Sequencing and detection of a new strain of grapevine leafroll-associated virus 3 in South Africa

by

Rachelle Bester

Thesis presented in partial fulfilment of the requirements for the degree Master of Science in Genetics at Stellenbosch University

Supervisor: Prof. J.T. Burger
Co-supervisor: Dr. H.J. Maree
Faculty of Science

December 2012

Declaration

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2012

Abstract

Grapevine leafroll-associated virus 3 (GLRaV-3) is the type member of the genus *Ampelovirus* in the family *Closteroviridae* and is considered to be the main contributing agent of grapevine leafroll disease (GLD) worldwide. A metagenomic sequencing study of a grapevine leafroll-diseased vineyard led to the discovery of a new variant of GLRaV-3 in South Africa. This new variant was most related to a New Zealand isolate, NZ-1. In this study, we sequenced two isolates, GH11 and GH30, of the new variant group of GLRaV-3. These isolates have less than 70% nucleotide (nt) identity to other known GLRaV-3 variants, indicating that they should be considered variants of a different strain of GLRaV-3. We propose that the GLRaV-3-like virus identified in this study be grouped together with NZ-1 and some Napa Valley isolates as Group VI of GLRaV-3. This study also provided further evidence that next-generation sequencing is an invaluable approach to identify novel viruses and variants, in that the draft sequence generated with bioinformatic tools in this study was 98% identical to the GH11 sequence generated using Sanger sequencing. The study further confirmed that the industry standard ELISA is still an effective GLRaV-3 diagnostic method and that it is able to detect all known variant groups of GLRaV-3. However, this assay is not able to differentiate between GLRaV-3 variant groups. In the current study therefore, a real-time RT-PCR was designed that is able to detect GLRaV-3 variant groups I, II, III and VI, using a single primer pair targeting the Hsp70h gene of GLRaV-3. If high-resolution melting (HRM) curve analysis is added to the real-time RT-PCR, it is possible to differentiate between variant groups based on three melting point intervals. The RT-PCR HRM assay provides a more sensitive and rapid tool to detect and differentiate between different GLRaV-3 variant groups. Finally, a multiplex RT-PCR was designed to differentiate between the variant groups present in South Africa. This multiplex RT-PCR offers a validation method for the RT-PCR HRM and provides an end-point PCR alternative for variant identification. In order to investigate the spread and impact of different GLRaV-3 variants in vineyards, sensitive diagnostic techniques are a necessity. The abovementioned tools will contribute to the understanding of the pathogenesis of GLD and aid epidemiological studies to investigate how these different GLRaV-3 variant groups are spreading, the association of specific GLRaV-3 variants to disease symptoms and the mealybug vector transmission efficiency for each GLRaV-3 variant.

Opsomming

Grapevine leafroll-associated virus 3 (GLRaV-3) is ’n lid van die genus *Ampelovirus* in die familie *Closteroviridae* en word beskou as die hoof bydraende faktor van wingerd-rolbladsiekte wêreldwyd. ’n Metagenomiese studie het bewys dat daar ’n nuwe variant van GLRaV-3 bestaan wat nog nie voorheen in Suid Afrika opgespoor kon word met die huidige opsporingsmetodes nie. Hierdie nuwe variant was naaste verwant aan ’n Nieu-Seelandse isolaat, NZ-1. In hierdie studie is die genoomvolgorde van twee isolate, GH11 en GH30, van hierdie nuwe GLRaV-3 variant groep bepaal. Hierdie twee isolate was minder as 70% identies aan ander GLRaV-3 variante, wat daarop dui dat hulle as variante van ’n nuwe virus-ras beskou behoort te word. Ons beveel aan dat hierdie GLRaV-3-verwante virus geklassifiseer word saam met die NZ-1 isolaat en ander isolate uit Kalifornië, as groep VI van GLRaV-3. Hierdie studie het ook verdere bewyse verskaf dat volgende-generasie volgordebepalingstegnologie ’n waardevolle benadering is om nuwe virusse en variante te identifiseer, deurdat die huidige studie gewys het dat die voorlopige volgorde, wat gegenereer is deur bioinformatika-instrumente, 98% identies was aan die GH11 volgorde wat met Sanger volgordebepaling verkry was. Hierdie studie het ook gevind dat die industrie-standaard ELISA, nog steeds ’n effektiewe GLRaV-3 diagnostiese metode is en wel infeksies, veroorsaak deur al die variant-groepe, sal kan identifiseer. Die ELISA toets is egter nie in staat om te onderskei tussen GLRaV-3 variant-groepe nie. In hierdie studie is ’n variant-identifiseerbare in-tyd tru-transkripsie polimerase ketting reaksie (PKR) ontwerp wat GLRaV-3 variant-groepe I, II, III en VI kan identifiseer deur middel van ’n enkele inleier-stel wat die GLRaV-3 Hsp70h-geen teiken. As hoë-resolusie smeltingskurwe-analise bygevoeg word by die in-tyd tru-transkripsie PKR, is dit moontlik om te onderskei tussen variant-groepe op grond van drie smeltingspunt intervalle. Die tru-transkripsie hoë-resolusie smeltingskurwe-toets verskaf meer sensitiewe en geoutomatiseerde metodes om GLRaV-3 variant-groepe te identifiseer en te onderskei. ’n Veelvuldige tru-transkripsie PKR is ook ontwerp om tussen variante wat tans in Suid-Afrika aangetref word, te onderskei en te dien as ’n valideringsmetode vir die in-tyd tru-transkripsie hoë-resolusie smeltingskurwe-toets. Sensitiewe en akkurate toetse, soos bogenoemde, is noodsaaklik vir die bestudering van die verspreiding en impak van die verskillende GLRaV-3 variante in wingerd. Hierdie metodes kan gebruik word om kennis ten opsigte van rolblad patogenese te verbreed en om by te dra tot epidemiologiese studies wat ondersoek hoe hierdie variant-groepe versprei, of daar ’n assosiasie bestaan tussen ’n spesifieke variant en siekte-simptome en of daar ’n verskil is in die witluisvektor oordragseffektiwitiet vir elke GLRaV-3 variant.

Acknowledgements

I would like to express my sincere gratitude to the following people and institutions for their various contributions to this study:

- Prof. J.T. Burger, for giving me the opportunity to be a part of an outstanding research group and for his supervision and guidance.
- Dr. H.J. Maree, for excellent leadership, intellectual input, support, guidance and suggestions throughout this study.
- Dr. D. Stephan, for critical reading of manuscripts.
- Dr. A.E.C Jooste, for the opportunity to be part of her research.
- Mr. T. Pepler, for for his assistance in the statistical analysis of data.
- Yolandi Espach, for her contribution to the survey.
- The staff of Vititec, for the use of their facilities and for their assistance.
- The owners of the wine farms, for allowing us to collect samples (Kanonkop and Chavonnes).
- Winetech for research funding.
- The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.
- Dried fruit technical services (DFTS), AgriSeta and Stellenbosch University, for their financial support
- SASEV, for travel funding to attend the 17th meeting of the ICVG in Davis, California USA.
- Everyone from the Vitis laboratory, for the enjoyable working environment.
- Friends and family.

Table of Contents

Chapter 4: Real-time RT-PCR HRM curve analysis and multiplex RT-PCR to detect and differentiate between GLRaV-3 variant groups I, II, III and VI

Chapter 5: Conclusion

List of Abbreviations

µl	microliter
µM	micromolar
aa	amino acid
ARC	Agricultural Research Council
AMV	Avian myeloblastosis virus
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
BYSV	beet yellow stunt virus
BYV	Beet yellow virus
CDD	Conserved Domain Database
cDNA	complementary cDNA
CP	coat protein
DAC	direct antigen-coating
DAS	double antibody sandwich
dCP	divergent coat protein
dNTP	Deoxynucleotide triphosphate
CTAB	cetyltrimethylammonium bromide
CTV	Citris trizteza virus
cv	cultivar
DNA	Deoxyribonucleic acid
DOI	Digital object identifier
dsRNA	double-stranded RNA
EDTA	Ethylenediamine tetra-acetic acid
ELISA	enzyme-linked immunosorbent assay
FRET	fluorescence resonance energy transfer
GDP	Gross Domestic Product
GLD	grapevine leafroll disease
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-5	Grapevine leafroll-associated virus 5
GLRaV-6	Grapevine leafroll-associated virus 6

GLRaV-7	Grapevine leafroll-associated virus 7
GLRaV-8	Grapevine leafroll-associated virus 8
GLRaV-9	Grapevine leafroll-associated virus 9
GLRaV-Pr	Grapevine leafroll-associated virus Pr
GLRaV-De	Grapevine leafroll-associated virus De
GLRaV-Car	Grapevine leafroll-associated virus Car
GSyV-1	Grapevine syrah virus-1
GVE	Grapevine virus E
HRM	high-resolution melting
Hsp70h	heat shock protein 70 homoloque
Hsp90h	heat shock protein 90 homoloque
IC-PCR	immunocapture PCR
ICTV	International Committee on Taxonomy of Viruses
ICVG	International Committee for Study of the Virus and Virus like Diseases of the Grapevine
IF	immunofluorescence
IQR	Interquartile range
kb	kilobases
LChV-2	Little cherry virus 2
LIYV	Lettuce infectious yellows virus
MAb	monoclonal antibody
MAQ	Mapping and Assembly with Quality
MEGA	Molecular Evolutionary Genetics Analysis
mM	millimolar
mRNA	messenger RNA
NCBI	National Center for Biotechnology Information
ng	nanogram
NGS	next-generation sequencing
nm	nanometer
NRF	National Research Foundation
nt	nucleotide
nts	nucleotides
ORF	open reading frame
PAb	polyclonal antibody
PAGE	polyacrylamide gel electrophoresis

PBNSPaV	Plum bark necrosis stem-pitting-associated virus
PCR	polymerase chain reaction
PMWaV-1	Pineapple mealybug wilt-associated virus 1
PMWaV-2	Pineapple mealybug wilt-associated virus 2
PPRI	Plant Protection Research Institute
PVP	Polyvinylpyrrolidone
RdRp	RNA-dependent RNA polymerase
RLM-RACE	RNA-ligase-mediated rapid amplification of cDNA ends
RNA	Ribonucleic acid
RT-PCR	reverse transcription PCR
SASEV	South African Society for Enology and Viticulture
SAWIS	South African Wine Industry Information and Systems
sg	subgenomic
SNP	single nucleotide polymorphism
SPSS	Statistical Package for the Social Sciences
ss	single-stranded
ssDNA	single stranded DNA
SSCP	single-strand conformation polymorphism
TAE	Tris-acetic acid EDTA
Taq	*Thermus aquaticus* DNA polymerase
TAS	triple antibody sandwich
TEM	transmission electron microscopy
Tris-HCL	2-amino-2-(hydroxymethyl)-1,3-propanediol-hydrochloric acid
U	unit
UTR	untranslated region

List of Figures

List of Tables

Chapter 1: Introduction

1.1 General introduction

Grapevine is economically the most important fruit crop in the world. More than 7.9 million hectares are planted in temperate and tropical climatic regions. The world grape production in 2007 could be divided into 65% for wine, 28% for table grapes and 7% for raisins (Bouquet, 2011). In 2011 South Africa was ranked as the 8th largest wine producing country contributing 3.6% to the total world production with 1.7% of the global wine grape surface area (http://www.sawis.co.za/info/annualpublication.php). The wine industry alone contributes over R26 billion to South Africa's annual GDP (http://www.sawis.co.za/info/macro_study2009.php). According to SA Wine Industry Information and Systems (SAWIS), if indirect and induced impacts are included, the wine industry supported more than 275 thousand jobs in 2009 (http://www.ara.co.za/uploads/DTI-convention-brochure%20final.pdf). In 2011, the South African wine industry produced an average of 775 liters of wine per gross ton of grapes from the nine grape producing regions in South Africa (Orange river, Olifants river, Malmesbury, Little Karoo, Paarl, Robertson, Stellenbosch, Worcester, Breedekloof) (http://www.sawis.co.za/info/annualpublication.php).

A great concern since 2005 is that the South African industry has not been able to replace the 5% of vineyards annually lost due to pests and disease-associated pathogens (http://www.sawis.co.za/info/download/Vineyards2012.pdf). Until 2006 the total surface area covered with grapevine annually increased, but since then a decrease of 1 578 hectare (ha) was observed. The 20-year lifespan of grapevines that served as a guideline until now is not attainable, especially with regard to red varieties and a life span of 15 years is considered to be more realistic according to role players in the industry (http://www.sawis.co.za/info/download/Vineyards2012.pdf). The age distribution of South African vineyards in 2011 showed that 49% of red varieties were between 11-15 years old, with 33% of white varieties between 4-10 years old (http://www.sawis.co.za/info/download/Vineyards2012.pdf).

The decline observed in surface area covered with grapevine and the drop in the average age of vineyards, emphasize the need for effective disease strategies. Numerous pests and pathogens like insects, viruses, bacteria and fungi, impact negatively on the global grapevine industry. In 2006, more

than 70 different infectious agents have been identified in grapevine, making it the highest number of pathogens encountered in a single woody species (Martelli and Boudon-Padieu, 2006). Although most pathogens, like insects and fungi, can be controlled by pesticides, managing and controlling virus infections relies greatly on controlling its spread. The intrinsic susceptibility of grapevines and the environments under which they are grown increase their risk for virus infections. One of the most important diseases of grapevine around the world is grapevine leafroll disease (GLD) that has been associated with a number of grapevine leafroll-associated viruses (GLRaVs). Grapevine leafroll-associated virus 3 (GLRaV-3) is considered the main contributing agent of GLD, with detrimental effects on both wine and table grapes. Different molecular variant groups of GLRaV-3 have been identified, but their individual contribution to leafroll disease is unknown. Studying the different GLRaV-3 variants at a molecular level can assist with elucidating GLD etiology. Focusing on epidemiology studies and developing effective detection techniques will aid in managing and controlling the spread of this disease.

1.2 Aims and objectives

The aim of this study was to sequence the new molecular variant of GLRaV-3 identified in the metagenomic sequencing study by Coetzee et al. (2010) and to evaluate and design detection assays associated with GLRaV-3. The following objectives were set out to achieve these aims:

- To identify GLRaV-3 infected grapevine plants, singly infected with only the new molecular variant group of GLRaV-3.
- To determine the complete genome sequence of a South African isolate of the new molecular variant group of GLRaV-3.
- To perform bioinformatic analysis to compare the sequence identity of the new molecular variant to other molecular variant groups of GLRaV-3.
- To evaluate the ELISA kit currently used by the industry for its effectiveness in detecting the new molecular variant group of GLRaV-3.
- To design and optimize a molecular variant identification real-time reverse transcription polymerase chain reaction (RT-PCR) high-resolution melting (HRM) assay that will be able to detect all known GLRaV-3 molecular variants (in collaboration with Dr. A.E.C. Jooste at the Agricultural Research Council (ARC) - Plant Protection Research Institute (PPRI) in Pretoria).

- To design and optimize an end-point RT-PCR protocol to be able to detect all GLRaV-3 variants.
- To determine the prevalence of the new molecular variant group of GLRaV-3.

1.3 Chapter layout

The thesis is divided into five chapters: a general introduction and literature overview, followed by two research chapters and a general conclusion. Each chapter is introduced and referenced separately.

Chapter 1: Introduction
General introduction, aims and objectives of the study and the chapter layout of the thesis is provided. The scientific outputs generated during the study are stated.

Chapter 2: Literature overview
An overview of the literature relating to grapevine leafroll disease, GLRaV-3 and detection techniques used, is given.

Chapter 3: Complete nucleotide sequence of a new strain of grapevine leafroll-associated virus 3.
In this chapter, the sequencing of the South African isolates, GH11 and GH30, of the new molecular variant group is described. Results from the industry standard ELISA evaluation and additional analysis of generated sequences, not included in the original publication, are also included in this chapter.

Chapter 4: Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate grapevine leafroll-associated virus 3 variant groups I, II, III and VI.
In this chapter the design and optimization of the GLRaV-3 molecular variant identification real-time RT-PCR HRM assay is described together with the design and optimization of the multiplex RT-PCR protocol used for the validation of the RT-PCR HRM assay.

Chapter 5: Conclusion
General concluding remarks and future prospects

1.4 Research outputs

The following papers, conference proceedings and conference posters were generated during the study.

1.4.1 Publications

- Bester, R., Maree. H.J. and Burger, J.T. (2012). Complete nucleotide sequence of a new strain of grapevine leafroll-associated virus 3 in South Africa. Archives of Virology. 157, 1815–1819.
  This paper forms the basis of Chapter 3.

- Bester, R., Jooste, A.E.C., Maree. H.J. and Burger, J.T. (2012). Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate grapevine leafroll-associated virus 3 variant groups I, II, III and VI. Virology Journal. 9:219.
  Chapter 4 was accepted for publication in Virology Journal.

1.4.2 Conference proceedings

- Bester, R., Jooste, A.E.C., Maree, H.J. and Burger, J.T. Sequencing of a new strain of GLRaV-3 and the establishment of detection techniques for the variant groups known to be present in South Africa.
  Presentation combining results from Chapter 3 and 4 delivered by R. Bester at the 34th South African Society for Enology and Viticulture (SASEV) Congress, 14-16 November 2012.

- Jooste, A.E.C., Bester, R., Maree, H.J., de Koker, W. and Burger, J.T. A survey of red and white cultivars to test an improved detection technique for grapevine leafroll-associated virus 3 (GLRaV-3) variants identified in South African vineyards.
  The work of Chapter 4 contributed to this study and was delivered by Dr. A.E.C. Jooste at the 17th meeting of the International Committee for Study of the Virus and Virus-like Diseases of the Grapevine (ICVG), Davis, California USA, 7-14 October 2012.

1.4.3 Posters

- Maree, H.J., Coetzee, B., Nel, Y., Bester, R., Freeborough, M-J. & Burger, J.T. Detection of a new strain of grapevine leafroll-associated virus 3 in South African vineyards. Poster presented

at the 32nd South African Society for Enology and Viticulture (SASEV) Congress, Stellenbosch, 18-19 November 2010.
The work of Chapter 3 contributed to this study and was presented by Dr. H.J. Maree.

- Bester, R., Jooste, A.E.C., Maree, H.J., Burger, J.T. Real-time RT-PCR high-resolution melting curve analysis and multiplex RT-PCR to detect and differentiate between grapevine leafroll-associated virus 3 variant groups.
  Poster summarizing Chapter 4 was presented by R. Bester at the 17th meeting of the International Committee for Study of the Virus and Virus-like Diseases of the Grapevine (ICVG), Davis, California USA, 7-14 October 2012.

1.5 References

Bouquet, A., 2011. Grapevines and viticulture, in: Adam-Blondon, A-F., Martinez-Zapater, J.M., Kole, C. (Eds.), Genetics, genomics, and breeding of grapes. Science publishers, Jersey, USA, pp. 1-29.

Martelli, G.P., Boudon-Padieu, E., 2006. Directory of infectious diseases of grapevines and viroses and viruslike diseases of grapevine: Bibliographic report 1998-2004. Opinions Mediterraneennes Serie B: Studies and Research.

1.6 Recommended internet resources

http://www.sawis.co.za/info/annualpublication.php

http://www.sawis.co.za/info/macro_study2009.php

http://www.ara.co.za/uploads/DTI-convention-brochure%20final.pdf

http://www.sawis.co.za/info/download/Vineyards2012.pdf

Chapter 2: Literature review

2.1 Introduction

Grapevine is one of the most widely grown horticultural crops and a highly valuable agricultural commodity. Grapes can be consumed directly (dried and fresh) or the juice, fresh or fermented to produce wine. The production of wine dates back to the beginning of civilization and has followed agricultural development through time. Chance and curiosity played a role in the origin of wine making as the first wine was produced from damaged grapes that spontaneously fermented. The science of winemaking was initiated as far back as 6000 BC based on little or no knowledge of the microbes that drive fermentation or the biology of grapevines (Chambers and Pretorius, 2010) (http://www.wosa.co.za/sa/history_beginning.php). Today, grapes are predominantly used for wine production.

Grapevines are exposed to a variety of pests and pathogens that threaten the development and health of the world's viticultural industry. Viruses, viroids, phloem- and xylem-limited prokaryotes, fungi and insects lead to heavy losses, shorten the productive lifespan of vineyards and reduce the quality and yield (Martelli and Boudon-Padieu, 2006). The importance of the grapevine industry and the magnitude of the problems caused by pathogens, have been the main motivation behind extensive research programs focused at disease prevention. To combat pathogens there is a need for resistant cultivars or control measures to prevent the spread of diseases. Until now, breeding of resistant cultivars against viral infections has not been successful and disease management relies on accurate and early diagnosis of pathogens.

Grapevines are known to be the host to a number of taxonomically diverse viruses that can cause severe damage to the host if mixed infections would occur (Prosser et al., 2007). The long history of cultivation, grafting between different varieties and the introduction of new viruses by means of vectors such as mealybugs, scale insects and nematodes are probably the reasons for the fact that grapevine is the host to the largest number of viruses of any crop plant (Martelli and Boudon-Padieu, 2006; Prosser et al., 2007). In 2006, the International Council for the Study of Viruses and Virus diseases of the Grapevine (ICVG) recognized 58 viruses that infect grapevine (Martelli and Boudon-Padieu, 2006) with the addition of grapevine virus E (GVE) in 2008 (Nakaune et al., 2008) and grapevine syrah virus-

1 (GSyV-1) in 2009 (Al Rwahnih et al., 2009). In 2011, the first DNA virus infecting grapevine was discovered (grapevine vein clearing virus (GVCV)) (Zhang et al., 2011) and in 2012, a novel circular DNA virus (grapevine cabernet franc-associated virus (GCFaV)) was identified in grapevine (Krenz et al., 2012). There are five major grapevine diseases caused by viruses that include leafroll disease, grapevine degeneration and decline, rugose wood complex, graft incompatibility and the fleck disease complex (Martelli and Boudon-Padieu, 2006; Martelli and Digiaro, 1999). Out of all these diseases, grapevine leafroll disease (GLD) caused by viruses in the family *Closteroviridae*, is considered the most important and widespread disease worldwide including South Africa (Martelli and Boudon-Padieu, 2006; Pietersen, 2004).

2.2 Grapevine leafroll disease

2.2.1 Grapevine leafroll disease distribution

Grapevine leafroll disease was first described more than a century ago and was initially portrayed as a nutrient deficiency. It was not until 1936 when the symptoms were first transmitted through grafting from symptomatic to healthy vines, that GLD could be distinguished from other factors and a pathogen was suspected (Charles et al., 2006a, Charles et al., 2006b; Over de Linden and Chamberlain, 1970). Grapevine leafroll disease has been known by different names around the world, "rougeau", "f1avescence", and "brunisure" in France, "Rollkrankheit" in Germany, and "red-leaf" and "White Emperor disease" in California USA (Over de Linden and Chamberlain, 1970). Grapevine leafroll disease was first identified over a century ago and is present in all grape-growing regions of the world including South Africa, New Zealand, South and North America, Europe, Africa and the Middle East (Akbas et al., 2007; Charles et al., 2006a; Charles et al., 2009; Fiore et al., 2008; Fuchs et al., 2009; Habili et al., 1995; Mahfoudhi et al., 2008; Maliogka et al., 2008; Maree et al., 2008; Sharma et al., 2011). The wide international distribution of the disease can probably be ascribed to the movement of infected plant material from one country to another. Grapevine leafroll disease is present in all grape producing regions of South Africa and is a serious threat to the grapevine industry.

2.2.2 Viruses associated with grapevine leafroll disease

Grapevine leafroll disease was first associated with closterovirus-like particles in the mid-1980s (Charles et al., 2006b). To date 11 different viruses have been found in leafroll-infected vines and are all members of the family *Closteroviridae.* Grapevine leafroll-associated viruses 1, 3, 4, 5, 6, 9 (GLRaV-1, -3, -4, -5, -6, -9) and the recently characterized GLRaV-Pr, GLRaV-De and GLRaV-Car,

are members of the genus A*mpelovirus* (Abou Ghanem-Sabanadzovic et al., 2010; Maliogka et al., 2008; Maliogka et al., 2009; Martelli et al., 2012; Martelli et al., 2002; Martelli and Boudon-Padieu, 2006). Grapevine leafroll-associated virus 2 (GLRaV-2) is considered the type species of the genus C*losterovirus* and GLRaV-7 has not yet been assigned to a genus (Abou Ghanem-Sabanadzovic et al., 2010; Martelli et al., 2002, Martelli and Boudon-Padieu, 2006). Even though a number of viruses are associated with GLD, GLRaV-3 is considered the main causative agent in South Africa (Pietersen, 2004).

2.2.3 Grapevine leafroll disease symptoms and impact

The existence of GLD and its effect on vine performance have been investigated for many years. Many studies have shown that this disease does not only adversely affect vine growth but it also has a detrimental effect on grape yield and juice quality. The disease symptoms are distinctive in red cultivars during autumn, but are less obvious in white cultivars (Pietersen, 2004). The early visual signs of GLD are the appearance of red and reddish-purple discolorations in the interveinal areas of mature leaves while the midrib and main veins remain green. In advanced stages the margins of infected leaves roll downward (Figure 2.1). In some white cultivars like Chardonnay, leaves may show yellowing or chlorotic mottling and in some cases the leaf margins will roll downwards, but in other white cultivars there might be no visual signs of infection (Figure 2.1) (Freeborough and Burger, 2008; Pietersen, 2004; Rayapati et al., 2008).

