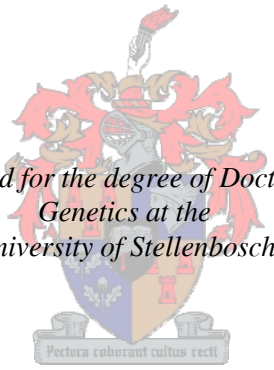
Identification and molecular characterization of three genetic variants of Grapevine leafroll-associated virus 3 (GLRaV-3) from South African vineyards and their spread in local vineyards

by

Anna Elizabeth Catharina Jooste



Dissertation presented for the degree of Doctor of Philosophy in Genetics at the University of Stellenbosch

Promoter: Prof. J.T. Burger
Co-promoter: Dr. D.E. Goszczynski
Faculty of Science
Department of Genetics

March 2011

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2011

Abstract

Grapevine diseases, in particular virus and virus-like diseases, are threatening grapevine industries worldwide; also in South Africa. Grapevine leafroll (GLR) is one of the most important diseases of grapevines, occurring in all grape-producing countries worldwide. Grapevine leafroll-associated virus 3 (GLRaV-3) is known to be closely associated with GLR disease and occurs commonly in South African vineyards. In this study three genetic variants of GLRaV-3 were identified in vineyards of the Western Cape, South Africaby single strand conformation polymorphism (SSCP) profiles generated from a region amplified in ORF5. A specific SSCP profile could be assigned to each variant group and these wereconfirmed by sequencing of the ORF5 regions.These results demonstrated that SSCP analysis on this region in ORF5 provides a fast and reliable indication of the GLRaV-3 variant status of a plant, which in many instances showed mixed infections. The full genome sequence of one representative of each variant group *i.e.* isolates 621 (group I), 623 (group II) and PL-20 (group III), was determined by sequencing overlapping cloned fragments of these isolates. The sequences of genomic 5' ends of these isolates were determined by RLM-RACE. Sequence alignment of the 5'UTRs indicated significant sequence and length variation in this region, between the three South African variant groups. Nucleotide sequence alignment of the Hsp70h and CP gene regions of these isolates with those of isolates from elsewhere in the world, followed by phylogenetic analysis, further supported the presence of three GLRaV-3 variants in South Africa, and that two or three additional variant groups occurs elsewhere in the world. We further investigated the prevalence of these three GLRaV-3 variants in mother blocksof different cultivars and from different vine growing regions, using SSCP analysis. The majority of the plants studied, were infected with the group II variant, similar to isolates 623 and GP18. The distribution of the three GLRaV-3 variants within a spatio-temporally recorded cluster of diseased plants was studied by means of SSCP profile analysis. We showed that different GLRaV-3 variants are transmitted to adjacent plants in an infection cluster. Results showed that, in some leafroll disease clusters, the variant that was present in the original GLRaV-3 infected plant of a cluster was transmitted to adjacent plants in a row and across rows. Some plants in the cluster were also infected with variants not present in the original plant. These infections could have been caused by mealybug vectors feeding on plants from surrounding areas and then infecting these plants.

The scientific information generated on GLRaV-3 variants in this project contributed to the advancement of our knowledge of genetic variability and provides a basis of further

epidemiology and vector-virus studies. The study showed for the first time that different GLRaV-3 variants were transmitted to adjacent plants in a row and across rows in a GLR disease cluster. The diversity detected in the 5'UTR between variants from the three genetic groups provides a platform for the further study of the biological characteristics of GLRaV-3 variants.

