed version
Robin M. MacDiarmid	Contributed to other sections of the manuscript and review the completed version

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Kar Mun Chooi		12/02/2018
Daniel Cohen		12/02/2018
Robin M. MacDiarmid		12/02/2018

Last updated: 19 October 2015

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies.	
Nature of contribution by PhD candidate	Arnaud Blouin: Research design, samples collection and preparation, data analysis and manuscript preparation
Extent of contribution by PhD candidate (%)	75%

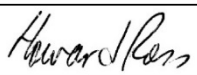
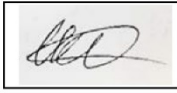
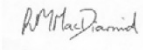
CO-AUTHORS

Name	Nature of Contribution
Howard A. Ross	Contributed to the phylogenetic analysis and comments to the manuscript
Jody Hobson-Peters	Constructive comments to the manuscript
Caitlin O'Brien	Contributed to the bead saturation experiment
Ben Warren	Assistance with the bioinformatics analysis
Robin M. MacDiarmid	Assistance in research design and review of the manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Howard A. Ross		2/03/2018
Jody Hobson-Peters		19/02/2018
Caitlin O'Brien		13/2/2018
Ben Warren		13/02/2018
Robin M. MacDiarmid		12/02/2018

Last updated: 19 October 2015

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. Improvement of double-stranded immunocapture for grapevine virus enrichment	
Nature of contribution by PhD candidate	Arnaud Blouin: Research design, samples collection and preparation, data analysis and manuscript preparation
Extent of contribution by PhD candidate (%)	90%

CO-AUTHORS

Name	Nature of Contribution
Kar Mun Chooi	Assistance in research design and review of the manuscript
Robin M. MacDiarmid	Assistance in research design and review of the manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Kar Mun Chooi		12/02/2018
Robin M. MacDiarmid		12/02/2018

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. Distinct isolates of grapevine rupestris vein feathering virus detected in Vitis vinifera in New Zealand	
Nature of contribution by PhD candidate	Arnaud Blouin: Research design, samples collection and preparation, data analysis and manuscript preparation
Extent of contribution by PhD candidate (%)	95%

CO-AUTHORS

Name	Nature of Contribution
Robin M. MacDiarmid	Assistance in research design and review of the manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Robin M. MacDiarmid		12/02/2018

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.	
Identification of a novel vitivirus from grapevines in New Zealand	
Nature of contribution by PhD candidate	Arnaud Blouin: Research design, samples collection and preparation, data analysis and manuscript preparation
Extent of contribution by PhD candidate (%)	80%

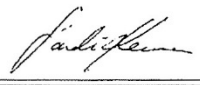
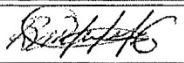
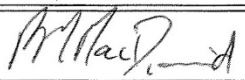
CO-AUTHORS

Name	Nature of Contribution
Sandi Keenan	sample collection and small RNA extraction
Kathryn R. Napier	small RNA data analysis (YABI Virus Surveillance and Diagnosis toolkit)
Roberto A. Barrero	small RNA data analysis (YABI Virus Surveillance and Diagnosis toolkit)
Robin M. MacDiarmid	Assistance in research design and review of the manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Sandi Keenan		20th September 2017
Kathryn R. Napier		20th September 2017
Roberto A. Barrero		20th September 2017
Robin M. MacDiarmid		20 September 2017

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand	
Nature of contribution by PhD candidate	Arnaud Blouin: Research design, samples collection and preparation, data analysis and manuscript preparation
Extent of contribution by PhD candidate (%)	75%

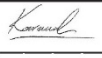
CO-AUTHORS

Name	Nature of Contribution
Kar Mun Chooi	Contributed in the survey (PCR)
Kathryn R. Napier	small RNA data analysis (YABI Virus Surveillance and Diagnosis toolkit)
Roberto A. Barrero	small RNA data analysis (YABI Virus Surveillance and Diagnosis toolkit)
Ben Warren	Assistance with the bioinformatics analysis
Robin M. MacDiarmid	Assistance in research design and review of the manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Kar Mun Chooi		12 February 2018
Kathryn R. Napier		12/02/2018
Roberto A. Barrero		12/02/2018
Ben Warren		13/02/2018
Robin M. MacDiarmid		12 February 2018

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

Co-Authorship Form

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work**. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work. Snapshots of a country's vineyard virome	
Nature of contribution by PhD candidate	Arnaud Blouin: Research design, samples collection and preparation, data analysis and manuscript preparation
Extent of contribution by PhD candidate (%)	70%

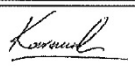
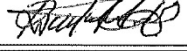
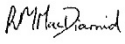
CO-AUTHORS

Name	Nature of Contribution
Kar Mun Chooi	Assistance in the RT-PCR validation of GVE-like and review of the manuscript
Ben Warren	Bioinformatics analysis
Roberto Barrero	small RNA data initial analysis (YABI Virus Surveillance and Diagnosis toolkit)
Kathryn Napier	small RNA data initial analysis (YABI Virus Surveillance and Diagnosis toolkit)
Robin M. MacDiarmid	Assistance in research design and review of the manuscript

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Kar Mun Chooi		1/03/2018
Ben Warren		1/03/2018
Roberto Barrero		2/03/2018
Kathryn Napier		
Robin M. MacDiarmid		1/03/2018

1 General introduction

Status of chapter:

The part 1.4.1 to 1.4.3 is an extract from an invited book chapter:

Blouin A. G., Chooi K. M., Cohen D., & MacDiarmid R. M. (2017a). Serological Methods for the Detection of Major Grapevine Viruses. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 409-429). Cham: Springer International Publishing.

The co-authorship forms are presented at the start of this thesis, after the acknowledgements.

This rest of this chapter has not been published and is not intended for publication.

1.1 New Zealand wine history

With approximately 1% of the global wine production from only 0.5% of the global vineyard acreage, New Zealand is a major contributor of quality wine worldwide (International Organisation of Vine and Wine, 2017). For instance, New Zealand recently made the top three wines imported into the United States of America (USA) by value (New Zealand Winegrowers, 2017a). However, its wine history is very short when compared with the millennium of wine production in the Middle East and European regions. The first record of grapevines in New Zealand is attributed to the missionary Samuel Marsden in Kerikeri (Northland) in 1819 (Danielmeier, 2008), while James Busby is credited with producing the first wine in 1832 from the same region, Waitangi (Danielmeier, 2008). Prior to being a consular representative in New Zealand and assisting William Hobson to draft the Treaty of Waitangi, James Busby was based in Australia where he is considered a founder of the Australian wine industry due, in part, to the impressive importation of *Vitis* cultivars he made mostly from Europe (Krake et

al., 1999). During the 19th century, several vineyards were planted in different regions of New Zealand but few were successful. The oldest existing winery in New Zealand, Mission Estate, was established in Hawke's bay by French missionaries in 1851 with the first commercial sale of wine recorded in 1870 (Howland, 2014).

At the end of the 19th century, the recently created Department of Agriculture requested the services of Romeo Bragato on loan from the Victorian government in order to stimulate the wine sector, with the aim of assessing the viticulture and wine making prospects of New Zealand. Bragato's report was extremely positive with a big prospect for viticulture in selected regions (especially Hawke's Bay, Wairarapa and Central Otago). He also encouraged the use of phylloxera resistant rootstocks following his identification of the pest (*Daktulosphaira vitifoliae*) in two vineyards in Auckland, and awareness of the massive impact that the insect had in European wine producing regions (Bragato, 1895). In 1892 the state Waerenga Experimental Station was established at Te Kauwhata, Waikato (North Island). It was then referred as the Te Kauwhata research station and in 1897, the first grapevines were planted (Dalley, 2008). In 1902 Bragato took the role of government viticulturist for the New Zealand Department of Agriculture and extended the size and content of the research station (Dalley, 2008, Danielmeier, 2008). In addition to his vision for the country's wine prospect, his awareness of the phylloxera threat, viticulture knowledge (Bragato, 1895) and his wine-making skills were recognised during the Franco-British wine exhibition in 1908 when five of the wines from the Te Kauwhata research station won gold medals (Howland, 2014). He is now celebrated every year by the New Zealand Winegrowers (NZW, the body representing the nation's grape and wine producers) who chose his name for their annual conference. However, the New Zealand wine industry remained very marginal for most of the 20th century. There was a low demand for wine as beer was more popular, and there was restriction on alcohol sale supported by prohibition advocates (Danielmeier, 2008) despite Bragato's plea:

"*It is now a widely-accepted fact, as proved by statistics, that in wine-producing countries drunkenness is less known and morality of the people stands at a higher level than in countries where wine is not the national beverage*." (Bragato, 1895).

After the Second World War, the wine production increased and with Gisborne and Hawke's Bay becoming the biggest wine producers. Marlborough's first vines were planted in 1973. However, the extension was too rapid and the higher tax on local wine, alongside abundant vintages, made the industry uneconomic. A 'vine-pull for money' scheme was run by the New Zealand government in the 1980s to reduce the acreage of the country's vineyards by 25% under the auspicious of responding to a wine excess. The Government's grape-vine extraction scheme removed a total of 1583 ha of grapevines nationwide (Parliamentary Debates, 1986)

However, with no requirements for re-use of that same land this action resulted in refreshment of vineyards around the country into promising cultivars such as Sauvignon blanc. By the mid to late 1980s, and with the first accolades for Sauvignon blanc, the New Zealand wine industry gained a strong reputation on the world stage. Since the 1990s, the industry has grown to become a substantial participant of the New Zealand economy; wine is now New Zealand's second biggest horticultural-based export, growing consistently to NZ$1.66 billion in 2017 (FreshFacts, 2016). The most planted cultivar in New Zealand is Sauvignon blanc (60% of the land), followed by Pinot noir (15%). Marlborough (South Island) is the biggest grape growing region representing 68% of the planted area, followed by Hawke's Bay (North Island) with 13%. Sauvignon blanc represents 86% of all exported New Zealand wine (New Zealand Winegrowers, 2017a).

1.2 A brief history of grapevine viruses in New Zealand

The old frescos from Pompei have been preserved for centuries by the ashes of the Mount Vesuvius eruption (79 AD). It is in some of these paintings that the oldest grapevine virus symptoms can be observed. Indeed, the distorted grapevine leaves depicted resemble those caused by the virus Grapevine fanleaf virus (GFLV) (Martelli, 2017). Vegetative propagation of grapevines has been a common practice for the last 6000 years (Reynolds, 2017), resulting in humans being the main vector of grapevine viruses with additional transmission routes superfluous to the point that some cultivars, such as the *Vitis vinifera* cv. Red globe, can be identified by the viruses they host.

If the records are accurate (Danielmeier, 2008), it is probable that the first grapevine viruses on New Zealand's shores landed in Samuel Marsden's vineyard, in 1819. Further evidence of the presence of virus disease in New Zealand grapevines can be traced to 1902; in the 10th report of the Department of Agriculture, Romeo Bragato described Cabernet Sauvignon that produced no fruit and were easily distinguished from fruit bearing vines by the early reddening of their leaves (Bragato, 1902). These descriptions could be attributed to the virus Grapevine leafroll-associated virus 3 (GLRaV-3). By the 1960s, such leafroll disease was known to be widespread (Chamberlain et al., 1970, McKissock, 1964). Subsequently, the virus has been recognised as the most detrimental in the country (Andrew et al., 2015, Bell, 2015, Bonfiglioli & Hoskins, 2006, Bonfiglioli et al., 2002, Charles et al., 2006). In the first official report that specifically describes leafroll disease, McKissock (1964) also reported GFLV, an important disease of vineyards at the time. The impact of GFLV was assessed by Chamberlain in 1970,

with high incidence reported in Auckland and Hawke's Bay, the two main wine regions at the time (Chamberlain et al., 1970). The alarming report by Mossop (1986), relating a widespread occurrence of GFLV and of the related virus Arabis mosaic virus (ArMV) in New Zealand, constitutes the last publication of viruses belonging to the genus *Nepovirus* in New Zealand grapevines. Ultimately, as a result of the absence of a vector in the country, and with a better control over planting material health, GFLV (and other nepoviruses) is now considered eradicated from commercial vineyards (MacDiarmid & Cohen, 2007). Since the mid-1980s, 17 other viruses have been reported in New Zealand grapevines (Veerakone et al., 2015 also in Table 1.1).

1.3 Grapevine viruses

By 2017 nearly 70 viruses had been reported worldwide in grapevines, the large majority with a single-stranded RNA genome, while those with a double-stranded RNA genome or a DNA genome virus have been recently reported (Martelli, 2017). A selection of the major viruses relevant to New Zealand are described below.

1.3.1 Grapevine leafroll-associated virus 3 (GLRaV-3)

GLRaV-3 is known to be the main contributor of leafroll disease and is by far the most damaging viral disease of grapevines in New Zealand (Andrew et al., 2015, Bonfiglioli & Hoskins, 2006) and worldwide (Burger et al., 2017).

1.3.1.1 Virus properties

GLRaV-3 is the type species of the genus *Ampelovirus*. The genus name is derived from the Greek "Ampelos" meaning grapevine. The virus particle is a flexuous filament of 1800 nm long and 12 nm wide. It has a monopartite single-stranded positive sense RNA genome of 18498 nt (Maree et al., 2013). The 5' end is believed to have a methylated nucleotide cap and the 3' end is not polyadenylated. The RNA has up to 13 Open Reading Frames (ORFs) (Table 1.1). The first ORF, named ORF1a, codes for a polyprotein containing the motifs of the leader papain-like protease, the methyltransferase the AlkB domain, and the helicase (Maree et al., 2013). The second ORF is called ORF1b because it is not initiated by a methionine start codon but instead is expressed via a +1 ribosomal frameshift mechanism to extend ORF1a

(Agranovsky et al., 1995). This +1 frameshift is observed in all the members of the family *Closteroviridae* (Martelli et al., 2012). ORF1b codes for the RNA-dependant RNA-polymerase (RdRp) that is required for the replication of viral RNA. ORF4 codes for a heat shock 70 homologue (HSP70h) that assembles onto the virion tail and has a role in cell-to-cell movement (Maree et al., 2013). ORF5 codes for a protein homologous to HSP70h, Heat shock 90 homologue (HSP90h), that, by analogy to research on *Beet yellows virus*, a relative in the genus *Closterovirus*, is also involved in the virion tail assembly and cell to cell movement (Alzhanova et al., 2007). ORF6 codes for the coat protein that represents the majority of the virion, whereas ORF7 codes for the minor coat protein that is a predominant component of the virion tail (Agranovsky et al., 1995). ORF10 encodes a protein with a viral suppressor of RNA silencing (VSR) function as demonstrated in studies undertaken in *Nicotiana benthamiana* (Gouveia et al., 2012). By contrast to these encoded proteins of known function, the ORFs 2, 3, 11 and 12 are very short (encoding proteins of 4 to 7 kDa) and of unknown function, with little homology across the GLRaV-3 groups and no homology to any other viruses. Yet, ORF3 p5 resembles similar hydrophobic transmembrane proteins found in other closteroviruses. It has been suggested that the hydrophobic p5 protein acts as a movement protein, where as a transmembrane protein it resides in the endoplasmic reticulum and is involved in virion movement from cell to cell (Dolja et al., 2006). The ORF2 is not essential and is not present in all GLRaV-3 strains. ORF8, 9 and 10 are specific to the *Ampelovirus* members, and code for potential VSRs as shown for the ORF10 (Maree et al., 2013).

1.3.1.2 Symptomatology

Some records from as early as the mid-19th century imply that GLRaV-3 was already problematic in Italy and France with the presence of premature leaf reddening (Martelli, 2014a). The symptoms of GLRaV-3 include downward leaf curl and the early onset of red leaves with green veins on red cultivars and leaf-rolling with leaf margins curling. The virus affects fruit maturity and even ripening of the bunch. The yield can be lowered by up to 66% (Over de Linden & Chamberlain, 1970). Visual identification of GLRaV-3 symptoms in red cultivars has shown to be very accurate and specific (Bell et al., 2017, Pietersen et al., 2017). In white cultivars, symptoms are less visible, but in some cases, leaf curling and uneven ripening can be visible in cultivars such as Pinot gris and Chardonnay (Charles et al., 2006, Rayapati et al., 2008). In the case of Sauvignon blanc, visual identification is impossible and the diagnostic relies on the serological method Enzyme-Linked Immunosorbent Assay (ELISA), or the molecular Reverse-Transcription Polymerase Chain reaction (RT-PCR).

When present, the symptoms are visible from the end of the summer and progressively more noticeable until post-harvest. The virus also does not express symptoms in most rootstock cultivars (Martelli, 2014a). The financial impact of GLRaV-3 is extensive and estimated to be US$40,000 for Cabernet franc growing in the Finger Lakes vineyard of New York (Atallah et al., 2012), and between $US29,902 to $226,405 per ha over the 25 year lifetime of a Californian vineyard (Ricketts et al., 2015). In New Zealand, the cost of GLRaV-3 was estimated to be between NZ$8172 per ha when a roguing regime was in place, to NZ$57,901 per ha with no control (Nimmo-Bell & Company Ltd, 2006). A whole block replant is considered in New Zealand when incidence of the virus exceeds 20% (Hoskins et al., 2011), while the threshold proposed in the USA is 25% (Almeida et al., 2013, Atallah et al., 2012, Ricketts et al., 2015).

1.3.1.3 Host range

Until recently GLRaV-3 was known to only infect *Vitis sp.*, but a recent report has added *N. benthamiana* as a host of GLRaV-3 (Prator et al., 2017). In their study, Prator and colleagues showed that the mealybug *Planococcus ficus* could transmit the virus from *V. vinifera* to *N. benthamina* and that the virus replicated to a high titre but was not mechanically transmissible from and to this host.

1.3.1.4 Genetic variability

Until the mid-2000s, the genetic diversity of GLRaV-3 was underestimated with the most divergent isolate described in 2005 being more than 90% nucleotide identity identical to the reference strain (Turturo et al., 2005). In 2007, Bonfiglioli and colleagues submitted to GenBank a sequence of an isolate (NZ-1) that only shared 70% nucleotide identity to the other members of GLRaV-3 known at the time (EF508151.1). The identification of this sequence had big implications on the diagnostics of the virus as most published RT-PCR protocols of the time would not detect this newly described strain. Over the years, additional sequence variabilities were identified within the conserved groups I to V (Fajardo et al., 2007, Jooste et al., 2010, Wang et al., 2011). Only in 2012 were three fully sequenced isolates described to demonstrate homology to the isolate NZ-1; i.e. clone 3 from California, in addition to GH11 and GH30 from South Africa (Bester et al., 2012, Seah et al., 2012). Later, isolate NZ2 from New Zealand was partially sequenced and revealed significant difference to the NZ-1 group (Chooi et al., 2013). It was proposed in an inclusive review of GLRaV-3 in 2013, that

the sequences could be classified into six groups, and the partial sequence of New Zealand NZ2 would remain un-grouped until further characterisation (Maree et al., 2013). The original five groups would remain the same as described by Gouveia (Gouveia et al., 2011), with group VI made of the isolates NZ-1 (New Zealand), clone 3 (USA), GH11 (South Africa) and GH30 (South Africa). It is important to note that there is less homology within the isolates of Group VI than within all the members of Group I–V. Recently, another isolate from South Africa (GH24) was sequenced that did not cluster with any of the previously described isolates (including members of Group VI or NZ2) (Maree et al., 2015). In that study, Maree and colleagues suggested a new classification of the species by assembling groups I–V in a supergroup A, group VI in a supergroup B, the newly describe GH24 as a new supergroup C and a supergroup D comprising all the non-classified strains represented by a collection of coat protein sequences submitted to GenBank that were identified during a survey in Portugal between 2007–2010) (Maree et al., 2015). This large genetic variation has had a significant consequence on diagnostic protocols over the years and the early diagnostics are only suitable to detect viruses within supergroup A (Chooi et al., 2013). Recent research has also shown that NZ-1 and NZ2 do not react strongly to the monoclonal antibodies MAbNY1.1 (Blouin et al., 2017a).

1.3.1.5 Management strategies

GLRaV-3 is transmitted by several species of mealybugs including *Pseudococcus ficus*; *P. maritimus*, *P. viburni*; *P. longispinus*, *P. calceolariae*, *P. comstocki*, *Planococcus citri*, *Phenacoccus aceris* and *Heliococcus bohemicus*. In addition, some soft scales have also been identified as a vector: *Pulvinaria vitis*, *Parthenolecanium corni*, *Ceroplastes rusci*, *Neopulvinaria innumerabilis*, *Coccus longulus*, *Parasaissetia nigra* and *Saissetia sp.* (reviewed in Almeida et al., 2013). The options for managing GLRaV-3 are limited since virus infected plants cannot be cured and the vectors are very common in most vineyards around the word. Additionally, the adjacent white cultivars, where the virus symptoms are often difficult to visually identify, can form a virus reservoir. Management strategies include a rigorous certification programme to ensure all plantings are free of the virus. Vector controls should be effective on mealybug nymphs, the most active stage for virus transmission, and should prevent the accumulation of large mealybug populations on grapevine parts (Sandanayaka et al., 2012). Mealybugs have been shown to survive in soil on remnant roots long after vine and stump removal and are an additional source of infection after replanting (Bell et al., 2009). As a result, it is important to remove as much of the roots as possible before replanting. Lastly,

a roguing programme based on removal of symptomatic vines every year has proven to be a successful strategy for the red cultivar, but it is a long and expensive process (Almeida et al., 2013, Bell, 2015, Pietersen et al., 2013, Pietersen et al., 2017).

1.3.1.6 Situation in New Zealand

As described above, GLRaV-3 is the most damaging virus to grapevines in New Zealand since the symptoms were first described by Bragato in 1902 (Andrew et al., 2015, Bell, 2015, Bonfiglioli & Hoskins, 2006, Bragato, 1902). The virus is widespread but is mostly problematic in the North Island and upper South Island red cultivars. The main production of the Marlborough region is Sauvignon blanc which is less impacted by the virus, with no visible symptoms and only minor changes in berry size and content (Montero et al., 2016). The main vectors of GLRaV-3 identified in New Zealand are *P. longispinus, P. calceolariae* and *P. viburni* (Charles et al., 2010). In 2005, NZW developed the Grafted Grapevine Standards (GGS) to make sure that planted material is GLRaV-3-free (New Zealand Winegrowers, 2017b). In addition, red cultivars are surveyed visually for the presence of the virus and infected vines are removed and replaced with certified GLRaV-3 virus free plants (Hoskins et al., 2011). Visual inspection, roguing and vector control has proven successful in reducing the level of virus infection from 11.8% to less than 3% in 3 years (Almeida et al., 2013, Bell, 2015). In regions where the red cultivars are dominant, some wineries have replaced their white cultivars in order to monitor the virus visually and remove the virus reservoir. In the Gimblett gravel (Hawke's bay) the winery Craggy Range spent more than NZ$1 M to replace the 8 ha of award-wining Chardonnay with Syrah (Gibb, 2014).

1.3.2 Grapevine leafroll-associated virus 1 (GLRaV-1)

1.3.2.1 Virus properties

In the same genus as GLRaV-3 (Ampelovirus), grapevine leafroll-associated virus 1 (GLRaV-1) was the first virus characterised from the Grapevine leafroll disease complex. The virus shares particle characteristics with GLRaV-3. The genome is 18,659 nt long and has 10 ORFs. It has the same structure as GLRaV-3 except for having two minor coat proteins (CPm1 and CPm2 encoded by ORFs 7 and 8), followed by only two ORFs. This is the only member of the *Closteroviridae* family to have one coat protein with two minor coat proteins (Martelli et al., 2012).

1.3.2.2 Symptomatology

It is hard to find symptom descriptions for GLRaV-1 alone as it is generally associated with other viruses such as GLRaV-3 to form the grapevine leafroll disease, although studies have reported that GLRaV-1 is responsible for significant leafroll symptoms (Martelli et al., 2012). A study showed that vines of *V. vinifera* cv. Nebbiolo infected with GLRaV-1, *Grapevine virus A* (GVA) and Grapevine rupestris stem pitting-associated virus (GRSPaV) produced the same quantity as a healthy vine with no difference in bunch number and number of clusters. Only the titratable acidity (TA) and the resveratrol content were significantly increased (Giribaldi et al., 2011). These results contradict previous observations on the *V. vinifera* cv. Savagnin rose where the addition of GLRaV-1 and GVA to a control infected with GRSPaV lowered the yield by 42–54% over 6 consecutive years. However, no difference in TA or soluble sugars was reported (Komar et al., 2010). These two examples illustrate the difficulty of assessing the biological impact of a virus because the results will be subject to multiple factors that include the number of viruses involved (both studies had co-infection of GVA and GRSPaV), the genetic variability of the viruses (those studied and their co-infectants), the grapevine cultivar genetics (scion and rootstock) and the environment (e.g. soil type and weather conditions).

1.3.2.3 Host range

GLRaV-1 was only known to infect *Vitis* species until a recent publication describing the detection of the virus from pomegranates (*Punica granatum L*) in Turkey (Çağlayan et al., 2016).

1.3.2.4 Genetic variability

There is a significant genetic variation between the GLRaV-1 sequences. They can be clustered into three groups that vary from each other by up to 20% nucleotide identity in the coat protein (Alabi et al., 2011).

1.3.2.5 Management strategies

GLRaV-1 is transmissible by mealybugs of the genera *Helicoccus*, *Phenacoccus* and *Planococcus,* and soft scale insects belonging to the genera *Pulvinaria*, *Neopulinaria* and *Parthenolecanium* (Le Maguet et al., 2013). *P. viburni, P. maritimus*, *P. comstocki* and P. *calceolariae* were later added to this list (Naidu et al., 2014). In the countries where GLRaV-

1 is problematic, it is detected alongside with GLRaV-3 and managed as one disease as described for GLRaV-3 (Pietersen et al., 2017).

1.3.2.6 Situation in New Zealand

Although the spread of GLRaV-1 was suspected by Bonfiglioli in New Zealand (Bonfiglioli et al., 2001), no further evidence has been reported since. Petersen and Charles showed that the common mealybugs found in New Zealand vineyards, *P. longispinus* and *P. calceolariae,* were not able to transmit GLRaV-1 under experimental conditions (Petersen & Charles, 1997), which is in agreement with field observations but inconsistent with what is described overseas for *P. calceolariae* (Naidu et al., 2014). Consequently, GLRaV-1 is relatively rare in New Zealand and has had limited impact. In 1997, it was found only in 3.5% of grapevines with leafroll disease symptoms (Petersen & Charles, 1997). However, in New Zealand the Chardonnay clone Mendoza is known to be infected with GLRaV-1 as the country's clonal material descends from the already infected clone imported from Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, by David Sheat in 1971 (Hoskins & Thorpe, 2010a).

1.3.2.7 Related viruses

Grapevine leafroll-associated virus 4 (GLRaV-4) is also an Ampelovirus. This species is an amalgamation of what was previously known as GLRaV-4, GLRaV-5, GLRaV-6 and GLRaV-9 based on their serological and biological relationships, and epidemiological characteristics (Martelli et al., 2012). GLRaV-4 has the smallest genome of the family *Closteroviridae* with only 13,700 nt and it codes for only seven ORFs (Ghanem-Sabanadzovic et al., 2012). GLRaV-4 appears to have a similar but milder symptomology than GLRaV-3 (Martelli et al., 2012). It is transmitted by *Pl. ficus* which is not recorded within New Zealand (Daane et al., 2012). In New Zealand, GLRaV-4 was previously reported as GLRaV-4, GLRaV-5 and GLRaV-9 (Habili et al., 2002, Pearson et al., 2006, Veerakone et al., 2015).

Grapevine leafroll-associated virus 7 (GLRaV-7) is a member of the newly recognised genus *Velarivirus,* family *Closteroviridae* (Adams et al., 2014, Al Rwahnih et al., 2012a). It is believed to be asymptomatic with no known vector (Al Rwahnih et al., 2017). GLRaV-7 is not known to occur in New Zealand.

1.3.3 Grapevine leafroll-associated virus 2 (GLRaV-2)

GLRaV-2 is a member of the genus *Closterovirus*, in the same family as GLRaV-1, and GLRaV-3 (Closteroviridae).

1.3.3.1 Virus properties

GLRaV-2 particles are filamentous and flexuous with a length of 1600 nm. Its genome is about 16,500 nts and it has nine ORFs. The ORF1a and 1b code for the protein responsible for the replication of the virus, then a quintuple gene block (ORF2- ORF6), is believed to be responsible of the intercellular movement, with the last genes (ORF7 and 8) possible VSRs. The coat protein is encoded by ORF5 and the minor coat protein by ORF4 (Angelini et al., 2017)

1.3.3.2 Symptomatology

Although responsible for causing leaf roll symptoms, GLRaV-2 is more notorious for causing graft incompatibility in some rootstocks, such as the strain GLRaV-2 RG originally detected in red globe cultivars that have no leafroll symptoms but are responsible for severe graft incompatibility (Martelli, 2014c). In Chardonnay, GLRaV-2 was shown to have a significant negative impact on yield (-22%), the most detrimental virus amongst grapevine virus B (GVB), GRSPaV, grapevine fleck virus (GFkV), and GLRaV-3 (Komar et al., 2007).

1.3.3.3 Host range

As opposed to the other grapevine leafroll viruses, GLRaV-2 is mechanically transmissible and can be transmitted to *N. benthamiana* (Angelini et al., 2017).

1.3.3.4 Genetic variability

There is a large genetic variability observed between the strains of GLRaV-2. The strains GLRaV-2 RG and GLRaV-2 Alphie (or BD) are believed to produce more severe graft incompatibility and are the main agents responsible for Rugose wood complex, expressed by malformation of the stem growth on susceptible indicator (Alkowni et al., 2011, Bonfiglioli et al., 2003). The nucleotide identity is lower than 75% between divergent strains and the strain

PV20 is the most divergent of all. There are at least six lineages of the virus, with clear evidence of serological variation (Alkowni et al., 2011, Angelini et al., 2017).

1.3.3.5 Management strategies

With no known natural vector, management of GLRaV-2 is based on planting GLRaV-2 free grapevines.

1.3.3.6 Situation in New Zealand

Only two different sources of GLRaV-2 have been reported in New Zealand: 1) the imported Bordeaux clones of Sauvignon blanc 316 and 317, released from quarantine in 1988 and propagated and distributed to the industry in 1992 (Hoskins & Thorpe, 2010b); and 2) the divergent strain Alfie identified in 2003 by Rod Bonfiglioli (Bonfiglioli et al., 2003) that is similar to the BD strain described since by Bertazzon and colleagues (2010). The detection of GLRaV-2 by Waite diagnostics in 2002 is likely to originate from the Bordeaux clones (Habili et al., 2002). Not documented is the presence of GLRaV-2 in New Zealand in Red globe grapevines planted for domestic use (Angelini et al., 2003).

1.3.4 Grapevine fanleaf virus (GFLV)

GFLV is a Nepovirus; a moniker standing for nematode-transmitted polyhedral virus. In the literature the fan leaf symptoms of this disease were described in 1841 and 1865 making it the oldest reported grapevine virus symptoms (Andret-Link et al., 2004, Oliver & Fuchs, 2011). The vector of GFLV is the nematode *Xiphinema index.*

1.3.4.1 Virus properties

The virion is polyhedral with 28 nm diameter. Like all Nepoviruses (family *Comoviridae*) the virus is bi-partite positive sense ssRNA. RNA1 is 7342 nt long with one single ORF coding for a polyprotein. RNA2 is 3774 nt long and also codes for a single polyprotein. The polyproteins are cleaved by the RNA1-encoded viral proteinase. Five mature proteins i.e. cleaved products are coded by RNA1: 1A (putative proteinase cofactor); $1B^{HEL}$ (putative helicase); $1C^{VPg}$ (virus genome linked protein); $1D^{Pro}$ (chymotrypsin-like cysteine proteinase); and $1E^{Pol}$ (RdRp). Three are encoded by RNA2: $2A^{HP}$ (homing protein); $2B^{MP}$ (movement protein); and $2C^{CP}$(coat protein) (Andret-Link et al., 2004).

1.3.4.2 Symptomatology

Symptoms of GFLV are extremely varied (Demangeat et al., 2005) but distorted and asymmetric leaves are characteristic of the virus (Chamberlain et al., 1970). In addition, sharp toothed margins (Andret-Link et al., 2004) and malformation of berries and cane (fasciation) are associated with GFLV (Liebenberg et al., 2009, Martelli, 2014b). Other forms of GFLV symptoms are a bright yellow mosaic that may affect the whole plant and are easy to diagnose (Martelli, 2014b). In general the symptoms tend to be more severe in early spring and fade during the summer (Martelli, 2014b). Fan leaf disease is considered one of the most devastating viral diseases of grapevine (Martelli, 2014b, Martelli & Boudon-Padieu, 2007, Olivier et al., 2010), if not the most for some countries (Andret-Link et al., 2004). It is widespread in most grape growing regions of the world and in most states of the USA apart from California. Europe is particularly affected by the virus, with two-thirds of French vineyards either moderately or severely affected by the virus (Demangeat et al., 2005). The crop loss can vary significantly (10 to >80%) (Andret-Link et al., 2004), with exceptional cases causing total loss of production for that season (Raski et al., 1983). The wine quality is also affected with poor berry set, smaller bunches and uneven ripening of berries (Andret-Link et al., 2004). Due to slow movement of the nematode vector, distribution of symptoms is usually patchy within a vineyard.

1.3.4.3 Host range

Although *Vitis* sp is the natural host for GFLV, the virus is also mechanically transmissible to 35 plant species across six families. The major hosts used in research laboratories include *Chenopodium quinoa*; *C. amaranticolor*; *N. occidentalis*; and *Gomphrena globosa* (Vigne et al., 2004).

1.3.4.4 Genetic variability

The GFLV population is genetically highly divergent. Genome plasticity was demonstrated by inoculation and several passages of the virus on the herbaceous host *C. quinoa*. The sequence was then compared between the original isolate (from grapevines) and the recovered (from *C. quinoa*), and a variation of 13% nt was observed (Naraghi-Arani et al., 2001). A population study of GFLV identified nine groups (A-I), with GFLV isolates in mixed infections and five GLFV recombinants out of 347 isolates (Vigne et al., 2004). Sokhandan-Bashir and Melcher

(2012) found that the GFLV genomes were highly correlated to their geographical location suggesting possible host adaptation and recombination.

1.3.4.5 Management strategies

Cross protection is a strategy available for the management of GFLV. The concept uses a mild stain of a virus to protect against a severe strain of the same virus. For GFLV this approach was trialled with mild strains GFLV-GHu (Huss et al., 1989) and the related virus ArMV strain ArMV-TA (Legin et al., 1993) in vineyards of *V. vinifera* cv. Gewurztraminer. After 8 years, the cross protection was proven effective especially with the GFLV-GHu mild strain. Unfortunately, the mild strain isolates chosen were shown to have a significant detrimental effect on fruit yield (-9–17%) but no effect on fruit quality (Komar et al., 2008).

In order to control for the vector, most vegetable systems incorporate crop rotation which is unfeasible in the case of the perennial crop grapevine. Leaving soil fallow for a few years before replanting is uneconomical especially since *X. index* can be detected more than 4 years after a fallow period (Martelli, 2014b). Soil fumigation was commonly used to reduce the impact of GFLV before replanting but this practice is being discontinued because of its high ecological impact (Martelli, 2014b). Biological control research is ongoing with some *Trichoderma* species showing good results in the laboratory to control *X. index* (Daragó et al., 2013). Rootstock resistance to the nematode is the best tool available to date. The locus XiR1 (*X. index* Resistance 1) derived from *V. arizonica* has been identified (Oliver & Fuchs, 2011).

1.3.4.6 Situation in New Zealand

The virus was so common in New Zealand in the 1970s that it was one of the most damaging viruses of grapevine (Chamberlain et al., 1970). However, the geographical isolation of New Zealand kept the nematode vector *X. index* from entering the country. This indicates that the virus was introduced into the country by infected material and the grafting/replanting of GFLV-infected plants was responsible for its spread. As a consequence, the virus was virtually eradicated by replanting and grafting non infected plants (MacDiarmid & Cohen, 2007). This GFLV example illustrates that without movement a virus is easy to manage. There is still anecdotal evidence of the virus in some old germplasm plants, and probably in some old table grapevine.

1.3.4.7 Related viruses

In the same genus (*Nepovirus*), only ArMV, *Tobacco ringspot virus* (TRSV), and *Tomato ringspot virus* (ToRSV) were reported in *Vitis* in New Zealand. These three viruses have different vectors to GFLV, *X. diversicaudatum* for ArMV, and *X. americanum* for TRSV and ToRSV. Like *X. index*, these nematodes are also absent from New Zealand have not been reported in New Zealand recently. Of the eleven additional nepoviruses recorded on grapevine, only Cherry leaf roll virus (CLRV) is known to be present in New Zealand on multiple hosts but not *Vitis* (Veerakone et al., 2015).

1.3.5 Grapevine rupestris stem pitting-associated virus (GRSPaV)

Of all the grapevine viruses described, GRSPaV is recognised as the most widespread virus of grapevine worldwide and is found in most *Vitis* species, including wine grapes, table grapes and rootstocks (Meng & Gonsalves, 2007). Generally, GRSPaV is a virus without important biological impact and it is one of the most difficult grapevine viruses to cure by thermotherapy (Gribaudo et al., 2006). Consequently, GRSPaV is accepted in most grapevine certification programmes, which can explain, in part, its wide distribution in commercial vineyards. Vegetative propagation is responsible for its distribution, and its ubiquity is evidence of the general lack of a negative impact on grapevines. However, this does not explain why the virus is also found in wild *Vitis* species, and isolated *V. vinifera* (Nolasco et al., 2006, Sevin et al., 2012).

1.3.5.1 Virus properties

The monopartite single stranded positive sense RNA virus is part of the genus *Foveavirus* in the family *Betaflexiviridae*. The reference sequence is 8725 nt long excluding the polyA tail. It has five ORFs. The first ORF is the replicase polyprotein with the motifs for metyltransferase, AlkB, Papain-like protease, RNA helicase, and RNA dependant-RNA polymerase. ORFs 2, 3 and 4 are relatively small (24, 13 and 8 kDa, respectively) and overlap to form the 'triple gene block' involved in movement of the virus. The last ORF codes for the coat protein (Meng et al., 1998).

1.3.5.2 Symptomatology

The ubiquity of GRSPaV makes it the prime suspect for novel diseases. It is associated with the disease Rupestris stem pitting (RSP), a disease of *V. rupestris* causing small pits on the stem that are only visible when the bark is removed. The virus has little known impact on the growth of most grapevine cultivars, although it may reduce vigour when a scion cultivar is grafted onto a rootstock of *V. rupestris* parentage (Alabi et al., 2010). GRSPaV is also associated with vein necrosis on the rootstock 'Richter 110' (Bouyahia et al., 2005). Additional diseases have been associated with GRSPaV, or a strain thereof. GRSPaV-SY strain was detected in Syrah vines showing decline and graft incompatibility in California (Lima et al., 2006). However, the diseases of Syrah are not associated with the same causal agent in different parts of the world. Lima et al. (2006) suggested that GRSPaV-SY could be responsible for the Syrah decline in California, and Beuve et al. (2013) showed the opposite in France. Meanwhile, GVA is associated with 'Syrah disease' in Australia and South Africa (Habili, 2013). All of these reports suggest that the Syrah cultivar is more sensitive to graft incompatibility, and a different causal agent or, in most cases, the coinfection of different viruses may result in a similar disease. A study of grapevine response to GRSPaV infection showed a moderate impact on plant physiology materialised in lower berry size and weight, a change in the transcript profiles with increased CO_2 fixation and a moderate reduction in the photosynthesis rate, as well as some defence mechanisms. This study also presented some similarities to other work looking at plant responses to water and salinity stresses (Gambino et al., 2012). In conclusion, the authors suggest the long coexistence between GRSPaV and the grapevine may have resulted in co-evolution of less severe viral infections. In addition, they also alert to the high degree of variability between virus and plant genomes.

1.3.5.3 Host range

GRSPaV is naturally restricted to *Vitis spp.* but can be, under certain conditions, mechanically inoculated onto herbaceous indicator (not between grapevine plants); it was detected in pollen and was shown to be transmitted by seed (Lima, 2006, Meng & Gonsalves, 2007, Meng et al., 2013, Rowhani et al., 2000).

1.3.5.4 Genetic variability

There is large genetic variability amongst the GRSPaV population. In 2006, Meng et al. (2006) suggested four genetically distinct groups: GRSPaV-1; GRSPaV-SG1; GRSPaV-BS and GRSPaV-

VS. Since this analysis more variant isolates have been sequenced: groups GRSPaV-SY (Lima et al., 2006); GRSPaV-PN (Lima et al., 2009), GRSPaV-MG (Morelli et al., 2011) and GRSPaV-MT (Terlizzi et al., 2011). The latest review available to date counted 15 groups (Meng & Rowhani, 2017).

1.3.5.5 Situation in New Zealand

Although GRSPaV is common in New Zealand, very little is known about the incidence and virus genetic variability in the country. GRSPaV was first reported in New Zealand in 1986 (Mossop, 1986). Grapevine vein necrosis disease was also reported based on the symptoms (Pearson et al., 2006)

1.3.6 Grapevine virus A (GVA)

Grapevine virus A (GVA) is the type species of the genus *Vitivirus* (Family *Betaflexiviridae*). Vitiviruses were first identified in grapevines but have since been detected in mint, Heracleum, Arracacha, Agave and kiwifruit (Adams et al., 2004, Blouin et al., 2012, Martelli et al., 1997, Tzanetakis et al., 2007).

1.3.6.1 Virus properties

GVA has flexuous filamentous particles of 800 nm. Its genome is single stranded positive sense monopartite RNA virus with 7349 nt and it encodes five ORFs. The replicase polyprotein is encoded from the first ORF, the ORF3 codes for the movement protein, the ORF4 for the coat protein and ORF5 for a VSR. The protein encoded by the ORF2 has no known function and is not present in the vitivirus detected in Agave (accession KY190215) (Adams et al., 2004).

1.3.6.2 Symptomatology

GVA is the putative agent for Grapevine Kober stem grooving; it is detected by graft indexing on the Kobber 5BB indicator (interspecific hybrid of *Vitis*) and is characterised by the grooving of the grafted wood (Martelli, 2014c). The GVA group of viruses is also the putative agent for the Shiraz disease in Australia and South Africa (Habili, 2013). Shiraz disease seems to mostly impact the cultivar Syrah; GVA is symptomless on the white grape cultivars. Shiraz disease symptoms consist of annual shoots that do not mature and leafroll-like symptoms on the

leaves along with their late fall from the vine compared with leaves from non-infected plants. The plant dies within years of the disease being evident (Goszczynski et al., 2008, Goszczynski & Habili, 2012, Minafra et al., 2017).

1.3.6.3 Host range

GVA is mechanically transmissible to *N. benthamiana*, *N. clevelandii, N. occidentalis, C. quinoa* and *C. amaranticolor.*

1.3.6.4 Genetic variability

Historically GVA was clustered into three groups, but with the constant addition of new sequences a new analysis revealed six groups based on the partial replicase gene, where the most sequences are available; many recombination events were observed (Alabi et al., 2014).

1.3.6.5 Management strategies

Since GVA shares the same mealybug vectors as GLRaV-1 and GLRaV-3 it can be managed alongside the leafroll diseases (Minafra et al., 2017). However, it was demonstrated that GVA is only acquired by the mealybug *Pl. ficus* in the presence of the GLRaV-3 in the donor plant; the ampelovirus is believed to be a 'helper virus' required by the vitivirus for its transmission. During the inoculation phase of dual virus carrying mealybug, only a small proportion of the plants are infected by GVA alone (2%). GLRaV-3 transmission alone is higher (24%) and a majority of plants get infected with both viruses (34%) or none at all (43%) (Blaisdell et al., 2012). Different mealybug vectors seem to have different ratios of virus transmission between the vitivirus and the ampeloviruses as was shown between *Pl. ficus* and the *Pl. citri* (Bertin et al., 2016). A recent publication showed that grapevines infected with GVA without the ampelovirus GLRaV-1 or GLRaV-3 is very rare, supporting the hypothesis of the 'helper virus' (Rowhani et al., 2018). However, the study also demonstrated additional benefits for the vitivirus as it was detected in higher concentration in the presence of the ampelovirus than in its absence. Similar concentration effects were observed with GVB and GLRaV-2. This difference of virus titre suggests that the benefits of a leafroll virus is occurring in the donor plant with better replication possibly as a result of strong VSR activity. It does not rule out the possibility of an interaction within the vector such as the one described between *Potato aucuba mosaic virus* (PAMV) and its 'helper-virus' Potyvirus and their transmission by aphids (Fereres & Raccah, 2015).

1.3.6.6 Situation in New Zealand

GVA is commonly found in New Zealand but no information is available about the genetic group it belongs to and incidence of the virus (Veerakone et al., 2015). The mealybug *P. longispinus* is common in New Zealand and is recognised as a vector of the virus.

1.3.6.7 Related viruses

The virus GVB has a similar morphology and genome to GVA, except for a lack of homology in ORF2 and ORF5. Some strains of GVB, such as GVB 935-1, are believed to be the responsible agent for Corky bark disease (Minafra et al., 2017). Corky bark is detected by biological indexing when grafting (chip budding) infected material onto the indicator LN33 (interspecific hybrid of *Vitis*); symptoms are the cracking of young shoots and swelling between the nodes. GVB is also vectored by mealybugs including *P. longispinus.* The virus was reported in New Zealand but not recently and appears to be rare (Mossop, 1986).

Grapevine virus D (GVD) is rare worldwide, poorly characterised and it has no known vector (Minafra et al., 2017). The homology between GVD and GVK, a virus recently described from Korea (Jo et al., 2017a), suggests that they are the same virus.

Grapevine virus E (GVE) was first discovered in Japan and then in South Africa (Coetzee et al., 2010a, Nakaune et al., 2008) where it is widespread. The mealybug *P. comstocki* is the only known vector to date from a donor plant also infected with GLRaV-3 (Nakaune et al., 2008).

Grapevine virus F (GVF) has been reported only in California, Tunisia and South Africa (Al Rwahnih et al., 2012b) and it was detected in multiple accessions from the United States Department of Agriculture, Agriculture Research Service (USDA-ARS) National Clonal Germplasm Repository collections in Davis, California that originated from multiple countries (Al Rwahnih et al., 2014). The virus may have been imported with the grapevine originally or transmitted within the collection. GVA, GVE and GVF were found to be widespread in a virus survey in South African vineyards (Jooste et al., 2015).

Grapevine virus C was removed from the virus list of grapevines as it was shown that the virus was a strain of GLRaV-2 (Minafra et al., 2017).

Four novel grapevine vitiviruses were reported in early 2018; Grapevine virus H was detected in Portugal (Candresse et al., 2018) and Grapevine virus J in California (Al Rwahnih, personal

communication), in addition to Grapevine virus G (GVG) and Grapevine virus I (GVI) identified during this study (Chapter 3).

1.3.7 The fleck-like viruses

The fleck-like viruses is a group made up of five evolutionary related viruses from the family *Tymoviridae*. Most of these were only fully sequenced recently using High Throughput Sequencing (HTS) technologies, and none are attributed to a major disease. The virus GFkV belongs to the genus *Maculavirus* and Grapevine red globe virus (GRGV) is a tentatively assigned to that same genus *Maculavirus.* Grapevine syrah virus-1 (GSyV-1) belongs to the genus *Marafivirus*; Grapevine asteroid mosaic-associated virus (GAMaV) and Grapevine rupestris vein feathering virus (GRVFV) are yet to be recognised in the same Marafivirus genus by the International Committee on Taxonomy of Viruses (ICTV). No insect vector of these viruses has been reported (Sabanadzovic et al., 2017). The increasing usage of HTS has revealed the high incidence of this group of viruses around the world. It was suggested that the fleck-like viruses could now be considered as stable components of the grapevine virome thus demonstrating their presence long before our comprehension of their existence (Saldarelli et al., 2017).

1.3.8 Additional viruses

The viruses described below are recent discoveries from overseas research, and have not been reported in New Zealand.

1.3.8.1 Grapevine Pinot gris virus (GPGV)

Grapevine Pinot gris virus (GPGV; genus *Trichovirus*, Family *Betaflexiviridae*) was recently identified in Italy. As its name implies, it was first detected in Pinot gris (Giampetruzzi et al., 2012). It can cause severe symptoms in some white cultivars (including Sauvignon blanc and Chardonnay), but also Pinot noir and table grapes. Symptoms include leaf chlorosis and deformation, stunting growth, low vigour and irregular ripening of the berries. A correlation between the virus strain, virus titre and symptoms has been identified (Bertazzon et al., 2016). The virus, spread by mites, appears to move at an alarming rate. Archived RNA samples extracted before 2005, showed that the virus was present in high frequency in some Eastern European countries compared to Western Europe. Subsequently, the virus incidence across

Europe was at least two fold higher from RNA prepared after 2010, including in Western Europe. This rapid spread is demonstrated by 27 out of 34 plants testing positive in 2014 compared to their 2002 status which was GPGV-free (Bertazzon et al., 2016). The virus has been reported in most wine production regions of Europe, South Korea, China, Canada, Turkey, USA and Australia (Giampetruzzi et al., 2012, Glasa et al., 2014, Morelli, 2014, Wu & Habili, 2017).

1.3.8.2 Grapevine vein clearing virus (GVCV)

Grapevine vein clearing virus (GVCV) is a Badnavirus from the family *Caulimoviridae* (Zhang et al., 2014). The virus was discovered in *V. vinifera* and a hybrid cultivar in the Midwest region of the USA. GVCV is associated with vein clearing and vine decline and the way the virus spreads suggests it has an insect vector that is not yet identified (Zhang et al., 2011). GVCV has not been identified outside of the USA. It was detected in a number of *Vitis* species and genetic analysis showed two distinct clades (Guo et al., 2014).

1.3.8.3 Grapevine roditis leaf discoloration-associated virus (GRLDaV)

GRLDaV is another virus from the genus *Badnavirus*. Although the virus was only sequenced in 2015 (Maliogka et al., 2015), the self-named disease with which it is associated was known for almost three decades (Rumbos & Avgelis, 1989). The original description of the virus was made in Greece (Maliogka et al., 2015), but it was detected in a single symptomless vine in Italy (Chiumenti et al., 2016) and in four plants in Turkey (Ulubaş Serçe et al., 2017).

1.3.8.4 Grapevine red blotch virus (GRBV) and other DNA viruses

Grapevine red blotch virus (GRBV) is another DNA virus known to infect grapevines. It is a member of the family *Geminiviridae* and forms the genus *Grablovirus* (Al Rwahnih et al., 2013, Krenz et al., 2012). It is associated with serious reddening and blotchiness of the leaves, a decrease of berry sugar accumulation and yield. The symptoms of GRBV can be mistaken for GLRaV-3 and the impact of the virus is comparable to leafroll disease. The economic losses attributed to GRBV have been estimated at up to US$68,548 per ha over the 25 productive years of a vineyard's life-span in the Napa County (Ricketts et al., 2016). GRBV is widespread in North America and has been reported in Canada, Switzerland and South Korea (Cieniewicz et al., 2017). Table grape accessions from various origins worldwide tested positive at the National USDA-ARD Clonal Germplasm Repository (Al Rwahnih et al., 2014, Sudarshana et al.,

2015). From this study, the significant degree of variability between isolates from Europe, Asia, the Middle East, Africa and South America held in the National USDA-ARD Clonal Germplasm Repository support the hypothesis that the virus is present in those countries but we know that the virus was present in California for a long enough time to develop this genetic variability. Indeed, a herbarium sample of *V. vinifera* collected in 1940 tested positive for GRBV (Al Rwahnih et al., 2015a). Moreover, the recognised vector of the virus (*Spissistilus festinus* (Hemiptera: Membracidae)) is native to south-eastern USA and is found in relatively high incidence in California (Bahder et al., 2016). More recently, while surveying for GRGaV, a second member of the genus *Grablovirus* was identified by rolling circle amplification (RCA) and cloning of non-cultivated grapevines. The virus was named Wild vitis virus 1 (WVV-1) and no visible symptoms were associated with its presence (Perry et al., 2018).

Another related virus from the same family was reported in 2016 and named Grapevine geminivirus A (GGVA). It was detected by HTS from two plants imported into the USA from South Korea (Al Rwahnih et al., 2016a). During the same HTS run, a novel virus from the genus *Fabavirus* (Family *Comovirinae*) was also detected and named Grapevine fabavirus from the same two plants (Al Rwahnih et al., 2016b).

Another DNA virus was detected from Brazil and was named Temperate fruit decay-associated virus (TFTDaV) because it was detected in apples, pears and grapevines. Its genome of 3.4 kb does not fit in any described family to date. It was only detected in one grapevine, two pears and one apple (Basso et al., 2015). Like WVV-1, TFTDaV was detected by RCA not using HTS.

The viruses described above were selected because of their impact on grapevines or their ubiquity worldwide. Other grapevine viruses have been detected but they are either rare occurrences or they are mostly opportunistic viruses, common in alternative hosts that on some rare occasions are detected in grapevines. This list of secondary grapevine viruses includes: potato virus X, alfalfa mosaic virus and cucumber mosaic virus (Martelli, 2017).

Table 1.1: A list of the major grapevine viruses detected in New Zealand and their characteristics prior to this study (Veerakone et al., 2015).

Name **Genus - Family**	**Acronym**	**Particle**	**Genome size** (in k nt)	**Number of ORFs***	**Vector**	**Major Symptoms**	**Alternative host**	**New Zealand**
Grapevine leafroll-associated virus 3 Ampelovirus - Closteroviridae	GLRaV-3	flexuous	18.5	13	mealybug	Leafroll	*Nicotiana benthamiana*	widespread
Grapevine leafroll-associated virus 1 Ampelovirus - Closteroviridae	GLRaV-1	flexuous	18.5	10	mealybug	Leafroll	*Punica granatum L*	rare - associated with Chardonnay Mendoza
Grapevine leafroll-associated virus 4 Ampelovirus - Closteroviridae	GLRaV-4	flexuous	13.6	7	mealybug	Mild leafroll	Not known	rare
Grapevine leafroll-associated virus 2 Closterovirus - Closteroviridae	GLRaV-2	flexuous	16.5	9	Not known	Graft incompatibility	*Nicotiana benthamiana*	associated with Sauvignon blanc import from France
Grapevine fanleaf virus Nepovirus - Comoviridae	GFLV	polyhedral	RNA1 7.3 RNA2 3.8	2	nematode	Leaf distortion	Multiple herbaceous	eradicated from commercial vineyards
Grapevine rupestris stem pitting-associated virus Foveavirus - Betaflexiviridae	GRSPaV	flexuous	8.7	5	seed	Mild	Multiple herbaceous	widespread
Grapevine virus A Vitivirus - Betaflexiviridae	GVA	flexuous	7.4	5	mealybug	Graft incompatibility	Multiple herbaceous	widespread
Grapevine fleck viruses Maculavirus - Tymoviridae	GFkV	Isometric	7.5	4	unknown	Not significant	Not known	present
Grapevine redglobe virus Tentative Maculavirus - Tymoviridae	GRGV	Isometric	6.7	3	unknown	Not significant	Not known	present
Grapevine rupestris vein feathering virus Tentative Marafivirus - Tymoviridae	GRVFV	Isometric	6.7	1	unknown	Not significant	Not known	present
Grapevine asteroid mosaic associated virus Tentative Marafivirus - Tymoviridae	GAMaV	Isometric	6.7	2	unknown	Not significant	Not known	present
Grapevine syrah virus-1 Marafivirus - Tymoviridae	GSyV-1	Isometric	6.5	2	unknown	Not significant	Not known	present

*ORFs: open reading frames

1.4 Diagnostics

With evidence of the detrimental effect of grapevine viruses, there has been a need for diagnostic tools that are sensitive, robust, timely and economical. Over the years different tools were developed to respond to these requirements but most failed in one of the parameters above. Biological indexing is a way to visualise viruses that may be latent in the host by grafting a bud of the assayed plant to a range of susceptible indicators. The assay has an important place in phytosanitary policies because it gives a visual confidence that if some pathogenic viruses are present they will give visible clues on these sensitive hosts. It can also detect diseases of unknown aetiology that cause necrotic union or necrosis distortion on 3309C or 110R *Vitis* indicators (Rowhani et al., 2017). However, in the quest for the best diagnostic tool, this method has a major drawback — time — as the assay can take up to 4 years.

Polymerase chain reaction (PCR) method responds faster to the presence or absence of a pathogen. It is sensitive, timely and is not overly expensive. The assumption when using PCR is that the nucleic acid extracted is representative of the plant; the time of sampling and tissue selected is critical. In addition, the detection by PCR relies on the design of the primer and the stability of the short primer binding site. This is particularly important for the diagnostics of pathogens with significant genetic variability as observed for grapevine RNA viruses.

Another method of diagnostic, ELISA, is heavily used for its robustness, price and speed. Since a major part of this thesis relates to serological assays, an extract from a book chapter I contributed to, is included here (1.4.1 to 1.4.3).

1.4.1 Serological methods for the detection of major grapevine viruses

Extracted from Blouin AG, Chooi KM, Cohen D, MacDiarmid RM (2017) In: Meng B, Martelli GP, Golino DA, Fuchs M (eds) Grapevine Viruses: Molecular Biology, Diagnostics and Management. Springer International Publishing, Cham, pp 409-429.

1.4.1.1 Introduction

The Enzyme-Linked Immunosorbent Assay (ELISA) is a test that uses antibody binding specificity to detect substances such as peptides, proteins, antibodies, and hormones. Its use

is now very common in plant virus detection, with antibodies specific to the coat protein of the target virus.

It was first described by Engvall and Perlmann (1971). Their new assay followed the principle of radio immuno assay, which was developed in 1960 (Yalow & Berson, 1960) with major modifications. Instead of measuring the antigen antibody reaction using radioactivity, the ELISA measured the reaction using the activity of an enzyme (alkaline phosphatase). Eliminating the use of a radioactive label transformed this immuno-assay into one that could be used in many diagnostic laboratories. In addition, a significant simplification was achieved by coating plastic with either the antigen or antibody instead of using cellulose particles as previously required. The introduction of the plastic medium reduced dramatically the number of washing steps required and excluded the centrifugation step. In 1974, the microtitre plate was introduced as a platform for ELISA against malaria (Voller et al., 1974). Introduction of the microtitre plate led to the small, economical and standardized form of the ELISA we know today. The rapid uptake of this new assay, combined with the universal platform, allowed further development of diagnostic laboratories' equipment, including multichannel pipettes, plate washer, plate reader and automated liquid handler, which greatly improved the throughput, simplicity and cost effectiveness of ELISA. In 1977, Clark and Adams used most of these developments to successfully detect and quantify plant viruses, i.e. arabis mosaic virus (ArMV) and plum pox virus (PPV), which led to the rapid uptake of companies to produce and commercialize ELISA reagents for plant virus diagnostics (Clark & Adams, 1977). Subsequently, ELISA has been the most popular assay in diagnostic laboratories for detection of human, animal and plant viral pathogens. Most viruses that cause serious diseases in plants can now be screened by ELISA.

For grapevines, commercial ELISA kits are available for the main viral pathogens, including the ampeloviruses grapevine leafroll virus-associated virus 1; -3 and -4 (GLRaV-1, -3, -4); the closterovirus grapevine leafroll virus-associated virus 2 (GLRaV-2); the nepoviruses ArMV and grapevine fanleaf virus (GFLV); the vitiviruses grapevine virus A and B (GVA and GVB); and the maculavirus grapevine fleck virus (GFkV).

1.4.1.2 Advantages and limitations of immuno-assays

ELISA provides either highly specific or broad detection capability depending on the antibody used. For example, the AP-conjugated antibodies for the detection of GFLV distributed by Bioreba detects all known GFLV strains. With the continual addition of sequences in GenBank,

genetic diversity of viruses is more fully understood and alignments of the coat protein sequences can be used to select a conserved amino acid (aa) sequence synthesized against which antibodies can be produced, as shown for GLRaV-1 (Esteves et al., 2013). A difficulty of this approach is that not all conserved sequences are in structurally accessible parts of the protein, explaining which is why some peptide MAb strategies fail. In order to increase the breadth of the ELISA across different species of viruses, it is possible to mix antibodies for several viruses. For example, Pietersen et al. (2013) mixed the antibodies against GLRaV-1 -2 and -3 for simultaneous multiplex detection. This approach has been used by several diagnostic companies (e.g. for the detection of GLRaV-1 and/or GLRaV-3, or GFLV and/or ArMV). However, a positive ELISA result does not indicate which one of the target viruses is present in the samples. Alternatively, despite fewer applications, it is technically possible to select MAbs that are specific to a virus strain, for instance to discriminate between GLRaV-4 strains (Gomez et al., 2015, Reynard et al., 2015).

The sensitivity of detection by ELISA depends on the affinity between the antibodies and antigen, the amount of conjugate-enzyme bound and the substrate used, but there is no amplification of the antigen itself. Therefore, molecular methods have proven to be more sensitive when compared directly with ELISA (Gambino & Gribaudo, 2006, Komínek & Bryxiová, 2004, Liebenberg et al., 2009) and they provide the opportunity to gain more information about the virus, i.e. sequence data. However, the lack of template amplification in ELISA means that the assay is less prone to contamination. In addition, ELISA is not affected by polysaccharides or polyphenols present in the sap that are often responsible for interfering with PCR reactions. These features contribute to the robustness of the ELISA when compared to the molecular assays.

By contrast to molecular-based methods, ELISA is technically less demanding as the sample preparation for ELISA is much simpler than for nucleic-acid based detection. The equipment required to run the test is very simple and relatively inexpensive. In addition, most reagents are affordable, with the antibodies remaining the major cost but their prices vary depending on type and supplier. A typical ELISA assay is completed within two days but a substantial part of this time involves incubation periods during which the laboratory operator can undertake other tasks.

A simple form of serological test is the lateral flow immunochromatographic assay that detects the presence of an antigen in an extremely short time (minutes) from the plant sap extract. Since these assays can be performed in the field, it removes the errors that can arise from mislabeling and handling. With this minimum point of care, the test is suited for rapid response

in the case of a new disease, and also a perfect tool during training of symptom identification. Although still rare in the grapevine virus field, there have been reports of their use (Liebenberg et al., 2009). Availability, price and sensitivity are the main challenges for this tool to become widely adopted.

When using sensitive antibodies, tissue from multiple plants can be combined into a composite test sample, thus reducing the cost while still testing large number of plants. This approach is most cost-effective when the incidence of virus is low and only a small number of samples needs to be re-analyzed to identify infected plants. This method is validated and accepted for the detection of GLRaV-3 in the New Zealand Grafted grapevine standards (New Zealand Winegrowers, 2017b).

In contrast to the development of a new PCR protocol, the development of a new ELISA protocol can be a slow and expensive process. The virus needs to be purified or its coat protein synthesized or expressed in a suitable cell culture system via recombinant DNA technology in full or partially. Purification of a specific grapevine virus can be challenging because, as is typical for many grapevine viruses, it may be in a mixed infection, as well as in a low titre, be phloem-limited and/or may be labile**.** ELISA kits are not yet available for the recently described viruses *Grapevine Pinot gris virus* (GPGV) and *Grapevine red blotch-associated virus* (GRBV). A GPGV PAb was described (Saldarelli et al., 2015) for use in western blotting, but it is not suitable for use in an ELISA format.

ELISA has decreased in popularity with the promise of more sensitive or specific molecular technologies, however these latter platforms also have inherent limitations such as the specificity of primers and problems of cross-contamination when handling large numbers of field samples. ELISA is reliable and sensitive and therefore continues to provide robust results.

1.4.1.3 Conclusions and future directions

Despite the increased demand for molecular tools, at a time where the cost of sequencing is in steady decline, alternative usage of antibodies are being developed. The Luminex xMAP is, for example, a variation to ELISA that improves the speed of diagnostic by the simultaneous detection of multiple viruses. It uses fluorescent microspheres (beads) as a support for antibodies (van der Vlugt et al., 2015). The beads used for one assay are labelled with a set ratio of two fluorochromes, the bead-address. By altering the ratio of the two fluorochromes, theoretically, up to 500 different beads can be distinguished. Individual bead addresses can be used for each assay and mixtures of bead addresses are used to simultaneously run assays

with multiple viral targets. The tests use a 96 well plate format and since the beads are magnetic, washing steps can be automated and provide very high stringency. At the end of the assay, samples are analyzed on a small flow cytometer (or LED-based image analyzer) that recognizes the individual bead-address. The total amount of fluorescence per bead-address is correlated to the titer of each virus for each sample. Beads can also be covalently coupled with oligonucleotides to make a molecular assay. Although xMAP has been available for many years (Vignali, 2000), its uptake by plant virology has been slow with only a handful of publications of multiplex detection of viruses by serological assay (Bergervoet et al., 2008, Charlermroj et al., 2013, Croft et al., 2008). Since this method can be used to assay for multiple viruses that infect a crop simultaneously in a single well, it provides a step-change in throughput that would be of significant benefit for routine detection in a systematic testing regime such as the quarantine testing for import/export, and the quality control of propagation material for vegetatively propagated crops such as grapes.

In parallel to the development of new assays, some progress have recently been made to design epitope-specific antibodies. Nanobodies, a novel class of antibodies with a single domain naturally produced by camelids were reported in 1993 (Hamers-Casterman et al., 1993). They present the advantage of being small with a high stability and high sensitivity. They can be produced in transgenic plants to a very high concentration (up to 30% of the total leaf protein in *Nicotiana benthamiana*) (Teh & Kavanagh, 2010). Once expressed in planta, the flexibility for storage and production is a great advantage for diagnostic use, as they can be stored as seed and sown upon request. A similar approach was taken by Cogotzi et al. (2009) who expressed a single-chain fragment variable (scFv) antibody specific to GLRaV-3 in *Escherichia coli*. Nanobodies, as well as the scFv, can be used in a similar way to MAbs produced from hybridoma cells in an ELISA format as shown in the medical lab by Zhu et al (2014). for the detection of influenza H5N1, but also for grapevine virus detection with the nanobody specific to GFLV (Ackerer et al., 2015) or the scFv raised against GLRaV-3 (Cogotzi et al., 2009). Another application of the nanobody is its direct expression within the host of a virus to generate a transgenic plant resistant to that virus (Ghannam et al., 2015). When applied to grapevine viruses, this method was demonstrated to protect against GFLV (Hemmer et al., 2015).