Grapevine leafroll viruses are phloem-limited and degeneration of phloem cells in leaves, stems and petioles have been reported. This degeneration obstructs carbohydrate export and results in starch accumulation that negatively impacts on the photosynthetic activity of the plant (Cabaleiro et al., 1999; Charles et al., 2006a; Charles et al., 2006b; Mannini et al., 1996). Many studies also showed that fruits from GLD vines had lower sugar levels due to a reduced capacity to accumulate sugars, delayed fruit maturity and fruit coloration (Cabaleiro et al., 1999; Charles et al., 2006a; Charles et al., 2006b; Mannini et al., 2011; Over de Linden and Chamberlain, 1970; Pietersen, 2004; Vega et al., 2011). Leafroll-infected vines were also shown to have higher levels of titratable acid and lower levels of anthocyanin (Charles et al., 2006a; Charles et al., 2006b; Rayapati et al., 2008; Vega et al., 2011). Grapevine leafroll disease agents, in conjunction with other viruses, can also cause graft incompatibility and young vine failure (Golino et al., 2002).

Figure 2.1. Grapevine leafroll disease (GLD) visual symptoms. A. *Vitis vinifera* cv. Cabernet Sauvignon vineyard with a GLD cluster of infected vines. B. Symptom expression in Cabernet Sauvignon (Photo: H.J. Maree). C. Symptom expression in Chardonnay (Photo: H.J. Maree).

Grapevine leafroll disease symptoms may vary within and among vineyards due to several factors regarding consistency in cultivar, age of vineyard, rootstock, period of virus infection, virus strain, climate and region of production, training and pruning systems (Cabaleiro et al., 1999; Freeborough and Burger, 2008; Mannini et al., 2011; Rayapati et al., 2008). The simultaneous comparison of these parameters makes final conclusions regarding the economic impact of GLD on grapevines difficult, but available studies have reported significant yield losses (Atallah et al., 2011; Charles et al., 2006a; Freeborough and Burger, 2008; Rayapati et al., 2008). In Washington State, estimated yield reductions attributed to GLD typically range between 5 and 10% (Rayapati et al., 2008). A recent study estimated

the economic impact of GLD ranging from $25,000 to $40,000 per hectare for yield reduction and quality penalty over a period of 25 years if left uncontrolled (Atallah et al, 2011). Grapevine leafroll disease does not only affect the yield of grapes produced but also the quality, reducing the ability of producers to maintain a competitive advantage in domestic and international wine markets.

2.2.4 Transmission

Grapevine leafroll disease is mainly transmitted by grafting or the vegetative propagation of infected plant material, but GLD transmission at a single location (within-row or vine-to-vine) is mediated by mealybug insect vectors (Charles et al., 2006b; Martelli and Boudon-Padieu, 2006). Mealybugs are small, phloem-feeding insects that transmit the viruses to grapevines in a semi-persistent manner (Charles et al., 2006b; Martelli and Boudon-Padieu, 2006; Tsai et al., 2008). The virus is usually acquired within 15 min-12 hours and the mealybugs can retain the virus for 12 hours to 5 days (Charles et al., 2006b; Tsai et al., 2008). First instars were shown to be more efficient vectors than adult mealybugs, but the virus cannot be transmitted trans-ovarially (Tsai et al., 2008). Data collected from a study in 2009 showed that GLD spread more rapidly in young Chardonnay vines than in young Merlot vines indicating that Chardonnay vines are more susceptible to the mealybug vector *Pseudococcus longispinus* and infections will spread faster in these vines (Charles et al., 2009).

The genus *Ampelovirus* contains viruses that only infect dicotyledonous hosts and none of the ampelovirus members are transmissible by sap inoculation (Martelli et al., 2002). Ampeloviruses can be transmitted by coccid (*Parthenolecanium, Pulvinaria, Neopulvinaria*) or pseudococcid (*Pseudococcus, Planococcus, Saccharicoccus, Dysmiococcus, Phenacoccus, Heliococcus*) mealybugs (Martelli et al., 2002). Mealybug vectors of GLRaV-3 specifically, were found to be *Planococcus ficus* (Figure 2.2), *Planococcus citri, Pseudococcus longispinus* (Figure 2.2), *Pseudococcus calceolariae, Pseudococcus maritimus, Pseudococcus affinis, Pseudococcus viburni* and *Pseudococcus comstocki,* while its other soft scale insect vectors were identified to be *Pulvinaria vitis* and *Neopulvinaria innumerabilis* (Charles et al., 2009; Fuchs et al., 2009; Martelli et al., 2002; Martelli and Boudon-Padieu, 2006; Petersen and Charles, 1997; Sforza et al., 2003; Tsai et al., 2008).

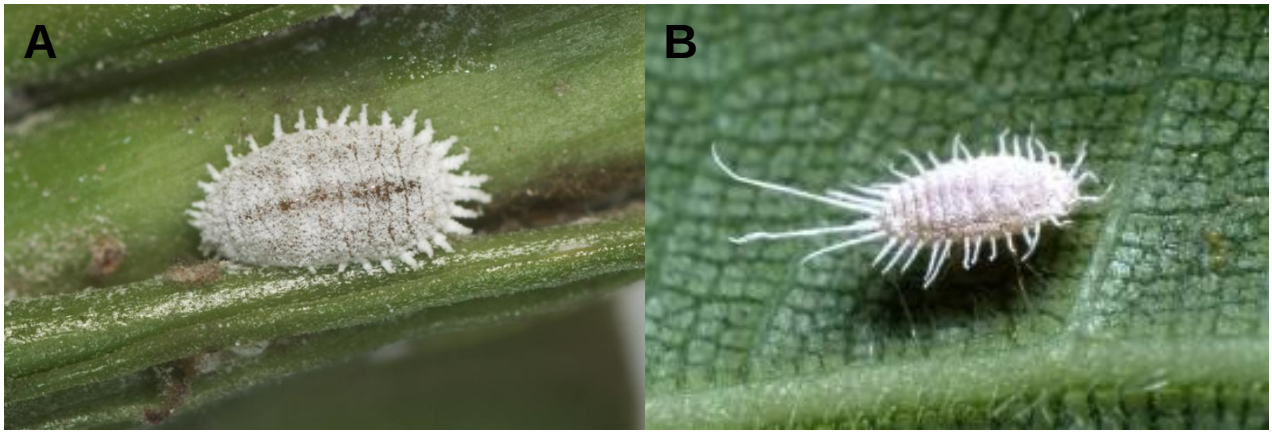

Figure 2.2. Photographs of grapevine mealybugs. A. *Planococcus ficus* (http://www.udec.ru/vrediteli/chervec.php) B. *Pseudococcus longispinus* (http://www.bcpcertis.com/Certis.bcp/English/Home/Our+Solutions/Biological+Controls/Pests+Find+out+more/Mealybug+find+out+more/page.aspx/1091).

2.2.5 Disease management

The spread of grapevine diseases throughout the wine growing regions of the world can mostly be attributed to the exchange and sale of virus-infected scion and rootstock material. As a result, many countries introduced certification programs to produce virus-free clonal plant material for grapevine growers. To date no natural resistance to GLRaV-3 in *Vitis vinifera* has been demonstrated and even though transgenic approaches are currently being assessed worldwide, they are unlikely to be available for field use in the near future (Charles et al., 2006a; Martelli and Boudon-Padieu, 2006; Pietersen, 2004). There are no agro-chemicals available that can selectively control virus infection in a plant, however heat therapy, meristem tip culture or somatic embryogenesis have been shown to eliminate viruses from infected plant material (Charles et al., 2006a; Pietersen, 2004). Grapevine leafroll disease is best managed by prevention, but if the disease is already present in vineyards a vector control approach has to be implemented. Mealybugs are the most important vectors of leafroll viruses, and therefore effective mealybug control is a requirement for slowing the spread of GLD. Various disease management strategies have been described, specifically in South Africa (Pietersen, 2004). Removal of infected vines (rogueing) is currently the best means of control to combat secondary spread. Atallah et al. (2011) found that early vine rogueing reduced the disease impact six-fold for Sauvignon blanc and seven-fold for Merlot when compared to a "no intervention" scenario. Rogueing and replanting of individual vines with virus-free material proved to be a more practical solution taking into account the economic implications of removing a whole vineyard (Rayapati et al., 2008). Thorough removal of old, infected vines and their roots and then lengthening the period that the site lies fallow between plantings, will also contribute to effective disease management (Bell et al., 2009; Charles et al., 2006a). Studies showed that remnant vine roots can host the virus for at least 12 months and it is therefore

likely that mealybugs will transfer the virus to healthy vines in newly replanted vineyards (Bell et al., 2009). Supplementing rogueing by applying systemic insecticides will also reduce the spread of leafroll disease by killing virus-infected mealybugs (Rayapati et al., 2008). Mealybug monitoring programs, biological control of mealybugs through natural enemies, managing ant populations and applying weed control can also aid in the control of the disease (Charles et al., 2006b; Pietersen, 2004).

The effective management of GLD depends on reliable accurate detection methods. Due to the long latent phase of new infections, some cultivars being symptomless and the similarity of the disease symptoms to magnesium and phosphorous deficiencies, visual inspection alone cannot be use to diagnose the disease (Constable et al., 2010). Accurate and sensitive detection techniques that are able to detect all viruses and their variants are therefore a necessity to control GLD.

2.3 Grapevine leafroll-associated virus 3

2.3.1 General

Worldwide GLRaV-3 has been closely associated with GLD and is considered the main causative pathogen of GLD in South Africa; so much so that leafroll diagnostics is focused at detecting GLRaV-3 (Charles et al., 2006a; Freeborough and Burger, 2008; Pietersen, 2004). The first closterovirus-like particles were purified from a GLRaV-3 isolate (NY-1) to produce an antiserum for the first indirect enzyme-linked immunosorbent assay (ELISA) (Zee et al., 1987). At the 10th Meeting of the International Council for the Study of Viruses and Virus diseases of the Grapevine (ICVG) in 1990, it was proposed to use the name grapevine leafroll-associated virus followed by roman numerals I to V, but in 1995 the International Committee on Taxonomy of Viruses (ICTV) determined that virus acronyms that have numbers are to be written in arabic numerals, separated by a hyphen from the letters (Boscia et al., 1995).

2.3.2 Taxonomy

Grapevine leafroll-associated virus 3 is the type member of the genus *Ampelovirus* in the family *Closteroviridae*. In 1998 when the family *Closteroviridae* was established, it comprised of only two genera distinguishing between viruses with monopartite and bipartite genomes (Martelli et al., 2000). In 2002 the ICTV revised the classification system to differentiate between viruses based on biological (type of vector) and molecular (conservation of genes) properties (Martelli et al., 2002) and the genus *Ampelovirus* was added to the family *Closteroviridae*. Other viruses belonging to this genus can be seen in Table 2.1. In 2000, a study reported the identification of GLRaV-8, but at the 16th Meeting of

the ICVG in 2009 GLRaV-8 was classified as a cloning artifact and it was concluded that GLRaV-8 does not exist (Dolja, 2009; Monis, 2000). The genus *Ampelovirus* is now split into two subgroups designated as I and II based on the genome size and structure of the present members (Table 2.1). Subgroup I contains viruses that are all serologically related, have the same genome structure and size and are all considered divergent variants of a single species, GLRaV-4 (Table 2.1) (Abou Ghanem-Sabanadzovic et al., 2010; Martelli et al., 2012). The other members of the genus *Ampelovirus* are classified as subgroup II (Table 2.1) (Abou Ghanem-Sabanadzovic et al., 2010; Martelli et al., 2012).

Table 2.1. Members of the genus *Ampelovirus.*

Genus *Ampelovirus*	
Grapevine leafroll-associated virus 4 (GLRaV-4)	Subgroup I
Grapevine leafroll-associated virus 5 (GLRaV-5)	
Grapevine leafroll-associated virus 6 (GLRaV-6)	
Grapevine leafroll-associated virus 9 (GLRaV-9)	
Grapevine leafroll-associated virus Pr (GLRaV-Pr)	
Grapevine leafroll-associated virus De (GLRaV-De)	
Grapevine leafroll-associated Carnelian virus (GLRaV-Car)	
Pineapple mealybug wilt-associated virus 1 (PMWaV-1)	
Pineapple mealybug wilt-associated virus 3 (PMWaV-3)	
Plum bark necrosis stem-pitting-associated virus (PBNSPaV)	
Grapevine leafroll-associated virus 1 (GLRaV-1)	Subgroup II
Grapevine leafroll-associated virus 3 (GLRaV-3)	
Pineapple mealybug wilt-associated virus 2 (PMWaV-2)	
Little cherry virus 2 (LChV-2)	

(Abou Ghanem-Sabanadzovic et al., 2010; Maliogka et al., 2008; Maliogka et al., 2009; Martelli et al., 2002)

2.3.3 Morphology and genome organization

Ampelovirus virions have long, flexuous and filamentous particles of 1400-2200 nm in length with GLRaV-3 particles being between 1800 and 2000 nm (Figure 2.3) (Ling et al., 1997; Martelli et al., 2002; Zee et al., 1987). Viruses belonging to the genus *Ampelovirus,* have a monopartite, linear, positive sense, ssRNA genome ranging in size from 16.9 – 19.5 kb (Martelli et al., 2002). The first complete genome sequence of GLRaV-3 (isolate NY-1) was reported in 2004 to be 17919 nucleotides (nts) in length consisting of 13 open reading frames (ORFs) with a 5’ and 3’ untranslated region (UTR) of 158 and 277 nts, respectively (Ling et al., 1998; Ling et al., 2004).

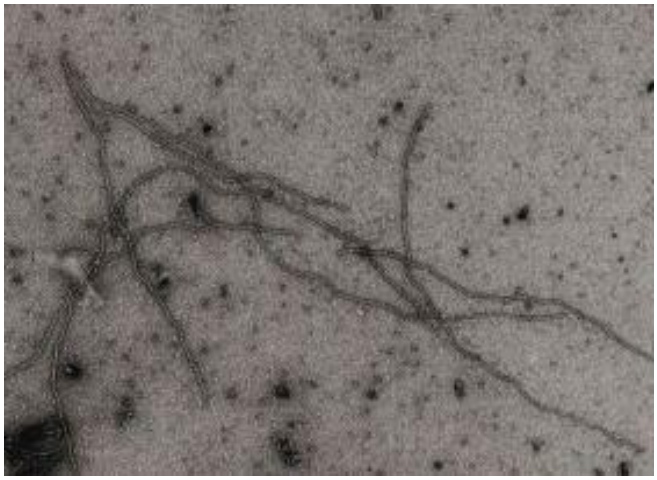

Figure 2.3. Electron micrograph of GLRaV-3 particles isolated from a leafroll-diseased vine (Pietersen, 2004).

In 2008, the genome of a South African isolate (GP18) of GLRaV-3 was sequenced, which showed a longer 5' UTR of 737 nts (Figure 2.4) (Maree et al., 2008). This was the longest 5' UTR reported for any member of the family *Closteroviridae* - compared to the 539 nts of little cherry virus 2 (LChV-2; AF531505), 105 nts of GLRaV-2 (AY881628), 107 nts of beet yellow virus (BYV; AF190581), 107 nts of citrus tristeza virus (CTV; DQ272579) and 97 nts of lettuce infectious yellows virus (LIYV; NC_003617). In 2010, three representative isolates of GLRaV-3 were sequenced, of which two (621 and 623) confirmed the 737 nts long 5' UTR, suggesting that the NY-1 sequence is incomplete (Jooste et al., 2010). To date the function of this longer 5' UTR is unknown and needs to be investigated. However, this region shows high variability between GLRaV-3 variants and is valuable for variation studies of GLRaV-3. The genome organization of GLRaV-3 is consistent with that expected for a typical monopartite closterovirus (Dolja et al., 1994). The conventional closterovirus description of Agranovsky et al (1994) is followed and the ORFs are designated as ORFs 1a, 1b and 2-12 (Figure 2.4).

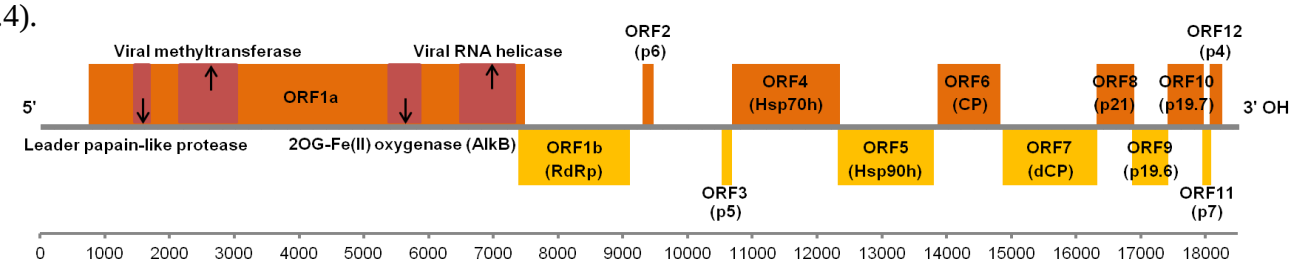

Figure 2.4. Schematic representation of the genome of GLRaV-3 (Isolate GP18) and positions (nt) of open reading frames (ORFs) and untranslated regions (UTRs) adapted from Maree et al., 2008.

The replication-associated proteins are encoded by replication gene domains situated within ORF1a and ORF1b. Open reading frame 1a encodes a large polypeptide with four conserved domains: leader papain-like protease, methyltransferase, AlkB and Helicase (Engel et al., 2008; Ling et al., 1998; Ling

et al., 2004; Maree et al., 2008). The RNA-dependent RNA polymerase (RdRp) is encoded by ORF1b that is predicted to be expressed by a +1 translational frameshift between ORF1a and ORF1b due to the last nts of ORF1a overlapping with ORF1b (Ling et al., 1998, Maree et al., 2008). The other genes are encoded in 11 ORFs expressed through a set of nested 3' co-termininal subgenomic (sg) mRNAs (Maree et al., 2010; Jarugula et al., 2010). Open reading frame 2 potentially encodes a small peptide (p6), but no counterpart ORFs have been found in other members of the family *Closteroviridae,* except for CTV, LIYV and beet yellow stunt virus (BYSV) where much larger unrelated ORFs were found at the same position (Karasev et al., 1995; Karasev et al., 1996, Klaasen et al., 1995). The validity of ORF2 is discussed further in chapter 3. Open reading frame 3 follows a long intergenic region and potentially codes for a small hydrophobic transmembrane protein (p5) (Ling et al., 1998). Sequence analysis of ORF4 (p59) indicated similarities to the heat shock protein 70 (Hsp70) family of chaperones with eight conserved motifs amongst members of the family *Closteroviridae* (Ling et al., 1998). The deduced protein of ORF5 (p55) showed limited sequence similarity (11.2-12.2%) to other members of the family *Closteroviridae* and two conserved regions previously defined in the heat shock protein 90 (Hsp90) homoloque of BYV and CTV could not be identified (Ling et al., 1998). Open reading frames 6 and 7 encode a coat protein (CP) and divergent coat protein (dCP, p53) respectively, based on conserved amino acids (N, R, G, D) identified in all closteroviruses (Ling 1997, 1998). Open reading frames 3-7 are thought to be involved in cell-to-cell movement of the virus (Dolja et al., 2006). No obvious similarities between ORF8 (p21), ORF9 (p19.6) and ORF10 (p19.7) and those of other closteroviruses could be found, other than the similar genome organization in terms of size and position (Ling et al., 1998). These ORFs are believed to be involved in systemic transport and suppressing RNA silencing (Dolja et al., 2006). Open reading frames 11 (p4) and 12 (p7) are unique to GLRaV-3 and are not found in other members of the family *Closteroviridae* (Ling et al., 1998). Chapter 3 elaborates on ORF11 and ORF12 of GLRaV-3 isolates GH11 and GH30.

2.3.4 Genome variability

To date, nine complete genome sequences of representative isolates of GLRaV-3 are available of which six are from South Africa (two from this study, Chapter 3). The complete genome of the Chilean isolate Cl-766 of GLRaV-3 was determined to be 17919 nts confirming the NY-1 sequence. Isolate Cl-766 showed 97.9% nucleotide (nt) identity to isolate NY-1, with most of the genetic diversity concentrated in ORF1a (Engel et al., 2008, Ling et al., 2004). The complete genome of GLRaV-3 isolate GP18 was 18498 nts long with a 5' UTR of 737 nts (Maree et al., 2008). This was the first report of the longer 5'

UTR and was confirmed by isolate 621, 623, WA-MR, GH11 and GH30 (Bester et al, 2012; Jarugula et al., 2010; Jooste et al., 2010). Isolate GP18 has a nt sequence similarity to isolate NY-1 of 93% over nts 580-18498 (Maree et al., 2008). Single-strand conformation polymorphism (SSCP) profiles, from a region amplified in ORF5, provided evidence that South African isolates of GLRaV-3 could be divided into three variant groups. As a result, isolates 621 (group I), 623 (group II) and PL-20 (group III) were sequenced as representative isolates of these groups. Isolate PL-20 was determined to be 18433 nts in length with a 5' UTR of 672 nts (Jooste et al., 2010). Variation between the group I (621) and group II (623) variants in the 5' UTR was 30%, while variant group II (623) and group III (PL-20) differed by 22%, and group I (621) and III (PL-20) by 33% (Jooste et al., 2010). Phylogenetic analysis based on the Hsp70h and CP sequences of GLRaV-3 confirmed the existence of the three molecular variant groups. In Portugal, two additional groups were identified based also on CP sequences (Gouveia et al., 2011). The most divergent isolates were first identified in New Zealand (isolate NZ-1) and then in the Napa Valley, California USA. The existence of these variants from other regions of the world and results from the studies by Jooste et al. (2010) and Maree (2010) suggested that more than three GLRaV-3 variants had to present in South African vineyards. The detailed description of isolates GH11 and GH30, and their comparison to other variants, is described in Chapter 3.

2.3.5 Virus replication and expression of ORFs

The mechanism by which GLRaV-3 replicates has not been studied specifically, but it is assumed that it follows a similar strategy to other closteroviruses like CTV and BYV, which have been studied comprehensively (Dolja et al., 2006; Maree, 2010). Most closteroviruses are transmitted to plants by insect vectors and therefore the replication cycle requires virion genome un-coating, followed by the translation of the viral genome (Dolja et al., 2006). The replication of positive stranded RNA viruses can be separated into four stages (Bustamante and Hull, 1998; Dolja et al., 2006): The uncoating of the virus, to expose the nucleic acid to the replication/translation machinery, translation to produce structural and non-structural proteins, replication of the genome, and the encapsidation of the progeny genomic strands. The translation process can be divided into primary translation of the replication proteins and secondary translation of the remaining proteins with later functions. During replication, which is catalyzed by the RdRp, a complementary negative RNA strand is synthesized that is used as the template for the synthesis of progeny genomic RNA and subgenomic RNAs (sgRNAs). After the encapsidation of the progeny genomic strands the virions can spread through the plant. Single stranded RNA viruses make use of sgRNAs to facilitate the expression of genes located at the 3' end of the

genome and the number of sgRNAs is approximately the same as the number of ORFs (Dolja et al., 2006; Maree et al., 2010). These sgRNAs are usually functionally monocistronic and each sgRNA will only express one protein from the 5'-proximal ORF (Dolja et al., 2006). The timing and quantity of sgRNA synthesis are regulated to optimize consecutive events, and therefore the sgRNAs that encode RNA silencing suppressors are produced during the early stages of infection to counteract a potent host defense system (Dolja et al., 2006). Genome translation will result in proliferation of replication complexes and the accumulation of structural proteins will in turn trigger virion assembly (Dolja et al., 2006). The study by Maree et al. (2010) predicted the existence of seven 3' co-terminal positive sense sgRNAs for the expression of ORFs 3-12 suggesting that ORF10-12 are possibly translated from the same sgRNA. The study by Jarugula et al (2010) confirmed the presence of sgRNAs in grapevine naturally infected with GLRaV-3. They found that four of the eleven putative 3' co-terminal sgRNAs, specific to ORF6, 8, 9 and 10, were present in higher levels. These results suggest that 3′ co-terminal sgRNAs accumulate at variable amounts, demonstrating their expression level differences in infected grapevine tissues.

2.4 Virus diagnostic techniques

Preventing and managing the spread of diseases caused by viruses relies mainly on planting virus-free material. In order to certify material as virus-free, sensitive, reliable and rapid identification of viruses is essential. The identification of disease-associated viruses has proven to be very difficult due to most diseased grapevines being infected with more than one virus. These different viruses can cause similar symptoms in a diseased vine and the low virus titer associated with new infections, the uneven virus distribution and the less noticeable symptoms in some white cultivars, complicate the identification even further. Several studies have investigated the ability to detect GLRaV-3 in different tissues collected from different time points over the course of a season (Ling et al., 2001; Monis et al., 1996; Teliz et al., 1987). These studies found that a high titer of GLRaV-3 is present in phloem tissue of the canes and trunk. With the onset of the growth season, the virus moves into the expanding shoot and can be detected in a range of tissues soon after budding. The accumulation of the virus in leaves occurs through the season and even though the virus was detectable in the leaf vein tissue collected late in the season, bark scrapings were found to be the most reliable source (Ling et al., 2001). To date several techniques have been applied to detect viruses in plant material including biological indexing, electron microscopy, serology, nucleic acid-based methods and next-generation sequencing.