Opsomming

Wingerdsiektes, veral virus siektes, bedreig wingerd industrieë wêreldwyd, asook die Suid Afrikaanse wingerdbedryf. Rolbladsiekte is een van die belangrikste siektes op wingerd en kom wêreldwyd voor. Die virus, grapevine leafroll-associated virus 3 (GLRaV-3), word sterk geassosieer met Rolbladsiekte en kom wydverspreid voor in Suid Afrikaanse wingerde. Tydens hierdie studie is drie genetiese variante van GLRaV-3 geïdentifiseer in wingerd moederblokke in die Wes-Kaap. Die GLRaV-3 variante is geïdentifiseer met 'n tegniek wat 'single-strand conformation polymorphism (SSCP)' genoem word. Die SSCP profiele was gegenereer vanaf PKR produkte van die ORF5 area op die genoom van GLRaV-3. Die geamplifiseerde produk van die ORF5 gebied is gebruik om die SSCP profiele te verkry en DNA-volgorde data in die gebied het die drie SSCP profiele gestaaf. Hierdie metode om virus variasie te bestudeer in plante is vinnig en betroubare resultate is verkry. Gemengde infeksies, wat gereeld in wingerd voorkom, kon ook met die tegniek opgespoor word. Die volledige nukleotied-volgorde van elkeen van die drie GLRaV-3 genome is volledig bepaal. Die isolate wat die drie variant groepe verteenwoordig is isolaat 621 (groep I), 623 (groep II) en PL-20 (groep III). Die nukleotiedvolgorde in die 5'UTR is bepaal met die RLM-RACE tegniek. Wanneer die 5'UTRs van die drie variante vergelyk is, het dit getoon dat daar verskille is in die volgordes en lengtes voorgekom het. Ander dele van die genoom, o.a. die dopproteïen (CP) en Hsp70 areas, is filogeneties vergelyk met isolate van regoor die wêreld. In die filogenetiese analise is bevind dat die drie GLRaV-3 variante saamgegroepeer het met ander isolate in die wêreld en dat daar elders ook twee to drie addisionele variant groepe van GLRaV-3 voorkom. Die verspreiding van die drie GLRaV-3 variante in wingerde is bestudeer in verskillende kultivars en in verskillende verbouingsgebiede. Die meerderheid van die plante in die studie was geïnfekteer met die groep II variant wat dieselfde is as isolate 623 en GP18. Die voorkoms van die drie variante in 'n siekte cluster is bestudeer d.m.v SSCP. Die studie het gewys dat verskillende GLRaV-3 variante versprei word na aangrensende plante in 'n ry en tussen rye. In sommige gevalle is die variant wat in die oorspronklik geïnfekteerde plant voorkom, oorgedra na naasliggende plante. Sommige van die plante in the infeksie area was ook met ander GLRaV-3 variante geïnfekteer wat moontlik deur wolluise oorgedra is vanaf naburige geïnfekteerde plante.

Die wetenskaplike inligting wat tydens hierdie studie beskryf word aangaande die identifikasie van GLRaV-3 variante, dra by tot die molekulêre kennis van GLRaV-3 en

verskaf ‘n basis vir verdure epidemiologiese -en insek oordragingstudies. Die studie het vir die eerste keer bewys dat verskillende GLRaV-3 variante na aanliggende plante in ‘n ry asook oor rye oorgedra word. Die diversiteit tussen die GLRaV-3 variant groepe in die 5’UTR moet verder ondersoek word en die deel van die genoom kan ‘n belangrike rol speel in die biologiese eienskappe van die variante.

Preface

This dissertation is divided into 5 chapters:

CHAPTER 1: General Introduction and Objectives of the Project

CHAPTER 2: Literature Review

CHAPTER 3: Three genetic grapevine leafroll-associated virus-3(GLRaV-3) variants identified from South African vineyards show high variability in their 5'UTR

CHAPTER 4: Distribution of grapevine leafroll associated virus 3 (GLRaV-3) variants in South African vineyards

CHAPTER 5: Conclusions

List of Abbreviations

aa	amino acid(s)
ArMV	Arabis mosaic virus
bp	base pair(s)
BPYV	Beet pseudo-yellows virus
BYSV	Beet yellow stunt virus
BYV	Beet yellows virus
CI	consistency index
CP	coat protein
CTV	Citrus tristeza virus
CYSDV	Cucurbit yellow stunting disorder virus
dCP	duplicate capsid protein
DNA	deoxyribonucleic acid
D-RNA	defective ribonucleic acid
dsRNA	double-stranded ribonucleic acid
ELISA	enzyme-linked immunosorbent assay
GFLV	Grapevine fanleaf virus
GLR disease	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-9	Grapevine leafroll–associated virus-9
GRSPaV	Grapevine rupestris stempitting-associated virus
GSyV-1	Grapevine Syrah virus-1
GVA	Grapevine virus A
GVB	Grapevine virus B
GVE	Grapevine virus E
HEL	helicase
Hsp70h	heat shock protein 70 homologue
ICVG	International Council for the study of Virus and Virus-like Diseases of the Grapevine
kb	kilobase(s)
LChV-1	Little cherry virus-2
LIYV	Lettuce infectious yellows virus
L-Pro	leader protease