Most of the antibodies are produced to detect a single viral strain or species, but some antibodies can detect most species in a viral genus [e.g. Potyvirus (Jordan & Hammond, 1991, Richter et al., 1995)]. However, the broadest detection of viruses can be achieved with a monoclonal antibody that binds double-stranded (ds)RNA produced by most RNA viruses

during genome replication. These antibodies were first developed 40 years ago (Moffitt & Lister, 1975) and have been tested for ELISA (Aramburu & Moreno, 1994, Aramburu et al., 1991, Garcia-Luque et al., 1986, Powell, 1991, Schonborn et al., 1991). However, the high level of background made these antibodies unsuitable for ELISA, but their specificity and avidity to one type of nucleic acid (O'Brien et al., 2015) showed great fit for IC-PCR (Nolasco et al., 1993). Recently, with the possibility of the untargeted sequencing at low cost using next generation sequencing, the dsRNA antibodies were shown to be an efficient method to enrich for virus nucleic acid (Blouin et al., 2016).

To conclude, more than 45 years after its development by Engvall et al. (1971), ELISA still has a place of choice in the diagnostic laboratory. When asked why ELISA has not been completely replaced by more modern assays, Eva Engvall (2010) replied “Few assays are as simple as the ELISA and require so little in terms of automation and equipment. There is beauty in simplicity”. Alongside with ELISA, several classic serological assays remain essential to detect or characterize viral disease; thus, the management and research of grapevine viruses will continue to benefit from new developments of antibody-based methods for years to come.

1.5 Sequencing

Grapevine viruses have co-evolved with their host for centuries. It is not rare to observe more than 30% nucleotide genetic difference between two virus isolates from the same species. The association between the long-lived host and the virus, combined with the lack of proofreading ability of the viral replicase, has contributed to this huge genome alteration between virus isolates. This brings colossal challenges for the detection of these virus genotypes. The list of known viruses is constantly increasing and the conventional tools cannot keep up with the new species and new variants of known viruses added every year. One solution to detect the full range of viruses present in the plant is to use the untargeted HTS.

1.5.1 History of DNA sequencing

In 1977, 24 years after the publication of the DNA structure by Watson and Crick (1953), Sanger et al. (1977) reported a new method to determine a nucleotide sequence in DNA. This method revolutionised biological research and for the following 30 years (1977–2007) it was the backbone of a giant leap in modern genetics. The Sanger sequencing method was the major contributor of sequencing until the mid-2000s. Despite recent advance in DNA sequencing technologies, this method still has an important place in genome research and molecular ecology.

1.5.2 Sanger sequencing

A few months before the Sanger publication, Maxam and Gilbert (1977) described a new method for sequencing DNA. Their sequencing process involved radioactive labelling one end of fragmented DNA followed by specific cleavage of the DNA by chemical reagents according to their base. Finally, gel electrophoresis was used to separate the fragments according to their length.

Rather than using chemicals to cleave the DNA at a specific base, Sanger and colleagues described a method that involved synthesis and the introduction of chain termination. Dideoxynucleoside triphosphate (ddNTP) analogues can be incorporated instead of deoxynucleoside triphosphates (dNTPs). The ddNTPs have a normal 5' triphosphate group and can be incorporated to a DNA fragment but they lack the 3' hydroxyl residue and therefore cannot extend the chain further.

In order to run a complete analysis, the termination of the DNA chain should happen at a low ratio and randomly. Four separate, radiolabelled reactions are run simultaneously and a small competing amount of one dideoxynucleoside (ddNTP: ddATP, ddCTP, ddGTP and ddTTP) added to each reaction. Every time a ddNTP is incorporated, that strand's synthesis is terminated. The four reactions are then separated by size on adjacent lanes of a polyacrylamide gel and the radiolabel bands are visualised by autoradiography. The information is read and recorded manually. If the shortest fragment of the four gel lanes had ddATP added, the first base sequenced is an Adenine. If the next longest fragment had ddTTP added, the next base sequenced is a Thymine, and so on.

The general principal of the Sanger sequencing method was used widely and improved regularly. Four fluorescent dyes specific to each ddNTPs allowed faster automated base calling (Smith et al., 1986). This improvement made the process faster and permitted all four reactions in one tube. This was commercialised by Applied Biosystems, Inc. (ABI). PCR (Mullis, 1986) also greatly improved the targeted amplification of DNA template, especially when a low template amount was available. In 1999, the first equipment using capillary separation of DNA fragments instead of the slab gels were used (Mardis, 2013, Marsh et al., 1997).

Sanger sequencing is still a very important method today. It has the advantage of generating relatively long fragments of 500–1000 bp. It is extremely accurate (about 99.999%), but this method is costly at about US$2.3/kb (Yoder, 2014). In New Zealand this method is available commercially for NZ$5–10/kb.

Sanger sequencing was used to complete many genomes, including the bacteriophage ϕX174 (Sanger et al., 1977) and then the bacteriophage λ (Sanger et al., 1982); the bacteria *Haemophilus influenzae Rd* (Fleischmann et al., 1995), the yeast *Saccharomyces cerevisiae* (Goffeau et al., 1996), the bacteria *Escherichia coli* (Blattner, 1997), the worm *Caenorhabditis elegans (*C. elegans Sequencing Consortium*, 1998)*, the common fruit fly *Drosophila melanogaster* (Myers et al., 2000), the plant *Arabidopsis thaliana* (Arabidopsis-Genome-Initiative, 2000) and of course the biggest sequencing effort, the Human Genome (International HapMap, 2005). A draft genome was completed in 2000 but a more polished sequence covering 99% of the gene-containing regions with an accuracy of 99.999% was made available in 2004 (International Human Genome Sequencing Consortium, 2001, 2004).

The enormous effort put into the human project, and its success led to new investments into technologies to reduce the cost of genome sequencing. This resulted in new chemistry used for sequencing, engineering and instrument development leading to a dramatic increase in

sequencing throughput and a significant decrease in cost. All the technologies that emerged from this research have been called “Next generation sequencing” or NGS as opposed to First generation sequencing that includes the Sanger and Maxam and Gilbert sequencing. Second generation sequencing or NGS uses amplified target DNA compared to third generation sequencing also known as single molecule real time sequencing. As opposed to first generation, both the second and the third platforms sequence many thousands of DNA molecules at a time. The term NGS is now commonly replaced by HTS.

1.5.3 Second generation sequencing

Second and third generation sequencing yield several thousand sequence reads at one time. In order to increase the sequencing signal for each molecule, second generation sequencing technologies incorporate a DNA amplification step. A random fragmentation of high quality and high purity DNA is followed by a ligation with linkers specific to the sequencing platform used. The molecules are then immobilised on a support (also specific to each platform) and amplified to form distinct clones of a single fragment. The last step requires the detection of each nucleotide incorporated during the synthesis (sequencing by synthesis).

The advantage of HTS is the speed and the number of data generated, as well as the possibility of relative quantification. However, its disadvantages include the shorter read length and the lower quality (overcome by the amount of reads).

1.5.3.1 Roche 454 pyrosequencer

The first HTS platform, commercially available in 2005, was the 454 pyrosequencer by Roche diagnostic (originally 454 Life sciences Corporation founded by Jonathan M. Rothberg). For this platform the library is amplified by an emulsion PCR. The DNA is first fragmented and adaptors are ligated. The DNA is denatured and added to a PCR that includes beads covalently linked to the complementary sequence of the adaptor primer, dNTPs and polymerase. Oil is added and emulsified to generate numerous emulsions that contain the PCR reagents with one bead and one DNA fragment. Each emulsion bubble becomes a microreactor for the PCR (emulsion PCR or emPCR). After the amplification, the emulsion is broken and the beads are checked for quality (beads with no amplified DNA or with too much DNA are removed). The DNA is denatured and the reaction is placed onto a PicoTiterPlate, where only one bead can fit a well. The base of the PicoTiterPlate is transparent (glass) and a charge-coupled device (CCD) camera records the light across the plate. From the top, enzymes are loaded in the

form of smaller beads bearing two immobilised enzymes, ATP sulfurylase and luciferase. The top of the PicoTiterPlate acts as a flow cell through which the nucleotides are flowed, one at a time. As one nucleotide is incorporated, a pyrophosphate is released and converted to ATP catalysing the luciferase, resulting in a chemiluminescent flash that is recorded by the camera. The quantity of light recorded is proportional to the number of nucleotides incorporated. However, this step represents one of the limits of this chemistry; as more nucleotides are incorporated, the less likely it is these will be recorded accurately. Pyrosequencing does not accurately measure homopolymers greater than six. In order to calibrate the instrument, the sequence starts with the adaptor sequence ACTG.

Output: This platform yields long fragments (400–650 nt), that are easy to assemble. Due to the chemistry, the reads have variable lengths as each DNA molecule will be sequenced at a different pace; a sequence with many homopolymers will incorporate more nucleotides in one step. Inaccurate measurement of homopolymers produces insertion/deletions that mostly contribute to the 1% error rate of this sequencing method. Roche have commercialised several versions of the 454 sequencer machines but to date the most powerful (454 FLX +) yields about 650 million bases (MB) per run at a cost of about $US9.5/MB. The benchtop 454 GS junior titanium yielded 50 MB per run at a cost of about US$19.54/MB at the time Roche stopped the support of the 454 sequencers in 2016.

1.5.3.2 Illumina sequencing by reversible terminator dye

Like Sanger sequencing, Illumina, Inc. also uses dye termination, but unlike Sanger the blocking elements are reversible. The platform was commercially available in 2007 by Solexa, Inc. and was subsequently purchased by Illumina. The library preparation is different to the 454 method. DNA fragments are cleaved and size selected before specific adaptors are ligated. The physical support for the amplification is a glass flow cell composed of eight channels decorated with adaptors of complementary sequence to those ligated to the DNA. The flow cell uses both surfaces. The target single-stranded DNA is increased by bridge amplification; DNA is hybridised to the adaptors on the surfaces of the cell, amplified using the adaptor as a primer and then denatured (the original fragment is washed away). The amplified fragment is then ‘bridged’ to another adaptor at its 3’ end while still attached at its 5’ to the glass. A second cycle of annealing, extension and denaturation is completed (using the adaptor to prime the reaction). The fragments also anneal at the 3’ end while still ‘bridged’ to the 5’. After 35 cycles, the cloned fragments form a physical cluster on the cell ready to sequence.

Four nucleotides each carrying a different fluorescent label are presented at each cycle of synthesis from one end of the 'bridge' cluster. However, because the 3' hydroxyl position of the ribose sugar is blocked, only one nucleotide can be incorporated. The signal is then read (using different wavelengths from each nucleotide). After each cycle, the fluor group is chemically cleaved to stop the fluorescence and 3'-OH is chemically deblocked. This process (incorporation, recording, fluor cleavage and 3'-OH cleavage) is repeated for each nucleotide read (typically 150 times for a HiSeq). In order to start a second read from the other end of the molecule (paired end sequencing), the amplified fragment is denatured and washed off. The limited bridge amplification is restarted and the sequencing cycle is repeated with a reverse primer.

Output: Illumina sequencing is 10 times more accurate than the 454 sequencing (0.1% error rate). The main error types are substitutions. The read quality decreases with the length of the molecule due to the increase of noise. The noise can be a result of polymerases that are not synchronised within a cluster (phasing errors). This happens when a blocking group is not included (incorporation of two or more nucleotides in one cycle) or not cleaved (no incorporation). The noise can also be due to incompletely cleaved fluor groups. Illumina have developed many sequencers with various sequencing capabilities. To date, the two most extreme sequencers are the benchtop MiSeq Nano that can generate 0.3 GB per run with a read length of 150 nt (x 2 when paired end) in 17 h, and the HiSeq X that can yield up to 900 GB per run (read length of 150 nt x2) in 3 days (Table 1.2).

1.5.3.3 Ion Torrent semiconductor sequencing

Ion Torrent is owned by Life Technology and uses the same library preparation as Roche 454 Pyrosequencing; it was also developed by Jonathan M. Rothberg. The library is amplified by emPCR and the fragments are primed with sequencing primers and deposited in an Ion chip (one bead per well). Unlike Illumina or Roche 454 pyrosequencing, the signal recorded is not light but pH. The Ion Chip forms a flow cell with the nucleotides added one at a time, from the top surface, while the bottom of the wells is a hydrogen ion detector. The nucleotides are used are natives (non-modified) with each incorporation releasing a hydrogen ion from the 3'-OH. Similar to the Roche 454 pyrosequencing, nucleotides are added one at a time and the platform records if one or more are incorporated.

Output: Similarly to other companies, Ion Torrent commercialises several instruments. The lower throughput is the PGM314. It can yield 0.22 GB per run in 4 h with a read length of

about 400 nt. The largest sequencer is the Proton I generating 16 GB per run and providing a read length of about 200 nt (Table 1.2). In the same way as the Roche 454 pyrosequencing, the read length is variable and the errors are mostly insertions/deletions of homopolymers. The rate of error is estimated to be 1%.

1.5.3.4 SOLiD Sequencing by ligation

SOLiD was developed by Applied Biosystem and, like Ion Torrent, is now owned by Life Technology. The preparation of the library is similar to that of Roche 454 pyrosequencing and Ion Torrent, using the amplification by emPCR. The chemistry of SOLiD is very different to the other platforms since it uses sequencing by ligation (Sequencing by Oligo Ligation Detection). The amplified beads are covalently attached to the SOLiD flowcell and are read by adding runs of eight base probes. Each probe has a ligation site specific to the first two bases (complementary to the target) and six degenerate bases, which includes a cleavage site at the 5^{th} base and a fluorescent dye linked to the last base. Once a probe is linked (base +1 from the primer), the fluorescence is measured and the last two bases are cleaved (including the fluorescent dye) and washed. Then the next probe is linked at base +6 from the primer, and the third one at base + 11. In order to read the gaps between the ligation sites, each fragment must be sequenced several times with primers shifting position by one nucleotide. Since the target site is includes two nucleotides there are 16 distinct probes (2^4) but only four dyes are used. In order to resolve the sequence, each base needs to be read twice. To achieve the complete coverage of a DNA strand with two ligations to each base, each fragment needs to be read five times with primers starting a base apart.

Output: The number of ligation runs required in this methods results in a slow run of eight days with a short read length (55 x 2 paired end). The advantage of this ligation system is the accuracy of 0.06% error rate. The only equipment commercialised is the SOLiD 5500xl with 155 GB per read (Table 1.2).

1.5.4 Third generation sequencing

1.5.4.1 Single molecule real time sequencing (SMRT) with PacBio (Pacific Biosciences)

Pacific Biosciences of California, Inc., (PacBio) was the first provider of single molecule real time sequencing. In this method, no amplification of the DNA library is made. The DNA polymerase binds to a single-stranded DNA molecule and is then fixed inside the well of the

sequencing support. Zero Mode Waveguide (ZMW) is the commercial name of the micro wells on the chip. There are 150,000 ZMWs on a chip and they represent the world's smallest light detection volume of about 20 zeptoliter (10^{-21} L). Each ZMW is about 70 nm in diameter by 100 nm deep, approximately the size of virus particle and small enough to filter light and therefore reduce noise. All four nucleotides are supplied in the ZMVs simultaneously, each base being individually phospholinked. Each time a nucleotide is incorporated, the fluorescent label is excited, diffusing light signal for a few milliseconds. The light is detected by a sensitive detector based under the ZMW (through the glass bottom). After incorporation, the phosphate and its label are cleaved off and diffused away. The light is recorded on a movie and converted into base calls with quality metrics. The incorporation rate is about 1 to 3 bases per second.

Output: PacBio has only one instrument called RS that can yield about 45 MB per run. The main advantage is the long read length that is very variable with an average of 3000 nt and maximum of about 10,000 nt. This long read makes PacBio an ideal platform to close genomes that are incomplete due to large regions with repeats or high GC content. The run time is very short and the preparation of library does not need amplification and is therefore rapid. However, the common error is insertions at the very high rate of 15%. A way to decrease the error rate is to multiply the depth of reads, but because there is no amplification this is not easy as each molecule can be different and therefore there is no way to distinguish an error from a different sequence. Because the insertions are random, a coverage of 15 fold gives an accuracy of about 99%. One way to increase the true coverage is to circularise the DNA using the SMRT Bell linkers. In this way, the same fragment can be read several times, but this is only available for short fragments (250 nt) which therefore defeats the purpose of long read sequencers. The run time is 10–14 h at a cost of US$7–38 per MB. Another advantage of SMRT technology is that it takes advantage of recording the polymerase pulse in real time. The signal timing is modified by a subtle modification in the DNA such as a DNA methylation variant. An accurate call on base modification would make this platform a real advantage in epigenetic research.

Table 1.2: Comparison of the main High Throughput Sequencing (HTS) platforms. Run time, reads, yield and costs are listed after the molecular ecologist 2016 Next Generation Sequencing (NGS) field guide (http://www.molecularecologist.com/next-gen-table-2-2016/ [accessed 15 February 2018]).

Instrument	Amplification	Run time	M Reads /run	Bases / read	Mbp/run	cost/Gb (US$)
Applied Biosystems 3730 (capillary)	PCR, cloning	2 hrs.	0.000 096	650	62.400	$2,307,69 2.31
Illumina MiniSeq - Mid	bridgePCR	17 hrs.	8	300	2,400,000,	$229.17
Illumina MiniSeq - High	bridgePCR	7 hrs.	25	75	1,875,000	$426.67
Illumina MiniSeq - High	bridgePCR	13 hrs.	25	150	3,750,000,	$249.33
Illumina MiniSeq - High	bridgePCR	24 hrs.	25	300	7,500,000	$200.00
Illumina MiSeq v2 Nano	bridgePCR	17 hrs.	1	300	300,000	$1,866.67
Illumina MiSeq v2 Nano	bridgePCR	28 hrs.	1	500	500,000,	$1,360.00
Illumina MiSeq v2 Micro	bridgePCR	19 hrs.	4	300	1,200,000	$708.33
Illumina MiSeq v2	bridgePCR	4 hrs.	15	50	750,000	$1,060.00
Illumina MiSeq v2	bridgePCR	24 hrs.	15	300	4,500,000	$225.56
Illumina MiSeq v2	bridgePCR	39 hrs.	15	500	7,500,000	$151.33
Illumina MiSeq v3	bridgePCR	21 hrs.	25	150	3,750,000	$233.33
Illumina MiSeq v3	bridgePCR	56 hrs.	25	600	15,000,000	$102.00
Illumina NextSeq 500 - Mid v2	BridgePCR	15 hrs.	130	150	19,500,000	$52.82
Illumina NextSeq 500 - Mid v2	BridgePCR	26 hrs.	130	300	39,000,000	$42.31
Illumina NextSeq 500 - High v2	BridgePCR	11 hrs.	400	75	30,000,000	$46.00
Illumina NextSeq 500 - High v2	BridgePCR	18 hrs.	400	150	60,000,000	$44.17
Illumina NextSeq 500 - High v2	BridgePCR	29 hrs.	400	300	120,000,000	$35.33
Illumina HiSeq 2500 - rapid run	BridgePCR	10 hrs.	300	50	15,000,000	$95.33
Illumina HiSeq 2500 - rapid run	BridgePCR	40 hrs.	300	200	60,000,000	$55.33
Illumina HiSeq 2500 - rapid run v2	BridgePCR	10 hrs.	300	50	15,000,000	$95.33
Illumina HiSeq 2500 - rapid run v2	BridgePCR	27 hrs.	300	200	60,000,000	$55.33
Illumina HiSeq 2500 - rapid run v2	BridgePCR	60 hrs.	300	500	150,000,000	$37.77
Illumina HiSeq 2500 - high output v3	BridgePCR	2 days	1500	50	75,000,000	$83.20
Illumina HiSeq 2500 - high output v3	BridgePCR	11 days	1500	200	300,000,000	$52.87

Instrument	Amplification	Run time	M Reads / run	Bases / read	Mbp / run	cost / Gb (US$)
Illumina HiSeq 2500 - high output v4	BridgePCR	40 hrs.	2000	50	100,000,000	$62.40
Illumina HiSeq 2500 - high output v4	BridgePCR	6 days	2000	250	500,000,000	$28.82
Illumina HiSeq 4000	BridgePCR	1 day	2500	50	125,000,000	$48.08
Illumina HiSeq 4000	BridgePCR	2 days	2500	150	375,000,000	$29.36
Illumina HiSeq 4000	BridgePCR	3.5 days	2500	300	750,000,000	$20.53
Illumina HiSeq X - Five	BridgePCR	< 3 days	3000	300	900,000,000	$10.63
Illumina HiSeq X - Ten	BridgePCR	< 3 days	3000	300	900,000,000	$7.08
Ion Torrent – PGM 314 chip v2	emPCR	4 hrs.	0.55	400	220,000	$2,154.55
Ion Torrent – PGM 316 chip v2	emPCR	5 hrs.	3	400	1,200,000	$561.67
Ion Torrent – PGM 318 chip v2	emPCR	7 hrs.	5.5	400	2,200,000	$397.27
Ion Torrent - Proton I	emPCR	4 hrs.	80	200	16,000,000	$62.50
Ion Torrent - S5 520 chip	emPCR	4 hrs.	5	400	2,000,000	$476.50
Ion Torrent - S5 530 chip	emPCR	4 hrs.	20	400	8,000,000	$139.13
Ion Torrent - S5 540 chip	emPCR	2.5 hrs.	80	200	16,000,000	$79.69
Oxford Nanopore MinION (fast mode; high volume user)	None - SMS	varies	0.026	10000	260,000	$2,307.69
Oxford Nanopore MinION (std. speed; low volume user)	None - SMS	varies	0.6	10000	6,000,000	$166.67
Pacific Biosciences RS II	None - SMS	≤6 hrs.	0.055	12000	660,000	$303.03
Pacific Biosciences Sequel	None - SMS	≤6 hrs.	0.385	10000	3,850,000	$181.82
SOLiD – 5500 (PI)	emPCR	8 days	700	110	77,000,000	$79.23
SOLiD – 5500xl (4hq)	emPCR	8 days	1410	110	155,100,000	$67.72

1.5.5 Future of HTS

Since 2005, the price for DNA sequencing has decreased dramatically with the introduction new chemistry and equipment, but the rate of decrease has reached a plateau since 2011 and is now reducing at a similar rate to that prior to 2005 in accordance with the Moore's Law. The new sequencing platforms have reached their optimum and the US$1000 per human genome is within reach. In the near future, improvements are likely to be in library preparation, sequencing for difficult templates, e.g. extreme GC content, and longer reads, as

well as improvement of the sequencing of single molecule that requires no library amplification. New players, such as the Oxford nanopore, offer detection by current fluctuation of DNA sequence passing through a nanopore (1 nm) in a synthetic membrane that should result in faster and cheaper sequencing methods.

1.6 HTS application in grapevine virology

Plant viruses have small genomes ranging from 1 kb (the individual component of nanoviruses) to about 20 kb (members of the genus *Closterovirus*). Using a HTS approach may appear excessive, however over time many publications have highlighted the advantages of such technologies in research, and diagnostics. HTS is non-targeted and very sensitive and as a consequence it represents the most sensitive tool to detect all viruses present in a tissue. HTS is now used to solve diseases of unknown aetiology, and the rate of virus discovery has significantly accelerated over the past 10 years. Since 2009 and the first HTS-discovered grapevine virus (GSyV-1), 19 additional viruses have been reported, 10 since 2017. Only three of the last 19 viruses were sequenced without the assistance of HTS: WVV-1, TFTDaV, both with ssDNA genome, were detected by RCA not using HTS, and grapevine leafroll-associated virus 13 (GLRaV-13) genome obtained by Sanger sequencing (Table 1.3).

The use of HTS in plant virology often includes a step of enrichment of viral nucleic acid. Many researchers use a dsRNA purification as input. During replication, an RNA plant virus generates long dsRNA molecules (Figure 1.1); DNA viruses produce transcripts with overlapping terminii. Within a plant (or animal) this feature is unique to viruses, therefore dsRNA is a good enrichment step for virus detection. Adding an enrichment step to sequence the small genome of the viruses, with the use of individual tags to each sample enables the pooling of many samples from different plants in one sequencing run. This has been used for large ecological studies (Coetzee et al., 2010b, Roossinck et al., 2010). To date, dsRNA is the most common target for HTS in grapevine virology (Table 1.3).

In the infected plant, the dsRNA originating from virus replication is recognised by the defence mechanism and cleaved by Dicer-like proteins. The guide strand of the small RNA (21–24 mers) integrates into the RNA-induced silencing complex (RISC) to specifically bind and cleave the target viral RNA. Sequencing the small RNA is like investigating the rubbish bin after a robbery to find the intruders (Figure 1.1). The system has been used in HTS to detect viruses

(Barrero et al., 2017, Li et al., 2012) and for research purposes (Hu et al., 2011, Miozzi et al., 2013, Singh et al., 2012), and it is also commonly used in grapevine virology (Table 1.3).

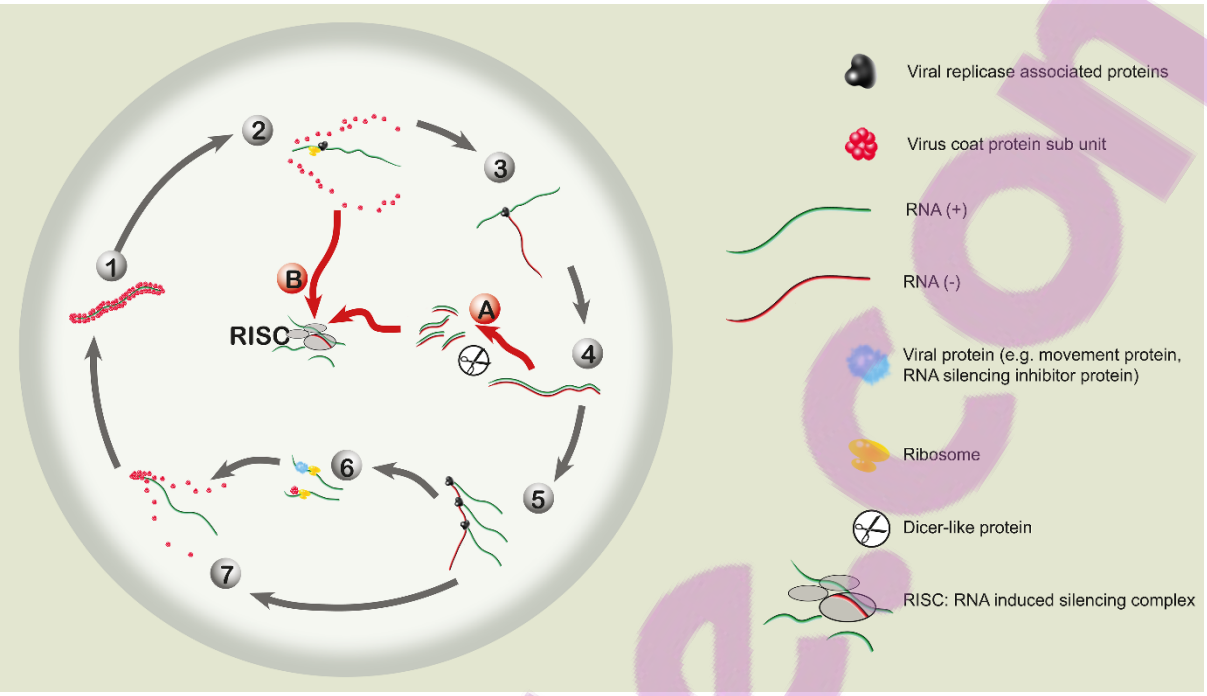


Figure 1.1: Typical replication cycle of single stranded RNA positive virus. The virus enters a plant cell to begin its replication cycle [1]. As the coat protein (CP) molecules are stripped away from the RNA [2], host ribosomes begin to translate the replicase and associated proteins. The replicase proteins generate the negative-sense (- sense) RNA template of the virus RNA [3]. This is used to generate both full-length positive-sense (+ sense) RNA [4] (making a full-length replicative form dsRNA), and the + sense subgenomic RNAs (sgRNAs) [5] that are used to express the viral proteins including the CP [6]. The + sense RNA is either encapsidated by the CP to form new virus particles [7] or moved to an adjacent cell for another round of replication. The full-length dsRNA [4] is targeted by the plant defence mechanism and cleaved with Dicer-like proteins [A]. The guide strand of the small RNA produced (21-24mers) integrates into the RNA-induced silencing complex (RISC). The RISC complex sequence-specifically binds and cleaves the target viral RNA [B].

The choice of the HTS platform for plant virology is selected depending on the aim of the sequencing run. The long reads and relative low throughput of Roche 454 pyrosequencing were advantageous for the analysis on a desktop computer. However, the chemistry used by Roche 454 pyrosequencing was the most expensive on the HTS market when support of the platform by Roche ceased in 2016. It was the platform used for the first publication of HTS in grapevine virology and the detection of GSyV-1 (Table 1.3). The high performing instruments such as the SOLiD 5500xl or Ion Proton I, are really useful for large genome research of plants

and animals, but not for virology. The third generation sequencing capable of very long reads may have some benefits in the long term to determine the true sequence of one virus molecule, but its inaccuracy prevents the distinction between error and real sequence variation in a situation of quasispecies.

Because of its accuracy, price and increased read length (up to 2 x 300 bp), the Illumina sequencing platforms are the most used sequencers in grapevine virology (Table 1.3). At least until the trivialization of the Oxford Nanopore.

Table 1.3: Impact of the High Throughput Sequencing (HTS) in grapevine virology since 2009.

Year	Known viruses detected	* New virus detected or - significant findings	Number of plant assayed	Sequencing Platform	Nucleic acid targeted	Country	Reference
2009	GRSPaV; GRVFV; GLRaV-4; HSVd; GYSVd; AGVd;	- First report of HTS for grapevine virus detection * Grapevine syrah virus-1 (Foveavirus: Quinvirinae: Betaflexiviridae: ssRNA+)	2	Life Sciences 454	Total RNA + dsRNA	USA	Al Rwahnih et al., 2009
2010	GRSPaV; GLRaV-3; GVA; GVE	- First large scale survey of grapevine virus	44 (pooled)	Illumina GA II	dsRNA	SA	Coetzee et al., 2010b
2010	GVE	- Second GVE genome	1	Illumina GA II	dsRNA	SA	Coetzee et al., 2010a
2010	GRSPaV; GFkV; GRGV;	- First report of virus detection by small RNA sequencing	1	Illumina Solexa platform	sRNA	Italy	Pantaleo et al., 2010
2011	multiple mycoviruses	- 26 putative fungal virus groups identified from five virus families from a single plant	1	Life Sciences 454	dsRNA	USA	Al Rwahnih et al., 2011
2011		* Grapevine vein clearing virus (Badnavirus; Caulimovirus; DNA)	2	Illumina GA II	sRNA	USA	Zhang et al., 2011
2012	GLRaV-7; GRSPaV; GRVFV; GSyV-1; GRGV	- First genome of GLRaV-7	1	Life Sciences 454 - Illumina GA II	dsRNA	USA	Al Rwahnih et al., 2012a
2012		* Grapevine virus F (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	1	Illumina GA II	dsRNA	USA	Al Rwahnih et al., 2012b
2012	GRSPaV; GVFV; GSyV-1	* Grapevine pinot gris virus (Trichovirus: Quinvirinae: Betaflexiviridae: ssRNA+)	2	Illumina GA II	sRNA	Italy	Giampetruzzi et al., 2012
2012		* Grapevine red blotch virus (Grablovirus; Geminiviridae; DNA)	3	Illumina GA IIx	dsRNA	USA	Al Rwahnih et al., 2013
2015	GVA; GVB; GLRaV-3	* Grapevine Roditis leaf discoloration-associated virus (Badnavirus: Caulimovirus: Badnavirus: DNA)	1	318 Ion Torrent PGM	sRNA	Italy	Maliogka et al., 2015
2015	GLRaV-1; GLRaV-2; GLRaV-3; GLRaV-4; GRSPaV; GVA; GVB; GFLV; GRBV; GFkV; GRVFV; GSyV-1; ToRSV; GCSV; GVE	- First comparison between HTS and Biological Indexing for the detection of grapevine virus	15	Illumina GA IIx	dsRNA	USA	Al Rwahnih et al, 2015
2015	GRSPaV	- The place of HTS in a certification scheme with regards to the virus status	20	Illumina HiScan SQ	sRNA	Italy	Giampetruzzi et al., 2015
2015	GRVFV; GAMaV; GRSPaV-1; GPGV; CMV; PVY; GLRaV-2	- Detections of viruses from publically available HTS data	11	HiSeq 1000	total RNA	Korea	Jo et al., 2015
2015		* Temperate fruit decay-associated virus (unclassified ssDNA viruses)		Sanger	RCA	Brazil	Basso et al., 2015
2016	GLRaV-1; GRSPaV; GPGV; GRVFV; GVA; GVB; GSyV-1	- HTS technology for a population survey	8	Illumina HiScan SQ	sRNA	Czech Republic	Eichmeier et al., 2016
2016	not mentioned	* Grapevine fabavirus (Fabavirus; Comovirinae; Secoviridae; Picornavirales	2	Illumina NextSeq 500	total RNA	USA	Al Rwahnih et al., 2016b

Year	Known viruses detected	* New virus detected or - significant findings	Number of plant assayed	Sequencing Platform	Nucleic acid targeted	Country	Reference
2016	Grapevine fabavirus	* Grapevine geminivirus A (Geminivirus; Geminiviridae; DNA)	2	Illumina NextSeq 500	Total NA	USA	Al Rwahnih et al., 2016a
2017	not mentioned	* Grapevine virus T (Foveavirus: Quinvirinae: Betaflexiviridae:ssRNA+)	1	Illumina GA IIx	Total RNA	Korea	Jo et al., 2017b
2017	Grapevine virus K	- First genome sequence of GVD (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	1	Illumina GA IIx	Total RNA	Korea	Yeonhwa Jo et al., 2017a
2017	GCSV; GVCV; GRGV; GLRaV-2; GLRaV-3; GVA; GRSPaV; GVA; GVB; GFkV; GRVFV	* Grapevine enamovirus-1 (Luteovirividea, enamovirus, ssRNA +)	17	HiSeq 2000	dsRNA	Brazil	Al Rwahnih et al. 2017
2017		* Wild Vitis virus 1 (Grablovirus; Geminiviridae; DNA)	9	Sanger	RCA	USA	Perry et al., 2018
2018	2 more viruses not presented	* Grapevine virus H (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	1	Illumina Hi-Seq3000	Total RNA	Portugal	Candresse et al, 2018
2018	Not available	* Grapevine virus J (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	Not available	Not available	Not available	USA	Al Rwahnih, pers com
2018	GVG	* Grapevine badnavirus 1 (Badnavirus; Caulimovirus; DNA)	4	Illumina HiSeq 4000	Total RNA	Czech Republic	Vončina & Almeida, 2017a
2018	GLRaV-2; GLRaV-3; GRSPaV; GRVFV; GRGV	* Grapevine virus G (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	1	Illumina HiSeq 2500	dsRNA	NZ	Blouin et al., 2018a*
2018	GLRaV-3; GVA; GRSPaV; GRVFV; GRGV; GVGV	* Grapevine virus I (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	1	Illumina HiSeq 2500	dsRNA	NZ	Blouin et al., 2018b*
2018	GAMaV; GFkV; GLRaV-1; GLRaV-2; GLRaV-3; GLRaV-4; GRGV; GRSPAV; GRVFV; GSyV-1; GVA; GVB; GVD; GVE-like; GVG; GVI, GGVA	* Grapevine virus E NZ variant (Vitivirus: Trivirinae: Betaflexiviridae: ssRNA+)	225	Illumina HiSeq 2500	dsRNA	NZ	This thesis*

*Study reported in this thesis

1.7 Aims and Objectives

The main aim of this thesis is to provide a deeper understanding of the virus population present in New Zealand grapevines.

More specifically, the objectives were to:

- Optimise new virus detection tools based on using HTS for grapevine virus identification; the tool should be practical, scalable and economical
- Use HTS to identify new viruses and understand their phylogenetic place in the virus kingdom
- Run a large scale screen of grapevines viruses in New Zealand's non-commercial and commercial plantings in order to provide the sector with a better understanding of the viruses already present in New Zealand, as well as offering new insights into those with the potential to cause problems in the future.

1.8 Chapter overview and publication of research

Chapter 1: Introduction

This chapter introduces grapevines and its viruses with specific emphasis on the New Zealand situation. The description of antibody based detection technologies for the detection of grapevine viruses formed a substantial part of an invited book chapter.

Blouin AG, Chooi KM, Cohen D, MacDiarmid RM 2017. Serological methods for the detection of major grapevine viruses. In: Meng B, Martelli G, Golino D, Fuchs M eds. Grapevine viruses: molecular biology, diagnostics and management. pp 409-429. Springer.

Chapter 2: Methodology

This chapter relates the development of a novel approach to extract double-stranded RNA in order to enrich for nucleic acids of a viral origin and sequence by HTS.

- This method development was published in the journal Molecular Ecology Resources:

Blouin A, Ross H, Peters H, O'Brien C, Warren B, MacDiarmid RM 2016. A new plant virus discovered by immunocapture of double stranded RNA; Assessment of a novel approach for viral metagenomics studies. Molecular Ecology Resources 16: 1255–1263.

- The method was then optimised to improve the viral dsRNA yield from recalcitrant hosts, including grapevines. The improved protocol was published in a conference proceedings and is described in the methods section of the survey manuscript (see Chapter IV).

Blouin AG, Chooi KC, MacDiarmid RM 2018. Improvement of double-stranded immunocapture for grapevine virus enrichment. Proceedings the 19th Congress of the International Council for the study of Virus and Virus-Like Diseases of the Grapevine (ICVG). April 2018.

Chapter 3: New viruses detected in New Zealand

The novel viruses discovered as part of this study are described in this chapter as well as their phylogenetic relationship with the related genus. This chapter comprises three publications that describe:

- The distinct strains of GRVFV in New Zealand.

Blouin AG, MacDiarmid RM 2017. Distinct strains of Grapevine rupestris vein feathering virus detected in Vitis vinifera in New Zealand. Plant Disease 101(12), 2156-2156.

- The identification of the new-to-science vitivirus GVG.

Blouin AG, Keenan S, Napier KR, Barrero RA, MacDiarmid RM 2017. Identification of a novel vitivirus from grapevines in New Zealand. Archives of virology. 163 (1), 281-284.

- The identification of a second new-to-science vitivirus, GVI, in co-infection with GVG.

Blouin AG, Chooi KM, Warren B, Napier KR, Barrero RA, MacDiarmid RM 2018. Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand. Archives of virology. doi: 10.1007/s00705-018-3738-5. [Epub ahead of print].

Chapter 4: Snapshots of a country's vineyard virome

This chapter summarises the results obtained from sequencing 225 grapevines from four different environments and discusses the significances of these results. The four groups are:

i. Four asymptomatic Sauvignon blanc vineyards in two wine regions.
ii. Plants from the high-health NZW germplasm.
iii. Plants from the low-health NZW germplasm.

iv. Plants collected in commercial vineyards from Hawke's Bay mostly showing symptoms.

At the time of thesis submission this chapter had not been published.

Chapter 5: General discussion

This chapter summarises the findings of the thesis, identifies strengths and weaknesses of the approaches used and discusses the direction for further research.

This chapter has not been published and is not intended for publication.

This project was presented in multiple issues of the New Zealand Winegrower magazines. These publications have been added in the Appendices 7.3. A publication describing a new virus referred to as 'Grapevine virus E-like' in this study was submitted with international collaborators to the journal Virus Gene and can be found in the Appendices 7.4.

2 Methodology

This chapter relates the development of a novel approach to extract double-stranded RNA, in order to enrich for nucleic acids of viral origin.

Status of chapter:

The first part of this chapter (2.1) has been published in the Molecular Ecology Resources Journal.

Blouin A. G., Ross H. A., Hobson-Peters J., O'Brien C. A., Warren B., & MacDiarmid R. (2016). A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies. *Molecular Ecology Resources, 16*(5), 1255-1263.

The second part of this chapter (2.2) has been submitted and accepted for the Proceedings the 19th Congress of the International Council for the study of Virus and Virus-Like Diseases of the Grapevine (ICVG), April 2018.

Blouin A. G., Chooi K. C., MacDiarmid R. (2018). Improvement of double-stranded immunocapture for grapevine virus enrichment. Paper presented at the Proceedings of the 18th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Chile.

The co-authorship forms are presented at the start of this thesis, after the acknowledgements.

2.1 A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies

2.1.1 Introduction

For rapid and efficient screening of large numbers of environmental, primary industry and medical samples for the presence of viruses, a virus nucleic acid enrichment step is required

prior to sequencing. Because nucleic acids are difficult to isolate from some plant tissue samples, the development of an effective virus detection assay would be an important contribution to the effectiveness of such screening. Many new viruses have been identified with the advance in sequencing technologies. To detect and sequence plant viruses using Next Generation Sequencing (NGS) platforms, four strategies have been adopted based on the replicative cycle of a virus. The first strategy involves the sequencing of total RNA and then using bioinformatics to distinguish the viral RNAs from the plant sequence (Seo et al., 2015, Verbeek et al., 2014). The second strategy is to sequence RNA extracted from virions purified from the sample (Tatineni et al., 2014, Thapa et al., 2012, Thapa et al., 2015), while the third method is to sequence the short, small interfering RNAs, some of which are derived from the plant antivirus defence mechanism (Giampetruzzi et al., 2012, Loconsole et al., 2012). Finally, the fourth method is to sequence the double-stranded RNA (dsRNA) produced during the replication of RNA viruses (Marais et al., 2015, Roossinck et al., 2010, Thapa et al., 2015). Each of these methods has its advantages and limitations, as reviewed by Massart et al. (2014).

Three-quarters of the plant viruses described to date have a positive-sense, single-stranded RNA genome (Hull, 2014). Most RNA viruses produce a dsRNA intermediate replicative form, and since viruses are the dominant source of long dsRNA molecules in plants, this is commonly exploited for enrichment of viral nucleic acid. A dsRNA extraction method described by Morris & Dodd in 1979 has since been adapted and used for plant virus discovery (Morris & Dodds, 1979). Recently, this traditional dsRNA enrichment method has been used for virus nucleic acid enrichment prior to NGS to diagnose viral infection in a plant (Al Rwahnih et al., 2012, Marais et al., 2014) but also in ecological studies (Coetzee et al., 2010b, Roossinck et al., 2010, Thapa et al., 2015). However, the traditional dsRNA extraction method based on chromatography on cellulose is a bottleneck for streamlined, large-scale analyses of viral populations as it is time-consuming and requires large amounts of sample tissue. A rapid and efficient method for dsRNA enrichment would enable very high throughput of samples and identification of virus sequences from environmental, primary industry and /or medical samples.

Moffitt and Lister (1975) made the first antibodies against dsRNA and developed a “simple and sensitive” serological test for dsRNA mycovirus after a phenol nucleic acid extraction from fungi. This was followed by several other publications of anti-dsRNA monoclonal antibodies (mAbs) being used in enzyme-linked immunosorbent assay (ELISA) (Aramburu & Moreno, 1994, Aramburu et al., 1991, Garcia-Luque et al., 1986, Powell, 1991, Schonborn et al., 1991).

Although Powell was “in some way disappointed” by the mAb that they had developed for use in ELISA, due to the high background it gave (Powell, 1991), the same mAb was found to be sensitive when used in an immunocapture reverse transcriptase polymerase chain reaction (IC-RT-PCR) by Nolasco et al. (1993). This was the first time the broad spectrum dsRNA binding method was combined with the sensitivity of PCR, enabling confirmation of the presence of eight viruses and one viroid from a range of plant species (Nolasco et al. 1993). The restrictions of this method for virus discovery were the specificity of the PCR primers and the prerequisite for knowing the viral target sequence prior to amplification. To our knowledge, only the mAb described by Schönborn et al. (1991) (J2) is available commercially and has been used in research, mostly to localise dsRNA in infected cells (Fontana et al., 2008, Triantafilou et al., 2012).

Recently, O’Brien et al. (2015) described two anti-dsRNA IgM mAbs (3G1 and 2G4) raised against virus purification products from an infected mosquito cell culture and subsequently referred to as Monoclonal Antibodies against Viral RNA Intermediates in Cells (MAVRIC). In the characterisation of the mAbs for ELISA and immunofluorescence, they showed the utility of these mAbs to detect arthropod-borne viruses from a diverse range of families.

In this study we have assessed a new protocol for the isolation of dsRNA in a pull-down experiment with the anti-dsRNA mAb 2G4 described by O’Brien et al. (2015). This protocol enables rapid dsRNA enrichment. Combined with a sensitive and sequence-independent PCR approach prior to NGS the utility of this protocol was confirmed by the detection of five viruses in three plant samples, including one virus which represents a new species of the Macluravirus genus.

2.1.2 Materials and methods

2.1.2.1 mAbs concentration

Two mAbs 2G4 and 3G1 were tested in the first assay. The specificity of 2G4 and 3G1 to long dsRNA (>40 and 50 nucleotides) and not to ssRNA or RNA:DNA hybrids was demonstrated (O’Brien et al., 2015). The optimal conditions for immunoprecipitation were established by determining the quantity of hybridoma supernatant required to saturate 10 µg of protein-L beads (Thermo Scientific Pierce™).

2.1.2.2 Preparation of synthetic dsRNA

A synthetic dsRNA molecule was generated from the sequence of the virus Actinidia virus X (AVX accession KC568202). The RNA from AVX-infected plant tissue was reverse transcribed and amplified by PCR using the primers T7 AVX 3048F (GAATTAATACGACTCACTATAGGGAGA CTGGTGATAGCCGTCAGTCC) and T7 AVX 5508R (GAATTAATACGACTCACTATAGGGA GATGGAAGTGATGTGACAGCCGA) using SuperScript® III One-Step RT-PCR System with Platinum® Taq DNA Polymerase (Thermo Fisher Scientific) following the manufacturer's protocol. From the amplicon, RNA was synthesised in both directions using a T7 RNA polymerase (Epicentre, Illumina) following the manufacturer's protocol and then treated with DNase I (Thermo Fisher Scientific) and RNAse H (Thermo Fisher Scientific) following the manufacturer's protocol. Quantification was performed with a Nanodrop spectrophotometer. The concentration of the 2460 bp molecule was measured at 175.8 ng/µL by Nanodrop, equivalent to about 6.78 x 1010 copies per µL as calculated by Endmemo.com (http://www.endmemo.com/bio/dnacopynum.php).

2.1.2.3 Pulldown assay for the recovery of synthetic dsRNA

Protein L Magnetic beads (Thermo Scientific Pierce™) were washed with TBST as per the manufacturer's protocol. Anti-dsRNA mAbs 2G4, 3G1 and negative IgM isotype control mAb 3D6 as hybridoma supernatant (Hobson-Peters et al., 2013, O'Brien et al., 2015) were coated onto the beads at saturating conditions as previously determined with 12.5 µg of beads used per sample. These were placed on a rotary tube suspension mixer for one hour at 37°C. The beads were then washed with TBS buffer as per manufacturer's recommendation and resuspended in 50 µL of PBS.

The synthetic dsRNA was diluted 1:100 in water to a final volume of 500 µL (dilution point A) and then twice further diluted 10-fold (dilution points B and C). For each dilution point, three replicates, each of 50 µL, were added to the coated beads (50 µL) along with 400 µL of PBST to a final volume of 500 µL. Samples were incubated for one 1.5 hours at 37°C on a rotary tube suspension mixer. The beads were recovered and washed twice with 400 µL TBST then resuspended in 50 µL of water.

Reverse transcription was performed using Superscript III. From the 50 µL resuspended beads, 1 µL was added to 5.3 µL of sterile deionised water and 0.2 µL of 10 µM primer AVX qR1 (AGTCTGGGTCAATGAGTTGTGGTG) and 0.5 mM of dNTPs. The tubes were heated at

65°C for 5 min then chilled on ice for one min. The Superscript III enzyme (100 U) was added with 2 µL of 5x buffer and 10 mM DTT, then incubated at 50°C for 30 min followed by a deactivation step of 15 min at 75°C. Real time PCR was performed using the primers AVX qR1 and AVX qF1 (TCTCTAATGCCGGGTTAAGTTTCCT) at a final concentration of 0.1 µM with 2 µL of HOT FIREPol® EvaGreen® HRM mix DNA polymerase (Solis BioDyne) and 5.8 µL of water using 2 µL of cDNA. The reaction was performed on an EcoStudy (Illumina). The thermal profile was 95°C for 15 min followed by 40 amplification cycles of 95°C for 15 sec, 60°C for 20 sec and 72°C for 20 sec then a melt curve of 95°C for 15 sec, 55°C for 15 sec and 95°C for 15 sec.

2.1.2.4 Plant sample preparation

A Māori potato leaf sample (*Solanum tuberosum* L.) known to be infected with Potato virus S (PVS) and Potato virus Y (PVY) by ELISA (data not shown), and a New Zealand native lily, rengarenga or rock lily (*Arthropodium cirratum* (Forst.f.) R.Br) showing virus-like symptoms (chlorotic streaking and mild necrosis) were collected from Lincoln (South Island, New Zealand). The leaf samples were freeze-dried. A sample of broad-leaved dock (*Rumex obtusifolius*) comprising eight leaves each from a distinct plant showing symptoms ranging from mild mottle to general necrosis was collected in the Auckland region (North Island, New Zealand) and stored frozen at -20°C since 2011 then freeze-dried in 2014; this pooled sample had tested positive for Cherry leaf roll virus (CLRV) by ELISA (data not shown).

For each sample, 100 mg leaf tissue was freeze-dried then ground to a fine powder in a plastic bag by a tissue homogeniser mounted on a drill press. Two mL of TBST were added to each sample and the extract was transferred to a microcentrifuge tube and spun at 17,000 g in a benchtop centrifuge for 5 min to clarify the extract.

2.1.2.5 Pulldown of dsRNA from plant specimens

Protein L Magnetic beads (Thermo Scientific Pierce™) were coated as described for the pulldown assay with the following modifications: 50 µg of beads saturated with 2G4 hybridoma supernatant per sample. These were placed on a rotary tube suspension mixer for 1 hour at 37°C. The beads were then washed with TBST buffer as per manufacturer's recommendation. The beads were resuspended in TBST, transferred into the plant extract and incubated for 1 hour at 37°C on a rotary tube suspension mixer. The beads were recovered and washed as previously and resuspended in 30 µL of sterile deionised water.

2.1.2.6 Amplification and barcoding

The RT reaction from dsRNA template was performed following the method described by Roossinck et al. (2010) using 6 µL of the resuspended beads-dsRNA, heated at 95°C for 2 min with 3 µL random primer (5' CCTTCGGATCCTCC N6-12 3') at a concentration of 20 µM (equimolar mix of N6, N8 and N12), and 9 µL of water and placed on ice for 1 min. The RT mix was then added as per the manufacturer's recommendation comprising 300 U of SuperScript III (Thermo Fisher Scientific), 0.5 mM dNTPs, 6 µL of 5x buffer, and 3 µL of DTT. Samples were incubated on ice for 15 min then at 50°C for 1 hour. After the incubation 0.75 µL of Ribonuclease A was then added (20 mg/mL Sigma) and the samples were incubated at room temperature for 15 min then 85°C for 2 min. Primers were removed using the Qiagen PCR purification column following the manufacturer's protocol. The samples were eluted in 30 µL 1:10 dilution elution buffer provided.

The PCR amplification protocol followed the method described by Roossinck (2010). Samples were amplified individually, using 17.5 µL of the RT product in a 50 µL reaction using Takara Prime script reagents (Clonetech) following the manufacturer's protocol for PCR (25 µL 2 x buffer, 2.5 µL of 50 mM MgCl2) with 5 µL of a single primer (10 uM, barcode-CCTTCGGATCCTCC). The barcodes used were CACGC, CAGAC, and CAGCG for the potato, rengarenga, and dock, respectively. The first cycle was 94°C for 5 min, 65°C for 30 sec, 72°C for 1 min and then 40 cycles consisting of 94°C for 15 sec, 45°C for 15 sec and 72°C for 5 sec and a final extension of 72 °C for 5 min. Samples were loaded on an 1% agarose gel and a section of gel between 500 and 800 bp was excised and purified using Zymoclean™ Gel DNA Recovery Kit following the manufacturer protocol.

Each sample was quantified using Qubit fluorimetric quantification (Thermo Fisher Scientific) and normalised to 5 ng per sample. Library preparation on the combined samples (TruSeq Nano DNA Library Prep Kit, Illumina) and sequencing was performed by Macrogen Korea on HiSeq 2000 (Illumina) using a partial run with 100 bp paired end.

2.1.2.7 Bioinformatics

All the bioinformatics analyses were performed in the Galaxy web-based platform running on a Plant & Food Research Limited internal server. The sequence data was de-multiplexed using barcode splitter (https://toolshed.g2.bx.psu.edu/view/devteam/fastx_barcode_splitter) with no mismatches or deletions allowed. The adaptors were removed and poor quality sequence

was removed using the Fastq-mcf tool with the following parameters: C=1000000; t% = 0.0001; l = 50; q = 25; and x=5). For each pool of data a *de novo* assembly was performed by Trinity on Galaxy (https://toolshed.g2.bx.psu.edu/view/bhaas/trinityrnaseq) using the two directions of sequence with no strand-specific library type, 1500 group pairs distance and 75 path reinforcement distance. The contigs obtained were compared online by BLAST to identify the virus present. Bowtie2 (version 2.2.4) was used to align each read-pair set separately (single-end mode), with alignment mode “end to end”, and the pre-set option “very sensitive” (Langmead & Salzberg, 2012). The pre-set option minimum alignment score to be considered mapped is -49.8 with the 82 nucleotides reads. The mapped reads were counted with Samtools Flagstat in Galaxy, and the numbers of reads mapped by the forward and reverse direction were added.

2.1.2.8 Alignment and tree construction

The virus tentatively named as Rumex virus Y was compared with members of Potyviridae on a partial coat protein fragment (most conserved region) at the aa level. The samples were trimmed to the Potyviridae coat protein motif (pfam00767) to a size ranging from 195 to 244 aa. The sequences were aligned using ClustalW (BLOSUM matrix, 10 gap opening, 0.1 gap extension). A maximum likelihood tree was built using PhyML v3 with the proportion of invariant sites estimated, empirical aa frequencies, variation in rates across sites modelled as a gamma distribution with 4 categories and 100 bootstrap replicates using the BLOSUM62 scoring matrix (Guindon et al., 2010, Guindon & Gascuel, 2003).

2.1.3 Results

2.1.3.1 The binding of dsRNA to mAb 2G4 in a pull-down assay is not affected by the concentration of dsRNA

Initial studies were performed to investigate the adaption of the anti-dsRNA mAbs to a pull-down assay platform. The saturation of the beads with the mAbs was equivalent to approximately 7 µg of purified antibody per 10 µg of beads (Appendix Figure 7.1). At this saturating concentration, more mAb 2G4 could be eluted from the beads than mAb 3G1 indicating that more of the former mAb was actually bound to the beads when saturated. To assess the binding capability of the mAb-saturated beads to dsRNA, and its direct detection by RT-PCR, dsRNA obtained by transcription of a 2460 nucleotide PCR product originating

from the ORF1 of the virus Actinidia virus X (AVX, GenBank KC568202) was recovered by mAb 2G4 and 3G1 in a pull-down assay using magnetic Protein L beads as assessed by reverse transcriptase quantitative PCR (RT-qPCR) (Figure 2.1). The non-dsRNA binding mAb 3D6 (isotype-matched control) was used as a negative control and had only one very late Ct value (>38) out of three biological replicates at the highest concentration confirming the dsRNA recovery observed with 2G4 and 3G1 treatments was due to the binding to the antibodies and not to non-specific binding to the beads. Both mAbs showed similar binding performance to the dsRNA solution.

2.1.3.2 NGS results from plant extracts

The efficiency of the viral enrichment with the anti-dsRNA mAb-saturated beads from virus-infected plants was measured by sequencing the recovered dsRNA. After the read deconvolution, the three plant host samples resulted in a similar number of reads: 127,898 for the potato sample, 198,597 for the rengarenga sample and 154,027 reads for the dock sample. From the total sequencing data, 95% of bases had the minimum quality score of 20 (Q20), and 87% of Q30.

From the Māori potato sample, 125,426 reads passed the quality control (98%). *De novo* assembly with Trinity yielded 96 contigs longer than 200 nucleotides. The two longest contigs were 8,469 and 5,015 nucleotides and both matched the virus Potato virus S (PVS) by nucleotide BLAST; however, the two sequences only shared 76% nucleotide identity across their common region (ORF1). In order to compare the two contigs, we examined the common 5015 nucleotides region in ORF1. The first contig matched PVS isolate WaDef-US (GenBank FJ813512) and isolate 09.369 (GenBank LN851191) equally with 98% identity both for the nucleotide and aa. When the isolate 09.369 was used as a reference for mapping using Bowtie2, 52,216 reads assembled (42% of the total number of reads after QC) forming a 8,452 nucleotides consensus sequence named NZ-O (GenBank KU058656) that is 98% identical to PVS isolates WaDef-US and 09.369. When compared with the isolate WaDef-US, NZ-O covers 99.6% of the genome with only the extremities not sequenced (33 nucleotides in total across both the 5' UTR and the 3' UTR). When using NZ-O as a reference it mapped to 53,909 reads (43% of the total number of reads after QC, Table 2.1).

The second contig matched the PVS isolate BB-AND (GenBank JQ647830) with 96% nucleotide and aa identity. When used as a reference the isolate BB-AND mapped 29,576 reads (24% of the total number of reads after QC) forming a consensus sequence called NZ-A of 8,390

nucleotides (GenBank KU058657). The isolate BB-AND was the closest match to NZ-A with 96% identical nucleotides and 97% identical aa. When compared with the isolate BB-AND genome, only the extremities were not sequenced representing a total of 129 nucleotides across both the 3' and 5' UTRs. When used as a reference, NZ-A was mapped by 35,983 reads (29% of the total number of reads after QC, Table 2.1). When compared, NZ-O and NZ-A shared only 80% nucleotide identity, confirming that they are different strains of PVS.

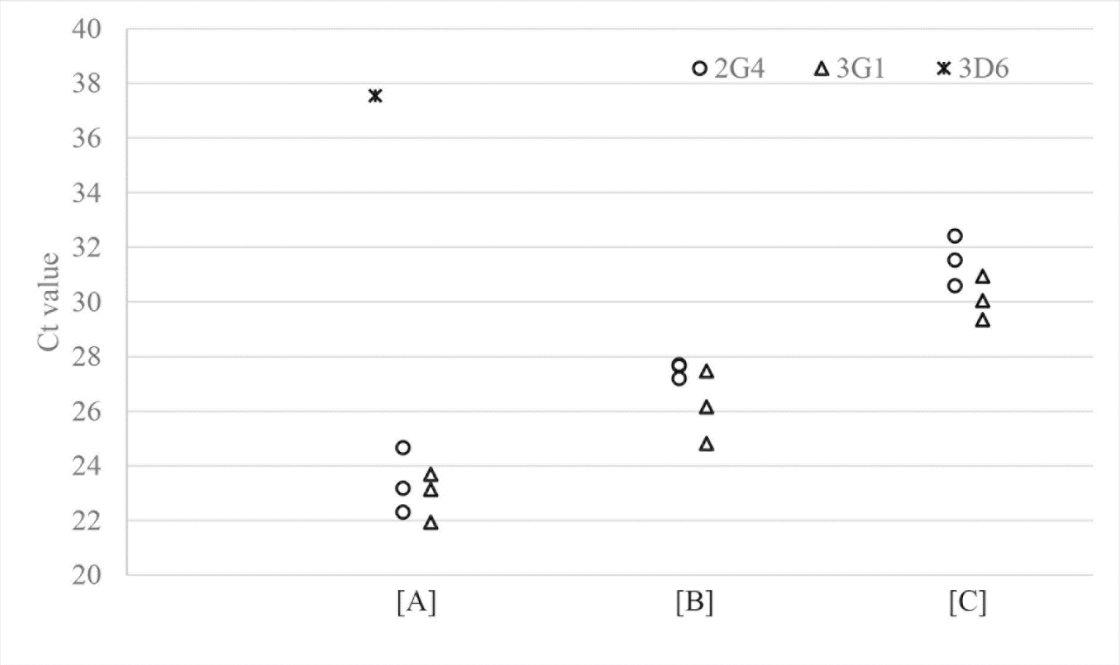


Figure 2.1: dsRNA recovered using three different antibodies. Using protein L magnetic beads and the mAbs 2G4, 3G1 or 3D6 hybridoma supernatant. Level of dsRNA measured by RT-qPCR. A tenfold dsRNA dilution is presented in the horizontal axis (point A, B and C), and the Ct value of the qPCR in the vertical axis.

In addition to PVS, 15 contigs were found to match Potato virus Y (PVY) with length varying between 208 and 1491 nucleotides that match PVY isolate N (GenBank D00441). When PVY isolate N was used as a reference, 3,405 reads mapped (3% of the total number of reads after QC, Table 2.1).

From the rengarenga sample, 193,322 reads passed the quality control (97%). The *de novo* assembly yielded 113 contigs longer than 200 nucleotides. The longest contig (6,241 nucleotides) as well as four additional contigs shared 99% nucleotide identity with the NZ290 isolate of Turnip mosaic virus (TuMV; GenBank AB093612). When this TuMV accession was

used as a reference it mapped 56,738 reads (29% of the total number of reads after QC). The consensus sequence obtained (GenBank KU053508) was at 9,787 nucleotides long and shared 99.5% nucleotide identity with TuMV isolate NZ290.

From the dock sample, 149,708 reads passed the quality control (97%). In total, 127 contigs larger than 200 nucleotides were obtained from *de novo* assembly. The longest contig (6,496 nucleotides) did not match any known sequence in the GenBank, but its translated sequence shared 52% aa identity with the polyprotein of the virus Chinese yam necrosis mosaic virus (CYNMC GenBank NC_018455). Another translated contig of 1,727 nucleotides also matched the same virus and these two contigs overlapped by 65 nucleotides (with three mismatches). When using the assembled contig as a reference, 58,305 reads mapped (39% of the total number of reads after QC, Table 2.1). The consensus sequence formed is 8,174 nucleotides with 11 ambiguities. This CYNMC-like virus is further characterised below as a novel Rumex virus. Out of the remaining contigs, seven matched CLRV sequences with the longest 685 nucleotides long. When mapped on the CLRV genome 441, 126 reads mapped the RNA1 (GenBank KC937025) and 244 mapped the RNA2 (GenBank KC937030).

2.1.3.3 The viral sequence detected in dock represents a new Macluravirus species

The novel Rumex virus is a monopartite RNA virus of 8,174 nucleotides extending from the 5' UTR (154 nucleotides) to the 3' UTR (208 nucleotides) that encodes one single polyprotein of 2,603 aa. The closest relative to this virus is CYNMC with only 52% aa identity over the coding region. The new virus genome encodes the Potyviridae motifs listed in the NCBI's conserved motif database (Marchler-Bauer et al. 2015); a peptidase C4; a helicase (GSGKSX3P and DEXH); an RNA dependant RNA polymerase (RdRp) motif; and the Potyviridae coat protein (CP) motifs. It also has the conserved motif GA6 responsible for a polymerase slippage resulting in an additional ORF (namely PIPO) and the production of the P3N-PIPO protein (Chung et al. 2008, Rodamilans et al., 2015, Olspert et al., 2015). A phylogenetic tree based on the CP encoded by members of the Potyviridae family shows that the putatively named Rumex virus Y (RVY) is in the Macluravirus genus (Figure 2.2). A phylogenetic tree estimated using the full length polyprotein indicates similar clustering (Figure 2.3).

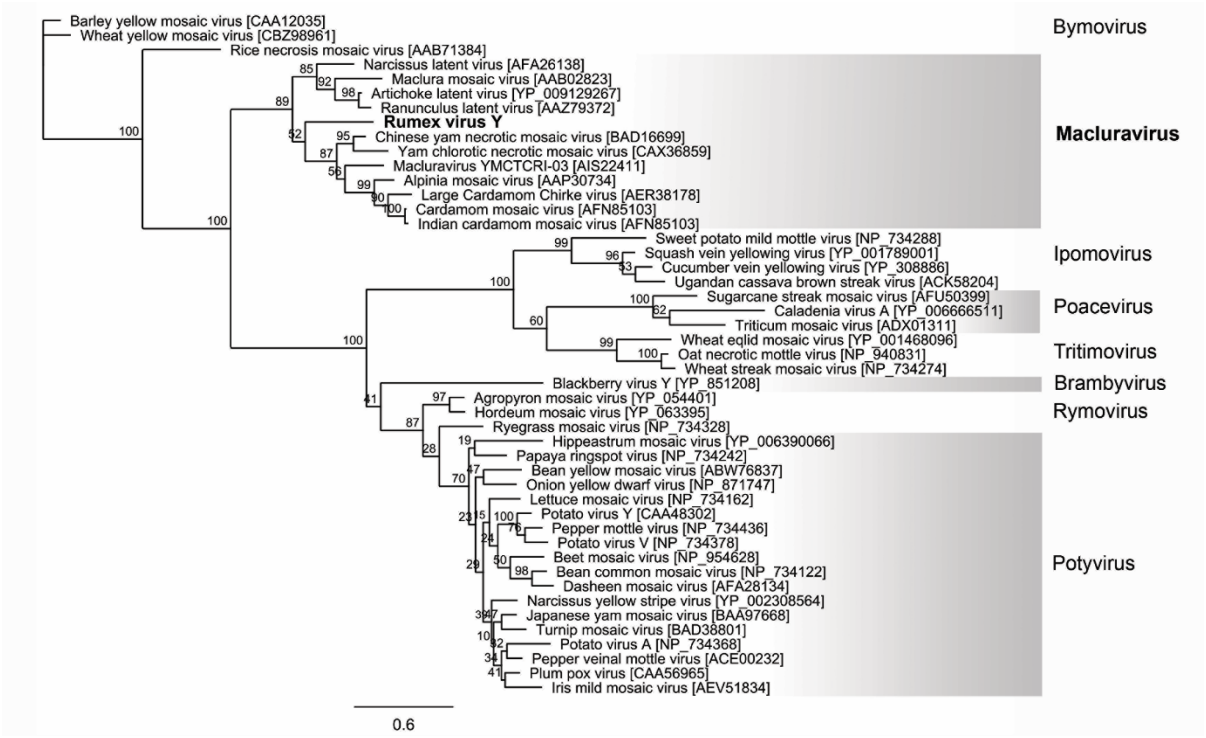


Figure 2.2: The relationship of a new Macluravirus detected in dock with representative members of the family *Potyviridae* in an aa conserved region of the coat protein (pfam00767). Alignment performed with ClustalW and maximum likelihood tree obtained with PhyML v3.