2.4.1 Biological indexing

Before molecular detection techniques became accessible to most laboratories, biological indexing was used for grapevine virus disease detection (Nicholas, 2006, Shanmuganathan and Fletcher, 1980; Weber et al., 2002). Biological indexing takes advantage of the plant's response to the presence of pathogens to indicate a virus infection (Constable et al., 2010). Standardized plants are used and inoculated with material from the test vine and are observed for characteristic symptoms. Two biological indexing methods are used for the detection of grapevine viruses: herbaceous indexing by sap inoculation and hardwood indexing that requires the grafting of buds from the vine being tested onto indicator grapevine varieties (Constable et al., 2010). Herbaceous indexing is performed in a greenhouse and involves rubbing extract from the test vine onto leaves of sensitive indicator plants (Weber et al., 2002). If viruses were present in the test plant, diagnostic symptoms will develop in several weeks. Grapevine leafroll-associated virus 3 is not sap-transmissible and requires hardwood indexing through grafting. The indicator plants with the grafted material are planted in the field and observed for at least two seasons for the development of virus disease symptoms (Weber et al., 2002). Woody indicators should be free of viruses and phytoplasmas and should also be geographically isolated from other grapevines to ensure that field infection does not occur and lead to false positive results (Constable et al., 2010). Biological indexing onto woody indicators is labor intensive, time-consuming and dependent on the successful inoculation of associated viruses (Weber et al., 2002). Inoculation of viruses is affected by graft take and the presence of viruses in the grafted chip buds. Uneven distribution of the virus, strain variation within the associated virus species, low virus titer and the lack of symptom expression can also affect the results observed with indexing (Constable et al., 2010, Rowhani et al., 1997). It was shown by Cirami et al (1988) that some viruses may not induce obvious symptoms on the selected biological indictors and that more than one grapevine variety should be used as indicator plants. Biological indexing detects the disease rather than the associated virus and therefore expression is non-specific and could indicate that the vine is stressed or infected with a virus other than the one associated with the disease (Monis et al., 1996). Although biological indexing can be a successful detection method, it requires highly skilled virologist for disease confirmation and remains a subjective observation.

2.4.2 Electron microscopy

The use of electron microscopy to identify virus particles in plants dates back to 1939 when the first commercial electron microscope became available (Corbett, 1974). The first electron micrographs of a

virus were published in 1939 and showed the rod-like nature of the particles of tobacco mosaic virus (TMV) (As referenced by Steere, 1964). Transmission electron microscopy (TEM) is used for the detection of virus types and diseases since size and ultrastructural features are specific for each group of viruses (Zechmann and Zellnig, 2009). The early use of TEM in plant virology involved morphological characterization of viral particles in crude or purified suspensions (Kitajima, 2004). Negative staining of viruses and following visualization by TEM can provide rapid and accurate results, but in some cases where ultrastructural changes induced by the virus are important for identification, it is also necessary to prepare tissue samples for ultrastructural investigation (Zechmann and Zellnig, 2009). This can be time consuming and labor intensive and may take several days to obtain the results. The efficiency of virus visualization can be improved by combining the specificity of serological assays with the visualization capabilities of electron microscopy (Naidua and Hughes, 2001). Immunosorbent electron microscopy can be used to localize several types of non-structural properties or viral proteins such as polymerases, helicases and movement proteins (Kitajima, 2004). Immunogold labeling can be used for visualization by using gold particles that are attached to antibodies (Alberts et al., 2002; Kitajima, 2004). Gold is used for its high electron density which increases electron scatter to give a high contrast (Alberts et al., 2002). Electron microscope observations of diseased tissues led to the discovery of many associations between specific virus-like particles and a disease. Electron microscopy also contributed to the isolation of the first GLRaV-3 particles in 1987 (Zee et al., 1987). Currently, the major disadvantage of electron microscopy is the need for skilled electron microscopists that is declining rapidly (Kitajima, 2004).

2.4.3 Serology

Serological assays were originally developed to detect viruses by utilizing antibodies to detect epitopes of protein antigens. The different formats for immunological diagnostic techniques include enzyme-linked immunosorbent assay (ELISA), immunofluorescence (IF) and immuno-strip tests (Schaad et al., 2003). Enzyme-linked immunosorbent assays are by far the most commonly used immunodiagnostic technique for virus detection since the 1970s (Clark and Adams, 1977; Engvall and Perlmann, 1971; Van Weemen and Schuurs, I971). Variations on this technique exist that differ from each other in the way the antigen-antibody complex is detected but the underlying mechanism is the same (Figure 2.5).

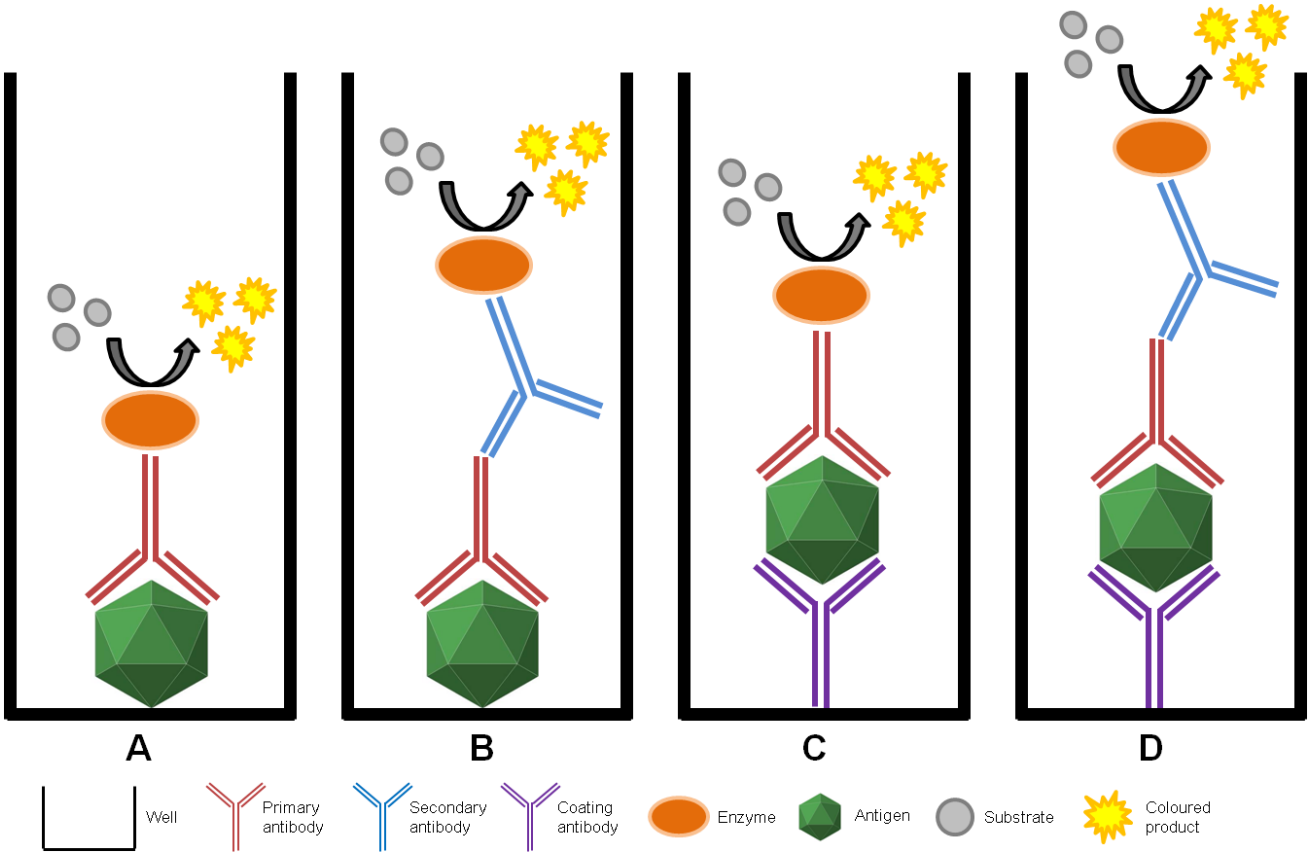

Figure 2.5. Four types of enzyme-linked immunosorbent assay (ELISA) commonly used for virus detection. **A. Direct antigen-coating (DAC) ELISA**. The microtiter plate wells are directly coated with the test sample and then incubated with a primary antibody, which binds to the target antigen if the virus is present in the test sample. The primary antibody is conjugated to an enzyme (e.g. alkaline phosphatase) that converts the added substrate (e.g. p-nitrophenyl phosphate) leading to the development of a color change. **B. Indirect ELISA**. Similar to DAC, but the primary antibody is detected via a secondary antibody that is conjugated to the enzyme. **C. Double antibody sandwich ELISA (DAS-ELISA)**. Similar to DAC, except coating antibodies are used to coat the microtiter plate, which then trap the target antigen within the test sample. **D. Triple antibody sandwich (TAS) ELISA**. Similar to DAS-ELISA, except that before adding the detecting antibody-enzyme conjugate, a monoclonal antibody is added (Constructed and adapted from Clark and Adams, 1977; Naidua and Hughes, 2001; O'Donnell, 1999; Ward et al., 2004).

Antibodies used in serological assays can either be polyclonal or monoclonal. Polyclonal antibodies are a heterogeneous mixture of antibodies directed towards different epitopes of the target protein (Naidua and Hughes, 2001). Polyclonal antibodies can be generated much more rapidly, at less expense, and with less technical skill than is required to produce monoclonal antibodies (Lipman et al., 2005). Monoclonal antibodies are produced from a clonal population of cells derived from a single hybridoma cell line and therefore, each monoclonal antibody has the same specificity and affinity for a single epitope (Lipman et al., 2005; Naidua and Hughes, 2001; Ward et al., 2004). Enzyme-linked immunosorbent assay is a robust, simple and cost effective detection method that can be used to process many samples at the same time (O'Donnell, 1999; Ward et al., 2004). The efficiency of the ELISA is independent of the ratio of antibody to antigen. Therefore detection will occur at all virus concentrations. The technique has high quantitative potential due to the retention of enzyme-labeled

antibody being proportional to the virus concentration (Clark and Adams, 1977). The disadvantages of the technique are that it has a high developmental cost and is not as sensitive as nucleic acid-based methods (O'Donnell, 1999). Enzyme-linked immunosorbent assays have been used routinely around the world for the detection of GLD-associated viruses. The first ELISA test used to screen field samples for this disease was developed using the polyclonal antiserum produced against closterovirus-like particles in 1987 (Teliz, 1987; Zee et al., 1987). Another polyclonal antiserum specific to grapevine leafroll-associated virus 3 (GLRaV-3) was developed using a recombinant coat protein expressed in *E. coli* (Ling et al., 2000; Ling et al., 2001). This antiserum was used to develop an effective double antibody sandwich ELISA (DAS-ELISA) for GLRaV-3 detection. In South Africa ELISA is the most popular serological test routinely used by the industry for the detection of grapevine viruses. An indirect ELISA for the specific or simultaneous detection of GLRaV-1, GLRaV-2 and GLRaV-3 in grapevines was developed for the South African industry by the Plant Protection Research Institute (PPRI) at the Agricultural Research Council (ARC).

2.4.4 Nucleic acid-based methods

Nucleic acid-based methods have increasingly been used in recent years to develop diagnostic assays for plant pathogens. These methods have the potential to be very sensitive and highly specific and are based on the unique nucleic acid sequence of the pathogen (O'Donnell, 1999; Ward et al., 2004). Inexpensive and effective nucleic acid extraction methods have already been described, including total RNA, double-stranded RNA (dsRNA) and DNA extractions from plant material (White et al., 2008; Valverde et al., 1990). Nucleic acid-based assays depend on the availability of sequence data and for many pathogens sequencing data is already available in public databases. Unique DNA or RNA diagnostic regions in this sequence data can be used to design primers or probes for detection through nucleic acid hybridization. Nucleic acid-based detection techniques have the advantage that any region of the pathogen genome can be targeted to develop the diagnostic test compared to serological tests where only specific proteins like the coat protein are targeted (Naidua and Hughes, 2001).

The most commonly used method in nucleic acid-based diagnostics is the polymerase chain reaction (PCR), which is based on the exponential amplification of pathogen DNA, using pathogen-specific DNA primers (O'Donnell, 1999). The introduction of PCR has provided a new and reliable approach for the detection of grapevine viruses because of its high sensitivity, specificity, and speed (Osman et al., 2008). The PCR is estimated to be 100–1000 times more sensitive than ELISA, but has a lower

throughput potential (Charles et al., 2006a; Osman et al., 2008,). Since most plant viruses have RNA genomes, reverse transcription PCR (RT-PCR) is used to detect these plant viruses (Ward et al., 2004).

One of the variations on RT-PCR, is multiplex RT-PCR (Dai et al., 2012; Dovas and Katis, 2003; Elnifro et al., 2000; Osman et al., 2007; Ward et al., 2004). Multiple primer sets directed at different pathogens or variants of a single pathogen can be included in a single reaction. However, the design of compatible sets of primers and detecting related viral sequences present in the same sample can be challenging. In chapter 4 the development of a multiplex RT-PCR, to detect different variant groups of GLRaV-3, is described. Nested PCR is another variation used to improve the sensitivity and/or specificity of the assay (Dovas and Katis, 2003; Ward et al., 2004). Nested PCR involves two consecutive PCR reactions of which the second reaction uses a primer set that recognize a region within the PCR product amplified by the first set of primers (Ward et al., 2004). Nested PCR has been widely used for the detection of phytoplasmas since the concentration of phytoplasmas can be very low in grapevine and the presence of compounds inhibiting the PCR can be high (Angelini et al., 2007). Nested PCR can be time-consuming, labor-intensive and the risk for cross contamination higher (Angelini et al., 2007). Another approach to PCR is to use immunocapture PCR (IC-PCR). Antibodies are immobilized on the surface of a microtiter plate or microfuge tube and used to bind the pathogen, which is then detected using PCR (Ward et al., 2004). Spot-PCR is another rapid procedure whereby a small drop of unbuffered sap from grapevine leaf petioles is placed on filter paper which can be used as the template for PCR. This procedure has successfully been applied for the detection of pathogens in woody plants (Dovas and Katis, 2003; La Notte et al., 1997; Osman and Rowhani, 2006).

Single-strand conformation polymorphism is an added extension to PCR that is based on the principle that partially denatured double-stranded DNA (dsDNA) migrates as two single-stranded DNA (ssDNA) bands in non-denaturing polyacrylamide gel electrophoresis (PAGE). The migration of the two strands depends on the sequence of nucleotides and their conformation under the electrophoresis conditions chosen (Palacio and Duran-Vila, 1999). Since it was first described by Orita et al. (1989), the analysis of SSCP has proven to be a simple and reliable method of detecting sequence differences between DNA fragments (Beier, 1993). This technique can be used for rapid identification of divergent molecular variants of a virus and has been used to investigate the population structure and genetic variability of GLRaV-3 (Jooste and Goszczynski, 2005; Jooste et al., 2010; Turturo et al., 2005).

The quantification of target DNA has been simplified with the introduction of real-time PCR where unknown samples are quantified absolutely or relatively by comparing it to a standard DNA sample or to a reference gene (Feng et al., 2008). This method requires no post-reaction processing since the amplified product is detected by a built-in fluorometer as it accumulates. Target DNA amplification is detected by using non-specific DNA binding dyes (e.g. SYBR Green) or specific fluorescent probes (Ward et al., 2004). DNA-binding dyes binds non-specifically to double-stranded DNA (dsDNA) and exhibits little fluorescence when it is free in solution, but high fluorescence when it binds dsDNA. DNA binding dyes are more prone to error since any non-specific PCR products, and primer-dimer artifacts can generate a signal, but specific fluorescent probes are more expensive and the design more complicated (Ward et al., 2004). Probes are single-stranded DNA or RNA molecules that are labeled with a reporter molecule such as a radioactive isotope, an enzyme or a fluorescent dye that can hybridize to complementary DNA/RNA sequences of the target (Ward et al., 2004).

Probe-based real-time PCR is very sensitive chemistry that utilizes fluorescence quenching to ensure that specific fluorescence is detected only when the product of interest is amplified. Different fluorescent probe-based chemistries have been develop, including: TaqMan® probes, molecular Beacons, hybridization probes (FRET-based) and Eclipse® probes (Figure 2.6). Real-time PCR probes offer two main advantages over DNA-binding dyes since they specifically detect the target sequence and no non-specific sequences and secondly, they allow multiplex reactions to be performed. Multiple probes directed at different targets can be added to a single reaction. Each probe will be labelled with a different flourophore that is detected at a specific wavelength unique to each flourophore. The best known probe-based method is TaqMan®. A TaqMan® assay was developed for the simultaneous detection of GLRaV-1-5 and -9. This study showed that TaqMan® RT-PCR was more sensitive than conventional one-step RT-PCR (Osman et al., 2007). TaqMan® low-density arrays have recently been introduced as a modified method of real-time TaqMan® PCR. This method uses microtiter plates with dried TaqMan® PCR primers/probes complexes added to the wells. This technique was applied to detect 13 different grapevine viruses in infected tissues by adding the cDNA of test samples and PCR master mix to the microtiter plate wells (Osman et al., 2008).

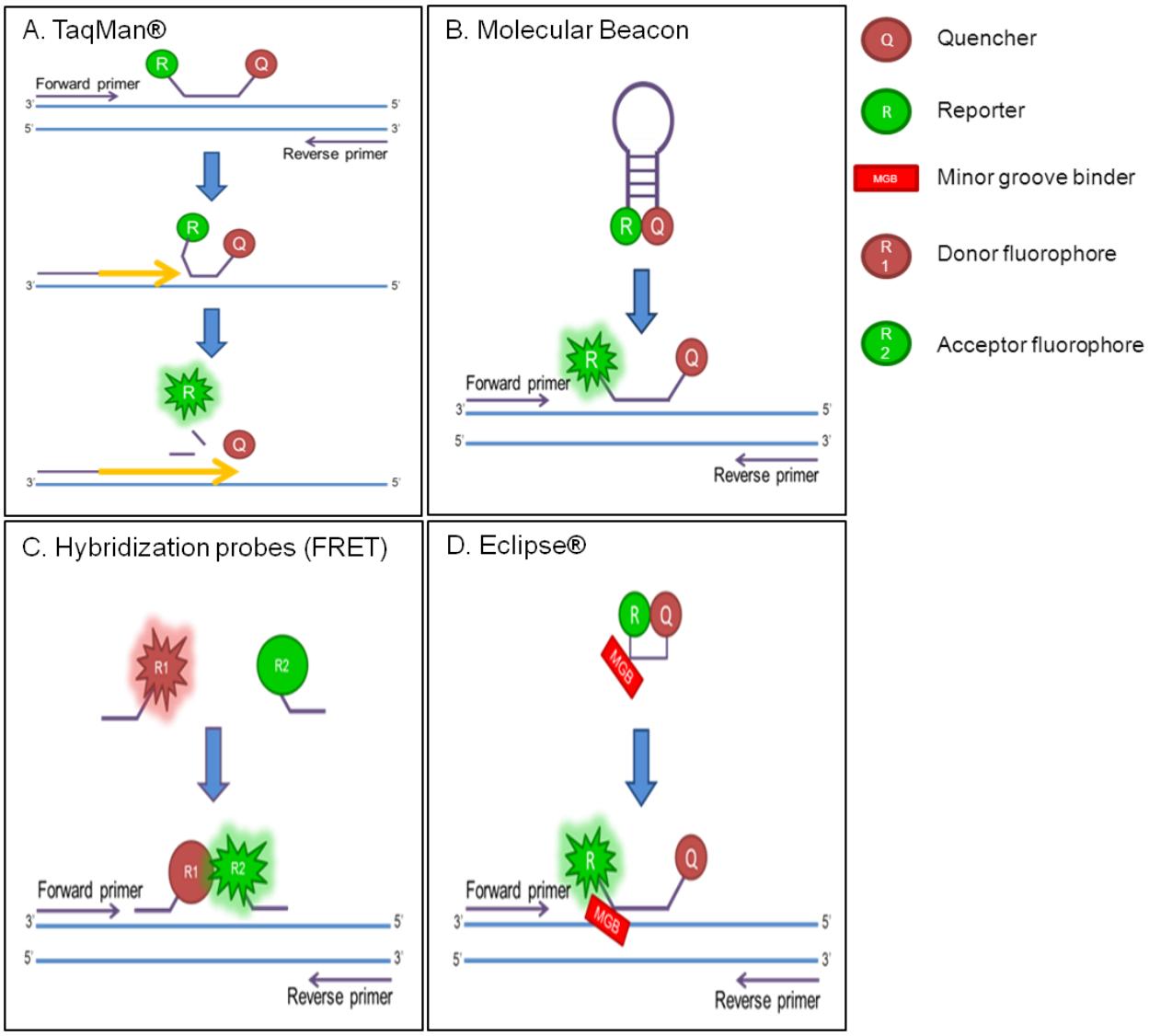

Figure 2.6. Different fluorescent probe-based chemistries. **A. TaqMan®** probes are sequence-specific, labelled oligonucleotides attached to a reporter and quencher molecule. During extension, the probe is partially displaced and the reporter is cleaved by Taq DNA polymerase. **B. Molecular Beacon** probes are dye-labelled oligonucleotides that form a hairpin structure with a fluorescent reporter molecule attached to the 5' end of the probe and a quencher attached to the 3' end. During annealing the probe binds to the target sequence and the reporter molecule is separated from the quencher. **C. Hybridization probes (FRET)** are two probes that are labelled with a pair of dyes that can engage in fluorescence resonance energy transfer (FRET). The donor dye is attached to the 3' end of the first probe, while the acceptor dye is attached to the 5' end of the second probe. During annealing the probes bind to the target in a head-to-tail orientation and the acceptor fluorophore can fluoresce. **D. Eclipse®** probes have a fluorescent reporter molecule at the 3' end and a quencher and a minor groove binder at the 5' end. The unhybridized probe has a random coil conformation that brings the reporter in close proximity to the quencher. During annealing, the probe will hybridize to the target utilizing the minor groove binder. The probe will unfold to become linearized, separating the reporter molecule from the quencher molecule (Adapted from http://www.bio-rad.com; http://www.sigmaaldrich.com; http://www.eurogentec.com).

Melting curve analysis or high-resolution melting (HRM) curve analysis is an extension of real-time PCR that can be used for genotyping based on PCR product melt profiles (Reed, 2007). Melting curve analysis or HRM curve analysis relies on the dissociation of DNA binding dyes from melting dsDNA. High-resolution melting is much more sensitive than conventional melting curve analysis and a

temperature resolution of 0.1°C/s is possible (Reed 2007). High-resolution melting requires a saturating concentration of an intercalating dye like SYTO 9, EVAGreen or LCGreen (Jeffery et al., 2007; Reed, 2007; Toi and Dwyer, 2008). The saturating concentration prevents dye molecule redistribution during melting, providing superior HRM resolution (http://www.bio-rad.com). Unlike normal DNA binding dyes, such as SYBR Green, saturating dyes can be used at high concentrations because they do not inhibit DNA polymerases (www.kapabiosystems.com). High-resolution melting analysis has been used primarily for the discovery and genotyping of single nucleotide polymorphisms (SNPs), but has also been used for precise amplicon verification (Mackay et al., 2008). High-resolution amplicon melting has already been used for the analysis of microsatellite markers in a number of rootstock and grapevine scion varieties commonly grown in New Zealand for varietal certification (Mackay et al., 2008). High-resolution melting curve analysis has not yet been applied to differentiate between grapevine viruses. In chapter 4 a real-time RT-PCR HRM assay is described that can detect and differentiate between different molecular variant groups of GLRaV-3.

Grapevines are prone to contain mixed infections of several viruses, making the use of most of the abovementioned techniques time-consuming. Oligonucleotide microarray analysis is a technique that can be used to detect several viruses or genes at the same time. A grapevine microarray, containing 570 unique probes designed against highly conserved and species-specific regions of 44 plant viral genomes could accurately detect 10 grapevine viruses (Engel et al., 2010). This approach provides a powerful tool for high throughput screening that can be useful for plant certification purposes. Microarray technologies are still expensive and require extensive data analysis. Recently, the successful application of macro-array methodology was demonstrated as an alternative to microarray technology. Thompson et al. (2012) provided an unbiased multiplex detection system using a single robust macro-array platform for grapevine leafroll-associated viruses. The relative simplicity and robustness of this methodology will be accessible to most molecular biology laboratories, but this platform can be limited by the availability of sequence data for certain virus species and differ in detection sensitivity in comparison to RT-PCR.

2.4.5 Next-generation sequencing

Present grapevine disease diagnostics rely on ELISA or RT-PCR to target viruses that have in the past been associated with diseases (Adams et al., 2009). Although these techniques can be very specific and reliable, they do not take into the account the contribution of other known or unknown viruses that may

be involved in the disease etiology. Although a number of viruses have been associated with specific diseases, viral diseases are often caused by a complex of viruses with more than one virus infecting a single plant (Prosser et al., 2007). Different virus variants can also exist that may go undetected if highly specific RT-PCR protocols are used. Both ELISA and RT-PCR require prior knowledge of the viruses. The use of metagenomic sequencing to establish the total viral complement of a sample has been shown to avoid the abovementioned limitations of current grapevine diagnostics (Adams et al., 2009; Al Rwahnih et al., 2009; Coetzee et al., 2010; Kreuze et al., 2009). New technologies which are able to sequence viruses from environmental samples can result in the generation of sequence information for the complete virome using an unbiased approach. Second generation or next-generation sequencing (NGS) instruments have been developed, avoiding the limitations associated with Sanger sequencing (Hall, 2007; Mardis, 2008). The use of universal adaptors, rather than sequence-specific primers, makes next-generation sequencing specifically suitable to sequence all the genetic material present in a sample without prior knowledge of the organisms present (Hall, 2007; Mardis, 2008; Tucker et al., 2009). Next-generation sequencing does not use chain termination chemistry and electrophoresis like Sanger sequencing, but relies on the massively parallel sequencing of clonally amplified or single DNA molecules that are spatially separated in a flow cell (Voelkerding, 2009). Different next-generation sequencing platforms exist that are capable of producing millions of DNA sequence reads in a single run. These include the Roche GS FLX Titanium XL+FLX system, the Applied Biosystems SOLiD 5500xl System and the Illumina HiSeq 2500 system (Table 2.2).

Table 2.2. Comparison of the latest available next-generation sequencing platforms.