MET	methyltransferase
nm	nanometer(s)
nt	nucleotide(s)
ORF	open reading frame
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
PMWaV-3	Pineapple mealybug wilt-associated virus-3
PNW	Pacific North West
RdRp	RNA-dependent RNA polymerase
RE	restriction enzyme
RFLP	restriction fragment length polymorphism(s)
RI	retention index
RLM-RACE	RNA ligase mediated rapid amplification of cDNA ends
RPA	ribonuclease protein assay
RSP	Rupestris stem pitting
RT-PCR	reverse transcription-polymerase chain reaction
RW	rugose wood
SAWIS	SA Wine Industry Information System
sgRNA	subgenomic ribonucleic acid
SSCP	single strand conformation polymorphism
ss-RNA	single-stranded ribonucleic acid
TBR	Tree bisection and reconnection
UTR	untranslated region
Winetech	Wine Industry Network of Expertise and Technology
WOSA	Wines of South Africa

Acknowledgements

I would like to thank the following people and institutions:

- My supervisors, Prof Johan T Burger and Dr. Dariusz E Goszczynski, for support and guidance throughout the study
- Gerhard Pietersen for guidance with the field work, critical reading of manuscripts and support since the start of my career
- Hano Maree for the assistance with lab work in Stellenbosch, critical reading of manuscript and the pleasant team work
- Prof Dirk U Bellstedt, Department of Biochemistry, SU, for assistance with the phylogenetic analysis of genomes and critical reading of manuscripts
- My colleagues and friends at ARC-PPRI: Marika, Kassie, Johan, Isabel, Susan, Ahmed, Teresa, for motivation, support and creating a pleasant working environment
- All the friends in the *Vitis* lab, Stellenbosch, for your support and making my stay in the lab pleasant
- Elsa van Niekerk, ARC-PPRI, who assisted with the graphics of the manuscripts
- The financial assistance of Winetech and the THRIP program of the NRF throughout the study
- SASEV, for funding the trip to attend the 16th ICVG meeting in Dijon, France in 2009
- The examiners of this dissertation, for your time and inputs
- My friends, for encouragement and endless support. It meant a lot to me, you are the best!
- My parents and sisters, your love and support kept me going, words are not enough!
- My best Friend, my Heavenly Father, who is always there to keep an eye on me.

Contents

Contents of chapters

CHAPTER 1

General Introduction and Objectives of Project

1.1 BACKGROUND INFORMATION OF PROJECT

1.1.1 General introduction

Wine making in South Africa has been ongoing for 300 years already and this industry is one of the best established in the country. The main centre for grape production in South Africa is in the Western Cape Province where long, warm summers provide ideal conditions for viticulture. Currently 108 000 hectares of wine grapes are under cultivation locally over an area of 800 kilometres in length.

A number of virus diseases are threatening the grapevine industry worldwide. Grapevine leafroll (GLR) is one of the most important viral diseases of grapevines, occurring in all grape-producing countries of the world. Grapevine leafroll-associated virus3 (GLRaV-3) is known to be closely associated with GLR disease and occurs commonly in South African vineyards (Pietersen 2004, 2006). Several epidemiological studies showed that GLR is spreading rapidly in vineyards. Despite the negative impact of GLRaV-3 on grapevine industries worldwide, the genetic variability of the virus, knowledge essential for developing effective control measures to the virus, is largely unknown. Recently the genetic variability of GLRaV-3 is being investigated more frequently in world wide vineyards.

GLRaV-3 is transmitted between grapevines by at least six species of pseudococcid mealybugs and four soft scale species. The interaction between the virus and the vector, not studied here, is an important aspect to consider in understanding the GLR disease complex. Figure 1 illustrates some of the research aspects discussed in this study, namely, 1) a field survey of GLR-infected mother blocks, 2) analysis of GLRaV-3 variants in infected plants using the SSCP technique to examine genetic variability, 3) identification of GLRaV-3 variants, 4) the full-length genome sequence determination of three GLRaV-3 variants, and 5) the interaction between the virus and the mealybug vector. GLRaV-3 is believed to be the major virus in GLR-infected plants in South African vineyards and the study of GLRaV-3 variants will be presented here.