Table 2.1: Number of viral reads mapped to a reference sequence using Bowtie2.

Host	Total read number after QC	Virus	Viral reads	Viral percentage
Potato	125426			
		PVS^{O}	53909	43.0%
		PVS^{A}	35983	28.7%
		PVY	3405	2.7%
Rengarenga	193322			
		TuMV	60738	31.4%
Dock	149708			
		Rumex virus Y	58305	39.0%
		CLRV	370	0.2%

PVS^{O} [KU058656]; PVS^{A} [KU058657]; PVY [D00441]; TuMV [KU053508]; Rumex virus Y [KU053507]; CLRV [KC937025 and KC937030].

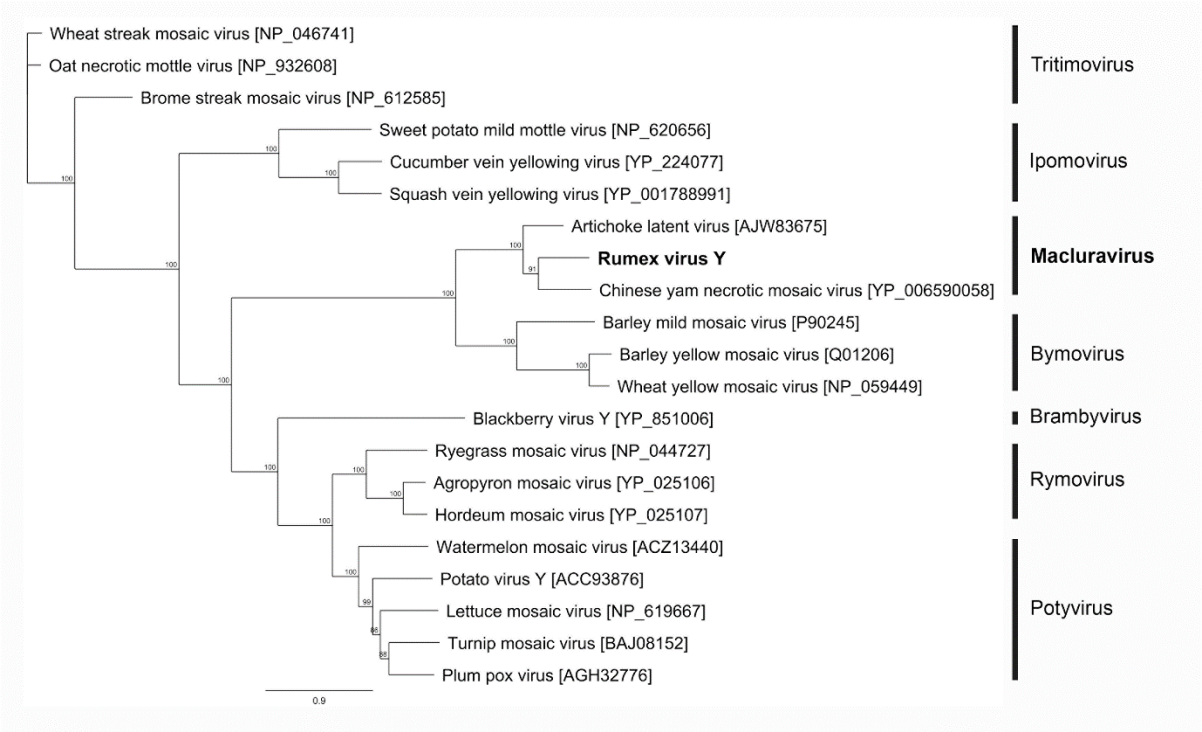


Figure 2.3: The relationship of a new Macluravirus detected in dock with representative members of the family *Potyviridae* in the full polyprotein using ClustalW alignment (BLOSUM matrix, 10 gap opening, 0.1 gap extension). A maximum likelihood tree was built using PhyML v3, with the proportion of invariant sites estimated, empirical aa frequencies, variation in rates across sites modelled as a gamma distribution with 4 categories, 100 bootstrap replicates and using the BLOSUM62 scoring matrix.

2.1.4 Discussion

The use of mAbs raised against dsRNA to enrich for viral sequences shows potential to increase both speed and cost efficiency for identification of virus genomes including those from difficult to process tissues. Purification of dsRNA on cellulose is a common enrichment step for plant virus, but it is performed over 2 days. In this study, we have demonstrated a cost- and time-efficient method to permit enrichment of viral dsRNA from plant tissues that can be completed in less than three hours. Using this method we successfully identified five viruses from three plants, and we were able to distinguish the different strains of a virus infecting a single plant as well as to retrieve the almost full genome of a previously uncharacterised virus.

Our method retrieved a high proportion (31–74%) of reads of viral origin, and substantial proportions of viral genomes. This level of viral recovery is comparable with the one obtained from cellulose chromatography by Minutillo et al. (2015) with 48–70%, but much higher than the results shown by Roossinck et al. (2010) with values ranging from 1% to 5%. However,

the host plants and virus types are highly likely to also contribute to the differences in dsRNA recovery between those studies rather than simply the dsRNA enrichment methods.

Using the dsRNA pull-down assay coupled with NGS, we were able to detect these five different viruses representing three families (*Betaflexiviridae*, *Potyviridae* and *Picornaviridae*) from three different hosts representing a commercial crop, a New Zealand native ornamental and a weed. From potato, two common viruses were identified; PVS that comprised 72% of the reads, and PVY with 3% of the reads. The PVS reads formed two distinct isolates matching the highly divergent PVSA and PVSO clades (Cox & Jones, 2010). It is the first time these two variant strains were detected in the same plant in New Zealand.

The potyvirus TuMV was recovered from rengarenga for the first time. It is now common to find this ubiquitous Potyvirus within New Zealand native plants (Fletcher et al., 2009, Fletcher et al., 2010, Veerakone et al., 2015).

From the dock composite sample we identified almost the full sequence of a novel Macluravirus. Since this sample represents eight dock plants showing diverse symptoms (mild mottle to general necrosis), and another virus was also detected (CLRV), it is not known which symptom is caused by the novel Macluravirus, and therefore it was tentatively named Rumex virus Y (RVY), the Y reflecting the Potyviridae family. The impact of this virus and its host range are unknown.

The mAb-based enrichment protocol was performed in microfuge tubes from only a small amount of freeze-dried samples (100 mg) demonstrating its convenient use in low tech laboratories that might be associated with processing of high numbers of 'field' samples for an ecological or medical study. In 2010, Roossinck presented the result of a large viral metagenomics study from 384 environmental samples that alongside the publications by Thapas et al. (2012 and 2015) represent the largest plant virus metagenomics studies published to date. The Roossinck protocol used 5 g of starting material followed by a phenol: chloroform extraction and chromatography using cellulose CF11 (Morris and Dodds, 1979). We believe that the anti-dsRNA antibodies used in the immunocapture protocol presented in this current report are a major advance over the traditional dsRNA enrichment due to their simplicity and rapidity that will enable similar studies at a reduced cost; our protocol used a crude extract, eliminating the need for phenol: chloroform extraction. In the future, this protocol could be streamlined for large-scale studies as long as the sequencing depth remains sufficient to detect multiple virus infections as the number of reads recovered was very variable between the co-infecting viruses of potato and the rumex, perhaps reflecting the

different rates of replication of the viruses at the time of sampling. In addition, multiplexing large number of samples may be restricted by the normalisation process and the tag jumps where a small number of sequences are not reported to the correct tag (Schnell et al., 2015). In the case of an ecological study a 'lawnmower' metagenomics approach is possible where no individual barcoding is required because the emphasis is on a comparison between different environments (Roossinck, 2012).

All the viruses recovered in the present study are positive sense single-stranded RNA viruses, by far the most common plant virus described to date (Hull, 2014) and the detection of replication intermediates from positive sense ssRNA viruses is also widely reported (Minutillo et al., 2015, Marais et al. 2015). Whether our approach can be further extended to detect negative sense single-stranded RNA viruses and DNA plant viruses should be investigated in future studies. While previous reports (O'Brien et al., 2015, Weber et al., 2006) suggest that mAbs to dsRNA cannot be used for the detection of negative sense single-stranded RNA viruses, we successfully detected a putative member of the genus *Emaravirus* (negative sense RNA virus) from a symptomatic plant using the pulldown assay described herein in a parallel study (data not shown). The detection of dsRNA from DNA viruses has been reported before (Al Rwahnih et al., 2013, Roossinck et al., 2010, Thapa et al., 2015) and is explained by the overlapping of convergent transcription. The sequencing of dsRNA is therefore a very powerful method capable of detecting all virus types, however, the routine use of dsRNA sequencing for negative sense RNA and DNA viruses warrants further validation.

Another area that needs to be explored is whether our protocol can be used to detect replicating virus in recalcitrant plants such as strawberry, banana or yam. Such plants can accumulate significant quantity of tannins, phenolics and polysaccharides that may interfere with the nucleic acids and prevent their binding to the anti-dsRNA mAbs. However, in this case, we believe that the pulldown experiment can be performed on diluted extract in order to reduce the concentration of these interfering compounds, or alternatively it could be used after total RNA extraction.

An alternative approach to dsRNA extraction was developed by Kobayashi and colleagues (2009) using the recombinant plant dsRNA-binding protein GST-DRB4*, derived from Arabidopsis thaliana. Similar to the method presented here, the DRB4* protein fused to a GST-tag was recovered with a Sepharose bead enabling a fast extraction. One aspect their publication does not cover is the plant weight required to detect the infected virus by RT-PCR following the extraction. In addition, their work was exclusively based on herbaceous indicator plants (*Chenopodium quinoa* and *Nicotiana benthamiana*) infected with a single virus. The

GST-DRB4* protein was used for an analysis of the dsRNA present in aquatic microbial communities (Decker & Parker, 2014). The dsRNA was extracted from total microbial RNA isolated from water and sequenced by NGS (Ion torrent). The method was able to detect unique dsRNA reads (about 30% of the total reads) including some that could be assembled into new RNA virus-like elements.

2.1.5 Conclusion

The results presented demonstrate enrichment for viral nucleic acids from soft plant tissues that is comparable with that achieved using the traditional dsRNA extraction method by cellulose chromatography, but with substantial gains in time and ease including the ability to translate to high throughput formats.

In conclusion, Son et al. (2015) suggest that dsRNA antibodies have a role in animal virus discovery using NGS and here, we present that this dsRNA enrichment method is efficient for virus detection and discovery in plants. In addition, recent publications suggest that dsRNA is a good target for virus identification in environmental samples (Decker & Parker 2015), invertebrate samples (O' Brien et al. 2015) and vertebrate samples (Son et al. 2015), and dsRNA enrichment with mAbs represents an appropriate tool for large-scale viral population genetic studies across different environments.

Acknowledgment

We would like to thank John Fletcher for the recommendation and collection of the South Island samples. We thank Mahmoud Khalifa and Simon Bulman for their useful comments on the manuscripts as well as Karmun Chooi and Dan Cohen for the constructive discussions during the research period. Finally we acknowledge Plant & Food Research for the funding and New Zealand Winegrowers for the financial support through the Rod Bonfiglioli memorial scholarship.

2.2 Improvement of double-stranded immunocapture for grapevine virus enrichment

2.2.1 Introduction

High throughput sequencing is a powerful tool for virus discovery but its usage seems disproportionate with regard to the tiny virus genome size. In order to take full advantage of this technology a viral enrichment method is used to increase the viral to plant sequence ratio. Double-stranded RNA (dsRNA) is the enrichment method we selected, using antibody as a bait and magnetic beads as a support. The method was proven very effective for the detection of highly replicating viruses, mostly from herbaceous plants (Blouin et al., 2016), however, when applied to grapevines, it lacked sensitivity, with the majority of the sequence obtained being of plant origin, in particular matching the ribosomal RNA (rRNA). The rRNA is highly structured and forms large numbers of potential targets for the dsRNA antibodies used in the dsRNA capture experiment. We assessed the effect of different buffers and clarification processes as well as the addition of ribonucleases on the capability to detect viruses and rRNA from grapevine tissue.

2.2.2 Materials and methods

Leaf tissue was selected from potted plants with known infection of grapevine leafroll-associated virus 3 (GLRaV-3) group I. Leaf samples were harvested (central part of the leaves and about 1 cm of petiole), chopped into small pieces (< 5 mm) then mixed and divided equally into 21 bags before being dehydrated. The wet weight of each sample was 2 g fresh tissue. This experiment was divided in seven treatments, each made of three replicates. The dsRNA capture protocol was constant across treatments, with 50 ug of protein L magnetic beads (Thermo Scientific Pierce™) coated with 200 uL of hybridoma supernatant (2G4 or 3G1) per sample as previously described (Blouin 2016). The buffers tested were Tris-buffered saline with tween (TBSt : 25mM Tris, 150 mM NaCl and 0.05% tween) or CTAB buffer described by White et al. (2008) (2 % CTAB; 2 % PVP K-40; 25 mM EDTA; 100 mM Tris-HCl (pH 8.0); 2 M NaCl; 0.5 g/L spermidine). The buffer, antibody and protocol used in each of the seven treatments is described in Table 2.2. The recovered beads were resuspended in water and a reverse transcriptase was performed using the Tetro Reverse Transcriptase (Bioline). The

virus and rRNA concentration was measured by hydrolysis probes in a duplex reaction using PerfeCTa Multiplex qPCR ToughMix (Quanta bio) on an EcoStudy (Illumina).

2.2.3 Results and discussion

The results showed no real difference between the two mAB lines (treatment A and B) but a drastic effect of the buffers and process was observed (Figure 2.4). The addition of 2% PVP improved significantly the detection of viral dsRNA with a detection of GLRaV-3 about 15 cycles earlier in treatment C versus A or B. The difference between buffers (treatment C: TBST or treatment D: CTAB) was marginal when the antioxidant and PVP was added (PVP is included in CTAB buffer), but the extract clarification, and the precipitation of the nucleic acids described in Tzanetakis and Martin (2008) helped the enhancement of the viral dsRNA capture. These modifications increased the consistency of the bead recovery. Overall, the CTAB buffer was the preferred buffer for its reliability and because it yielded a smaller and cleaner pellet after the isopropanol precipitation providing the possibility to resuspend in smaller volume and work in microfuge tubes from that step. The pellet can even be resuspended in 1 mL without loss simplifying even more the ensuing steps (data not shown). The level of rRNA recovered was also increased by the modified treatment (C and D) but the addition of RNAse T1 (ssRNA specific) helped reduce the rRNA concentration as shown by a PCR detection delayed by 10 cycles. The increased concentration of RNase T1 (F and G) did not decrease further the rRNA concentration suggesting than the remaining rRNA was highly structured.

The total extraction, from homogenisation to the recovery of the beads, was completed in less than 6 hours and does not require the use of organic solvents. The method was developed for virus discovery and large ecological surveys of grapevines by high throughput sequencing (HTS). Preliminary results show a much increase level of viral reads (up to 75%).

Acknowledgements

This work was supported by The New Zealand Institute for Plant & Food Research Limited and the Rod Bonfiglioli scholarship from New Zealand Winegrowers. We would like to thank John Mackay (dnature diagnostics & research ltd) for his assistance with probe design. We are grateful to Dr Jody Hobson-Peters (University of Queensland, Australia) for providing the dsRNA antiserum.

Table 2.2: Different protocols used for the seven treatments. All treatments were made of three replicates of 2g of grapevine leaf tissue infected with GLRaV-3.

Treatment	A	B	C	D	E	F	G
Extraction buffer	TBSt		TBSt + 2% PVP + 50 mM Na_2SO_3	CTAB + 50 mM Na_2SO_3			
Incubation	X		+ 1mL 20% SDS per 20 mL extract → 20 min at 65ºC + 5mL 5M KAc per 20 mL extract → 20 min on ice				
Clarification	8,500 g for 15 min at 4ºC and filtered						
Nucleic acids precipitation	X		+ 0.8V isopropanol followed → centrifugation at 20,000 g for 20 min at 4ºC. Pellet washed with 70% EtOH and resuspended in 10 mL TBS				
RNAse treatment	X		X		125 U RNAseT1	550 U RNAseT1	500 U RNAseT1
					→30 min @ 37ºC		
Antibodies	2G4	3G1	2G4				
Incubation	1 hour @ 37ºC						

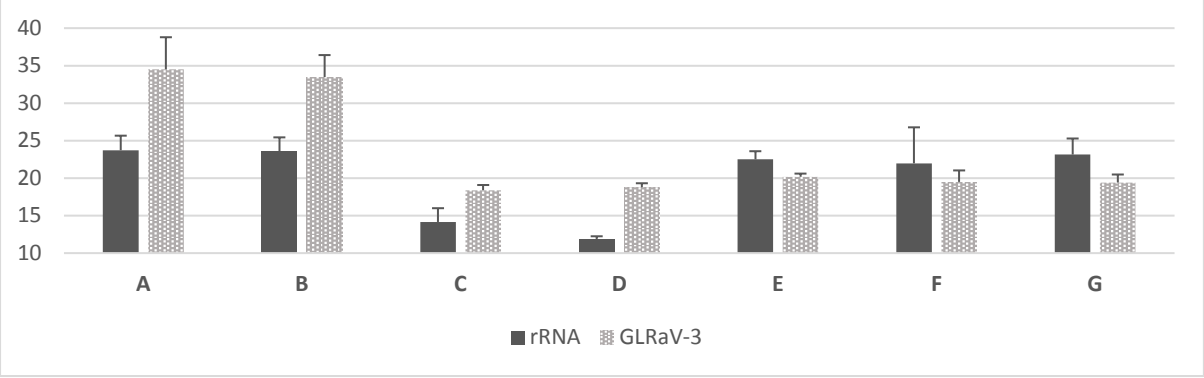


Figure 2.4: Effect of the different protocols on the yield of ribosomal RNA (black) and GLRaV-3 (grey) measured by Taqman assay and expressed in Ct value (low Ct value correspond to high concentration of target sequence) from leaf samples of *Vitis vinifera* infected. Each treatment is made of 3 x 2g replicates and two technical replicates. Treatments are described in Table 2.2.

3 New viruses detected in New Zealand

The novel viruses discovered as part of this study are described in this chapter as well as their phylogenetic relationship with the related genus. This chapter comprises three publications.

Status of chapter:

The first part of this chapter (3.1) has been published in Plant disease Journal as a disease note.

Blouin A. G., & MacDiarmid R. M. (2017b). Distinct Isolates of Grapevine rupestris vein feathering virus Detected in Vitis vinifera in New Zealand*. Plant Disease*, 101(12), 2156-2156.

The second part of this chapter (3.2) has been submitted in Archives of Virology. April 2018.

Blouin A. G., Keenan S., Napier K. R., Barrero R. A., & MacDiarmid R. M. (2018a). Identification of a novel vitivirus from grapevines in New Zealand. *Archives of Virology, 163*(1), 281-284.

The second part of this chapter (3.3) has been submitted in Archives of Virology. April 2018.

Blouin A. G., Chooi K. M., Warren B., Napier K. R., Barrero R. A., & MacDiarmid R. M. (2018b). Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand. *Archives of Virology*, 1-4 (ahead of print).

The co-authorship forms are presented at the start of this thesis, after the acknowledgements.

3.1 Distinct isolates of Grapevine rupestris vein feathering virus detected in Vitis vinifera in New Zealand

Grapevine (*Vitis* sp.) is host to more than 65 viruses (Martelli, 2014d) including grapevine rupestris vein feathering virus (GRVFV); from the genus *Marafivirus*, family *Tymoviridae*. As

part of a virus investigation in New Zealand, GRVFV was detected from three samples of *Vitis vinifera*. RNA was extracted using Spectrum Plant Total RNA Kit (Sigma) and treated with deoxyribonuclease RQ1. A cDNA library was constructed and sequenced by High throughput sequencing (HTS) by a paired-end method (2x125 bp) at the Australian Genome Research Facility on an Illumina HiSeq sequencer. The sequences were assembled with Trinity (Grabherr et al., 2011) in the Galaxy web-based platform. The first sample was a Chardonnay clone 8021 collected in November 2016 in the New Zealand Winegrower's germplasm collection, Lincoln, South Island, New Zealand. From the *de novo* assembly of the total sequence (93M) one contig (isolate NZ Ch8021; accession no. MF000325) of 6708 nt comprising 1533 reads (average coverage of 28.5 fold) representing a near full length GRVFV genome was detected. This contig shared 80% nt identity with GRVFV isolate CHASS (KY513702). From the same HTS run, GRVFV was also detected from two additional grapevines on two different libraries using the same extraction and sequence analysis methods. The first one was obtained from another Chardonnay collected the same day in the same germplasm collection. From the total 92M reads, a contig (isolate ChTK0004; accession no. MF000326) of 6544 nt was assembled from 1416 reads (27 fold average coverage), with the closest sequence from GenBank being GRVFV isolate Mauzac (KY513701, 86% nt identity). The third GRVFV sequence was detected from a lyophilized sample of a Syrah plant showing early reddening symptoms collected in March 2015 from a commercial vineyard (Hawke's Bay, North Island). From the *de novo* transcriptome assembly (99 M reads total) a contig (isolate NZ Sy047; accession no. MF000327) of 2696 nt comprised 299 reads, was detected, with a closest match in GenBank to GRVFV isolate CHASS (KY513702, 84% nt identity). When aligned with the GRVFV reference genome (isolate Mauzac KY513701), the contig aligned from nt 2676 to 5371.The presence of the virus was confirmed in the three samples by RT-PCR and Sanger sequencing using the primers GRVFV_6156F and GRVFV_6600R described in Reynard et al. (2017) and the amplicon showed 100% identity to the sequences described above when available (the two Chardonnay samples MF481199 and MF481200). The addition of these three new sequences of GRVFV increased the knowledge of the virus genetic diversity. Between them, the three isolates shared 80.9 - 83.7% nt identity in their 2696 nt common region (93.2-97% aa identity) and the closest genomes available in GenBank shared only 86.8% nts (isolate Mauzac, KY513701) in the common region. The impact of GRVFV on the health of the grapevine was not assessed, as the three plants tested were co-infected with multiple other viruses and viroids. This is the first report of GRVFV sequence from New Zealand.

Acknowledgments

Plant & Food Research for funding and New Zealand Winegrowers for financial support through the Rod Bonfiglioli Memorial Scholarship.

3.2 Identification of a novel vitivirus from grapevines in New Zealand

3.2.1 Introduction

The genus *Vitivirus* was named after the host *Vitis* of its type member, Grapevine virus A. The genus belongs to the family *Betaflexiviridae*, subfamily *Trivirinae*. Members of this genus have filamentous virions of about 725-825 nm by 12 nm with a positive sense RNA genome of 7400 to 7600 nt with five Open Reading Frames (ORFs) (Adams et al., 2004). The polyprotein, encoded by ORF1, comprises the domains required for replication, the putative movement protein is encoded by the ORF3, and the coat protein (CP) by ORF4. The ORF5 codes for a putative RNA-binding protein. In the last International Committee on Taxonomy of Viruses (ICTV) master species list (2016 v1.3 available from https://talk.ictvonline.org/) nine viruses are officially members of the genus Vitivirus, namely, grapevine viruses A, B, D, E and F; Actinidia virus A and B; mint virus 2; and Heracleum latent virus (Bem & Murant, 1979, Blouin et al., 2012, Minafra et al., 2017, Tzanetakis et al., 2007). Three recent accessions on GenBank suggest additional species, arracacha virus V (Oliveira et al., 2017), Agave tequilana leaf virus (NCBI accession KY190215), and grapevine virus K (MF072319).

3.2.2 Material and Methods

During a small scale untargeted virus survey, a novel vitivirus was identified from a *Vitis vinifera* Chardonnay clone 8021 (VID561 - TK06562) sampled in the New Zealand Winegrowers' germplasm collection, Lincoln, New Zealand. This plant was imported in 1988 from France and entered the germplasm collection which itself was subsequently moved within New Zealand several times. The origin of the virus infection is undetermined as it may have hitchhiked through importation or be a new infection from within New Zealand.

The initial virus sequence information was obtained from small RNA (sRNA) sequencing. Leaf samples were collected in January 2016, and sRNA was extracted using the mirVana microRNA isolation kit (Ambion, Thermo Fischer Scientific). The library was prepared and sequenced on an Illumina HiSeq by BGI. The sRNA analysis was performed using the YABI Virus Surveillance and Diagnosis (VSD) toolkit (Barrero et al., 2017, Hunter et al., 2012) and identified 10 contigs (between 239 to 1333 nt in length), which showed homology to accessions of grapevine virus E (GVE) on GenBank, but only at the protein level (BlastX) with values ranging between 41 to 72% aa homology. This vitivirus-like sequence information was used to design RT-PCR primers and Sanger sequencing of RT-PCR products confirmed the presence and the sequence of the virus. Four viruses and two viroids were also detected from the sample by the VSD toolkit (grapevine leafroll-associated virus 2, grapevine leafroll-associated virus 3 (strain NZ2), grapevine rupestris stem pitting virus, grapevine red globe virus, hop stunt viroid and grapevine yellow speckle viroid).

The plant was resampled in November 2016, and total RNA was extracted using Spectrum Plant Total RNA Kit (Sigma), treated with deoxyribonuclease RQ1 and submitted to RNASeq (125 bp pair-end) at the Australian Genome Research Facility on a HiSeq 2500 sequencer. Assembly of the data (93 M reads in total) was performed with Trinity (Grabherr et al., 2011) in the Galaxy web-based platform. All the viruses and viroids detected previously were confirmed with the addition of grapevine rupestris vein feathering virus (Blouin & MacDiarmid, 2017a). Following assembly, the two largest contigs with homology to vitiviruses were 5551 and 1943 nt long and they could be assembled with a 20 nt overlap (100% identical) forming a final contig of 7474 nt that matched the Sanger sequences previously obtained. When used as a reference, this 7474 nt sequence was mapped using Bowtie2 (version 2.2.4) (Langmead & Salzberg, 2012) to 1279 reads from the total RNA sequencing (21 x coverage) and 418685 reads from the sRNA sequencing (1176 x coverage). The genome was completed with the 5´ UTR sequence obtained with the SMARTer® RACE 5´/3´ Kit (Clontech Laboratories, Inc. A Takara Bio Company) and the 3´ UTR by RT-PCR using an oligo(dT) anchored reverse primer.

3.2.3 Results and discussion

The complete genome of the virus is 7496 nt long (MF405923) and contains five ORFs coding for proteins containing all the conserved domains expected in a vitivirus when analysed using the conserved domain database (Figure 3.1a) (Marchler-Bauer et al., 2011). The ORF1 (position 65:5176 nt) encodes a polyprotein of 1710 aa that contains four recognised domains (from the N terminus): metyltransferase; helicase; 2OG-Fe(II) oxygenase superfamily (AlkB);

and RNA-dependent RNA-Polymerase (RdRp) at its C terminus (Figure 3.1a). The second ORF starts 14 nt after the stop codon of the ORF1 (position 5191-5655 nt) and encodes 154 aa long protein. It has no recognised domains and no known function, as observed for the other vitiviruses (Martelli et al., 1997, Minafra et al., 2017). The third ORF starts 20 nt after the ORF2 (position 5676-6536 nt) and encodes 286 aa long protein with viral movement protein domain. The ORF4 overlaps ORF3 by 89 nt (position 6448:7053 nt) and encodes a 201 aa protein with the recognised Tricho CP domain. ORF5 starts 36 nt after the stop codon of the ORF4 (position 7090:7443 nt) and is the shortest ORF coding for 117 aa with a viral nucleic acid binding domain. The 5´ and 3´ UTRs are 64 and 53 nt long, respectively (excluding the 3´ poly A tail).

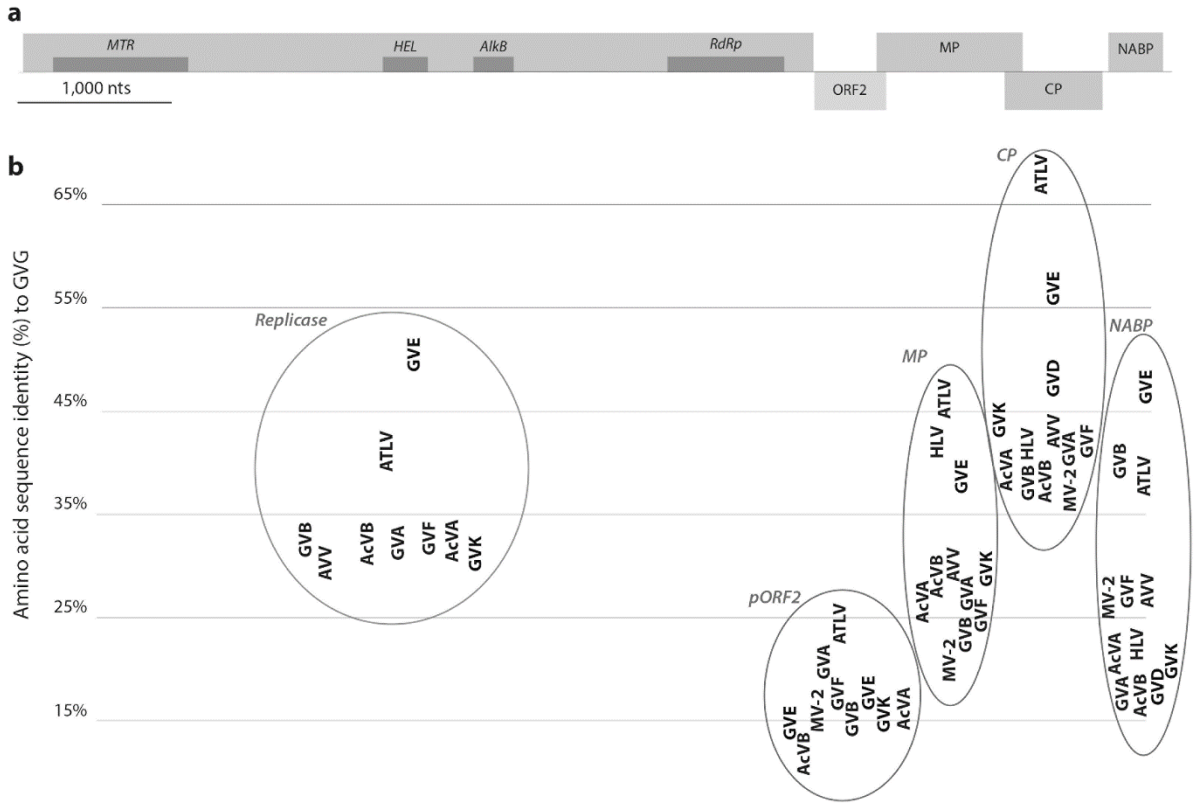


Figure 3.1: a. Schematic representation of the genome organisation for grapevine virus G (GVG). Open reading frames are represented by boxes with the conserved domains italicised. Acronyms used are for the methyl transferase domain (MTR); the helicase domain (HEL), the 2OG-Fe(II) oxygenase domain (AlkB), the RNA-dependant RNA-polymerase domain (RdRp), the movement protein (MP), the coat protein (CP) and the nucleic acid binding protein (NABP). **b.** A graphic representation of the aa sequence identity (in percentage) in the five open reading frames (when available) between GVG and other vitiviruses. The aa percentage identity is set at the base of the first letter of the virus acronym. The protein of the viruses

compared originate from the sequences of grapevine virus A (GVA) DQ855086; grapevine virus B (GVB) EF583906; grapevine virus D (GVD) KX828708; grapevine virus E (GVE) GU903012; grapevine virus F (GVF) JX105428; grapevine virus K (GVK) MF072319, Agave tequilana leaf virus (ATLV) KY190215; Actinidia virus A (AcVA) JN427014; Actinidia virus B (AcVB) JN427015; arracacha virus V (AVV) KY392781; mint virus 2 (MV-2) AY913795; and Heracleum latent virus (HLV) CAA55855.

Within the vitivirus, the closest sequences available on GenBank are those of GVE and Agave tequilana leaf virus (ATLV) (Figure 3.1b). On the full sequence pairwise comparison, the new virus shares 52 and 53% nts identity with ATLV and GVE respectively. The replicase gene shares 51% nts (41% aa) and 56% nts (51% aa) to ATLV and GVE respectively. The highest homology is found in the coat protein encoded in the ORF4, with 65% nts (67% aa) and 59% nts (55 % aa) with ATLV and GVE respectively.

Phylogenetic analysis was performed on the protein sequences of the replicase gene and the coat protein with the members of the genus Vitivirus. ClustalW was used for sequence alignment (10 gap opening, 0.1 gap extension) and a neighbour joining tree (Jules-Cantor distance model) using citrus leaf blotch virus (CLBV, JN983456) and apple chlorotic leaf spot virus (CAE52495) as outgroup for the analysis of the replicase (Figure 3.2a) and coat protein (Figure 3.2b) respectively. Both analyses confirm a close relationship of the new virus with GVE and ATLV. A similar phylogenetic tree was obtained from the full genome alignment (data not shown). The sequence of grapevine virus T (GVT, MF095096), deposited in GenBank as a possible vitivirus, was not included in these analysis as it does not show enough homology with members of this group. The taxonomy of this accession should be reviewed as its genome structure and sequence homology suggest GVT to be a member of the genus Foveavirus. The analysis also provides indication that GVK (MF072319) is a variant of GVD (MF774336).

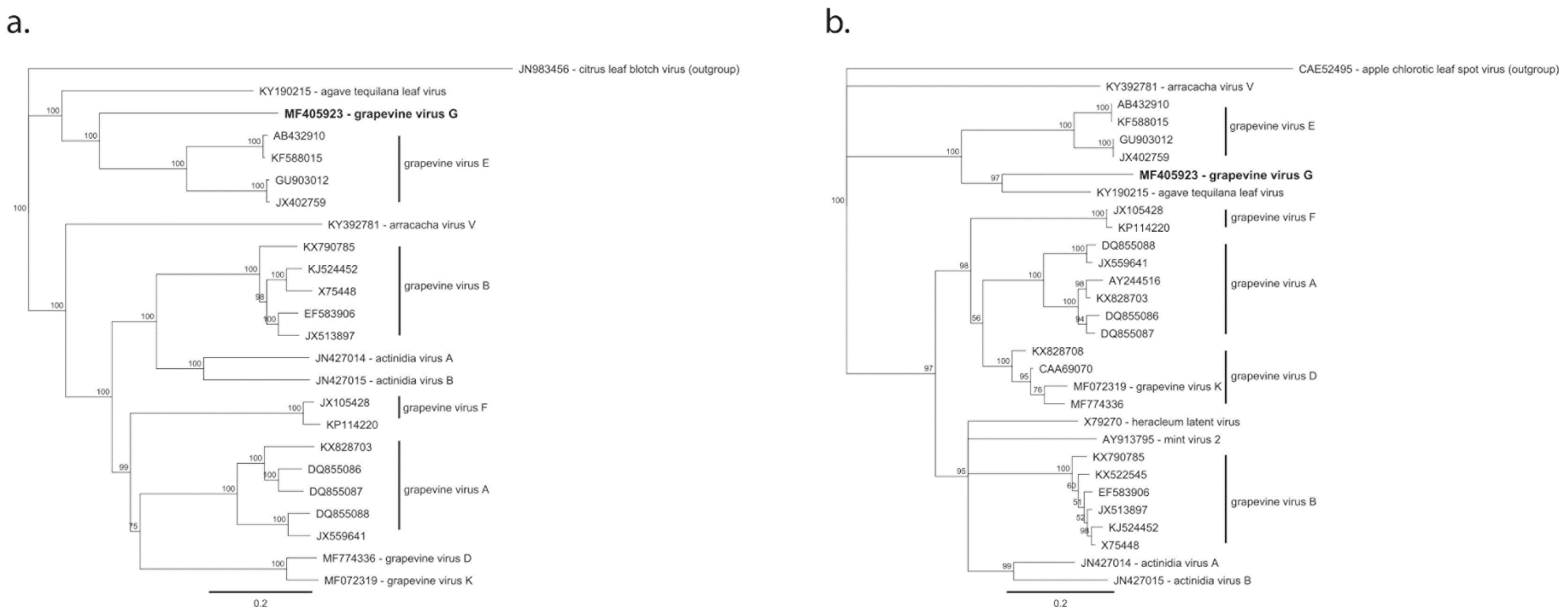


Figure 3.2: Neighbor-Joining trees (1000 bootstrap replicates using the Jukes-Cantor distance model) of two proteins (a. replicase polyprotein, b. coat protein) of representative members of the genus Vitivirus. Protein alignment (translated from the accession number indicated) performed with ClustalW (BLOSUM cost matrix with a gap opening cost set at 10, and a gap extension cost at 0.1). Consensus support is shown in percentage on the branch. Citrus leaf blotch virus and apple chlorotic leaf spot virus used as outgroups for the phylogenetic analysis of a. (replicase) and b. (coat protein) respectively.

During the original sRNA HTS run, two additional genomes of the same virus were obtained (sequence accessions MF405924 and MF405925). The limited genetic variation (< 1% nt across the genome) suggests a single infection origin. Furthermore, a limited survey made of 30 plants from the same germplasm collection and 26 from various commercial sources was performed from dsRNA extracted by immunocapture following a protocol adapted from Blouin et al (2016). The reverse transcriptase followed the same method as Blouin et al (2016) and the PCR used the primers GVG_4951F (GGC AAG TTG GAG GTG GAT ATG AC) and GVG_5285R (ACT GAT GTT AAG AGG TAG CTT GCA C). Nine additional plants were found positive from the germplasm, and two from a commercial vineyard. Sequence analysis of this 282 amplicon showed less than 4% nucleotide variation between the sequences.

3.2.4 Conclusion

Based on the species demarcation criteria proposed by the ICTV, distinct species is considered within the genus Vitivirus if it shares less than 80% aa identity or 72% nt identity for the CP and the RdRp with its closest relative (Adams et al., 2004). This new sequence is clearly below that threshold and should be considered as a representative of a new species in the genus Vitivirus. We propose to name this new virus grapevine virus G (GVG). Assessment of the biological impact of this virus on the plant is challenging due to difficulty to transmit the virus via inoculation with virus particles (Minafra et al., 2017) and because of the presence of co-infecting viruses. However, future research to extend the survey to additional material to estimate the extent of the virus spread may identify a vine infected with only GVG.

Acknowledgements

We would like to acknowledge The New Zealand Institute for Plant & Food Research Limited for funding, New Zealand Winegrowers for the financial support through the Rod Bonfiglioli Memorial Scholarship, and Plant Biosecurity Cooperative Research Centre for the funding of the small RNA sequencing and analysis (PBCRC2064). We would like to thank our Plant & Food Research colleagues, John Fletcher for the collection of samples, Ben Warren for his assistance in the bioinformatics, and Kieren Arthur, Kar Mun Chooi, and Dan Cohen for the useful comments on the manuscripts. We would like to acknowledge Dr Maher Al Rwahnih for alerting us of the new genome of GVD available on GenBank.

3.3 Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand

3.3.1 Introduction

The use of High Throughput Sequencing (HTS) has had a massive effect on the rate of discovery of previously overlooked obligate parasites. The decrease of the technology cost combined with the uptake of the methodology by more diagnostic laboratories has resulted in an unprecedented level of detection of novel virus genomes (reviewed in Roossinck, 2017). We reported earlier the finding of a new vitivirus named grapevine virus G (GVG) from small RNA (sRNA) sequencing and total RNA sequencing (Blouin et al., 2018a). While characterising GVG, we identified from the same HTS run a second, related vitivirus. In order to maintain the accepted taxonomical consistency we propose to name this new virus grapevine virus I (GVI) and we use this name hereafter.

3.3.2 Material and Methods

A sample of Vitis vinifera cv Chardonnay (VID499 – TK0004) was collected from the New Zealand Winegrowers' germplasm collection, Lincoln, New Zealand, in November 2016 and total RNA was submitted to RNASeq at the Australian Genome Research Facility on a HiSeq 2500 sequencer after a DNase RQ1 treatment alongside with the sample VID561 - TK06562 in which the virus GVG was described. Using the bioinformatics pipeline previously described (Blouin et al., 2018a). A vitivirus-like sequence of 7439 nt was retrieved from the *de novo* analysis. In light of this new sequence, we examined the small RNA (sRNA) data obtained previously (Blouin et al., 2018a). from the same plant and we found that the contig was mapped by 217532 reads (579 x coverage). The sequence was confirmed by Sanger sequencing. The genome was completed with the 5´ UTR sequence by using the SMARTer® RACE 5´/3´ Kit (Clontech Laboratories, Inc. A Takara Bio Company) and the 3´ UTR by RT-PCR with an oligo(dT) anchored reverse primer. The full genome of 7507 nt, excluding the polyA tail, was deposited in GenBank, as grapevine virus I, under the accession number MF927925. In addition to this virus, the plant was found infected with several viruses including grapevine leafroll-associated virus 3, grapevine virus A, grapevine rupestris stem pitting virus,

grapevine rupestris vein feathering virus (MF000326), grapevine redglobe virus, grapevine virus G (MF405924), and the viroids hop stunt viroid and grapevine yellow speckle viroid, and these were confirmed from the sRNA data using the YABI Virus Surveillance and Diagnosis (VSD) toolkit (Barrero et al., 2017, Hunter et al., 2012).

3.3.3 Results and discussion

The genome of GVI is comparable to the description of vitiviruses (Adams et al., 2004, Minafra et al., 2017) with a single positive single-stranded RNA molecule containing five Open Reading Frames (ORFs). The ORF1 encodes a 1696 aa polyprotein (nt position 69-5159) that contains the recognised domains of methyltransferase; helicase; 2OG-Fe(II) oxygenase superfamily (AlkB); and RNA-dependent RNA-Polymerase (RdRp). The closest relative on GenBank is grapevine virus E (GVE, isolate SA94, GU903012) with 65% aa and nt identity. The second ORF (ORF2) overlaps with the ORF1 by 11 nt (nt position 5149-5652) and codes for a 167 aa putative protein with poor homology to known proteins and no recognised domains, as observed in previously characterised viruses from that genus (Minafra et al., 2017). The third ORF (ORF3) starts 32 nt downstream of the ORF2 and codes for a 264 aa protein (nt position 5685-6479) containing a viral movement protein domain. The movement protein of GVE is its closest relative with 63% aa identity (65% nt). The next ORF (ORF4) overlaps with ORF3 by 70 nt and codes for a 199 aa protein (nt position 6409-7008) containing the tricho coat super family domain. This protein shares 65% aa identity to the coat protein of agave tequilina leaf virus (ATLV) (68% nt); 63% aa with GVE (66% nt) and 62% aa with GVG (61% nt). The ORF5 starts 29 nt downstream of the ORF4 and codes for a 121 aa protein (nt position 7038-7403) with a recognised viral nucleic acid binding protein (NABP). This protein is the most conserved with GVE the closest match on GenBank (72% aa and 70% nt identity). It is interesting to note that the NABP of Grapevine virus B (GVB) is the second closest match with 66% aa identity as none of the other proteins of the vitivirus GVB group with the GVE clade (Figure 3.3).

A phylogenetic analysis was conducted on the replicase and the coat protein genes from a ClustalW aligment (BLOSUM cost matrix with gap opening cost set at 10 and gap extend cost at 0.1) and made with Neighbor-joining method using the Jukes-Cantor genetic distance model. The citrus leaf blotch virus replicase (JN983456) and apple chlorotic leaf spot virus coat protein (CAE52495) were used as outgroups. The replicase phylogenetic tree (Figure 3.3a) shows GVI branching off the GVE cluster before GVG and ALTV. The coat protein,

although not strongly supported, GVI clusters with GVG and ATLV within the GVE group (Figure 3.3b).

From a limited survey of old vine accessions located in the same germplasm collection, Lincoln, New Zealand, we identified eight GVI positive vines (including the original VID499-TK0004) from 18 plants tested by sRNA HTS with between 196 to 8209 reads per million mapped to the genome MF927925 (coverage of 92 to 100% of the genome; 17- to 842-fold coverage). Five positive samples were *V. vinifera* (Sylvaner, Chardonnay, Dolcetto or Shiraz), one was a *V. labrusca* (Fredonia) and two were interspecific hybrids (Chelois and Pinard). In order to evaluate the establishment of the virus, 58 additional vines (34 from the germplasm and 24 from a commercial vineyard) were tested for the presence of GVI using cDNA synthetized from immunocaptured dsRNA (Blouin et al., 2016), followed by a PCR using the primers GVG-GVI 4595F (TTY TCT CAG AAG ART TAY GAT GAT C) and GVI 5212R (TAT GTT CAG CTC ATG AAG GTG CTC) and sequencing. Two additional infections were detected from the germplasm but the virus was not detected in the plants sourced from commercial vineyards.

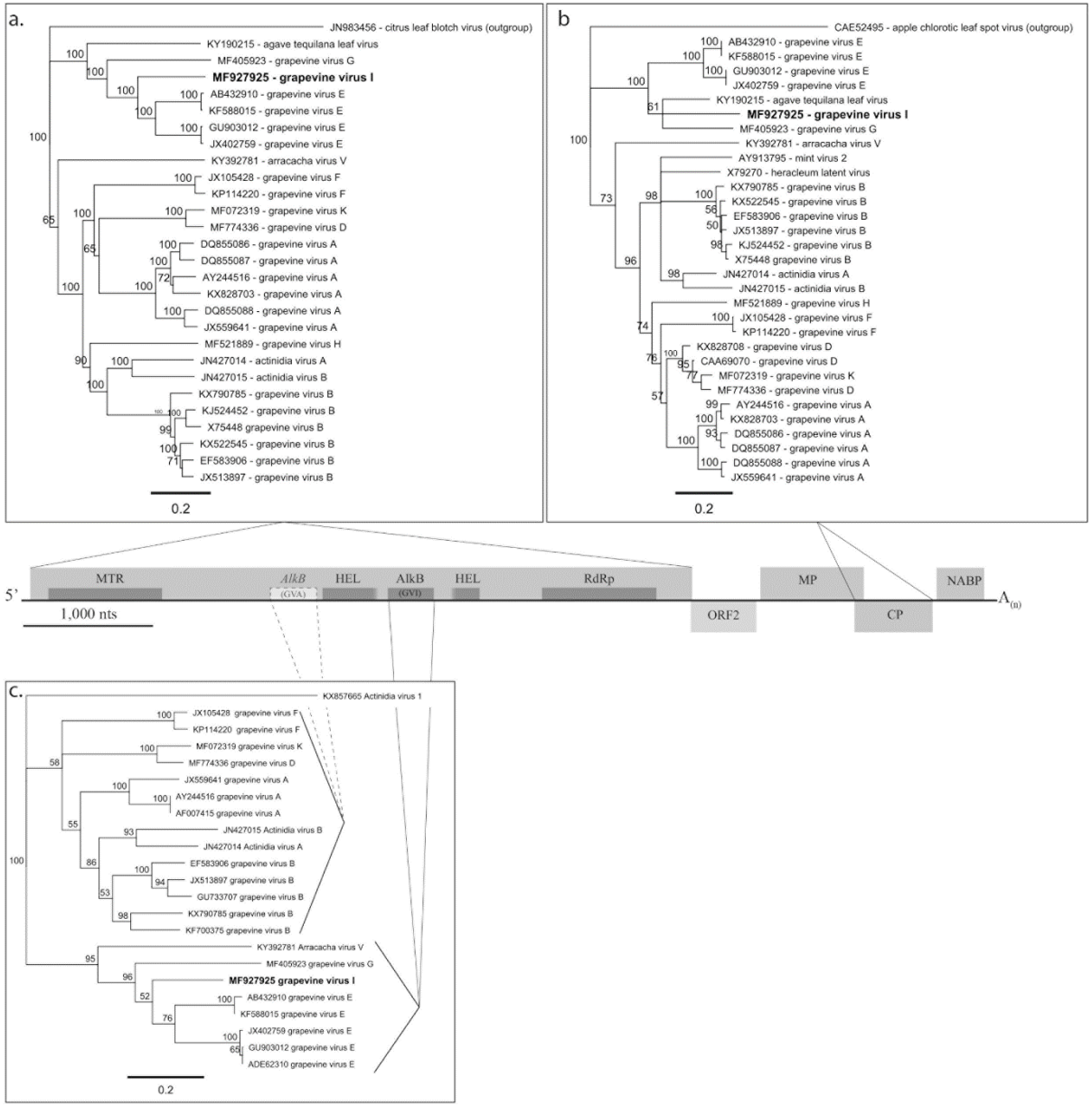


Figure 3.3: Schematic representation of the genome organisation for grapevine virus I (GVI). Open reading frames are represented by boxes with the conserved domains italicised. Acronyms used are for the methyl transferase domain (MTR); the helicase domain (HEL), the 2OG-Fe(II) oxygenase domain (AlkB), the RNA-dependant RNA-polymerase domain (RdRp), the movement protein (MP), the coat protein (CP) and the nucleic acid binding protein (NABP). Numbers at the edges of the boxes represent the first and last nucleotide postion of the ORF. The white box with AlkB (GVA) represents the alternative location of this domain for some other members of the vitivirus (GVA clade). Insets represent the Neighbor-Joining trees (1000 bootstrap replicates using the Jukes-Cantor distance model) of a. replicase polyprotein, b. coat protein and c. AlkB domain of representative members of the genus Vitivirus. Protein alignments (translated from the accession number indicated) were performed with ClustalW (BLOSUM cost matrix with a gap opening cost set at 10, and a gap extension cost at 0.1). Consensus support is shown as a percentage on the branch. Citrus leaf blotch virus, apple chlorotic leaf spot virus and Actinidia virus-1 were used as outgroups for the phylogenetic analysis of a. (replicase) b. (coat protein) and c. (AlkB) respectively.

The replicase gene structure of the vitiviruses shows a shift in position of the AlkB domain. AlkB proteins are widely distributed in cellular organisms but are only found in viruses infecting perennial hosts. Their role may involve protection against methylation damage. In the replicase polyprotein, the AlkB domain is found within the helicase domain for GVE, GVG, arracacha virus V (AVV) and GVI , as opposed to the other members of the vitivirus (GVA clade) that have their AlkB domain located upstream of the helicase (Figure 3.3). This alteration of the genome arrangement supports the hypothesis that GVE and relatives have gained their AlkB domains horizontally, independent from the GVA clade (Dolja et al., 2017). This explains the differences among the sequences of the domains, where GVE, GVG, AVV and GVI cluster together (Figure 3.3c). Although, AVV is genetically related with the members of the GVA clade (Figure 3.3a and Figure 3.3b), the position and aa sequence of its AlkB domain is related to members of the GVE clade suggesting a similar origin (Figure 3.3c). ATLV is the only vitivirus without an AlkB domain and this absence explains its shorter genome by about 350 nt, a difference that corresponds to the length of this motif. The phylogenetic analysis of the replicase (Figure 3.3a) suggests that the virus never incorporated the motif (more basal and thereby older branching). Despite its close relationship with GVE, GVG and GVI, ATLV is not only unique by the lack of the AlkB domain but it also lacks the ORF2. Agave is the only known monocotyledon infected with a vitivirus to date, and does not undergo traditional secondary growth with production of secondary phloem; it may provide a clue for the function of the AlkB domain and ORF2 protein.

3.3.4 Conclusion

The genome of the new virus described here falls below the threshold for species demarcation from its closest relative within the genus Vitivirus of 80% aa identity or 72% nt identity for the coat protein and the RdRp (Adams et al., 2004). Therefore we propose that it is considered as a new species in the genus under the name grapevine virus I.

In the 20-year interval between its establishment (1997) (Adams et al., 2004) and its review in the last ICTV update in 2016 the genus *Vitivirus* gained five species from the four original members (GVA, B, grapevine D (GVD) and heracleum latent virus) to nine (GVE; grapevine virus F; Actinidia virus A and B; mint virus 2) (Al Rwahnih et al., 2012b, Blouin et al., 2012, Nakaune et al., 2008, Tzanetakis et al., 2007). Eighteen months since that last ICTV release, four new viruses that fit the description of the genus have been reported or deposited on GenBank: AVV, ATLV (KY190215), GVG and grapevine virus H (Blouin et al., 2018a, Candresse et al., 2018, Oliveira et al., 2017). In addition, the full genome of GVD was released

(MF774336). With the description of GVI, and assuming they are all accepted within the genus in the next ICTV release, the number of vitiviruses will increase to 14, eight of which have been described in grapevine.

Acknowledgements

This work was funded and supported by the Rod Bonfiglioli Scholarship from New Zealand Winegrowers Inc., and The New Zealand Institute for Plant & Food Research Limited Wine Research programme, funded by the Ministry for Business, Innovation and Employment Strategic Science Investment Fund. The Plant Biosecurity Cooperative Research Centre funded the small RNA sequencing and analysis (PBCRC2064). We would like to thank our Plant & Food Research colleagues: John Fletcher for the collection of samples, Sandi Keenan for her assistance in the sRNA extraction, and Kieren Arthur, and Dan Cohen as well as Mahmoud Khalifa (Damietta University, Egypt) for the useful comments on the manuscripts.

4 Snapshots of a country's vineyard virome

This chapter summarises the results obtained from sequencing 225 grapevines from four different environments and discusses the significances of these results.

Status of chapter:

At the time of thesis submission this chapter had not been published.

4.1 Introduction

New Zealand's year-on-year increasing wine export industry is currently worth more than $NZ 1.6 B per annum with a $NZ 2 B target by 2020. Despite wine making dating back to the colonial days of the mid-1800s, its development to the current form is through relatively recent plantings with a major expansion of vineyard area over the past 30 years including the flagship Marlborough Sauvignon blanc vineyards. Representing 60% of the vineyard surface, this cultivar sustains 86% of the wine exported (New Zealand Winegrowers, 2017a).

The viral disease pressure on the New Zealand vineyards was first surveyed back in 1970, and at that time only two viral diseases were detected: fanleaf caused by grapevine fanleaf virus (GFLV genus *Nepovirus*), and leafroll disease associated with the leafroll viruses (Chamberlain et al., 1970). In the last review of plant viruses recorded in New Zealand, 17 viruses of grapevine were reported from the 70 known to infect grapevines worldwide (Veerakone et al., 2015, Martelli, 2017). But from the New Zealand grape grower's perspective, only one virus is a major cause of concern: grapevine leafroll-associated virus 3 (GLRaV-3). For its detrimental effects on the red grape cultivars, the industry and government have invested to reduce its incidence (Andrew et al., 2015). However, in Sauvignon blanc, viruses are mostly overlooked as a result of the lack of obvious symptoms. Here we assess whether New Zealand grapevines host more viruses that could be a potential future threat, or benefit, to the vineyards under environmental changes (e.g. climate change, or new virus or virus-vector introductions).

Historically, detection of plant viruses has been a challenging task, due to their invisibility (without an expensive transmission electron microscope), and inability to culture isolates (obligatory pathogens). The advance in technologies such as ELISA helped diagnostic laboratories identify known strains, but only with the development of high throughput sequencing (HTS) has *de novo* discovery of viruses become simplified. Even so, to sequence all DNA or RNA from a plant in order to detect the small contribution of viruses' tiny genomes is expensive in term of reagents, computing and time. To reduce these expenses, it is possible to enrich for the viral target as reviewed by Massart et al (2014). One method is to deplete the ribosomal RNA that represents 80% of the total RNA sequenced thereby increasing the viral ratio in the sequence data. By contrast, rolling circle amplification is used to amplify the signal of circular DNA viruses (Shepherd et al., 2008). Virion-Associated Nucleic Acids (VANA) consists of sequencing the nucleic acid obtained after a purification of virus particles, thus removing most of the plant genetic content during the process (Filloux et al., 2015, Blouin et al., 2010). Another viral sequencing enrichment is obtained with the small interfering RNAs (siRNA); in part siRNA is a degradation product of virus RNA by the plant cell and can be detected by sequencing small RNA (21-24mers) (Kreuze et al., 2009) (Chapter 1, Figure 1.1). Finally, double stranded RNA (dsRNA) is a replicative form of all RNA viruses. Viruses with positive sense genomes represent the vast majority of the plant viruses described to date. Furthermore, reports show that dsRNA sequencing have been used to detect some DNA viruses (Al Rwahnih et al., 2015b, Roossinck et al., 2010, Rott et al., 2017, Simmonds et al., 2017). Because the plant does not produce many long dsRNA molecules, it has been a method of choice for virus nucleic acid detection over many decades (Morris & Dodds, 1979). Historically, despite being able to isolate viral nucleic acids, the identification of the virus was still challenging until the era of the HTS that brought the capability to sequence DNA, without *a priori* knowledge of its coding i.e. untargeted sequencing. As a consequence dsRNA extraction again became a popular technique amongst plant virologists. The tedious aspect of the extraction with the use of solvents and the large scale of the operation due to the low titre of dsRNA (especially in woody host) made the method time consuming. Recently we proposed an alternative method that exploits immunocapture using dsRNA specific antibodies (Blouin et al., 2016).

Here we present a modified immunocapture method to enrich viral genomes from grapevines and its use to review the virome of New Zealand's vineyard. We measured the virus status of grapevines from four distinct categories. The first group, SB, comprised 167 vines from four commercial Sauvignon blanc vineyards (SB-1, SB-2, SB-3 and SB-4) planted in two major wine

regions and was selected to assess the background virome of the New Zealand large-scale grape growing; the second and third categories, GC, comprised a total of 35 vines collected in the New Zealand Winegrowers' (NZW) germplasm collection, 19 from a "low-health" block (GC-LH) and 16 from a "high-health" block (GC-HH), that collectively form a set of historic virus origins into New Zealand; the fourth group, 'disease response' (DR), comprised 25 samples from grapevines with atypical phenotypes, some of which were collected and submitted by viticulturists. The virus detection results obtained from sixteen of the GC-LH were compared using small RNA HTS. In total, this study reports the viral status of 225 vines from different origins to give today's picture of the New Zealand grapevine virome.

4.2 Materials and Methods

4.2.1 Plant material

First, four vineyards planted with Sauvignon blanc were sampled in early April 2016 (autumn), two from Hawke's Bay (New Zealand North Island) blocks SB-1 and SB-2 and two blocks, SB-3 and SB-4, from Marlborough (South Island). The well-established vineyards were about 20 years old in the South Island (SB-4 was planted in 1995 and SB-3 in 1998) and 10 years old in the Hawke's Bay (SB-1 and SB-2 were planted in 2006). All the vineyards used the same rootstock (SO4) except for SB-2 (Schwarzmann). For each location, 45 plants were collected across the vineyard regardless of vine phenotype. The location of each plant was recorded as the row, bay and vine numbers. Leaf and mature cane material were collected. Leaf samples were stored freeze-dried and cane material was stored into plastic bags held in a 4^{o}C cold room until processed.

Samples from the second and third categories were collected from the NZW germplasm blocks (Lincoln, South Island) in June 2017 (winter) as cane material. GC-LH material, 19 samples from the low-health germplasm block, and GC-HH material, 16 samples from the high-health block were stored in a 4^{o}C cold room until processed. The low-health block is a repository of various *Vitis* species assembled from multiple grapevine collections made in New Zealand over the past century including the first one, originated from the Te Kauwhata research station established in 1897, under the name Waeranga (Bragato, 1903). The high-health block is a collection made of grapevines that had historically been subjected to virus-curing by thermotherapy and more recent imports that were subjected to biosecurity scrutiny (1980s onward). Most of the grapevines from the high-health block were tested for common viruses

(nepoviruses, leafroll 1 2 3 and 4, and grapevine fleck virus) before the establishment of the block in the early 2000s.

Figure 4.1: Picture of Grapevine VID280 (AB537) from the germplasm collection (picture taken in February 2016)

The disease response group (DR) was collected by a range of people at different times. They all originated from commercial vineyards in the Hawke's Bay region and were collected because of the symptoms, unusual phenotype, or as control for a neighbouring symptomatic grapevine (see Table 7.3 and Figure 4.2). Collections of these mature leaves were performed either in March 2015 or April 2017 and the samples were preserved as freeze-dried leaves until processed.

Sixteen of the 19 GC-LH samples were also collected for small RNA sequencing. Only GC-LH samples AB384; AB546 and AB549 were not tested by the sRNA HTS methods. Mature leaves from the middle of the canopy were sampled in January 2016 (summer). Three samples were collected from plant AB541, one at the base of the canopy, one at the middle as for the other samples, and one on the top of the canopy.

Figure 4.2: Picture of the reddening symptoms observed on some AB045 sample (picture taken April 2015)

Figure 4.3: Picture of Grapevine SB-3 U169, first plant on the left (picture take in December 2017 by Bex Whoolley)

4.2.2 Double-stranded RNA Sample processing

At sample preparation time, each cane from the SB environment was wiped with 70% ethanol to remove potential fungal growth. The bark was removed from the cane material (SB, GC-LH and GC-HH samples) and the phloem tissue was scraped with a sterile scalpel then placed into a Bioreba universal extraction bag. The weight of the phloem tissue harvested was noted and the bags were frozen (-80^{o}C) and were freeze-dried upon processing (48 hours in a FTS Flex-Dry freeze dryer). The bags were sealed with masking tape after lyophilisation. A total of 177 SB samples were prepared, three samples from vineyard SB-4 were not prepared as the tissue was either dead or too small; 168 samples were prepared from more than 2.5 g fresh weight (FW) of tissue, nine samples had a FW varying from 1.5 and 2.4 g. All the samples from GC-LH and GC–HH were prepared from more than 2.5 g FW. For the DR samples, about 1 g of dry weight (DW) of leaf tissue (mostly leaf veins) was transferred into a Bioreba universal extraction bag and sealed with masking tape. The tissue was then ground into a fine powder using a tissue homogeniser mounted on a drill press.

Total nucleic acid extraction was performed using CTAB buffer (2 % CTAB; 2 % PVP K-40; 25 mM EDTA; 100 mM Tris-HCl (pH 8.0); 2 M NaCl; 0.5 g/L spermidine with 50 mM Na_2SO_3), as described by White et al. (2008). In a fume hood the tape seal was carefully removed from the sample bag and CTAB buffer was added to each sample in the ratio 10 mL per g FW for SB, GC-LH and GC-HH samples, and 30 mL per g DW for DR samples. The fume hood was wiped with tissue moistened with 70% ethanol between each sample to reduce the risk of cross contamination. A maximum of 25 mL of extract was recovered into a 50 mL conical centrifuge tube, after filtration through the inner bag synthetic intermediate layer. Additional extract, when available, was preserved frozen.

The following protocol is described for a 25 mL extract but was adapted for a lower volume when required by keeping the same proportions. A sodium dodecyl sulfate and potassium acetate (SDS/KAc) extraction (DePaulo & Powell, 1995) was performed as follows: 1.25 mL of 20% SDS solution was added to the 25 mL extract followed by an incubation at 65^{o}C for 20 min. Then 6.25 mL of 5 M KAc was added and the samples were left on ice for a further 20 min. The tubes were then centrifuged at ~12,000 x *g* for 15 min at 4^{o}C. The extract was passed through a miracloth into a sterile 50 mL conical centrifuge tube and 0.8 volume of isopropanol was added followed by a gentle mix prior to centrifugation at ~20,000 x g for 20 min, at 4^{o}C. The pellet was washed twice with 75% EtOH and resuspended in Tris-buffered saline (TBS, 25 mM Tris, 0.15M NaCl, pH 7.5) and transferred into a microfuge tube; 125 U

of RNase T1 (Applied Biosystems Business) was added and the tubes were incubated at 37°C for 1 hour on a rotary tube suspension mixer.

Simultaneously, protein L magnetic beads (Thermo Scientific Pierce™) were washed and coated with anti-dsRNA antibody as previously described (Blouin et al., 2016) at the following concentration: for each sample, 50 µg of magnetic beads were coated with 200 µL of monoclonal antibody 2G4 hybridoma supernatant (O'Brien et al., 2015) at 37°C for 1 hour. The coated beads were then washed with TBS buffer and added to the RNase T1 digested extract and placed again at 37°C for a further hour. The beads were finally washed three times on a magnetic rack holder with 1 mL, 500 µL then 200 µL TBS, then air dried for a few minutes and resuspended in 10 or 15 µL ultra-pure water.

RT-PCR: The reverse transcription (RT) was performed following the method described previously (Blouin et al., 2016) with the following modifications: the total reaction volume was 10 µL including 4 µL resuspended (homogenised) beads; denaturation of the dsRNA (and antibodies resulting in their separation) was performed at 99°C for 2 min; and RNase inhibitor (40 U RNasOUT, Life Technologies) was added to the enzymatic reaction. Following the RT, RNase A treatment was performed as previously described then the cDNA was cleaned using Agentcourt Ampure XP (Beckman Coulter Inc) according to the manufacturer's protocol; the cDNA was eluted in 15 µL ultra-pure water. Polymerase chain reaction (PCR) was performed using the CloneAmp HiFi (Clonetech, Takara) in a 20 µL reaction following the manufacturer's instructions, with 1 µM of a single primer for each reaction and 4 µL sample cDNA. The first cycle was 98°C for 2 min, 65°C for 1 min and 72°C for 1 min followed by 40 cycles consisting of 98°C for 5 sec, 48°C for 5 sec and 72°C for 5 sec followed by a final extension of 72°C for 5 min. PCR products were visualised on a 1 % agarose gel.

A positive control was extracted alongside each extraction batch (13 samples). The sample was made of a freeze dried grapevine leaf (1.5g FW) infected with GLRaV-3 and GRSPaV. After the RT reaction samples were tested for the GRSPaV by Taqman PCR using GRSPaV_8479F CCT TGG AGA GAT TAG TGG TGG AA and GRSPaV_8659R AGC ATG GAA AGG GAA TAC TAT TAG TAC with the CAL Fluor Orange 560 – BHQ-1 probe GRSPaV_8630P TAT TCC AGC GAA CAG GCT TAA CCC AGC in duplex with ribosomal RNA reaction Vv_rRNA-26s_2874F TCG ATG TCG GCT GTT CCT ATC and Vv_rRNA-26s_2953R CAG CTC ACG TTC CCT ATT GGT with the CAL Fluor Red 610 – BHQ-2 probe Vv_rRNA-26S_2897P TGT GAA GCA GAA TTC ACC AAG TGT TGG A. The rRNA PCR primers and probes were designed on the 26S ribosomal RNA of *Vitis* (accession AF479207.1) because of its abundance in a previous HTS run using the original immunocapture protocol.