	Sequencing platforms		
	Roche GS FLX Titanium XL+FLX system	**Applied Biosystems SOLiD 5500xl System**	**Illumina HiSeq system**
Read Length	Up to 1,000 bp	Mate-paired: 2 x 60 bp Paired-end: 75 bp x 35 bp Fragment: 75 bp	2 × 100 bp
Throughput/output	700 Mb per run	10–15 Gb/day 180 Gb per run	600 Gb per run
Reads per Run	~1,000,000 shotgun	Single reads: 800 Million Paired-end reads: 2.8 Billion	Single reads: 3 Billion total (187 Million/lane) Paired-end reads: 6 Billion (374 Million/lane)

Data were obtained from the respective websites:
Roche: http://454.com/products/gs-flx-system/index.asp
Applied Biosystems: http://www.appliedbiosystems.com/absite/us/en/home/applications-technologies/solid-next-generation-sequencing/next-generation-systems.html
Illumina: http://www.illumina.com/systems/sequencing.ilmn

Several *in silico* software tools and algorithms have been developed for the extensive data analysis associated with next-generation sequencing. The choice of software will depend on the characteristics of the output data and will include *de novo* sequence assemblers (e.g. Velvet), mapping assemblers (e.g. MAQ) and alignment viewers (MapView) (Zerbino and Birney, 2008; Li et al., 2008). Numerous studies have applied next-generation sequencing successfully to identify known and novel viruses from diseased plant material. From a single grapevine plant displaying typical Syrah decline symptoms an unknown marafivirus was identified (Al Rwahnih et al., 2009). A new cucumovirus, with the suggested name gayfeather mild mottle virus, was isolated from *Liatris spicata* (Adams et al., 2009). High-throughput parallel sequencing of small RNAs from diseased, as well as symptomless plants, yielded the expected full-length sequences of the inoculated viruses (Sweetpotato feathery mottle virus and Sweetpotato chlorotic stunt virus), as well as that of unknown viruses belonging to the Badnavirus and Mastrevirus genera (Kreuze et al., 2009). Coetzee et al. (2010) established the viral profile of a severely diseased vineyard and identified grapevine virus E (GVE) for the first time in South Africa. Chapter 3 also describes the identification and confirmation of a new GLRaV-3 variant that was first identified using next-generation sequencing. These studies indicate the usefulness of next-generation sequencing technologies as a diagnostic tool to identify a plant virus when no prior knowledge of the virus is available. Next-generation sequencing is still relatively expensive to be used for routine diagnostics. However, data generated can be used to develop more accurate diagnostic assays since next-generation sequencing can provide information regarding disease complexes, dominant variants of viral species and an indication of the frequency of viruses found in infected material.

2.5 Conclusion

Grapevine leafroll disease is considered to be the most economically destructive disease of grapevine and a major constraint to the production of premium wine grapes. Final conclusions regarding the economic impact of GLD are complicated by numerous factors, but significant yield losses are the common denominator in all studies. Symptom expression of GLD is highly variable among cultivars and not identifiable based on visual indications alone. Accurate virus screening is the cornerstone of GLD management to ensure high-quality grapevine certification programs. Grapevine leafroll disease has been more closely associated with GLRaV-3 than any other GLRaV, and has therefore been the focus of GLD diagnostics. The genetic variation observed in the GLRaV-3 genome has been studied more intensively ever since the publishing of the first complete genome sequence in 2004. Many phylogenetic studies based on partial sequences of different GLRaV-3 isolates, showed the existence of

at least five molecular variant groups with some of these sequences, like isolate NZ-1, clustering separate from these groups. All of these variants were identified in a number of grapevine varieties and from different countries around the world, supporting the view that they are genuine genomic variants of GLRaV-3. It is therefore necessary to not only have sensitive and rapid detection methods to test for commonly occurring viruses, but also to have techniques available to detect all virus variants and new emerging viruses that may influence disease etiology.

2.6 References

Abou Ghanem-Sabanadzovic, N., Sabanadzovic, S., Uyemoto, J., Golino, D., Rowhani, A., 2010. A putative new ampelovirus associated with grapevine leafroll disease. Arch.Virol. 155, 1871-1876.

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., Boonham, N., 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. Mol. Plant. Pathol. 10(4), 537-545.

Agranovsky, A.A., Koonin, E.V., Boyko, V.P., Maiss, E., Frotschl, R., Lunina, N.A., Atabekov, J.G., 1994. Beet yellows closterovirus: Complete genome structure and identification of leader papain-like thiol protease. Virology. 198, 311-324.

Akbas, B., Kunter, B., Ilhan, D., 2007. Occurrence and distribution of grapevine leafroll-associated viruses 1, 2, 3 and 7 in Turkey. J. Phytopathol. 155, 122-124.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Looking at the Structure of Cells in the Microscope, in: Molecular Biology of the Cell. 4th edition. Garland Science, New York, available from: http://www.ncbi.nlm.nih.gov/books/NBK26880/

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, 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. Arch. Virol. 156, 397-403.

Angelini, E., Luca Bianchi, G., Filippin, L., Morassutti, C., Borgo, M., 2007. A new TaqMan® method for the identification of phytoplasmas associated with grapevine yellows by real-time PCR assay. J. Microbiol. Methods. 68, 613-622.

Atallah, S.S., Gómez, M.I., Fuchs, M.F., Martinson, T.E., 2011. Economic impact of grapevine leafroll disease on *Vitis vinifera* cv. Cabernet franc in Finger Lakes Vineyards of New York. Am. J. Enol. Viticult. DOI: 10.5344/ajev.2011.11055

Beier, D.R., 1993. Single-strand conformation polymorphism (SSCP) analysis as a tool for genetic mapping. Mamm. Genome. 4(11), 627-631.

Bell, V.A., Bonfiglioli, R.G.E., Walker1, J.T.S., Lo1, P.L., Mackay, J.F., McGregor, S.E. 2009. Grapevine leafroll-associated virus 3 persistence in *Vitis vinifera* remnant roots. J. Plant. Pathol. 91(3), 527-533.

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. Arch. Virol. 157, 1815-1819.

Boscia, D., Greif, C., Gugerli, P., Martelli, G.P., Walter, B., Gonsalves, D., 1995. Nomenclature of grapevine leafroll-associated putative closteroviruses. Vitis. 34, 171-175.

Bustamante, P.I., Hull, R., 1998. Plant virus gene expression strategies. Electron. J. Biotechn. 1(2), DOI: 10.2225/vol1-issue2-fulltext-3.

Cabaleiro, C., Segura, A., García-Berrios, J.J., 1999. Effects of grapevine leafroll-associated virus 3 on the physiology and must of *Vitis vinifera* L. cv. Albariño following contamination in the field. Am. J. Enol. Viticult. 50, 40-44.

Chambers, P.J., Pretorius, I.S., 2010. Fermenting knowledge: the history of winemaking, science and yeast research. EMBO Rep. 11, 914-920.

Charles, J.G., Cohen, D., Walker, J.T.S., Forgie, S.A., Bell, V.A., Breen, K.C., 2006a. A review of Grapevine leafroll-associated virus type 3 (GLRaV-3) for the New Zealand wine industry. Report to New Zealand Winegrowers. HortResearch Client Report. 18447, 1-79.

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

Charles, J., Froud, K., van, d.B., Allan, D., 2009. Mealybugs and the spread of grapevine leafroll-associated virus 3 (GLRaV-3) in a New Zealand vineyard. Australas. Plant. Pathol. 38, 576-583.

Cirami, R.M., Velsen, R.J., Niejalke, J., 1988. Grapevine virus indexing in the South Australian vine improvement scheme, 1974-1987. Aust. J. Exp. Agric. 28, 645-649.

Clark, M.F., Adams, A.N., 1977. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. J. Gen. Virol. 34, 475-483.

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

Constable, F.E., Nicholas, P., Rodoni, B.C. 2010. Development and validation of diagnostic protocols for the detection of endemic and exotic pathogens of grapevines. Final report to grape and wine research & development corporation. Department of primary industries, DPI 05/04.

Corbett, M.K., 1974. Detection of viruses and diagnosis of plant viral diseases by electron microscopy. Acta Hort. (ISHS). 36, 141-188.

Dai, J., Cheng, J., Huang, T., Zheng, X., Wu, Y., 2012. A multiplex reverse transcription PCR assay for simultaneous detection of five tobacco viruses in tobacco plants. J. Virol. Methods. 183, 57-62.

Dolja, V.V., Karasev, A.V., Koonin, E.V., 1994. Molecular biology and evolution of closteroviruses: Sophisticated build-up of large RNA genomes. Annu. Rev. Phytopathol. 32, 261-285.

Dolja, V.V., Kreuze, J.F., Valkonen, J.P.T., 2006. Comparative and functional genomics of closteroviruses. Virus Res. 117, 38-51.

Dolja, V.V., 2009. New developments in understanding gene function and evolution of the grapevine closteroviruses. Extended abstracts 16th ICVG conference, Dijon, France. 29-31.

Dovas, C.I., Katis, N.I., 2003. A spot multiplex nested RT-PCR for the simultaneous and generic detection of viruses involved in the aetiology of grapevine leafroll and rugose wood of grapevine. J. Virol. Methods. 109, 217-226.

Elnifro, E.M., Ashshi, A.M., Cooper, R.J., Klapper, P.E., 2000. Multiplex PCR: Optimization and application in diagnostic virology. Clin. Microbiol. Rev. 13, 559-570.

Engel, E.A., Escobar, P.F., Rojas, L.A., Rivera, P.A., Fiore, N., Valenzuela, P.D.T., 2010. A diagnostic oligonucleotide microarray for simultaneous detection of grapevine viruses. J. Virol. Methods. 163, 445-451.

Engel, E., Girardi, C., Escobar, P., Arredondo, V., Dominguez, C., Perez-Acle, T., Valenzuela, P., 2008. Genome analysis and detection of a Chilean isolate of grapevine leafroll-associated virus-3. Virus Genes. 37, 110-118.

Engvall, R., Perlmann, P., 1971. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. Immunochemistry. 8, 871-874.

Feng, J., Zeng, R., Chen, J., 2008. Accurate and efficient data processing for quantitative real-time PCR using a tripartite plant virus as a model. BioTechniques. 44, 901-912.

Fiore, N., Prodan, S., Montealegre, J., Aballay, E., Pino, A.M., Zamorano, A., 2008. Survey of grapevine viruses in Chile. J. Plant. Pathol. 90, 125-130.

Freeborough, M.J., Burger, J.T., 2008. Leafroll: Economic implications. Wynboer, http://www.wynboer.co.za/recentarticles/200812-leafroll.php3

Fuchs, M., Marsella-Herrick, P., Loeb, G.M., Martinson, T.E., Hoch, H.C., 2009. Diversity of Ampeloviruses in Mealybug and Soft Scale Vectors and in Grapevine Hosts from Leafroll-Affected Vineyards. Phytopathology. 99, 1177-1184.

Fuchs, M., Martinson, T.E., Loeb, G.M., Hoch, H.C., 2009. Survey for the three major leafroll disease-associated viruses in Finger Lakes vineyards in New York. Plant Dis. 93, 395-401.

Golino, D., Sim, S., Gill, R., Rowhani, A. 2002. California mealybugs can spread grapevine leafroll disease. Calif. Agr. 56(6), 196-201.

Gouveia, P., Santos, M., Eiras-Dias, J., Nolasco, G., 2011. Five phylogenetic groups identified in the coat protein gene of grapevine leafroll-associated virus 3 obtained from Portuguese grapevine varieties. Arch. Virol. 156, 413-420.

Habili, N., Fazeli, C., Ewart, A., Hamilton, R., Cirami, R., Saldarelli, P., Minafra, A., Rezaian, M., 1995. Natural spread and molecular analysis of grapevine leafroll-associated virus 3 in Australia. Phytopathology. 85, 1418-1422.

Hall, N., 2007. Advanced sequencing technologies and their wider impact in microbiology. J. Exp. Biol. 210, 1518-1525.

Holland, P.M., Abramson, R.D., Watson, R., Gelfand, D.H., 1991. Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of Thermus aquaticus DNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 88(16), 7276-7280.

Jarugula, S., Gowda, S., Dawson, W., Naidu, R., 2010. 3' co-terminal subgenomic RNAs and putative cis-acting elements of Grapevine leafroll-associated virus 3 reveals 'unique' features of gene expression strategy in the genus Ampelovirus. Virology Journal. DOI:10.1186/1743-422X-7-180.

Jeffery, N., Gasser, R.B., Steer, P.A., Noormohammadi, A.H., 2007. Classification of Mycoplasma synoviae strains using single-strand conformation polymorphism and high-resolution melting-curve analysis of the vlhA gene single-copy region. Microbiology. 153, 2679-2688.

Jooste, A.E.C., Goszczynski, D.E. 2005. Single-strand conformation polymorphism (SSCP), cloning and sequencing reveals two major groups of divergent molecular variants of grapevine leafroll-associated virus 3 (GLRaV-3). Vitis. 44, 39-43.

Jooste, A.E.C., Maree, H.J, Bellstedt, D., 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. Arch. Virol. 155, 1997-2006.

Karasev, A.V., Boyko, V.P., Gowda, S., Nikolaeva, O.V., Hilf, M.E., Koonin, E.V., Niblett, C.L., Cline, K., Gumpf, D.J., Lee, R.F., Garnsey, S.M., Lewandowski, D.J., Dawson, W.O., 1995. Complete sequence of the citrus tristeza virus RNA genome. Virology. 208, 511-520.

Karasev, A.V., Nikolaeva, O.V., Mushegian, A.R., Lee, R.F., Dawson, W.O., 1996. Organization of the 3'-terminal half of beet yellow stunt virus genome and implications for the evolution of closteroviruses. Virology. 221, 199-207.

Kitajima, E.,W., 2004. Electron microscopy in plant virology: Past, present and future. Microsc. Microanal. 10(2), 212-213.

Klaassen, V.A., Boeshore, M.L., Koonin, E.V., Tian, T., Falk, B.W., 1995. Genome structure and phylogenetic analysis of lettuce infectious yellows virus, a whitefly-transmitted, bipartite closterovirus. Virology. 208, 99-110.

Krenze, B., Thompson, J.R., Fuchs, M., Perrya, K.L., 2012. Complete Genome Sequence of a New Circular DNA Virus from Grapevine. J. Virol. 86(14), 7715.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., 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, 1-7.

La Notte, P., Minafra, A., Saldarelli, P., 1997. A spot-PCR technique for the detection of phloem-limited grapevine viruses. J. Virol. Methods. 66, 103-108.

Li, H., Ruan, J., Durbin, R., 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. 18, 1851-1858.

Ling, K.S., Zhu, H.Y., Alvizo, H., Hu, J.S., Drong, R.F., Slightom, J.L., Gonsalves, D., 1997. The coat protein gene of grapevine leafroll-associated closterovirus-3: cloning, nucleotide sequencing and expression in transgenic plants. Arch. Virol. 142, 1101-1116.

Ling, K.S., Zhu, H.Y., Drong, R.F., Slightom, J.L., McFerson, J.R., Gonsalves, D., 1998. Nucleotide sequence of the 3'-terminal two-thirds of the grapevine leafroll-associated virus-3 genome reveals a typical monopartite closterovirus. J. Gen. Virol. 79, 1299-1307.

Ling, K.S, Zhu, H.Y., Gonsalves, D., 2004. Complete nucleotide sequence and genome organization of Grapevine leafroll-associated virus 3, type member of the genus Ampelovirus. J. Gen. Virol. 85, 2099-2102.

Ling, K., Zhu, H., Jiang, Z., Gonsalves, D., 2000. Effective application of DAS-ELISA for detection of grapevine leafroll-associated closterovirus-3 using a polyclonal antiserum developed from recombinant coat protein. Eur. J. Plant Pathol. 106, 301-309.

Ling, K., Zhu, H., Petrovic, N., Gonsalves, D., 2001. Comparative effectiveness of ELISA and RT-PCR for detecting grapevine leafroll-associated closterovirus-3 in field samples. Am. J. Enol. Vitic. 52, 21-27.

Lipman, N.S., Jackson, L.R., Trudel, L.J., Weis-Garcia, F., 2005. Monoclonal versus polyclonal antibodies: Distinguishing characteristics, applications, and information resources. ILAR J. 46(3), 285-268.

Mackay, J., Wright, C., Bonfiglioli, R., 2008. A new approach to varietal identification in plants by microsatellite high-resolution melting analysis: application to the verification of grapevine and olive cultivars. Plant methods. DOI:10.1186/1746-4811-4-8.

Mahfoudhi, N., Digiaro, M., Dhouibi, M.H., 2008. Incidence and distribution of grapevine leafroll-associated viruses in Tunisian vineyards. J. Phytopathol. 156, 556-558.

Maliogka, V.I., Dovas, C.I., Katis, N.I., 2008. Evolutionary relationships of virus species belonging to a distinct lineage within the Ampelovirus genus. Virus Res. 135, 125-135.

Maliogka, V., Dovas, C., Lotos, L., Efthimiou, K., Katis, N., 2009. Complete genome analysis and immunodetection of a member of a novel virus species belonging to the genus *Ampelovirus*. Arch. Virol. 154, 209-218.

Mannini, F., Argamante, N., Credi, R., 1996. Improvement in the quality of grapevine 'Nebbiolo' clones obtained by sanitation. Acta Hort. 427, 319-324.

Mannini, F., Mollo, A., Credi, R., 2011. Field performance and wine quality modification in a clone of *Vitis vinifera* cv. Dolcetto after GLRaV-3 elimination. Am. J. Enol. Vitic. DOI: 10.5344/ajev.2011.11020

Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. 24(3), 133-141.

Maree, H.J., Freeborough, M-J., Burger, J.T., 2008. Complete nucleotide sequence of a South African isolate of grapevine leafroll-associated virus 3 reveals a 5' UTR of 737 nucleotides. Arch. Virol. 153, 755-757.

Maree, H.J., Gardner, H.F.J., Freeborough, M-J., Burger, J.T., 2010. Mapping of the 5' terminal nucleotides of Grapevine leafroll-associated virus 3 sgRNAs. Virus Res. 151, 252-255.

Maree, H.J., 2010. Identification and characterisation of Grapevine leafroll-associated virus 3 genomic and subgenomic RNAs. PhD thesis, Stellenbosch University.

Martelli, G.P., Abou Ghanem-Sabanadzovic, N., Agranovsky, A.A., Al Rwahnih, M., Dolja, V.V.,

Dovas, C.I., Fuchs, M., Gugerli, P., Hu, J.S., Jelkmann, W., Katis, N.I., Maliogka, V.I., Melzer, M.J., Menzel, W., Minafra, A., Rott, M.E., Rowhani, A., Sabanadzovic, S., 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. J. Plant. Pathol. DOI: 10.4454/jpp.fa.2012.022

Martelli, G.P., Agranovsky, A.A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R.H.A., Dolja, V.V., Falk, B.W., Gonsalves, D., Jelkmann, W., Karasev, A.V., Minafra, A., Namba, S., Vetten, H.J., Wisler, G.C., Yoshikawa, N., 2000. Family *Closteroviridae*, in: Van Regenmortel, M.H.V., Fauquet, C.L., Bishop, D.H.L., Carstens, E., Estes, M.K., Lemon, S., Maniloff, J., Mayo, M.A., McGeoch, D., Pringle, C.R., Wickner, R.B. (Eds.), Virus taxonomy. Seventh report of the international committee on taxonomy of viruses. Academic Press, San Diego, pp. 943–952.

Martelli, G.P., Agranovsky, A.A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R.H.A., Dolja, V.V., Falk, B.W., Gonsalves, D., Jelkmann, W., Karasev, A.V., Minafra, A., Namba, S., Vetten, H.J., Wisler, G.C., Yoshikawa, N., 2002. The family *Closteroviridae* revised. Arch. Virol. 147, 2039-2044.

Martelli, G.P., Boudon-Padieu, E., 2006. Directory of infectious diseases of grapevines and viroses and viruslike diseases of grapevine: Bibliographic report 1998-2004. Opinions Mediterraneennes Serie B: Studies and Research.

Martelli G.P., Digiaro M. 1999. Proceedings of the Mediterranean network on grapevine closteroviruses 1992-1997 and the viroses and virus-like diseases of the grapevine a bibliographic report 1985-1997. Options Méditerranéennes série B: Etudes et Recherches.

Monis, J., 2000. Development of monoclonal antibodies reactive to a new grapevine leafroll-associated closterovirus. Plant Dis. 84, 858-862.

Monis, J., Bestwick, R.K., 1996. Detection and localization of grapevine leafroll-associated closteroviruses in greenhouse and tissue culture grown plants. Am. J. Enol. Vitic. 47, 199-205.

Naidua, R.A., Hughes, J.d'A. 2001. Methods for the detection of plant virus diseases. Plant virology in sub-Saharan Africa**:** Proceedings of a conference organized by international institute of tropical agriculture, Ibadan, Nigeria. 233-260.

Nakaune, R., Toda, S., Mochizuki, M., Nakano, M., 2008. Identification and characterization of a new vitivirus from grapevine. Arch. Virol. 153, 1827-1832.

Nicholas, P., 2006. Selection of clones for the Australian national nuclear grapevine collection. The Australian & New Zealand Grapegrower & Winemaker. May, 31-33.

O'Donnell, K., 1999. Plant pathogen diagnostics: present status and future developments. Potato Res. 42, 437-447.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., Sekiya, T., 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. U.S.A. 86(8), 2766-2770.

Osman, F., Leutenegger, C., Golino, D., Rowhani, A., 2008. Comparison of low-density arrays, RT-PCR and real-time TaqMan® RT-PCR in detection of grapevine viruses. J. Virol. Methods. 149, 292-299.

Osman, F., Leutenegger, C., Golino, D., Rowhani, A., 2007. Real-time RT-PCR (TaqMan®) assays for the detection of grapevine leafroll-associated viruses 1-5 and 9. J. Virol. Methods. 141, 22-29.

Osman, F., Rowhani, A., 2006. Application of a spotting sample preparation technique for the detection of pathogens in woody plants by RT-PCR and real-time PCR (TaqMan®). J. Virol. Methods. 133, 130-136.

Over de Linden, A.J., Chamberlain, E.E., 1970. Effect of grapevine leafroll virus on vine growth and fruit yield and quality. N. Z. J. Agric. Res. 13, 689-698.

Palacio, A., Duran-Vila, N., 1999. Single-strand conformation polymorphism (SSCP) analysis as a tool for viroid characterisation. J. Virol. Methods. 77, 27-36.

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

Pietersen, G., 2004. Spread of grapevine leafroll disease in South Africa - a difficult, but not insurmountable problem. Wynboer, http://www.wynboer.co.za/recentarticles/articles.php3.

Prosser, S.W., Goszczynski, D.E., Meng, B., 2007. Molecular analysis of double-stranded RNAs reveals complex infection of grapevines with multiple viruses. Virus Res. 124, 151-159.

Rayapati, N., O'Neal, S., Walsh, D., 2008. Grapevine leafroll disease. Washington State University Extension. EB2027E, http://cru.cahe.wsu.edu/CEPublications/eb2027e/eb2027e.pdf.

Reed, G.H., 2007. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. Pharmacogenomics. 8, 597-608.

Rowhani, A., Uyemoto, J.K., Golino, D.A., 1997. A comparison between serological and biological assays in detecting grapevine leafroll-associated viruses. Plant Dis. 81, 799-801.

Schaad, N.W., Frederick, R.D., Shaw, J., Schneider, W.L., Hickson, R., Petrillo, M.D., Luster, D.G., 2003. Advances in molecular-based diagnostics in meeting crop biosecurity and phytosanitary issues. Annu. Rev. Phytopathol. 41, 305-324.

Sforza, R., Boudon-padieu, E., Greif, C., 2003. New mealybug species vectoring grapevine leafroll-associated viruses-1 and -3 (GLRaV-1 and -3). Eur. J. Plant Pathol. 109, 975-981.

Shanmuganathan, N., Fletcher, G., 1980. Indexing grapevine clones in the fruit variety foundation of Australia for virus and virus-like diseases. Aust. J. Exp. Agric. 20, 115-118.

Sharma, A.M., Wang, J., Duffy, S., Zhang, S., Wong, M.K., Rashed, A., Cooper, M.L., Daane, K.M., Almeida, R.P.P., 2011. Occurrence of grapevine leafroll-associated virus complex in Napa Valley. PLoS ONE. 6, e26227.

Steere, R.L., 1964. Electron microscopy of plant viruses. Bot. Rev. 30, 629-666.

Teliz, D., 1987. Field serological detection of viral antigens associated with grapevine leafroll disease. Plant Dis. 71(8), 704-709.

Thompson, J. R., Fuchs, M., Fischer, K.F., Perry, K.L., 2012. Macroarray detection of grapevine leafroll-associated viruses. J. Virol. Methods. 183, 161-169.

Toi, C.S., Dwyer, D.E., 2008. Differentiation between vaccine and wild-type varicella-zoster virus genotypes by high-resolution melt analysis of single nucleotide polymorphisms. J. Clin. Virol. 43, 18-24.

Tsai, C.W., Chau, J., Fernandez, L., Bosco, D., Daane, K.M., Almeida, R.P.P., 2008. Transmission of grapevine leafroll-associated virus 3 by the vine mealybug (*Planococcus ficus*). Phytopathology. 98, 1093-1098.

Tucker, T., Marra, M., Friedman, J.M., 2009. Massively parallel sequencing: The next big thing in genetic medicine. Am. J. Hum. Genet. 85, 142-154.

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. J. Gen. Virol. 86, 217-224.

Van Weemen, B.K., Schuurs, A.H.W.M., 1971. Immunoassay using antigen-enzyme conjugates. FEBS Lett. 15, 232-236.

Valverde, R.A., Nameth, S.T., Jordan, R.L., 1990. Analysis of double-stranded-RNA for plant-virus diagnosis. Plant Dis. 74(3), 255-258.