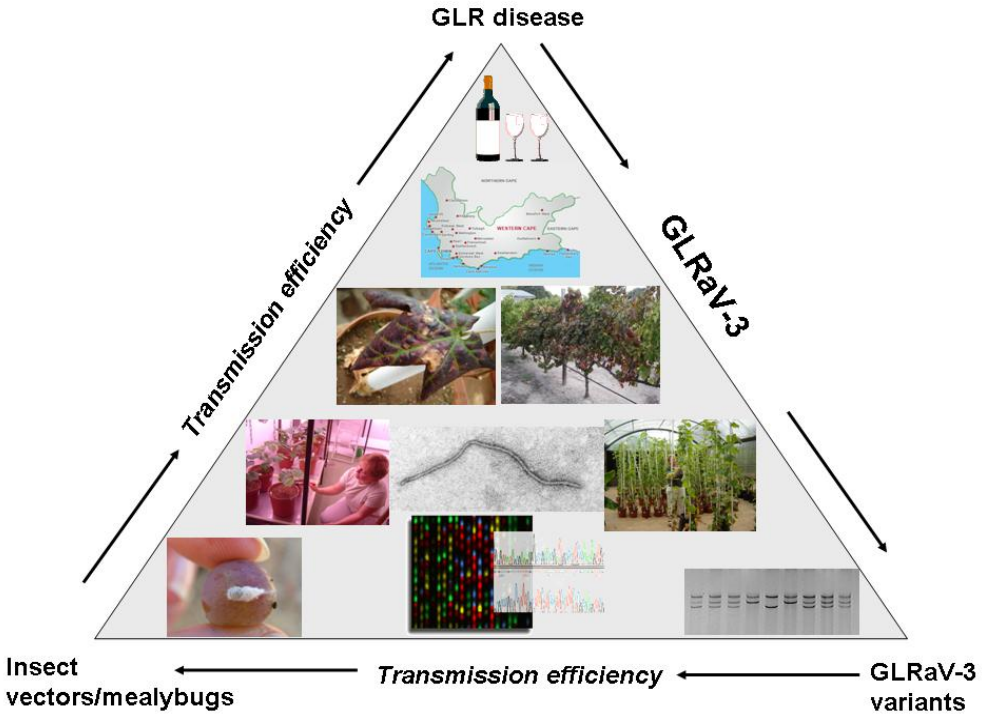


Figure 1. A graphic presentation of the research aspects in this study, with the focus on the identification of GLRaV-3 variants.

The molecular characterization of a GLRaV-3 full length genome was first published in 2004 and for four years only one full length sequence of a GLRaV-3 isolate, NY-1, was known (Ling et al., 2004). In 2008, two additional full length sequences were published, a Chilean isolate Cl-766 (Engel et al., 2008) and a South African isolate GP18 (Maree et al., 2008). In a previous study of South African isolates of GLRaV-3, single strand conformation polymorphism (SSCP), restriction enzyme (RE) SSCP, cloning and sequencing techniques were used to identify two clearly divergent molecular groups of the virus (Jooste & Goszczynski, 2005). The first molecular variant, represented by isolate 621, was very similar to the NY-1 isolate of Ling et al. (2004) while sequence data of the second molecular variant, represented by isolate 623, was very similar to the complete genome sequence of the South African isolate GP18 (Maree et al., 2008). The molecular divergence between these two variant groups was especially high in the 5' terminal part (partial sequences of the 5'UTR and ORF1a) of the virus where nucleotide sequences differed by 35%. Sequence data of the remaining coding regions showed nucleotide similarities above 90% between the variant groups. The two variant groups could be distinguished by unique SSCP profiles generated from an amplified region in ORF 5. This project is a continuation of the initial study published by Jooste & Goszczynski (2005).

1.1.2 Value of work

There are currently no published data on the association of biological properties of GLRaV-3 variants to a specific pathogenic characteristic. *i.e.* symptom expression or spread of GLRaV-3 in plants. To be able to study biological properties of a virus the genetic variability of a virus must be known. The full length sequence data published to date showed clear variability between the genomes of different GLRaV-3 isolates. These studies revealed genomic regions where nucleotide changes can have a significant impact, for example, the variable 5'UTR. In this study the identification of GLRaV-3 variants in South African vineyards will be discussed. For the successful control of GLR disease it is important to know the variant status of a plant as well as the interaction between the vector and virus variants. It is also important to have specific and universal detection methods in place to detect all GLRaV-3 variants. The lack of mechanical transmissibility of GLRaV-3 has impaired the molecular and biological characterization of the virus.