4.2.3 High throughput sequencing

From each sample, 5 µL barcoded PCR was pooled into a tube according to its vineyard of origin, resulting in four tubes to each generate a library for the SB group; one library was made of the 35 samples from GC (including both LH and HH components) and another from 24 DR samples. The pooled samples were cleaned using Agentcourt Ampure XP (Beckman Coulter Inc.) following the manufacturer's protocol (40 µL at a time). Samples were sent to the Australian Genome Research Facility where four libraries of SB were sequenced on one run, and the two libraries, one from each GC and DR, were sequenced on a second run (alongside an additional unrelated library) on a HiSeq 2500 sequencer (125 bp paired-end).

4.2.4 Bioinformatic pipeline

Raw read quality was checked with `FastQC`. Pooled Illumina reads were deconvoluted using `fastq-multx` with the barcodes for each sample. Individual samples' reads were checked again with FastQC, which informed trimming and filtering strategy. Reads were hard-trimmed at the 5' end with `fastx_trimmer`, then filtered on base and overall read quality with `fastq-mcf`. A final post-processing quality check was performed, and filtering statistics were summarised. Cleaned samples were re-pooled. Host genome (*Vitis vinifera*) associated reads were removed via short read mapping, yielding the non-host reads for use in viral assembly. The viral reads from each pool were assembled with `Trinity`, a `Quast` report was produced for the assemblies. Viral contigs were identified by Blastn search, and BlastX against virus database for the larger contigs (>1 kb). The list of virus identified was used to make the reference sequences used for the mapping step, in addition with some important virus sequences not identified (including the nepoviruses and DNA viruses recently identified). Reads were mapped to all the assembled viral sequences using `bowtie2` short read mapper. The bioinformatics analysis is available in the appendices Figure 7.4. Alignments were summarised, counted. A counts summary table was also produced. The small RNA data pre-processing was performed using the YABI VSD (Barrero et al., 2017) and mapped on the same reference using the `bowtie2`.

4.2.5 Cut-off

A cut-off was set for each virus group. The cut-off values are defined in Table 4.1: Cut off value used to distinguish the positives in each of the virus detected. The row are expressing

the cut-off value in read mapped per million (RMPM), the average RMPM for the samples considered positive [Mean (+)], the number of positive samples [n (+)], and the standard deviation between the positive samples [SD (+)]; the average RMPM for the samples considered negative [Mean (-)], the number of positive samples [n(-)], and the standard deviation between the positive samples [SD(-)]. with the average number of mapped reads of the samples considered positive, and the standard deviation; the same information is also presented for the samples considered negative.

4.2.6 New viruses

The sequence of a novel vitivirus sequence was confirmed by RT-PCR of resample tissue (August and Decembre 2017) and by PCR of the cDNA prepared previously. Primers used were GVE-li_4543F (GCA CAT CAT GAT CTT TTC ACA G) and GV-li_5212R (ATC ACT ATC TTC CTA ATC AAC TCT T); and GVE-li_4543F with GVE-li MP-R (GGA TGT TTT GAG CAC CGT TGA); GVE-li_MP-F (TGT GGG CTG CAT AGC AGT AG) with GVE-li_CP-R (AAC AGT AAC ACC TAC CAT ACC T); and GVE-li_NABP-F (CAA GGC CAA AAT CTA TAC TAG C) with an anchored oligodT primer (GAT TTA GGT GAC ACT ATA G T_{17}V).

The presence of Grapevine geminivirus A (GGVA) was confirmed with the primer pairs GGVAv541 / GGVAc924 and GGVAv2097 / GGVAc239 (Al Rwahnih et al., 2016a) from a DNA extraction (Qiagen DNeasy) of resampled leaf petiole and veins (February 2018).

4.3 Results

4.3.1 Data output

The SB group comprised a total of 177 samples prepared on a vineyard basis and sequenced in four libraries (SB-1, SB-2, SB-3 and SB-4) through one HiSeq 2500 run. After data quality trimming, and demultiplexing, the number of reads per sample varied between 292 and 679,414 reads. Eleven samples with less than 60,000 reads were removed from the analysis, leaving 166 samples with between 67,704 and 676,414 reads each and an average read number of about 350,000 reads; this comprised 44 samples in SB-1; 43 in SB-2; 40 in SB-3 and 39 in SB-4. From the other HiSeq 2500 sequencing run the GC environment (a total of 35 samples from GC-LH and GC-HH) yielded between 345,314 and 1,270,868 reads (average of ~640,000 reads) and the DR environment (24 samples) gave between 331,682 and 1,785,984

reads with an average of ~862,700 reads. From the sRNA run, all the libraries yielded more than 27 M reads with the exception of one sample (sample AB539 with 19 M reads), making an average of 33 M between the 18 samples.

4.3.2 Virus detection

Regardless of the environment, no viruses were detected in 48 samples (21%), one virus only in 124 (55%), two viruses only in 23 samples and three or more viruses were detected in 30 samples. However, the virus load per plant varied greatly between the four categories (Figure 4.4). The SB environment being the most homogeneous environment with 87.3% of the plants showing no (24.7%) or only one virus infection (62.7%). As expected, the opposite was observed from the low-health germplasm (GC-LH) with only 10.5% of the plants with either no (1 plant) or one (1 plant) virus infection, and 84% samples (16 plants) with four or more viruses (up to seven).

A total of 17 viruses were detected from the different environments. The viruses identified can be grouped into five categories, i.e. GRSPaV; the four leafroll viruses belonging to the family *Closteroviridae*; the five fleck-like viruses from the family *Tymoviridae*; the members of the genus *Vitivirus;* and the DNA virus GGVA.

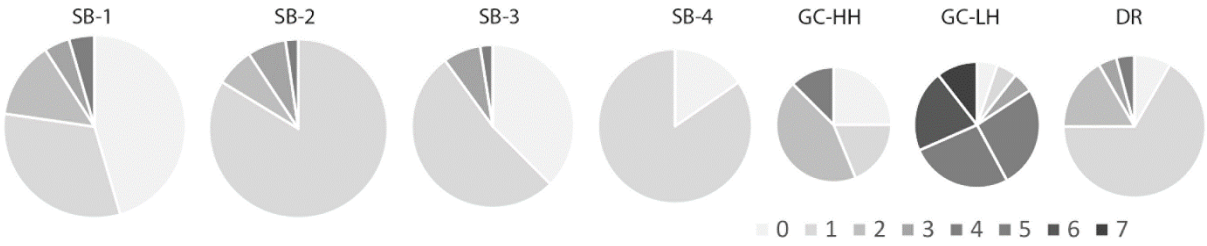


Figure 4.4: Virus load in each of the seven groups. The lighter colour represent the plants with no virus detected, the darkest represent plant infected with seven virus species. Size of the pie chart is correlated to the number of samples in each group. SB-1 to SB-4 are the commercial Sauvignon blanc samples (n= 44, 43, 40 and 39 in SB-1, -2, -3 and -4 respectively); GC-HH represent the samples collected in the high health block of the germplasm (n= 16) and the GC-LH represent the low-health block (n= 19). The DR group represent the samples collected/submitted because of symptoms (n=24).

Multiple contigs obtained showed homology to dsRNA viruses, including members of the family totiviridae, or unclassified dsRNA viruses. These viruses were not included in the results analysis and were considered to originate from fungal contributions within the plant samples. Sequences related to the Mitoviruses similar to Sclerotinia sclerotiorum mitovirus 2 and 4 were detected but also considered to be of fungal origin and not included here.

The detection of Hop stunt viroid (HSVd) was extremely common but resulted in highly variable mapped read numbers. The vineyards SB-1; -3 and -4 group showed a very high rate of mapping to HSVd with an average of 3.5% total reads; in SB-2, an average of 0.4% total reads mapped to HSVd but only 7 samples were above 1% reads mapped. From the second sequencing run (GC and DR samples), only six samples (four GC-HH and two DR) showed very high levels of mapping to HSVd (> 2%) and the rest were below 1%. Grapevine yellow speckle viroid (GYSVd-1) was less common with only one sample in the SB run with more than 2% reads mapped (sample SB-1-130) while the rest showed less than 0.4% reads mapped to GYSVd-1.

4.3.3 Grapevine ruspestris stem pitting virus (GRSPaV)

The virus GRSPaV was detected in 100 plants, which represents 44% of the samples. The virus was detected from each of the four categories but only in two of the four SB vineyards; GRSPaV was not detected in SB-3 and SB-4. It was detected in all the plants from SB-2 and 32% from SB-1. The virus was detected in 63% in GC-LH, 69% plants tested in GC-HH, and 83% in DR. Multiple strains of GRSPaV were detected, often in mixed infections. The majority of the reads mapped to the GRSPaV-SY strain (isolates AY368590.1 and KX274275), but some samples mapped GRSPaV-MG, JF. LSL, and SK704C.

4.3.4 Leafroll viruses

The leafroll group is defined by the viruses detected from the Closteroviridae and is made of the ampeloviruses GLRaV-1, -3, -4, and the closterovirus GLRaV-2. GLRaV-1 was only detected in two plants from the GC-LH block (11%), both plants being from a chardonnay cultivar. GLRaV-4 was also only detected in two plants, a Sauvignon blanc (AB554) in GC-HH, and in a interspecific hybrid (Seibel 10096; AB536) in GC-LH. Both samples harbour a different strain of GLRaV-4; one GLRaV-4 strain similar to isolate LR106 (mapped to FJ467503) was detected

in the sample AB554 while GLRaV-4 strain 5 (mapped to JX513893) was detected in sample AB536.

GLRaV-3 was detected in all the environments but only in three plants from two of the four SB vineyards (one in SB-1 and two in SB-3).The virus was found in the germplasm with an incidence of 25% in in GC-HH and 74% in GC-LH. One quarter of the DR plants were found positive for this virus. Different strains of GLRaV-3 were detected, 16 GLRaV-3 positive plants had strains belonging to group I, 18 from group VI and 17 from NZ2. A single strain of GLRaV-3 infection was detected in only eight plants from a total of 27, nine were found to have the three strains of the virus (all from GC-LH).

The virus GLRaV-2 was detected in four plants from the North Island SB vineyards, three in SB-1 and one in SB-2. The virus was detected in more than half the plants sampled in GC-LH (53%), in one plant from the GC-HH, but was absent from the DR samples. The three positive plants in SB-1 were found in the same row and they shared the same virus genotype, with 98% nt identity to isolate BD (DQ286725) with 96-98% coverage. A plant infected with GLRaV-2 in vineyard SB-2 harboured a strain with a match to isolate OR1 (FJ436234) with 99% genome coverage and 99% nt identity. The GLRaV-2 positive samples from the GC-LH and GC-HH were all infected with a single isolate related to isolate OR1 (FJ436234) with the exception of sample AB536 that had, in addition to OR1 isolates (95.9% coverage, 99% nt identity), reads that mapped to the isolate 93/955 (99% genome coverage; 99% nt identity, AF039204).

Table 4.1: Cut off value used to distinguish the positives in each of the virus detected. The row are expressing the cut-off value in read mapped per million (RMPM), the average RMPM for the samples considered positive [Mean (+)], the number of positive samples [n (+)], and the standard deviation between the positive samples [SD (+)]; the average RMPM for the samples considered negative [Mean (-)], the number of positive samples [n(-)], and the standard deviation between the positive samples [SD(-)].

Virus	GLRaV-1	GLRaV-3	GLRaV-4	GLRaV-2	GRVFV	GSyV-1	GAMaV	GRGV	GFkV	GRSPaV	GVA	GVB	GVD	GVE-like	GVG	GVI
cut-off value	6000	6000	6000	6000	1000	4000	1000	1500	1000	5000	1000	5000	500	5000	2000	2000
Mean (+)	16008	368507	76235	90241	10381	22466	1257	19614	24047	198990	9072	45690	849	6277	4180	13438
n (+)	2	27	2	15	19	72	1	17	6	100	13	1	1	1	15	9
SD (+)	1197	277842	43647	90978	8182	24891	-	23866	19192	156865	5987	-	-	-	1821	9766
mean (-)	2	357	13	46	22	405	3	47	14	431	17	0	0	5	11	6
n (-)	223	198	223	200	206	153	224	208	219	125	212	224	224	224	210	216
SD (-)	8	820	39	127	61	844	37	118	37	618	83	3	2	17	56	20

Table 4.2 Percentage of plant infected with one of the 17 viruses detected in the seven different categories: Sauvignon blanc vineyards in Hawke's bay (SB-1 and SB-2) or Marlborough (SB-3 and SB-4), samples received mostly symptomatic (DR) and samples collected from the germplasm high health (GC-HH) and low health block (GC-LH). The number of samples per category is listed in column n.

	n	GLRaV-1	GLRaV-3	GLRaV-4	GLRaV-2	GRVFV	GSyV-1	GAMaV	GRGV	GFkV	GRSPaV	GVA	GVB	GVD	GVE-like	GVG	GVI	GGVA
SB-1	*44*	0%	2%	0%	7%	14%	20%	0%	9%	7%	32%	0%	0%	0%	0%	0%	0%	0%
SB-2	*43*	0%	0%	0%	2%	7%	7%	0%	5%	7%	100%	0%	0%	0%	0%	0%	0%	0%
SB-3	*40*	0%	5%	0%	0%	0%	60%	0%	8%	0%	0%	0%	0%	0%	3%	3%	0%	0%
SB-4	*39*	0%	0%	0%	0%	0%	85%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
DR	*24*	0%	25%	0%	0%	0%	0%	0%	4%	0%	83%	8%	4%	0%	0%	4%	0%	0%
GP-HH	*16*	0%	25%	6%	6%	6%	13%	6%	25%	0%	69%	0%	0%	0%	0%	0%	0%	0%
GP-LH	*19*	11%	74%	5%	53%	47%	5%	0%	16%	0%	63%	58%	0%	5%	0%	68%	47%	5%

4.3.5 Fleck-like viruses

The marafivirus virus grapevine syrah virus-1 (GSyV-1) was the second most common virus detected with an overall incidence of 32% (72 plants). The virus was very common in SB with incidence of 7% (SB-2), 20% (SB-1), 60% (SB-3) and 85% (SB-4) but was more scarce in the germplasm (one plant GSyV-1 infected from the GC-LH and two plants from the GC-HH) and was not detected within the DR samples. Grapevine rupestris vein feathering virus (GRVFV, marafivirus genus) was detected in SB-1 and SB-2 with incidences of 14 and 7%, respectively, but not in SB-3 and SB-4. GRVFV was detected in 47% plants from the GC-LH but only one plant in the GC-HH and not detected in the DR environment. Grapevine red globe virus (GRGV) is tentatively assigned to genus maculavirus. It was not detected in SB-4 and was found in low numbers (4-16%) in all the other libraries except for GC-LH where the incidence of this virus was higher (25%). The namesake virus of the fleck-like group, grapevine fleck virus (GFkV, marafivirus), was the first virus characterised of these five. However, in this study GFkV was only detected in three vines in SB-1 (7%) and three in SB-2 (7%). GFkV was not detected in the GC-LH, GC-HH or DR environments. Grapevine asteroid mosaic-associated virus (GAMaV), another virus from the marafivirus genus, was only detected in one plant from the GC-HH group. This sample was detected from a *de novo* assembled contig similar to GAMaV isolate GV30 (KX354202.1) with 90-96% nucleotide identity.

4.3.6 Vitivirus

From the vitivirus genus, GVA was only detected in the GC-LH and DR libraries. In GC-LH (low health), the virus was detected in 58% of the plants tested. In the DR environment, the virus was detected in two samples. GVB was only detected in one plant (Pinot gris) from the DR environment. The GVB sequence obtained from this sample matched the divergent Chinese isolate GVB-QMWH (KF700375.1) with 96% nt identity. The same plant was also infected with GVA, GLRaV-3 and GRSPaV. GVD was detected in only one plant, from the GC-LH, but only by mapping contigs obtained by *de novo* assembly instead of the GVD genomes available (MF774336 and MF072319). Two vitiviruses were detected from the same plant (SB-3-169) within the SB environment, a variant strain of GVE (termed GVE-like) and GVG. GVG was also detected in 13 additional plants from the GC-LH (68%) and one plant from the DR environment. The recently characterised GVI was detected in nine plants, all from the GC-LH environment (47%).

4.3.7 New viruses

4.3.7.1 GVE-like

The genome of the GVE-like virus was analysed and compared with other vitiviruses. The partial ORF1 (4907 nt) codes for a polyprotein that contains common vitivirus conserved domains (viral methyltransferase; viral helicase; 2OG-Fe(II) oxygenase (alkB domain); a RNA dependant-RNA-polymerase (RdRp)); a carlavirus endopeptidase (C23 pfam05379) was detected between the methyltransferase and the helicase. The closest match on Genbank for ORF1 is GVE isolate SA94 with 68% nt identity (71% aa). The ORF2 is 528 nt long and codes for a putative protein of 175 aa with only 40% aa identity to GVE. The third ORF is 846 nt long (281 aa) and has a recognised viral movement protein domain. The movement protein of GVE is the closest sequence present in GenBank with 73% nt identity (72% aa). The ORF4 is 600 nt long (199 aa) and codes for a recognised coat protein domain. GVE coat protein (CP) is the closest sequence available with 74% nt identity (79% aa). The last ORF, with 354 nt codes for a 117 aa protein with two recognised overlapping domains, a viral nucleic acid binding domain and the Citrus tristeza virus p23 protein. This short sequence shares 85% of its nucleotides sequence, 81% aa with the last ORF/protein of GVE (AB432910). Phylogenetic analysis of both the ORF1 (RdRp) and the CP genes (Appendix Figure 7.2 and Figure 7.3) shows that the GVE-like virus is a divergent strain of GVE.

4.3.7.2 Grapevine geminivirus A

The geminivirus GGVA was detected in one plant (AB537) by mapping 163,611 small RNA reads to the GGVA genome (KX618694) (99.7% genome coverage; 99.1% nt identity). This represents 5017 reads mapped per million reads sequenced. The virus was not detected in the other samples. From the dsRNA HTS only one read and its pair from sample of AB537 mapped to the GGVA genome however these did share 100% homology to that reference sequence. No reads mapped this virus from the additional 224 samples obtained from dsRNA extraction. The sequence of GGVA obtained was confirmed by RT-PCR and Sanger sequencing.

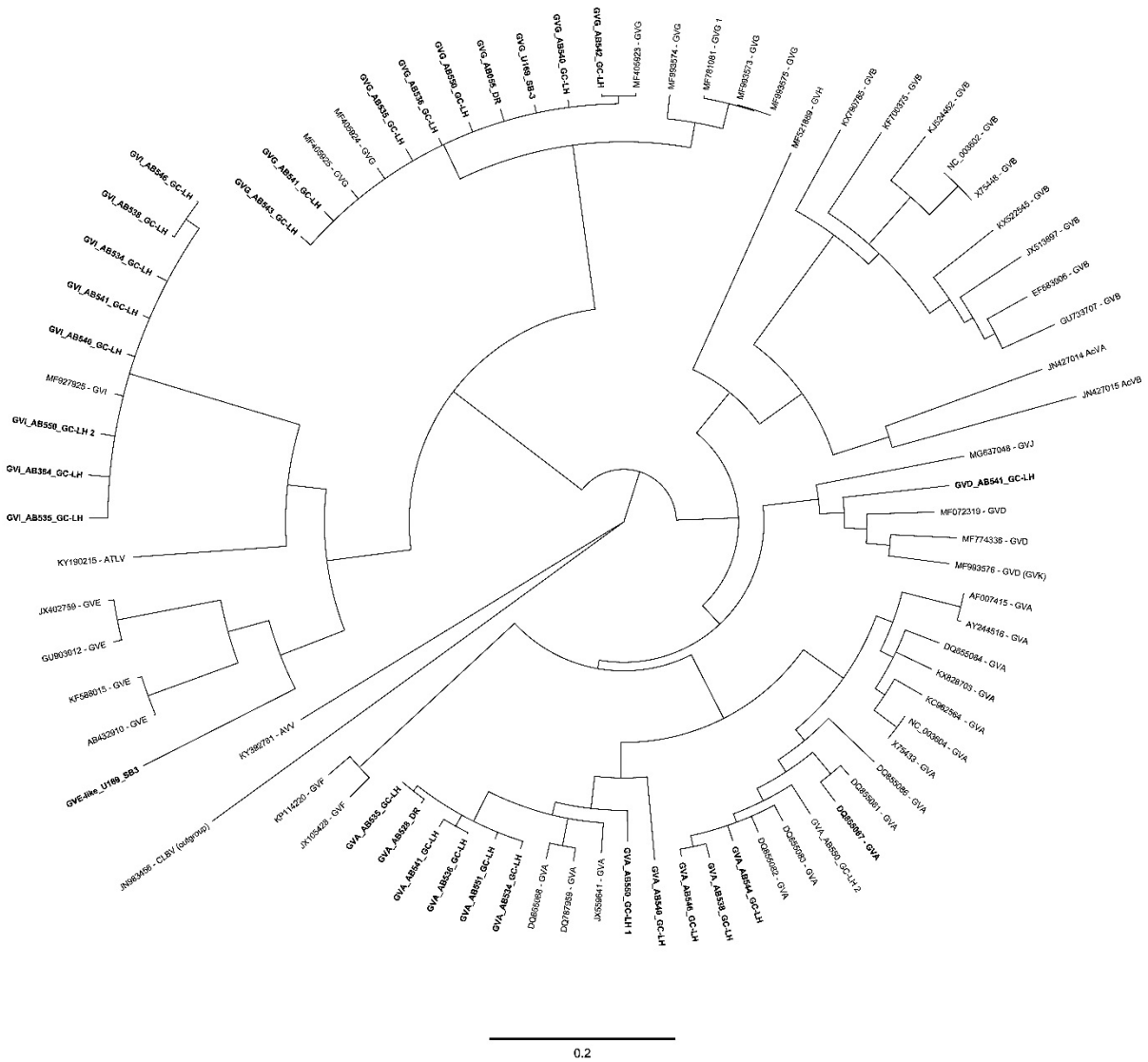


Figure 4.5: Neighbor-Joining tree (1000 bootstrap replicates using the Jukes-Cantor distance model) of the coat protein of representative members of the genus Vitivirus including 31 samples obtained by mapping from this study when the region was covered (12 GVA, 1 GVD, 1 GVE-like, 9 GVG and 8 GVI). Nucleotide alignment (extracted from the accession number indicated) was performed with ClustalW (BLOSUM cost matrix with a gap opening cost set at 10, and a gap extension cost at 0.1). Citrus leaf blotch virus was used as the outgroup.

4.3.8 Small RNA HTS validation

Sixteen samples from the GC-LH were sequenced by small RNA HTS. Only samples AB384, AB546, and AB549 were not sequenced by small RNA HTS from the GC-LH samples. The sRNA reads were mapped onto the same reference sequences. The viruses identified were identical for all the samples described above with the exception of the viruses from the fleck-like group. The sRNA data was used to establish the cut-off point for the leafroll and the vitivirus groups (Appendix Table 7.2).

By small RNA HTS two additional GRVFV positive plants were detected (AB535; AB536) and three samples positive by dsRNA were yielded relatively low sRNA count; one additional GSyV-1 positive plant was detected (AB551) but the GSyV-1 positive plant identified by dsRNA (AB544) had low sRNA reads mapped. Sample AB543 mapped relatively low number with both methods. For GRGV the same results were obtained using both methods but one positive sample of AB541 was only detected from the top canopy sample and had low yield by the dsRNA HTS method (below the cut-off value); conversely the sample AB542 that was positive for GRGV by dsRNA HTS and yielded low number of read mapped by the sRNA HTS. Two GFkV samples (AB535 and AB536) were found positive by sRNA HTS with mapped reads from dsRNA HTS below the cutoff value.

The results from the three samples sequenced from AB541 revealed that the viral sRNA yield decreased drastically at the top of the canopy for all the viruses detected (GLRaV-4, GRSPaV, GVA, GVD, GVG and GVI) except for the fleck-like viruses where the yield was higher for GRVFV and GRGV in the top canopy leaf. The same observations of top canopy leaves providing more abundant reads was made for the viroid reads.

4.3.9 Additional validation

From the 18 samples of the GC-LH, the status of GLRaV-1, -2 and -3 confirmed previous ELISA indexing. All the samples positive for the vitiviruses GVD, GVG and GVI were confirmed by RT-PCR. The SB samples were tested by RT-qPCR for the presence of GRSPaV. Out of 55 samples with more than 5000 reads mapped to GRSPaV, three samples were not detected (no Ct value or Ct > 36). Out of 111 samples with less than 5000 reads mapped to GRSPaV, two samples had a Ct < 36. The qPCR data and the sRNA HTS results were used to set the cut-off value at 5000 reads for GRSPaV. The cut off threshold for the GSyV-1 was set by RT-PCR. Six samples from the SB category with reads mapping to GSyV-1 (>4000 reads per sample)

were all positive. From 21 samples with less than 4000 reads mapped to GSyV-1, 15 tested negative and six positive.

4.4 Discussion

4.4.1 Method and approach

The new sequencing technologies available today allow the detection of all viruses present in the plant. By using the immunocapture of dsRNA we have enriched for viral RNA in a relatively simple and rapid method from grapevines, a tissue that typically only harbours low virus titres. From 2.5 g of tissue, the extraction was completed in about 6 hours. The ratio of viral and viroid reads that mapped to our reference sequences ranged per sample between 0.3% and 87.4% (Table 7.1) with an average of 17%. From the 16 samples sequenced by sRNA HTS, this ratio was between 0.2% and 33.6%. By comparison, we have obtained total RNA sequencing data from sample AB541 and AB542 as part of another study; the viral read represents respectively 0.11% and 0.02% (without ribosomal depletion). The same two samples yielded 75 and 69% viral RNA reads by dsRNA and 16 and 26% viral RNA by sRNA HTS. This illustrates the power of the viral sequence enrichment methods for the detection of the RNA viruses and as a consequence the reduced sequencing depth required (by 100- to 1000-fold per sample) to detect the same viruses.

Due to the notorious ubiquity of GRSPaV, it was used as a target for assessing the quality of the dsRNA extraction but this assumption was challenged by the absence of GRSPaV in two vineyards. A negative correlation could be observed between the incidence of the GSyV-1 and GRSPaV with higher GRSPaV read numbers observed in SB-2, lower incidence in SB-1, and no detection in SB-3 and SB-4. Considering the demonstrated incidence of both GRSPaV and GSyV-1, a triplex assay should be developed to target these two very common viruses for use as an “internal control” for dsRNA extraction alongside with the rRNA probe already used in this study. With the cut-off value selected in this study such a duplex assay would have been positive for 72% samples.

Using the dsRNA antibody method adapted for grapevines, for the first time worldwide we have been able to assess the virome of a country’s vineyard using an untargeted approach for virus or viroid sequencing. We have compared the virome obtained in four distinct categories: (1) four homogeneous vineyards planted in Sauvignon blanc in two different

regions that depicted the background virome of the New Zealand vineyard (2) grapevines from an old germplasm collection from a low heath block and (3) a high health block that together illustrated the origin of the viruses to New Zealand (4) samples collected from commercial vineyards mostly from symptomatic grapevines that explored potential viral-diseases.

4.4.2 Cut-off

At the cost of sensitivity a conservative approach was chosen to reduce the chance of calling false positives. The opposite approach would have been taken in a phytosanitary assay in order to avoid missing samples that are genuinely positive for viruses. For instance, in their comparison of HTS and biological indexing, Al Rwahnih and colleagues detected positive GVA with less than 30 reads (Al Rwahnih et al., 2015b). Viruses with low homology to the GenBank accession are the hardest to detect. The samples infected with GAMaV and GVD are in this hard-to-match category where the mapping was only detected on the contigs obtained by *de novo* assembly (GAMaV) or by RT-PCR (GVD). Due to the limited length of the reference genomes available, the number obtained were low. Similar observations can be made of GFkV samples that were below the positive threshold in the GC-LH yet were positive by sRNA HTS; the GFkV reads from AB535 only mapped contigs sharing less than 90% nt identity with the reference sequences form GenBank but it did not map the GFkV genomes (AJ309022 and AJ309022).

Although the results obtained from the sRNA HTS are not directly comparable as they were sampled from different tissues and at different sampling dates, the two strategies gave the same results for all the viruses except those viruses in the fleck-like group. The disparities in detection of fleck-like viruses may be due to their uneven distribution within the plant as indicated by the three different canopy level samples assessed by small RNA HTS. For instance, the results of the sRNA concentration (reads per million) in sample AB541 reveals more fleck-like virus in the top of the canopy during the summer. This contrasts with the known distribution of GFkV that is reported to be consistent all year except for the petioles in late summer (Fiore et al., 2009). The results between the sRNA and dsRNA HTS suggest that the cut-off value used in this study was overly high and thus the current report underestimates the incidence of these viruses. The selected cut-off was maintained constant between libraries where the cross-talk between samples varied depending on the total number of reads present for one virus. A high total can be achieved when the majority of the samples are infected (e.g.

for GRSPaV or GSyV-1) or when a few samples yield extremely abundant reads (e.g. GLRaV-3). This would suggest a different cut-off selection for each library.

Perhaps a single time-point, is insufficient to provide full confidence in the virus status of a plant. As more viruses have been identified using more generic methods (i.e. untargeted) that also provide improved sensitivity, the attention to virus-host biology has decreased. Obviously the best tissue and time selection seems to be opposite between the fleck-like viruses and the leafroll viruses. In this study, the sRNA samples were made of multiple mature leaves from the same part of the canopy and the dsRNA samples made of two or three canes from dormant wood; note that sample AB541 was the exception as it was sampled from three different canopy locations for the sRNA HTS. The sample size used by Al Rwahnih and colleagues for dsRNA extraction is 30 g of scrapping, which suggests a large amount of cane from different parts of the vine (Al Rwahnih et al., 2015b). Likewise, when sampling fruit trees Rott and colleagues pooled two dsRNA time points (Rott et al., 2017). The study supports these two approaches to provide high levels of confidence of the virus status, especially in a phytosanitary assay with consequence on movement or propagation of the plants tested. However for survey purposes the current approach was sufficient as every significant finding was confirmed by an alternative method.

4.4.3 Origins of the viruses in New Zealand

The high virus load detected in the germplasm collection is explained by its historical origin. Most of the vines deposited in this collection originated from the Te Kauwhata Viticultural Research Station, a national reference collection located in the Waikato region (North Island) that was initially established in the late 1800s then increased through imports of new grapevine cultivars. Over the ensuing decades, the cultivars were assessed and the best were distributed to New Zealand grape growers. Some of the accessions can be traced back to Romeo Bragato, one of the pioneers of modern New Zealand viticulture practices. The Te Kauwhata collection was not maintained *in situ* after the 1980s and was moved to different locations before being incorporated to the New Zealand Winegrowers germplasm collection in its current location, Lincoln (South Island). Lincoln is a region that is regarded as having low pressure from the insect vectors of viruses, in particular mealybugs. The plants are now self-rooted. Some of the plants, treated by thermotherapy to remove virus infections, were moved to an isolated high-health block (GC-HH) alongside the most recent imports that have been screened through post-entry quarantine processes (1980s onward). From the 16 GC-HH plants sampled an average of 1.6 viruses per plant were detected, as opposed to 4.6 virus infection

per plant sampled in the low-health (GC-LH, 19 plants). Thus, the terminology used to name those two blocks is justified (Figure 4.4).

If we include GFkV that was only detected by sRNA HTS, a total of 15 different viruses were detected from the GC-LH or GC-HH germplasm collection, proving that these older vines provide an insight into the full extent of virus diversity in grapevines within New Zealand. GVB and GVE were the only two viruses not detected in germplasm, but instead detected in SB or DR environments. Importantly, only a small fraction of the germplasm collection was sequenced (39 out of a total of 953 accessions) and it is likely that these two viruses and perhaps others are also present in those non-tested plants within germplasm collection. Indeed, corky bark, a disease caused by GVB, was reported from this collection in 1985 (Smart, 1985). Overall, the range of viruses detected in the collection is comparable with those reported by Al Rwahnih et al. (2015) where they compared the detection of viruses by HTS and biological indexing. From 15 samples tested they identified had only one plant with no virus (only one viroid) and a virus load of 4.6 viruses per plant (excluding the viroid). A total of 15 viruses were identified, the main difference is the presence of some nepoviruses, Grapevine Cabernet Sauvignon reovirus and the DNA virus GRBV as well as the absence of the vitiviruses GVG and GVI and the geminivirus GGVA which are new-to-science viruses reported since 2015; very similar infection rates were obtained for GLRaV-2, -3 and GRSPaV (Al Rwahnih et al., 2015b).

Predictably for an insect-vectored virus, GLRaV-3 was detected at a high incidence in the GC-LH (74%). The three strains of the virus known to be common in New Zealand were detected in similar proportions. The large number of multiple GLRaV-3 strain infections (nine plants with triple infection) reflects the long exposure time and/or the virus accumulation by grafting. GLRaV-1 movement under the New Zealand environment is rare or absent and the presence of two GLRaV-1 positive plants can be best explained by vegetative propagation. They are likely to originate from the GLRaV-1 infected Chardonnay Mendoza imported into New Zealand in 1971 (Hoskins & Thorpe, 2010a). Although, their virus infection profiles suggest that the two plants (AB543 and AB544) took two independent routes to the germplasm collection as evidenced by distinct profiles of co-infecting viruses. Four plants from the GC-HH had no virus detected in them (25%). Viruses detected in the high-health block include GLRaV-2, GLRaV-3 and GLRaV-4 and the four fleck-like viruses (GRVFV, GSyV-1, GAMaV and GRGV).

The closterovirus GLRaV-2 is not known to have a vector, so its high incidence in the GC-LH collection can only be explained by the vegetative propagation including top-working. Ten plants of the GC-LH (53%) and one in the GC-HH (6%) were found GLRaV-2 positive. The

high level of this virus was also observed in other virome study of germplasm in the USA (53%) and in Brazil (40%) (Al Rwahnih et al., 2015b, Fajardo et al., 2017). In our study, most sequences obtained are closely related to the OR1 isolate (FJ436234) similar to the strain PN, one of the original GLRaV-2 genomes identified worldwide (Angelini et al., 2017). It is probable that this virus was introduced in New Zealand on imported Sauvignon blanc clone 316 and clone 317 in the mid-1980s (Hoskins & Thorpe, 2010b). The same strain of GLRaV-2 was detected in the SB-2 vineyard but the three positive vines from the same row of SB-1 were infected by a virus with high homology to the BD strain. The BD strain of GLRaV-2 was already described in New Zealand under the isolate named 'Alfie' (Bonfiglioli et al., 2003); this strain was not detected in the germplasm samples.

The diversity of vitiviruses detected in the GC library (comprising GC-LH and GC-HH samples) is novel and originated exclusively from the samples from the low-health block. With 13 plants infected, GVG is the most commonly detected vitivirus. It was also detected in the commercial vineyards SB-3 and the DR environments. GVI too, was detected frequently with nine plants positive, all of which were also co-infected with GVG and GVA. GVD was only detected once, in a plant co-infected with three additional vitiviruses (GVA, GVG and GVI). GVD and GVI were only detected in the GC-LH samples. GVG and GVI were detected with low genetic variability but the GVA sequences obtained were highly variable (Figure 4.5). Overall, a high level of co-infection was also observed in the vitivirus-infected grapevines. In total, vitiviruses were detected in 18 plants; all these plants were also infected by GLRaV-3. This observation supports the hypothesis that vitivirus replication is enhanced in the presence of a closterovirus (Rowhani et al., 2018).

4.4.4 Virome background (SB environment)

A HTS survey of viruses in vineyards of this scale has not been previously reported. In South Africa, 44 grapevine samples from a symptomatic vineyard were pooled together and sequenced however the virus incidence could not be assessed (Coetzee et al., 2010). Despite this drawback, four viruses were detected from these symptomatic vines: GLRaV-3, GRSPaV and the vitiviruses GVA and GVE (disregarding the mycoviruses). It was the first report of GVE in South Africa. Coincidentally, this current study reports a virus similar to GVE for the first time in New Zealand. However, the incidence of GVE appears to be very low in New Zealand, as opposed to reports of widespread incidence in South African (Jooste et al., 2015).

The SB vineyards were each found to be homogeneous; SB-4 vineyard is most uniform with GSyV-1 the only virus detected. Across the four commercial Sauvignon blanc vineyards, the virus pressure was limited with 88% of the plants hosting either zero or only a single virus (Figure 4.4). The most common virus detected from the four SB blocks was the GSyV-1 especially in SB-1; SB-3 and SB-4. These three vineyards with high GSyV-1 incidence all use the rootstock SO4 while SB-2 that used the rootstock Schwarzmann.

GRSPaV is considered the most prevalent virus in grapevines worldwide (Meng et al., 2017), therefore it is unexpected to not detect the virus within two vineyards. The two vineyards are from the same region, and both planted in 2006. It is possible that they used the same source material, although the rootstock used was different. The incidences of GRSPaV detected in the GC-LH, GC-HH and DR samples varied between 63 and 83% which is more in line with previous reports (Meng et al., 2017). When assessing the strain of GRSPaV detected, the GRSPaV-SY was by far the most common in all the environments. However, we found multiple strains in most samples in line with a recent study in Slovakia (Glasa et al., 2017). Between the 225 samples of this study, only the GRSPaV isolate Tannat-Rspav1 (KR528585.1) was not present (no individual sample showed more than 26% coverage to that isolate by mapping) with the other 22 reference genomes mapping with more than 60% coverage to one or more samples.

4.4.5 Potential viral-disease

The very heterogeneous DR group was made from different cultivars, vineyards and plant ages. Due to the nature of the samples (disease response), it was expected to detect pathogenic viruses such as GLRaV-3 would be detected (25%). The multiple samples showing early reddening, different to the GLRaV-3 symptoms, did not detect any candidate virus that might account for the symptoms. GRSPaV was present in most samples except for AB056 where no viruses were detected. The reddening may be a biotic or abiotic stress response from the plant, with no virus associated in some cases. Three species of vitiviruses were detected GVA, GVB and GVG in three plants from the DR group, a Pinot gris infected with GVA and GVB (AB044), GVG in a Gewürztraminer (AB055) and GVA in a Syrah (AB528). GVA is the causal agent of Syrah disease in Australia and South Africa, it is therefore interesting to find GVA in a plant showing early reddening (Goszczynski & Habili, 2012, Habili et al., 2016). The DR group Syrah sample was infected with GVA belonging to group I (Figure 4.5) that is distinct from the group II associated with Syrah disease (Goszczynski & Habili, 2012). In addition, several Syrah samples from the same vineyard (AB526 to AB533) were compared but only

AB526 to AB530 showed the symptoms and GVA was only detected in one symptomatic samples. A caveat is that the virus was not systemic at the time of sampling however such a virus distribution is unlikely to be associated with the symptoms. It is important to note that the DR samples were prepared from leaves, as opposed to cane phloem scraping used in the SB and germplasm samples. The Pinot gris infected with both GVA and GVB did not display unusual symptoms. The fleck-like viruses were fairly rare in the DR group compared to the SB with only one infection detected for each of GRVFV, GSyV-1 and GRGV.

4.4.6 New to New Zealand viruses

The GVE-like virus reported here represents to date the only GVE present in New Zealand. It is different to all previously described GVE sequences and is on the cusp of a being defined as a different species. The threshold for species demarcation within the genus Vitivirus is set at 80% aa identity or 72% nt identity for the coat protein and the RdRp (Adams et al., 2004). The GVE-like virus described in this study is 66% nt (71% aa) to the closest replicase (GU903012), and 74% nt (79% aa) in the coat protein. Despite the distance, the phylogenetic analysis support GVE classification as it places the virus described here on the same clade as those GVE sequences previously described (Figure 4.5; Figure 7.2 and Figure 7.3). No apparent symptoms were recorded (Figure 4.3). This detection demonstrates the benefit of the untargeted HTS technologies for surveying viruses and the detection of unknown viruses.

GGVA was detected in one of the GC-LH samples not showing any symptoms at the time of collection (Figure 4.1). The plant infected by the virus is an interspecific *Vitis* hybrid named Siebel 7052 with an original accession of TK00184 probably imported in New Zealand from the USA in 1957 alongside the grapevines Siebel 6339 (TK00183) and Siebel 10096 (TK00188). The virus was clearly identified by the sRNA HTS but was missed by the dsRNA HTS, although one read and its pair mapped the genome perfectly (no other GGVA reads were identified from any other sample). The dsRNA antibodies used for the enrichment are specific to dsRNA and do not bind to any other form of nucleic acid (O'Brien et al., 2015). Previous publications have reported the detection of DNA viruses by dsRNA HTS. The difference here may be due to the cellulose protocol being not as specific as the mAb, or the amount of the dsRNA potentially formed by the overlapping RNA transcript being extremely low and therefore detected only when sequencing with more depth. In his publication, Rott et al., used an average depth between 6 and 10 M reads, which is more than 20 fold the depth used in this report (Rott et al., 2017). GGVA is the first DNA virus detected in New Zealand grapevines.

4.5 Conclusion

This study has taken advantage of an antibody-based enrichment for untargeted virus detection by HTS to update the virome of New Zealand vineyards. The diverse origins of the samples draw a picture of the country's current virus diversity. The fleck-like viruses are very common in the vineyards and the high incidence of GSyV-1 may be linked with the SO4 rootstock. This virus is not known to cause any disease or symptoms within the vineyard and its high incidence in healthy-looking grapevines endorses the notion that it is not detrimental to the plant. As expected, the virome from commercial vineyards is relatively homogeneous, comprising the fleck-like viruses and GRSPaV. The stable nature of this virome is explained by the uniform planting material resulting in a distinctive virus inoculum. In addition, the limited number of viruses capable of movement, beside vegetative propagation, under New Zealand environment increases this stability. This contrasts with the high virus load and diversity observed in the old germplasm. Similar observations and comparable viruses were detected in a recent study from South Africa between young and old grapevines (Oosthuizen, 2017). The overall incidence of GRSPaV was high but surprisingly absent from two commercial blocks. However, the variability of that virus revealed that most reported strains of the virus worldwide are present in this analysis. Vitivirus incidence was low in the commercial environment (SB and DR) but common in the GC-LH; the diversity observed in this genus, spanning six species, was unexpected.

The excellent RNA virus detection obtained from the dsRNA HTS tool developed contrasts with the poor performance with DNA viruses. GGVA confirms that the dsRNA tool is not well suited to detect DNA viruses. At a time when more new-to-science DNA grapevine viruses are being discovered, including some pathogenic ones (Al Rwahnih et al., 2016a, Al Rwahnih et al., 2013, Basso et al., 2015, Maliogka et al., 2015, Perry et al., 2018, Vončina et al., 2017a) it is important to integrate a detection method specifically for this genome type, e.g. RCA. The sRNA HTS offers a very good alternative but it is more costly and limited by the detection of new-to-science viruses due to the short read length, although considering the exponential rate of new virus sequences available on GenBank, this drawback shrinks every sequencing run. The sRNA and dsRNA HTS are complementary. The scale of this study was only possible with the enrichment of dsRNA allowing to reduce the sequencing cost. An economical addition would be to do a RCA HTS alongside the dsRNA to complement the virus detection with the detection of both the DNA and the RNA viruses.

This study constitutes the largest HTS survey of grapevines to date and unveils the virome of the New Zealand vineyard. In the last review of plant virus records in New Zealand, Veerakone (2015) reported 17 viruses, 13 of which were encountered in the present study. Four viruses belonging to the genus *Nepovirus* were reported in New Zealand grapevines but were not detected in this survey thereby supporting that they are regarded as eradicated from commercial vineyard (MacDiarmid & Cohen, 2007). In addition to the Veerakone records, this study adds two recently identified vitiviruses (GVG, and GVI), the first report to New Zealand of a highly divergent isolate of GVE, and the first report in New Zealand of GGVA.

Acknowledgments

The authors would like to acknowledge The New Zealand Institute for Plant & Food Research Limited for funding, New Zealand Winegrowers for the financial support through the Rod Bonfiglioli Memorial Scholarship, and Plant Biosecurity Cooperative Research Centre for the funding of the small RNA sequencing and analysis (PBCRC2064). We would like to acknowledge the funding from the NZW NZWE1401 – Resilient & Profitable Wine Industry – towards the SB sequencing. We would like to thank our Plant & Food Research colleagues, John Fletcher for the collection of samples, Sandi Keenan for the assistance with sampling and sRNA extraction and Shamini Pushparajah for her help with the sequencing arrangement. Thank you to Nick Hoskins for the sample of Pinot gris. We are very grateful to Gardette Valmont for the sample preparation. We would like to acknowledge Jody Hobson-Peter for kindly providing the dsRNA antibodies.

5 General discussion

Status of chapter:

This chapter has not been published and is not intended for publication.

5.1 Context

The number of grapevine virus species has changed during the course of this study. In 2014, Professor Giovanni Martelli (University of Bari, Italy), considered as the leading expert in the field of grapevine virology, counted 65 viruses in grapevine (Martelli, 2014d). Thereafter, eight additional viruses were recorded, including GVG and GVI discovered as part of this study (Chapter 1, Table 1.3). The pace of virus discovery has increased with the availability to most laboratories of the HTS technologies, whereas they were inaccessible to many research groups until recently due to cost and the expertise required. This doctoral research aimed at 1) optimising the virus detection tools based using HTS for grapevine virus identification with a practical, scalable and economical tool; 2) use HTS, to identify new viruses and understand their phylogenetic relationships; and 3) run a large scale screen of grapevines viruses in New Zealand's non-commercial and commercial plantings in order to provide the industry with a better understanding of the viruses already present in New Zealand as well as offering new insights into those with the potential to cause disease in future. This research gave light to the presence of four viruses not recorded in New Zealand before (GVE-like and GGVA) including two viruses new to science (GVG and GVI).

5.2 HTS: The tool, its faults and its roles

Diagnostics of characterised viruses in grapevine is well described but is not always trivial due to the extent of the genetic variability observed in all RNA viruses. The lack of proof-reading capability of the viral replicase that therefore introduces a high rate of mutations is a partial explanation of this genetic blur, but it is combined with extended isolation of same virus species in their long lived, perennial host (Elena & Sanjuán, 2005). This phenomena is further amplified by the vegetative nature of the propagation of *Vitis* over the past millennia. This

genetic variability of viruses can impact negatively on diagnostic tools that do not detect variant strains or different serotypes (Chooi et al., 2013).

Nowadays, HTS technologies are used typically to diagnose a small set of plants that display symptoms but the cost of the technologies generally prohibit large scale surveys (Chapter 1, Table 1.3). From different viral enrichment strategies described to date, dsRNA provides the highest yield of viral nucleic acids, and the modified enrichment tool developed as part of Chapter 2 was shown to be very efficient during the survey (Chapter 4) where the ratio of viral (and viroid) reads to total reads ranged between 0.3 and 91.5% with an average of 21.4% (when excluding the samples negative for virus infection). This high enrichment allowed the confident diagnosis of plants that are positive for RNA virus infection when as few as 60,000 reads were available. Secondly, this enrichment increased the economy of virus detection such that a large scale plant survey (225 plants) could be undertaken at a cost for sequencing typically used for only a few plant samples. The speed and ease of the protocol is also beneficial as the extraction from about 2-3 g of plant tissue is completed within a day. Ultimately, the scheduled commercialisation of the dsRNA antibodies by the University of Queensland (mAb 2G4) and Merk Millipore (mAb 3G1) will support usage by other laboratories of the tool developed in this study. Yet, there are some restrictions of the new method that should not be ignored.

1) In general, the dsRNA HTS tool is specific to RNA viruses

All of the viruses encountered from grapevine were positive sense RNA, either single of double stranded, although the double-stranded RNA viruses were not included in Chapter 4 (4.4.2) as their origin could not be established and are likely to be of fungal origin similarly to the Rott et al. study (2017). Short contigs of a negative sense RNA virus were detected from a non-grape host (karaka, *Corynocarpus laevigatus*) that was used at an early stage of the development of this tool. There was very poor detection of the DNA virus GGVA (two reads in the GGVA positive sample but no reads in other samples, Chapter 4), thus highlighting the restriction of this tool for DNA virus detection without high read depth.

2) The lack of an internal control of dsRNA precludes assessment of sample quality

Without the use of an internal control the absence of virus reads from a sample can equally be interpreted as a plant that is free of any RNA viruses or a plant from which there was only poor dsRNA extraction. The addition of the rRNA 26S by RT-qPCR prior to sequencing helped to measure the success of the nucleic acid extraction based on off-target binding of this plant derived RNA by the dsRNA antibodies. However, a viral target is a better control for the sample

preparation as is routinely used for RT-PCR (Chooi et al., 2016). One alternative is to spike the sample with a small amount of a plant hosting cryptic dsRNA viruses (e.g. *Phaseolus vulgaris* (bean) expressing two endornaviruses), and then assessing by RT-qPCR the quality of the dsRNA extraction by quantifying the cryptic viruses extracted (Kesanakurti et al., 2016). However, this approach can encounter issues with a predominance of 'spiked reads' and also ignore differences in ease of sample preparation between 'spiked' bean and the sample tissue (grape in the case of this present research). Prior to the results from the survey (Chapter 4), GRSPaV was used as the internal control, but surprisingly GRSPaV was not present in the majority of the grapevines and was even absent from two entire vineyard blocks (GRSPaV was absent in blocks SB-3 and -4, Table 4.2). The high incidence of GSyV-1 (Chapter 4) makes it another candidate for an internal control and in the future GSyV-1 might be targeted in a triplex hydrolysis probe assay alongside GRSPaV and the rRNA 26S prior to sample sequencing. Of note, the majority of the samples sequenced were infected with GSyV-1 or GRSPaV (72%, Chapter 4, Appendix Table 7.1).

3) Cross-talk between samples

Another difficulty of HTS analysis, is the cross-talk observed between samples, i.e. when a sequence is incorrectly assigned; this problem is not restricted to the dsRNA tool, in fact it was one of the main talking points of the European and Mediterranean Plant Protection Organisation (EPPO) Workshop on the use of Next Generation Sequencing technologies for plant pest diagnostics held in Bari, Italy in November 2017. The cross-talk experienced with dsRNA HTS is amplified by the large number of samples included in each library. Although this cross-contamination can originate from the laboratory, it is most often attributed to the library preparation when a sample is assigned to the wrong barcode (Schnell et al., 2015). In the case of the present research, it was observed when a sample generated a very high level of similar reads (e.g. GLRaV-3) or when a very high number of similar reads was obtained in the library by a majority of the samples in relatively high concentration (e.g. GSyV-1 or GRSPaV). The best solution to minimise cross-talk is to reduce the number of samples per library (and increase the number of libraries, although cross-talk can also be observed between libraries). However, such an increase in the required number of library preparations would negatively impact the cost of the method. Also, as suggested in Bari, it is important that each library preparation includes a negative control to measure the amount of cross-talk received and possibly, an unrelated positive control to assess the level of cross-talk generated. In addition, and in all cases, each significant finding should be confirmed by another tool (e.g. RT-qPCR).

Bioinformatics is a crucial part of the virus identification process. Currently, from HTS data, only viruses that have already been reported and viruses similar to known ones can be identified, such as the new virus belonging to the genus *Macluravirus* identified in potato in Chapter 2 (2.1.4.3), the two novel viruses belonging to the genus *Vitivirus* described in Chapter 3 (3.2.4 and 3.3.4) or a novel virus in the family *Closteroviridae* identified in kiwifruit as part of an unrelated study (Blouin et al., 2018c). The BLAST tool was identified as the limiting factor but some virus bioinformatics tools include the Hidden Markov Models (HMMs) option that search for protein motifs associated with plant viruses (Uricaru et al., 2006). This addition is a clear advantage and is included in the Virtool used by Rott et al. (2017). Screening of the HTS data from this current research would benefit from future scrutiny using this new Virtool, when technically possible, and may reveal additional virus detections beyond what is reported here. In addition, data generated during this study will be accessible for further data mining in the future with advanced bioinformatics tools. In addition, if an emerging virus disease was to occur, the data will remain available to assess its presence at the time of collection (2015-2017).

5.3 Role of HTS in plant virus detection

The role of HTS plant virus detection can be divided into three categories: 1) phytosanitary; 2) disease response; and 3) surveys. The dsRNA enrichment protocol developed in this current research, linked with existing HTS and bioinformatics tools, could be used in all three situations.

5.3.1 Phytosanitary

Before movement between countries or jurisdictions it is essential to know and reduce/remove inoculum of undesired viruses; likewise, before propagation of plant material that is distributed within a region. HTS (following any of the nucleic acid preparation methods reviewed in Chapter 4.2) offers a non-targeted approach that, in theory, can detect all viruses present in a plant. Phytosanitary measures require an integrated diagnostic approach comprising multiple tools. In this manner, the current strategy includes ELISA and RT-PCR assays combined with biological indexing (Golino et al., 2017), and new testing regimes should likewise be developed with the use of multiple assays, including HTS. The dsRNA HTS tool offers a relatively economical method to first screen samples. It could be followed or concurrently

complemented with a second HTS method (sRNA or total RNA) or/and more conventional approaches (RT-PCR, ELISA, biological indexing). If the budget was not restricted, HTS on ribosome-depleted total RNA to screen for all microbes present in the plan, is a comprehensive, albeit costly, approach. The cheaper sRNA HTS is an alternative approach that allows the detection of most viruses present, but it is not yet suited to detect highly divergent viruses.

With the improvement of the detection technologies over time, the priority of virus biology has decreased. Previously, the sampling time and tissue was very well defined in order to target the optimal tissue/time to detect one virus. With its improved sensitivity, and its broad spectrum of sequence targets, HTS can, in theory, detect all the viruses in one plant, at one time-point. The contrasting results obtained for the fleck-like viruses between two HTS methods highlights that the biology of the virus is important and multiple sampling points in a plant improves the detection of non-systemic or unevenly distributed viruses; sampling may occur during the early stage of infection or of the foliar season, or from an uninfected part of the plant. Multiple sampling times will further improve the chance of detecting all the viruses as we known that the ideal sampling tissue and time is different for different viruses targeted (Constable et al. 2012, Fiore et al., 2009).

5.3.2 Disease response

In the case of disease responses, especially of unknown aetiology, the dsRNA HTS tool is very economical to screen multiple plants regardless of whether they show the diseased phenotype or not (Table 1.3). Once again, any findings should be followed with a different and confirmatory test, and the user should be aware that DNA viruses may not be detected. Ideally, a run of sRNA HTS would also be performed simultaneously. If the number of diseased plants is limited, and the budget not a constraint, a total RNA HTS is preferred for its ability to detect pathogens beyond viruses and viroids. However, the improvement in the third generation sequencer MinION (Oxford Nanopore Technologies) has the potential to transform the way we respond to a disease of unknown aetiology. The MinION is a very small sequencer device that can be plugged into a computer and generate 10 to 20 Gb of DNA sequence data in a very short time. The US$1000 device is single use (the flow-cell is replaced every run) but with this amount of data generated, multiplex can be achieved. The input is DNA and the data is generated instantly. In theory, the dsRNA preparation (RT-PCRed) could be used in this platform. The MinION error rate is still high but in the case of a disease outbreak, the priority is to identify the causal agent, not its exact sequence. The simplified bioinformatics is

fast because of the long read length (up to several kb). The MinION is one tool that can be stored in the laboratory in preparedness for use in a disease outbreak, although the manufacturer recommends using the flow-cell within 8 weeks. The long lag between sample preparation and downloading of the sequence data typically experienced with current sequencing platforms (during this current study it was between 3 and 12 weeks) is the obstacle that is overcome by the MinION which can provide data the day of collection and therefore has tremendous potential for at, or near, point of care diagnosis (Adams et al., 2017, Thomas et al., 2017).

5.3.3 General surveillance

The potential of the advance in technology for real-time HTS will very soon change the disease diagnostic practises. However, in the case of surveys, we have shown that dsRNA HTS is a very powerful tool to detect RNA viruses including those not previously identified. The low cost of the dsRNA HTS makes it the first option but a DNA virus detection element should be added when DNA viruses are also of interest. Since all known plant DNA viruses are circular they can be detected by RCA (Shepherd et al., 2008). It would be opportune to include a RCA step using the DNA remaining from the total RNA extraction that is the input for dsRNA enrichment step. Together, these two methods should give a complete and economical snapshot of the virome of surveyed plants. Although the dsRNA and sRNA HTS are complementary for the identification of all viruses, the price of the sRNA process is still too high to be used in a large survey.

In addition to use in phytosanitary, disease response and survey purposes, the dsRNA HTS method could be used to localise and/or quantify viral replication within a tissue or a cell, and could be compared with that of sRNAs, total RNA, and/or capsid protein (or other virus-encoded proteins if antibodies or labelled proteins are available). Localisation of replicating viral RNA (or dsRNA formed during any other stage) might be detected using labelled dsRNA antibodies and compared with the pattern from uninfected tissue or cells (Lee et al., 2017, Triantafilou et al., 2012, Weber et al., 2006). The uninfected control is important as the antibodies can bind to rRNA, as demonstrated in this current research. However, translation processes, and therefore the location and amount of rRNA, may still be unique in infected cells as virus infection can modify translation (Hafrén et al., 2015). Quantification of the immunocaptured dsRNA could be undertaken using RT-qPCR (viral or plant) and compared with RT-qPCR of targeted sRNA, RT-qPCR of specific transcripts within total RNA or the capsid

as measured also by immunocapture RT-PCR. Such research might provide insights into the cell biology of virus replication and virus-host interactions.

5.4 New viruses detected

During this current study, 17 viruses were identified in *Vitis* including four new records in New Zealand of which two are new-to-science viruses. Three of these new records comprised detection of RNA viruses and one a DNA virus.

The vitivirus GVG was discovered first in the NZW germplasm, where it appears to be common (68% of the plants tested in the GC-LH). The virus was then detected in one plant within the Sauvignon blanc survey in Marlborough and another plant in a commercial vineyard in the Hawke's Bay region. These single plant findings suggest that the virus was propagated outside the collection. A recent report including sequence data showed the same virus was detected in Croatia (Vončina & Almeida, 2017a) although there are substantial genetic differences to the one detected in New Zealand (Chapter 4, Figure 4.5). The second new-to-science virus described is GVI, which is related to GVG within the vitivirus genus. GVI was detected in nine plants and these were all co-infected with GVG. This second and most recently reported vitivirus GVI has not yet been detected outside the grapevine germplasm collection, and has not yet been reported outside New Zealand. A virus related to GVE was also detected in the survey of Sauvignon blanc from Marlborough (Chapter 4.4.7.1). This was the first report of GVE (or a GVE-like virus) in the country. The only known infected plant was host to GVG. All these viruses belong to the genus *Vitivirus* with GVA, GVB and GVD that were also detected during this survey and GVF, GVH and GVJ reported in the same host but not yet identified in New Zealand.

Amongst grapevine vitiviruses, only GVA, GVB, GVE are known to be transmitted by mealybugs (Minafra et al., 2017). Vitiviruses are often associated with a member of the family *Closteroviridae*, in grapevine and other hosts such as mint and kiwifruit (Blouin et al., 2013, Blouin et al., 2018c, Tzanetakis et al., 2010). In grapevine it is believed that the vitiviruses require the presence of a leafroll virus (GLRaV-1, -2, or -3) for its acquisition (Blaisdell et al., 2012). However, a recent publication described that the vitivirus benefited from the presence of a leafroll virus to increase its replication in the *Vitis* host and therefore its chance of transmission rather than being dependent on transmission co-factors provided by the leafroll virus (Rowhani et al., 2018). This new observation may explain the association between the

two viruses and the increased transmission observed in the presence of the leafroll virus. In this current survey, all the plants infected by one of the six detected vitiviruses were also infected with GLRaV-3. The close interaction between these two viruses groups is not yet fully understood and should be investigated further.

The detection of GGVA constitutes the only finding of a DNA virus in the current study. This is the first report of the virus in New Zealand. The virus was originally described from the USA on imported vines from Korea. The imported plants displayed virus symptoms and were infected by multiple viruses. A subsequent survey of the USDA-ARD Clonal Germplasm Repository found 15 additional GGVA infected plants with no correlation with symptoms. The authors also reported graft transmissibility of the virus (Al Rwahnih et al., 2016a). The virus was subsequently reported in Korea and China where it may be widespread, but no association with symptoms was established (Fan et al., 2017, Jo et al., 2018). All the GenBank sequences available to date are very similar to each other (less than 3% difference at the nucleotide level). The single known plant positive for GGVA in New Zealand (sample # AB537; collection # VID280 TK00184) is an interspecific crossing, Seibel 7052. The plant was also found to be infected with GLRaV-2, GRSPaV, and HVd and no symptoms were reported at the time of collection (summer 2016) or re-collection (summer 2018). In the New Zealand grapevine variety register, the source and year of importation of this plant is absent but it was logged between TK00183 (Siebel 6339) and TK00188 (Siebel 10096) that were both imported in 1957 from the USA Department of Agriculture. From this information available to date, we can conclude that GGVA in New Zealand is not causing severe symptoms and its spread is very limited; after possibly 60 years it has not spread to any of the other 15 plants sampled from the same germplasm collection. The impact and the spread of the virus therefore appears to be negligible. However, this finding should generate more interest in DNA viruses and their place in the New Zealand virome.

5.5 Survey

The grapevine survey in this current study comprised 225 plants and detected a total of 17 viruses. From the Table 1.3 within Chapter 1 that summarises all known HTS studies of plant viruses, no other HTS study has previously detected as many viruses in grapevines. Such a detection rate could insinuate that New Zealand vineyards are highly infected. However, closer examination demonstrates the exact opposite. The New Zealand commercial vineyards have a low virus incidence, with less than one virus detected per vine between the four Sauvignon

blanc vineyards (average detection of 0.96 viruses per vine, from Table 7.1). In most cases, the viruses detected were GRSPaV or GSyV-1 and without these two prevalent viruses, the average number of viruses detected per plant drops to 0.2. The assortment of viruses that were detected at low incidence was surprising as seven additional viruses were detected from the Sauvignon blanc vineyards. The relatively high health of the vineyards can be credited to the studious work of the nurseries to propagate clean material and the GGS established and managed by the NZW that determines grapevine quality of plants sold by the nurseries, including the absence of GLRaV-3 (New Zealand Winegrowers, 2017b). This low virus incidence is also the result of strict import health regulations for *Vitis* since the Biosecurity Act in 1993 (New Zealand Parliamentary, 1993). This legal document reformed the laws related to pest and unwanted organisms to New Zealand. The viruses reported in this thesis were not only reported to the Ministry for Primary Industries (MPI) as new to New Zealand reports of an infectious agent, but also prompted discussions with members of the NZW about the consequences of the high incidence of GSyV-1 and whether more viruses should be tested in addition to GLRaV-3 as part of the GGS evolution.

To date, no adverse biological impacts of the fleck-like viruses (members of the family *Tymoviridae* that include GFkV, GRVFV, GAMaV, GSyV-1 and GRGV) have been reported, and it was even suggested that these viruses could be added to GRSPaV, HSVd and GYSVd as the "virome background" of a "healthy-looking" grapevine and excluded from sanitary measures (Saldarelli et al., 2017). This statement clearly resonates with the results of this current survey. The ecological impact of these viruses is still unclear but according to their incidence, GRSPaV and GSyV-1 could inhabit the same ecological niche. In fact, Spearman's rank-order correlation provided evidence of a negative correlation between GSyV-1 and GRSPaV ($r_s(175) = -0.602$, $p = <0.001$). The detection of multiple species of vitivirus is more problematic than the fleck-like viruses as vitiviruses can be associated with diseases (reviewed in Minafra et al., 2017 and discussed in Chapter 1.3.6.2 and 1.3.6.7). However, the management of GLRaV-3 is likely to also remove co-infecting vitiviruses from nurseries and vineyards resulting in only low vitivirus incidence; the results of the survey confirm the strong association between GLRaV-3 and the presence of a vitivirus under New Zealand conditions.

This survey confirms that besides GLRaV-3, there is a lack of, or at most very low, movement of the other ampeloviruses (GLRaV-1 and GLRaV-4) as they were only detected in very few plants and were not found outside the germplasm collection. In the case of GLRaV-1, the two positive plants were likely to have been propagated from the same Chardonnay Mendoza imported in 1971. In contrast, the vitiviruses were more common, especially GVA and GVG.

These two viruses were detected outside the germplasm. These findings would suggest that under New Zealand environments the vitiviruses GVA and GVG are vectored.

To date, this is the largest survey of commercial grapevines using HTS. This study sheds light on the virus diversity within New Zealand vineyards, the potential route of various viruses into New Zealand, and the projected impact of the viruses based on the current knowledge of their biology.

5.6 Future directions

The revolution that began with the detection of plant viruses by HTS about ten years ago is now accessible to most laboratories around the world, contributing to the exponential rise of new virus reports (Table 1.3). The increased number of new sequence genomes available helps the detection of related viruses around the world and further increases the speed of discovery. Recent reports of new viruses include many DNA viruses and suggest previous underestimation of their incidence and their contribution in virus ecology (Cieniewicz et al., 2017). Management strategies for DNA viruses can vary from the ones described for the RNA viruses due to, sometime, different insect vectors whose behaviour in vineyards is not fully understood (Basso et al., 2015, Cieniewicz et al., 2017, Qiu & Schoelz, 2017). The detection of GGVA by HTS of sRNA has shown that the dsRNA HTS tool would have missed the presence of DNA viruses in the commercial vineyards. Further development of DNA virus detection either alone or to complement RNA virus detection in surveys is a future direction for investigation.