Vega, A., Gutiérrez, R., Pena-Neira, A., Cramer, G., Arce-Johnson, P., 2011. Compatible GLRaV-3 viral infections affect berry ripening decreasing sugar accumulation and anthocyanin biosynthesis in *Vitis vinifera*. Plant. Mol. Biol. 77, 261-274.

Voelkerding, K.V., Dames, S.A., Durtschi, J.D., 2009. Next-generation sequencing: From basic research to diagnostics. Clin. Chem. 55, 641-58.

Ward, E., Foster, S.J., Fraaije, B.A., McCartney, H.A., 2004. Plant pathogen diagnostics: immunological and nucleic acid-based approaches. Ann. Appl. Biol. 145, 1-16.

Weber, E., Golino, D., Rowhani, A., 2002. Laboratory testing for grapevine diseases. Practical winery and vineyard.
http://fpms.ucdavis.edu/WebSitePDFs/Articles/LaboratoryTestingForGrapevineDiseasesPWJanFeb2002.pdf.

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. Biotechnol. J. 3, 1424-1428.

Zechmann, B., Zellnig, G., 2009. Rapid diagnosis of plant virus diseases by transmission electron microscopy. J. Virol. Methods. 162, 163-169.

Zee, F., Gonsalves, D., Goheen, A., Kim, K.S., Pool, R., Lee, R.F., 1987. Cytopathology of leafroll-diseased grapevines and the purification and serology of associated closteroviruslike particles. Phytopathology. 77, 1427-1434.

Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821-829.

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.

2.7 Recommended internet resources

http://www.wosa.co.za/sa/history_beginning.php

http://www3.bio-rad.com/B2B/vanity/gexp/content.do?BV_SessionID=@@@@1369559164.1346146328@@@@&BV_EngineID=ccccadfhjegkiefcfngcfkmdhkkdfll.0&root=/Product%20Family/GX/Home&pcatoid=-35468&ccatoid=-36527&country=US&language=English&BV_SessionID=@@@@1369559164.1346146328@@@@&BV_EngineID=ccccadfhjegkiefcfngcfkmdhkkdfll.0

http://www.sigmaaldrich.com/etc/medialib/docs/Sigma/Brochure/fluorescent_dna_probes.Par.0001.File.tmp/fluorescent_dna_probes.pdf

http://www.eurogentec.com/uploads/qPCR-guide.pdf

http://www.kapabiosystems.com/public/pdfs/kapa-hrm-fast-pcr-kits/Introduction_to_High_Resolution_Melt_Analysis_Guide.pdf

http://454.com/products/gs-flx-system/index.asp

http://www.appliedbiosystems.com/absite/us/en/home/applications-technologies/solid-next-generation-sequencing/next-generation-systems.html

http://www.illumina.com/systems/sequencing.ilmn

Chapter 3: Complete nucleotide sequence of a new strain of grapevine leafroll-associated virus 3

3.1 Introduction

Grapevine leafroll disease (GLD) is an economically important disease of grapevine worldwide that reduces yield and quality of both wine and table grapes. The main causative agent in South African vineyards is grapevine leafroll-associated virus 3 (GLRaV-3) (Pietersen, 2004). It is a positive-sense single-stranded RNA virus that is phloem-limited and is the type member of the genus *Ampelovirus* in the family *Closteroviridae* (Martelli et al., 2002).

In South Africa, three molecular variant groups of GLRaV-3 have been identified, and these are supported by complete nucleotide (nt) sequences. They include isolates 621 (group I), 623 and GP18 (both group II), and PL-20 (group III) (Jooste et al., 2010; Maree et al., 2008). Two additional variant groups have been identified in Portugal, but sequencing data of these groups are limited (Gouveia et al., 2010). To date a variety of techniques have been applied to detect, identify and characterize novel viruses. Metagenomic sequencing or next-generation sequencing is an effective alternative to conventional RT-PCR and Sanger sequencing to identify viral pathogens occurring at extremely low titers, without the necessity of any prior knowledge (Adams et al., 2009; Al Rwahnih et al., 2009; Coetzee et al., 2010; Kreuze et al., 2009). Next-generation sequencing has been applied in a number of studies showing the viability of this approach to detect and identify known and novel viruses. Evidence from a metagenomic sequencing study of a diseased vineyard showed the existence of at least two GLRaV-3 genetic variants (Coetzee et al., 2010). One of these variants was similar to the group II variants, which is well characterized and widespread in South African vineyards, while the other variant was not previously detected (Coetzee et al., 2010; Maree, 2010). This new South African variant was closely related to the New Zealand isolate, NZ-1.

Accurate virus detection and identification is essential for preventing the spread of diseases. It is for this reason very important to have reliable detection techniques that are able to detect all variants of a specific virus. The discovery of the new GLRaV-3 variant in South Africa raised concerns regarding the efficacy of the South African wine industry's diagnostic approach. Enzyme-linked immunosorbent assay (ELISA) is the wine industry's method of choice for virus diagnostics due to the low cost and

potential for high throughput (Charles et al., 2006). However, to establish an ELISA for a newly identified virus has a high development cost and is time consuming (Ward et al., 2004). Hence, it was important to evaluate the industry's current diagnostic ELISA for its ability to detect the newly identified GLRaV-3 variant.

This chapter describes the sequencing of two representative isolates of the new molecular variant group of GLRaV-3 (GH11 and GH30) in order to validate the draft sequence generated from the metagenomic data. Additionally, the industry standard ELISA diagnostic is evaluated for its efficacy in detecting the newly identified variant group of GLRaV-3.

3.2 Materials and methods

3.2.1 Bioinformatics

The sequencing data from the metagenomic study (Coetzee et al., 2010) was used to assemble a near complete draft sequence of the new GLRaV-3 variant using the *de novo* assembler Velvet 0.7.49 (Zerbino and Birney, 2008). Different parameters were experimented with in order to improve the GLRaV-3-specific scaffolds. The parameters used to assemble the 19,247,026 reads into longer GLRaV-3 scaffolds were a hash length of 31, a coverage cut-off of 100 and an expected coverage of 7000. Scaffolds resembling full-length sequences of GLRaV-3 were identified using BLAST against the NCBI nucleotide and protein databases. The scaffold most related to the New Zealand isolate was selected and insertion-deletion errors observed in the open reading frames (ORFs) of this scaffold were manually removed by comparing the ORFs to those of other complete genomes of GLRaV-3 (GP18 [EU259806.1] and NY-1 [AF037268.2]) using BioEdit 7.0.5.3 (Hall, 1999). This edited sequence was designated as the draft of the new GLRaV-3 variant's genome and used as the reference sequence for mapping the original reads. Read mapping was performed in order to calculate the read depth and coverage using the "easyrun" command of the Mapping and Assembly with Quality (MAQ) assembler version 0.7.1 (Li et al., 2008).

3.2.2 Preliminary survey

Existing diagnostic primer sets were unable to detect the new variant and therefore new diagnostic primers were designed to specifically detect this variant. The primers were designed from the assembled draft sequence to amplify the 5' untranslated region (UTR) and used to screen field samples. These primers were designated, LR3-Like_(26)_For and LR3-Like_(775)_Rev (Table 3.1). The known

variants of GLRaV-3 were detected using the LC1/LC2 primers (Table 3.1) (Osman and Rowhani, 2006).

Table 3.1. Primers sets used to sequence the isolates of the new variant of GLRaV-3 (GH11 and GH30) and the primer sets used for GLRaV-1, GLRaV-2 and GLRaV-3 diagnostics.

Amplicon /Purpose	Primer name	Primer sequence (5'-3')	Target position*	Size (bp)	Reference
Genome sequencing 1	SNZ45F	TCACCAGAACTTTCCTCCT	46	1640	This study
	SNZ1675R	AACCAGCACCACCCTCTT	1685		
Genome sequencing 2	SNZ1599F	GTTGCCTTGCCCGATATT	1599	1528	This study
	SNZ3143R	GACGACTAACGCCCACAT	3126		
Genome sequencing 3	SNZ3072F	GTTATCACGGGAAGTCGC	3072	1748	This study
	SNZ4845R	TCGGCAAAATGTCCTCAG	4819		
Genome sequencing 4	SNZ3558F	CCGCTCCCAATAACTGAAGA	3558	581	This study
	SNZ4155R	TACCGCTGAAGATAACCCAC	4138		
Genome sequencing 5	SNZ4770F	CAGGTGGTTAAGGACAAAGC	4761	1852	This study
	SNZ6628R	CTACATTCGCCGTTATCAC	6612		
Genome sequencing 6	SNZ5439F	GCTGGGTGAAGGCTATACG	5441	486	This study
	SNZ5943R	CTGGGGCTTAGTGACTGGCA	5926		
Genome sequencing 7	SNZ6576F	GGATGAGTTCGTTAGGGC	6578	1740	This study
	SNZ8333R	CACCAAGCGTCAAGAGTAT	8317		
Genome sequencing 8	SNZ8272F	CGACAAATCTCAGAGTGCC	8274	1806	This study
	SNZ10095R	GAAAAAGCGAGCCACCTA	10079		
Genome sequencing 9	SNZ10015F	CAAGCAGGGTGGGTAAAT	10016	1718	This study
	SNZ11756R	CTAAATGGTGCGACTATGG	11733		
Genome sequencing 10	SNZ11664F	CGCTATAACGCAACACCTC	11659	1509	This study
	SNZ13189R	GGGCAACATCCTAACCAT	13167		
Genome sequencing 11	SNZ12549	AACGCTGCTAGGTCGGAAG	12540	1723	This study
	SNZ14290	GTGCTGGTGTCGGTCTAACT	14262		
Genome sequencing 12	SNZ13135F	CACTGTGTCCGCTTATGC	13130	1847	This study
	SNZ14999R	TCGCCTCTACTTCTTTTGC	14976		
Genome sequencing 13	SNZ14938F	CAGTGACCGCCGATAATA	14933	1862	This study
	SNZ16817R	CTCACAAATCCCGAATACG	16794		
Genome sequencing 14	SNZ16744F	ATGATGTGGGGTCTAGTTTC	16739	1546	This study
	SNZ18303R	TTCGCTCCTCGCTATACAG	18284		
Genome sequencing 15	SNZ18217F	TCTGCTCTGTATCTGCCG	18214	386	This study
	GP18_18448_Rev	CCAAACTTTGATTGGATTTTGGC	18599		H.J. Maree
5' RACE	SNZ73R	GCAAGGTACAAGGAGGAAAG	55		This study
5' RACE	SNZ106R	CAACGAAACTTAACGAGATT	88		This study
3' Poly(A) tailing	SNZ18505_F	GACATAAACTCTACCTCACGG	18505		This study
3' Poly(A) tailing	GVA-dT(17)	TACGATGGCTGCAGTTTTTTTTTTTTTTTT			Meng et al., 2005
GLRaV-3 group I, II, III diagnostics	LC1	CGCTAGGGCTGTGGAAGTATT		546	Osman and Rowhani, 2006
	LC2	GTTGTCCCGGGTACCAGATAT			
GLRaV-3 group VI diagnostics	LR3-Like (26) For	TAAATGCTCTAGTAGGTATCG		750	H.J. Maree
	LR3-Like (775) Rev	CGAATGTAATCCATGACCTTA			
GLRaV-1 diagnostics	GLRaV-1_419_LQV1-H47	GTTACGGCCCTTTGTTTATTATGG		397	Osman and Rowhani, 2006
	GLRaV-1426_LEV1-C447	CGACCCCTTTATTGTTTGAGTATG			
GLRaV-2 diagnostics	GLRAV-2 rooi F	TATGAGTTCCAACACAAGCGTGC		681	M. Engelbrecht
	GLRaV-2 rooi-R	ACACCGTGCTTAGTACCTCC			

*Reference to isolate GH11

Plant material was collected from 96 randomly selected vines from two separate vineyards (*Vitis vinifera* cvs Merlot and Cabernet Sauvignon, 48 samples each) in the Stellenbosch wine-growing region of South Africa and screened for the presence of GLRaV-3. Total RNA was extracted from the phloem tissue of the 96 plants using a modified cetyltrimethylammonium bromide (CTAB) method (White et al., 2008) and RNA was subjected to RT-PCR using Avian myeloblastosis virus (AMV) reverse transcriptase (Fermentas) and Taq DNA polymerase (KAPA Biosystems).

3.2.3 Sequencing of isolate GH11 and GH30

In order to design primer pairs to sequence the new GLRaV-3 variant, suitable regions in the draft sequence were identified using the read depth and coverage calculated using the MAQ read assembly (Figure 3.1). Fifteen primer pairs were designed to produce overlapping amplicons spanning the GLRaV-3 draft sequence (Table 3.1). Cabernet Sauvignon grapevine material was screened and two plants (GH11 and GH30) were identified to be singly infected with only the new variant of GLRaV-3. Total RNA was extracted from the phloem tissue using a modified CTAB method (White et al., 2008). Avian myeloblastosis virus (AMV) reverse transcriptase (Fermentas) was used for cDNA synthesis from total RNA, and high-fidelity DNA polymerase (Ex Taq, TaKaRa) was used in all PCRs. Amplicons were cloned, and at least three clones from isolate GH11 were sequenced. One clone from each amplicon was sequenced for isolate GH30, except when there were significant differences between GH11 and GH30, in which case two or more additional clones were sequenced. A consensus sequence was generated using Vector NTI 10 (Invitrogen). To determine the 5' terminal sequence of the GH11 and GH30 isolates, two genome-specific nested reverse primers were designed, and total RNA was subjected to RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) using a First-Choice RLM-RACE Kit (Ambion, USA) according to the manufacturer's instructions. The resulting amplicons were cloned and sequenced. Poly(A) tailing on dsRNA with yeast poly(A) polymerase (Affymetrix) was used to determine the 3' end of the genome. Double-stranded RNA was extracted from phloem tissue using an adapted cellulose extraction protocol (Valverde et al., 1990), and cDNA was synthesized using an oligo(dT) primer (Meng et al., 2005). A genome-specific forward primer was designed and used in combination with the oligo(dT) primer in PCR amplifications. Amplicons were cloned and sequenced.

3.2.4 Genome analysis

To identify sequence similarities between isolates, multiple sequence alignments were performed using Bio-Edit 7.0.5.3 (Hall, 1999). A phylogenetic tree was constructed using a multiple alignment of a 415 nt segment of several GLRaV-3 coat protein nt sequences available in GenBank. The tree was constructed by means of the neighbor-joining method with 1000 bootstrap replications using the Molecular Evolutionary Genetics Analysis (MEGA) software program version 5.0 (Tamura et al., 2007). The nt and amino acid (aa) sequence identities for GH11 were compared to other isolates using Vector NTI 10 (Invitrogen). The ORFs were identified using the NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and compared to the known GLRaV-3 ORFs.

3.2.5 ELISA

The indirect ELISA developed for the simultaneous detection of GLRaV-1, GLRaV-2 and GLRaV-3 (Plant Protection Research Institute (PPRI), Agricultural Research Council (ARC)) was used to evaluate if the newly identified variant group of GLRaV-3 can be detected. Twenty GLRaV-3 positive samples were used and subjected to ELISA to evaluate the indirect ELISA kit. These 20 samples were selected from the preliminary survey conducted on *V. vinifera* cvs. Merlot and Cabernet Sauvignon and could be divided into two test groups; samples positive only for the newly identified variant group of GLRaV-3 (11 samples) and samples positive for other GLRaV-3 variants, but not the newly identified variant group (9 samples). The GLRaV-3 positive and negative controls supplied by Vititec (South African commercial vine improvement facility) were also included in the assay. The indirect ELISA was carried out according to the manufacturer's instructions and performed in a microtiter plate. One gram of a representative sample of petioles from each plant was ground in 5 ml of extraction buffer and 100 µl was analyzed in duplicate for each sample. The color development was measured in a microtiter plate spectrophotometer at 405 nm after incubation periods of 15 and 30 minutes. An independent sample t-test was used to calculate if a significant difference in absorption values were observed between the samples positive for the newly identified variant group and the samples positive for the other variant groups using the Statistical Package for the Social Sciences (SPSS) software package 19 (IBM). All twenty samples were screened for GLRaV-1 and GLRaV-2 using RT-PCR since the industry standard ELISA is a simultaneous detection assay for GLRaV-1, GLRaV-2 and GLRaV-3.

3.3 Results and discussion

3.3.1 Draft assembly and preliminary survey

The Velvet assembly resulted in 179 scaffolds with the longest scaffold being 18645 nucleotides (nts) in length. This scaffold (Node 247) was closely related to the NZ-1 sequence according to the NCBI BLAST results and only 65-66% identical to the other variant groups found in South Africa (Table 3.2). The draft sequence was generated from Node 247 by comparing the ORFs to those of other complete genomes of GLRaV-3 and manually removing the insertion-deletion errors. This sequence could only be regarded as a draft, since it was generated using a pooled sample of 44 vines (Coetzee et al., 2010). Furthermore, the parameters used for this draft assembly were out of the normal range and incorporated insertion-deletion errors when compared, in multiple alignments, to other variants. In the MAQ read mapping with the draft sequence as reference, the reads were assembled to an average depth of 1978 reads covering 97% of the draft sequence (Figure 3.1). In order to confirm the existence of this new variant in South Africa, 96 field samples were screened. Fifty-two percent of the samples were infected with the new variant of GLRaV-3, 50% with known variants and 79% with a mixture of variants. Sequencing of the amplicons from the diagnostic RT-PCRs confirmed the draft sequence generated by the *de novo* assembly of the new GLRaV-3 variant.

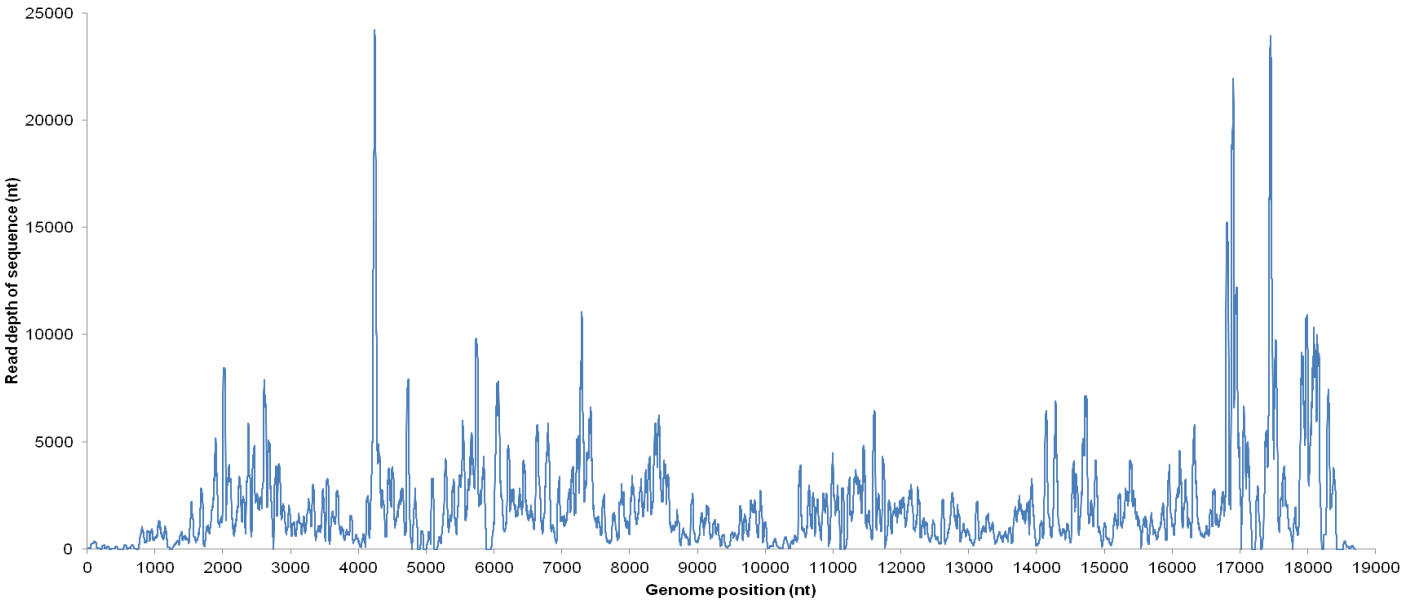

Figure 3.1. Mapping and Assembly with Quality (MAQ) assembly of reads to an average depth of 1978 reads covering 97% of the new South African variant draft sequence.

3.3.2 Sequence and phylogenetic analysis

To construct an accurate genome sequence of the new GLRaV-3 molecular variant, Sanger sequencing was used. *Vitis vinifera* Cabernet Sauvignon grapevine material from a virus isolate collection (Vitis Laboratory, Stellenbosch University, South Africa), maintained in a greenhouse was screened and two plants (GH11 and GH30) were identified to be singly infected with only the new variant of GLRaV-3. The complete genome sequences of GLRaV-3 isolates GH11 and GH30 consist of 18,671 and 18,576 nt, respectively. The complete sequences were deposited in the GenBank database [accessions JQ655295 and JQ655296]. GH11 and GH30 are 99.1% identical at the nt level (Table 3.2), but the 5' UTR of GH30 was found to be 95 nts shorter in length. Both isolates' genomes are larger than the 18,498 nts reported for isolates GP18, 623 and 621, since the region between ORF1b and ORF3 is 188 nts longer. This region has only 44% nt identity when compared to other South African isolates and 87.5% nt identity to the NZ-1 isolate from New Zealand.

Table 3.2. Pairwise nucleotide (nt) sequence comparisons between the original assembled Node 247, the draft sequence used for primer design and the complete nt sequences of representative isolates from the four variant groups known to be present in South Africa.

Sequences	Node 247	Draft	Isolate GH11	Isolate GH30	Isolate 621	Isolate 623	Isolate GP18
Draft	98.66						
Isolate GH11	98.85	98.45					
Isolate GH30	98.44	98.1	99.05				
Isolate 621	66.41	66.76	66.5	66.29			
Isolate 623	66.05	66.58	66.12	65.92	90.97		
Isolate GP18	65.9	66.44	65.96	65.75	90.56	99.04	
Isolate PL-20	65.13	65.5	65.19	64.95	85.89	86.03	85.65

The phylogenetic tree shows the three South African molecular variant groups (Jooste et al., 2010) and the two additional Portuguese groups (Gouveia et al., 2010) described previously (Figure 3.2). Isolates GH11 and GH30 cluster with the isolate identified in New Zealand, NZ-1 (Figure 3.2). There is only 6416 nts of the NZ-1 sequence available, spanning a section of ORF1b (RdRp) up to a section of ORF6 (CP). The NZ-1 sequence is 91% identical to GH11. Recently, GLRaV-3 variants similar to NZ-1 were also identified in vineyards in the Napa Valley (California, USA) (Sharma et al., 2011).

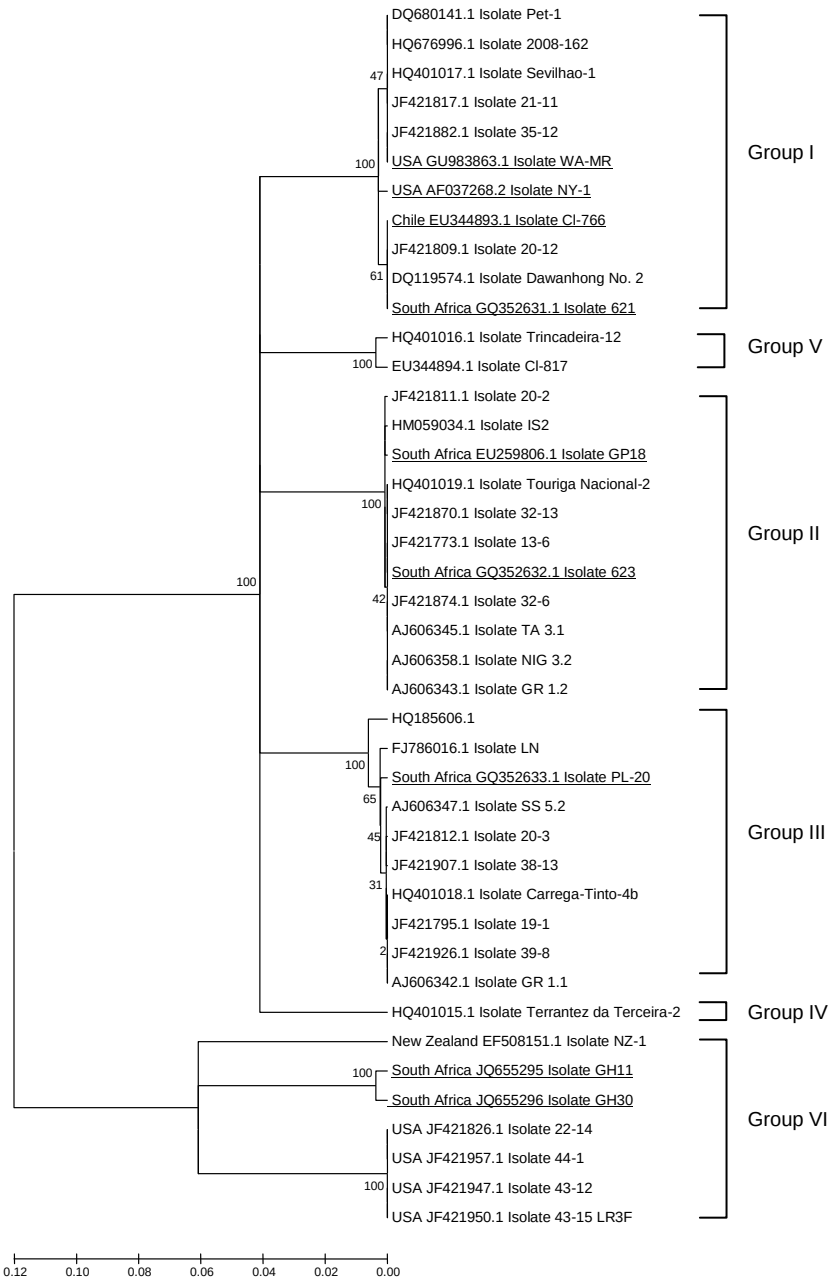

Figure 3.2. Phylogenetic tree constructed from partial nucleotide (nt) sequences of the CP gene of grapevine leafroll-associated virus 3 (GLRaV-3) isolates by neighbor-joining with 1000 bootstrap replicates. Bootstrap values are indicated on branch nodes. The scale bar represents the number of nt substitutions per position. Isolates with complete genome sequences available are underlined. Previously described molecular variant groups are indicated (Gouveia et al., 2010; Jooste et al., 2010) as well as the new proposed group VI which includes the NZ-1 isolate and the Napa Valley isolates.