1.2 OBJECTIVES OF PROJECT

The rapid spread of GLR in South African vineyards (Pietersen, 2004) is of major concern to the industry. Molecular variability, which determines biological properties of a virus, and the virus-vector interactions, are the most important aspects to consider to advance our knowledge of disease epidemiology and devise efficient management strategies. The ***first objective*** of this project was to obtain the full genome sequences of the two variants, represented by isolates 621 (group I) and 623 (group II), described in Jooste & Goszczynski (2005). ***Objective two***was to investigate the presence and interaction of the two GLRaV-3 variants in the GLR disease clusters from mother blocks in different regions. A field survey was done; firstly to determine which of the variants occurred predominantly in the selected mother blocks and secondly to determine if there are any differences in the distribution patterns of the two GLRaV-3 variants. A related objective was to assess if the distribution of GLRaV-3 variants correlate with the spread of GLR in vineyards. What transpired from the field surveys done in mother blocks in 2007 and 2008 was the identification of a third molecular variant and this lead to ***objective 3*** to obtain the full genome sequence of this third variant, represented by PL-20 (group III).

1.3 PUBLICATIONS AND PRESENTATIONS

The work presented in this dissertation was published in the following journals and presented at the following meetings:

Peer reviewed publications:

Jooste AEC, Maree HJ, Bellstedt D, Goszczynski DE, Pietersen G, Burger JT (2010) Three genetic grapevine leafroll-associated virus-3(GLRaV-3) variants identified from South African vineyards show high variability in their 5'UTR. DOI: 10.1007/s00705-010-0793-y. Arch Virol 155 (12):1997-2006

Jooste AEC, Pietersen G, Burger JT (2010) Distribution of grapevine leafroll associated virus 3 (GLRaV-3) variants in South African vineyards. *Submitted to European Journal of Plant Pathology, accepted 15 November 2010*

Popular publication:

Jooste E (2008) A serious disease threatening the South African wine industry, PPRINews 74: pp 14-15.

International Conferences: Oral presentations

Jooste AEC, Goszczynski DE (2006) Differentiation between two distinct molecular variants of GLRaV-3. 15th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG). In: Extended abstracts. 3-7 April 2006, Stellenbosch, South Africa.

Jooste E, Maree H, Pietersen G, Goszczynski DE, Burger J (2009)Identification and distribution of three divergent molecular variants of grapevine leafroll-associated virus-3(GLRaV-3) in South African vineyards. 16th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG). In: Extended abstracts, 31 August- 4 September 2009, Dijon, France.

Maree HJ, **Jooste E,** Stephan D, Freeborough M-J, Burger JT (2009) Characterisation of the genomic and subgenomic RNA of grapevine leafroll-associated virus-3 (GLRaV-3). 16th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG). In: Extended abstracts, 31 August- 4 September 2009, Dijon, France.

International Conference: Poster presentation

Pietersen G, Oosthuizen T, **Jooste E**, Filippin L, Bertazzon N, Angelini E (2009) Shiraz disease and grapevine yellows in South Africa. 16th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG). In: Extended abstracts, 31 August- 4 September 2009, Dijon, France.

National Conference: Oral presentation

Jooste AEC, Goszczynski DE, Pietersen G, Burger JT (2009) Identification of a third molecular variant of grapevine leafroll-associated virus-3(GLRaV-3) associated with Leafroll Disease in vineyards of the Western Cape. 46th Congress of the Southern African Society for Plant Pathology,25-28 January, 2009, Gordons Bay, South Africa.

South African industry programme: Winetech(Yearly progress meetings: Oral presentations)

Jooste AEC(2006) Progress on the molecular and biological characterisation of GLRaV-3 variants. Winetech Grapevine Virus Workshop held, Olive Grove, Infruitech, Stellenbosch, 15 Augustus 2006.

Jooste AEC (2007) GLRaV-3 variants: Sequence results, biological experiment and field survey. Winetech Grapevine Virus Workshop held, Olive Grove, Infruitech, Stellenbosch, 2 May, 2007.

Jooste AEC(2008)Variability on the GLRaV-3 genome. Winetech Virus Workshop held, Olive Grove, Infruitech, Stellenbosch, 19 August 2008.

Jooste AEC(2009) Final report on GLRaV-3 variant status in South African vineyards, Winetech Virus Workshop held, Olive Grove, Infruitech, Stellenbosch, 15 September 2009.