The grapevine germplasm block is no longer used to propagate *Vitis* for use by grapegrowers and its high level of virus infection is notorious in New Zealand. The collection represents a snapshot of the past both in terms of imported and propagated *Vitis* cultivars, and grapevine viruses. The virus diversity in the grapevine germplasm was much higher than the commercial vineyards, as expected for old germplasm collections and as observed in reports of other examples (Al Rwahnih et al., 2015b, Vončina et al., 2017b). With 14 viruses (15 when including GFkV that was only detected by sRNA) this limited study of 39 plants sequenced (GC-LH plus GC-HH) was able to detected most of the grapevine viruses known in the country including new-to-New Zealand and new-to-science viruses. It is important to note, only 4% of the germplasm blocks was assessed. The remaining 96% of the germplasm is likely to host additional viruses and likely represents most of the viruses present in the New Zealand

commercial vineyards. It is essential to fully assess this collection in order to increase the knowledge of the virus diversity in the country. In addition to the historical virome of the germplasm, the present commercial virome should be further understood through a more comprehensive survey across regions, cultivars, scion combinations, vineyard management regimes, and planting dates.

The new genetic data generated in this current study along with other grapevine research should be used to understand the biology of grapevine infecting viruses and their place in the vineyard environment. For example, the ubiquity of GRSPaV and GSyV-1 is likely to be the result of human propagation of healthy appearing grapevines thus further supporting the notion that these viruses are not pathogenic under 'typical' conditions. GRSPaV infected grapevines have even been shown to infer some benefits such as a possible improved drought tolerance (Gambino et al., 2012). However, the biology resulting in an impact on the plant can change when the plant environment changes, for instance by the presence of an additional virus, a mutation in the virus' or plant's genome, different rootstock / scion combinations, or changes in the physical environment. This is what happened in Australia when Chardonny vines were replaced by Shiraz on the same rootstock; the 'latent' GVA virus became virulent on the new scion (Habili et al., 2016). Therefore, there may be some circumstances under which a mutualistic relationship (whereby both partners benefit) changes to a commensal (whereby one partner benefits while the other is not affected), or an antagonistic one (whereby one partner benefits but at the expense of the other and thus a diseased plant may result), or vice versa (Roossinck, 2011).

Generally, the impact of a virus on an industry is correlated to the amount of research on the virus, hence there is little knowledge available for those putative "latent" viruses. Furthermore, "latent" viruses are generally more recently discovered as a consequence of HTS and have had less time for research to be undertaken on them. It is important to study the biology of plant viruses, in particular newly described viruses and "latent" viruses, to understand the potential threat, or benefit, to the industry. For instance, there is a lack of understanding of the relationship between the vitivirus and GLRaV-3. Such a knowledge gap needs to be filled to better manage grapevine production. From the latest research, the vitivirus benefits from the presence of the leafroll viruses through increased replication and a greater transmission rate (Rowhani, 2017). Is it the presence of a viral suppressor of RNA silencing (VSR) in the leafroll virus enabling the vitivirus to replicate more and therefore help its acquisition by the vector? What is the impact of the relationship for the leafroll virus?

Since the start of this project, several publications have shaken the dogma suggesting that most grapevine viruses are specialists i.e. have only a single host. In the past two years, the detection of non-vitis alternative hosts have been reported for GLRaV-1, GLRaV-3 and GPGV (Çağlayan et al., 2016, Prator et al., 2017, Gualandri et al., 2017). These findings are important to understand in the ecological equilibrium of the virus in the environment and should be taken into account when determining the best disease management strategy. We believe GLRaV-1 is not vectored well in New Zealand and GPGV is not known to be present in New Zealand (Veerakone et al., 2015), but GLRaV-3 is a significant disease that is vectored by extant New Zealand mealybugs (Charles et al., 2009). It would be interesting to extend the next survey to include known or potential alternative hosts in search of leafroll viruses, in addition to other grapevine viruses. HTS is now sufficiently economical to undertake a survey of the alternative hosts of viruses growing within vineyards. For instance, a "lawn-mower" approach (that uses no-barcoding for a large number of samples) could be used to assess whether some grapevine viruses can be detected outside what is/was considered as their primary host. A comparison between vineyards with known varying viromes (e.g. SB-2 and SB-4) as well as vineyards with high GLRaV-3 pressure would help identify alternative virus host plants, if any and to compare groundcover plant and the grapevine viromes under different conditions.

Thanks to HTS the increased speed of discovery of new microorganisms has the potential for a massive impact on biosecurity and its regulations (MacDiarmid et al., 2013, Massart et al., 2017). Regulators are aiming at a moving target of biosecurity alerts often combined with a dearth of biology associated with new-to-science viruses. Due to the lag between the generation of the molecular and biological data, the phytosanitary decisions on plant movement or propagation have to be made on assumptions that a new virus is behaving the same as the most closely described virus and therefore poses the same risks, or to stop any movement and propagation as a preventative measure. New virus detections within a country or other jurisdiction could potentially be used as a trade barrier. The impact on trade is of significant importance to both importing and exporting industries, and the best approach is to persuade and finance research groups to do more biological studies, to understand the real risks (for each individual virus on its plant host(s)), and to convince governments to update their list of viruses present in their country by promoting more untargeted surveys such as the one presented in this thesis (Massart et al., 2017).

Lastly, scientists and regulators should not underestimate the role of the growers and individuals working closely in the vineyards. They are the first to witness unusual phenotypes

that may be caused by a pathogen. The vigilance of the growers and the close communication with them will help avoid epidemics such as GRBV detected in the USA or GPGV in Europe.

5.7 Conclusion

This research enabled the development of a dsRNA enrichment tool that could result in more than 70% HTS reads being of viral origin (Chapter 2.1). With further development of this tool, especially for use on recalcitrant tissues and reduction of rRNA contamination, dsRNA enrichment from grapevines resulted as much as 90% HTS reads being of viral origin (Chapter 2.2). Using this improved method, the largest survey of grapevines reported to date using HTS was undertaken. The New Zealand grapevine virome comprises 21 viruses of which 17 were detected in this project; four viruses belonging to the genus *Nepovirus* have not been reported recently, nationwide. Five fleck-like viruses, four leafroll viruses, GRSPaV and six vitiviruses including two new-to-science viruses and one new-to-New Zealand virus were detected in addition to GGVA, the first grapevine DNA virus identified in New Zealand.

One of the main outcomes from this research is a better understanding of the high incidence, low impact or "background" virome in New Zealand grapevines which is predominantly either of two latent viruses, GRSPaV or GSyV-1. Most other viruses detected had a low incidence in the commercial vineyards, with GLRaV-3 remaining as the major viral disease agent. The detected vitiviruses varied by species but were only detected as co-infections with GLRaV-3 thus providing a simple 'remove leafroll virus' management tool for viticulturists (Bell, 2015). Knowledge of the status of the grapevine germplasm and commercial vineyard virome provides baseline data for regulators to set phytosanitary measures and to be more prepared in a disease response. This is also valuable information for nurseries, and it highlights the effective work of sanitation and risk awareness historically conveyed by scientists such as Dr Rod Bonfiglioli and Dr Richard Smart. It also re-enforces the importance of using clean plant stocks in the nurseries to avoid spreading new pathogens.

This thesis presents an up to date assessment of the viruses present in New Zealand grapevines. The information from this survey is reassuring to the New Zealand grape growing industry and its regulators as no viral threat was detected and its supports the management of GLRaV-3 remaining the highest priority of the industry. However, the detection, for the first time in New Zealand, of a DNA virus in grapevine suggests that there will be more DNA viruses

present than anticipated at the start of this project. The ecological niche of DNA viruses in viticulture and the threat they could present should be investigated.

6 References

Ackerer L., Monsion B., Hemmer C., Komar V., Vigne E., Gertz C., & Demangeat G. (2015). *Nanobody-based products as diagnostic tool for Grapevine fanleaf degeneration viruses.* Paper presented at the Proceedings of the 18th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Ankara, Turkey.

Adams I., Braidwood L., Stomeo F., Phiri N., Uwumukiza B., Feyissa B., Mahuku G., Wangai A., Smith J. & Mumford R. (2017). Characterising maize viruses associated with maize lethal necrosis symptoms in sub Saharan Africa. bioRxiv:161489.

Adams M., Lefkowitz E., King A., & Carstens E. (2014). Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2014). *Archives of Virology, 159*(10), 2831-2841.

Adams M. J., Antoniw J. F., Bar-Joseph M., Brun A. A., Candresse T., Foster G. D., & Fauquet C. M. (2004). The new plant virus family Flexiviridae and assessment of molecular criteria for species demarcation. *Archives of Virology, 149*(5), 1045-1060.

Agranovsky A. A., Lesemann D. E., Maiss E., Hull R., & Atabekov J. G. (1995). " Rattlesnake" structure of a filamentous plant RNA virus built of two capsid proteins. *Proceedings of the National Academy of Sciences, 92*(7), 2470-2473.

Al Rwahnih M., Daubert S., Golino D., & Rowhani A. (2009). Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. *Virology, 387*(2), 395-401.

Al Rwahnih M., Daubert S., Urbez-Torres J. R., Cordero F., & Rowhani A. (2011). Deep sequencing evidence from single grapevine plants reveals a virome dominated by mycoviruses. *Archives of Virology, 156*(3), 397-403.

Al Rwahnih M., Dolja V. V., Daubert S., Koonin E. V., & Rowhani A. (2012a). Genomic and biological analysis of Grapevine leafroll-associated virus 7 reveals a possible new genus within the family Closteroviridae. *Virus Research, 163*(1), 302-309.

Al Rwahnih M., Sudarshana M. R., Uyemoto J. K., & Rowhani A. (2012b). Complete Genome Sequence of a Novel Vitivirus Isolated from Grapevine. *Journal of Virology, 86*(17)

Al Rwahnih M., Dave A., Anderson M. M., Rowhani A., Uyemoto J. K., & Sudarshana M. R. (2013). Association of a DNA virus with grapevines affected by red blotch disease in California. *Phytopathology, 103*(10), 1069-1076.

Al Rwahnih M. A., Rowhani A., Golino D. A., Islas C. M., Preece J. E., & Sudarshana M. R. (2014). Detection and genetic diversity of Grapevine red blotch-associated virus isolates in table grape accessions in the National Clonal Germplasm Repository in California. *Canadian Journal of Plant Pathology*, 37(1), 130-135.

Al Rwahnih M., Rowhani A., & Golino D. (2015a). First Report of Grapevine red blotch-associated virus in Archival Grapevine Material From Sonoma County, California. *Plant Disease, 99*(6), 895-895.

Al Rwahnih M., Daubert S., Golino D., Islas C., & Rowhani A. (2015b). Comparison of Next-Generation Sequencing Versus Biological Indexing for the Optimal Detection of Viral Pathogens in Grapevine. *Phytopathology, 105*(6), 758-763.

Al Rwahnih M., Alabi O. J., Westrick N. M., Golino D., & Rowhani A. (2016a). Description of a Novel Monopartite Geminivirus and Its Defective Subviral Genome in Grapevine. *Phytopathology, 107*(2), 240-251.

Al Rwahnih M., Alabi O. J., Westrick N. M., Golino D., & Rowhani A. (2016b). Near-complete genome sequence of grapevine fabavirus, a novel putative member of the genus Fabavirus. *Genome announcements, 4*(4), pp.e00703-16.

Al Rwahnih M., Saldarelli P., & Rowhani A. (2017). Grapevine leafroll-associated virus 7. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 221-228). Cham: Springer International Publishing.

Alabi O. J., Martin R. R., & 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*, 91(1), 265-276.

Alabi O. J., Al Rwahnih M., Karthikeyan G., Poojari S., Fuchs M., Rowhani A., & Naidu R. A. (2011). Grapevine leafroll-associated virus 1 Occurs as Genetically Diverse Populations. *Phytopathology, 101*(12), 1446-1456.

Alabi O. J., Rwahnih M. A., Mekuria T. A., & Naidu R. A. (2014). Genetic Diversity of Grapevine virus A in Washington and California Vineyards. *Phytopathology, 104*(5), 548-560.

Alkowni R., Zhang Y. P., Rowhani A., Uyemoto J. K., & Minafra A. (2011). Biological, molecular, and serological studies of a novel strain of grapevine leafroll-associated virus 2. *Virus Genes, 43*(1), 102-110.

Almeida R. P., Daane K. M., Bell V. A., Blaisdell G. K., Cooper M. L., Herrbach E., & Pietersen G. (2013). Ecology and management of grapevine leafroll disease. *Frontiere of Microbiology, 4*, 94.

Alzhanova D. V., Prokhnevsky A. I., Peremyslov V. V., & Dolja V. V. (2007). Virion tails of Beet yellows virus: coordinated assembly by three structural proteins. *Virology, 359*(1), 220-226.

Andret-Link P., Laporte C., Valat L., Ritzenthaler C., Demangeat G., Vigne E., Fuchs M. (2004). Grapevine fanleaf virus: Still a major threat to the grapevine industry. *Journal of Plant Pathology, 86*(3), 183-195.

Andrew R., Bell V., Hoskins N., Pietersen G., & Thompson C. (2015). *Leafroll 3 virus and how to manage it.* Auckland: New Zealand Winegrowers. 108p.

Angelini E., Bertazzon N., & Borgo M. (2003). Occurrence of the Redglobe strain of Grapevine leafroll-associated virus 2 in table and wine grape varieties in Italy. *Vitis-Geilweilerhof, 42*(4), 203-204.

Angelini E., Aboughanem-Sabanadzovic N., Dolja V., & Meng B. (2017). Grapevine leafroll-associated virus 2. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 141-165). Cham: Springer International Publishing.

Arabidopsis-Genome-Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature, 408*(6814), 796-815.

Aramburu J., & Moreno P. (1994). Detection of Double-Stranded-Rna (Dsrna) in Crude Extracts of Virus-Infected Plants by Indirect Elisa. *Journal of Phytopathology-Phytopathologische Zeitschrift, 141*(4), 375-385.

Aramburu J., Navascastillo J., Moreno P., & Cambra M. (1991). Detection of double-stranded-RNA by ELISA and dot immunobinding assay using an antiserum to synthetic Polynucleotides. *Journal of Virological Methods, 33*(1-2), 1-11.

Atallah S. S., Gomez M. I., Fuchs M. F., & Martinson T. E. (2012). Economic Impact of Grapevine Leafroll Disease on Vitis vinifera cv. Cabernet franc in Finger Lakes Vineyards of New York. *American Journal of Enology and Viticulture, 63*(1), 73-79.

Bahder B. W., Zalom F. G., Jayanth M., & Sudarshana M. R. (2016). Phylogeny of Geminivirus Coat Protein Sequences and Digital PCR Aid in Identifying Spissistilus festinus as a Vector of Grapevine red blotch-associated virus. *Phytopathology, 106*(10), 1223-1230.

Barrero R. A., Napier K. R., Cunnington J., Liefting L., Keenan S., Frampton R. A., & Bellgard M. I. (2017). An internet-based bioinformatics toolkit for plant biosecurity diagnosis and surveillance of viruses and viroids. *BMC Bioinformatics, 18*(1), 26.

Basso M. F., da Silva J. C. F., Fajardo T. V. M., Fontes E. P. B., & Zerbini F. M. (2015). A novel, highly divergent ssDNA virus identified in Brazil infecting apple, pear and grapevine. *Virus Research, 210*, 27-33.

Bell V. A., Bonfiglioli R. G. E., Walker J. T. S., Lo P. L., Mackay J. F., & McGregor S. E. (2009). Grapevine leafroll-associated virus 3 persistence in Vitis vinifera remnant roots. *Journal of Plant Pathology, 91*(3), 527-533.

Bell V. A. (2015). *An integrated strategy for managing Grapevine leafroll-associated virus 3 in red berry cultivars in New Zealand vineyards.* (PhD), Victoria University of Wellington.

Bell V. A., Blouin A. G., Cohen D., Hedderley D. I., Oosthuizen T., Spreeth N., & Pietersen G. (2017). Visual symptom identification of Grapevine leafroll-associated virus 3 in red berry cultivars supports virus management by roguing. *Journal of Plant Pathology, 99*(2), 477-482.

Bem F., & Murant A. (1979). Host range, purification and serological properties of heracleum latent virus. *Annals of Applied Biology, 92*(2), 243-256.

Bergervoet J. H. W., Peters J., van Beckhoven J. R. C. M., van den Bovenkamp G. W., Jacobson J. W., & van der Wolf J. M. (2008). Multiplex microsphere immuno-detection of potato virus Y, X and PLRV. *Journal of Virological Methods, 149*(1), 63-68.

Bertazzon N., Borgo M., Vanin S., & Angelini E. (2010). Genetic variability and pathological properties of Grapevine leafroll-associated virus 2 isolates. *European Journal of Plant Pathology, 127*(2), 185-197.

Bertazzon N., Filippin L., Forte V., & Angelini E. (2016). Grapevine Pinot gris virus seems to have recently been introduced to vineyards in Veneto, Italy. *Archives of Virology, 161*(3), 711-714.

Bertazzon N., Forte V., Filippin L., Causin R., Maixner M., & Angelini E. (2017). Association between genetic variability and titre of Grapevine Pinot gris virus with disease symptoms. *Plant Pathology*, 66(6), 949-959.

Bertin S., Cavalieri V., Gribaudo I., Sacco D., Marzachì C., & Bosco D. (2016). Transmission of Grapevine virus A and Grapevine leafroll-associated virus 1 and 3 by Heliococcus bohemicus (*Hemiptera: Pseudococcidae*) Nymphs From Plants With Mixed Infections. *Journal of Economic Entomology*.

Bester R., Maree H. J., & Burger J. T. (2012). Complete nucleotide sequence of a new strain of grapevine leafroll-associated virus 3 in South Africa. *Archives of Virology*, 157(9), 1815-1819.

Beuve M., Moury B., Spilmont A. S., Sempé-Ignatovic L., Hemmer C., & Lemaire O. (2013). Viral sanitary status of declining grapevine Syrah clones and genetic diversity of Grapevine Rupestris stem pitting-associated virus. *European Journal of Plant Pathology, 135*(2), 439-452.

Blaisdell G. K., Zhang S., Daane K. M., & Almeida R. P. P. (2012). Patterns of Virus Transmission from Hosts with Mixed Infections. *Proceedings of the 17th Congress of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG) Davis, California, USA, October 7-14, 2012*(2012), 178-179.

Blattner F. R., Plunkett G., Bloch C. A., Perna N. T., Burland V., Riley M., Collado-Vides J. Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y. (1997) "The complete genome sequence of Escherichia coli K-12." science 277, no. 5331 1453-1462.

Blouin A. G., Chavan R. R., Pearson M. N., Macdiarmid R. M., & Cohen D. (2012). Detection and characterisation of two novel vitiviruses infecting Actinidia. *Archives of Virology, 157*(4), 713-722.

Blouin A.G. , Pearson M.N. , Chavan R. R., Woo E. N. Y., Lebas B. S. M., Veerakone S., & Cohen D. (2013). Viruses of Kiwifruit (*Actinidia* species). *Journal of Plant Pathology, 95*(2), 221-235.

Blouin A. G., Ross H. A., Hobson-Peters J., O'Brien C. A., Warren B., & MacDiarmid R. (2016). A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies. *Molecular Ecology Resources, 16*(5), 1255-1263.

Blouin A. G., Chooi K. M., Cohen D., & MacDiarmid R. M. (2017a). Serological Methods for the Detection of Major Grapevine Viruses. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 409-429). Cham: Springer International Publishing.

Blouin A. G., & MacDiarmid R. M. (2017b). Distinct Isolates of Grapevine rupestris vein feathering virus Detected in Vitis vinifera in New Zealand*. Plant Disease*, 101(12), 2156-2156.

Blouin A. G., Keenan S., Napier K. R., Barrero R. A., & MacDiarmid R. M. (2018a). Identification of a novel vitivirus from grapevines in New Zealand. *Archives of virology, 163*(1), 281-284.

Blouin A. G., Chooi K. M., Warren B., Napier K. R., Barrero R. A., & MacDiarmid R. M. (2018b). Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand. *Archives of virology*, *163*(5), 1371-1374.

Blouin A. G., Biccheri R., Khalifa M. E., Pearson M. N., Poggi Pollini C., Hamiaux C., Cohen D. & Ratti C. (2018c). Characterization of Actinidia virus 1, a new member of the family Closteroviridae encoding a thaumatin-like protein. *Archives of Virology* 163(1):229-234.

Bonfiglioli R. G. E., Hoskins N., & Edwards F.(2001).Grapevine leafroll viruses in New Zealand viticulture. *Riversun Technical Bulletin.* pp. 1-8.

Bonfiglioli R., Hoskins N., & Edwards F. (2002). Grapevine leafroll virus type 3 spreading in New Zealand. *The Australian and New Zealand Grapegrower and Winemaker, 457,* 58-61.

Bonfiglioli R., Edwards F., Hoskins N., Pantaleo A. (2003). Graft incompatibility syndrome in New Zealand Merlot vines involves another possible variant of GLRaV-2. *Australian & New Zealand Grapegrower & Winemaker, 476,* 50-54.

Bonfiglioli R., & Hoskins N. (2006). Managing virus in New Zealand vineyards. *The Australian and New Zealand Grapegrower and Winemaker, 515,* 43-46.

Bouyahia H., Boscia D., Savino V., La Notte P., Pirolo C., Castellano M., & Martelli GP. (2005). Grapevine rupestris stem pitting-associated virus is linked with grapevine vein necrosis. *Vitis-Geilweilerhof, 44*(3), 133.

Bragato R. (1895). *Report on the prospects of viticulture in New Zealand: Together with instructions for planting and pruning*: S. Costall.

Bragato R. (1902). Viticulture *The tenth report of the Department of Agriculture* (pp. 450-455).

Bragato R. (1903). Viticultural branch *The eleventh report of the Department of Agriculture*.

Burger J., Maree H.J., Gouvea P., & Naidu R. A. (2017). Grapevine leafroll-associated virus 3. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 167-195). Cham: Springer International Publishing.

Caenorhabditis elegans Consortium, (1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science, 282*(5396), 2012-2018.

Çağlayan K., Elçi E., & Gazel M. (2016). Detection and partial characterization of grapevine leafroll-associated virus 1 in pomegranate trees in Turkey. *European Journal of Plant Pathology, 145*(1), 199-202.

Candresse T., Theil S., Faure C., & Marais A. (2018). Determination of the complete genomic sequence of grapevine virus H, a novel vitivirus infecting grapevine. *Archives of Virology, 163*(1), 277-280.

Chamberlain E. E., Over De Linden A. J., & Berrysmith F. (1970). Virus diseases of grapevines in New Zealand. *New Zealand Journal of Agricultural Research, 13*(2), 338-358.

Charlermroj R., Himananto O., Seepiban C., Kumpoosiri M., Warin N., Oplatowska M., Elliott C. T. (2013). Multiplex detection of plant pathogens using a microsphere immunoassay technology. *PloS ONE, 8*(4), e62344.

Charles J. G., Cohen D., Walker J. T. S., Forgie S. A., Bell V. A., & Breen K. C. (2006). A review of the ecology of grapevine leafroll associated virus type 3 (GLRaV-3). *New Zealand Plant Protection, 59*, 330-337.

Charles J. G., Froud K. J., van den Brink R. & Allan D. J. (2009). Mealybugs and the spread of grapevine leafroll-associated virus 3 (GLRaV-3) in a New Zealand vineyard. *Australasian Plant Pathology* 38(6):576-583.

Charles J. G., Bell V. A., Lo P. L., Cole L. M., & Chhagan A. (2010). Mealybugs (Hemiptera: Pseudococcidae) and their natural enemies in New Zealand vineyards from 1993-2009. *New Zealand Entomologist, 33*, 84-91.

Chiumenti M., Giampetruzzi A., Morelli M., Savino V. N., Martelli G. P., La Notte P., Saldarelli P. (2016). Detection and molecular characterization of a Grapevine Roditis leaf discoloration-associated virus (GRLDaV) variant in an autochthonous grape from Apulia (Italy). *Virus Genes, 52*(3), 428-431.

Chooi K. M., Cohen D., & Pearson M. N. (2013). Molecular characterisation of two divergent variants of grapevine leafroll-associated virus 3 in New Zealand. *Archives of Virology, 158*(7), 1597-1602.

Chooi K. M., Cohen D. & Pearson M. N. (2016). Differential distribution and titre of selected grapevine leafroll-associated virus 3 genetic variants within grapevine rootstocks. *Archives of Virology* 161(5):1371-1375.

Cieniewicz E. J., Pethybridge S. J., Loeb G., Perry K., & Fuchs M. (2017). Insights Into the Ecology of Grapevine red blotch virus in a Diseased Vineyard. *Phytopathology, 108*(1), 94-102.

Clark M. F., & Adams A. N. (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Journal of General Virology, 34*(3), 475-483. doi: 10.1099/0022-1317-34-3-475

Coetzee B., Maree H. J., Stephan D., Freeborough M.-J., & Burger J. T. (2010a). The first complete nucleotide sequence of a grapevine virus E variant. *Archives of Virology, 155*(8), 1357-1360.

Coetzee B., Freeborough M. J., Maree H. J., Celton J. M., Rees D. J. G., & Burger J. T. (2010b). Deep sequencing analysis of viruses infecting grapevines: Virome of a vineyard. *Virology, 400*(2), 157-163.

Cogotzi L., Giampetruzzi A., Nölke G., Orecchia M., Elicio V., Castellano M. A., Saldarelli P. (2009). An assay for the detection of grapevine leafroll-associated virus 3 using a single-chain fragment variable antibody. *Archives of Virology, 154*(1), 19-26.

Constable F. E., Connellan J., Nicholas P., & Rodoni B. C. (2012). Comparison of enzyme-linked immunosorbent assays and reverse transcription-polymerase chain reaction for the reliable detection of Australian grapevine viruses in two climates during three growing seasons. *Australian Journal of Grape and Wine Research, 18*(2), 239-244.

Cox B. A., & Jones R. A. (2010). Genetic variability in the coat protein gene of Potato virus S isolates and distinguishing its biologically distinct strains. *Archives of Virology, 155*(7), 1163-1169.

Croft H., Malinowski T., Krizbai L., Mikec I., Kajic V., Reed C., James D. (2008). Use of Luminex xMAP-derived Bio-Plex bead-based suspension array for specific detection of PPV W and characterization of epitopes on the coat protein of the virus. *Journal of Virological Methods, 153*(2), 203-213.

Daane K. M., Almeida R. P., Bell V. A., Walker J. T., Botton M., Fallahzadeh M., Walton V. M. (2012). Biology and management of mealybugs in vineyards *Arthropod Management in Vineyards:* (pp. 271-307): Springer.

Dalley B. (2008). Wine - Industry expansion, 1890s–1910s. Retrieved from http://www.TeAra.govt.nz/en/wine/page-3 [Accessed 14 February 2018]

Danielmeier T. (2008). The Winescape of New Zealand. *Pacific News, 29*, 11-13.

Daragó Á., Szabó M., Hrács K., Takács A., & Nagy P. I. (2013). In vitro investigations on the biological control of Xiphinema index with Trichoderma species. *Helminthologia, 50*(2), 132-137.

Decker Carolyn J. & Parker R. (2014). Analysis of Double-Stranded RNA from Microbial Communities Identifies Double-Stranded RNA Virus-like Elements. *Cell Reports, 7*(3), 898-906.

Demangeat G., Esmenjeaud D., Voisin R., Bidault J.-M., & Claverie M. (2005). Le court-noué de la vigne: Etat des connaissances sur cette maladie. *Phytoma, 587,* 38-42.

DePaulo J., & Powell C. (1995). Extraction of double-stranded RNA from plant tissues without the use of organic solvents. *Plant disease, 79*(3), 246-248.

Dolja V. V., Kreuze J. F., & Valkonen J. P. (2006). Comparative and functional genomics of closteroviruses. *Virus Research, 117*(1), 38-51.

Dolja V., Meng B., & Martelli G. (2017). Evolutionary Aspects of Grapevine Virology. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 659-688). Cham: Springer International Publishing.

Eichmeier A., Komínková M., Komínek P., & Baránek M. (2016). Comprehensive Virus Detection Using Next Generation Sequencing in Grapevine Vascular Tissues of Plants Obtained from the Wine Regions of Bohemia and Moravia (Czech Republic). *PLOS ONE, 11*(12), e0167966.

Engvall E., Jonsson K., & Perlmann P. (1971). Enzyme-linked immunosorbent assay. II. Quantitative assay of protein antigen, immunoglobulin g, by means of enzyme-labelled antigen and antibody-coated tubes. *Biochimica et Biophysica Acta (BBA) - Protein Structure, 251*(3), 427-434.

Engvall E. (2010). Citation Classic The ELISA, Enzyme-Linked Immunosorbent Assay. [Editorial Material]. *Clinical Chemistry, 56*(2), 319-320.

Esteves F., Teixeira Santos M., Eiras-Dias J. E., & Fonseca F. (2013). Molecular data mining to improve antibody-based detection of Grapevine leafroll-associated virus 1 (GLRaV-1). *Journal of Virological Methods, 194*(1–2), 258-270.

Fajardo T. V. M., Dianese E. C., Eiras M., Cerqueira D. M., Lopes D. B., Ferreira M. A. S. V., & Martins C. R. F. (2007). Variability of the coat protein gene of Grapevine leafroll-associated virus 3 in Brazil. *Fitopatologia Brasileira, 32*(4), 335-340.

Fajardo T. V. M., Silva F. N., Eiras M., & Nickel O. (2017). High-throughput sequencing applied for the identification of viruses infecting grapevines in Brazil and genetic variability analysis. *Tropical Plant Pathology*, *42*(4), 250-260.

Fan X., Zhang Z., Ren F., Hu G., Li Z., & Dong Y. (2017). First Report of Grapevine geminivirus A from Grapevines in China. *Plant Disease, 101*(7), 1333-1333.

Fereres A., & Raccah B. (2015). Plant virus transmission by insects. *eLS*.

Filloux D., Dallot S., Delaunay A., Galzi S., Jacquot E., & Roumagnac P. (2015). Metagenomics Approaches Based on Virion-Associated Nucleic Acids (VANA): An Innovative Tool for Assessing Without A Priori Viral Diversity of Plants. In C. Lacomme (Ed.), *Plant Pathology: Techniques and Protocols* (pp. 249-257). New York, NY: Springer New York.

Fiore N., Prodan S., & Pino A. (2009). Monitoring grapevine viruses by ELISA and RT-PCR throughout the year. *Journal of Plant Pathology*, 489-493.

Fleischmann R. D., Adams M. D., White O., Clayton R. A., Kirkness E. F., Kerlavage A. R., et al. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science, 269*(5223), 496-512.

Fletcher J., Bulman S., Fletcher P., & Houliston G. (2009). First record of Turnip mosaic virus in Cook's scurvy grass (Lepidium oleraceum agg.)—an endangered native plant in New Zealand. *Australasian Plant Disease Notes, 4*(1), 9-11.

Fletcher J., Lister R., Bulman S., & Heenan P. (2010). First record of Turnip mosaic virus in Pachycladon spp.(Brassicaceae): an endangered native plant species in New Zealand. *Australasian Plant Disease Notes, 5*(1), 9-10.

Fontana J., López-Montero N., Elliott R. M., Fernández J. J., & Risco C. (2008). The unique architecture of Bunyamwera virus factories around the Golgi complex. *Cellular Microbiology, 10*(10), 2012-2028.

FreshFacts. (2016). FreshFacts New Zealand Horticulture. Retrieved from http://www.freshfacts.co.nz/files/freshfacts-2016.pdf [Accessed 5 February 2018]

Gambino G., & Gribaudo I. (2006). Simultaneous detection of nine grapevine viruses by multiplex reverse transcription-polymerase chain reaction with coamplification of a plant RNA as internal control. *Phytopathology, 96*(11), 1223-1229.

Gambino G., Cuozzo D., Fasoli M., Pagliarani C., Vitali M., Boccacci P., & Mannini F. (2012). Co-evolution between Grapevine rupestris stem pitting-associated virus and Vitis vinifera L. leads to decreased defence responses and increased transcription of genes related to photosynthesis. *Journal of Experimental Botany, 63*(16), 5919-5933.

Ghanem-Sabanadzovic A. N., Sabanadzovic S., Gugerli P. & Rowhani A. (2012). Genome organization, serology and phylogeny of Grapevine leafroll-associated viruses 4 and 6: Taxonomic implications. *Virus Research 163*(1):120-8.

Ghannam A., Kumari S., Muyldermans S., & Abbady A. Q. (2015). Camelid nanobodies with high affinity for broad bean mottle virus: a possible promising tool to immunomodulate plant resistance against viruses. *Plant Molecular Biology, 87*(4), 355-369.

Garcia-Luque I., Brieva A., Diazruiz J. R., & Rubio N. (1986). Isolation and Partial Characterization of a Monoclonal-Antibody Specific for a Naturally-Occurring Double-Stranded-RNA. *Virology, 152*(1), 252-255.

Giampetruzzi A., Roumi V., Roberto R., Malossini U., Yoshikawa N., La Notte P., & Saldarelli P. (2012). A new grapevine virus discovered by deep sequencing of virus- and viroid-derived small RNAs in Cv Pinot gris. *Virus Research, 163*(1), 262-268.

Giampetruzzi A., Morelli M., Chiumenti M., Savino V., Martelli G., La Notte P., & Saldarelli P. (2015). *Towards the definition of the absolute sanitary status of certified grapevine clones and rootstocks. grapevine clones and rootstocks.* Paper presented at the Proceedings of the 18th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Ankara, Turkey.

Gibb R. (2014). Beaux Cailloux: It's not goodbye, it's see you later. Retrieved from http://rebeccagibb.com/beaux-cailloux-its-not-goodbye-its-see-you-later/ [Accessed 26 February 2018]

Giribaldi M., Purrotti M., Pacifico D., Santini D., Mannini F., Caciagli P., & Marzachi C. (2011). A multidisciplinary study on the effects of phloem-limited viruses on the agronomical performance and berry quality of Vitis vinifera cv. Nebbiolo. *Journal of Proteomics, 75*(1), 306-315.

Glasa M., Predajna L., Kominek P., Nagyova A., Candresse T., & Olmos A. (2014). Molecular characterization of divergent grapevine Pinot gris virus isolates and their detection in Slovak and Czech grapevines. *Archives of Virology, 159*(8), 2103-2107.

Glasa M., Predajňa L., Šoltys K., Sihelská N., Nagyová A., Wetzel T., & Sabanadzovic S. (2017). Analysis of Grapevine rupestris stem pitting-associated virus in Slovakia Reveals Differences in Intra-Host Population Diversity and Naturally Occurring Recombination Events. *The Plant Pathology Journal, 33*(1), 34-42.

Goffeau A., Barrell B. G., Bussey H., Davis R. W., Dujon B., Feldmann H., & Oliver S. G. (1996). Life with 6000 genes. *Science, 274*(5287), 546, 563-547.

Golino D., Fuchs M., Al Rwahnih M., Farrar K., Schmidt A. & Martelli G. (2017). Regulatory aspects of grape viruses and virus diseases: certification, quarantine, and harmonization. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 581-598). Cham: Springer International Publishing.

Gomez Talquenca G. S., Moyano S., & Lanza Volpe M. (2015). *Serological relationships among GLRaV-4 strains reflects the genetic variability of its coat protein gene.* Paper presented at the Proceedings of the 18th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Ankara, Turkey.

Goszczynski D., Du Preez J., & Burger J. (2008). Molecular divergence of Grapevine virus A (GVA) variants associated with Shiraz disease in South Africa. *Virus Research, 138*(1-2), 105-110.

Goszczynski D., & 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, 61*(1), 205-214.

Gouveia P., Santos M. T., Eiras-Dias J. E., & 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, 156*(3), 413-420.

Gouveia P., Dandlen S., Costa Â., Marques N., & Nolasco G. (2012). Identification of an RNA silencing suppressor encoded by Grapevine leafroll-associated virus 3. *European Journal of Plant Pathology, 133*(1), 237-245.

Grabherr M. G., Haas B. J., Yassour M., Levin J. Z., Thompson D. A., Amit I., & Regev A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology, 29*(7), 644-652.

Gribaudo I., Gambino G., Cuozzo D., & Mannini F. (2006). Attempts to eliminate Grapevine rupestris stem pitting-associated virus from grapevine clones. *Journal of Plant Pathology*, 293-298.

Gualandri V., Asquini E., Bianchedi P., Covelli L., Brilli M., Malossini U., & Si-Ammour A. (2017). Identification of herbaceous hosts of the Grapevine Pinot gris virus (GPGV). *European Journal of Plant Pathology, 147*(1), 21-25.

Guindon S., & Gascuel O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology, 52*(5), 696-704.

Guindon S., Dufayard J.-F., Lefort V., Anisimova M., Hordijk W., & Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology, 59*(3), 307-321.

Guo Q., Honesty S., Xu M. L., Zhang Y., Schoelz J., & Qiu W. (2014). Genetic Diversity and Tissue and Host Specificity of Grapevine vein clearing virus. *Phytopathology, 104*(5), 539-547.

Habili N., Symons R. H., & Cohen D. (2002). First detection of Grapevine leafroll virus types 2 and 5 in New Zealand by Waite Diagnostics. *Australian and New Zealand Grapegrower and Winemaker* (459), 23-27.

Habili N. (2013). Australian Shiraz Disease: an emerging virus disease of Vitis vinifera cv. Shiraz. *Wine and Viticulture Journal, 28*(1), 59-61.

Habili N., Wu Q., & Pagay V. (2016). Virus-associated Shiraz Disease may lead Shiraz to become an endangered variety in Australia. *Wine & Viticulture Journal, 31*(1), 47.

Hafrén A.; Lõhmus A.& Mäkinen K. (2015). Formation of Potato virus A-induced RNA granules and viral translation are interrelated processes required for optimal virus accumulation. *PLoS Pathogens* 11(12):e1005314.

Hamers-Casterman C., Atarhouch T., Muyldermans S., Robinson G., Hammers C., Songa E. B., & Hammers R. (1993). Naturally occurring antibodies devoid of light chains. *Nature, 363*(6428), 446-448.

Hemmer C., Ackerer L., Marmonier A., Hleibieh K., Schmitt-Keichinger C., Vigne E., & Ritzenthaler C. (2015). Nanobodies: getting over the hump of fanleaf degeneration? Paper presented at the Proceedings of the 18th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Ankara, Turkey.

Hobson-Peters J., Yam A. W. Y., Lu J. W. F., Setoh Y. X., May F. J., Kurucz N., & Weir R. (2013). A new insect-specific flavivirus from northern Australia suppresses replication of West Nile virus and Murray Valley encephalitis virus in co-infected mosquito cells. *PloS ONE, 8*(2), e56534.

Hoskins N., & Thorpe G. (2010a). Clones of Classic Varieties : The Chardonnay Portfolio. Retrieved from http://www.riversun.co.nz/assets/Uploads/Library/Riversun-articles/2010-The-Chardonnay-Portfolio.pdf [Accessed 4 February 2018]

Hoskins N., & Thorpe G. (2010b). Clones of Classic Varieties : The Sauvignon blanc Portfolio. Retrieved from http://www.riversun.co.nz/assets/Uploads/Library/Riversun-articles/2010-The-Sauvignon-Blanc-Portfolio.pdf [Accessed 6 February 2018]

Hoskins N., Thompson C., & Andrew R. (2011). New Zealand Winegrowers working hard to eliminate leafroll. *Australian & New Zealand Grapegrower & Winemaker 569,* 23-25.

Howland P. J. (2014). *Social, cultural and economic impacts of wine in New Zealand*: Routledge.

Hu Q. A., Hollunder J., Niehl A., Korner C. J., Gereige D., Windels D., & Heinlein M. (2011). Specific Impact of Tobamovirus Infection on the Arabidopsis Small RNA Profile. *Plos ONE, 6*(5).

Hull R. (2014). Chapter 8 - Origins and Evolution of Plant Viruses. In R. Hull (Ed.), *Plant Virology (Fifth Edition)* (pp. 423-476). Boston: Academic Press.

Hunter A., Macgregor A. B., Szabo T. O., Wellington C. A., & Bellgard M. I. (2012). Yabi: An online research environment for grid, high performance and cloud computing. *Source Code Biol. Med., 7*.

Huss B., Walter B., & Fuchs M. (1989). Cross-protection between arabis mosaic virus and grapevine fan leaf virus isolates in Chenopodium quinoa. *Annals of Applied Biology, 114*(1), 45-60.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860.

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature, 431*(7011), 931.

International HapMap C. (2005). A haplotype map of the human genome. *Nature, 437*(7063), 1299-1320.

International Organisation of Vine and Wine. (2017). Elements de conjoncture mondiale, . Retrieved from http://www.oiv.int/public/medias/5264/oiv-noteconjmars2017-fr.pdf [Accessed 5 February 2018]

Jo Y., Choi H., Cho J. K., Yoon J.-Y., Choi S.-K., & Cho W. K. (2015). In silico approach to reveal viral populations in grapevine cultivar Tannat using transcriptome data. *Scientific Reports, 5*, 15841.

Jo Y., Song M.-K., Choi H., Park J.-S., Lee J.-W., Lian S., & Cho W. K. (2017a). Genome Sequence of Grapevine Virus K, a Novel Vitivirus Infecting Grapevine. *Genome Announcements, 5*(37), e00994-00917.

Jo Y., Song M.-K., Choi H., Park J.-S., Lee J.-W., Lian S., Lee B. C.& Cho W. K. (2017b). Genome Sequence of Grapevine Virus T, a Novel Foveavirus Infecting Grapevine. *Genome Announcements* 5(37):e00995-17.

Jo Y., Choi H., Song M., Park J., Lee J., & Cho W. (2018). First report of Grapevine geminivirus A in diverse Vitis species in Korea. *Plant Disease, 102*(1), 255.

Jooste A. E., Maree H. J., Bellstedt D. U., Goszczynski D. E., Pietersen G., & Burger J. T. (2010). Three genetic grapevine leafroll-associated virus 3 variants identified from South African vineyards show high variability in their 5'UTR. *Archives of Virology*, 155(12), 1997-2006.

Jooste A. E., Molenaar N., Maree H. J., Bester R., Morey L., de Koker W. C., & Burger J. T. (2015). Identification and distribution of multiple virus infections in Grapevine leafroll diseased vineyards. *European Journal of Plant Pathology, 142*(2), 363-375.

Jordan R., & Hammond J. (1991). Comparison and Differentiation of Potyvirus Isolates and Identification of Strain-Specific, Virus-Specific, Subgroup-Specific and Potyvirus Group-Common Epitopes Using Monoclonal-Antibodies. *Journal of General Virology, 72*, 25-36.

Kesanakurti P., Belton M., Saeed H., Rast H., Boyes I., & Rott M. (2016). Screening for plant viruses by next generation sequencing using a modified double strand RNA extraction protocol with an internal amplification control. *Journal of Virological Methods, 236*, 35-40.

Kobayashi K., Tomita R., & Sakamoto M. (2009). Recombinant plant dsRNA-binding protein as an effective tool for the isolation of viral replicative form dsRNA and universal detection of RNA viruses. *Journal of General Plant Pathology, 75*(2), 87-91.

Komar V., Vigne E., Demangeat G., & Fuchs M. (2007). Beneficial effect of selective virus elimination on the performance of Vitis vinifera cv. Chardonnay. *American Journal of Enology and Viticulture, 58*(2), 202-210.

Komar V., Vigne E., Demangeat G., & Lemaire O. (2008). Cross-Protection as Control Strategy Against Grapevine fanleaf virus in Naturally Infected Vineyards. *Plant Disease, 92*(12), 1689-1694.

Komar V., Vigne E., Demangeat G., Lemaire O., & Fuchs M. (2010). Comparative Performance of Virus-Infected Vitis vinifera cv. Savagnin rose Grafted onto Three Rootstocks. *American Journal of Enology and Viticulture, 61*(1), 68-73.

Komínek P., & Bryxiová M. (2004). Comparison of three techniques for detection of grapevine leafroll-associated virus 1. *Acta virologica, 49*(1), 37-43.

Krake L. R., Scott N. S., Rezaian M., & Taylor R. (1999). *Graft-transmitted diseases of grapevines*: CSIRO PUBLISHING.

Krenz B., Thompson J. R., Fuchs M., & Perry K. L. (2012). Complete genome sequence of a new circular DNA virus from grapevine. *Journal of Virology, 86*(14), 7715.

Kreuze J. F., Perez A., Untiveros M., Quispe D., Fuentes, Barker, & Simon R. (2009). Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. *Virology, 388*.

Langmead B., & Salzberg S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods, 9*(4), 357-359.

Le Maguet J., Fuchs J. J., Chadoeuf J., Beuve M., Herrbach E., & Lemaire O. (2013). The role of the mealybug Phenacoccus aceris in the spread of Grapevine leafroll-associated virus-1 (GLRaV-1) in two French vineyards. *European Journal of Plant Pathology, 135*(2), 415-427.

Lee M., Etebari K., Hall-Mendelin S., van den Hurk A. F., Hobson-Peters J., Vatipally S., Schnettler E., Hall R.& Asgari S. (2017). Understanding the role of microRNAs in the interaction of Aedes aegypti mosquitoes with an insect-specific flavivirus. *Journal of General Virology* 98(7):1892-1903.

Legin R., Bass P., Etienne L., & Fuchs M. (1993). Selection of mild virus-strains of fanleaf degeneration by comparative field performance of infected grapevines. *Vitis, 32*(2), 103-110.

Li R., Gao S., Hernandez A. G., Wechter W. P., Fei Z., & Ling K. S. (2012). Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *Plos ONE, 7*(5), e37127.

Liebenberg A., Freeborough M.-J., Visser J., Bellstedt D. U., & Burger J. T. (2009). Genetic variability within the coat protein gene of Grapevine fanleaf virus isolates from South Africa and the evaluation of RT-PCR, DAS-ELISA and ImmunoStrips as virus diagnostic assays. *Virus Research, 142*(1–2), 28-35.

Lima M. (2006). *Detection of rupestris stem pitting-associated virus in seedlings of virus-infected maternal grapevine plants.* Paper presented at the Proceedings of the 15th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine, Stellenbosch, 3-7 April, 2006.

Lima F., Alkowni, Uyemoto J. K., Golino D., Osman F., & Rowhani A. (2006). Molecular analysis of a California strain of Rupestris stem pitting-associated virus isolated from declining Syrah grapevines. *Archives of Virology, 151*(9), 1889-1894.

Lima M., Alkowni R., Uyemoto J. K., & Rowhani A. (2009). Genomic study and detection of a new variant of Grapevine rupestris stem pitting associated virus in declining california pinot noir grapevines. *Journal of Plant Pathology, 91*(1), 155-162.

Loconsole G., Saldarelli P., Doddapaneni, Savino, Martelli G. P., & Saponari M. (2012). Identification of a single-stranded DNA virus associated with citrus chlorotic dwarf disease, a new member in the family Geminiviridae. *Virology, 432*(1), 162-172.

Morelli, de Moraes Catarino A., Susca, Saldarelli P., Gualandri V. Martelli G.P. (2014). First Report of Grapevine Pinot Gris Virus from Table Grapes in Southern Italy. *Journal of Plant Pathology*, 96(2).

MacDiarmid R., Cohen (2007). Viruses in New Zealand Grapevines. *New Zealand Winegrower, 11,* 40-43.

Macdiarmid R., Rodoni B., Melcher U., Ochoa-Corona F., & Roossinck M. (2013). Biosecurity implications of new technology and discovery in plant virus research. *PLoS Pathogens, 9*(8), e1003337-e1003337.

Maliogka V. I., Olmos A., Pappi P. G., Lotos L., Efthimiou K., Grammatikaki G., & Avgelis A. D. (2015). A novel grapevine badnavirus is associated with the Roditis leaf discoloration disease. *Virus Research, 203*, 47-55.

Marais A., Faure C., Couture C., Bergey B., Gentit P., & Candresse T. (2014). Characterization by Deep Sequencing of Divergent Plum bark necrosis stem pitting-associated virus (PBNSPaV) Isolates and Development of a Broad-Spectrum PBNSPaV Detection Assay. *Phytopathology, 104*(6), 660-666.

Marais A., Faure C., Mustafayev E., & Candresse T. (2015). Characterization of New Isolates of Apricot vein clearing-associated virus and of a New Prunus-Infecting Virus: Evidence for Recombination as a Driving Force in Betaflexiviridae Evolution. *PLoS ONE, 10*(6), e0129469.

Marchler-Bauer A., Lu S., Anderson J. B., Chitsaz F., Derbyshire M. K., DeWeese-Scott C., & Bryant S. H. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research, 39*, D225-D229.

Mardis E. R. (2013). Next-generation sequencing platforms. *Annual Review of Analytical Chemistry, 6*, 287-303.

Maree H.J., Almeida R.P., Bester R., Chooi K.M., Cohen D., Dolja V.V., Fuchs M.F., Golino D.A., Jooste A.E., Martelli G.P. & Naidu R.A., (2013). Grapevine leafroll-associated virus 3. *Frontiers in Microbiology*, 4, p.82.

Maree H. J., Pirie M. D., Oosthuizen K., Bester R., Rees D. J., & Burger J. T. (2015). Phylogenomic analysis reveals deep divergence and recombination in an economically important grapevine virus. *PLoS ONE, 10*(5), e0126819.

Marsh M., Tu O., Dolnik V., Roach D., Solomon N., Bechtol K., & Bashkin J. (1997). High-throughput DNA sequencing on a capillary array electrophoresis system. *Journal of Capillary Electrophoresis, 4*(2), 83-89.

Martelli, GP, Minafra A., & Saldarelli P. (1997). Vitivirus, a new genus of plant viruses. *Archives of Virology, 142*(9), 1929-1932.

Martelli, GP. & Boudon-Padieu E. (2007). Directory of infectious diseases of grapevines. *International Council for the Study of Virus and Virus-like Diseases of Grapevine (ICVG)*.

Martelli, GP., Abou Ghanem-Sabanadzovic N., Agranovsky A. A., Al Rwahnih M., Dolja V. V., Dovas C. I., & Saldarelli P. (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, 94*(1), 7-19.

Martelli GP (2014a). Grapevine leafroll. *Journal of Plant Pathology, 96*(1sup), S51-S70.

Martelli, GP. (2014b). Infectious degeneration (Grapevine fanleaf virus). *Journal of Plant Pathology, 96*(1sup), S11-S27.

Martelli GP (2014c). Rugose wood complex. *Journal of Plant Pathology, 96*(1sup), S73-S88.

Martelli GP (2014d). Grapevine-infecting viruses. *Journal of Plant Pathology*, 96(1sup), S7-S8.

Martelli, GP. (2017). An Overview on Grapevine Viruses, Viroids, and the Diseases They Cause *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 31-46): Springer.

Massart S., Olmos A., Jijakli H., & Candresse T. (2014). Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Research, 188*(0), 90-96.

Massart S., Candresse T., Gil J., Lacomme C., Predajna L., Ravnikar M., & Wetzel T. (2017). A Framework for the Evaluation of Biosecurity, Commercial, Regulatory, and Scientific Impacts of Plant Viruses and Viroids Identified by NGS Technologies. [Perspective]. *Frontiers in Microbiology, 8*(45).

Maxam A., & Gilbert W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America, 74*(2), 560-564.

McKissock A. (1964). Two important virus diseases of grapevines in New Zealand. *New Zealand Journal of Agricultural, 108*, 332-339.

Meng B., Pang S., Forsline P. L., McFerson J. R., & Gonsalves D. (1998). Nucleotide sequence and genome structure of grapevine rupestris stem pitting associated virus-1 reveal similarities to apple stem pitting virus. *Journal of General Virology, 79*(8), 2059-2069.

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. J *Journal of General Virology 87*(Pt 6):1725-33.

Meng B., & Gonsalves D. (2007). Grapevine rupestris stem pitting-associated virus: a decade of research and future perspectives. *Plant Viruses, 1*(1), 52-62.

Meng B. Z., Venkataraman S., Li C. H., Wang W. Z., Dayan-Glick C., & Mawassi M. (2013). Construction and biological activities of the first infectious cDNA clones of the genus Foveavirus. *Virology, 435*(2), 453-462.

Meng B., & Rowhani A. (2017). Grapevine rupestris stem pitting-associated virus *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 257-287): Springer.

Minafra A., Mawassi M., Goszczynski D., & Saldarelli P. (2017). Grapevine Vitiviruses. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 229-256). Cham: Springer International Publishing.

Minutillo S. A., Marais A., Mascia T., Faure C., Svanella-Dumas L., Theil S., & Candresse T. (2015). Complete Nucleotide Sequence of Artichoke latent virus Shows it to be a Member of the Genus Macluravirus in the Family Potyviridae. *Phytopathology, 105*(8), 1155-1160.

Miozzi L., Gambino G., Burgyan J., & Pantaleo V. (2013). Genome-wide identification of viral and host transcripts targeted by viral siRNAs in Vitis vinifera. *Molecular Plant Pathology, 14*(1), 30-43.

Moffitt E., & Lister R. (1975). Application of a serological screening test for detecting double-stranded RNA mycoviruses [from Penicillium chrysogenum]. *Phytopathology, 65*.

Montero R., Mundy D., Albright A., Grose C., Trought M., Cohen D., & Bota J. (2016). Effects of Grapevine Leafroll associated Virus 3 (GLRaV-3) and duration of infection on fruit composition and wine chemical profile of Vitis vinifera L. cv. Sauvignon blanc. *Food Chemistry, 197*, 1177-1183.

Morelli M., Minafra A., Boscia D., & Martelli G. P. (2011). Complete nucleotide sequence of a new variant of grapevine rupestris stem pitting-associated virus from southern Italy. *Archives of Virology, 156*(3), 543-546.

Morris T. J., & Dodds J. A. (1979). Isolation and analysis of double-stranded-rna from virus-infected plant and fungal tissue. *Phytopathology, 69*(8), 854-858.

Mossop D. W. (1986). Grapevine viruses. Grapevine fleck joins the bunch. *Southern Horticulture, Grapegrower & Winemaker, 3,* 30-31.

Mullis K. F., Scharf S., Saiki R., Horn G., Erlich H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol, 51*(1), 263-273.

Myers E. W., Sutton G. G., Delcher A. L., Dew I. M., Fasulo D. P., Flanigan J., & Venter J. C. (2000). A whole-genome assembly of Drosophila. *Science, 287*(5461), 2196-2204.

Naidu R., Rowhani A., Fuchs, Golino D., & Martelli G. P. (2014). Grapevine Leafroll: A Complex Viral Disease Affecting a High-Value Fruit Crop. *Plant Disease, 98*(9), 1172-1185

Nakaune R., Toda S., Mochizuki M., & Nakano M. (2008). Identification and characterization of a new vitivirus from grapevine. *Archives of Virology, 153*(10), 1827-1832.

Naraghi-Arani P., Daubert S., & Rowhani A. (2001). Quasispecies nature of the genome of Grapevine fanleaf virus. *Journal of General Virology, 82*(7), 1791-1795.

New Zealand Parliamentary (1993) Biosecurity Act 1993. No. 95 119p.

New Zealand Winegrowers. (2017a). Annual report Retrieved from https://www.nzwine.com/media/6600/nzw-annual-report-2017.pdf [Accessed 3 February 2018]

New Zealand Winegrowers. (2017b). New Zealand Winegrowers Grafted Grapevine Standard version 3.1. Retrieved from https://www.nzwine.com/media/6066/grafted-grapevine-standard-v31-1617.pdf [Accessed 11 September 2017]

Nimmo-Bell & Company Ltd. (2006). The economic effects and financial impact of GLRaV3. *Report for New Zealand Winegrowers*.

Nolasco G., Deblas C., Torres V., & Ponz F. (1993). A Method Combining Immunocapture and Pcr Amplification in a Microtiter Plate for the Detection of Plant-Viruses and Subviral Pathogens. *Journal of Virological Methods, 45*(2), 201-218.

Nolasco G., Santos C., Petrovic N., Teixeira Santos M., Cortez I., Fonseca F., & Sequeira O. (2006). Rupestris stem pitting associated virus isolates are composed by mixtures of genomic variants which share a highly conserved coat protein. *Archives of Virology, 151*(1), 83-96.

O'Brien C. A., Hobson-Peters J., Yam A. W., Colmant A. M., McLean B. J., Prow N. A., & Hall R. A. (2015). Viral RNA intermediates as targets for detection and discovery of novel and emerging mosquito-borne viruses. *PLoS Neglected Tropical Diseases, 9*(3), e0003629.

Oliveira L. M., Orilio A. F., Inoue-Nagata A. K., Nagata T., & Blawid R. (2017). A novel vitivirus-like sequence found in Arracacia xanthorrhiza plants by high throughput sequencing. *Archives of Virology, 162*(7), 2141-2144.

Oliver J. E., Vigne E., & Fuchs M. (2010). Genetic structure and molecular variability of Grapevine fanleaf virus populations. *Virus Research, 152*(1–2), 30-40.

Oliver J. E., & 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, 62*(4), 438-451

Oosthuizen K. (2017). Characterising the viral and microbial diversity in old and young grapevines. Available http://scholar.sun.ac.za/bitstream/handle/10019.1/102827/oosthuizen_characterising_2017.pdf?sequence=1 [Access 28 February 2018].

Over de Linden A. J., & Chamberlain E. E. (1970). Effect of grape-vine leafroll virus on vine growth and fruit yield and quality. *New Zealand Journal of Agricultural Research, 13*(3), 689-698.

Pantaleo V., Saldarelli P., Miozzi L., Giampetruzzi A., Gisel A., Moxon S., & Burgyan J. (2010). Deep sequencing analysis of viral short RNAs from an infected Pinot Noir grapevine. *Virology, 408*(1), 49-56.

Parliamentary Debates. (1986). House of Representatives, Volume 473, 31st July 1986. Retrieved from https://books.google.co.nz/books?id=wlGSFDf-DV8C&pg=PA3264&lpg=PA3264&dq=grapevine+removal+scheme+in+New+Zealand&source=bl&ots=N8sIMUgrqn&sig=TUMF5XMSQu0KaTTwRrPpaIuKL5Q&hl=en&sa=X&ved=0ahUKEwia4q67pZzZAhUDbrwKHQjmAo8Q6AEIPDAE#v=onepage&q=grapevine%20removal%20scheme%20in%20New%20Zealand&f=false [Accessed 15 February 2018]

Pearson M. N., Clover G. R. G., Guy P. L., Fletcher J. D., & Beever R. E. (2006). A review of the plant virus, viroid and mollicute records for New Zealand. *Australasian Plant Pathology, 35*(2), 217-252.

Perry K. L., McLane H., Thompson J. R., & Fuchs M. (2018). A novel grablovirus from non-cultivated grapevine (Vitis sp.) in North America. *Archives of Virology, 163*(1), 259-262.

Petersen C. L., & Charles J. G. (1997). Transmission of grapevine leafroll-associated closteroviruses by Pseudococcus longispinus and P-calceolariae. *Plant Pathology, 46*(4), 509-515.

Pietersen G., Bell V., & Krüger K. (2017). Management of Grapevine Leafroll Disease and Associated Vectors in Vineyards. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 531-560). Cham: Springer International Publishing.

Pietersen G., Spreeth N., Oosthuizen T., van Rensburg A., van Rensburg M., Lottering D., & Tooth D. (2013). Control of Grapevine Leafroll Disease Spread at a Commercial Wine Estate in South Africa: A Case Study. *American Journal of Enology and Viticulture, 64*(2), 296-305.

Powell C. (1991). Preparation and characterization of monoclonal antibodies to double-stranded RNA. *Phytopathology, 81*(2), 184-187.

Prator C. A., Kashiwagi C. M., Vončina D., & Almeida R. P. (2017). Infection and Colonization of Nicotiana benthamiana by Grapevine leafroll-associated virus 3. *Virology, 510*, 60-66.

Qiu W. & Schoelz J. (2017). Grapevine vein clearing virus: Diagnostics, Genome, Genetic Diversity, and Management. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 315-330). Cham: Springer International Publishing.

Raski D., Goheen A., Lider L., & Meredith C. (1983). Strategies against grapevine fanleaf virus and its nematode vector. *Plant Disease, 67*(3), 335-338.

Rayapati N., O'Neal S., & Walsh D. (2008). Grapevine leafroll disease. Retrieved from http://cru.cahe.wsu.edu/CEPublications/eb2027e/eb2027e.pdf [Accessed 28 February 2018]

Reynard J.-S., Schneeberger P. H., Frey J. E., & Schaerer S. (2015). Biological, Serological, and Molecular Characterization of a Highly Divergent Strain of Grapevine leafroll-associated virus 4 Causing Grapevine Leafroll Disease. *Phytopathology, 105*(9), 1262-1269.

Reynard J. S., Brodard J., Dubuis N., Yobregat O., Kominek P., Schumpp O., & Schaerer S. (2017). First Report of Grapevine rupestris vein feathering virus in Swiss Grapevines. *Plant Disease*, PDIS-01-17-0140-PDN.

Reynolds A. (2017). The Grapevine, Viticulture, and Winemaking: A Brief Introduction. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 3-29). Cham: Springer International Publishing.

Richter J., Rabenstein F., Proll E., & Vetten H. J. (1995). Use of Cross-Reactive Antibodies to Detect Members of the Potyviridae. *Journal of Phytopathology, 143*(8), 459-464.

Ricketts K. D., Gomez M. I., Atallah S. S., Fuchs M. F., Martinson T. E., Battany M. C., & Smith R. J. (2015). Reducing the Economic Impact of Grapevine Leafroll Disease in California: Identifying Optimal Disease Management Strategies. *American Journal of Enology and Viticulture, 66*(2), 138-147.

Ricketts K. D., Gómez M. I., Fuchs M. F., Martinson T. E., Smith R. J., Cooper M. L., & Wise A. (2016). Mitigating the Economic Impact of Grapevine Red Blotch: Optimizing Disease Management Strategies in U.S. Vineyards. *American Journal of Enology and Viticulture*.

Roossinck M. J., Saha P., Wiley G. B., Quan J., White J. D., Lai H., & Roe B. A. (2010). Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology, 19*, 81-88.

Roossinck M. J. (2011). The good viruses: viral mutualistic symbioses. *Nature Reviews Microbiology* 9(2):99-108.

Roossinck M. J. (2012). Plant Virus Metagenomics: Biodiversity and Ecology. *Annual Review of Genetics, 46*, 359-369.

Roossinck M. J. (2017). Deep sequencing for discovery and evolutionary analysis of plant viruses. *Virus Research, 239*, 82-86.

Rott M., Xiang Y., Boyes I., Belton M., Saeed H., Kesanakurti P., & Rast H. (2017). Application of Next Generation Sequencing for Diagnostic Testing of Tree Fruit Viruses and Viroids. *Plant Disease, 101*(8), 1489-1499.

Rowhani A., Daubert S., Arnold K., Al Rwahnih M., Klaassen V., Golino D., & Uyemoto J. K. (2018). Synergy between grapevine vitiviruses and grapevine leafroll viruses. *European Journal of Plant Pathology*.

Rowhani A., Zhang Y., Minafra A., Golino DA, & Uyemoto J. (2000). Grapevine rupestris stem pitting associated virus: population diversity, titer in the host and possible transmission vector. Paper presented at the Proceedings of the 13th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Adelaide, Australia.

Rowhani A., La Notte P., Uyemoto J., Daubert S., & Savino V. (2017). Biological Assays. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp.395-407). Cham: Springer International Publishing.

Rumbos I. C., & Avgelis A. (1989). Roditis Leaf Discoloration-A New Virus Disease of Grapevine: Symptomatology and Transmission to Indicator Plants. *Journal of Phytopathology, 125*(3), 274-278.

Sabanadzovic S., Aboughanem-Sabanadzovic N., & Martelli G. (2017). Grapevine fleck and similar viruses *Grapevine Viruses.* In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 331-349). Cham: Springer International Publishing.

Saldarelli P., Gualandri V., & Martelli G. P. (2015). *Grapevine leaf mottling and deformation and Grapevine Pinot gris virus: an update on an emerging Mediterranean diseas and a new virus.* Paper presented at the Proceedings of the 18th Meeting of the International Council for the Study of Viruses and Virus Diseases of the Grapevine (ICVG), Ankara, Turkey.

Saldarelli P., Giampetruzzi A., Maree H., & Al Rwahnih M. (2017). High-Throughput Sequencing: Advantages Beyond Virus Identification. In B. Meng G. P. Martelli D. A. Golino & M. Fuchs (Eds.), *Grapevine Viruses: Molecular Biology, Diagnostics and Management* (pp. 625-642) Cham: Springer International Publishing.

Sandanayaka W. R. M., Blouin A. G., Prado E., & Cohen D. (2012). Stylet penetration behaviour of *Pseudococcus longispinus* in relation to acquisition of grapevine leafroll virus 3. *Arthropod-Plant Interactions, 7*(2), 137-146.

Sanger F., Air G. M., Barrell B. G., Brown N. L., Coulson A. R., Fiddes J. C., & Smith M. (1977). Nucleotide sequence of bacteriophage [phi]X174 DNA. *Nature, 265*(5596), 687-695.

Sanger, Coulson A. R., Hong G. F., Hill D. F., & Petersen G. B. (1982). Nucleotide sequence of bacteriophage λ DNA. *Journal of Molecular Biology, 162*(4), 729-773.

Schnell I. B., Bohmann K., & Gilbert T. P. (2015). Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources, 15*(6), 1289-1303.

Schonborn, Oberstrass J., Breyel E., Tittgen J., Schumacher J., & Lukacs N. (1991). Monoclonal-Antibodies to Double-Stranded-Rna as Probes of Rna Structure in Crude Nucleic-Acid Extracts. *Nucleic Acids Research, 19*(11), 2993-3000.

Seah Y. M., Sharma A. M., Zhang S. M., Almeida R. P. P., & Duffy S. (2012). A divergent variant of Grapevine leafroll-associated virus 3 is present in California. *Virology Journal, 9*(235).

Seo J.-K., Kang M., Kwak H.-R., Kim M.-K., Kim C.-S., Lee S.-H., & Choi H.-S. (2015). Complete genome sequence of motherwort yellow mottle virus, a novel putative member of the genus Torradovirus. *Archives of Virology, 160*(2), 587-590.

Sevin A., Boursiquot J., & Spilmont A. (2012). *GRSPaV, a prevalent virus in most of the Vitis species*. Paper presented at the Proceedings of the 17th Congress of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG) Davis, California, USA.

Shepherd D. N., Martin D. P., Lefeuvre P., Monjane A. L., Owor B. E., Rybicki E. P., & Varsani A. (2008). A protocol for the rapid isolation of full geminivirus genomes from dried plant tissue. *Journal of Virological Methods, 149*(1), 97-102.

Silva J. M. F., Al Rwahnih M., Blawid R., Nagata T., & Fajardo V. M. (2017). Discovery and molecular characterization of a novel enamovirus, Grapevine enamovirus-1. *Virus Genes, 53*(4), 667-671.