Due to the high degree of nt sequence identity between GH11 and GH30, further analysis focused on isolate GH11. The genome organization of GH11 for 12 of the 13 ORFs is similar to that of previously described GLRaV-3 isolates (Ling et al., 2004; Maree et al., 2008). The exception is ORF2, where, similar to NZ-1, no ORF could be identified at the appropriate position. However, a small ORF is

predicted between nts 9138 and 9296, which has low nt sequence identity to known variants of GLRaV-3. This poor alignment indicates that there are frameshifts in the GH11 sequence relative to other known variants, resulting in the putative translation of unrelated aa sequence. Due to the low similarity, this small putative ORF cannot be considered a GLRaV-3 ORF2 equivalent. Open reading frame 2 of GLRaV-3 potentially encodes a small peptide, but no counterpart ORFs have been found in other viruses belonging to the family *Closteroviridae* (Ling et al., 1998). The absence of a GLRaV-3 ORF2 in isolate GH11 suggests that the GLRaV-3 ORF2 will probably not harbor a conserved function. It is interesting to note that even though the GH11 ORFs 11 and 12 are predicted at the correct positions, their sequence alignments indicate that they also have frameshifts compared to other known variants. The polypeptide of ORF11 is the same size as those of other variants, but a frameshift changes the sequence from amino acid five onwards to unrelated sequence. The frameshift in GH11 ORF12 results in a premature stop codon and a polypeptide that is six amino acids shorter.

Nucleotide sequence identities between GH11 and known South African isolates ranged from 65.2% to 66.5% when the complete genomes were compared (Table 3.2). This confirms that GH11 is a new molecular variant, and possibly a new strain of GLRaV-3, given that the previously described molecular variants of GLRaV-3 are 91% similar when 621 (group I) is compared to 623 (group II), and 85.9% when 621 is compared to PL-20 (group III) (Table 3.2) (Jooste et al., 2010). Pairwise sequence comparison of GH11 ORFs (excluding ORF2) to other South African isolates showed that GH11 shared between 46.2% and 79.7% nt sequence identity, and between 24.3% and 92% aa identity (Table 3.3). The region with the lowest sequence identity was identified within ORF1a between nts 4583 and 5300 (Figure 3.3). This region has only 34.9% nt sequence identity, resulting in significant aa sequence differences. The sequence was considered accurate, since amplicons were sequenced from different RT-PCR events, and the sequences as a result were not derived from a single clone. The ORF1a sequence of GH11 is also confirmed by the metagenomic draft sequence and the GH30 sequence. The aa differences in ORF1a of GH11 do not affect the conserved domains identified using the NCBI Conserved Domain Database (CDD) (Marchler-Bauer et al., 2008). The viral methyltransferase was identified at aa 456-780, the viral (superfamily 1) RNA helicase at aa 1938-2197, and the 2OG-Fe(II) oxygenase superfamily (AlkB) at aa 1602-1691.

Table 3.3. Positions of untranslated regions (UTRs) and open reading frames (ORFs) on isolate GH11 and pairwise comparisons to the UTRs and ORFs of other South African isolates.

UTR or ORF	Position on the genome (GH11)	% identity to isolate GH11							
		Isolate 621		Isolate 623		Isolate GP18		Isolate PL20	
		nt	aa	nt	aa	nt	aa	nt	aa
5' UTR	1-737	58		55.2		55.1		54.3	
1a	738-7451	67.9	70.7	68.2	70.7	68	70.9	67.9	70.7
1b	7351-9072	73.3	82.6	76.8	84.1	76.5	59.8	52.7	83.4
2?	9138-9296	39.5	8.8	39.4	7.4	40.2	7.4	41.3	9.3
3	10703-10840	73.2	76.1	73.9	78.3	73.9	73.9	72.5	78.3
4	10859-12508	74.7	86.2	75.2	86.4	74.9	85.6	74.4	85.5
5	12501-13952	68.9	72.3	68.7	73.1	68.7	73.1	69.1	73.6
6	14047-14988	79.7	92	78.6	90.1	78.7	90.8	79	89.8
7	15051-16484	71.4	78	71.4	77.4	71.3	78.9	72.9	77
8	16495-17052	76.7	78.5	76.2	77.4	76.2	76.3	76	77.4
9	17049-17582	62.5	55.6	61	54.5	61	56.2	61.3	54.5
10	17589-18128	63	62.8	63.7	63.9	63.7	65	64.3	64.4
11	18131-18241	46.2	24.3	51.7	27	48.3	27	49.6	27
12	18243-18407	56.3	52.5	55.2	54.1	55.7	52.5	60.1	54.1
3' UTR	18408-18671	74.6		72.1		72.1		73.5	

? No ORF could be identified at the appropriate position, but a small ORF is predicted between nucleotides (nts) 9138 and 9296 which has low nucleotide (nt) sequence identity to known variants of GLRaV-3.

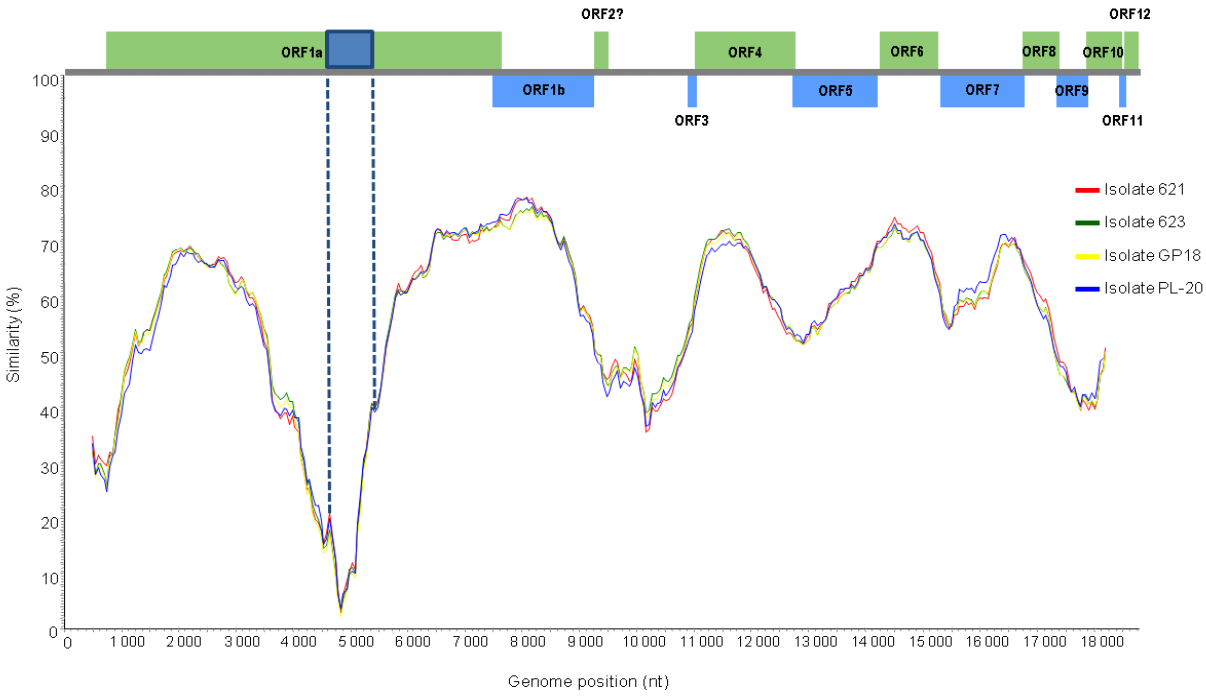

Figure 3.3. Schematic representation of the genome organization of GLRaV-3 GH11 (drawn to scale) illustrating the positions of the untranslated regions (UTRs) and the open reading frames (ORFs). The complete genome of isolate GH11 is also compared at the nucleotide (nt) level to the four other South African isolates (621 (GQ352631.1), 623 (GQ352632.1), GP18 (EU259806.1) and PL-20 (GQ352633.1)) in a similarity plot by constructing a multiple alignment of the five isolates and using a window size of 1000 nucleotides (nts) and a step size of 50 nts (SimPlot, PHYLIP (Pylogeny Inference Package) v3.5). The region with the lowest sequence identity was found within ORF1a between nts 4583-5300 and is indicated.

3.3.3 ELISA

From the 96 grapevine samples screened for different GLRaV-3 variants using RT-PCR, 11 samples were singly infected with only the newly identified variant group (test group 1) and 9 samples were positive with other variants of GLRaV-3, but not the newly identified variant group (test group 2). These 20 samples were screened for GLRaV-1 and GLRaV-2 using RT-PCR and subjected to the simultaneous indirect ELISA screening for GLRaV-1, 2 and 3. The 20 samples were negative for GLRaV-1 and GLRaV-2 according to the RT-PCR screenings and therefore a positive result with the ELISA could only be due to a GLRaV-3 infection. The yellow color development could be observed in all 20 samples and all samples were regarded as positive due to the absorption values being greater than two times the negative control's absorption value. Normal distribution was assumed for both test groups' data points for the 15 (reading 1) and 30 minutes (reading 2) incubation period according to the Shapiro-Wilk algorithm (Table 3.4). Consequently, an independent sample t-test was used to calculate if a significant difference was observed between the two test groups. The average absorption value for each sample was used for all analyses. Equal variances were assumed for readings 1 and 2 since the Levene's test for equality of variances was not significant (Table 3.4). Both p-values of the independent sample t-test of reading 1 and 2 were greater than 0.05 and consequently no significant difference could be observed between the newly identified variant group and other variant groups of GLRaV-3 (Table 3.4). It is thus concluded that the standard ELISA used for most of the industry GLRaV-3 routine diagnostics and for the screening of nuclear-stock grapevine material is efficient to detect all known variant groups of GLRaV-3.

3.4 Conclusion

In this study, we sequenced two isolates, GH11 and GH30, of a new variant of GLRaV-3. These isolates are less than 70% similar to other known GLRaV-3 variants, suggesting that they should be considered variants of a different strain of GLRaV-3. We propose that the GLRaV-3-like virus identified in this study be grouped together with NZ-1 and the Napa Valley isolates (Sharma et al., 2011) as Group VI of GLRaV-3 (Fig. 2). The draft sequence generated with bioinformatic tools was 98% identical to the GH11 sequence generated using Sanger sequencing (Table 3.2). This is further evidence that next-generation sequencing is a valuable approach to identify novel viruses and variants. This study also found that the industry standard ELISA is an effective GLRaV-3 diagnostic method and will be able to detect all known variant groups of GLRaV-3. However, this assay is not able to

differentiate between GLRaV-3 variant groups and a variant identification RT-PCR needs to be established for that purpose.

Table 3.4. Statistical analysis of the indirect enzyme-linked imunosorbent assay (ELISA) developed for the simultaneous detection of GLRaV-1, GLRaV-2 and GLRaV-3 (Plant Protection Research Institute, Agricultural Research Council) to evaluate if the newly identified variant group of GLRaV-3 can be detected.

Reading 1: 15 min incubation						
Sample	GLRaV-3 variant group	Average absorption at 405 nm	Average of test group	Shapiro-Wilk test of normality (p)	Levene's test for equality of variances (p)	Independent sample t-test (p)
M2.1	Group VI	0.423	0.435	0.325	0.224	0.213
M7.20		0.323				
C22.33		0.500				
C23.17		0.325				
C24.20		0.462				
C24.24		0.444				
C25.29		0.526				
C26.6		0.444				
C26.19		0.510				
C27.9		0.454				
C27.25		0.379				
M10.12	Group I, II or III	0.475	0.470	0.413		
M11.6		0.443				
M11.9		0.470				
M12.2		0.422				
M13.7		0.450				
M14.16		0.565				
M17.12		0.505				
M21.8		0.418				
C19.20		0.483				
Positive		0.331				
Negative		0.087				
Reading 2: 30 min incubation						
Sample	GLRaV-3 variant group	Average absorption at 405 nm	Average of test group	Shapiro-Wilk test of normality (p)	Levene's test for equality of variances (p)	Independent sample t-test
M2.1	Group VI	0.744	0.766	0.410	0.214	0.260
M7.20		0.551				
C22.33		0.880				
C23.17		0.563				
C24.20		0.808				
C24.24		0.776				
C25.29		0.941				
C26.6		0.782				
C26.19		0.912				
C27.9		0.808				
C27.25		0.665				
M10.12	Group I, II or III	0.818	0.824	0.471		
M11.6		0.772				
M11.9		0.825				
M12.2		0.735				
M13.7		0.796				
M14.16		0.993				
M17.12		0.885				
M21.8		0.731				
C19.20		0.863				
Positive		0.577				
Negative		0.094				

Significant level at $p<0.05$

3.5 References

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., Boonham, N., 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. Mol. Plant. Pathol. 10(4), 537-545.

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.

Charles, J.G., Cohen, D., Walker, J.T.S., Forgie, S.A., Bell, V.A., Breen, K.C., 2006. A review of grapevine leafroll-associated virus type 3 (GLRaV-3) for the New Zealand wine industry. Report to New Zealand Winegrowers. HortResearch Client Report. 18447, 1-79.

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

Gouveia, P., Santos, M.T., Eiras-Dias, J.E., Nolasco, G., 2010. Five phylogenetic groups identified in the coat protein gene of grapevine leafroll-associated virus 3 obtained from Portuguese grapevine varieties. Arch. Virol. 156(3), 413–420.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Jooste, A.E.C., 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. Arch. Virol. 155, 1997–2006.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., 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(1), 1-7.

Li, H., Ruan, J., Durbin, R., 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. 18(11), 1851-1858.

Ling, K.S., Zhu, H.Y., Drong, R.F., Slightom, J.L., Gonsalves, D., 1998. Nucleotide sequence of the 39-terminal two-thirds of the grapevine leafroll-associated virus-3 genome reveals a typical monopartite closterovirus. J. Gen. Virol. 79, 1299–1307.

Ling, K.S., Zhu, H.Y., Gonsalves, D., 2004. Complete nucleotide sequence and genome organization of grapevine leafroll-associated virus 3, type member of the genus Ampelovirus. J. Gen. Virol. 85, 2099–2102.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., 2008. CDD: specific functional annotation with the conserved domain database. Nucleic Acids Res. 37, 205–210.

Maree, H.J., 2010. Identification and characterisation of Grapevine leafroll-associated virus 3 genomic and subgenomic RNAs. PhD thesis, Stellenbosch University.

Maree, H.J., Freeborough, M-J., Burger, J.T., 2008. Complete nucleotide sequence of a South African isolate of grapevine leafroll-associated virus 3 reveals a 5' UTR of 737 nucleotides. Arch. Virol. 153, 755–757.

Martelli, G.P., Agranovsky, A.A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R.H.A., Dolja, V.V., Falk, B.W., Gonsalves, D., Jelkmann, W., Karasev, A.V., Minafra, A., Namba, S., Vetten, H.J., Wisler, G.C., Yoshikawa, N., 2002. The family Closteroviridae revised. Arch. Virol. 147(10), 2039–2044.

Meng, B., Li, C., Goszczynski, D.E., Gonsalves, D., 2005. Genome sequences and structures of two biologically distinct strains of Grapevine leafroll-associated virus 2 and sequence analysis. Virus Genes. 31, 31–41.

Osman, F., Rowhani, A., 2006. Application of a spotting sample preparation technique for the detection of pathogens in woody plants by RT-PCR and real-time PCR (TaqMan). J. Virol. Methods. 133, 130-136.

Pietersen,G., 2004. Spread of grapevine leafroll disease in South Africa - a difficult, but not insurmountable problem. Wynboer. http://www.wynboer.co.za/recentarticles/articles.php3.

Sharma, A.M., Wang, J., Duffy, S., Zhang, S., Wong, M.K., Rashed, A., Cooper, M.L., Daane, K.M., Almeida, R.P.P., 2011. Occurence of grapevine leafroll-associated virus complex in Napa valley. PLoS ONE. 6(10), DOI:10.1371/journal.pone.0026227.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Valverde, R.A., Nameth, S.T., Jordan, R.L., 1990. Analysis of double-stranded-RNA for plant-virus diagnosis. Plant Dis. 74(3), 255–258.

Ward, E., Foster, S.J., Fraaije, B.A., McCartney, H.A., 2004. Plant pathogen diagnostics: immunological and nucleic acid-based approaches. Ann. Appl. Biol. 145, 1-16.

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. Biotechnol. J. 3(11), 1424–1428.

Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18(5), 821–829.

Chapter 4: Real-time RT-PCR HRM curve analysis and multiplex RT-PCR to detect and differentiate between GLRaV-3 variant groups I, II, III and VI

4.1 Introduction

Grapevine leafroll-associated virus 3 (GLRaV-3) is a positive-sense single-stranded RNA virus that is the type member of the genus *Ampelovirus* in the family *Closteroviridae* (Martelli et al., 2002). This virus is phloem-limited and is considered the main contributing agent of leafroll disease worldwide with detrimental effects on both wine and table grapes. Six variant groups of GLRaV-3 have been identified of which four are known to be present in South Africa (Bester et al., 2012; Engel et al., 2008; Gouveia et al., 2010; Jarugula et al., 2010; Jooste et al., 2010; Ling et al., 2004; Maree et al., 2008). The genomes of at least one representative isolate of variant groups I, II, III and VI have been sequenced. These are isolates 621, WA-MR NY-1 and CI-766 (group I) (Engel et al., 2008; Jarugula et al., 2010; Jooste et al., 2010; Ling et al., 2004), 623 and GP18 (group II) (Jooste et al., 2010; Maree et al., 2008), and PL-20 (group III) (Jooste et al., 2010). Recently, isolates GH11 and GH30 (group VI), were identified, and showed less than 70% nucleotide (nt) identity to other GLRaV-3 variant groups (Bester et al., 2012). Limited sequence information for GLRaV-3 variant groups IV and V is available and isolates from these groups are only represented by coat protein gene sequences in the GenBank database (Gouveia et al., 2010). All these genetic variants commonly occur as mixed infections. However, no specific disease symptoms or geographic distribution could so far be assigned to a specific variant group or cluster of variant groups. It is therefore necessary to develop an effective method that can detect all GLRaV-3 variants and differentiate between them.

Previously, single-strand conformation polymorphism (SSCP) profiles have been used to investigate the population structure and genetic variability of GLRaV-3 variants (Jooste et al., 2010; Turturo et al., 2005). Although SSCP analysis is fast and cost-effective for variant typing based on sequence heterogeneity, the technique is not as sensitive as RT-PCR and requires sequencing to verify new variants. Metagenomic sequencing or next-generation sequencing is the most sensitive diagnostic tool available to detect and identify known and novel viruses (Adams et al., 2009; Al Rwahnih et al., 2009; Coetzee et al., 2010; Kreuze et al., 2009). Next-generation sequencing can identify viral pathogens occurring at extremely low titers without the necessity of any prior sequence knowledge. Although this

technique is unbiased, it is still too expensive to use for routine diagnostics. Reverse transcription polymerase chain reaction (RT-PCR) is a diagnostic tool capable of detecting virus sequences at low concentrations and can be designed to be genus-, species-, or isolate-specific (Charles et al., 2006; Ward et al., 2004). The design of optimal RT-PCR primers requires accurate sequence information. The recently sequenced GLRaV-3 group VI was found to be less than 70% similar to other GLRaV-3 variant groups and warrants a re-evaluation of existing GLRaV-3 RT-PCR diagnostic primers.

Real-time RT-PCR is another technique that has been successfully utilized to detect various plant viruses, including GLRaV-3 (Osman et al., 2007). It is a rapid, reliable and quantitative detection method that is more sensitive than conventional RT-PCR. It has the potential for multiplexing and is therefore able to detect several pathogens in the same reaction (Osman et al., 2007). The development of high-resolution melting (HRM) curve analysis, as an extension to real-time RT-PCR, provides a rapid, high-throughput, cost effective and single tube approach to discriminate and genotype strains of bacteria and viruses (Gori et al., 2010). The genotyping of variants does not require a labeled probe and sequence variants can be distinguished from each other based on their individual melting temperatures (Reed et al., 2007). High-resolution melting curve analysis was effectively applied in diagnostics for viruses affecting humans (Toi and Dwyer, 2008) as well as for phytopathogenic bacteria (Gori et al., 2010).

This chapter describes the development of a simple and reliable one-step real-time RT-PCR assay with HRM curve analysis (RT-PCR HRM) for the simultaneous detection and identification of GLRaV-3 variants of groups I, II, III and VI, all four previously detected in South African vineyards. To achieve this, a universal primer set, able to detect and differentiate these variant groups, was designed. A multiplex RT-PCR was also developed to validate the RT-PCR HRM. The application of these protocols will aid in the understanding of the molecular epidemiology of GLRaV-3 and leafroll disease and assist programmes focused at managing and controlling the spread of GLRaV-3.

4.2 Materials and methods

4.2.1 Virus source and sample preparation

Plant material from 173 grapevine plants was used to establish and validate the RT-PCR HRM. Forty vines from a study in 2008, where the distribution of GLRaV-3 variants in disease clusters were investigated, were re-collected from a vineyard in the Worcester vine growing region (Jooste et al.,

2011). Ninety grapevine plants were randomly selected during a field survey in 2008 from two severely infected vineyards in the Stellenbosch area and 39 grapevine samples were from a virus isolate collection (Vitis Laboratory, Stellenbosch University, South Africa), maintained in *V. vinifera*, grown in the greenhouse. An additional GLRaV-3 positive sample for each variant group, singly infected with only that variant (Group I, II, III and VI), was obtained from a virus isolate collection (ARC-Plant Protection Research Institute, Pretoria, South Africa). Phloem scrapings were prepared from cane material collected during winter. Total RNA was extracted from 2.5 g phloem tissue using an adapted Cetyltrimethylammonium bromide (CTAB) method (2% CTAB, 2.5% PVP-40, 100 mM Tris-HCL pH8, 2 M NaCl, 25 mM EDTA pH8 and 3% β-mercaptoethanol) (White et al., 2008).

4.2.2 Primer design

Conserved regions in the GLRaV-3 genome were used to design primer pairs that are able to detect the four GLRaV-3 variant groups found in South Africa. These conserved regions had to be in close proximity to result in amplicons with lengths of 150-300 base pairs (bp). Representative isolates of GLRaV-3 variant groups with complete genome sequences available [GenBank: GQ352631.1, GenBank: EU259806.1, GenBank: GQ352632.1, GenBank: GQ352633.1, GenBank: JQ655295, GenBank: GU983863.1, GenBank: AF037268.2, GenBank: EU344893.1] were used to identify the conserved regions by constructing a multiple sequence alignment using BioEdit 7.0.5.3 (Hall, 1999). The partial isolate NZ-1 sequence was also included in the multiple sequence alignment [GenBank: EF508151.1]. Six primers pairs were identified targeting ORF1a, ORF1b, ORF4 and ORF6 (Table 4.1). The six primers pairs were tested on samples singly infected with a specific GLRaV-3 variant group using the real-time RT-PCR to identify which primer can most effectively detect all variants and possibly differentiate between them by using HRM curve analysis.

4.2.3 Verification of one-step real-time RT-PCR assay with melting curves generated from plasmid DNA

Real-time RT-PCR amplicons of GLRaV-3 variant groups I, II, III and VI were cloned into a pGEM-T-easy Vector (Promega) and sequenced to obtain variant-specific plasmid DNA. Artificial *in vitro* mixed infections between the variant-specific plasmid DNA were made to determine whether the chosen primer pair could differentiate between variants if mixed infections would be present in field plants. Duplex infections were made in a 1:3, 1:1 and 3:1 ratio for each combination of two variant groups. Reaction mixtures of all variant-specific plasmid DNA PCR HRM assays contained 1x KAPA Taq

Buffer A (KAPA Biosystems), 0.4 µM reverse primer (IDT), 0.4 µM forward primer (IDT), 0.2 mM dNTP mix (Fermentas), 1 µM SYTO 9 (Invitrogen), 0.04 U/µl KAPA Taq DNA polymerase (KAPA Biosystems) and 0.01 ng plasmid DNA. Cycle conditions included an initial denaturation step at 94°C for 5 minutes, followed by 45 cycles of 94°C for 10 seconds, annealing at 55°C for 10 seconds and elongation at 72°C for 20 seconds. Acquisition on the green channel was recorded at the end of the extension step. High-resolution melting curves of PCR amplicons were obtained with temperatures ranging from 70°C to 90°C with a 0.1°C increase in temperature every two seconds.