CHAPTER 2

Literature Review

2.1 INTRODUCTION

2.1.1 Grapevine cultivation: Now and Then

The wine grapevine is part of the plant genus *Vitis*, with species name *vinifera,* meaning wine-bearing, and has the unique ability to accumulate sugar in its grapes up to a third of its volume making its juice a clean and lively drink (Johnson, 2005). The earliest evidence of wine-making dates back to archeological findings of grape pips from as early as 7000-5000BC in Georgia. Excavations in Turkey, Damascus in Syria, Byblos in Lebanon and Jordan have produced grape pips from the Stone Age, about 8000BC (Johnson, 2005). Recent physical evidence from China showed that a stem from *Vitis vinifera*, discovered in the Yanghai Tombs, Turpan District in Xinjiang, proved to be nearly 2300 years old, which suggests that there was grape cultivation at least from that time in China (Jiang et al., 2009). Wine drinking was enjoyed by ancient Egyptians and is well documented in their paintings. The cultivation of grapevine and olive in the Mediterranean cultures made Greece one of the economic strongholds in earlier times. Grapevine production soon spread from Greece onto Italian shores when the Tuscany of today belonged to the Etruscans, who were keen grapevine growers and wine producers. The grapevine production culture spread to most of Europe and followed western civilization.

The history of wine making in South Africa goes back to the first cultivation of vineyards in the Western Cape in 1655 by Jan van Riebeeck who planted the first vineyard and his successor, Simon van der Stel, who planted a vineyard on his farm Constantia a few years later. The wines from this elite farm are still famous today. Additional expertise regarding vine growing and wine production came when the French Huguenots arrived and settled at the Southern tip of Africa between 1680 and 1690. Their wine making skills left a permanent impression on the South African wine culture.

Grapevine cultivation in California followed in 1697 and Australia and New Zealand in 1813 (WOSA, http://www.wosa.co.za/sa/history_beginning.php). The so-called New World vineyards, including Australia, New Zealand, South Africa, Chile, Argentina, Mexico and the United States soon produced wines comparable to the finest French wines. The distinction between Old World (Europe) and New World wine lies with different philosophies of winemaking (Swart & Smit, 2006). Old World wine making is based on tradition and Nature is the key factor; wine is viewed primarily as an expression of *terroir* (a combination of

topography, climate, geology and soil variations) rather than individual varieties (Swart & Smit, 2006). Old World wines tend to have lower alcohol levels (Alc. 11-12%), fruit flavours relating to each variety are less pronounced and Old World wines have a greater maturation potential. New World wines are characterized by the application of new technologies, innovative cultivation and exploration of new ideas. These wines are created to be consistent in quality, are defined by varietal characteristics and the expression of a wine's fruit characteristics (Swart & Smit, 2006). New World wines have higher alcohol levels (up to Alc. 16%) and tend to have a more pronounced fruitiness because they are grown in warmer regions and sugar levels are higher. The maturation potential of these wines is not as high as the Old World wines. South African wines are often described as lying somewhere between these two worlds, with the structure and restraint of the Old World and the fruit intensity of the New (Swart & Smit, 2006).

2.1.2 Grapevine Cultivation: Regions and Varieties

There are five main wine production regions in the Western Cape (Figure 1), namely BreedeRiverValley, Coastal, Little Karoo, OlifantsRiver and Boberg, which cover21 diverse districts and some 64 smaller wards. The vineyards included in this study were from Stellenbosch-, Paarl-, Wellington-, Rawsonville-, Worcester- and Somerset West grape production districts. A new wine production area has recently been developed in KwaZulu-Natal that stretches from Greytown to Oribi Flats and the Midlands where altitudes reach up to 1500 metres.