Singh K., Talla A., & Qiu W. P. (2012). Small RNA profiling of virus-infected grapevines: evidences for virus infection-associated and variety-specific miRNAs. *Functional & Integrative Genomics, 12*(4), 659-669.

Smart R. E. (1985). Oveseas specialist gives 'thumbs up' to NZ V.I. scheme. *Southern Horticulture, Grapegrower & Winemaker, 2,* 65.

Smith L. M., Sanders J. Z., Kaiser R. J., Hughes P., Dodd C., Connell C. R., & Hood L. E. (1986). Fluorescence detection in automated DNA sequence analysis. [10.1038/321674a0]. *Nature, 321*(6071), 674-679.

Sokhandan-Bashir N., & Melcher U. (2012). Population genetic analysis of grapevine fanleaf virus. *Archives of Virology, 157*(10),

Son K. N., Liang Z., & Lipton H. L. (2015). Double-Stranded RNA Is Detected by Immunofluorescence Analysis in RNA and DNA Virus Infections, Including Those by Negative-Stranded RNA Viruses. *Journal of Virology, 89*(18), 9383-9392.

Sudarshana M., Perry K. L., & Fuchs M. (2015). Grapevine red blotch-associated virus, an emerging threat to the grapevine industry. *Phytopathology*.

Tatineni S., McMechan A. J., Wosula E. N., Wegulo S. N., Graybosch R. A., French R., & Hein G. L. (2014). An eriophyid mite-transmitted plant virus contains eight genomic RNA segments with unusual heterogeneity in the nucleocapsid protein. *Journal of Virology, 88*(20), 11834-11845.

Teh Y.-H. A., & Kavanagh T. A. (2010). High-level expression of Camelid nanobodies in Nicotiana benthamiana. *Transgenic Research, 19*(4), 575-586.

Terlizzi F., Li C., Ratti C., Qiu W., Credi R., & Meng B. (2011). Detection of multiple sequence variants of Grapevine rupestris stem pitting-associated virus using primers targeting the polymerase domain and partial genome sequencing of a novel variant. *Annals of Applied Biology, 159*(3), 478-490.

Thapa V., Melcher U., Wiley G. B., Doust A., Palmer M. W., Roewe K., & Kamath N. (2012). Detection of members of the Secoviridae in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA. *Virus Research, 167*(1), 34-42.

Thapa V., McGlinn D. J., Melcher U., Palmer M. W., & Roossinck M. J. (2015). Determinants of taxonomic composition of plant viruses at the Nature Conservancy's Tallgrass Prairie Preserve, Oklahoma. *Virus Evolution, 1*(1), vev007.

Thomas J. E.; Wood T. A.; Gullino M. L. & Ortu G. (2017). Diagnostic Tools for Plant Biosecurity. Pages 209-226 in: Practical Tools for Plant and Food Biosecurity: Results from a European Network of Excellence. Gullino, M L, Stack, J P, Fletcher, J & Mumford, J D, eds. Springer International Publishing, Cham.

Triantafilou K., Vakakis E., Kar S., Richer E., Evans G. L., & Triantafilou M. (2012). Visualisation of direct interaction of MDA5 and the dsRNA replicative intermediate form of positive strand RNA viruses. *Journal of Cell Science, 125*(Pt 20), 4761-4769.

Turturo C., Saldarelli P., Yafeng D., Digiaro M., Minafra A., Savino V., & Martelli G. P. (2005). Genetic variability and population structure of Grapevine leafroll-associated virus 3 isolates. *Journal of General Virology, 86*, 217-224.

Tzanetakis I. E., Postman J. D., & Martin R. R. (2007). Identification, detection and transmission of a new vitivirus from Mentha. *Archives of Virology, 152*(11), 2027-2033.

Tzanetakis I. E., & Martin R. R. (2008). A new method for extraction of double-stranded RNA from plants. *Journal of Virological Methods, 149*(1), 167-170.

Tzanetakis I. E., Postman J. D., Samad A., & Martin R. R. (2010). Mint Viruses: Beauty, Stealth, and Disease. *Plant Disease, 94*(1), 4-12.

Ulubaş Serçe Ç., Altan B., Bolat V., Ayyaz M., Çifçi O., Önder S., & Maliogka V. I. (2017). First Report of Grapevine Roditis leaf discoloration-associated virus Infecting Grapevine (Vitis vinifera) in Turkey. *Plant Disease, 102*(1), 256.

Uricaru R., Rivals E., & Bréhélin L. (2006). Hidden Markov Models for the Detection of Motif Repeats in Protein Sequences. *IPG, Lyon, novembre-décembre*.

van der Vlugt R. A., van Raaij H., de Weerdt M., & Bergervoet J. H. (2015). Multiplex detection of plant pathogens through the luminex magplex bead system. *Plant Pathology: Techniques and Protocols*, 283-299.

Veerakone S., Tang J. Z., Ward L. I., Liefting L. W., Perez-Egusquiza Z., Lebas B. S. M., & Guy P. L. (2015). A review of the plant virus, viroid, liberibacter and phytoplasma records for New Zealand. *Australasian Plant Pathology, 44*(5), 463-514.

Verbeek M., Dullemans A., van Raaij H. G., Verhoeven J. J., & van der Vlugt R. A. (2014). Lettuce necrotic leaf curl virus, a new plant virus infecting lettuce and a proposed member of the genus Torradovirus. *Archives of Virology, 159*(4), 801-805.

Vignali D. A. A. (2000). Multiplexed particle-based flow cytometric assays. *Journal of Immunological Methods, 243*(1–2), 243-255.

Vigne E., Bergdoll M., Guyader S., & Fuchs M. (2004). Population structure and genetic variability within isolates of Grapevine fanleaf virus from a naturally infected vineyard in France: evidence for mixed infection and recombination. *Journal of General Virology, 85*(8), 2435-2445.

Voller A., Bidwell D., Huldt G., & Engvall E. (1974). A microplate method of enzyme-linked immunosorbent assay and its application to malaria. Research Support, U.S. Gov't, Non-P.H.S. *Bulletin of The World Health Organization, 51*(2), 209-211.

Vončina D., & Almeida R. (2017a). *Two novel virus-like sequences from grapevine.* Paper presented at the EPPO Workshop on the use of NGS technologies for plant diagnostics, Bari (Italy) 22-23 November 2017.

Vončina D., Al Rwahnih M., Rowhani A., Gouran, & Almeida R. (2017b). Viral Diversity in Autochthonous Croatian Grapevine Cultivars. *Plant Disease* 101(7), 1230-1235.

Wang J. B., Sharma A. M., Duffy S., & Almeida R. P. P. (2011). Genetic Diversity in the 3 ' Terminal 4.7-kb Region of Grapevine leafroll-associated virus 3. *Phytopathology, 101*(4), 445-450.

Watson J. D., & Crick F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature, 171*(4356), 737-738.

Weber F., Wagner V., Rasmussen S. B., Hartmann R., & Paludan S. R. (2006). Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *Journal of Virology, 80*(10), 5059-5064.

White E. J., Venter M., Hiten N. F., & Burger J. T. (2008). Modified cetyltrimethylammonium bromide method improves robustness and versatility: the benchmark for plant RNA extraction. *Biotechnology Journal: Healthcare Nutrition Technology, 3*(11), 1424-1428.

Wu Q., & Habili N. (2017). The recent importation of Grapevine Pinot gris virus into Australia. *Virus Genes, 53*(6), 935-938.

Yalow R. S., & Berson S. A. (1960). Immunoassay of endogenous plasma insulin in man. *The Journal of Clinical Investigation, 39*(7), 1157-1175.

Yoder J. (2014). 2014 NGS Field Guide: Resistance is Futile (mostly, at least for a while). *The Molecular ecologist*. http://www.molecularecologist.com/2014/03/2014-ngs-field-guide-resistance-is-futile-mostly-at-least-for-a-while/ [Accessed 19 March 2014]

Zhang Y., Singh K., Kaur R., & Qiu W. (2011). Association of a novel DNA virus with the grapevine vein-clearing and vine decline syndrome. *Phytopathology, 101*(9), 1081-1090.

Zhang Z., Qi S., Tang N., Zhang X., Chen S., Zhu P., & Wu Q. (2014). Discovery of Replicating Circular RNAs by RNA-Seq and Computational Algorithms. *PLoS Pathogens, 10*(12), e1004553.

Zhu M., Gong X., Hu Y., Ou W., & Wan Y. (2014). Streptavidin-biotin-based directional double Nanobody sandwich ELISA for clinical rapid and sensitive detection of influenza H5N1. *Journal of Translational Medicine, 12*(1), 1-10.

7 Appendices

7.1 Supplementary information for Chapter 2

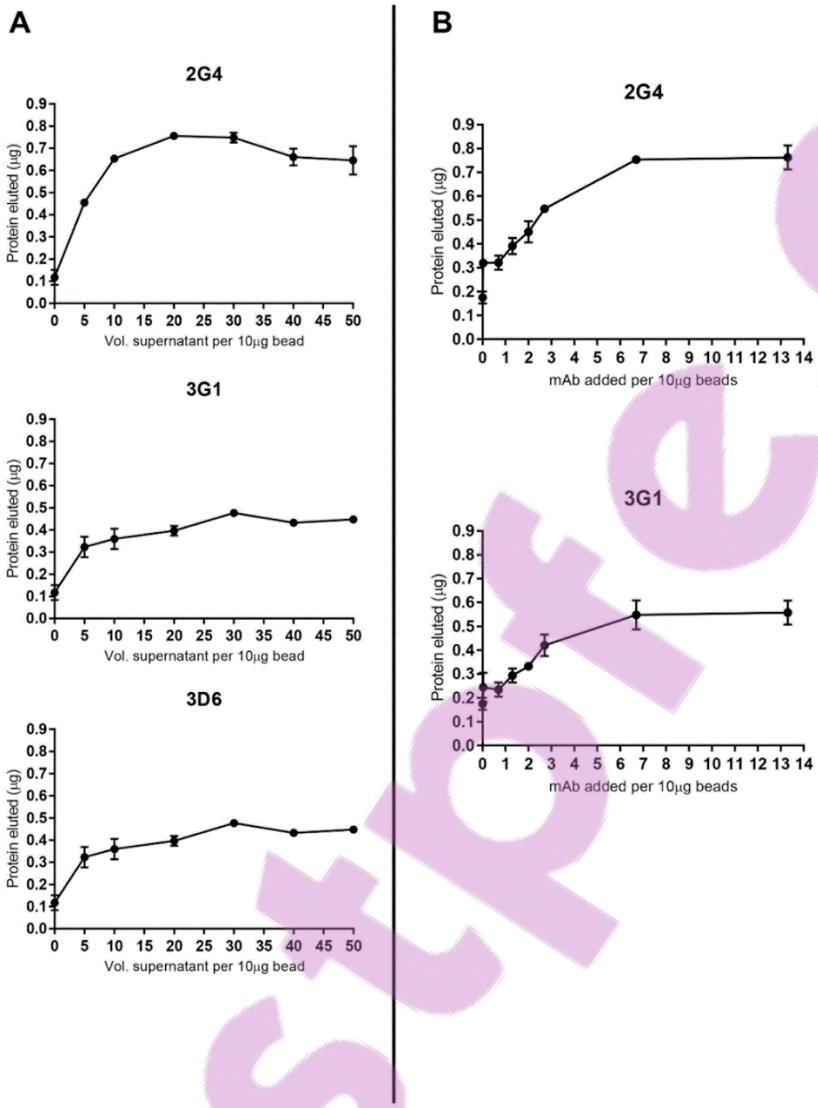


Figure 7.1: Bead saturation experiment: Dilutions of (**A.**) purified mAbs (2.5–100 µg) or (**B.**) hybridoma supernatant (25 – 500 µL) were in TBST. 500 µL of each dilution was added to 7.5–10 µL magnetic protein-L beads and incubated for 1 hour at room temperature with mixing. Beads were washed three times with TBST and once with TBS. Bound proteins were eluted with 20 µL 0.1M glycine (pH 2.0) and neutralised with 1.0M Tris (pH 8.5). Eluted protein was quantified with 3-6 repeated measurements using the NanoDrop ND-1000 with the protein A280 application.

7.2 Supplementary information for Chapter 4

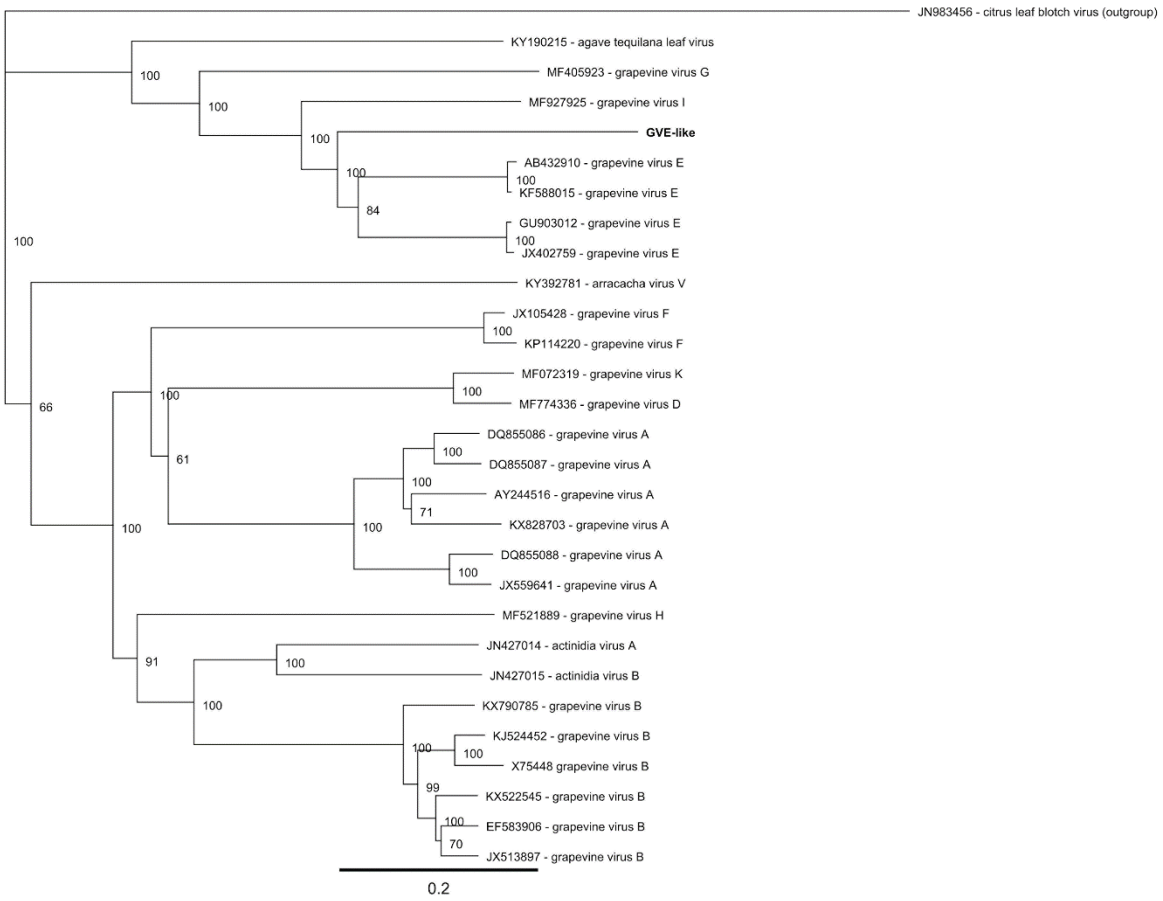


Figure 7.2: Phylogenetic analysis of the GVE-like vititivirus (replicase). Neighbor-Joining trees (1000 bootstrap replicates using the Jukes-Cantor distance model) of replicase protein of representative members of the genus Vitivirus. Protein alignments (translated from the accession number indicated) were performed with ClustalW (BLOSUM cost matrix with a gap opening cost set at 10, and a gap extension cost at 0.1). Consensus support is shown as a percentage on the branch. Citrus leaf blotch was used as outgroups for the phylogenetic analysis.

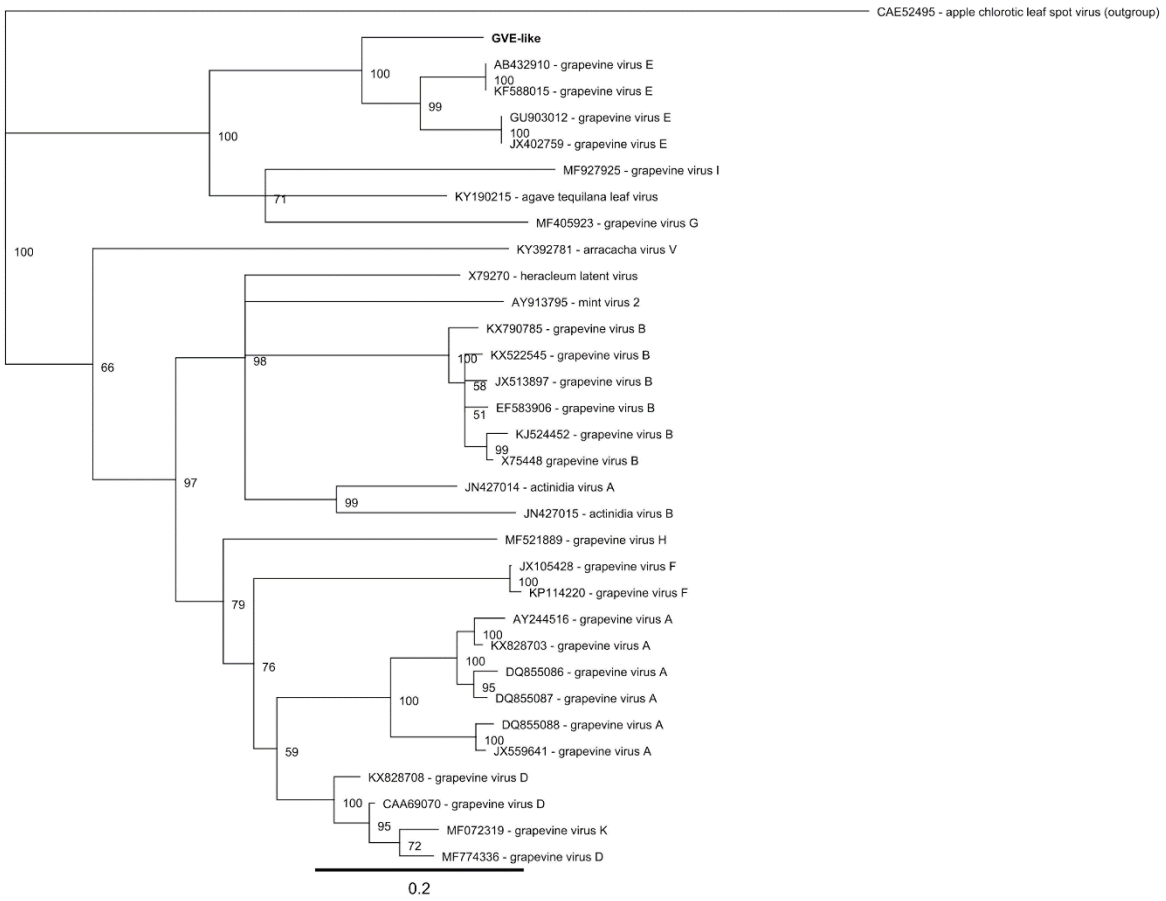


Figure 7.3: Phylogenetic analysis of the GVE-like vititivirus (coat protein).Neighbor-Joining trees (1000 bootstrap replicates using the Jukes-Cantor distance model) of coat protein of representative members of the genus Vitivirus. Protein alignments (translated from the accession number indicated) were performed with ClustalW (BLOSUM cost matrix with a gap opening cost set at 10, and a gap extension cost at 0.1). Consensus support is shown as a percentage on the branch. Apple chlorotic leaf spot virus was used as outgroups for the phylogenetic analysis.

Table 7.1 Virus status for each of the 225 grapevine sequenced. The GYSVd and HSVd column list the number of reads mapped to the viroids grapevine yellow speckle viroid-1 and Hop stunt viroid respectively normalised per million reads sequenced. The last two columns indicate the number of virus considered positive and the percentage of viral and viroid reads detected for that sample.

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-1	A-017	SB-1	Sauvignon blanc	GRSPaV	0	12562	1	9.2%
SB-1	A-018	SB-1	Sauvignon blanc		1502	61944	0	6.4%
SB-1	A-023	SB-1	Sauvignon blanc	GRSPaV	103	38375	1	13.8%
SB-1	A-028	SB-1	Sauvignon blanc	GSyV-1	2554	58688	1	6.6%
SB-1	A-029	SB-1	Sauvignon blanc		1361	65574	0	6.7%
SB-1	A-034	SB-1	Sauvignon blanc	GRGV; GRSPaV	1489	59071	2	15.4%
SB-1	A-039	SB-1	Sauvignon blanc	GLRaV-2; GFkV; GRSPaV	1377	28003	3	19.2%
SB-1	A-040	SB-1	Sauvignon blanc	GLRaV-2; GSyV-1; GFkV; GRSPaV	511	34369	4	21.4%
SB-1	A-050	SB-1	Sauvignon blanc	GLRaV-2; GRGV; GFkV; GRSPaV	807	15975	4	22.0%
SB-1	A-051	SB-1	Sauvignon blanc	GRGV	23	2654	1	0.6%
SB-1	A-079	SB-1	Sauvignon blanc		203	21246	0	2.4%
SB-1	A-080	SB-1	Sauvignon blanc		1005	33155	0	3.4%
SB-1	A-085	SB-1	Sauvignon blanc	GRSPaV	7	8950	1	8.2%
SB-1	A-090	SB-1	Sauvignon blanc		118	31575	0	3.4%
SB-1	A-091	SB-1	Sauvignon blanc		19	19079	0	2.2%
SB-1	A-108	SB-1	Sauvignon blanc		104	22052	0	2.5%
SB-1	A-109	SB-1	Sauvignon blanc		8	4197	0	0.6%
SB-1	A-114	SB-1	Sauvignon blanc	GRSPaV	11	297	1	1.6%
SB-1	A-119	SB-1	Sauvignon blanc	GRVFV	7	7483	1	1.4%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-1	A-120	SB-1	Sauvignon blanc		331	13728	0	1.7%
SB-1	A-130	SB-1	Sauvignon blanc		22790	100705	0	12.4%
SB-1	A-131	SB-1	Sauvignon blanc	GRVFV	280	17160	1	4.3%
SB-1	A-136	SB-1	Sauvignon blanc		434	25396	0	2.9%
SB-1	A-150	SB-1	Sauvignon blanc	GSyV-1	636	40594	1	5.4%
SB-1	A-151	SB-1	Sauvignon blanc		487	6261	0	0.8%
SB-1	A-156	SB-1	Sauvignon blanc	GSyV-1	1244	30132	1	3.9%
SB-1	A-171	SB-1	Sauvignon blanc	GRVFV; GRSPaV	378	15800	2	4.5%
SB-1	A-176	SB-1	Sauvignon blanc		1481	31712	0	3.4%
SB-1	A-177	SB-1	Sauvignon blanc	GRVFV; GRGV	888	13439	2	3.7%
SB-1	A-198	SB-1	Sauvignon blanc		25	5236	0	0.6%
SB-1	A-199	SB-1	Sauvignon blanc	GSyV-1	197	9422	1	2.0%
SB-1	A-204	SB-1	Sauvignon blanc		208	16486	0	1.7%
SB-1	A-211	SB-1	Sauvignon blanc		241	14486	0	1.5%
SB-1	A-212	SB-1	Sauvignon blanc	GSyV-1; GRSPaV	50	2856	2	1.4%
SB-1	A-227	SB-1	Sauvignon blanc		254	46020	0	4.9%
SB-1	A-228	SB-1	Sauvignon blanc	GSyV-1	286	55921	1	6.9%
SB-1	A-233	SB-1	Sauvignon blanc	GLRaV-3	227	21993	1	12.3%
SB-1	A-238	SB-1	Sauvignon blanc		11	4388	0	0.5%
SB-1	A-239	SB-1	Sauvignon blanc		671	25478	0	2.7%
SB-1	A-251	SB-1	Sauvignon blanc	GRVFV; GSyV-1; GRSPaV	580	36480	3	12.0%
SB-1	A-252	SB-1	Sauvignon blanc	GSyV-1; GRSPaV	730	28517	2	5.3%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-1	A-257	SB-1	Sauvignon blanc		438	31364	0	3.2%
SB-1	A-262	SB-1	Sauvignon blanc	GRVFV; GRSPaV	399	22540	2	7.2%
SB-1	A-263	SB-1	Sauvignon blanc	GRSPaV	161	28633	1	5.8%
SB-2	D-012	SB-2	Sauvignon blanc	GRSPaV	0	10969	1	62.3%
SB-2	D-013	SB-2	Sauvignon blanc	GRSPaV	18	1947	1	34.1%
SB-2	D-018	SB-2	Sauvignon blanc	GRSPaV	18	1028	1	33.6%
SB-2	D-029	SB-2	Sauvignon blanc	GRSPaV	124	10137	1	46.1%
SB-2	D-034	SB-2	Sauvignon blanc	GRSPaV	18	29593	1	30.4%
SB-2	D-039	SB-2	Sauvignon blanc	GRSPaV	0	4327	1	19.5%
SB-2	D-040	SB-2	Sauvignon blanc	GRSPaV	0	10884	1	35.2%
SB-2	D-046	SB-2	Sauvignon blanc	GRSPaV	0	138	1	74.8%
SB-2	D-063	SB-2	Sauvignon blanc	GRSPaV	25	878	1	40.1%
SB-2	D-064	SB-2	Sauvignon blanc	GRSPaV	79	2312	1	38.7%
SB-2	D-085	SB-2	Sauvignon blanc	GRSPaV	19	950	1	26.9%
SB-2	D-086	SB-2	Sauvignon blanc	GSyV-1; GRSPaV	10	381	2	40.2%
SB-2	D-091	SB-2	Sauvignon blanc	GRSPaV	0	290	1	27.2%
SB-2	D-096	SB-2	Sauvignon blanc	GRSPaV	14	228	1	15.7%
SB-2	D-097	SB-2	Sauvignon blanc	GRSPaV	8	359	1	23.7%
SB-2	D-109	SB-2	Sauvignon blanc	GRSPaV	95	1584	1	14.5%
SB-2	D-114	SB-2	Sauvignon blanc	GRSPaV	50	650	1	42.2%
SB-2	D-115	SB-2	Sauvignon blanc	GSyV-1; GRSPaV	0	1494	2	36.7%
SB-2	D-136	SB-2	Sauvignon blanc	GRSPaV	266	16428	1	26.5%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-2	D-137	SB-2	Sauvignon blanc	GRSPaV	18	3173	1	39.7%
SB-2	D-142	SB-2	Sauvignon blanc	GRSPaV	55	1137	1	38.1%
SB-2	D-152	SB-2	Sauvignon blanc	GRSPaV	0	15338	1	40.6%
SB-2	D-153	SB-2	Sauvignon blanc	GRSPaV	0	842	1	1.4%
SB-2	D-158	SB-2	Sauvignon blanc	GRSPaV	0	2995	1	58.2%
SB-2	D-163	SB-2	Sauvignon blanc	GRSPaV	0	3196	1	53.2%
SB-2	D-164	SB-2	Sauvignon blanc	GRSPaV	0	461	1	45.7%
SB-2	D-169	SB-2	Sauvignon blanc	GRSPaV	62	1786	1	40.0%
SB-2	D-170	SB-2	Sauvignon blanc	GRVFV; GFkV; GRSPaV	0	2260	3	39.8%
SB-2	D-187	SB-2	Sauvignon blanc	GRVFV; GFkV; GRSPaV	8	194	3	11.8%
SB-2	D-188	SB-2	Sauvignon blanc	GRSPaV	10	989	1	21.2%
SB-2	D-202	SB-2	Sauvignon blanc	GRSPaV	29	537	1	17.6%
SB-2	D-207	SB-2	Sauvignon blanc	GRSPaV	6	1179	1	25.4%
SB-2	D-208	SB-2	Sauvignon blanc	GLRaV-2; GSyV-1; GRGV; GRSPaV	46	457	4	27.6%
SB-2	D-227	SB-2	Sauvignon blanc	GRSPaV	264	5352	1	21.5%
SB-2	D-228	SB-2	Sauvignon blanc	GRVFV; GFkV; GRSPaV	0	350	3	53.0%
SB-2	D-240	SB-2	Sauvignon blanc	GRSPaV	206	5418	1	19.8%
SB-2	D-241	SB-2	Sauvignon blanc	GRSPaV	107	5719	1	20.8%
SB-2	D-246	SB-2	Sauvignon blanc	GRSPaV	24	1351	1	29.5%
SB-2	D-251	SB-2	Sauvignon blanc	GRSPaV	0	1049	1	32.1%
SB-2	D-256	SB-2	Sauvignon blanc	GRSPaV	25	376	1	33.3%
SB-2	D-257	SB-2	Sauvignon blanc	GRGV; GRSPaV	270	2675	2	32.4%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-2	D-269	SB-2	Sauvignon blanc	GRSPaV	604	13765	1	32.5%
SB-2	D-270	SB-2	Sauvignon blanc	GRSPaV	15	3838	1	15.3%
SB-3	U-012	SB-3	Sauvignon blanc		89	27415	0	2.8%
SB-3	U-018	SB-3	Sauvignon blanc		0	71684	0	7.2%
SB-3	U-028	SB-3	Sauvignon blanc		83	32094	0	3.3%
SB-3	U-029	SB-3	Sauvignon blanc		0	74081	0	7.4%
SB-3	U-034	SB-3	Sauvignon blanc	GSyV-1	7	28797	1	4.0%
SB-3	U-040	SB-3	Sauvignon blanc	GSyV-1	125	96372	1	10.4%
SB-3	U-047	SB-3	Sauvignon blanc	GSyV-1	28	65507	1	8.6%
SB-3	U-048	SB-3	Sauvignon blanc		0	65730	0	6.6%
SB-3	U-065	SB-3	Sauvignon blanc	GSyV-1; GRGV	0	7385	2	6.3%
SB-3	U-083	SB-3	Sauvignon blanc	GSyV-1	0	32881	1	11.2%
SB-3	U-096	SB-3	Sauvignon blanc	GSyV-1	0	10359	1	2.5%
SB-3	U-097	SB-3	Sauvignon blanc	GLRaV-3; GSyV-1	169	9397	2	16.4%
SB-3	U-116	SB-3	Sauvignon blanc		102	6161	0	0.6%
SB-3	U-117	SB-3	Sauvignon blanc		13	10925	0	1.1%
SB-3	U-122	SB-3	Sauvignon blanc	GSyV-1	11	13959	1	4.9%
SB-3	U-127	SB-3	Sauvignon blanc		14	9955	0	1.0%
SB-3	U-132	SB-3	Sauvignon blanc	GRGV	41	10600	1	3.3%
SB-3	U-133	SB-3	Sauvignon blanc		17	13741	0	1.4%
SB-3	U-145	SB-3	Sauvignon blanc	GSyV-1	58	11473	1	2.8%
SB-3	U-146	SB-3	Sauvignon blanc	GSyV-1	89	12524	1	6.7%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-3	U-163	SB-3	Sauvignon blanc	GSyV-1	53	10626	1	2.5%
SB-3	U-164	SB-3	Sauvignon blanc		11	10688	0	1.4%
SB-3	U-169	SB-3	Sauvignon blanc	GLRaV-3; GSyV-1; GVE-like; GVG	8	12178	4	10.6%
SB-3	U-183	SB-3	Sauvignon blanc	GSyV-1	5	16245	1	2.7%
SB-3	U-184	SB-3	Sauvignon blanc		8	7371	0	0.8%
SB-3	U-198	SB-3	Sauvignon blanc	GSyV-1	110	17461	1	3.8%
SB-3	U-199	SB-3	Sauvignon blanc	GSyV-1; GRGV	0	5609	2	7.6%
SB-3	U-204	SB-3	Sauvignon blanc		25	11768	0	1.2%
SB-3	U-216	SB-3	Sauvignon blanc		0	8452	0	0.9%
SB-3	U-217	SB-3	Sauvignon blanc	GSyV-1	84	16009	1	2.4%
SB-3	U-222	SB-3	Sauvignon blanc	GSyV-1	2	10339	1	3.9%
SB-3	U-223	SB-3	Sauvignon blanc	GSyV-1	25	17195	1	5.9%
SB-3	U-238	SB-3	Sauvignon blanc	GSyV-1	24	13420	1	2.4%
SB-3	U-239	SB-3	Sauvignon blanc	GSyV-1	142	10505	1	2.6%
SB-3	U-244	SB-3	Sauvignon blanc	GSyV-1	29	16838	1	4.1%
SB-3	U-249	SB-3	Sauvignon blanc	GSyV-1	21	37141	1	6.0%
SB-3	U-250	SB-3	Sauvignon blanc		104	30671	0	3.1%
SB-3	U-257	SB-3	Sauvignon blanc		204	29008	0	2.9%
SB-3	U-262	SB-3	Sauvignon blanc	GSyV-1	30	20715	1	3.5%
SB-3	U-263	SB-3	Sauvignon blanc	GSyV-1	28	15107	1	2.2%
SB-4	W-017	SB-4	Sauvignon blanc	GSyV-1	147	69454	1	9.6%
SB-4	W-018	SB-4	Sauvignon blanc	GSyV-1	13	54264	1	10.5%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-4	W-023	SB-4	Sauvignon blanc	GSyV-1	1473	182916	1	19.3%
SB-4	W-024	SB-4	Sauvignon blanc	GSyV-1	2120	235141	1	28.7%
SB-4	W-029	SB-4	Sauvignon blanc	GSyV-1	28	70598	1	12.5%
SB-4	W-047	SB-4	Sauvignon blanc	GSyV-1	0	33418	1	19.6%
SB-4	W-052	SB-4	Sauvignon blanc	GSyV-1	44	85259	1	12.6%
SB-4	W-053	SB-4	Sauvignon blanc	GSyV-1	216	105543	1	14.2%
SB-4	W-066	SB-4	Sauvignon blanc	GSyV-1	3666	174862	1	21.4%
SB-4	W-078	SB-4	Sauvignon blanc	GSyV-1	3730	106022	1	22.5%
SB-4	W-083	SB-4	Sauvignon blanc	GSyV-1	3625	285237	1	31.7%
SB-4	W-084	SB-4	Sauvignon blanc	GSyV-1	164	200195	1	23.4%
SB-4	W-111	SB-4	Sauvignon blanc	GSyV-1	11	12533	1	3.1%
SB-4	W-121	SB-4	Sauvignon blanc	GSyV-1	174	11725	1	1.7%
SB-4	W-122	SB-4	Sauvignon blanc	GSyV-1	417	37281	1	5.8%
SB-4	W-134	SB-4	Sauvignon blanc		230	112334	0	11.4%
SB-4	W-135	SB-4	Sauvignon blanc		35	31100	0	3.4%
SB-4	W-145	SB-4	Sauvignon blanc	GSyV-1	304	8963	1	1.5%
SB-4	W-150	SB-4	Sauvignon blanc	GSyV-1	225	21741	1	4.9%
SB-4	W-151	SB-4	Sauvignon blanc	GSyV-1	363	43190	1	5.3%
SB-4	W-158	SB-4	Sauvignon blanc	GSyV-1	72	15314	1	5.8%
SB-4	W-159	SB-4	Sauvignon blanc	GSyV-1	39	9624	1	2.2%
SB-4	W-164	SB-4	Sauvignon blanc		211	15624	0	1.6%
SB-4	W-185	SB-4	Sauvignon blanc	GSyV-1	22	8148	1	1.4%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
SB-4	W-186	SB-4	Sauvignon blanc	GSyV-1	78	17922	1	2.8%
SB-4	W-191	SB-4	Sauvignon blanc	GSyV-1	318	21826	1	4.0%
SB-4	W-192	SB-4	Sauvignon blanc	GSyV-1	34	7497	1	1.5%
SB-4	W-207	SB-4	Sauvignon blanc	GSyV-1	54	21126	1	2.8%
SB-4	W-208	SB-4	Sauvignon blanc	GSyV-1	424	69621	1	9.7%
SB-4	W-213	SB-4	Sauvignon blanc	GSyV-1	4	8603	1	2.3%
SB-4	W-218	SB-4	Sauvignon blanc		0	4791	0	0.7%
SB-4	W-219	SB-4	Sauvignon blanc	GSyV-1	374	23740	1	4.6%
SB-4	W-238	SB-4	Sauvignon blanc	GSyV-1	281	29074	1	5.3%
SB-4	W-239	SB-4	Sauvignon blanc	GSyV-1	137	29707	1	4.4%
SB-4	W-251	SB-4	Sauvignon blanc		16	8285	0	0.9%
SB-4	W-252	SB-4	Sauvignon blanc	GSyV-1	50	17856	1	2.9%
SB-4	W-257	SB-4	Sauvignon blanc	GSyV-1	10	6049	1	2.4%
SB-4	W-269	SB-4	Sauvignon blanc	GSyV-1	10	7705	1	5.2%
SB-4	W-270	SB-4	Sauvignon blanc		43	18608	0	2.3%
DR	AB044	Hawke's bay	Pinot gris	GLRaV-3; GRSPaV; GVA; GVB	0	236	4	60.6%
DR	AB045	Gimblett gravels	Syrah	GRSPaV	16	62913	1	45.6%
DR	AB047	Gimblett gravels	Syrah	GRSPaV	0	22027	1	25.6%
DR	AB048	Gimblett gravels	Malbec	GRGV; GRSPaV	452	2858	2	4.1%
DR	AB049	Gimblett gravels	Malbec		0	68	0	0.3%
DR	AB051	Gimblett gravels	Malbec	GRSPaV	1	332	1	1.1%
DR	AB052	Gimblett gravels	Malbec	GRSPaV	3	110	1	17.6%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
DR	AB053	Gimblett gravels	Cabernet sauvignon	GRSPaV	0	14	1	36.8%
DR	AB054	Gimblett gravels	Gewürztraminer	GLRaV-3	3	422	1	18.1%
DR	AB055	Gimblett gravels	Gewürztraminer	GLRaV-3; GVG	101	5035	2	42.8%
DR	AB056	Gimblett gravels	Merlot		2	3841	0	1.0%
DR	AB057	Gimblett gravels	Chardonnay	GRSPaV	0	693	1	33.6%
DR	AB058	Gimblett gravels	Grenache	GRSPaV	0	1794	1	4.5%
DR	AB059	Bridge Pa Triangle	Merlot	GRSPaV	0	6325	1	20.4%
DR	AB060	Bridge Pa Triangle	Merlot	GRSPaV	0	503	1	20.4%
DR	AB062	Gimblett gravels	T3	GRSPaV	31	4616	1	10.3%
DR	AB526	Hawke's bay	Syrah	GLRaV-3; GRSPaV	0	7869	2	20.1%
DR	AB527	Hawke's bay	Syrah	GRSPaV	0	3892	1	11.4%
DR	AB528	Hawke's bay	Syrah	GLRaV-3; GRSPaV; GVA	0	8910	3	23.3%
DR	AB529	Hawke's bay	Syrah	GRSPaV	0	6324	1	20.7%
DR	AB530	Hawke's bay	Syrah	GLRaV-3; GRSPaV	0	147	2	58.0%
DR	AB531	Hawke's bay	Syrah	GRSPaV	0	659	1	7.6%
DR	AB532	Hawke's bay	Syrah	GRSPaV	1	978	1	4.0%
DR	AB533	Hawke's bay	Syrah	GRSPaV	2	2945	1	11.6%
GC-HH	AB552_VID891	Lincoln Uni NZW collection	Rupestris St George	GRVFV; GAMaV	0	411	2	2.2%
GC-HH	AB553_VID894	Lincoln Uni NZW collection	Rupestris St George	GLRaV-3; GRSPaV	0	1123	2	6.8%
GC-HH	AB554_VID1027	Lincoln Uni NZW collection	Sauvignon Blanc	GLRaV-4; GRSPaV	0	658	2	11.1%
GC-HH	AB555_VID1030	Lincoln Uni NZW collection	Sauvignon Blanc 316	GLRaV-2; GRSPaV	15	3920	2	16.0%
GC-HH	AB556_VID1032	Lincoln Uni NZW collection	Sauvignon Blanc		6	3181	0	0.5%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
GC-HH	AB557_VID1034	Lincoln Uni NZW collection	Sauvignon Blanc	GRSPaV	2	265	1	17.7%
GC-HH	AB558_VID1062	Lincoln Uni NZW collection	Gewurtztraminer	GRSPaV	0	3759	1	15.6%
GC-HH	AB559_VID1064	Lincoln Uni NZW collection	Gewurtztraminer	GLRaV-3; GSyV-1; GRGV; GRSPaV	49	72994	4	20.4%
GC-HH	AB560_VID1066	Lincoln Uni NZW collection	Gewurtztraminer	GLRaV-3; GSyV-1; GRGV; GRSPaV	42	63630	4	29.7%
GC-HH	AB561_VID1068	Lincoln Uni NZW collection	Gewurtztraminer	GRGV; GRSPaV	983	24436	2	22.0%
GC-HH	AB562_VID1070	Lincoln Uni NZW collection	Gewurtztraminer	GLRaV-3; GRSPaV	0	19994	2	31.3%
GC-HH	AB563_VID1072	Lincoln Uni NZW collection	Gewurtztraminer	GRSPaV	4	1657	1	9.7%
GC-HH	AB564_VID1074	Lincoln Uni NZW collection	Gewurtztraminer	GRGV; GRSPaV	0	1044	2	12.2%
GC-HH	AB565_VID1076	Lincoln Uni NZW collection	Chenin Blanc		0	93	0	0.5%
GC-HH	AB567_VID1080	Lincoln Uni NZW collection	Chenin Blanc		0	526	0	0.4%
GC-HH	AB568_VID1082	Lincoln Uni NZW collection	Chenin Blanc		5	298	0	0.5%
GG-LH	AB384_VID836	Lincoln Uni NZW collection	Shiraz	GLRaV-3; GLRaV-2; GVA; GVG; GVI	0	177	5	55.0%
GG-LH	AB534_VID236	Lincoln Uni NZW collection	Chelois	GLRaV-3; GLRaV-2; GRVFV; GRSPaV; GVA; GVG; GVI	0	27	7	78.4%
GG-LH	AB535_VID253	Lincoln Uni NZW collection	Pinard	GLRaV-3; GRSPaV; GVA; GVG; GVI	0	60	5	85.7%
GG-LH	AB536_VID262	Lincoln Uni NZW collection	Seibel 10096	GLRaV-3; GLRaV-4; GLRaV-2; GRSPaV; GVA; GVG	0	42	6	91.5%
GG-LH	AB537_VID280	Lincoln Uni NZW collection	Seibel 7052	GLRaV-2; GRSPaV; GGVA	0	222	3	17.7%
GG-LH	AB538_VID378	Lincoln Uni NZW collection	Sylvaner	GLRaV-3; GRVFV; GVA; GVG; GVI	3	1447	5	77.3%
GG-LH	AB539_VID417	Lincoln Uni NZW collection	Mt Albert 10		0	97	0	0.9%
GG-LH	AB540_VID493	Lincoln Uni NZW collection	Chardonnay	GLRaV-3; GVA; GVG; GVI	1	1915	4	59.5%
GG-LH	AB541_VID499	Lincoln Uni NZW collection	Chardonnay	GLRaV-3; GRVFV; GRSPaV; GVA; GVD; GVG; GVI	0	1225	7	75.1%

Library	Plant ID	Vineyard name or location	Cultivar	Virus detected	GYSVd	HSVd	# virus	ratio
GG-LH	AB542_VID561	Lincoln Uni NZW collection	Chardonnay 8021	GLRaV-3; GLRaV-2; GRVFV; GRGV; GRSPaV; GVG	10	1255	6	68.7%
GG-LH	AB543_VID567	Lincoln Uni NZW collection	Chardonnay	GLRaV-1; GLRaV-3; GRVFV; GVG	0	96	4	51.3%
GG-LH	AB544_VID568	Lincoln Uni NZW collection	Chardonnay	GLRaV-1; GLRaV-3; GRVFV; GSyV-1; GRSPaV; GVA	0	314	6	18.6%
GG-LH	AB545_VID576	Lincoln Uni NZW collection	Chardonnay	GRSPaV	0	78	1	2.2%
GG-LH	AB546_VID580	Lincoln Uni NZW collection	Chenin Blanc	GLRaV-3; GRVFV; GVA; GVG; GVI	0	200	5	86.5%
GG-LH	AB547_VID709	Lincoln Uni NZW collection	Sauvignon Blanc 316	GLRaV-2; GRVFV; GRGV; GRSPaV	0	251	4	37.8%
GG-LH	AB548_VID710	Lincoln Uni NZW collection	Sauvignon Blanc 316	GLRaV-2; GRVFV; GRGV; GRSPaV	0	320	4	39.6%
GG-LH	AB549_VID712	Lincoln Uni NZW collection	Sauvignon Blanc 317	GLRaV-3; GLRaV-2; GRSPaV; GVG	0	28	4	71.6%
GG-LH	AB550_VID769	Lincoln Uni NZW collection	Dolcetto	GLRaV-3; GLRaV-2; GRSPaV; GVA; GVG; GVI	0	46	6	83.4%
GC-LH	AB551_VID835	Lincoln Uni NZW collection	Shiraz	GLRaV-3; GLRaV-2; GVA; GVG; GVI	0	14	5	38.2%

Table 7.2: Read mapped per million (RMPM) for each sample for each virus from the GC-LH sequenced by dsRNA and sRNA HTS. The reads were only counted once and the RMPM value indicated is the sum of all the reads mapped when more than one reference sequences was available for a virus.

	GLRaV-1		GLRaV-3		GLRaV-4		GLRaV-2		GRVFV		GSyV-1		GAMaV		GRGV		GFkV		GRSPaV	
	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA
AB534	90	0	1079	443872	2	171	20735	288275	497	1545	11	7	2	4	2	84	3	0	15186	22232
AB535	79	20	4388	819038	4	89	160	479	1144	185	25	6	7	2	127	54	3858	370	7962	7634
AB536	75	15	2642	633259	229	107098	33292	148020	987	145	61	13	28	3	57	78	6190	57	3919	8704
AB537	77	8	102	3563	2	130	31388	96223	7	69	2	2	370	554	7	85	3	0	3929	76183
AB538	63	7	15144	735333	6	111	161	464	3307	3833	3	7	1	0	2	86	2	3	23	703
AB539	50	11	92	5189	3	268	149	769	6	121	0	35	0	15	1	234	2	7	22	1370
AB540	89	15	38858	552109	8	194	166	775	4	117	1	7	0	1	1	125	2	6	29	1176
AB541 (base)	93	0	81676	-	15	-	197	-	305	-	4	-	2	-	382	-	3	-	21410	-
AB541 (middle)	124	10	101219	691020	13	216	205	706	595	1394	4	1	1	3	340	420	4	4	24108	44357
AB541 (top)	125	0	9481	-	2	-	209	-	4498	-	29	-	5	-	1453	-	9	-	1369	-
AB542	78	12	55493	584892	1	86	185330	63748	158	2863	1	5	2	2	175	1537	4	7	11419	29440
AB543	249140	16855	26637	491732	15	40	176	184	230	1152	108	432	1	2	211	95	2	2	27	331
AB544	276961	15162	16189	107766	1	110	180	378	290	13899	169	4515	0	3	3	72	4	4	9596	36458
AB545	75	6	87	1077	2	58	227	174	3	43	1	3	0	0	1	17	1	0	5153	20639
AB547	249	0	93	3445	2	113	144850	196853	828	17985	4	4	1	4	620	2564	2	0	12407	156352
AB548	88	0	87	2457	2	133	93158	262905	1355	11175	4	5	1	0	734	1898	2	2	12551	116833
AB550	102	5	57207	713347	3	8	211632	48755	5	13	0	5	0	0	2	24	2	0	2419	27033
AB551	89	2	28226	341901	4	87	72978	28923	5	40	1129	3	1	1	2	44	2	0	54	432

	GVA		GVB		GVD		GVE-like		GVG		GVI		GGVA		GYSVd		HSVd	
	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA	sRNA	dsRNA
AB534	128	10792	4	0	4	0	0	0	762	8292	194	8294	2	0	2	0	1555	27
AB535	3857	10316	24	0	3	2	1	2	608	3822	690	15071	1	0	495	0	2515	60
AB536	424	13100	8	0	1	13	0	0	652	4593	4	83	1	0	3	0	1410	42
AB537	19	141	14	0	8	4	0	0	8	30	6	81	5017	2	4	0	2243	222
AB538	5505	5806	9	0	7	0	0	0	1072	3789	1739	21212	2	0	2370	3	2965	1447
AB539	15	247	9	0	0	11	1	0	9	108	4	140	1	0	1528	0	7	97
AB540	1034	20468	4	0	1	1	1	0	1587	6312	1028	11445	2	0	3	1	2894	1915
AB541 (base)	7843	-	19	-	500	-	1	-	3700	-	2843	-	2	-	360	-	1556	-
AB541 (middle)	16119	2165	58	0	1474	849	0	0	4077	2166	5782	6500	1	0	587	0	1867	1225
AB541 (top)	1214	-	7	-	21	-	1	-	817	-	923	-	1	-	1084	-	2818	-
AB542	18	161	5	0	20	3	1	0	11742	2814	5	97	3	0	1232	10	856	1255
AB543	25	83	3	0	1	4	0	0	13441	2049	4	45	1	0	345	0	1277	96
AB544	28340	7195	58	0	13	1	0	0	11	55	6	53	1	0	2243	0	1393	314
AB545	19	35	4	0	0	6	0	0	8	12	4	46	1	0	3161	0	1774	78
AB547	23	102	3	0	0	8	0	0	16	61	4	73	1	0	351	0	1665	251
AB548	22	109	2	0	1	5	0	0	9	62	6	45	1	0	257	0	1376	320
AB550	4560	5749	11	0	84	2	8	0	5513	2756	8194	35951	1	0	3	0	1025	46
AB551	1811	1268	5	0	24	0	0	0	2127	3338	3124	5568	1	0	179	0	1476	14

Table 7.3: Symptoms recorded and virus detected from the DR samples

Plant ID	cultivar	collection date	Symptoms recorded	virus detected
AB044	Pinot gris	Mar-15	symptoms not recorded	GRSPaV; GLRaV-3 (I, VI); GVA; GVB
AB045	Syrah	Mar-15	Early reddening symptoms (not GLRaV-3 symptoms)	GRSPaV
AB047	Syrah	Mar-15	Early reddening symptoms (not GLRaV-3 symptoms)	GRSPaV; GRVFV
AB048	Malbec	Mar-15	Early reddening symptoms (dark red-not GLRaV-3 symptoms)	GRSPaV
AB049	Malbec	Mar-15	Early reddening symptoms (not GLRaV-3 symptoms)	GRSPaV
AB051	Malbec	Mar-15	no symptoms	GRSPaV; GRGV
AB052	Malbec	Mar-15	young leave crinkle similar to herbicide damage	GRSPaV
AB053	Cabernet sauvignon	Mar-15	Strong GLRaV-3 symptoms (small leaves)	GRSPaV
AB054	Gewürztraminer	Mar-15	pale yellow leaves (small leaves)	GLRaV-3 (I, VI)
AB055	Gewürztraminer	Mar-15	pale yellow leaves (small leaves)	GLRaV-3 (VI); GVG
AB056	Merlot	Mar-15	basal leave reddening (segmentation - not GLRaV-3 symptoms)	no virus detected
AB057	Chardonnay	Mar-15	Cane wilting (trunk disease)	GRSPaV

Plant ID	cultivar	collection date	Symptoms recorded	virus detected
AB058	Grenache	Mar-15	mild chlorotic spekle	GRSPaV
AB059	Merlot / 3306	Mar-15	light reddening basal leave (not GLRaV-3)	GRSPaV
AB060	Merlot / 3306	Mar-15	redening leave border (not GLRaV-3 symptoms)	GRSPaV
AB062	T3	Mar-15	fasciation of the cane	GRSPaV
AB526	Syrah	Apr-17	leaf reddening (patchy - GLRaV-3)	GRSPaV; GLRaV-3 (NZ2)
AB527	Syrah	Apr-17	leave reddening (segmentation - not GLRaV-3 symptoms)	GRSPaV
AB528	Syrah	Apr-17	leaf reddening (GLRaV-3)	GRSPaV; GLRaV-3 (NZ2); GVA
AB529	Syrah	Apr-17	leaf marging reddening (segmentation - not GLRaV-3 symptoms)	GRSPaV; GSyV-1
AB530	Syrah	Apr-17	leaf reddening (patchy - GLRaV-3)	GRSPaV; GLRaV-3 (I)
AB531	Syrah	Apr-17	no symptoms	GRSPaV
AB532	Syrah	Apr-17	no symptoms	GRSPaV
AB533	Syrah	Apr-17	no symptoms	GRSPaV

Figure 7.4: Bioinformatics programming SB samples

README

February 22, 2018

1 Map Reads to Assembled Viral References
The Viral reads filtered from previous step will now be mapped to the contigs assembled with Trinity.

metadata

species Vitis vinifera
scientist Arnaud Blouin bioinformatician Ben Warren location /workspace, /output

1.1 01 Map to References
Use reference generated from other analysis.
In [3]: OUTPUT=01.map_to_references mkdir -p $OUTPUT

cp ../../../2017_run_3/05.map_to_assembled_refs/01.map_to_A_B_libs_reference/bt2*
$OUTPUT/

In [2]: cp ../../../2017_run_3/05.map_to_assembled_refs/Reference_sequence_feb_2018b.fa ./

1.1.1 Map Reads
In [4]: INPUT=../01.preprocessing/04.2.filter_hard_trimmed_per_sample/*_R1.fq.gz OUTPUT=01.map_to_references
INDEX=01.map_to_references/bt2idx module load bowtie2/2.2.9
module load samtools/1.2

for FILE1 in $INPUT do
FILE2=${FILE1//R1/R2}
NAME=`basename $FILE1 .fq.gz | perl -pe 's/^(\d+).*\.(.*)_R1/\1_\2/'`
echo -n "bowtie2 -p 8 --sensitive --end-to-end -x $INDEX -1 $FILE1 -2 $FILE2 " echo " | samtools view -Shu -F
4 - | samtools sort - ${OUTPUT}/$NAME"
echo
done | asub -n 8 -j ${OUTPUT}/map_reads

Job <85570> is submitted to default queue <normal>.

1.1.2 Stats
In [5]: INPUT=01.map_to_references/*.bam module load samtools/1.2
for BAM in $INPUT do
echo "samtools flagstat $BAM > ${BAM/.bam/.stats}" done | asub -j ${OUTPUT}/bam_stats

Job <85571> is submitted to default queue <normal>.

1.2 02 Counts
Counts from the mapping.
We want to count all primary alignments, regardless of pairing, so we want these flags to be
FALSE.

4 0x4 segment unmapped; we only want mapped reads

256 0x100 secondary alignment; no secondary alignments
2048 0x800 supplementary alignment; no supplimentary alignments
---
2308 0x904 combined flags

So we need to use the following samtools negated flags: -F 2308

1.2.1 Raw mapping counts, including zero counts for missing data.
In [6]: INPUT=01.map_to_references/*.bam

OUTPUT=02.counts mkdir -p $OUTPUT

SAM_FLAGS="-F 2308"
REF=./Reference_sequence_feb_2018b.fa module load samtools/1.2
# Create reference ID list
grep -Po '(?<=^>)\S+' $REF | sort > $OUTPUT/all_reference_ids.txt

# Count the reads as per the SAM flags
# This method adds zeros where there are no alignments to a reference
for BAM in $INPUT do
NAME=`basename $BAM .bam`

echo "samtools view $SAM_FLAGS $BAM | \ awk '{print \$3}' | \
sort | uniq -c | awk '{print \$2\"\t\"\$1}' | \ join -t $'\t' -a 1 -j 1 -e '0' -o '1.1,2.2' \
${OUTPUT}/all_reference_ids.txt - > ${OUTPUT}/${NAME}.counts" done | asub -j ${OUTPUT}/count_bams

Job <85572> is submitted to default queue <normal>.

1.2.2 Collate Counts
In [7]: INPUT=02.counts/*.counts OUTPUT=02.counts/counts_summary.txt
READ_FILES=../01.preprocessing/04.2.filter_hard_trimmed_per_sample

Echo header
echo -ne "sample\ttotal_reads" > $OUTPUT
grep -Po '^>\S+' ./Reference_sequence_feb_2018b.fa | tr -d '>' | sort | awk
'{printf("\t%s", $1)}' >> $OUTPUT echo >> $OUTPUT

for FILE in $INPUT do
echo -ne "`basename $FILE .counts`\t" SAMPLE="`basename $FILE .counts | cut -d_ -f1`" BARCODE="`basename
$FILE .counts | cut -d_ -f2-`"
READ_FILE=`ls ${READ_FILES}/${SAMPLE}*${BARCODE}*R1.fq.gz`
NUM_READS=`zcat $READ_FILE | awk 'NR % 4 == 1' | wc -l`
echo -ne "$((NUM_READS * 2))" # Becuase we count each pair individually
awk '{printf("\t%d", $2)}' $FILE echo
done >> $OUTPUT

#head $OUTPUT

1.3 Outputs
In [10]: # Create output dir
OUTPUT_STORAGE=/output/genomic/viral/Vitis/vinifera/vineyard_ecosystem/combine_with_seco
nd_run/04.map_to_assembled_reads
mkdir -p $OUTPUT_STORAGE

Symlink to output if [! -L ./output] then

```
ln -s $OUTPUT_STORAGE output
fi

PDF_TEMPLATE=~/.jupyter/nbconvert/templates/better-article.tplx

# Render notebook
module load pfr-python3
jupyter nbconvert --to markdown README.ipynb jupyter nbconvert --to html README.ipynb
jupyter nbconvert --to pdf --template $PDF_TEMPLATE README.ipynb module unload pfr-python3

# Copy and notebook renderings to output (overwrite read-only permissions)
cp  -f  README.*  $OUTPUT_STORAGE/

[NbConvertApp] Converting notebook README.ipynb to markdown [NbConvertApp] Writing 30650 bytes to
README.md [NbConvertApp] Converting notebook README.ipynb to html [NbConvertApp] Writing 289838 bytes
to README.html [NbConvertApp] Converting notebook README.ipynb to pdf [NbConvertApp] Writing 53279
bytes to notebook.tex [NbConvertApp] Building PDF
[NbConvertApp] Running xelatex 3 times: ['xelatex', 'notebook.tex'] [NbConvertApp] Running bibtex 1 time:
['bibtex',  'notebook']
[NbConvertApp] WARNING | bibtex had problems, most likely because there were no citations
[NbConvertApp] PDF successfully created [NbConvertApp] Writing 56623 bytes to README.pdf

In [9]: # Copy Results with rsync, removing source files
RESULTS='
01.map_to_references 02.counts
'

# Enable write permissions
chmod +w -R ./output/*

# Sync outputs, use '-n' for dry-run testing
nice -3 rsync -hlrmv --relative --remove-source-files --stats \
--exclude  '*.out'  \
--exclude  '*.err'  \
--exclude  '*.bt2'  \
$RESULTS ./output/

# Make output contents read only
chmod -w -R ./output/*

building  file list ... done 01.map_to_references/ 01.map_to_references/1083_D-012_O01.bam
01.map_to_references/1083_D-012_O01.stats 01.map_to_references/1083_D-013_BC2014-2.bam
01.map_to_references/1083_D-013_BC2014-2.stats 01.map_to_references/1083_D-018_BC2014-3.bam
01.map_to_references/1083_D-018_BC2014-3.stats 01.map_to_references/1083_D-028_B01.bam
01.map_to_references/1083_D-028_B01.stats 01.map_to_references/1083_D-029_D01.bam
01.map_to_references/1083_D-029_D01.stats 01.map_to_references/1083_D-034_E01.bam
01.map_to_references/1083_D-034_E01.stats 01.map_to_references/1083_D-039_C01.bam
01.map_to_references/1083_D-039_C01.stats 01.map_to_references/1083_D-040_A01.bam
01.map_to_references/1083_D-040_A01.stats 01.map_to_references/1083_D-045_P01.bam
01.map_to_references/1083_D-045_P01.stats 01.map_to_references/1083_D-046_N01.bam
01.map_to_references/1083_D-046_N01.stats 01.map_to_references/1083_D-063_BC2014-5.bam
01.map_to_references/1083_D-063_BC2014-5.stats 01.map_to_references/1083_D-064_BC2014-7.bam
01.map_to_references/1083_D-064_BC2014-7.stats 01.map_to_references/1083_D-085_BC2014-13.bam
01.map_to_references/1083_D-085_BC2014-13.stats 01.map_to_references/1083_D-086_BC2014-10.bam
01.map_to_references/1083_D-086_BC2014-10.stats 01.map_to_references/1083_D-091_BC2014-8.bam
01.map_to_references/1083_D-091_BC2014-8.stats 01.map_to_references/1083_D-096_BC2014-6.bam
```

01.map_to_references/1083_D-096_BC2014-6.stats 01.map_to_references/1083_D-097_BC2014-4.bam
01.map_to_references/1083_D-097_BC2014-4.stats 01.map_to_references/1083_D-109_BC2014-14.bam
01.map_to_references/1083_D-109_BC2014-14.stats 01.map_to_references/1083_D-114_BC2014-9.bam
01.map_to_references/1083_D-114_BC2014-9.stats 01.map_to_references/1083_D-115_BC2014-12.bam
01.map_to_references/1083_D-115_BC2014-12.stats

01.map_to_references/1083_D-136_M01.bam 01.map_to_references/1083_D-136_M01.stats
01.map_to_references/1083_D-137_K01.bam 01.map_to_references/1083_D-137_K01.stats
01.map_to_references/1083_D-142_BC2014-16.bam 01.map_to_references/1083_D-142_BC2014-16.stats
01.map_to_references/1083_D-152_I01.bam 01.map_to_references/1083_D-152_I01.stats
01.map_to_references/1083_D-153_G01.bam 01.map_to_references/1083_D-153_G01.stats
01.map_to_references/1083_D-158_J01.bam 01.map_to_references/1083_D-158_J01.stats
01.map_to_references/1083_D-163_F01.bam 01.map_to_references/1083_D-163_F01.stats
01.map_to_references/1083_D-164_H01.bam 01.map_to_references/1083_D-164_H01.stats
01.map_to_references/1083_D-169_BC2014-15.bam 01.map_to_references/1083_D-169_BC2014-15.stats
01.map_to_references/1083_D-170_L01.bam 01.map_to_references/1083_D-170_L01.stats
01.map_to_references/1083_D-187_BC2014-22.bam 01.map_to_references/1083_D-187_BC2014-22.stats
01.map_to_references/1083_D-188_BC2014-18.bam 01.map_to_references/1083_D-188_BC2014-18.stats
01.map_to_references/1083_D-202_BC2014-23.bam 01.map_to_references/1083_D-202_BC2014-23.stats
01.map_to_references/1083_D-207_BC2014-19.bam 01.map_to_references/1083_D-207_BC2014-19.stats
01.map_to_references/1083_D-208_BC2014-17.bam 01.map_to_references/1083_D-208_BC2014-17.stats
01.map_to_references/1083_D-227_BC2014-27.bam 01.map_to_references/1083_D-227_BC2014-27.stats
01.map_to_references/1083_D-228_BC2014-25.bam 01.map_to_references/1083_D-228_BC2014-25.stats
01.map_to_references/1083_D-240_BC2014-35.bam 01.map_to_references/1083_D-240_BC2014-35.stats
01.map_to_references/1083_D-241_BC2014-33.bam 01.map_to_references/1083_D-241_BC2014-33.stats
01.map_to_references/1083_D-246_BC2014-34.bam 01.map_to_references/1083_D-246_BC2014-34.stats
01.map_to_references/1083_D-251_BC2014-28.bam 01.map_to_references/1083_D-251_BC2014-28.stats
01.map_to_references/1083_D-256_BC2014-24.bam 01.map_to_references/1083_D-256_BC2014-24.stats
01.map_to_references/1083_D-257_BC2014-26.bam 01.map_to_references/1083_D-257_BC2014-26.stats
01.map_to_references/1083_D-269_BC2014-32.bam 01.map_to_references/1083_D-269_BC2014-32.stats
01.map_to_references/1083_D-270_BC2014-29.bam 01.map_to_references/1083_D-270_BC2014-29.stats
01.map_to_references/1092_A-017_I01.bam 01.map_to_references/1092_A-017_I01.stats
01.map_to_references/1092_A-018_G01.bam 01.map_to_references/1092_A-018_G01.stats
01.map_to_references/1092_A-023_J01.bam 01.map_to_references/1092_A-023_J01.stats
01.map_to_references/1092_A-028_D01.bam 01.map_to_references/1092_A-028_D01.stats
01.map_to_references/1092_A-029_B01.bam

01.map_to_references/1092_A-029_B01.stats 01.map_to_references/1092_A-034_E01.bam
01.map_to_references/1092_A-034_E01.stats 01.map_to_references/1092_A-039_A01.bam
01.map_to_references/1092_A-039_A01.stats 01.map_to_references/1092_A-040_C01.bam
01.map_to_references/1092_A-040_C01.stats 01.map_to_references/1092_A-050_F01.bam
01.map_to_references/1092_A-050_F01.stats 01.map_to_references/1092_A-051_H01.bam
01.map_to_references/1092_A-051_H01.stats 01.map_to_references/1092_A-079_L01.bam
01.map_to_references/1092_A-079_L01.stats 01.map_to_references/1092_A-080_N01.bam
01.map_to_references/1092_A-080_N01.stats 01.map_to_references/1092_A-085_O01.bam
01.map_to_references/1092_A-085_O01.stats 01.map_to_references/1092_A-090_M01.bam
01.map_to_references/1092_A-090_M01.stats 01.map_to_references/1092_A-091_K01.bam
01.map_to_references/1092_A-091_K01.stats 01.map_to_references/1092_A-108_BC2014-4.bam
01.map_to_references/1092_A-108_BC2014-4.stats 01.map_to_references/1092_A-109_BC2014-2.bam
01.map_to_references/1092_A-109_BC2014-2.stats 01.map_to_references/1092_A-114_BC2014-5.bam
01.map_to_references/1092_A-114_BC2014-5.stats 01.map_to_references/1092_A-119_P01.bam
01.map_to_references/1092_A-119_P01.stats 01.map_to_references/1092_A-120_BC2014-3.bam
01.map_to_references/1092_A-120_BC2014-3.stats 01.map_to_references/1092_A-130_BC2014-7.bam
01.map_to_references/1092_A-130_BC2014-7.stats 01.map_to_references/1092_A-131_BC2014-9.bam
01.map_to_references/1092_A-131_BC2014-9.stats 01.map_to_references/1092_A-136_BC2014-10.bam
01.map_to_references/1092_A-136_BC2014-10.stats 01.map_to_references/1092_A-150_BC2014-15.bam