4.2.4 Real-time RT-PCR and HRM analysis

The primer pair that could most effectively detect and differentiate between GLRaV-3 variant groups I, II, III and VI was used to screen the 173 samples to optimize the assay. Each reaction was performed in duplicate using the RT-PCR HRM on a Qiagen Rotor-Gene Q thermal cycler. Reaction mixtures contained 1x KAPA Taq Buffer A (KAPA Biosystems), 0.4 µM reverse primer (IDT), 0.4 µM forward primer (IDT), 0.2 mM dNTP mix (Fermentas), 1 µM SYTO 9 (Invitrogen), 0.04 U/µl KAPA Taq (KAPA Biosystems), 0.08 U/µl Avian Myeloblastosis Virus (AMV) reverse transcriptase (Fermentas) and 100 ng RNA. Optimized cycle conditions were a cDNA synthesis step at 48°C for 30 minutes, an initial denaturation step at 94°C for 5 minutes, followed by 45 cycles of 94°C for 10 seconds, annealing at 55°C for 10 seconds and elongation at 72°C for 20 seconds. Acquisition on the green channel was recorded at the end of the extension step. High-resolution melting curves of PCR amplicons were obtained with temperatures ranging from 70°C to 90°C with a 0.1°C increase in temperature every two seconds. HRM curve analysis was performed using the Rotor-Gene software version 1.7. In order to use the RT-PCR HRM to differentiate between variants, a melting point confidence interval had to be determined for each variant group. The data generated for each variant group were tested for normality using the Shapiro-Wilk algorithm and descriptive statistics were calculated using the Statistical Package for the Social Sciences (SPSS) software package 19 (IBM).

Table 4.1. List of primers used with the real-time RT-PCR HRM assay and the end-point multiplex RT-PCR protocol.

Primer pair	Sequence (5'-3')	Target region	Amplicon size (bp)
LR3.HRM1.F	TAGACGTTAAAGATGTGAAGCG	GLRaV-3 ORF1a	167
LR3.HRM1.R	TCGTACACATCCACCATA		
LR3.HRM2.F	GTCCTAGATTCGGATTTTGTCG	GLRaV-3 ORF1a	231
LR3.HRM2.R	GAATACTCTTCGCCCTATC		
LR3.HRM3.F	CTGGTTGCTTTCGAGGTATATGAG	GLRaV-3 ORF1b	295
LR3.HRM3.R	CACTTCAAGGTGTTGCGCTT		
LR3.HRM4.F	TAATCGGAGGTTTAGGTTCC	GLRaV-3 ORF4	226
LR3.HRM4.R	GTCGGTTCGTTAACAACAC		
LR3.HRM5.F	TGTGTAAGAAGGTTATGGG	GLRaV-3 ORF6	224
LR3.HRM5.R	TACTGCCTTACCGGGTTTTC		
LR3.HRM6.F	GTCACCAGGTGTTCCAAACC	GLRaV-3 ORF1a	305
LR3.HRM6.R	AACGCCCTGTATGTCCTCTC		
LR3_Universal_F	TAAATGCTCTAGTAGGATTC	GLRaV-3 5' UTR	
621_430R	TAACCCAACACGACGATGAG	GLRaV-3 5' UTR	429[a]
623_564R	CTCACGCTAAACACACCAAG	GLRaV-3 5' UTR	563[a]
PL20_315R	GTTTGTAACAAAGAAACACG	GLRaV-3 5' UTR	314[a]
GH11_180R	CCAAAACGAAGACGAAAAGAAGAG	GLRaV-3 5' UTR	179[a]
LR_ORF1aR	CGTCCGCTTCACCCCTTTGG	GLRaV-3 ORF1a	
Vv_Actin_F[b]	CTTGCATCCCTCAGCACCTT	*V. vinifera* predicted actin-7	82
Vv_Actin_R[b]	TCCTGTGGACAATGGATGGA		

[a]Amplicon size if used together with LR_Universal_F

[b] Reid et al., 2006

4.2.5 Variant status conformation using multiplex RT-PCR

Variant-specific RT-PCR reverse primers (Table 4.1) targeting the 5' UTR of the GLRaV-3 variant groups I, II, III and VI were designed to be used in a single reaction with one forward primer. This multiplex RT-PCR was designed to validate the HRM analysis and assign each sample to a specific variant group. A primer pair targeting the *V. vinifera* actin gene was also included in the multiplex RT-PCR to act as an RNA-specific internal control. A two-step RT-PCR multiplex protocol was used and approximately 1000-1500 ng of total RNA was denatured at 65°C for 5 minutes with 2 µM of LR_ORF1aR primer (IDT) and 2 µM of Vv_Actin_R (IDT) (Reid et al., 2006) (Table 4.1) and incubated for 2 minutes on ice (5 µl final volume). The RNA was reverse-transcribed by incubation at 48°C for 1 h in a reaction mixture (10 µl final volume) containing 1x Avian Myeloblastosis Virus (AMV) reverse transcriptase buffer (Fermentas), 1 mM dNTP mix (Fermentas), 1U/µl Ribolock (Fermentas) and 0.5 U/µl AMV reverse transcriptase (Fermentas). A 2.5 µl aliquot of cDNA was subjected to PCR in a 25 µl reaction mixture containing 1x KAPA Taq buffer B (KAPA Biosystems), 0.4 mM dNTP mix (Fermentas), 0.4 µM LR_universal_F primer (IDT), 0.28 µM Vv_Actin F (IDT)

(Reid et al., 2006), 0.28 µM Vv_Actin R (IDT) (Reid et al., 2006), 0.4 µM of each variant-specific reverse primer (IDT) (Table 4.1), 0.5µg/µl Bovine Serum Albumin (BSA) (Roche) and 0.08 U/µl KAPA Taq DNA polymerase (KAPA Biosystems). Cycle conditions included an initial denaturation step at 94°C for 5 minutes, followed by 35 cycles of 94°C for 30 seconds, annealing at 58°C for 20 seconds and elongation at 72°C for 40 seconds. Final extension was at 72°C for 7 minutes. Amplicons were visualized on an ethidium bromide-stained 2% TAE-agarose gel (2 M Tris, 1M glacial acetic acid, 0.05M Na_2EDTA, pH 8).

4.3 Results and discussion

4.3.1 Primer design and evaluation

Six primer pairs were evaluated for their ability to detect and differentiate between GLRaV-3 variant groups I, II, III and VI, utilizing the RT-PCR HRM (Figure 4.1). From the six primer pairs evaluated for the RT-PCR HRM, primer pairs LR3.HRM1 and LR3.HRM3 were eliminated since they were unable to differentiate between groups III and VI, and groups I and II, respectively (Figure 4.1A and Figure 4.1C). The amplification efficiency for group II variants by primer pairs LR3.HRM2 and LR3.HRM5 were sub-optimal (Figure 4.1B and Figure 4.1E). Only primer pairs LR3.HRM4 (Figure 4.1D) and LR3.HRM6 (Figure 4.1F) showed reproducible results, with primer pair LR3.HRM4 yielding equal amplification for all variant groups, based on electrophoretic analysis. Primer pair LR3.HRM4 (Figure 4.1D) produced a single PCR product of 226 bp for each variant group when visualized on a 1.5% TAE agarose gel (Figure 4.2). After HRM curve analysis the LR3.HRM4 primer pair produced one melting peak each for group I and II variants with average melting points of 83.60°C and 83.77°C, respectively (Table 4.2). Variant groups III and VI both produced a major peak together with a smaller shoulder peak. The average melting points of the major melt peak for groups III and VI were 85.44°C and 85.97°C, respectively (Table 4.2). The shoulder peaks produced by the LR3.HRM4 primer pair for groups III and VI were not regarded as unspecific amplification since only one band was detected after gel electrophoresis. Therefore, it is likely that the shoulder peaks can be the result of uneven G/C distribution throughout the targeted RNA of groups III and VI (Dujic et al., 2011; Jeffery et al., 2007). The last 126 nucleotides (nts) at the 3’ end of the 226 bp amplicon for groups III and VI have an average GC content of 53% and 56% respectively, compared to the 47.5% and 46% for groups I and II. No discriminatory difference could be detected between the average melting points of groups I and II (Table 4.2). Pairwise nt sequence comparisons showed that there are only up to 11 nt differences between GLRaV-3 group I and group II within the targeted region, whereas the other variant groups

had 24-61 nt differences (Table 4.3). Although primer pair LR3.HRM6 was unable to detect group VI variants, it could efficiently differentiate between group I and II variants (Figure 4.1F). It produced a single melting peak on a derivative melting curve for both variant groups I and II with average melting points of 85.03°C and 86.41°C, respectively (Table 4.2). Sequence analysis performed on the Hsp70h gene sequences available on GenBank, spanning the LR3.HRM4 primer pair target region, indicates that the LR3.HRM4 primer pair will be able to detect all variants from groups I, II, III and VI. Unfortunately groups IV and V could not be included in this study as only coat protein sequences of these variant groups are available.

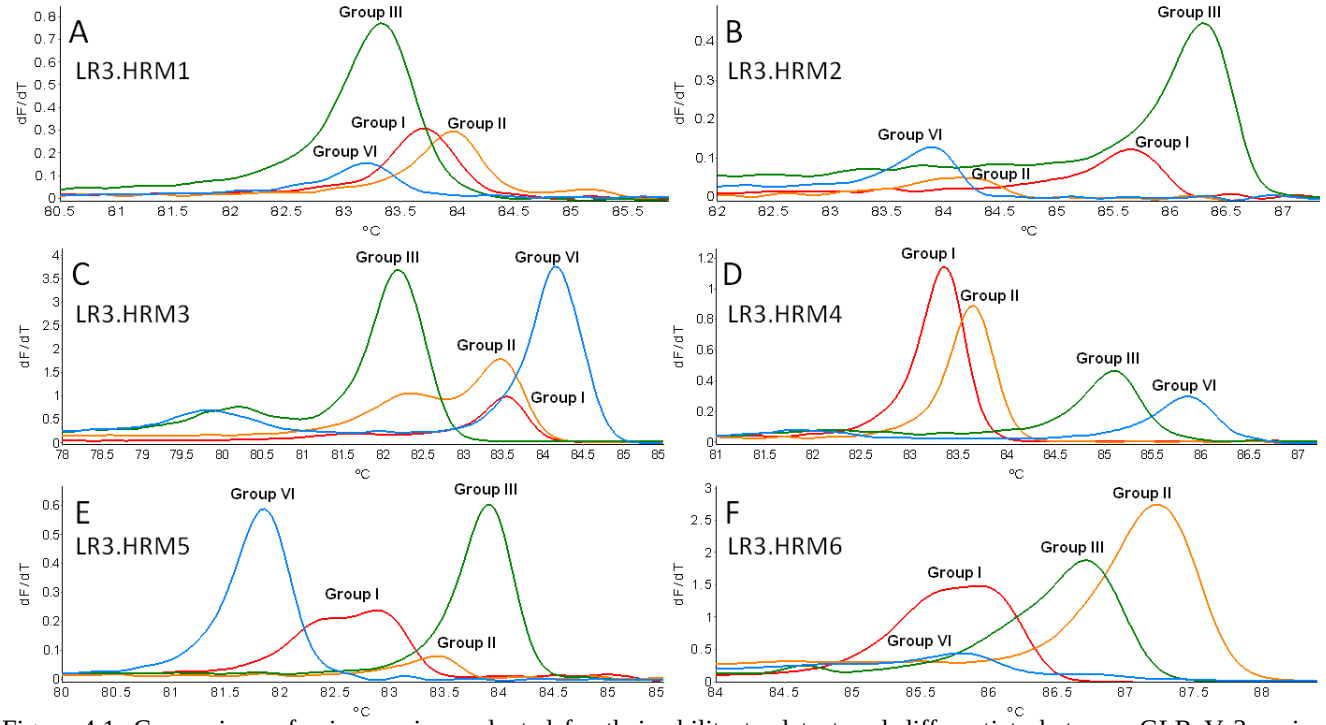

Figure 4.1. Comparison of primer pairs evaluated for their ability to detect and differentiate between GLRaV-3 variant groups. Derivative HRM curves (dF/dT) obtained using RNA extracted from plants singly infected with only one variant group of GLRaV-3 in the real-time RT-PCR HRM assay.

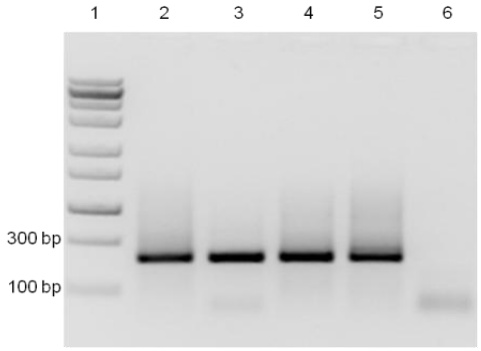

Figure 4.2. Agarose gel electrophoresis of RT-PCR HRM amplicons. Visualization of amplicons generated with the real-time RT-PCR HRM assay with primer pair LR3.HRM4, separated on a 1.5% TAE agarose gel with ethidium bromide staining. Lane 1: Fermentas Zipruler Express DNA ladder 2, Lane 2: Group I, Lane 3: Group II, Lane 4: Group III, Lane 5: Group VI, Lane 6: RNA negative control.

Table 4.2. Descriptive statistics of melting points generated by real-time RT-PCR HRM assays with primer pairs LR3.HRM4 and LR3.HRM6.

Variant group	Number of data points[a]	Min	Max	Mean	Temperature range between upper and lower limit
LR3.HRM4					
Group I	31	83.20	83.98	83.60	0.78
Group II	203	83.15	84.25	83.77	1.10
Group III	73	84.87	85.70	85.44	0.83
Group VI	142	85.30	86.37	85.97	1.07
LR3.HRM6					
Group I	27	84.78	85.42	85.03	0.64
Group II	187	85.95	86.90	86.41	0.95

[a]Number of melting point temperatures generated per variant group from the 121 samples. More than one melting point temperature per sample was generated due to mix infections and duplex reactions.

Table 4.3. Pairwise comparison of LR3.HRM4 amplicon (226 nucleotide (nt) segment of Hsp70h) for each variant group. The upper comparison in bold is the number of nucleotide (nt) differences between variant sequences and the lower comparison is the percent identity (%) between variant sequences.

		Variant group representative isolates									
		1	2	3	4	5	6	7	8	9	10
Group I_GU983863.1_GLRaV-3_Isolate_WA-MR	1		**0**	**1**	**1**	**9**	**10**	**24**	**50**	**50**	**53**
Group I_AF037268.2_GLRaV-3_Isolate_NY-1	2	100		**1**	**1**	**9**	**10**	**24**	**50**	**50**	**53**
Group I_EU344893.1_GLRaV-3_Isolate_CI-766	3	99.56	99.56		**2**	**9**	**10**	**24**	**50**	**50**	**53**
Group I_GQ352631.1_GLRaV-3_Isolate_621	4	99.56	99.56	99.12		**10**	**11**	**25**	**49**	**49**	**52**
Group II_GQ352632.1_GLRaV-3_Isolate_623	5	96.02	96.02	96.02	95.58		**2**	**27**	**49**	**49**	**53**
Group II_EU259806.1_GLRaV-3_Isolate_GP18	6	95.58	95.58	95.58	95.13	99.12		**27**	**51**	**51**	**53**
Group III_GQ352633.1_GLRaV-3_Isolate PL-20	7	89.38	89.38	89.38	88.94	88.05	88.05		**58**	**58**	**61**
Group VI_JQ655295_GLRaV-3_Isolate GH11	8	77.88	77.88	77.88	78.32	78.32	77.43	74.34		**0**	**10**
Group VI_JQ655296_GLRaV-3_Isolate_GH30	9	77.88	77.88	77.88	78.32	78.32	77.43	74.34	100		**10**
Group VI_EF508151.1_GLRaV-3_Isolate_NZ-1	10	76.55	76.55	76.55	76.99	76.55	76.55	73.01	95.58	95.58	

4.3.2 Verification of one-step real-time RT-PCR HRM assay

Variant-specific plasmids containing the amplicons from primer pairs LR3.HRM4 and LR3.HRM6 were constructed. The derivative HRM curves (dF/dT) and normalized HRM curves generated by using the variant-specific plasmid DNA in the real-time PCR HRM assay (PCR HRM) (Figure 4.3) verified the relative positions of the melting curves observed when singly infected plant RNA samples were screened (Figure 4.1D and Figure 4.1F).

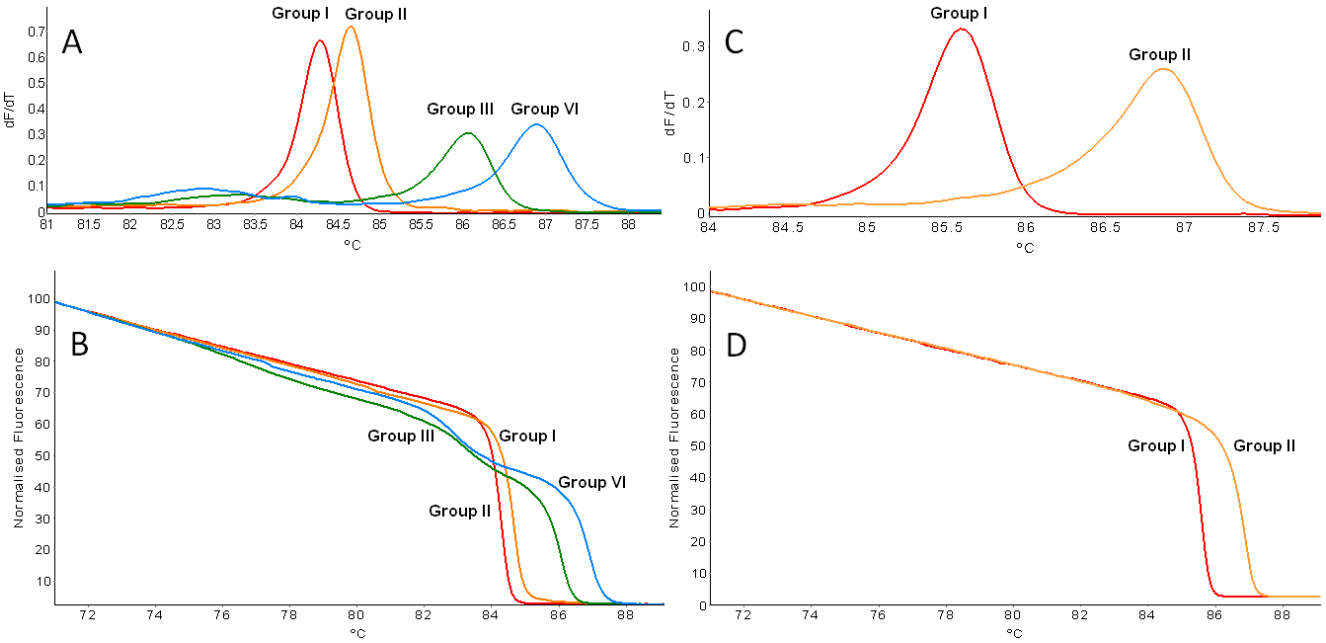

Figure 4.3. High-resolution melting (HRM) curve analysis using variant-specific plasmid DNA in real-time PCR HRM assays. Derivative HRM curves (dF/dT) (A and C) and normalized HRM curves (B and D) obtained using SYTO 9 for the detection of GLRaV-3 variants. Primer pair LR3.HRM4 is represented by A and B and primer pair LR3.HRM6 is represented by C and D.

Duplex artificially mixed infections between the variant-specific plasmid DNA confirmed that primer pair LR3.HRM4 can differentiate between mixed infections (Figure 4.4B-F), except for mixed infections of variants from groups I and II (Figure 4.4A). The duplex artificial mix with the variant-specific plasmid DNA of groups I and II illustrated that a single melting peak is produced on the derivative melting curve (Figure 4.4A). This melting peak was not distinguishable from either the singly infected group I or II melting peaks, based on the confidence intervals calculated. In order to classify a sample as group I and/or II it was concluded that an additional RT-PCR HRM assay with primer pair LR3.HRM6 is necessary. Primer pair LR3.HRM6 could differentiate between variants of groups I and II based on the duplex artificial mixed infection analysis (Figure 4.4G).

4.3.3 Real-time RT-PCR and HRM analysis

One hundred and sixty nine grapevine samples were screened with the LR3.HRM4 primer pair of which 48 samples tested negative for GLRaV-3. From the remaining 121 samples positive for GLRaV-3, 35 were positive for group III variants and 87 positive for group VI variants. One hundred and two samples were positive for group I and/or group II variants. These 102 samples were screened with the LR3.HRM6 primer pair to determine their variant status and 14 samples were found to be infected with

group I variants and 88 samples infected with group II variants. Of the 121 GLRaV-3 positive samples, 73 samples had multiple infections (Table 4.4).

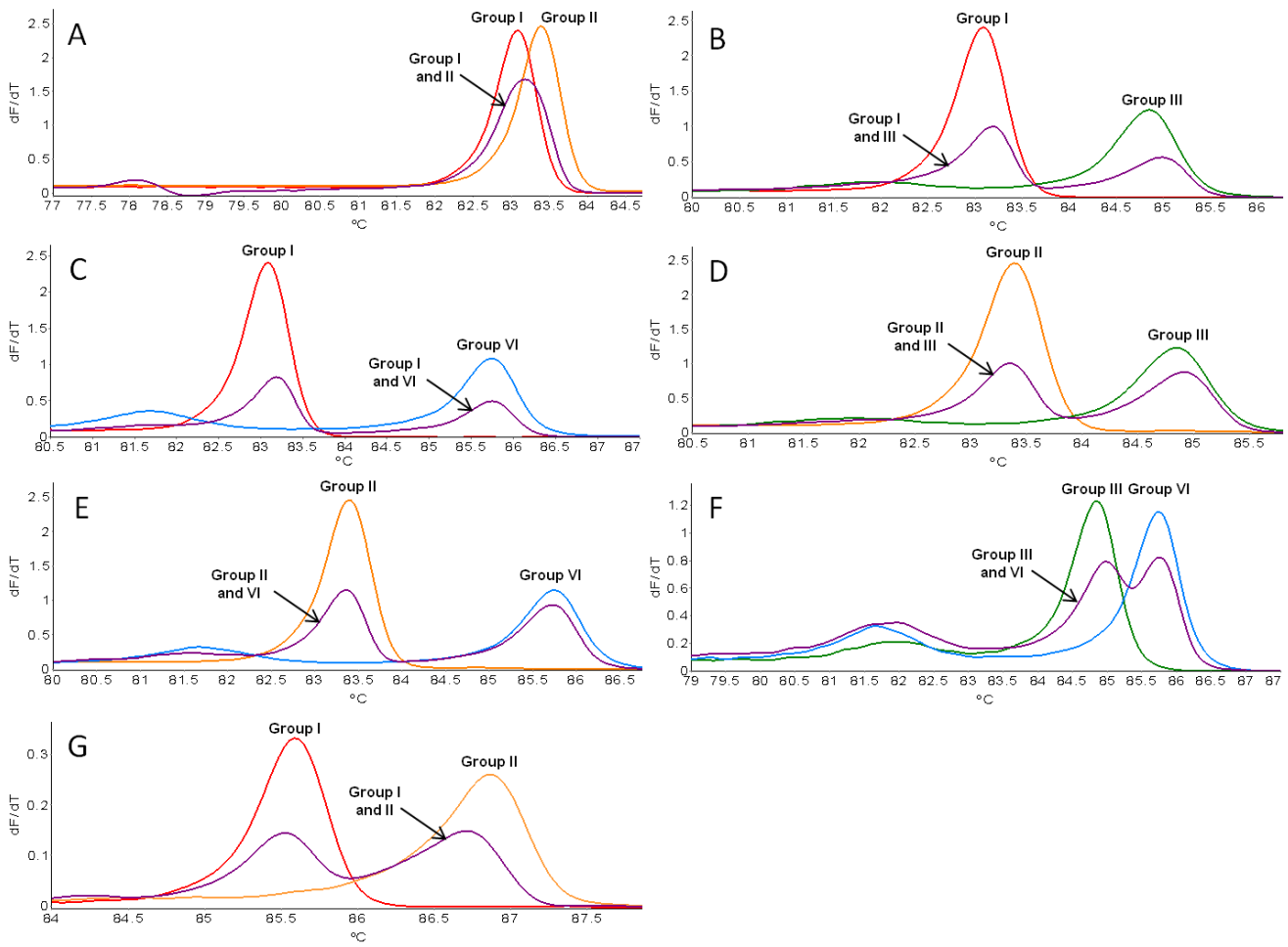

Figure 4.4. Variant-specific plasmid DNA duplex infections. Comparison of the different duplex infections possible between variant groups I, II, III and VI. Derivative HRM curves (dF/dT) obtained using primer pairs LR3.HRM4 (A-F) and LR3.HRM6 (G) in real-time PCR HRM assays using variant specific plasmid DNA. All mixed infections shown are the 1:1 duplex artificial mix compared to melting curves generated from single variant reactions.

In order to use RT-PCR HRM with the LR3.HRM4 primer pair for variant differentiation, a confidence interval for each variant group's melting point was determined. The sample melting points for groups II, III and VI were not normally distributed (Table 4.5), resulting in less than 95% of the melting points to fall within the interval ±1.96 standard deviations from the mean. All RT-PCR HRM reactions were performed in duplicate and yielded consistent melting points per sample, however, significant variation was observed between samples from the same variant group. This can probably be explained by the existence of quasispecies that arose from the high mutation rate of the viral genome (Holland et al., 1982; Vignuzzi et al., 2006). An average melting point was therefore not adequate to differentiate GLRaV-3 variant groups. A melting point temperature interval was consequently calculated to include

95% of the melting points observed. The 2.5th and 97.5th percentiles for each variant group (Table 4.5) were used to calculate the limits of the interval to include 95% of the data. The largest possible interval for each variant group was also determined to include the highest number of melting points for each variant group without overlapping with the adjacent interval (Table 4.5). Intervals where all data points (100% confidence) fell within the maximum range, the limits were adjusted to the 2.5th to 97.5th percentile to incorporate a margin of error to ensure accurate classification. The confidence intervals of groups I and II overlapped almost completely and therefore differentiation was not possible. The group I and II intervals could not be separated and a joint interval from 83.22°C to 84.18°C (95% confidence) was calculated for samples from variant groups I and/or II. For groups III and VI the intervals were calculated as 84.57°C to 85.64°C (95.89% confidence) and 85.65°C to 86.37°C (92.96% confidence), respectively. To differentiate groups I and II, primer pair LR3.HRM6 was used. The melting points of both groups I and II were also not normally distributed and the confidence intervals were calculated using the 2.5th to 97.5th percentile range. The group I interval was calculated from 84.79°C to 85.39°C (95% confidence) and for group II from 86.01°C to 86.78°C (95% confidence). Outliers were identified within variant groups III and VI for primer pair LR3.HRM4 and within variant group I for primer pair LR3.HRM6 (Table 4.5). The comparatively high number of outliers identified within variant group VI resulted in a lower confidence level for this variant group compared to the other groups.