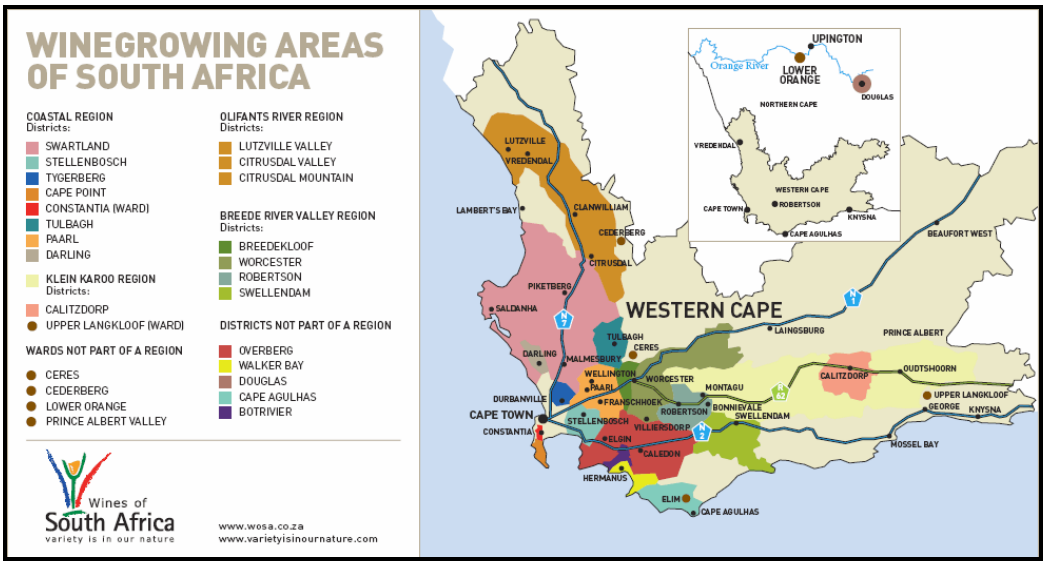


Figure 1. The winegrowing regions in the Western Cape (www.wosa.co.za).

According to WOSA over 40% of vineyards were replanted in recent years to ensure that the South African market competes globally, moving from quantity to noble cultivars and quality wines. The shift to planting more white cultivars than red cultivars in the past 4 years is a reversal of the 10-year trend of planting more red cultivars.

Noble varieties which have been cultivated increasingly in the past few years include Sauvignon Blanc and Chardonnay, which produce top-class white wines, and Shiraz and Pinot noir. Although most of the vine varieties were imported material, up to now six crossings have been released. The best known of these is a red variety, Pinotage, a hybrid of Pinot Noir and Hermitage, which is cultivated on a fairly large scale. In total, 21 red and 20 white varieties are grown in South Africa (Table 1).

Table 1. The white –and red wine varieties grown in South Africa

WHITE-WINE VARIETIES	RED-WINE VARIETIES
Bukettraube	Cabernet Franc
CapeRiesling (Crouchen Blanc)	Cabernet Sauvignon
Chardonnay	Carignan
Chenel	Cinsaut
Chenin Blanc (Steen)	Gamay (Noir)
Clairette Blanche	Grenache (Noir)
Colombar(d)	Malbec
Emerald Riesling	Merlot
Gewűrztraminer	Mourvèdre
Grenache (Blanc)	Muscadel
Muscat d'Alexandrie (Hanepoot)	Nebbiolo
Muscadel	Petit Verdot
Nouvelle	Pinot Noir
Palomino (White French Grape)	Pinotage
Pinot Gris	Roobernet
Sauvignon Blanc	Ruby Cabernet
Semillon (Green Grape)	Shiraz
Ugni Blanc (Trebbiano)	Souzào
Viognier	Tinta Barocca
Weisser Riesling (Rhine Riesling)	Touriga Nacional
	Zinfandel

2.1.3 Economic importance of the South African wine industry

The economic importance of the wine industry is shown by the 348 500 people being employed directly and indirectly in the wine industry. According to a study commissioned by the SA Wine Industry Information System (SAWIS), the wine industry contributes 9.7% to the Western Cape's gross geographic product. The study concluded that of the R14.6 billion contributed by the wine industry to the regional economy, some R3 billion was generated indirectly through wine-tourism activities centered in the winelands. Although local vineyards account for just 1.5% of the world's vineyards, South Africa ranks as number eight in volume

production of wines and produces 3% of the world's wine (WOSA website http://www.wosa.co.za).

2.1.4 Grapevine virus diseases

A wide range of viruses and virus–like diseases are threatening the grapevine industry worldwide as well as locally. The grapevine diseases, grapevine leafroll (GLR), shiraz disease, syrah decline, rugose wood (RW) complex including rupestris stem pitting (RSP) disease, corky bark, kober stem grooving and LN33 stem grooving syndrome, cause economic losses in worldwide grapevine production areas. A report written in 2006 listed 58 plant viruses that infect grapevine (Martelli & Boudon-Padieu, 2006). These viruses represent eight families and 18 plant virus genera. At the 16th International Council for the study of Virus and Virus-like Diseases of the Grapevine (ICVG) meeting, held in Dijon, France, two newly characterized viruses were added to this list, namely *Grapevine virus E*(GVE) (Nakaune et al.,2008) and *Grapevine syrah virus-1* (GSyV-1) (Al Rwahnih et al., 2009). The identification of new viruses in vineyards is important to understand the interaction between viruses and disease complexes. With new technologies emerging every day, this task is becoming easier and the characterization of viruses at the molecular level is much faster. However, to associate a specific virus with a disease is still a challenge and requires precise studies, including biological studies.

The focus in this dissertation will be on grapevine leafroll (GLR) disease, with the main focus on grapevine leafroll associated virus 3 (GLRaV-3).