01.map_to_references/1092_A-150_BC2014-15.stats 01.map_to_references/1092_A-151_BC2014-13.bam
01.map_to_references/1092_A-151_BC2014-13.stats 01.map_to_references/1092_A-156_BC2014-8.bam
01.map_to_references/1092_A-156_BC2014-8.stats 01.map_to_references/1092_A-157_BC2014-6.bam
01.map_to_references/1092_A-157_BC2014-6.stats 01.map_to_references/1092_A-171_BC2014-16.bam
01.map_to_references/1092_A-171_BC2014-16.stats 01.map_to_references/1092_A-176_BC2014-12.bam
01.map_to_references/1092_A-176_BC2014-12.stats 01.map_to_references/1092_A-177_BC2014-14.bam
01.map_to_references/1092_A-177_BC2014-14.stats 01.map_to_references/1092_A-198_BC2014-18.bam
01.map_to_references/1092_A-198_BC2014-18.stats 01.map_to_references/1092_A-199_BC2014-22.bam
01.map_to_references/1092_A-199_BC2014-22.stats 01.map_to_references/1092_A-204_BC2014-23.bam
01.map_to_references/1092_A-204_BC2014-23.stats 01.map_to_references/1092_A-211_BC2014-19.bam
01.map_to_references/1092_A-211_BC2014-19.stats

01.map_to_references/1092_A-212_BC2014-17.bam 01.map_to_references/1092_A-212_BC2014-17.stats
01.map_to_references/1092_A-227_BC2014-27.bam 01.map_to_references/1092_A-227_BC2014-27.stats
01.map_to_references/1092_A-228_BC2014-25.bam 01.map_to_references/1092_A-228_BC2014-25.stats
01.map_to_references/1092_A-233_BC2014-28.bam 01.map_to_references/1092_A-233_BC2014-28.stats
01.map_to_references/1092_A-238_BC2014-24.bam 01.map_to_references/1092_A-238_BC2014-24.stats
01.map_to_references/1092_A-239_BC2014-26.bam 01.map_to_references/1092_A-239_BC2014-26.stats
01.map_to_references/1092_A-251_BC2014-32.bam 01.map_to_references/1092_A-251_BC2014-32.stats
01.map_to_references/1092_A-252_BC2014-34.bam 01.map_to_references/1092_A-252_BC2014-34.stats
01.map_to_references/1092_A-257_BC2014-35.bam 01.map_to_references/1092_A-257_BC2014-35.stats
01.map_to_references/1092_A-262_BC2014-33.bam 01.map_to_references/1092_A-262_BC2014-33.stats
01.map_to_references/1092_A-263_BC2014-29.bam 01.map_to_references/1092_A-263_BC2014-29.stats
01.map_to_references/1093_W-017_F01.bam 01.map_to_references/1093_W-017_F01.stats
01.map_to_references/1093_W-018_D01.bam 01.map_to_references/1093_W-018_D01.stats
01.map_to_references/1093_W-023_B01.bam 01.map_to_references/1093_W-023_B01.stats
01.map_to_references/1093_W-024_A01.bam 01.map_to_references/1093_W-024_A01.stats
01.map_to_references/1093_W-029_C01.bam 01.map_to_references/1093_W-029_C01.stats
01.map_to_references/1093_W-047_H01.bam 01.map_to_references/1093_W-047_H01.stats
01.map_to_references/1093_W-052_E01.bam 01.map_to_references/1093_W-052_E01.stats
01.map_to_references/1093_W-053_G01.bam 01.map_to_references/1093_W-053_G01.stats
01.map_to_references/1093_W-065_I01.bam 01.map_to_references/1093_W-065_I01.stats
01.map_to_references/1093_W-066_K01.bam 01.map_to_references/1093_W-066_K01.stats
01.map_to_references/1093_W-078_M01.bam 01.map_to_references/1093_W-078_M01.stats
01.map_to_references/1093_W-083_J01.bam 01.map_to_references/1093_W-083_J01.stats
01.map_to_references/1093_W-084_L01.bam 01.map_to_references/1093_W-084_L01.stats
01.map_to_references/1093_W-105_N01.bam 01.map_to_references/1093_W-105_N01.stats
01.map_to_references/1093_W-106_P01.bam 01.map_to_references/1093_W-106_P01.stats
01.map_to_references/1093_W-111_BC2014-3.bam 01.map_to_references/1093_W-111_BC2014-3.stats
01.map_to_references/1093_W-121_BC2014-4.bam 01.map_to_references/1093_W-121_BC2014-4.stats
01.map_to_references/1093_W-122_BC2014-6.bam 01.map_to_references/1093_W-122_BC2014-6.stats
01.map_to_references/1093_W-134_O01.bam

01.map_to_references/1093_W-134_O01.stats 01.map_to_references/1093_W-135_BC2014-2.bam
01.map_to_references/1093_W-135_BC2014-2.stats 01.map_to_references/1093_W-145_BC2014-8.bam
01.map_to_references/1093_W-145_BC2014-8.stats 01.map_to_references/1093_W-150_BC2014-5.bam
01.map_to_references/1093_W-150_BC2014-5.stats 01.map_to_references/1093_W-151_BC2014-7.bam
01.map_to_references/1093_W-151_BC2014-7.stats 01.map_to_references/1093_W-158_BC2014-9.bam
01.map_to_references/1093_W-158_BC2014-9.stats 01.map_to_references/1093_W-159_BC2014-12.bam
01.map_to_references/1093_W-159_BC2014-12.stats 01.map_to_references/1093_W-164_BC2014-15.bam
01.map_to_references/1093_W-164_BC2014-15.stats 01.map_to_references/1093_W-185_BC2014-16.bam
01.map_to_references/1093_W-185_BC2014-16.stats 01.map_to_references/1093_W-186_BC2014-18.bam
01.map_to_references/1093_W-186_BC2014-18.stats 01.map_to_references/1093_W-191_BC2014-10.bam
01.map_to_references/1093_W-191_BC2014-10.stats 01.map_to_references/1093_W-192_BC2014-14.bam
01.map_to_references/1093_W-192_BC2014-14.stats 01.map_to_references/1093_W-207_BC2014-23.bam
01.map_to_references/1093_W-207_BC2014-23.stats 01.map_to_references/1093_W-208_BC2014-25.bam

01.map_to_references/1093_W-208_BC2014-25.stats 01.map_to_references/1093_W-213_BC2014-22.bam
01.map_to_references/1093_W-213_BC2014-22.stats 01.map_to_references/1093_W-218_BC2014-17.bam
01.map_to_references/1093_W-218_BC2014-17.stats 01.map_to_references/1093_W-219_BC2014-19.bam
01.map_to_references/1093_W-219_BC2014-19.stats 01.map_to_references/1093_W-238_BC2014-24.bam
01.map_to_references/1093_W-238_BC2014-24.stats 01.map_to_references/1093_W-239_BC2014-26.bam
01.map_to_references/1093_W-239_BC2014-26.stats 01.map_to_references/1093_W-251_BC2014-27.bam
01.map_to_references/1093_W-251_BC2014-27.stats 01.map_to_references/1093_W-252_BC2014-29.bam
01.map_to_references/1093_W-252_BC2014-29.stats 01.map_to_references/1093_W-257_BC2014-33.bam
01.map_to_references/1093_W-257_BC2014-33.stats 01.map_to_references/1093_W-269_BC2014-28.bam
01.map_to_references/1093_W-269_BC2014-28.stats 01.map_to_references/1093_W-270_BC2014-32.bam
01.map_to_references/1093_W-270_BC2014-32.stats 01.map_to_references/1094_U-012_D01.bam
01.map_to_references/1094_U-012_D01.stats 01.map_to_references/1094_U-013_B01.bam
01.map_to_references/1094_U-013_B01.stats 01.map_to_references/1094_U-018_E01.bam
01.map_to_references/1094_U-018_E01.stats 01.map_to_references/1094_U-028_F01.bam
01.map_to_references/1094_U-028_F01.stats 01.map_to_references/1094_U-029_G01.bam
01.map_to_references/1094_U-029_G01.stats 01.map_to_references/1094_U-034_H01.bam
01.map_to_references/1094_U-034_H01.stats

01.map_to_references/1094_U-039_I01.bam 01.map_to_references/1094_U-039_I01.stats
01.map_to_references/1094_U-040_J01.bam 01.map_to_references/1094_U-040_J01.stats
01.map_to_references/1094_U-047_A01.bam 01.map_to_references/1094_U-047_A01.stats
01.map_to_references/1094_U-048_C01.bam 01.map_to_references/1094_U-048_C01.stats
01.map_to_references/1094_U-065_O01.bam 01.map_to_references/1094_U-065_O01.stats
01.map_to_references/1094_U-066_M01.bam 01.map_to_references/1094_U-066_M01.stats
01.map_to_references/1094_U-078_P01.bam 01.map_to_references/1094_U-078_P01.stats
01.map_to_references/1094_U-083_L01.bam 01.map_to_references/1094_U-083_L01.stats
01.map_to_references/1094_U-084_N01.bam 01.map_to_references/1094_U-084_N01.stats
01.map_to_references/1094_U-096_BC2014-5.bam 01.map_to_references/1094_U-096_BC2014-5.stats
01.map_to_references/1094_U-097_BC2014-3.bam 01.map_to_references/1094_U-097_BC2014-3.stats
01.map_to_references/1094_U-116_BC2014-8.bam 01.map_to_references/1094_U-116_BC2014-8.stats
01.map_to_references/1094_U-117_BC2014-10.bam 01.map_to_references/1094_U-117_BC2014-10.stats
01.map_to_references/1094_U-122_BC2014-12.bam 01.map_to_references/1094_U-122_BC2014-12.stats
01.map_to_references/1094_U-127_BC2014-6.bam 01.map_to_references/1094_U-127_BC2014-6.stats
01.map_to_references/1094_U-132_BC2014-2.bam 01.map_to_references/1094_U-132_BC2014-2.stats
01.map_to_references/1094_U-133_BC2014-4.bam 01.map_to_references/1094_U-133_BC2014-4.stats
01.map_to_references/1094_U-145_BC2014-9.bam 01.map_to_references/1094_U-145_BC2014-9.stats
01.map_to_references/1094_U-146_BC2014-7.bam 01.map_to_references/1094_U-146_BC2014-7.stats
01.map_to_references/1094_U-163_BC2014-14.bam 01.map_to_references/1094_U-163_BC2014-14.stats
01.map_to_references/1094_U-164_BC2014-16.bam 01.map_to_references/1094_U-164_BC2014-16.stats
01.map_to_references/1094_U-169_BC2014-17.bam 01.map_to_references/1094_U-169_BC2014-17.stats
01.map_to_references/1094_U-183_BC2014-15.bam 01.map_to_references/1094_U-183_BC2014-15.stats
01.map_to_references/1094_U-184_BC2014-13.bam 01.map_to_references/1094_U-184_BC2014-13.stats
01.map_to_references/1094_U-198_BC2014-19.bam 01.map_to_references/1094_U-198_BC2014-19.stats
01.map_to_references/1094_U-199_BC2014-23.bam 01.map_to_references/1094_U-199_BC2014-23.stats
01.map_to_references/1094_U-204_BC2014-24.bam 01.map_to_references/1094_U-204_BC2014-24.stats
01.map_to_references/1094_U-216_BC2014-35.bam 01.map_to_references/1094_U-216_BC2014-35.stats
01.map_to_references/1094_U-217_BC2014-33.bam 01.map_to_references/1094_U-217_BC2014-33.stats
01.map_to_references/1094_U-222_BC2014-22.bam

01.map_to_references/1094_U-222_BC2014-22.stats 01.map_to_references/1094_U-223_BC2014-18.bam
01.map_to_references/1094_U-223_BC2014-18.stats 01.map_to_references/1094_U-238_BC2014-28.bam
01.map_to_references/1094_U-238_BC2014-28.stats 01.map_to_references/1094_U-239_BC2014-26.bam
01.map_to_references/1094_U-239_BC2014-26.stats 01.map_to_references/1094_U-244_BC2014-29.bam
01.map_to_references/1094_U-244_BC2014-29.stats 01.map_to_references/1094_U-249_BC2014-25.bam
01.map_to_references/1094_U-249_BC2014-25.stats 01.map_to_references/1094_U-250_BC2014-27.bam
01.map_to_references/1094_U-250_BC2014-27.stats 01.map_to_references/1094_U-257_BC2014-36.bam

01.map_to_references/1094_U-257_BC2014-36.stats 01.map_to_references/1094_U-262_BC2014-32.bam
01.map_to_references/1094_U-262_BC2014-32.stats 01.map_to_references/1094_U-263_BC2014-34.bam
01.map_to_references/1094_U-263_BC2014-34.stats 01.map_to_references/bam_stats.sh
01.map_to_references/map_reads.sh
02.counts/
02.counts/1083_D-012_O01.counts 02.counts/1083_D-013_BC2014-2.counts 02.counts/1083_D-018_BC2014-3.counts 02.counts/1083_D-028_B01.counts 02.counts/1083_D-029_D01.counts 02.counts/1083_D-034_E01.counts 02.counts/1083_D-039_C01.counts 02.counts/1083_D-040_A01.counts 02.counts/1083_D-045_P01.counts 02.counts/1083_D-046_N01.counts 02.counts/1083_D-063_BC2014-5.counts 02.counts/1083_D-064_BC2014-7.counts 02.counts/1083_D-085_BC2014-13.counts 02.counts/1083_D-086_BC2014-10.counts 02.counts/1083_D-091_BC2014-8.counts 02.counts/1083_D-096_BC2014-6.counts 02.counts/1083_D-097_BC2014-4.counts 02.counts/1083_D-109_BC2014-14.counts 02.counts/1083_D-114_BC2014-9.counts 02.counts/1083_D-115_BC2014-12.counts 02.counts/1083_D-136_M01.counts 02.counts/1083_D-137_K01.counts 02.counts/1083_D-142_BC2014-16.counts 02.counts/1083_D-152_I01.counts 02.counts/1083_D-153_G01.counts 02.counts/1083_D-158_J01.counts 02.counts/1083_D-163_F01.counts 02.counts/1083_D-164_H01.counts 02.counts/1083_D-169_BC2014-15.counts 02.counts/1083_D-170_L01.counts 02.counts/1083_D-187_BC2014-22.counts 02.counts/1083_D-188_BC2014-18.counts 02.counts/1083_D-202_BC2014-23.counts 02.counts/1083_D-207_BC2014-19.counts 02.counts/1083_D-208_BC2014-17.counts 02.counts/1083_D-227_BC2014-27.counts 02.counts/1083_D-228_BC2014-25.counts

02.counts/1083_D-240_BC2014-35.counts 02.counts/1083_D-241_BC2014-33.counts 02.counts/1083_D-246_BC2014-34.counts 02.counts/1083_D-251_BC2014-28.counts 02.counts/1083_D-256_BC2014-24.counts 02.counts/1083_D-257_BC2014-26.counts 02.counts/1083_D-269_BC2014-32.counts 02.counts/1083_D-270_BC2014-29.counts 02.counts/1092_A-017_I01.counts 02.counts/1092_A-018_G01.counts 02.counts/1092_A-023_J01.counts 02.counts/1092_A-028_D01.counts 02.counts/1092_A-029_B01.counts 02.counts/1092_A-034_E01.counts 02.counts/1092_A-039_A01.counts 02.counts/1092_A-040_C01.counts 02.counts/1092_A-050_F01.counts 02.counts/1092_A-051_H01.counts 02.counts/1092_A-079_L01.counts 02.counts/1092_A-080_N01.counts 02.counts/1092_A-085_O01.counts 02.counts/1092_A-090_M01.counts 02.counts/1092_A-091_K01.counts 02.counts/1092_A-108_BC2014-4.counts 02.counts/1092_A-109_BC2014-2.counts 02.counts/1092_A-114_BC2014-5.counts 02.counts/1092_A-119_P01.counts 02.counts/1092_A-120_BC2014-3.counts 02.counts/1092_A-130_BC2014-7.counts 02.counts/1092_A-131_BC2014-9.counts 02.counts/1092_A-136_BC2014-10.counts 02.counts/1092_A-150_BC2014-15.counts 02.counts/1092_A-151_BC2014-13.counts 02.counts/1092_A-156_BC2014-8.counts 02.counts/1092_A-157_BC2014-6.counts 02.counts/1092_A-171_BC2014-16.counts 02.counts/1092_A-176_BC2014-12.counts 02.counts/1092_A-177_BC2014-14.counts 02.counts/1092_A-198_BC2014-18.counts 02.counts/1092_A-199_BC2014-22.counts 02.counts/1092_A-204_BC2014-23.counts 02.counts/1092_A-211_BC2014-19.counts 02.counts/1092_A-212_BC2014-17.counts 02.counts/1092_A-227_BC2014-27.counts 02.counts/1092_A-228_BC2014-25.counts 02.counts/1092_A-233_BC2014-28.counts 02.counts/1092_A-238_BC2014-24.counts 02.counts/1092_A-239_BC2014-26.counts 02.counts/1092_A-251_BC2014-32.counts 02.counts/1092_A-252_BC2014-34.counts 02.counts/1092_A-257_BC2014-35.counts 02.counts/1092_A-262_BC2014-33.counts 02.counts/1092_A-263_BC2014-29.counts 02.counts/1093_W-017_F01.counts 02.counts/1093_W-018_D01.counts 02.counts/1093_W-023_B01.counts 02.counts/1093_W-024_A01.counts 02.counts/1093_W-029_C01.counts 02.counts/1093_W-047_H01.counts

02.counts/1093_W-052_E01.counts 02.counts/1093_W-053_G01.counts 02.counts/1093_W-065_I01.counts 02.counts/1093_W-066_K01.counts 02.counts/1093_W-078_M01.counts 02.counts/1093_W-083_J01.counts 02.counts/1093_W-084_L01.counts 02.counts/1093_W-105_N01.counts 02.counts/1093_W-106_P01.counts 02.counts/1093_W-111_BC2014-3.counts 02.counts/1093_W-121_BC2014-4.counts 02.counts/1093_W-122_BC2014-6.counts 02.counts/1093_W-134_O01.counts 02.counts/1093_W-135_BC2014-2.counts 02.counts/1093_W-145_BC2014-8.counts 02.counts/1093_W-150_BC2014-5.counts 02.counts/1093_W-151_BC2014-7.counts 02.counts/1093_W-158_BC2014-9.counts 02.counts/1093_W-159_BC2014-12.counts 02.counts/1093_W-164_BC2014-15.counts 02.counts/1093_W-185_BC2014-16.counts 02.counts/1093_W-186_BC2014-18.counts 02.counts/1093_W-191_BC2014-10.counts 02.counts/1093_W-192_BC2014-14.counts 02.counts/1093_W-207_BC2014-23.counts 02.counts/1093_W-208_BC2014-25.counts 02.counts/1093_W-213_BC2014-22.counts 02.counts/1093_W-218_BC2014-17.counts 02.counts/1093_W-219_BC2014-19.counts 02.counts/1093_W-238_BC2014-24.counts 02.counts/1093_W-239_BC2014-26.counts 02.counts/1093_W-251_BC2014-27.counts 02.counts/1093_W-252_BC2014-29.counts 02.counts/1093_W-257_BC2014-33.counts 02.counts/1093_W-269_BC2014-28.counts 02.counts/1093_W-270_BC2014-32.counts 02.counts/1094_U-

012_D01.counts 02.counts/1094_U-013_B01.counts 02.counts/1094_U-018_E01.counts 02.counts/1094_U-028_F01.counts 02.counts/1094_U-029_G01.counts 02.counts/1094_U-034_H01.counts 02.counts/1094_U-039_I01.counts 02.counts/1094_U-040_J01.counts 02.counts/1094_U-047_A01.counts 02.counts/1094_U-048_C01.counts 02.counts/1094_U-065_O01.counts 02.counts/1094_U-066_M01.counts 02.counts/1094_U-078_P01.counts 02.counts/1094_U-083_L01.counts 02.counts/1094_U-084_N01.counts 02.counts/1094_U-096_BC2014-5.counts 02.counts/1094_U-097_BC2014-3.counts 02.counts/1094_U-116_BC2014-8.counts 02.counts/1094_U-117_BC2014-10.counts 02.counts/1094_U-122_BC2014-12.counts 02.counts/1094_U-127_BC2014-6.counts 02.counts/1094_U-132_BC2014-2.counts 02.counts/1094_U-133_BC2014-4.counts

02.counts/1094_U-145_BC2014-9.counts 02.counts/1094_U-146_BC2014-7.counts 02.counts/1094_U-163_BC2014-14.counts 02.counts/1094_U-164_BC2014-16.counts 02.counts/1094_U-169_BC2014-17.counts 02.counts/1094_U-183_BC2014-15.counts 02.counts/1094_U-184_BC2014-13.counts 02.counts/1094_U-198_BC2014-19.counts 02.counts/1094_U-199_BC2014-23.counts 02.counts/1094_U-204_BC2014-24.counts 02.counts/1094_U-216_BC2014-35.counts 02.counts/1094_U-217_BC2014-33.counts 02.counts/1094_U-222_BC2014-22.counts 02.counts/1094_U-223_BC2014-18.counts 02.counts/1094_U-238_BC2014-28.counts 02.counts/1094_U-239_BC2014-26.counts 02.counts/1094_U-244_BC2014-29.counts 02.counts/1094_U-249_BC2014-25.counts 02.counts/1094_U-250_BC2014-27.counts 02.counts/1094_U-257_BC2014-36.counts 02.counts/1094_U-262_BC2014-32.counts 02.counts/1094_U-263_BC2014-34.counts 02.counts/all_reference_ids.txt 02.counts/count_bams.sh 02.counts/counts_summary.txt

Number of files: 538
Number of files transferred: 536 Total file size: 435.27M bytes
Total transferred file size: 435.27M bytes Literal data: 435.27M bytes
Matched data: 0 bytes File list size: 15.70K
File list generation time: 0.945 seconds File list transfer time: 0.000 seconds Total bytes sent: 435.36M
Total bytes received: 10.20K

sent 435.36M bytes received 10.20K bytes 13.82M bytes/sec total size is 435.27M speedup is 1.00

7.3 Supplementary information: publications in New Zealand Winegrower Magazine

7.3.1 Virus diversity in New Zealand grapevines: sequence, ecology and impact. Overview of the Rod Bonfiglioli scholarship research project. NZ Winegrower 2014 86:72-73

Virus diversity in New Zealand grapevines: sequence, ecology and impact

Arnaud Blouin[1,2], Howard Ross[2], Robin MacDiarmid[1,2]

[1] The New Zealand Institute for Plant & Food Research Limited, [2] School of Biological Sciences, The University of Auckland
arnaud.blouin@plantandfood.co.nz

13-115

Grapevine viruses threaten both quality and quantity of grapes and, to date, 63 viruses have been described in grapevines worldwide.

The most recent survey of viruses in New Zealand grapevines was conducted in the 1970s using techniques that have now been superseded.

This doctoral research (undertaken by Arnaud Blouin) aims to:

Develop an efficient, broad-spectrum protocol to identify viruses and viroids in New Zealand grapevines based on the most recent DNA sequencing technologies.

Undertake a large-scale survey of grapevines in New Zealand to determine the full range of grape viruses and viroids present.

Quantify the symptom expression as well as the spatial and temporal distribution of two selected viruses within distinct rootstocks and scions of grafted plants to reveal how the rootstock affects virus movement and replication in the scion.

Research outcomes will provide the winegrape sector with a better understanding of the viruses already present in New Zealand as well as offering new insights into those that threaten to cause problems in the future.

New tools for identification and genetic characterisation of viruses are being developed with the aim to increase testing volumes and speed.

We are particularly interested in grapevines that show unusual symptoms. If you have an unusual-looking vine or a non-symptomatic vine within a heavily infected block, please contact Arnaud at arnaud.blouin@plantandfood.co.nz. Samples may be collected from your selections and the viruses within them sequenced

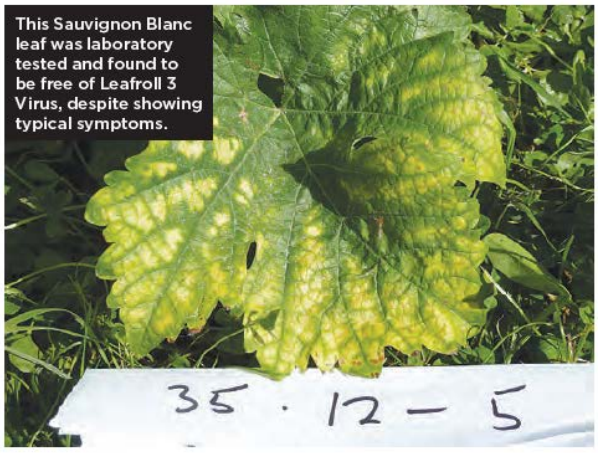
to add to this study.

This doctoral research is supported by New Zealand Winegrowers' Rod Bonfiglioli scholarship and Plant & Food Research, with supervision from Drs MacDiarmid and Ross. Dr Rod Bonfiglioli was a passionate grapevine virologist, who was amongst the first to report the extent of the genetic variability within grapevine viruses.

This project aims to respond to Dr Bonfiglioli's aspiration to assess the current status of all grapevine viruses in New Zealand and to determine their adverse impacts on quantitative and/or qualitative parameters of wine production.

This project complements alternative research approaches explored under the Grape and Wine Research Programme, which delivers research that aligns with both Plant & Food Research's and New Zealand Winegrowers' strategies.

7.3.2 Sports and spots for survey NZ Winegrower Magazine 2014 89:57



This Sauvignon Blanc leaf was laboratory tested and found to be free of Leafroll 3 Virus, despite showing typical symptoms.

SPORTS AND SPOTS FOR SURVEY

Assessing the level of virus infection across New Zealand vineyards is the subject of Arnaud Blouin's doctoral research – and he's looking for industry participation during the run-up to vintage 2015.

Blouin, a senior staff member at Plant & Food Research Ltd in Auckland, is conducting a survey of grapevine viruses already present in the country as part of his studies at the University of Auckland. He also hopes that growers and viticulturists will contact him if they spot unusual symptoms in the vineyard during the period from late January to April.

Symptoms of interest include:

- Leaf distortion, leaf spots, leaf scorch or unusual leaf colouration (early yellowing or reddening, for example)
- Extremely low vigour
- Unusual branching or uneven wood maturation on canes
- Early leaf drop
- Graft incompatibility symptoms

'The symptoms may be an indication of infection from one or more grapevine viruses, which can be identified using laboratory diagnostics,' observes Blouin. 'Some viruses might be harmless but the goal of my survey is to learn as much as possible about their presence and impact – and growers can help with that.'

Blouin is the current recipient of *New Zealand Winegrowers' Rod Bonfiglioli Scholarship*, and he is working closely with team members from the Virus Elimination Project to stop the spread of Grapevine leafroll-associated virus 3 in New Zealand vineyards. He can be contacted by email at *Arnaud.Blouin@plantandfood.co.nz* ■

7.3.3 Virus diversity in New Zealand grapevines. NZ Winegrower Magazine 2015 90:72-74

Virus diversity in New Zealand grapevines

Arnaud Blouin, Howard Ross and Robin MacDiarmid

Arnaud Blouin, a PhD candidate at the University of Auckland, is based at Plant & Food Research in Auckland, supervised by Dr Robin MacDiarmid and Dr Howard Ross.

This work is funded by New Zealand Winegrowers) with the Rod Bonfiglioli Memorial Scholarship and Plant & Food Research.

13 - 115

What do we know?

Based on the description of symptomatic vines by Bragato in 1902, there is little doubt that Grapevine leafroll disease (GLRD) has been present in and detrimental to New Zealand vineyards for more than 100 years. In 1970, Chamberlain conducted what is still the most thorough survey of viruses present in New Zealand grapevines. The survey was limited by the knowledge of grapevine viruses at the time, and the laborious methods for virus isolation and identification by graft inoculation to grapevines and mechanical inoculation to herbaceous indicator plants. He concluded that, despite no natural vector being known at the time, GLRD and Grapevine fanleaf virus (fanleaf) were the most damaging viruses for the industry, because of poor management practices used (top-grafting). With the improved knowledge and better management practices, fanleaf is now almost eradicated from commercial vineyards. We know that the main contributor to GLRD is leafroll 3 (Grapevine leafroll-associated virus 3), spread by mealybug vectors, and that this is the most economically destructive virus to New Zealand vineyards and grapevines worldwide. Because of its adverse impacts, leafroll 3 has been the focus of much research in New Zealand over recent years.

In acknowledging the importance of leafroll 3, we cannot overlook other viruses and the impact these might have on production and wine quality. Some viruses present in New Zealand have serious impacts on grapevines overseas. For example, Grapevine corky bark was reported in New Zealand until 1986. It is an important disease associated with some strains of Grapevine virus B (GVB). Grapevine corky bark causes graft union disorders, with a degree of severity mainly influenced by the rootstock selection. Symptoms include delayed maturity, lower yield and in some cases vine death. Although GVB was reported by Dr Rod Bonfiglioli in 2006, no Grapevine corky bark symptoms have been reported in New Zealand since 1986.

Furthermore, in 2003, Dr Bonfiglioli reported a new strain of leafroll 2 (Grapevine leafroll-associated virus 2) that caused graft incompatibility. A concerning aspect of his report is that this virus was apparently symptomless in some high-performing grape varieties and he concluded with the statement, "Obviously, the time has come for a full appraisal of the role of the GLRaV-2 group of viruses in high-performance viticulture".

Viruses we don't have and don't want

Recently a virus was identified in the USA causing the reddening of the leaves. The virus was named Red blotch (Grapevine red blotch-associated virus). It is now thought that the virus may have been undetected for a long time, as the symptoms can be seen as leafroll 3. The impact of the virus is similar to the one of leafroll 3 in California, with decrease of sugar accumulation and delayed ripening. Red blotch can also reduce colour in red grapes and their wine (a decrease in the anthocyanin content). The virus can be transmitted by a leafhopper (in glasshouse conditions), but there are still only few data about the rate of the virus transmission. Red blotch has been found in most grape cultivars (white, red, and table) across all the wine regions of the USA. (Figure 1).

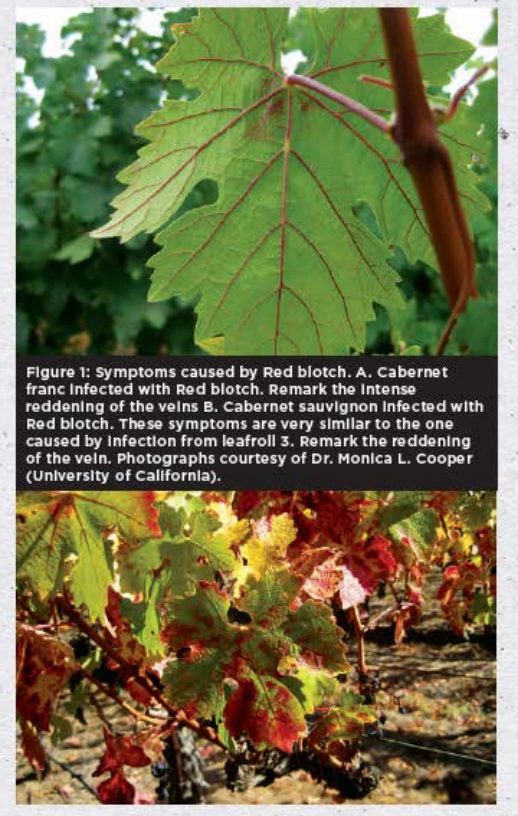


Figure 1: Symptoms caused by Red blotch. A. Cabernet franc infected with Red blotch. Remark the intense reddening of the veins B. Cabernet sauvignon infected with Red blotch. These symptoms are very similar to the one caused by infection from leafroll 3. Remark the reddening of the vein. Photographs courtesy of Dr. Monica L. Cooper (University of California).

Grapevine Pinot Gris virus was first described in 2012 in Italy. As its name implies, it was first identified in Pinot Gris. It appears to be problematic on white culti-

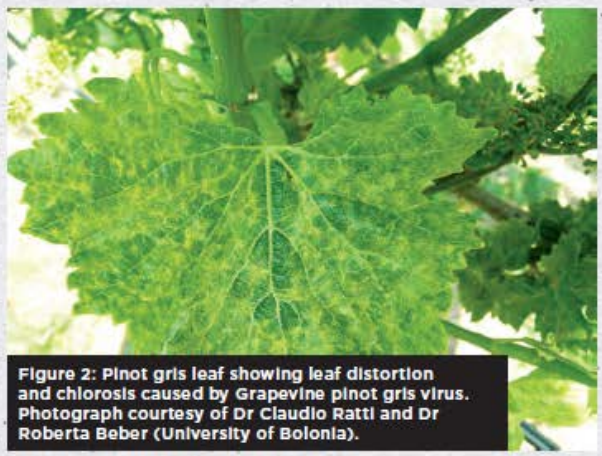


Figure 2: Pinot gris leaf showing leaf distortion and chlorosis caused by Grapevine pinot gris virus. Photograph courtesy of Dr Claudio Ratti and Dr Roberta Beber (University of Bolonia).

vars (including Sauvignon Blanc and Chardonnay), but also Pinot Noir and table grapes. Some of the symptoms can be confused with the one caused by fanleaf, including the leaf chlorosis and deformation, stunting growth and low vigour, and irregular ripening of the berries. The grapeleaf blister mite (Colomeurs vitis), a suspected vector, is common in New Zealand. In Italy, the virus appears to move at an alarming rate (this may just be a consequence of a better testing regime). The virus was recently reported in Slovakia and in South Korea. (FIGURE 2)

Nepoviruses were common to New Zealand when Chamberlain carried out his survey (1970) but have mostly been eradicated from commercial vineyards since then. Fortunately, the nematode vectoring these viruses (Xiphinema index) is absent from New Zealand soils. Some of the older plants may have survived with the virus, or the virus may have been transmitted by grafting. Symptoms vary between viruses, but they include yellow mosaic, leaf distortion, and/or asymmetrical leaves. Sometime, infected plants can be bright yellow in spring. The main virus is fanleaf (Figure 3).

Figure 3: A grapevine shot showing severe fanleaf symptoms. Photograph courtesy of Professor Giovanni Martelli (University of Bari).

What do we want to learn? - Sequence diversity – ecological impact

Recent studies have revealed extreme sequence diversity amongst grapevine viruses.

In New Zealand, Dr Karmun Chooi showed that the different strains of leafroll 3 vary by more than 20% across their genome. In other countries, high sequence variability has been observed in most grapevine viruses including Rupestris stem pitting virus (Grapevine rupestris stem pitting-associated virus), leafroll 1 (Grapevine leafroll-associated virus 1), and leafroll 2.

Viruses (especially those with RNA genomes) have high mutation rates.

The vegetative propagation of the grapevine means that a virus can infect (multiply and mutate) the same host for hundreds of years.

The consequences of the virus's genetic diversity on the grapevine are mostly unknown. In other crops we have observed that a small change in the genetic code can have spectacular consequences.

In tamarillo for example, less than 0.5% variation in the genome of Tamarillo mosaic virus is enough to make the difference between a healthy looking plant and one showing severe symptoms (Figure 4).

Figure 4: Tamarillo plants infected with two strains of Tamarillo mosaic virus sharing 99.5% of the genome. On the left, a plant infected with a severe strain showing stunted growth, necrosis on the cane and chlorosis of the leaves. On the right, a plant infected with a mild strain with no symptoms. Photograph Tim Holmes/ research project Sam Edwards.

Plant virus research has so far focused on viruses that are harmful to their host.

This harmfulness is biased to what humans desire (good looking fruits or vegetables, no obvious symptoms, and a good yield).

Cucumber mosaic virus is a common pathogen of vegetables responsible for a reduced yield and misshaped vegetables, making them unmarketable.

Recently, the virus was shown to confer drought and cold tolerance to the plant.

This detrimental virus becomes beneficial under severe conditions.

Returning now to grapevine viruses, one of the main impacts of leafroll 3 on grapevines is the delayed maturity and uneven ripening of the bunch (especially in red cultivars).

But from the plant's perspective, this phenotype allows birds to feed on the berries over a longer period, a possible advantageous feature for the dispersion of the seed.

Understanding the viruses that infect New Zealand grapes and their full, not just obviously human-centric, impacts on grapes may lead to prevention of new and debilitating grapevine diseases and new management control strategies of existing virus infections.

Survey of New Zealand grapes : what tool to use, where to sample?

Research on plant virus diversity and ecology is in its infancy mainly because of the high cost of genetic sequencing to identify new virus isolates or species in individual plants. The current knowledge of the virus diversity in New Zealand grapevines is largely based on pioneering studies carried out 40 to 50 years ago; this knowledge needs to be revised to be informative for the increasingly complex management needs of the national vineyard.

Most diagnostic tests available to detect grapevine viruses are targeted to specific known viruses and therefore do not reveal the full spectrum of viruses that are probably in our vineyards.

New technologies are now able to identify a plant virus and its genetic sequence without a priori knowledge of the host or pathogen. These techniques have proven to be very powerful but to date they have not been used for a large-scale survey. We need to assess the multiple options available, and adapt the techniques already published to construct the most useful tool for virus detection in a large-scale survey of vineyards.

This project aims to develop a practical tool to sample, extract nucleic acid, sequence and identify viruses from grapevines by analysing each step, including the grapevine sample methods and timing, the nucleic acid choices and preparations, the sequencing platforms and providers, and the bioinformatic approaches and software. The robust, high throughput, and economic tool generated will subsequently be used to undertake a large survey of New Zealand grapevines.

The survey will include old vines originating from the Te Kauwhata Viticultural Research station. When working on this station, Bragato reported replacing unsuitable varieties with new ones by top-grafting. This action resulted in the accumulation of the viruses in the rootstocks and their spread thereafter to the new scions. The Te Kauwhata Viticultural Research station has been distributing vine cuttings since the late 1800s. The survey of these older vines from the collection and the older vines from commercial vineyards will provide new insights into the extent of virus diversity in grapevines in New Zealand.

The second part of the survey will include any non-conforming vines (unusual phenotypes, or non-symptomatic vines within a heavily infected block), as well as identifying the causative agent(s) for any new disease outbreak.

As part of this research, you could help by contacting Arnaud Blouin (*arnaud.blouin@plantandfood.co.nz*) if you see any vines looking unusual.

In particular, some of the "harmful" viruses could include leaf distortion; leaf spots, leaf scorch or unusual leaf colouration (early yellowing or reddening, for example); extremely low vigour; unusual branching; uneven wood maturation on canes; early leaf drop; and/or graft incompatibility symptoms.

Why is it important?

The development of a new virus detection tool from this research will enable the New Zealand grape industry to respond rapidly and incisively should a disease causal agent with unusual symptoms be detected in any vineyard.

This ability to monitor actively is significant and timely as the industry prepares a Government Industry Agreements (GIA) on Biosecurity *Readiness and Response*.

The survey of the New Zealand vineyard will provide the industry with an in-depth and up-to-date record of the grapevine viruses and viroids in New Zealand. Based on reviews of the scientific literature, the risks associated with each microorganism will also be described.

As more virus variation is detected in many virus families, the new sequence data from the survey will be transferred to the scientific community by loading into the science web-based database GenBank. Doing so will help diagnostic laboratories to improve their testing methods in order to detect all known virus variants.

This research will also improve New Zealand's capability in grapevine virus detection and maintain the country's high profile on the international stage in this important field which is rapidly progressing with the vast amount of new information flowing from an array of new technologies. ■

Pegasus Bay, North Canterbury.
PHOTO SUPPLIED BY NZW.

7.3.4 The upsides of viruses. NZ Winegrower Magazine 2016 97: 86-87

The upsides of viruses

13-115

Arnaud G. Blouin, Vaughn A. Bell, Robin M. MacDiarmid - Plant & Food Research

VIRUSES are small infectious agents that reproduce in living cells. We are all aware of the downsides of plant viruses - the diseases they cause, the crop losses, and in some cases plant deaths. But there are upsides too – viruses can have positive effects on plant growth and wellbeing.

Since viruses were discovered in the late 19th century, virus diagnoses have focused on cultivated plants showing symptoms; as a result, the vast majority of the ~1000 plant virus species described are pathogens in crop plants. However, with new detection methods, numerous latent (or symptomless) viruses are being discovered in cultivated plants. Moreover, recent surveys for viruses in natural ecosystems show that the majority of these plants are healthy but virus-infected. This raises the question, what is the role of these viruses in seemingly healthy plants?

Several beneficial aspects of plant viruses have been documented, involving interactions with insects, fungi, other viruses, or their host plants.

The New Zealand Winegrowers Rod Bonfiglioli Memorial Scholarship, awarded to Arnaud Blouin has enabled the development of a detection tool capable of identifying all viruses present in a grapevine. With that tool, we can detect all the viruses present in a plant, not only the pathogenic ones. In addition, by looking at healthy vines, we may detect new viruses that can be friendly to their host. There are several ways in which a virus may be beneficial to its host, and it is possible that some of the viruses we can now detect will be part of the grapevine management in the future. For example, one obvious candidate would be to find an isolate (strain) of Grapevine leafroll-associated virus 3 (or leafroll 3) that is not harmful to the vine and that could be used as to protect it against the common pathogenic leafroll 3.

Are there good viruses known in grapevines?

From more than 60 viruses known to infect grapevine, less than five are very nasty and mobile, leafroll-3 being on the top of that unwanted list. The rest are, at worst, pathogenic under certain condition or only limited number of cultivar/rootstock. Grapevine rupestris stem pitting virus (GRSPaV) is the most common virus of grapevine and is probably present in the majority of the grapevines worldwide. How can this virus, with no known veetor, be so common? The answer is probably down to centuries of vegetative propagation increasing the distribution of the virus, which in turn, probably selected against the most harmful strains.

Recently, a laboratory in Italy demonstrated that GRSPaV has a small physiological and yield cost to the plant, while the benefit includes an improved tolerance to water stress. The long co-existence between grapevines and GRSPaV may have resulted in the evolution of a form of mutual adaptation between the virus and its host. This may be an example where the vineyard manager also benefits from a virus-infected grapevine.

The same Italian scientists also studied the effect on wine of Grapevine virus B (GVB), a vitivirus that is often latent but is sometimes associated with graft incompatibly. Virus-infected V. vinifera Albarossa vines yielded 25% less berry weight despite having larger berry size, but there were no significant differences in juice pH, titratable acidity or soluble solids content. The wine produced from these virus-infected vines is higher in total anthocyanins and total flavonoids, which resulted in a more highly coloured and fuller bodied wine. This was highlighted by the overall judgment of a sensory panel, who showed a preference for the wine made from the virus-infected grapes.

What other benefits can viruses provide to plants?

Virus-virus interactions

When two viruses infect the same plant, each can either positively or negatively affect the other. Virus-virus interactions can also involve the insects that transport them to the next plant (their vectors). In grapevines for example, the Vitivirus Grapevine virus A can be transmitted only by a mealybug in the presence of a grapevine leafroll-associated virus in the donor plant. Alternatively, virus-virus interactions can be beneficial to the cultivated crop when used in cross protection. This interaction works like immunisation: a plant infected with a mild strain of the virus becomes protected to a severe virus infection. Although not perfect, this strategy has been successfully used in different horticultural systems, including in commercial citrus production, where non-virulent strains of the Citrus tristeza virus protect against severe strains of the disease. As explained above, this is a disease management we would like to trial for leafroll 3, if we detect a mild strain.

Host Manipulation

Insect viruses such as baculoviruses can manipulate their hosts to move to the top of the plant just before they die. On their way up the plant, the insects spread on to the leaves their heavy viral load, ready to infect more insects. The likelihood of the insects being eaten by birds and the viruses being dispersed long distances also increases.

Plant viruses can also manipulate their hosts to increase their movement. Some have a tight relationship with their insect vectors. For example, geminiviruses significantly modify the fitness (lifespan and egg-laying potential) of whiteflies (Bemisia tabaci) feeding on virus-infected plants. These whiteflies transport the virus between plants, and increased insect longevity is beneficial for virus dispersal. On the other hand, the same whiteflies have a decreased fitness when feeding on a pepper plant infected with Tomato spotted wilt virus (TSWV). The whitefly cannot spread TSWV. In this case, TSWV can be used to reduce the impact of whiteflies (and any viruses that it may carry), even though TSWV is detrimental to the pepper plant. Similarly, clover infected with White clover mosaic virus reduces infection by the non-vector herbivore fungus gnat (Bradysia sp.). In a way, the virus protects its host plant against non-vectored competition. There are also plant

Figure 1 Semper Augustus Tulip 17th century" by Unknown - Norton Simon Museum. Licensed under Public Domain via Commons.

Well before viruses were known as such, the chance occurrence of a virus that gave rise to one of the most expensive plants in history. In the early part of the 17th century in Holland, a single tulip bulb (Figure 1) was sold for 5200 guilders.

This sum was truly significant, at the time equivalent to 400 tons of herring (13 guilders per ton) or 15600 L of French brandy (1.5 guilders per gallon).

So what was so special about this tulip? It was a speculative time for tulip traders in Holland, and colour-breaking flowers were highly sought after.

The spectacular flower-breaking phenomenon was not understood, as it was not transmissible to the seeds, but it could be propagated through bulb offsets, although these offsets were smaller and their growth was delayed relative to that of the normal bulb.

The interest was so great that bulb growers could not keep up with demand.

Three centuries later, a virus was identified as being responsible for the flower pattern, and was named Tulip breaking virus (TBV).

viruses (e.g. Turnip yellows virus) that help their aphid vectors to reduce the negative impacts of parasitoids of the vector. Complex interactions among the virus, the plant and the insect (differentially for virus vector and non-vector) have recently been observed.

Mycovirus biocontrol

Viruses are so small they can also infect other microorganisms such as fungi or bacteria. Viruses that infect fungi can in some cases reduce the pathogenicity of the fungi, and thus be used as biocontrol agents (the use of a natural enemy of pest or pathogen to manage it). Chestnut blight, caused by the pathogenic fungus Cryphonectria parasitica, was introduced to the USA from Asia at the beginning of the 20th century, and destroyed an estimated four billion American chestnut trees. Cryphonectria parasitica hypovirus 1 (€HV1) is a virus that infects the fungus and reduces its growth. CHV1 has been used to control the chestnut blight in orchards in North America and in Europe. We now know that most fungi have their own, often specific, viruses. Botrytis sp., for example, can be infected by several viruses, some reducing its fitness (hypovirulence); however, the lack of transmission of those viruses between different fungal populations has historically limited their use as biocontrol tools.

Direct benefits to the plant

Viruses can be directly beneficial to infected plants. Several common viruses that are mostly pathogenic under normal growing conditions can confer drought resistance to their hosts. This physiological change provides the infected plant with an advantage for surviving harsh conditions, and the virus with a more sustainable host.

Another ubiquitous virus – Cucumber mosaic virus – confers freezing tolerance to beet. A virus in white clover suppresses the nitrogen-fixing nodulation that captures the atmospheric nitrogen when there is already sufficient nitrogen in the soil. For the clover, the advantage is the economy of producing an organ only when it is essential to its survival.

A more dramatic example of beneficial virus is the heat-tolerant panic grass (Dichanthelium lanuginosum) growing in Yellowstone National Park in the USA. Curvularia thermal tolerance virus (CThTV), is an essential virus for panic grass survival in soils where the temperature can reach 65°C. But CThTV does not infect the plant directly; instead, the virus is specific to the fungus (Curvularia protuberata) growing on the grass. The fungus or the grass alone are unable to tolerate temperatures above 40°C. Hence, survival is entirely reliant on all components in what is a three-way symbiosis or microbial "ménage à trois".

Conclusion

These examples demonstrate how complex microbial ecology is, and how viruses can affect their hosts and vectors. Viruses are present in every part of the ecosystem, and the plant-virus interaction fluctuates between antagonism (pathogenicity) and mutualism (symbiosis) under different environmental conditions. Only good knowledge of these interactions can give us options to adjust some environmental parameters to take advantage of the positive sides of viruses.

7.3.5 Exploring NZ's vineyard virome. New Zealand Winegrower magazine 2018 108: 22-23

Exploring NZ's vineyard virome

Dr Edwin Massey

This month's column examines the recent research conducted by Plant & Food Research scientist Arnaud Blouin on the New Zealand vineyard virome and considers some of its implications for the wine industry. This research, sponsored by New Zealand Winegrowers, provides insight on the prevalence and impact of viruses in the commercial vineyard estate; helps inform useful conclusions on vineyard biosecurity and raises a number of important research questions that could help to protect the wine industry's long-term sustainability.

Grapevine infected with grapevine red blotch virus.

Virus prevalence

One of the most significant results of Blouin's work is to illustrate just how common plant viruses are in the vineyard. Of the 18 samples taken from the New Zealand Winegrowers reference collection, 17 samples indicated the presence of one or more grapevine viruses. One plant was infected with eight different grapevine viruses and two viroids! Similarly, in the commercial estate, samples collected in specific Sauvignon Blanc blocks from both Marlborough and Hawke's Bay indicated the presence of:

- One infection of grapevine leaf roll-associated virus 2
- Three infections of grapevine leaf roll-associated virus 3
- One new vitivirus in one plant
- High incidence of Maculaviruses - 10% of total vines sampled
- Very high incidence of Marafiviruses - 59% of total vines sampled
- High incidence of grapevine rupestris stem pitting virus
- Very high viroids incidence

These results are likely to differ from one block to another.

These results highlight the significant variety in the New Zealand vineyard virome and show, with the diagnostic tools that are now available, further sampling of the commercial estate will likely lead to more viruses being discovered. Grapevines and the grapevine viruses they contain have most likely co-evolved over a long time period and it is a false expectation to consider that all the vines in your vineyard will be virus free. Most importantly, these results highlight that the presence of many specific viruses does not equal disease and that, in many cases, plants which contain some viruses can be considered healthy, and still capable of producing the high quality grapes that go into making outstanding wine.

Virus management remains a high priority throughout the wine industry. The New Zealand Winegrowers Grafted Grapevine Standard seeks to minimize the probability of infected material being released to the industry. The potential impact of these pathogenic viruses is further minimised by the quality control systems in place at many grapevine nurseries; and the standard operating procedures for virus management in many vineyards.

What viruses are not in New Zealand?

This research also highlights that two important grapevine viruses; grapevine red blotch virus and grapevine Pinot Gris virus; are not present in New Zealand.

Grapevine red blotch virus has not been regularly reported outside North America. Symptoms generally occur in late summer as irregular red blotching in leaf blades. The veins of affected leaves can turn partially

or fully red. Symptoms are often confused with grapevine leafroll disease caused by another virus, grapevine leafroll associated virus 3. The primary impact of grapevine red blotch is on the accumulation of total soluble solids. Typically, infected vines can be as much as four to five units lower than healthy vines. In North America this virus primarily affects red wine cultivars such as Cabernet Sauvignon, Pinot Noir, Cabernet Franc and Merlot but has also been detected in white wine cultivars such as Chardonnay, Riesling, and Viognier.

Grapevine Pinot Gris virus is common in many international wine regions and was first identified in Australia in 2017. The symptoms associated with infection include delayed budburst, leaf distortion and mottling, shortened internodes, increased berry acidity and yield loss (reports of up to 80%). These symptoms are most pronounced in spring and may be confused with cold or herbicide damage. Grapevine Pinot Gris virus can affect a wide range of cultivars including Pinot Gris and Chardonnay.

Californian grapevine infected with grapevine Pinot Gris virus.

Next steps – further research important for improved biosecurity

Blouin's research is a great step to better understanding the New Zealand vineyard virome, yet is far from being comprehensive. Further research, and a larger sample size, would help to confirm the prevalence of a range of viruses and the absence of others, like grapevine red blotch virus or grapevine Pinot Gris virus. Further evidence of absence is important for wine industry biosecurity as it would suggest that if either of these destructive viruses are detected in New Zealand, that this would be most likely from a recent grapevine introduction. From a biosecurity perspective, it is much more cost effective to respond to a recent introduction than a virus that has been undetected for many years.

Conclusion – being aware of the unusual

Often, it can be difficult to tell whether the unusual symptoms you see in your vineyard are caused by plant viruses, or a range of other biotic and abiotic factors. A clear diagnosis is critical. That's why if you do see something that just doesn't look right you should catch it, snap it, report it. Call the MPI biosecurity hotline 0800 80 99 66 and call New Zealand Winegrowers Biosecurity and Emergency Response Manager Ed Massey, 0211924924 edwin.massey@nzwine.com. After all, it's your asset, and your responsibility to protect it. ■

7.3.6 What is your leafroll number? New Zealand Winegrower magazine 2018 110: 85-88

PROGRESS REPORTS

What is your leafroll number?

Arnaud G Blouin and Robin M MacDiarmid

13-115

THE OLD frescos from Pompeii have been preserved for centuries by the ashes of the Mount Vesuvius eruption (79 AD). It is in some of these paintings that the oldest grapevine virus symptoms can be observed. Indeed, the distorted grapevine leaves depicted resemble those caused by the virus Grapevine fanleaf virus (fanleaf). Vegetative propagation of grapevines has been a common practice for the last 6000 years, resulting in humans being the main vector of grapevine viruses with additional transmission routes superfluous to the point that some cultivars, such as the Vitis vinifera Red globe, can be identified by the viruses they host (leafroll 2).

The missionary Samuel Marsden is often credited with having planted the first grapevine in New Zealand in 1819. With the current knowledge of the ubiquity of the grapevine viruses, we can probably also attribute to him the first grapevine viruses in New Zealand. Further evidence of the presence of virus disease in New Zealand grapevines can be traced to 1902; in the 10th report of the Department of Agriculture, Romeo Bragato described Cabernet Sauvignon vines that produced no fruit and were easily distinguished from fruit-bearing vines by the early reddening of their leaves. These descriptions could be attributed to the virus Grapevine leafroll-associated virus 3 (GLRaV-3 or leafroll 3). By the 1960s, such leafroll disease was known to be widespread. Subsequently, the virus has been recognised as the most detrimental in the country. In the first official report that specifically describes leafroll disease, McKissock also reported fanleaf, an important disease of vineyards at the time. The impact of fanleaf was assessed by Chamberlain in 1970, with high incidence reported in Auckland and Hawke's Bay, the two main wine regions at the time. The alarming report by Mossop in 1986, relating a widespread occurrence of fanleaf and of the related virus Arabis mosaic virus (ArMV) in New Zealand, constitutes the last publication of viruses belonging to the genus Nepovirus in New Zealand grapevines. Ultimately, as a result of the absence of a vector in the country, and with a better control over planting material health, fanleaf, and other nepoviruses, are now considered eradicated from commercial vineyards. Since the mid-1980s to the beginning of this study, 17 other viruses have been reported in New Zealand grapevines.

Figure 1. Picture of the reddening symptoms observed on Syrah sample AB045 infected with *Grapevine rupestris stem pitting-associated virus* (GRSPaV) and two common viroids but no candidate virus that might account for the symptoms. Picture taken by Arnaud Blouin (PFR) in April 2015.

Over the last four years, Arnaud Blouin has been undertaking his doctoral research titled "Virus diversity in New Zealand grapevines: Sequence, history and impact" under the supervision of Robin MacDiarmid at the University of Auckland and Plant and Food Research. The research that was funded and supported by the Rod Bonfiglioli Memorial Scholarship as well as Plant and Food Research has been reported previously in easy to read articles (Blouin AG, Bell VA; MacDiarmid RM 2016 The upsides of viruses. New Zealand Winegrower 97: 86-87, Anon. 2015 Sports and spots for survey. New Zealand Winegrower 89:57, Blouin AG, Ross H, MacDiarmid RM 2014 Virus diversity in New Zealand grapevines: sequence, ecology and impact. Overview of the Rod Bonfiglioli scholarship research project. New Zealand Winegrower 86:72-73, Massey E 2018. Exploring NZ's vineyard virome. New Zealand Winegrower 108: 22-23). More detailed information about the research can be found within the seven research papers that comprise his thesis. Now that the thesis is submitted, we have updated its outcomes in this article.

Plants tested

The "virome" is defined by the totality of the viruses present in one environment, in this case the New Zealand environment. The first part of this study was to establish a reliable assay to detect all virtuses present in a sample. The second part was to select which plants to test. We have collected 225 grapevines from different cultivars and regions that could be divided into four groups: First, the background virome of the commercial vineyard was assessed from 166 Sauvignon

blanc samples collected from four vineyards within two major wine regions of New Zealand (labelled as the SB group). Secondly, the historical introduction of viruses into New Zealand was assessed from 19 plants collected from the New Zealand Winegrowers' germplasm collection, the so-called "low-health" block (GC-LH group). The third source sampled 16 grapevines that had undergone a virus elimination process and were planted in the "high-health" block of the germplasm collection (GC-HH group). Lastly, potential virus-disease correlations were assessed from 24 commercial grapevine samples collected over three years by growers across New Zealand concerned about their symptomatic vines (DR group, Figure 1).

Viruses detected

From the 17 viruses reported to be present in New Zealand at the start of this study, we have detected 14. In addition to fanleaf virus that was absent from New Zealand prior to this study, the three viruses not detected in this research belong to the Nepoviruses and include and ArMV, Tobacco ringspot virus (TRSV), and Tomato ringspot virus (ToRSV). These viruses are vectored through the soil by nematodes soil that are not present in New Zealand.

The 14 viruses already known to be present in the country are the leafroll viruses 1, 2, 3 and 4, the vitiviruses GVA, GVB and GVD, the ubiquitous Grapevine rupestris stem-pitting associated virus (GRSPaV) and five members of the family Tymoviridae that we group under the name 'fleck-like viruses'; they are Grapevine fleck virus (GFkV), Grapevine rupestris vein feathering virus (GVFV), Grapevine red globe virus (GRGV), Grapevine asteroid mosaic-associated virus (GAMaV) and Grapevine Syrah virus-1 (GSyV-1).

In addition, we have detected three novel viruses belonging to the Vitivirus genus, the same genus as Grapevine virus A (GVA) and Grapevine virus B (GVB). Two of these viruses, Grapevine virus G (GVG) and Grapevine virus I (GVI) were not known to science before this research, the third one is novel and a distant relative to Grapevine virus E (GVE). Lastly, a virus named Grapevine geminivirus A (GGVA) was also detected during this study. This is the first report for of this virus in the country.

Viruses distribution

As predicted, a high virus load was detected in the germplasm collection. Most of the vines deposited in this collection originated from the Te Kauwhata Viticultural Research Station, a national reference collection located in the Waikato region that was initially established in the late 1800's, then increased through imports of new grapevine cultivars. Over the ensuing decades, the cultivars were assessed and the best were distributed to New Zealand grape growers. Some of the accessions can be traced back to Romeo Bragato. The Te Kauwhata collection was not maintained in situ after the 1980s and was moved to different locations before being incorporated into the New Zealand Winegrowers germplasm collection in its current location in Lincoln. Lincoln is a region that is regarded as having low pressure from the insect vectors of viruses, in particular mealybugs. The plants are now self-rooted. Some of the plants, treated by thermotherapy to remove virus infections, were isolated in the GC-HH block alongside the most recent imports that have been screened through post-entry quarantine processes (1980s onward). From the 16 GC-HH plants sampled, an average of 1.6 viruses per plant were detected, as opposed to 4.6 virus infection per plant sampled in the GC-LH block (19 plants). Thus, the high-health and low-health terminology used to name those two blocks is justified.

The vitivirus GVG was discovered first in the New Zealand Winegrowers' germplasm where it appears to be common (68% of the plants tested in the GC-LH). The virus was then detected in one plant within the Sauvignon blanc survey in Marlborough and another plant in a commercial vineyard in the Hawke's Bay region. These single plant findings suggest that the virus was propagated outside the collection. A recent report including sequence data showed the same virus was detected in Croatia although there are substantial genetic differences to the one detected in New Zealand. The second new-to-science virus described is GVI, which is related to GVG within the vitivirus genus. GVI was detected in nine plants and these were all co-infected with GVG. This vitivirus GVI has not yet been detected outside the grapevine germplasm collection, and has not yet been reported outside New Zealand. A virus related to GVE was also detected in the survey of Sauvignon blanc from Marlborough (Figure 2). This was the first report of GVE (or a GVE-like virus) in the country. The plant was also co-infected with GVG. All these viruses belong to the genus Vitivirus with GVA, GVB and GVD that were also detected during this survey and GVF, GVH and GVJ reported in the same host overseas but not yet identified in New Zealand.

All the plants infected by one vitivirus in this research (GVA, GVB, GVD, GVE-like, GVG and GVI) were also infected with leafroll 3. In grapevine it is believed that the vitiviruses require the

Table 1: Percentage of plants infected with one of the 17 viruses detected in the seven different categories: Sauvignon Blanc vineyards in Hawke's Bay (SB-1 and SB-2) or Marlborough (SB-3 and SB-4), mostly symptomatic samples received (DR), and samples collected from the germplasm high-healt (GC-HH) and low-health (GC-LH) blocks. The number of samples per category is listed in column n.

	n	GLRaV-1	GLRaV-3	GLRaV-4	GLRaV-2	GRVFV	GSyV-1	GAMaV	GRGV	GFkV	GRSPaV	GVA	GVB	GVD	GVElike	GVG	GVI	GGVA
SB-1	44	0%	2%	0%	7%	14%	20%	0%	9%	7%	32%	0%	0%	0%	0%	0%	0%	0%
SB-2	43	0%	0%	0%	2%	7%	7%	0%	5%	7%	100%	0%	0%	0%	0%	0%	0%	0%
SB-3	40	0%	5%	0%	0%	0%	60%	0%	8%	0%	0%	0%	0%	0%	3%	3%	0%	0%
SB-4	39	0%	0%	0%	0%	0%	85%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
DR	24	0%	25%	0%	0%	0%	0%	0%	4%	0%	83%	8%	4%	0%	0%	4%	0%	0%
GP-HH	16	0%	25%	6%	6%	6%	13%	6%	25%	0%	69%	0%	0%	0%	0%	0%	0%	0%
GP-LH	19	11%	74%	5%	53%	47%	5%	0%	16%	0%	63%	58%	0%	5%	0%	68%	47%	5%

presence of a leafroll virus (GLRaV-1, -2, or -3) for its acquisition. However, a recent publication described that the vitivirus benefited from the presence of a leafroll virus to increase its replication in the Vitis host and therefore its chance of transmission rather than being dependent on transmission co-factors provided by the leafroll virus. The presence of the vitivirus does not seem to impact the leafroll virus concentration or transmission.

The detection of the geminivirus GGVA in one GC-LH grapevine constitutes the only finding of a DNA virus in the current study. This is the first report of the virus in New Zealand. The virus was originally described from the USA on imported vines from Korea. The imported plants displayed virus symptoms and were infected by multiple viruses. A subsequent survey of the US Department of Agriculture – Agricultural Research Services' Clonal Germplasm Repository – found 15 additional GGVA infected plants with no correlation with symptoms. The virus was subsequently reported in Korea and China where it may be widespread, but no association with symptoms was established.

The single known plant positive for GGVA in New Zealand is an interspecific cross, Seibel 7052. The plant was also found to be infected with leafroll-2, GRSPaV, GAMaV, and a viroid, and no symptoms were observed at the time of collection (early 2016, Figure 3) or re-collection (January 2018). In the New Zealand grapevine variety register, the source and year of importation of this plant is absent but it was logged between TK00183 (Siebel 6339) and TK00188 (Siebel 10096) that were both imported in 1957 from the US Department of Agriculture. From the information available to date, we can conclude that GGVA in New Zealand is not causing severe symptoms and its spread is very limited; after possibly 60 years it has not spread to any of the other 15 plants sampled from the same germplasm collection. The impact and the spread of the virus therefore appears to be negligible. However, this finding should generate more interest in DNA viruses and their place in the New Zealand virome.

The grapevine survey in this current study comprised 225 plants and detected a total of 17 viruses. This study represents the largest survey worldwide for grapevine viruses and uses novel sequencing technologies that have identified the most viruses reported from a single study. Such a detection rate could insinuate that New Zealand vineyards are highly infected. However, closer examination demonstrates the exact opposite. The New Zealand commercial vineyards have a low virus incidence, with less than one virus detected per vine between the four Sauvignon blanc vineyards (average detection of 0.96 viruses per vine). In most cases, the viruses detected were GRSPaV or GSyV-1 and these are considered of low to no negative impact and may even be favourable (see below). Without these two prevalent viruses, the average number of viruses detected per plant drops to 0.2 per vine. The assortment of viruses that were detected at low incidence was surprising as seven additional viruses (leafroll-2, leafroll-3, GVE-like, GVG, GRVFV, GRGV and GFkV) were detected from the Sauvignon blanc vineyards. The relatively high health of the vineyards can be credited to the studious work of the nurseries to propagate clean material and the New Zealand Winegrowers' Grafted Grapevine Standard (GGS) established and managed by the New Zealand Winegrowers that determines grapevine quality of plants sold by the nurseries, including the absence of GLRaV-3. This is also valuable information for nurseries, and it highlights the effective work of sanitation and risk awareness historically conveyed by scientists such as Dr Rod Bonfiglioli and Dr

Figure 2. Picture of Grapevine SB-3 U169 (first plant on left) infected with the *Grapevine vitivirus E* (GVE), *Grapevine vitivirus G* (GVG), and leafroll 3. Picture taken by Bex Woolley (PFR) December 2017.

Richard Smart. It also reinforces the importance of using clean plant stocks in the nurseries to avoid spreading new pathogens.

The low virus incidence found in Sauvignon blanc surveyed in this study is also the result of strict import health regulations for Vitis since the Biosecurity Act in 1993. This legal document reformed the laws related to pest and unwanted organisms to New Zealand. The new viruses reported in this research were reported to the Ministry for Primary Industries (MPI) as reports of an infectious agent new to New Zealand. In addition they also prompted discussions with members of New Zealand Winegrowers about the consequences of the high incidence of GSyV-1 and whether more viruses should be tested in addition to GLRaV-3 as part of the evolution of the grafting standards.