Table 4.4. Analysis of grapevine leafroll-associated virus 3 (GLRaV-3) single and mixed variant group infections.

		Variant group	Number of infections
Single infections	48	I	7
		II	17
		III	0
		VI	24
Mixed infections	73	I+II	0
		I+III	0
		I+VI	1
		II+III	10
		II+VI	33
		III+VI	1
		I+II+III	0
		I+II+VI	4
		II+III+VI	22
		I+II+III+VI	2
Total	121		

Table 4.5. Calculation of the melting point confidence interval for each variant group based on real-time RT-PCR HRM curve analysis using LR3.HRM4 or LR3.HRM6 primer pairs.[a]

Variant group	2.5th percentile[b]	97.5th percentile[c]	Interquartile range (IQR) (75%-25%)[d]	Number of outliers (>±1.5xIQR)	Shapiro- Wilk test of normality (p)[e]	Melting point interval without overlaps		
								Confidence (%)
LR3.HRM4								
Group I	**83.22**	84.08	0.43	0.00	0.103	83.20	83.70	67.74
Group II	83.22	**84.18**	0.45	0.00	0.000	83.15	84.56	100
Group III	84.91	85.65	0.13	6.00	0.000	**84.57**	**85.64**	95.89
Group VI	85.35	86.28	0.15	12.00	0.000	**85.65**	**86.37**	92.96
LR3.HRM6								
Group I	**84.79**	**85.39**	0.09	7.00	0.002	84.78	85.69	100
Group II	**86.01**	**86.78**	0.42	0.00	0.000	85.70	86.90	100

[a]The data generated for each variant group was tested for normality in order to calculate the largest interval with the highest confidence without overlaps between variant groups. These intervals are indicated in bold. Intervals where all data points (100% confidence) fell within the maximum range, the limits were adjusted to the 2.5th to 97.5th percentile to incorporate a margin of error to ensure accurate classification.
[b]2.5th percentile is the melting point temperature where 2.5% of data points is less than or equal to that temperature.
[c]97.5th percentile is the melting point temperature where 2.5% of data points is greater than or equal to that temperature.
[d]Interquartile range is the interval where the middle 50% of melting point temperatures can be expected.
[e]Assume a normal distribution if p>0.05, meaning approximately 95% of melting point temperatures of the variant group will be within ±1.96 standard deviations of the mean.

The Rotor-Gene software can perform automated variant classification based on the melting point interval calculated from the derivative melting curve (dF/dT) profile for each sample. Bins were programmed based on the data set for each variant group that consisted of a calculated midpoint with a 95% confidence interval width. This allows the software to automatically classify each melting peak observed, according to the bins programmed. To avoid unnecessary peak calling, the temperature threshold can be set at 83°C, because none of the variant groups is expected to have a melting point below 83°C. These confidence intervals for both primer pairs LR3.HRM4 and LR3.HRM6 were calculated, based on data generated from RNA extracted using the CTAB method. It was observed that when a different RNA extraction protocol was used, the melting points for each variant group shifted proportionally (unpublished data). This is probably the result of the interaction of the intercalating SYTO 9 dye which is influenced by inhibitors and salt concentration in the RNA extract.

In this study, preliminary data on the incidence of GLRaV-3 variants in the Western Cape of South Africa were collected using RT-PCR HRM analysis. A previous study, using SSCP, identified variant group II as the most prevalent, with 54% of a sample size of 80 being infected by this variant (Jooste et al., 2011). In the present study, variant groups II and VI were equally distributed with a 39% infection rate each. Of the 224 infections detected in 121 positive samples, 21% were single variant infections, with half of these classified as group VI. These preliminary data are not necessarily an indication of the distribution of GLRaV-3, since more than half of the grapevine samples came from only three severely

infected vineyards and the rest were from greenhouse isolate collections that decreases the complexity of mix infections. However, it confirms the presence of four GLRaV-3 variant groups in South Africa and that the technique can successfully be applied to study the distribution of GLRaV-3 variants.

4.3.4 Variant status confirmation using multiplex RT-PCR

The multiplex RT-PCR was optimized to detect GLRaV-3 variant groups I, II, III and VI in a single reaction (Figure 4.5). Two reverse primers targeting GLRaV-3 ORF1a and the *V. vinifera* actin gene, respectively, were used for the cDNA synthesis. The PCR was optimized to produce a single amplicon for each variant group and the internal control. The reaction was tested with and without the addition of bovine serum albumin (BSA), but without BSA the amplification was sub-optimal for GLRaV-3 variant groups I and II. The addition of BSA has previously been shown to enhance the amplification efficiency of targeted DNA by stabilizing enzymes and neutralizing inhibitory contaminants (Kreader, 1996; Nagai et al., 1998; Ralser et al., 2006).

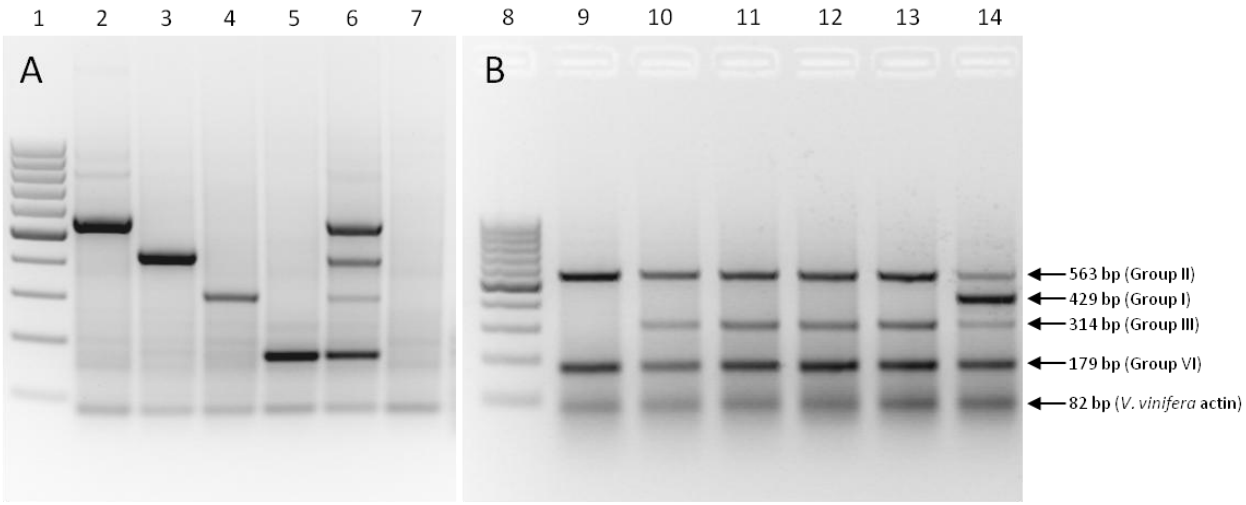

Figure 4.5. Agarose gel electrophoresis of multiplex RT-PCR amplicons. Visualization of multiplex RT-PCR amplicons separated on a 2% TAE agarose gel with ethidium bromide staining. Figure 4.5A represents grapevine samples singly infected with one GLRaV-3 variant group and Figure 4.5B represents field samples with multiple infections. Lane 1: Fermentas GeneRuler 100 bp DNA ladder, Lane 2: Group II variant (563 bp) and *V. vinifera* actin (82 bp), Lane 3: Group I variant (429 bp) and *V. vinifera* actin, Lane 4: Group III variant (314 bp) and *V. vinifera* actin, Lane 5: Group VI variant (179 bp) and *V. vinifera* actin, Lane 6: Group I, II, III and VI variants and *V. vinifera* actin, Lane 7: Negative control, Lane 8: Fermentas GeneRuler 100 bp DNA ladder, Lane 9: Group II, VI and *V. vinifera* actin, Lane 10: Group II, III, VI and *V. vinifera* actin, Lane 11: Group II, III, VI and *V. vinifera* actin, Lane 12: Group II, III, VI and *V. vinifera* actin, Lane 13: Group II, III, VI and *V. vinifera* actin, Lane 14: Group I, II, III, VI variants and *V. vinifera* actin.

One hundred and twenty one GLRaV-3 positive samples were screened using the multiplex RT-PCR protocol. Thirteen of these samples were positive for group I variants, 87 samples positive for group II variants, 32 samples positive for group III samples and 80 samples positive for group VI variants. The

multiplex RT-PCR validated 94% of the infections detected by the combined LR3.HRM4 and LR3.HRM6 RT-PCR HRM assays, indicating that the RT-PCR HRM is more sensitive than the multiplex RT-PCR. This is not unexpected, because of the specificity of the instrument used and the primer target region selected. The multiplex RT-PCR was designed to target the 5' UTR of the GLRaV-3 genome due to the high variability in this region. Insertions and deletions in this region made it an ideal target for the design of variant-specific primers. The 5'UTR is only represented in the genomic RNA whereas the Hsp70h is also represented in sub-genomic RNAs produced during GLRaV-3 replication. This implies an increased number of templates for the 3' half of the genome (Jarugula et al., 2010; Maree et al., 2010), making the Hsp70h region a better-suited target for viral diagnostics by improving sensitivity. Another advantage of the RT-PCR HRM is that it will be possible to identify a new variant group if a distinct melting curve profile is produced. With the multiplex RT-PCR a new variant group will remain undetected or unidentified if the primers are also specific for the new variant.

4.4 Conclusion

In order to investigate the spread and impact of different GLRaV-3 variants in vineyards, sensitive diagnostic techniques are a necessity. A serological test like ELISA is one of the preferred detection methods for plant viral disease diagnostics due to its simplicity and effectiveness (Ling et al., 2001). However, as viral sequences become available, virus-specific primers can be designed to be used in RT-PCR or real-time RT-PCR that is more sensitive than serological tests. In this study, a real-time RT-PCR was designed that is able to detect GLRaV-3 variant groups I, II, III and VI, using a single primer pair targeting the Hsp70h gene of GLRaV-3. If HRM curve analysis is added to the real-time RT-PCR, it is possible to differentiate between variant groups based on three melting point intervals. An additional primer pair was identified that is able to differentiate between variant groups I and II. The RT-PCR HRM assay provides a more sensitive and rapid tool to detect and differentiate between different GLRaV-3 variant groups. The multiplex RT-PCR offers an end-point PCR alternative to differentiate between the variant groups present in South Africa or to be used as a validation method for the RT-PCR HRM. The abovementioned tools will contribute to the understanding of the pathogenesis of leafroll disease and aid epidemiology studies to investigate how these different GLRaV-3 variant groups are spreading.

4.5 References

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., Boonham, N., 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. Mol. Plant. Path. 10(4), 537-545.

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.

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. Arch. Virol. 157, 1815-1819.

Charles, J.G., Cohen, D., Walker, J.T.S., Forgie, S.A., Bell, V.A., Breen, K.C., 2006. A review of grapevine leafroll-associated virus type 3 (GLRaV-3) for the New Zealand wine industry. Report to New Zealand Winegrowers. HortResearch client report. 18447, 1-79.

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

Dujic, T., Ostanek, B., Marc, J., Causevic, A., Minar, B., 2011. Optimization of high-resolution melting analysis simultaneous genotyping of two 11β-hydoxysteroid dehydrogenase type 1 gene polymorphisms. Genet. Test. Mol. Bioma. 15(1-2), 43-49.

Engel, E.A., Girardi, C., Escobar, P.F., Arredondo, V., Dominguez, C., Perez-Acle, T., Valenzuela, P.D., 2008. Genome analysis and detection of a Chilean isolate of Grapevine leafroll-associated virus-3. Virus Genes. 37(1), 110-118.

Gori, A., Cerboneschi, M., Tegli, S., 2010. High-resolution melting analysis as a powerful tool to discriminate and genotype pseudomonas savastanoi pathovars and strains. PLoS one. 7(1), 1-13.

Gouveia, P., Santos, M.T., Eiras-Dias, J.E., Nolasco, G., 2010. Five phylogenetic groups identified in the coat protein gene of grapevine leafroll-associated virus 3 obtained from Portuguese grapevine varieties. Arch. Virol. 156(3), 413-420.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95-98.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S., 1982. Rapid evolution of RNA genomes. Science. 215, 1577-1585.

Jarugula, S., Gowda, S., Dawson, W.O., Naidu, R.A., 2010. 3'-coterminal subgenomic RNAs and putative cis-acting elements of grapevine leafroll-associated virus 3 reveals 'unique' features of gene expression strategy in the genus Ampelovirus. Virol. J. 7(180), 1-14.

Jeffery, N., Gasser, R.B., Steer, P.A., Noormohammadi, A.H., 2007. Classification of mycoplasma synoviae strains using single-strand conformation polymorphism and high-resolution melting-curve analysis of the vlhA gene single-copy region. Microbiology. 153, 2679-2688.

Jooste, A.E.C., 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. Arch. Virol. 155, 1997–2006.

Jooste, A.E.C., Pietersen, G., Burger, J.T., 2011. Distribution of grapevine leafroll-associated virus 3 variants in South African vineyards. Eur. J. Plant. Pathol. 131, 371-381.

Kreader, C.A., 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. Appl. Environ. Microbiol. 62(3), 1102-1106.

Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., 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(1), 1-7.

Ling, K-S., Zhu, H-Y., Gonsalves, D., 2004. Complete nucleotide sequence and genome organization of grapevine leafroll-associated virus 3, type member of the genus Ampelovirus. J. Gen. Virol. 85, 2099–2102.

Ling, K.S., Zhu, H.Y., Petrovic, N., Gonsalves, D., 2001. Comparative Effectiveness of ELISA and RT-PCR for detecting grapevine leafroll-associated closterovirus-3 in field samples. Am. J. Enol. Viticult. 52(1), 21-27.

Maree, H.J., Freeborough, M-J., Burger, J.T., 2008. Complete nucleotide sequence of a South African isolate of grapevine leafroll-associated virus 3 reveals a 5' UTR of 737 nucleotides. Arch. Virol. 153, 755-757.

Maree, H.J., Gardner, H.F.J., Freeborough, M-J., Burger, J.T., 2010. Mapping of the 5' terminal nucleotides of grapevine leafroll-associated virus 3 sgRNAs. Virus Res. 151:252-255.

Martelli, G.P., Agranovsky, A.A., Bar-Joseph, M., Boscia, D., Candresse, T., Coutts, R.H.A., Dolja, V.V., Falk, B.W., Gonsalves, D., Jelkmann, W., Karasev, A.V., Minafra, A., Namba, S., Vetten, H.J., Wisler, G.C., Yoshikawa, N., 2002. The family Closteroviridae revised. Arch. Virol. 147(10), 2039–2044.

Nagai, M., Yoshida, A., Sato, N., 1998. Additive effects of bovine serum albumin, dithiothreitol, and glycerol on PCR. Biochem. Mol. Biol. Int. 44(1), 157-163.

Osman, F., Leutenegger, C., Colino, D., Rowhani, A., 2007. Real-time RT-PCR (TaqMan®) assays for the detection of grapevine leafroll-associated viruses 1–5 and 9. J. Virol. Methods. 141, 22–29.

Ralser, M., Querfurth, R., Warnatz, H.J., Lehrach, H., Yaspo, M.L., Krobitsch, S., 2006. An efficient and economic enhancer mix for PCR. Biochem. Biophys. Res. Commun. 347, 747–751.

Reed, G.H., Kent, J.O., Wittwer, C.T., 2007. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. Pharmacogenomics. 8(6), 597-608.

Reid, K.E., Olsson, N., Schlosser, J., Peng, F., Lund, S.T., 2006. An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. BMC Plant Biol. 6(27), DOI: 10.1186/1471-2229-6-27.

Toi, C.S., Dwyer, D.E., 2008. Differentiation between vaccine and wild-type varicella-zoster virus genotypes by high-resolution melting analysis of single nucleotide polymorphisms. J. Clin. Virol. 43, 18-24.

Turturo, C., Salderelli, 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. J. Gen. Virol. 86, 217–224.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R., 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature. 439, 344-348.

Ward, E., Foster, S.J., Fraaije, B.A., McCartney, H.A., 2004. Plant pathogen diagnostics: immunological and nucleic acid-based approaches. Ann. Appl. Biol. 145, 1-16.

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. Biotechnol. J. 3(11), 1424-1428.

Chapter 5: Conclusion

5.1 Conclusion and future prospects

Virus-associated diseases of grapevines constitute a major limiting factor to the development and security of the world viticultural industry. Grapevine leafroll disease (GLD) has been reported in most grapevine-growing countries in the world and has a significant effect on the profitability of vineyards over the lifespan of a vineyard if no intervention strategies are implemented (Freeborough and Burger, 2008). Even though other ampeloviruses have been associated with GLD, grapevine leafroll-associated virus 3 (GLRaV-3) is considered to be the main causative agent since it is most often associated with typical leafroll disease symptoms and damage. The studies by Jooste et al. (2010) and Gouveia et al. (2010) identified five molecular variant groups for GLRaV-3 of which three are present in South Africa. A metagenomic sequencing study provided the viral profile of a severely diseased vineyard with GLRaV-3 (group II) as the dominant virus present in the sample (Coetzee et al., 2010). The metagenomic study also identified a potential novel GLRaV-3 variant. The first aim of this study was therefore, to confirm this new molecular variant of GLRaV-3, using conventional Sanger sequencing.

The sequencing data from the metagenomic study (Coetzee et al., 2010) was used to assemble scaffolds resembling complete genome sequences of GLRaV-3. The longest scaffold was closely related to the NZ-1 sequence according to the NCBI BLAST results. This sequence could only be regarded as a draft, since it was generated using a pooled sample of 44 vines (Coetzee et al., 2010). The *de novo* assembly of next-generation sequencing reads most likely only result in chimeric scaffolds consisting of a range of templates that are closely related. Even if this chimeric scaffold was not a direct representative of the new variant group, it still provided a sequence that was closely related to the new variant. Two plants (GH11 and GH30) were identified to be singly infected with only the new variant and used to sequence the complete genome of the new molecular variant. The complete genome sequences of GLRaV-3 isolates GH11 and GH30 consist of 18,671 and 18,576 nt, respectively. GH11 and GH30 are 99.1% identical at the nucleotide (nt) level, but the 5' UTR of GH30 was found to be 95 nucleotides (nts) shorter in length. Samples infected with the new variant of GLRaV-3 with 5' UTRs of one or the other or both lengths were identified and confirmed this anomaly between isolates GH11 and GH30. The 5' UTR is highly variable between GLRaV-3 molecular variants and the 5' UTR of isolate GH11 is only between 54.3-58% similar to other South African GLRaV-3 variants. Even though the 5' UTR diversity

is high amongst variant groups, the 5' UTRs are conserved within a particular variant group, providing an ideal region for variation studies amongst variants. The significance of the extended 5' UTR of GLRaV-3 is not yet known and future studies are needed to determine the possible function of this region. The role of recombination in the emergence of new GLRaV-3 variants and possibly the existence of the 5' UTR length difference between GH11 and GH30 needs to be investigated.

Phylogenetic analysis on a partial sequence of the coat protein gene confirmed the five molecular variant groups identified previously (Gouveia et al., 2010; Jooste et al., 2010), with isolates GH11 and GH30 clustering separately. Isolates GH11 and GH30 are less than 70% similar to other known GLRaV-3 variants, suggesting that they should be considered variants of a different strain of GLRaV-3. The draft sequence compiled with bioinformatic tools was 98% identical to the GH11 sequence generated using Sanger sequencing, providing evidence that next-generation sequencing is a valuable approach to identify novel viruses and variants. No ORF2 could be identified at the appropriate position for GH11 and GH30 and although ORFs 11 and 12 of GH11 and GH30 are predicted at the correct positions, their sequence alignments showed low nt identity compared to other known variants. The fact that no counterpart ORFs have been found in other viruses belonging to the family *Closteroviridae*, suggests that GLRaV-3 ORF2, 11 and 12 will probably not harbor conserved functions. Additional research and whole-genome sequencing of viral isolates from other regions of the world will have to be compared with these identified in the current study to make further conclusions regarding these ORFs. The successful construction of an infectious clone of GLRaV-3 will provide a platform to study viral replication and ORF expression to determine the function of ORF2, 11 and 12.

The study also evaluated the industry standard enzyme-linked immunosorbent assay (ELISA) (ARC-PPRI, Pretoria) for its effectiveness to detect the newly identified variant of GLRaV-3. It was confirmed that this standard ELISA is an effective GLRaV-3 diagnostic method and detects all known variant groups of GLRaV-3, present in South Africa. This is important to the South African viticulture industry, since grapevine certification schemes rely on this ELISA to certify grapevine material as GLRaV-3 free. This ELISA can only detect GLRaV-3 infection and is not able to differentiate between variant groups. For this purpose, this study aimed to establish a simple and reliable one-step real-time RT-PCR assay with high-resolution melting (HRM) curve analysis (RT-PCR HRM) for the simultaneous detection and identification of GLRaV-3 variants of groups I, II, III and VI. The existence of quasispecies complicated the calculation of an average variant-specific melting point to differentiate

GLRaV-3 variant groups. A melting point temperature interval was calculated to include ±95% of the melting points observed without overlapping with the adjacent intervals. These intervals will become more refined as more samples are screened. An end-point multiplex RT-PCR was also developed to validate the RT-PCR HRM and to offer an end-point PCR alternative to differentiate between the variant groups present in South Africa. It will be interesting to investigate if the RT-PCR HRM assay can be expanded to include the detection of groups IV and V as more sequence data become available. Presently, the RT-PCR HRM relies on the extraction of high quality RNA, but if the assay can be optimized to work on crude plant extracts, it will also reduce the time and technical skills needed for this assay. The application of the abovementioned protocols will aid in the understanding of the molecular epidemiology of GLRaV-3 to investigate the distribution of variants and possible associations of specific GLRaV-3 variants to geographic distribution, cultivar, or specific disease symptoms. Presently all the techniques used routinely to detect a GLRaV-3 infection, including the techniques of this study, are all dependent on the virus titre and low virus titre associated with new infections and seasonal variation, can lead to false negative results. It may be valuable to design a quantitative real-time RT-PCR assay for GLRaV-3, to assess the efficacy of the abovementioned RT-PCR HRM assay to detect a low virus titre infection. An assessment of the plant's response to a GLRaV-3 infection, independent of the virus titre, might also prove to be an alternative detection method for GLRaV-3. Infection of most viruses in plants leads to the modulation of gene expression and is regulated by the presence of endogenous, small RNA (sRNA) molecules (Singh et al., 2012). Therefore, by investigating the plant's response to a GLRaV-3 infection through comparative sRNA profiling and differentially targeted genes can provide a marker signature that can possibly lead to earlier diagnosis of the disease and as a result improve disease management strategies.

This study does not necessarily provide an indication of the distribution of GLRaV-3 variants in South Africa, since the grapevine samples were collected from a limited amount of severely infected vineyards and from greenhouse isolate collections that decreased the complexity of mixed infections. However, it confirms the presence of four GLRaV-3 variant groups in South Africa and that the RT-PCR HRM assay can be valuable to study the distribution of GLRaV-3 variants. The abundance of the group II and VI variants observed compared to the other two variant groups, may either be the result of the accidental planting of infected plant material in vineyards, or the variant fitness that increases their frequency in a population. Previous studies have reported the dominant occurrence of group II variants

in South Africa (Coetzee et al., 2010; Jooste et al., 2010; Jooste et al., 2011), but whether this is still the situation after the discovery of group VI, warrants further investigation.

In order to determine if even more new variants exist, or if groups IV and V are present in South Africa, would require the mining of next-generation sequencing data for potential new variants and more whole-genome sequencing of viral isolates from South Africa and other regions of the world. It will also be interesting to continue the search for GLRaV-3 variants not only in wine grape cultivars, but also in table grape varieties. A comparative study between the GLRaV-3 variant composition of different cultivars (wine versus table grapes or white versus red) can provide valuable information regarding the distribution of variants. It will also be of importance to investigate the interaction between the mealybug vector and specific GLRaV-3 variants to determine if mealybug vectors transmit specific GLRaV-3 variants more efficiently.

5.2 References

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

Freeborough, M.J., Burger, J.T., 2008. Leafroll: Economic implications. Wynboer, http://www.wynboer.co.za/recentarticles/200812-leafroll.php3

Gouveia, P., Santos, M., Eiras-Dias, J., Nolasco, G., 2011. Five phylogenetic groups identified in the coat protein gene of grapevine leafroll-associated virus 3 obtained from Portuguese grapevine varieties. Arch. Virol. 156, 413-420.

Jooste, A.E.C., Maree, H.J, Bellstedt, D., 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. Arch. Virol. 155, 1997-2006.

Jooste, A.E.C., Pietersen, G., Burger, J.T., 2011. Distribution of grapevine leafroll-associated virus-3 variants in South African vineyards. Eur. J. Plant. Pathol. 131, 371-381.

Singh, K., Talla, A., Qiu, W., 2012. Small RNA profiling of virus-infected grapevines: evidences for virus infection-associated and variety-specific miRNAs. Funct. Integr. Genomics. DOI: 10.1007/s10142-012-0292-1.