Impact

To date, no adverse biological impacts of the fleck-like viruses (members of the family Tymoviridae that include GFkV, GRVFV, GAMaV, GSyV-1 and GRGV) have been reported, and it was even suggested by one international expert that these viruses could be added to GRSPaV as the "virome background" of a "healthy-looking" grapevine and excluded from sanitary measures. This statement clearly resonates with the results of this current survey. The ecological impact of these viruses is still unclear but according to their incidence, GRSPaV and GSyV-1 could inhabit the same ecological niche.

The detection of multiple species of vitivirus is more problematic than the fleck-like viruses as vitiviruses can be associated with diseases. However, due to their linked transmission, the management of GLRaV-3 is likely to also remove co-infecting vitiviruses from nurseries and vineyards resulting in only low vitivirus incidence; the results of the survey confirm the strong association between GLRaV-3 and the presence of a vitivirus under New Zealand conditions.

This survey confirms that besides GLRaV-3, there is a lack of, or at most very low, movement of the other ampeloviruses (GLRaV-1 and GLRaV-4) as they were only detected in very few plants and were not found outside the germplasm collection. In the case of GLRaV-1, the two positive plants were likely to have been propagated from the same Chardonnay Mendoza imported in 1971. In contrast, the vitiviruses were more common, especially GVA and GVG. These two viruses were detected outside the germplasm. These findings would suggest that under New Zealand environments the vitiviruses GVA and GVG are vectored.

The new sequencing technologies enable the understanding of the complete virome of an environment. That holistic view has changed our understanding of the place of viruses in the environment and challenge the automatic association of virus presence with disease. Historically, virus research has focused on the diseased plant and in grapevine viral-like symptoms are distinctive. This study highlights the presence of multiple viruses in healthy-looking plants.

To date, this is the largest survey of commercial grapevines using high throughput sequencing technologies and is larger than any other similar published virus survey for any plant host. This study sheds light on the virus diversity within New Zealand vineyards, the potential route of various viruses into New Zealand, and the projected impact of the viruses based on the current knowledge of their biology.

Acknowledgements

It was an honour to be awarded the Rod Bonfiglioli Memorial Scholarship from the New Zealand Winegrowers. I now better understand the impact Rod had on the industry; his communication about the importance of vineyard health, the impact of leafroll virus and his role in the development of the grafted grapevine standards. I hope he would have appreciated my research. I smiled when I realised 'Alfie' (leafroll-2), first described by him, was still present in the vineyards.■

Figure 3. Picture of the Seibel 7052 grapevine VID280 (AB537) from the low-health germplasm collection (GC-LH) infected with *Grapevine geminivirus A* (GGVA), leafroll-2, *Grapevine rupestris stem pitting-associated virus* (GRSPaV), *Grapevine asteroid mosaic-associated virus* (GAMaV), and a viroid. Picture taken by Arnaud Blouin (PFR) in February 2016.

7.4 Supplementary information: manuscript submitted

7.4.1 Debat HJ, Zavallo D, Brisbane RS, Voncina D, Almeida RPP, Blouin AG, Al Rwahnih M, Gomez-Talquenca S, Asurmendi S. 2018. Grapevine virus L: a Novel Vitivirus in Grapevine. bioRxiv 314674; doi: https://doi.org/10.1101/314674 - submitted to Virus Gene (August 2018)

Grapevine virus L: a Novel Vitivirus in Grapevine

Humberto Debat[1Ω*], Diego Zavallo[2Ω], Reid Soltero Brisbane[3], Darko Vončina[4], Rodrigo P.P. Almeida[5], Arnaud G. Blouin [6], Maher Al Rwahnih[7], Sebastian Gomez-Talquenca[8*] and Sebastian Asurmendi[2,9]

[1]Instituto de Patología Vegetal, Centro de Investigaciones Agropecuarias, Instituto Nacional de Tecnología Agropecuaria (IPAVE-CIAP-INTA), Córdoba, Argentina, X5020ICA.

[2]Instituto de Biotecnología, Centro de Investigación en Ciencias Veterinarias y Agronómicas (IB-CICVyA-INTA), Buenos Aires, Argentina, 1686.

[3]Foundation Plant Services, Davis, CA 95616, USA

[4]Department of Plant Pathology, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia

[5]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA.

[6]The New Zealand Institute for Plant & Food Research Limited, Private Bag 92169, Auckland, 1142, New Zealand

[7]Department of Plant Pathology, University of California, Davis, CA, USA

[8]Estación Experimental Agropecuaria Mendoza, Instituto Nacional de Tecnología Agropecuaria (EEA-Mendoza-INTA), Luján de Cuyo, Mendoza, Argentina, 5534.

[9]CONICET, Argentina.

[Ω] These authors contributed equally to this work

ORCID ID:

HD, 0000-0003-3056-3739; DZ, 0000-0002-9021-2175; DV, 0000-0001-6093-5471; RPPA, 0000-0003-2888-9617; AGB, 0000-0003-1360-1529; SGT, 0000-0001-9265-8333; SA, 0000-0001-9516-5948

* Correspondence:

Humberto Debat, debat.humberto@inta.gob.ar

Sebastian Gomez-Talquenca, gomez.talquenca@inta.gob.ar

Abstract

Vitivirus are ssRNA(+) viruses in the family *Betaflexiviridae* (subfamily *Trivirinae*). There are currently ten ICTV recognized virus species in the genus; nevertheless, the extended use of NGS technologies is rapidly expanding their diversity and six more have been proposed recently. Here, we present the characterization of a novel virus from grapevines, which fits the genomic architecture and evolutionary constraints to be classifiable within the *Vitivirus* genus. The detected virus sequence is 7,607 nt long, including a typical genome organization of ORFs encoding a replicase (RP), a 22 kDa protein, a movement protein, a coat protein (CP) and a nucleic acid binding protein. Here, we present the characterization of a novel virus from grapevines. Phylogenetic analyses based on the predicted RP and CP protein unequivocally places the new virus within the *Vitivirus* genus. Multiple independent RNAseq data confirmed the presence of the detected virus in berries at diverse developmental stages. Additionally, we detected, confirmed, and assembled virus sequences from grapevine samples of distinct cultivars from America, Europe, Asia and Oceania, sharing 74.9%-97.9% nt identity, suggesting that the identified virus is widely distributed and diverse. We propose the name grapevine virus L (GVL) to the detected *Vitivirus*.

Keywords

Vitivirus, Grapevine, virus discovery, *Betaflexiviridae*

Accession numbers

Grapevine virus L sequences have been deposited in GenBank under accession numbers: MH248020 (GVL-RI), MH643739 (GVL-KA), MH681991 (GVL-VL), MH686191 (GVL-SB)

Introduction

Vitiviruses have flexuous, non-enveloped, filamentous virus particles of 725-785 nm and 12 nm in length and diameter, respectively, with a nucleocapsid that is cross-banded and diagonally striated. Vitiviruses have a linear ssRNA(+) genome (~7.3-7.6 kb), with a methylated nucleotide cap at the 5' end and a 3' poly (A) tail [1-2]. There are ten species of vitiviruses recognized by the ICTV, nevertheless, six new species have been proposed recently, five of those infecting grapevine (*Vitis vinifera*) [3-7]. Grapevine is the most prevalent natural host of vitiviruses, but they have also been found to infect several crops such as mint (*Mentha x glaciaris*), arracacha (*Arracacia xanthorrhiza*) blue agave (*Agave tequilana*), kiwi (*Actinidia chinense*) and blackberry (Rubus spp) [8-11]. Vitiviruses appear to be latent in *V. vinifera* cultivars, and so far, only *Grapevine virus A* and *Grapevine virus B* have been consistently associated to grapevine diseases of the rugose wood complex (Grapevine vitiviruses) or Shiraz disease (reviewed by Minafra [12]). The etiological role of the recently described vitiviruses should be assessed in order to establish its relationships with known or unknown viral diseases. In addition, the synergistic effects between vitiviruses and other grapevine viruses appears to be significant [13]. The availability of complete sequences of these viruses could allow the development of full-length infectious clones, fulfillment of the Koch's postulates, and improve our understanding of the biological role of these viruses [14]. Here, we present the characterization of a novel vitivirus for which we tentatively propose as grapevine virus L (GVL). We were able to detect and characterize GVL in several distinctive *Vitis vinifera* cultivars from four continents.

Material and Methods

Virus discovery, confirmation, and annotation were implemented as described in [4-5, 15-16]. RT-PCR assays were deployed as reported in Al Rwahnih et al. [17]. Raw RNA data from several Sequence Read Archive (SRA) accessions was downloaded from the National Center for Biotechnology Information database (NCBI). These files was processed in the following form: trimmed and filtered with the Trimmomatic tool as implemented in http://www.usadellab.org/cms/?page=trimmomatic, and the resulting reads of each library were assembled *de novo* with Trinity v2.6.6 release with standard parameters [18]. The obtained transcripts were subjected to bulk local blastx searches (E-value < 1e-5) against a refseq virus protein database available at ftp://ftp.ncbi.nlm.nih.gov/refseq/release/viral/viral.1.protein.faa.gz. The resulting hits were explored manually and curated by iterative mapping using Bowtie2 available at http://bowtie-bio.sourceforge.net/bowtie2/. ORFs were predicted by ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) and annotated with the NCBI conserved domain search tool as implemented in https://www.ncbi.nlm.nih.gov/Structure/cdd. Phylogenetic insights were generated by multiple amino acid alignments of replicase proteins (BLOSUM62 scoring matrix) using as best-fit algorithm E-INS-i, which drives local alignment with generalized affine gap costs, and thus is applicable for RNA polymerases where diverse domains are dispersed among several highly divergent regions, by MAFTT v7.394 as implemented in https://mafft.cbrc.jp/alignment/software/ . The aligned proteins were subsequently used as input for FastTree 2.1.5 http://www.microbesonline.org/fasttree/ maximum likelihood phylogenetic trees (best-fit model = JTT-Jones-Taylor-Thorton with single rate of evolution for each site = CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1,000 resamples. SNPs were detected by the FreeBayes v0.9.18S tool with standard parameters implemented in https://github.com/ekg/freebayes.

Results and Discussion

In order to explore the potential grapevine viral landscape on High Throughput Sequencing (HTS) publicly available libraries, we downloaded raw RNA data from several SRA accessions on the NCBI database, finding known virus in the vast majority of tested samples. The raw data were assembled *de novo* and the obtained transcripts were subjected to bulk local blastx searches (E-value < 1e-5) against a refseq virus protein database. The resulting hits were explored manually. One dataset, specifically NCBI Bioproject PRJNA400621, SRX3144921-SRX3144956, composed of RNA libraries from grape berry samples of *V. vinifera* cv. Riesling and cv. Cabernet sauvignon from Beijing, China [19], presented a 6,155 nt long transcript from the assembled transcriptome of the SRX3144956 library. This sequence obtained a highly significant hit (E-value = 0; sequence identity 69%) with the replicase protein of *Grapevine virus E* (GVE, YP_002117775.1, [20]). Iterative read mapping allowed to extend and polish the assembled contig into a 7,607 nt long sequence, supported by 25,112 reads and a mean coverage of 161.7X, sharing the genome organization of vitiviruses (**Figure 1.A**; **Table 1**). The assembled virus sequence has a pairwise identity with GVE of 68.2% at the genomic RNA level. The predicted structure of the detected virus is conformed by a 67 nt 5´UTR, followed by five ORFs and a 66 nt 3´UTR, excluding the poly (A) tail of at least three adenosines, and unknown length. Both predicted UTRs are highly similar in length and size in relation to other *Vitivirus*, more precisely GVE and the proposed GVI and GVG (**Supp. Figure 1**). ORF1 (67-5,158 nt coordinates) encodes a replicase protein (RP) of 1,696 aa (192.1 kDa), sharing a 70.2% with GVE replicase (NC_011106). We employed the NCBI conserved domain search tool to annotate the RP, which presents a characteristic domain architecture of *Vitivirus*. The 5' region of RP presents a viral methyltransferase domain (E-value = 1.76e-56; Pfam = pfam01660; 46-337 aa coordinates), which has been involved in mRNA capping, followed by a DEXDc type viral RNA helicase (E-value = 1.47e-12; Pfam = pfam01443; 750-1047 aa coordinates). Interestingly, an Alkylated DNA repair dioxygenase domain (AlkB), of the2OG-Fe(II) oxygenase superfamily, was identified within the helicase domain (E-value = 2.34e-06; COG = COG3145; 917-1032 aa coordinates) as previously described for an American isolate of GVE [21]. Finally, as expected, an RNA dependent RNA polymerase domain was found at the 3' region of the RP (RdRP_2; E-value = 3.47e-27; Pfam = pfam00978; 1313-1637 aa coordinates).

Although vitiviruses utilize a series of sub-genomic RNAs (sgRNA) as general strategy for expression of ORF2-5, it is worth mentioning that ORF2 (5,155-5,766 nt) of the detected virus overlaps in the tetranucleotide AUGA with ORF1. This overlapping cue might probably function as a termination–reinitiation strategy signal, which was determined for caliciviruses and victoriviruses for coupled start-stop translation [22]. This putative AUGA signal between ORF1-ORF2, although not mentioned, is present in the genome sequence of the recently proposed grapevine virus J [5]. The predicted ORF2 encodes a leucine rich (15.3% L) 22.8 kDa protein, of unknown function, which is the most highly divergent protein of the identified virus (44% similarity and 98% coverage to the ORF2 encoded hypothetical protein of GVE, AHB08905). ORF3 (5,795-6,607 nt) encodes a putative movement protein (MP) of 270 aa (29.7 kDa) with a viral movement protein domain (E-value = 6.61e-07; Pfam = pfam01107) between the 26-195 aa coordinates. ORF4 (6,522-7,121 nt), partially overlapping with ORF3, encodes a coat protein (CP) 199 aa long (21.4 kDa) presenting a *Trichovirus* coat protein-like domain (Tricho_coat; E-value = 4.26e-45; Pfam = pfam05892; 37-199 aa coordinates). The predicted CP of the identified virus shares 78.4% similarity with GVE-CP. Lastly, ORF5 (7,139-7,540 nt)

encodes a 133 aa putative nucleic acid binding protein (15 kDa) with a nucleic acid binding domain (Viral_NABP; E-value = 1.40e-06; Pfam = pfam05515) between the 5-76 aa coordinates. This region shares significant similarity with the CTV_P23 domain (E-value =3.79e-06) of *Citrus tristeza virus* (*Closteroviridae*) P23 protein which is involved in asymmetrical RNA Accumulation [23].

The criteria for species demarcation in the family *Betaflexiviridae*, proposed by ICTV, is less than ca. 72% overall nucleotide identity and below 80% amino acid identity of CP or replicase predicted proteins [1]. Given that the detected virus shares 68.2% similarity with GVE at the genomic RNA level, and a 70.2% and 78.2% identity with the respective RP and CP proteins (**Supp. Figure 2**), is tentatively suggested that the detected virus sequence corresponds to a new species, for which we propose the name grapevine virus L (GVL). In order to entertain this hypothesis, based on genetic criteria, we gathered evolutionary insights of the putative GVL sequence identified in cv. Riesling grapevines from China (GVL-RI) in the context of the *Betaflexiviridae* family. We generated a phylogenetic tree based on RP proteins corresponding to the 93 refseq *Betaflexiviridae* RP sequences available at NCBI. Unequivocally, the putative GVL RP clustered within the *Vitivirus* genus (**Supp. Figure 3**). To provide local tree topology and additional phylogenetic insights we generated maximum likelihood phylogenetic trees of both the RP and CP protein of accepted and proposed members of the *Vitivirus* genus. In both cases, putative GVL clustered within one of the two major sub-clades, conformed by GVE and the proposed species GVG, GVI, and branching more distantly, the agave tequilana leaf virus (**Figure 1.B-C**). These results support the tentative assignment of the detected virus as a probable member of the *Vitivirus* genus.

To validate the putative GVL detection in SRX3144956 library, we took advantage of the data generated by Chen et al [19]. We mapped the 1,014,833,524 raw reads available at the PRJNA400621 project to the assembled GVL sequence, using as threshold two nt mismatches and a 22 nt seed and a cut-off value >10 reads hits per library with the Bowtie2 tool. None of the 18 RNA libraries of cv. Cabernet Sauvignon had the putative GVL. Interestingly, we were able to detect the virus in six additional cv Riesling libraries of the same study, obtained from independent biological samples conformed by grape berries sampled in three distinct developmental stages (E-L 35, E-L 36 and E-L 38) (**Supp. Table 1**). Moreover, while analyzing the intrinsic variability of the detected GVL sequences, we noticed that two of the samples presented 24 highly supported SNPs at frequencies >0.5 as detected by the FreeBayes v0.9.18S tool (**Supp. Table 2**). Given the low p-values of the variants and that 70% of the detected SNPs were silent, we speculate that these SNPs are not artifactual.

In order to explore the distribution of GVL, we gathered data from grapevine samples from Croatia [24], USA, and New Zealand. A collection of HTS libraries obtained from petioles RNA samples of autochthonous Croatian grapevines (cv Babica, cv Vlaska, cv Dobricic and cv Ljutun; Bioproject PRJNA415169) had virus RNA reads similar to GVL. In turn, *de novo* assembly and curation of the putative GVL transcript by iterative mapping resulted in a 7,584 nt virus genome (GVL-VL, mean coverage of 675X), sharing the genomic organization and a 75.1 % nt identity with GVL-RI (**Supp. Table 3**). The presence of this virus was subsequently confirmed by targeted RT-PCR (see below) of the HTS sequenced samples. Two PCR products were bi-directionally Sanger-sequenced and the obtained sequences shared a 100 % identity with GVL-VL.

The sample from USA corresponded to a quarantine selection Katelin (KA), which was received in 2012 for inclusion in the Foundation Plant Services (FPS, UC-Davis, CA) collection. The vine was grown in a screenhouse and assayed for known grapevine viruses as described previously [16]. In addition, total RNA from

the positive sample was high-throughput sequenced and the obtained ca. 11.5 million raw reads were filtered and *de novo* assembled with parameters described in Al Rwahnih et al. 2018 [16]. A contig with high similarity with GVL was identified by BLAST and subsequently refined, resulting in a 7,591 nt virus draft genome (GVL-KA, MH643739), sharing the genomic architecture and a 97.9% nt identity with GVL-RI (**Supp. Table 3**).

Lastly, the presence of GVL was confirmed in an HTS library derived from total RNA isolated from cv. Sauvignon Blanc, which was originally collected in April 2016 in the region of Marlborough, South Island, New Zealand. The virus was detected in only one of the 225 Vitis sp. surveyed by HTS. The presence of the virus was confirmed by RT-PCR and Sanger sequencing. A draft genome with a partial truncation at 5' was assembled, encompassing 7,365 nt, lacking both the 5'UTR and a ca. 184 nt coding region of the replicase (GVL-SB, mean coverage of 166X). The obtained sequence shared a 75.2% nt identity with GVL-RI. All draft genome sequences were annotated both structurally and functionally, showing consistent signatures of vitiviruses and a shared genome organization (**Supp. Figure 4**). In turn, sequence identity values were calculated indicating high sequence similarity in the predicted coat protein (92.5% to 99.5% aa identity) and significant divergence in the 22 kDa predicted protein (53.2% to 96.6% aa identity) (**Supp. Table 3**).

Further, an RT-PCR assay was deployed as reported in Al Rwahnih et al. [17] to investigate the prevalence of GVL in grapevine samples from USA, using forward [AGTTGAAGTCTAGGTGCACAC] and reverse [GTACTCAGACTTCCCCGATCTA] primers designed to target the MH248020 sequence of GVL-RI. The specificity of this primer pair in the detection of GVL was confirmed by RT-PCR using Total nucleic acid (TNA) from vines infected with GVA, GVB, GVD, GVE, GVH, GVG, GVJ. TNA was freshly extracted from leaf petiole tissue using the MagMax 96 Viral RNA isolation kit with the MagMax Express-96 magnetic particle processor (Thermo Fisher Scientific, USA). Using the above primers in those assays, none of the grapevines that were infected with the non-target vitiviruses were found to be positive, whereas a sample of grapevine cv. Katelin tested positive. The development of this specific RT-PCR assay will allow for the detection of GVL in field tests, and in clean-stock programs, facilitating a more effective control of this virus.

Given the intrinsic significant variability observed among the diverse GVL identified isolates, in the context of an overall interspecies affinity of GVL and GVE, we proceeded with additional phylogenetic analyses. We integrated the predicted products of four assembled GVL sequences and the five available, tentatively complete, genome sequences of GVE in multiple amino acid alignments (isolates TvAQ7, of *V. labruscana* from Japan; WAHH2, of *V. vinifera* cv Cabernet sauvignon from USA; SA94, of *V. vinifera* cv. Shiraz from South Africa; VVL-101, *V. vinifera* cv. Vlaska from Croatia and GFMG-1 from China). The maximum likelihood phylogenetic trees based both on RP and CP alignments mirrored the preceding tree topology presented in **Figure 1.B** and supported clade branching between GVL and GVE in the context of clustering within *Vitiviruses* (**Figure 2.A-B**). The availability of more RNA sequences corresponding to GVL would provide insights into the evolutionary trajectory and phylodynamics of GVL. The simultaneous detection of GVL in the Americas, Europe, Asia and Oceania in diverse *V. vinifera* cultivars suggest that this novel virus is widely distributed but not widely spread in any of the environment surveyed. The emerging biological properties of GVL remain elusive. Given that the samples harboring GVL also presented several grapevine viruses and viroids, including *Grapevine rupestris stem pitting-associated virus* (*Foveavirus*), *Grapevine Syrah virus-1* (*Marafivirus*), *Grapevine leafroll-associated virus 3* (*Ampelovirus*), *Grapevine yellow speckle viroid* 1

& 2 (*Apscaviroid*), and *Hop stunt viroid* (*Hostuviroid*), future studies should unravel whether GVL presence is associated to any specific symptoms in grapevine.

Acknowledgements

This work was supported by project PE1131022 of the Instituto Nacional de Tecnología Agropecuaria (INTA) and by ANPCyT PICT 2015-1532 and PICT 2016-0429. The funders had no role in study design and analysis, decision to publish, or preparation of the manuscript. The research from New Zealand is funded by New Zealand Winegrowers Inc. and The Ministry of Business, Innovation & Employment. Many thanks to Alfredo Diaz Lara, Department of Plant Pathology, University of California, Davis, CA for his technical assistance.

Author contributions

HD, SGT and SA designed the study. RSB, DV, RPPA, AGB and MAR, performed experiments, generated and analyzed data. HD integrated the data together with SA. HD wrote the initial draft of the manuscript and DZ, SGT, AGB, RSB, DV, RPPA, MAR and SA revised the manuscript. All authors approved the final version.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors. Therefore, informed consent was not required for this work.

References

1. Adams MJ, Candresse T, Hammond J, Kreuze JF, Martelli GP, Namba S, Pearson MN, Ryu KH, Saldarelli P, Yoshikawa N (2012) Family Betaflexiviridae. In: Virus Taxonomy – Ninth Report on the International Committee on Taxonomy of Viruses. AMQ King, MJ Adams, EB Carstens, EJ Lefkowitz, eds. Elsevier Academic Press, London, pp 920-941

2. Du Preez J, Stephan D, Mawassi M, Burger JT (2011) The grapevine-infecting vitiviruses, with particular reference to grapevine virus A. Arch Virol. 156:1495.

3. Blouin AG, Chooi KM, Warren B, Napier KR, Barrero RA, MacDiarmid RM (2018) Grapevine virus I, a putative new vitivirus detected in co-infection with grapevine virus G in New Zealand. Arch Virol. 163:1371-1374.

4. Blouin AG, Keenan S, Napier KR, Barrero RA, MacDiarmid RM (2018) Identification of a novel vitivirus from grapevines in New Zealand. Arch Virol.163:281–284.

5. Diaz-Lara A, Golino D, Al Rwahnih M (2018) Genomic characterization of grapevine virus J, a novel virus identified in grapevine. Arch Virol. https://doi.org/10.1007/s00705-018-3793-y

6. Candresse T, Theil S, Faure C, Marais A (2018) Determination of the complete genomic sequence of grapevine virus H, a novel vitivirus infecting grapevine. Arch Virol. 163:277–280.

7. Jo Y, Song MK, Choi H, Park JS, Lee JW, Lian S, Cho WK (2017) Genome Sequence of Grapevine Virus K, a Novel Vitivirus Infecting Grapevine. Genome Announc. 5:e00994-17.

8. Tzanetakis IE, Postman JD, Martin RR (2007) Identification, detection and transmission of a new vitivirus from Mentha. Arch Virol.152:2027–2033.

9. Oliveira LM, Orilio AF, Inoue-Nagata AK, Nagata T, Blawid R (2017) A novel vitivirus-like sequence found in Arracacia xanthorrhiza plants by high throughput sequencing. Arch Virol. 162: 2141-2144.

10. Blouin AG, Chavan RR, Pearson MN, Macdiarmid RM, Cohen D (2012) Detection and characterisation of two novel vitiviruses infecting Actinidia. Arch Virol. 157:713–722.

11. Hassan M, Shahid MS, Tzanetakis IE. (2018). Molecular characterization and detection of a novel vitivirus infecting blackberry. Arch Virol. https://doi.org/10.1007/s00705-018-3931-6

12. Minafra A, Mawassi M, Goszczynski D, Saldarelli P (2017) Grapevine Vitiviruses. In: Grapevine Viruses. Meng B, Martelli GP, Golino DA, Fuchs M. eds. Springer International Publishing AG.

13. Rowhani A, Daubert S, Arnold K, Al Rwahnih M, Klaassen V, Golino D, Uyemoto JK (2018) Synergy between grapevine vitiviruses and grapevine leafroll viruses. Eur J Plant Pathol. https://doi.org/10.1007/s10658-018-1426-7

14. Goszczynski DE (2015) Brief report of the construction of infectious DNA clones of South African genetic variants of Grapevine virus A and Grapevine virus B. Springer Plus. 4:739.

15. Al Rwahnih M, Daubert S, Golino D, Islas C, Rowhani A (2015). Comparison of next-generation sequencing versus biological indexing for the optimal detection of viral pathogens in grapevine. Phytopathology. 105:758-763.

16. Al Rwahnih M, Rowhani A, Westrick N, Stevens K, Diaz-Lara A, Trouillas FP, Preece J, Kallsen C, Farrar K, Golino D (2018). Discovery of viruses and virus-like pathogens in pistachio using high-throughput sequencing. Plant Dis. 102:1419-1425.

17. Al Rwahnih M, Daubert S, Islas C, Golino DA, Rowhani A (2014). Characterization of a fifth Vitivirus in grapevine. J Plant Pathol. 96:2019-222.

18. Haas B J, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, MacManes MD (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8:1494.

19. Chen WK, Bai XJ, Cao MM, Cheng G, Cao XJ, Guo RR, Duan CQ (2017) Dissecting the Variations of Ripening Progression and Flavonoid Metabolism in Grape Berries Grown under Double Cropping System. Front Plant Sci. 8:1912.

20. Nakaune R, Toda S, Mochizuki M, Nakano M (2008) Identification and characterization of a new vitivirus from grapevine. Arch Virol. 153:1827.

21. Alabi OJ, Poojari S, Sarver K, Martin RR, Naidu RA (2013) Complete genome sequence analysis of an American isolate of Grapevine virus E. Virus genes. 46:563-566.

22. Li H, Havens WM, Nibert ML, Ghabrial SA (2011) RNA sequence determinants of a coupled termination-reinitiation strategy for downstream open reading frame translation in Helminthosporium victoriae virus 190S and other victoriviruses (Family Totiviridae). J Virol. 85: 7343-7352.

23. Satyanarayana T, Gowda S, Ayllón MA, Albiach-Martí MR, Rabindran S, Dawson WO (2002) The p23 protein of Citrus tristeza virus controls asymmetrical RNA accumulation. J Virol. 76:473-483.

24. Vončina D & Almeida RPP (2018). Screening of some Croatian autochthonous grapevine varieties reveals a multitude of viruses, including novel ones. Arch Virol. https://doi.org/10.1007/s00705-018-3850-6

Figure and Table Legends

Figure 1. **(A)** Molecular characterization of grapevine virus L (GVL). Mapping pattern of virus RNA reads from library SRX3144956 to the consensus putative GVL assembly indicating coverage per position. The genome architecture of GVL is characterized by a monosegmented ssRNA(+) virus encoding for five ORFs arranged as 5'-UTR-RP-HP-MP-CP-NABP-UTR-3'. The predicted product of each ORF is depicted including the associated domains determined by NCBI-CDD. Abbreviations: RP, replicase protein; 22kDa, 22k Dalton hypothetical protein; MP, movement protein; CP, coat protein; NABP, nucleic acid binding protein. See the text for additional domain information. **(B)** Phylogenetic insights of putative grapevine virus L (GVL) in relation to accepted and proposed members of the genus *Vitivirus* based on MAFFT alignments (BLOSUM62 scoring matrix using as best-fit algorithm E-INS-i) of the predicted replicase and capsid protein **(C)** followed by maximum likelihood trees generated by FastTree (best-fit model = JTT-Jones-Taylor-Thorton with single rate of evolution for each site = CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1,000 resamples). Accession numbers of the corresponding sequences: actinidia virus A (JN427014), actinidia virus B (NC_016404), mint virus 2 (AY913795), grapevine virus B (NC_003602), grapevine virus H (MF521889), grapevine virus F (NC_018458), grapevine virus A (NC_003604), grapevine virus K (NC_035202), arracacha virus V (NC_034264), grapevine virus E (NC_011106), agave tequilana leaf virus (NC_034833), grapevine virus D (KX828708), grapevine virus G (MF405923), grapevine virus I (MF927925), heracleum latent virus (X79270), and grapevine Pinot gris virus (*Trichovirus*, NC_015782) used as outgroup. Scale bar represents substitutions per site. Node labels are FastTree support values. Asterisk indicates GVL.

Figure 2. Phylogenetic insights integrating four isolates of grapevine virus L (GVL) and five isolates of grapevine virus E in the context of members of the genus *Vitivirus* based on MAFFT alignments (BLOSUM62 scoring matrix using as best-fit algorithm E-INS-i) of the predicted replicase **(A)** and capsid protein **(B)** followed by maximum likelihood trees generated by FastTree (best-fit model = JTT-Jones-Taylor-Thorton with single rate of evolution for each site = CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1,000 resamples). Accession numbers of the corresponding sequences: grapevine virus E (GVE) isolate VVL-101(MF991950), GVE isolate WAHH2 (JX402759), GVE isolate SA94 (GU903012), GVE isolate GFMG-1 (KF588015), GVE isolate TvAQ7 (NC_011106). Additional sequences are indicated in Figure 1.B legend. Pinot gris virus was used as outgroup. Scale bar represents substitutions per site. Node labels are FastTree support values. GVL isolates are depicted in bold.

Table 1. Diverse structural highlights of putative grapevine virus L in the context of proposed and ICTV recognized (in italics) *Vitivirus* species. *Grapevine Pinot gris virus* (genus *Trichovirus*) is included as outgroup for phylogenetic analyses. Abbreviations: Accession #: NCBI Genbank accession number; Genome: Genome size (nt), 5´U: 5´UTR length (nt); RP: replicase protein length (aa); HP: hypothetical protein length (aa); MP: movement protein length (aa); CP: coat protein length (aa); NABP: nucleic acid binding protein length (aa); 3´UTR: 3´UTR length excluding the poly A tail (nt).

Supplementary Figure 1. ClustalW alignment of 5'-UTR (**A**) and 3'-UTR (**B**) of GVL, GVE and the proposed GVI and GVG. GVL 5'-UTR terminal 20 nt are identical to that of GVI. At the 3'-UTR in an equidistal position of 3' terminus a ten nucleotides monomer "GGCGAAAUAA" is conserved among the four virus sequences.

Supplementary Figure 2. Genetic distances expressed as percentage values and heat-maps of predicted RP (A) and CP (B) of accepted and proposed members of the genus *Vitivirus* and GVL, based on MAFFT alignments (BLOSUM62 scoring matrix using as best-fit algorithm E-INS-I). See Table 1 for abbreviations.

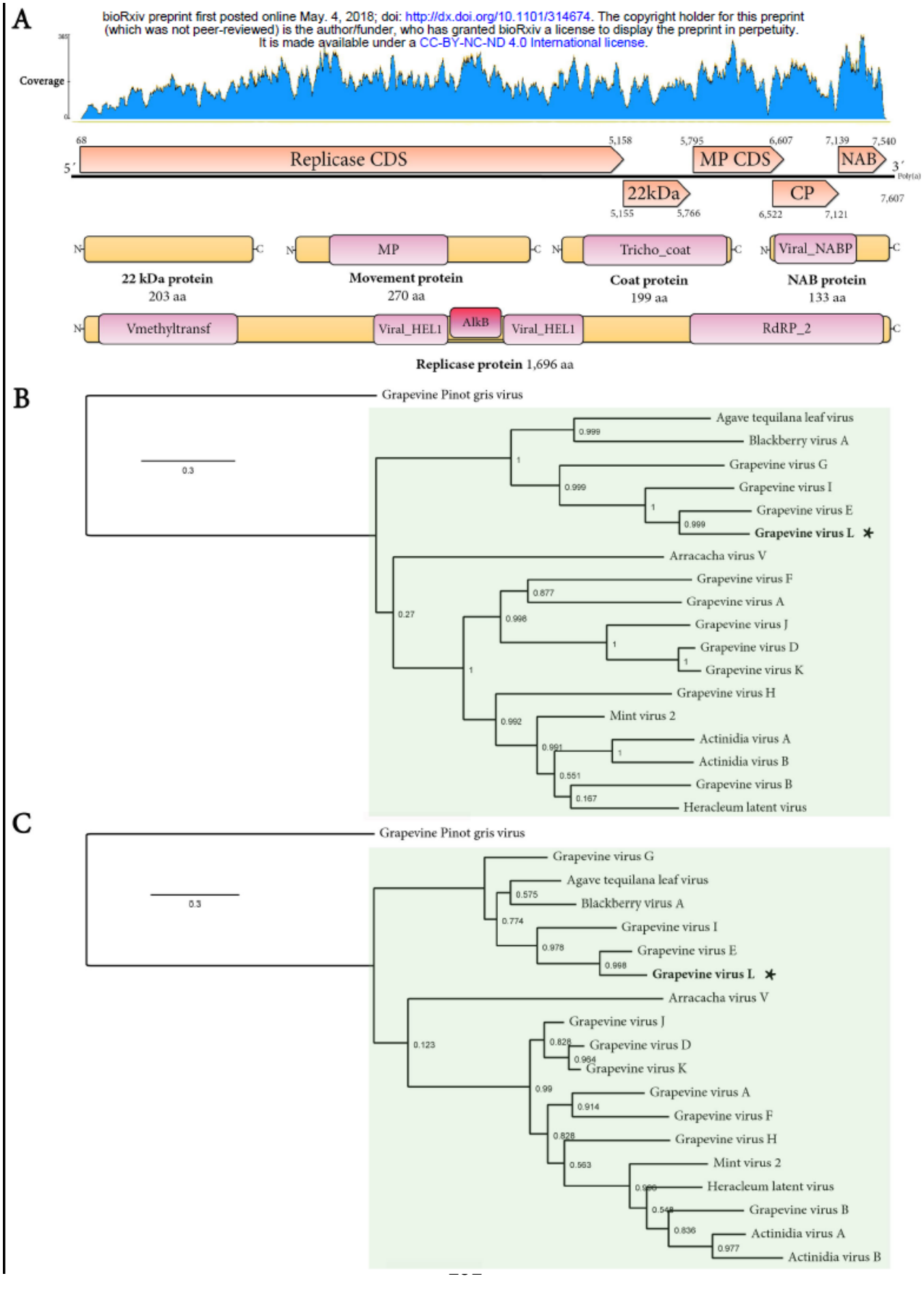
Supplementary Figure 3. Phylogenetic insights of putative grapevine virus L (GVL) in the context of family *Betaflexiviridae*, based on MAFFT alignments (BLOSUM62 scoring matrix using as best-fit algorithm E-INS-i) of NCBI refseq replicase proteins of *Betaflexiviridae* members, followed by maximum likelihood radial trees generated by FastTree (best-fit model = JTT-Jones-Taylor-Thorton with single rate of evolution for each site = CAT) computing local support values with the Shimodaira-Hasegawa test (SH) and 1,000 resamples). Scale bar represents substitutions per site. *Betaflexiviridae* family is depicted in blue and the monophyletic clade of vitiviruses is shown in red. GVL is indicated with an asterisk.

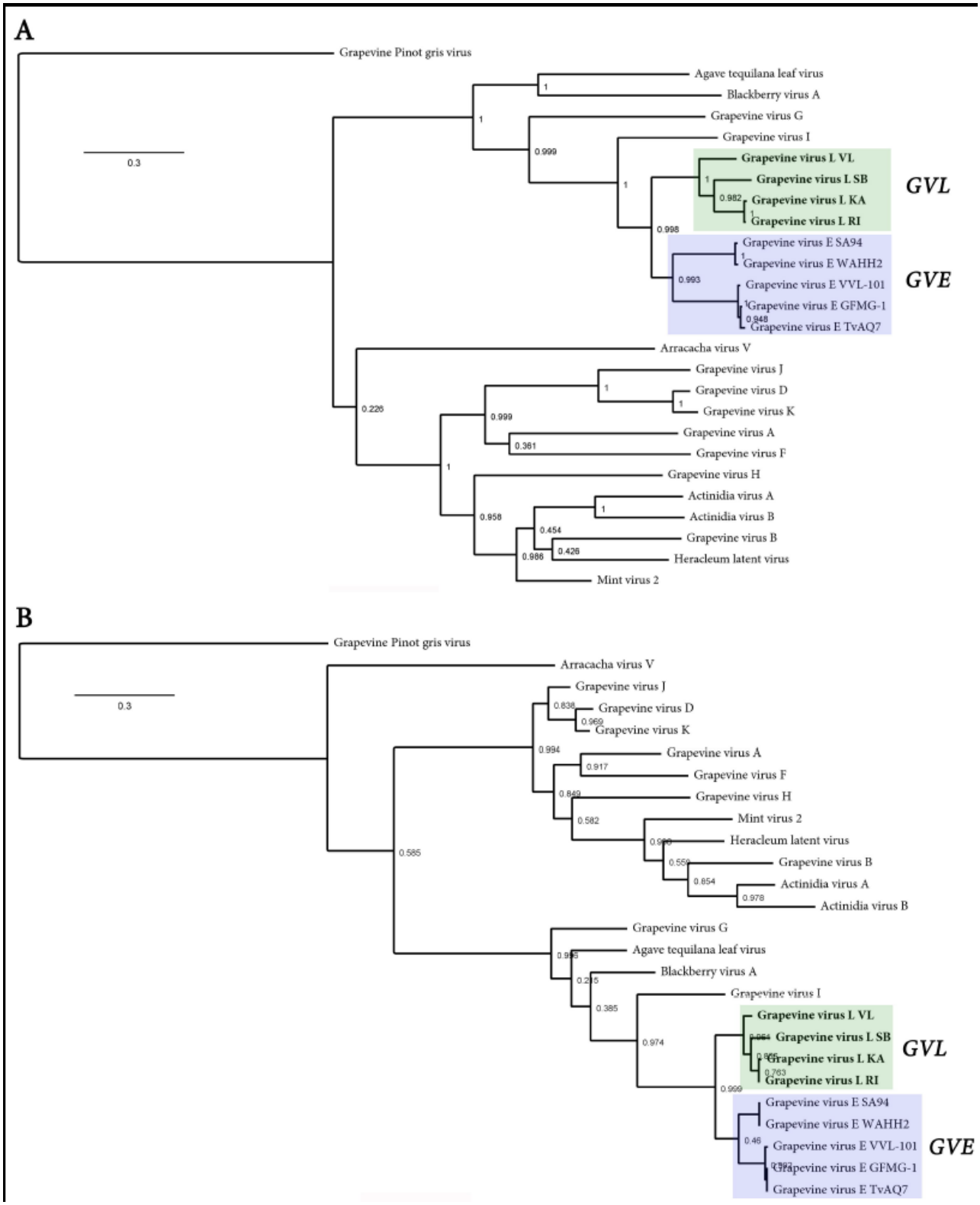
Supplementary Figure 4. Genomic architecture of the draft genomes of the identified GVL isolates characterized by a monosegmented ssRNA(+) virus encoding for five ORFs arranged as 5'-UTR-RP-HP-MP-CP-NABP-UTR-3'. The predicted product of each ORF is depicted including the associated domains determined by NCBI-CDD. Abbreviations: RP, replicase protein; 22kDa, 22k Dalton hypothetical protein; MP, movement protein; CP, coat protein; NABP, nucleic acid binding protein.

Supplementary Table 1. NGS RNA libraries used in this study generated by Chen et al [10]. Stage: sampled grapeberries developmental stage; Total reads: Total raw 50 nt reads generated for each library; GVL virus reads: viral reads detected in each sample; GVL RPM: virus reads per million total library reads.

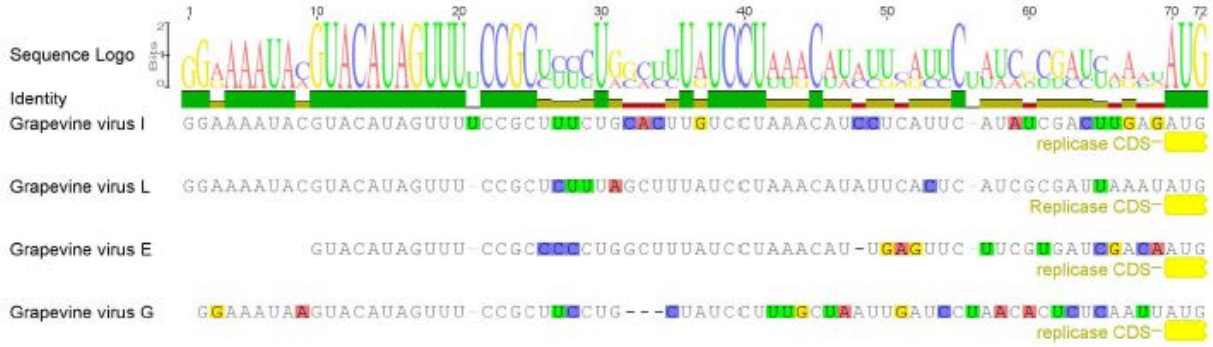
Supplementary Table 2. Variable sites of grapevine virus L identified in the virus RNA reads detected in SRX3144939- SRX3144940 (E-L 35 cv Riesling grapeberries). Polymorphism are indicated in relation with the consensus sequence of the E-L-38 cv Riesling strain determined for SRX3144956. Only variable sites with a frequency >0.5 of virus reads are shown.

Supplementary Table 3. Structural highlights and genetic distance of grapevine virus L isolates. Abbreviations: Accession #: NCBI Genbank accession number; GS: Genome size (nt), 5´U: 5´UTR length (nt); RP cds: replicase cds length (nt); HP cds: hypothetical cds length (nt); MP cds: movement protein cds length (nt); CP cds: coat protein cds length (nt); NABP: nucleic acid binding protein cds length (aa); 3´UTR: 3´UTR length excluding the poly A tail (nt). RPd-NABd: genetic distance of genomic region (nt) / predicted gene product (aa) of corresponding isolate with GVL-RI.

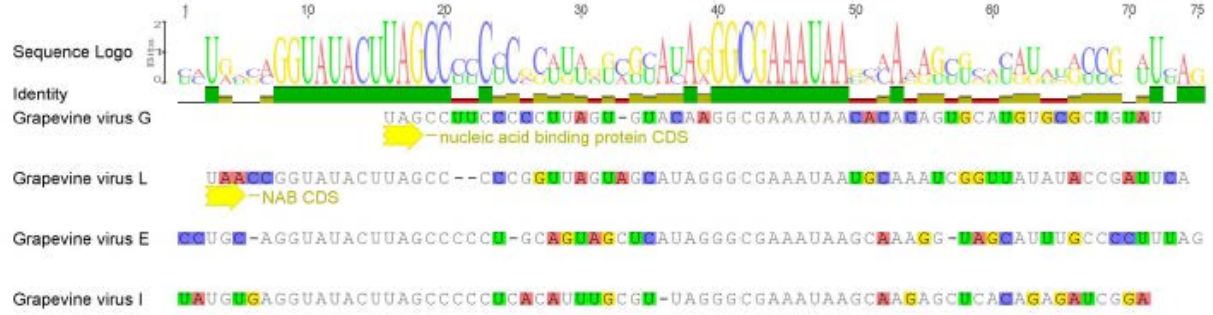




A



B

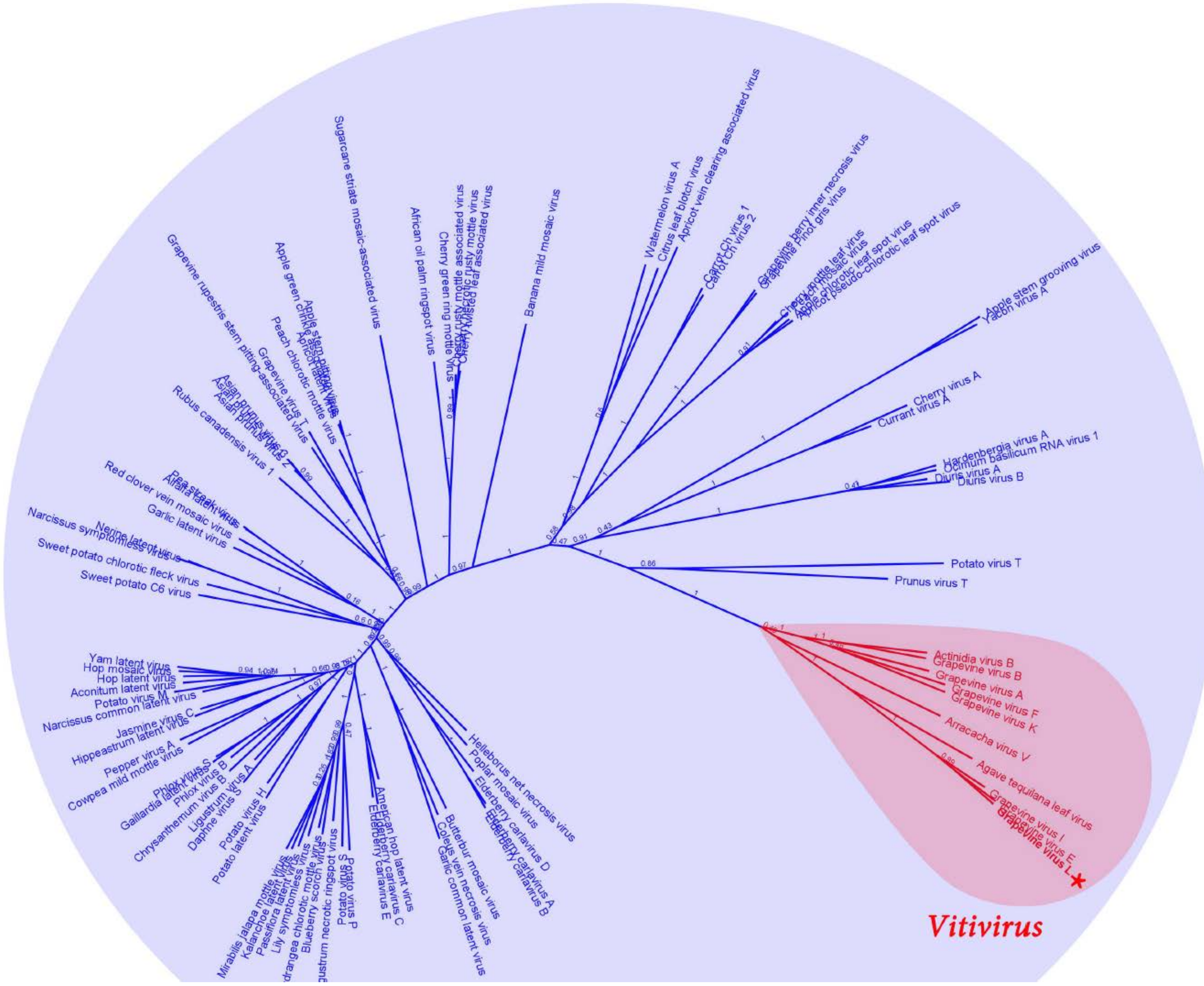


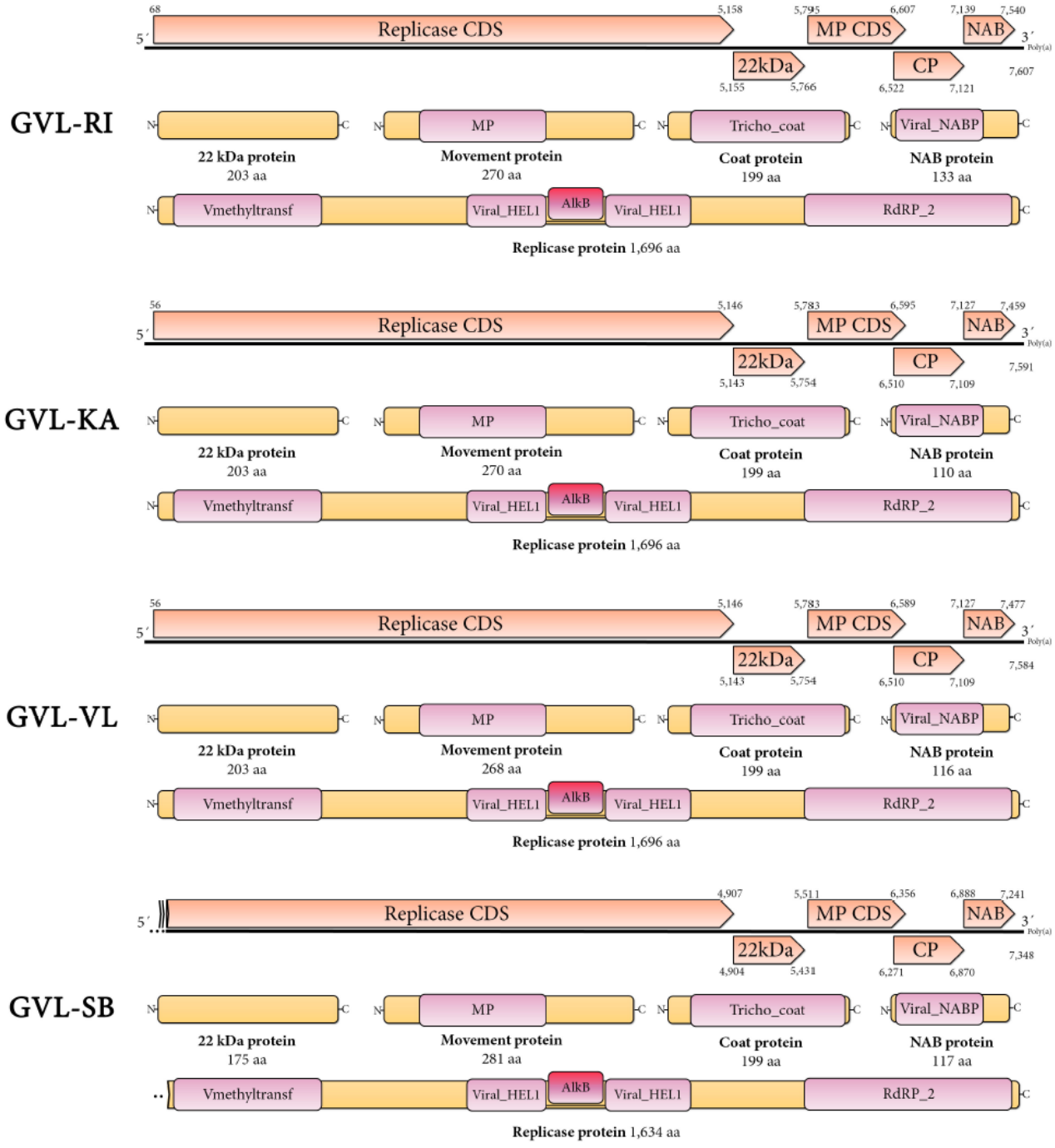
A

	HLV	MV-2	AcVB	AcVA	GVB	GVH	GVF	GVA	GVD	GVK	GVJ	GVI	* GVL	GVE	GVG	ATLV	AVV
HLV		34.6%	9.8%	9.9%	9.5%	9.1%	9.1%	8.9%	9.0%	8.9%	9.2%	8.0%	8.2%	8.2%	7.6%	8.4%	7.2%
MV-2	34.6%		72.8%	73.2%	70.6%	67.3%	66.2%	66.2%	64.4%	63.6%	66.1%	54.8%	56.3%	56.0%	54.9%	56.0%	51.9%
AcVB	9.8%	72.8%		65.9%	56.0%	50.1%	46.7%	47.1%	45.8%	45.9%	46.2%	30.5%	31.1%	31.5%	30.3%	35.1%	37.4%
AcVA	9.9%	73.2%	65.9%		54.8%	50.7%	45.5%	46.8%	45.9%	46.0%	45.5%	30.2%	31.0%	31.2%	29.0%	33.9%	38.0%
GVB	9.5%	70.6%	56.0%	54.8%		48.1%	46.7%	47.7%	46.1%	46.4%	46.6%	31.7%	31.4%	32.0%	30.6%	34.4%	37.6%
GVH	9.1%	67.3%	50.1%	50.7%	48.1%		45.8%	46.5%	45.6%	45.5%	44.9%	30.4%	30.9%	31.7%	30.3%	35.0%	37.8%
GVF	9.1%	66.2%	46.7%	45.5%	46.7%	45.8%		49.0%	48.5%	48.6%	48.6%	29.7%	30.2%	30.5%	30.8%	35.0%	36.4%
GVA	8.9%	66.2%	47.1%	46.8%	47.7%	46.5%	49.0%		50.2%	50.0%	50.1%	30.3%	30.7%	30.3%	30.7%	34.1%	37.2%
GVD	9.0%	64.4%	45.8%	45.9%	46.1%	45.6%	48.5%	50.2%		88.9%	65.6%	30.0%	30.6%	30.6%	29.3%	35.4%	37.9%
GVK	8.9%	63.6%	45.9%	46.0%	46.4%	45.5%	48.6%	50.0%	88.9%		65.1%	30.2%	30.6%	30.7%	29.1%	35.1%	37.4%
GVJ	9.2%	66.1%	46.2%	45.5%	46.6%	44.9%	48.6%	50.1%	65.6%	65.1%		29.9%	31.2%	31.0%	30.1%	35.1%	38.4%
GVI	8.0%	54.8%	30.5%	30.2%	31.7%	30.4%	29.7%	30.3%	30.0%	30.2%	29.9%		64.7%	64.2%	51.6%	41.7%	28.3%
* GVL	8.2%	56.3%	31.1%	31.0%	31.4%	30.9%	30.2%	30.7%	30.6%	30.6%	31.2%	64.7%		70.2%	51.0%	41.5%	28.8%
GVE	8.2%	56.0%	31.5%	31.2%	32.0%	31.7%	30.5%	30.3%	30.6%	30.7%	31.0%	64.2%	70.2%		50.7%	40.8%	28.9%
GVG	7.6%	54.9%	30.3%	29.0%	30.6%	30.3%	30.8%	30.7%	29.3%	29.1%	30.1%	51.6%	51.0%	50.7%		41.5%	27.3%
ATLV	8.4%	56.0%	35.1%	33.9%	34.4%	35.0%	35.0%	34.1%	35.4%	35.1%	35.1%	41.7%	41.5%	40.8%	41.5%		33.6%
AVV	7.2%	51.9%	37.4%	38.0%	37.6%	37.8%	36.4%	37.2%	37.9%	37.4%	38.4%	28.3%	28.8%	28.9%	27.3%	33.6%	

B

	AVV	GVB	AcVA	AcVB	HLV	MV-2	GVD	GVK	GVJ	GVA	GVH	GVF	ATLV	GVG	GVI	GVE	* GVL
AVV		44.6%	45.6%	43.5%	45.1%	43.0%	45.1%	44.0%	44.6%	41.2%	43.5%	39.2%	41.5%	39.5%	39.0%	38.0%	38.0%
GVB	44.6%		68.0%	60.9%	65.5%	62.4%	58.9%	57.9%	57.4%	55.1%	54.8%	54.0%	41.5%	36.6%	35.5%	37.5%	37.5%
AcVA	45.6%	68.0%		75.3%	67.5%	60.8%	58.4%	59.9%	57.9%	58.6%	56.9%	55.6%	41.0%	37.6%	37.0%	39.0%	37.5%
AcVB	43.5%	60.9%	75.3%		65.5%	57.4%	57.4%	58.4%	57.4%	54.5%	54.3%	56.6%	37.0%	34.7%	34.0%	36.0%	36.0%
HLV	45.1%	65.5%	67.5%	65.5%		66.5%	62.4%	59.4%	64.0%	55.6%	56.3%	57.6%	38.5%	38.6%	39.0%	38.5%	38.0%
MV-2	43.0%	62.4%	60.8%	57.4%	66.5%		58.4%	58.4%	59.4%	58.1%	53.8%	54.0%	38.0%	36.6%	35.5%	36.0%	36.5%
GVD	45.1%	58.9%	58.4%	57.4%	62.4%	58.4%		91.9%	84.3%	68.7%	62.9%	62.6%	43.5%	44.1%	39.5%	43.5%	43.0%
GVK	44.0%	57.9%	59.9%	58.4%	59.4%	58.4%	91.9%		84.3%	69.2%	63.5%	64.1%	43.0%	43.1%	38.5%	43.0%	42.5%
GVJ	44.6%	57.4%	57.9%	57.4%	64.0%	59.4%	84.3%	84.3%		69.2%	62.9%	66.2%	46.0%	44.6%	40.0%	44.5%	44.0%
GVA	41.2%	55.1%	58.6%	54.5%	55.6%	58.1%	68.7%	69.2%	69.2%		58.1%	61.6%	43.8%	39.4%	38.8%	41.8%	41.8%
GVH	43.5%	54.8%	56.9%	54.3%	56.3%	53.8%	62.9%	63.5%	62.9%	58.1%		58.6%	42.0%	41.1%	39.5%	39.0%	39.0%
GVF	39.2%	54.0%	55.6%	56.6%	57.6%	54.0%	62.6%	64.1%	66.2%	61.6%	58.6%		39.3%	39.9%	37.8%	38.8%	39.8%
ATLV	41.5%	41.5%	41.0%	37.0%	38.5%	38.0%	43.5%	43.0%	46.0%	43.8%	42.0%	39.3%		66.8%	65.3%	60.3%	56.8%
GVG	39.5%	36.6%	37.6%	34.7%	38.6%	36.6%	44.1%	43.1%	44.6%	39.4%	41.1%	39.9%	66.8%		61.8%	58.3%	54.8%
GVI	39.0%	35.5%	37.0%	34.0%	39.0%	35.5%	39.5%	38.5%	40.0%	38.8%	39.5%	37.8%	65.3%	61.8%		62.8%	62.3%
GVE	38.0%	37.5%	39.0%	36.0%	38.5%	36.0%	43.5%	43.0%	44.5%	41.8%	39.0%	38.8%	60.3%	58.3%	62.8%		78.4%
* GVL	38.0%	37.5%	37.5%	36.0%	38.0%	36.5%	43.0%	42.5%	44.0%	41.8%	39.0%	39.8%	56.8%	54.8%	62.3%	78.4%	


Sugarcane striate mosaic-associated virus
African oil palm ringspot virus
Cherry green ring mottle virus
Banana mild mosaic virus
Watermelon virus A
Citrus leaf blotch virus
Apricot vein clearing associated virus
Carrot Ch virus 1
Carrot Ch virus 2
Grapevine berry inner necrosis virus
Grapevine Pinot gris virus
Cherry mottle leaf virus
Peach mosaic virus
Apricot chlorotic leaf spot virus
Apricot pseudo-chlorotic leaf spot virus
Apple stem grooving virus
Yacon virus A
Cherry virus A
Currant virus A
Hardenbergia virus A
Ocimum basilicum RNA virus 1
Diuris virus A
Diuris virus B
Potato virus T
Prunus virus T
Actinidia virus B
Grapevine virus B
Grapevine virus A
Grapevine virus F
Grapevine virus K
Arracacha virus V
Agave tequilana leaf virus
Grapevine virus I
Grapevine virus E
Grapevine virus L
Vitivirus
Grapevine rupestris stem pitting-associated virus
Apple green crinkle associated virus
Peach chlorotic mottle virus
Grapevine virus T
Rubus canadensis virus 1
Red clover vein mosaic virus
Garlic latent virus
Narcissus symptomless virus
Nerine latent virus
Sweet potato chlorotic fleck virus
Sweet potato C6 virus
Yam latent virus
Hop mosaic virus
Hop latent virus
Aconitum latent virus
Potato virus M
Narcissus common latent virus
Jasmine virus C
Hippeastrum latent virus
Pepper virus A
Cowpea mild mottle virus
Phlox virus S
Phlox virus B
Gaillardia latent virus
Chrysanthemum virus B
Ligustrum virus A
Daphne virus S
Potato virus H
Potato latent virus
Mirabilis jalapa mottle virus
Kalanchoe latent virus
Passiflora latent virus
Lily symptomless virus
Blueberry scorch virus
Potato virus P
Potato virus S
American hop latent virus
Elderberry carlavirus C
Elderberry carlavirus E
Butterbur mosaic virus
Coleus vein necrosis virus
Garlic common latent virus
Helleborus net necrosis virus
Poplar mosaic virus
Elderberry carlavirus D
Elderberry carlavirus A
Elderberry carlavirus B


GVL-RI
68
5,158
5,795
6,607
7,139
7,540
5′
Replicase CDS
MP CDS
NAB
3′
Poly(a)
22kDa
5,155
5,766
CP
6,522
7,121
7,607
N
C
MP
Tricho_coat
Viral_NABP
22 kDa protein
203 aa
Movement protein
270 aa
Coat protein
199 aa
NAB protein
133 aa
Vmethyltransf
Viral_HEL1
AlkB
Viral_HEL1
RdRP_2
Replicase protein 1,696 aa
GVL-KA
56
5,146
5,783
6,595
7,127
7,459
5′
Replicase CDS
MP CDS
NAB
3′
Poly(a)
22kDa
5,143
5,754
CP
6,510
7,109
7,591
N
C
MP
Tricho_coat
Viral_NABP
22 kDa protein
203 aa
Movement protein
270 aa
Coat protein
199 aa
NAB protein
110 aa
Vmethyltransf
Viral_HEL1
AlkB
Viral_HEL1
RdRP_2
Replicase protein 1,696 aa
GVL-VL
56
5,146
5,783
6,589
7,127
7,477
5′
Replicase CDS
MP CDS
NAB
3′
Poly(a)
22kDa
5,143
5,754
CP
6,510
7,109
7,584
N
C
MP
Tricho_coat
Viral_NABP
22 kDa protein
203 aa
Movement protein
268 aa
Coat protein
199 aa
NAB protein
116 aa
Vmethyltransf
Viral_HEL1
AlkB
Viral_HEL1
RdRP_2
Replicase protein 1,696 aa
GVL-SB
4,907
5,511
6,356
6,888
7,241
5′
Replicase CDS
MP CDS
NAB
3′
Poly(a)
22kDa
4,904
5,431
CP
6,271
6,870
7,348
N
C
MP
Tricho_coat
Viral_NABP
22 kDa protein
175 aa
Movement protein
281 aa
Coat protein
199 aa
NAB protein
117 aa
Vmethyltransf
Viral_HEL1
AlkB
Viral_HEL1
RdRP_2
Replicase protein 1,634 aa

Virus	*Abbreviation*	*Accession #*	*Genome*	*5'UTR*	*RP*	*HP*	*MP*	*CP*	*NABP*	*3'UTR*
Actinidia virus A	AcVA	JN427014	7566	NA	1714	219	297	198	105	137
Actinidia virus B	AcVB	NC_016404	7488	70	1707	231	290	198	106	83
Agave tequilana leaf virus	ATLV	NC_034833	6958	52	1539	151	267	199	126	86
Arracacha virus V	AVV	NC_034264	7398	51	1705	NA	293	192	110	73
Blackberry virus A	BVA	MG254193	7285	59	1676	148	272	199	109	18
Grapevine virus A	GVA	NC_003604	7351	86	1707	177	278	198	90	68
Grapevine virus B	GVB	NC_003602	7599	48	1707	179	322	197	123	148
Grapevine virus D	GVD	MF774336	7479	79	1698	161	273	197	91	87
Grapevine virus E	GVE	NC_011106	7564	57	1698	191	265	199	108	134
Grapevine virus F	GVF	NC_018458	7551	90	1727	159	271	198	109	82
Grapevine virus G	GVG	MF405923	7496	64	1703	154	286	201	117	53
Grapevine virus H	GVH	MF521889	7446	97	1717	162	265	197	106	82
Grapevine virus I	GVI	NC_037058	7507	68	1696	167	264	199	121	104
Grapevine virus J	GVJ	MG637048	7390	95	1702	164	273	197	103	69
Grapevine virus K	GVK	NC_035202	7476	80	1700	161	273	197	91	77
Heracleum latent virus	HLV	X79270	3006	NA	253	NA	343	197	115	56
Mint virus 2	MV-2	AY913795	3897	NA	500	209	329	209	111	85
Grapevine virus L	GVL	MH248020	7607	67	1696	203	270	199	133	67
Vitivirus x□	NA	NA	7438	72	1692	177	286	198	108	89
Grapevine Pinot gris virus	GPGV	NC_015782	7275	94	1845	NA	369	195	NA	99

Table 1 Diverse structural highlights of grapevine virus L in the context of proposed and ICTV recognized (in italics) *Vitivirus* species. *Grapevine Pinot gris virus* (genus *Trichovirus*) is included as outgroup for phylogenetic analyses. Accession #: NCBI Genbank accession number; Genome: Genome size (nt), 5´U: 5´UTR length (nt); RP: replicase protein length (aa); HP: hypothetical protein length (aa); MP: movement protein length (aa); CP: coat protein length (aa); NABP: nucleic acid binding protein length (aa); 3´UTR: 3´UTR length excluding the poly A tail (nt).