2.2 GRAPEVINE LEAFROLL DISEASE

2.2.1 Symptoms

Grapevine leafroll disease (GLR) is one of the most important diseases of grapevines, occurring in all grape-producing countries worldwide, including South Africa (Pietersen, 2004).The disease delays ripening of grapevine berries, decreases the accumulation of sugar and ultimately influences the quality of the wine. The expression of GLR symptoms is variable among cultivars, and environmental conditions play a role as well. In red-berried cultivars the leaf blade areas turn red, whereas leaf yellowing of the same leaf occurs in white wine cultivars (Carstens, 2002). Some white cultivars may show no visual signs of infection (Rayapati et al., 2009). Symptoms are best observed in the period between harvesting and

shedding of leaves (late summer and early autumn). Typical GLR symptoms appear as distinctive downward rolling of leaves with leaf veins that stay green (Figure 2).

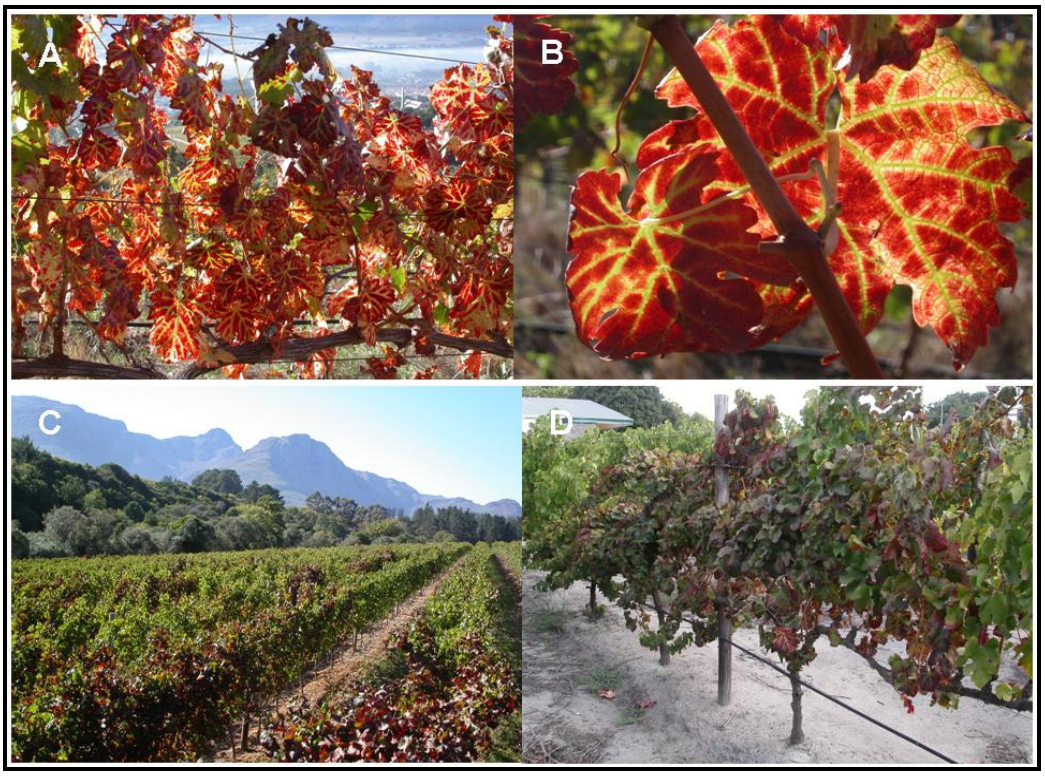


Figure 2A-D. Typical GLR symptoms (A, B) with areas between the veins turning red including downward curl of leaves and leafroll infection visible in rows and across rows (C, D) (Photos: G. Pietersen).

2.2.2 Viruses involved in GLR disease

Viruses from the *Ampelovirus* and *Closterovirus* genera are known to cause leafroll disease. Several phloem-limited filamentous viruses, identified as grapevine leafroll-associated viruses(GLRaVs), have been characterized from leafroll infected grapevines (Fuchs et al., 2009b). These viruses are from the genera *Closterovirus* (GLRaV-2), *Ampelovirus* (GLRaV-1, GLRaV-3, GLRaV-4, GLRaV-5, GLRaV-6 and GLRaV-9) and GLRaV-7 is not yet assigned to a genus (Fuchs et al., 2009b).

2.2.3 Epidemiology of GLR disease

Crop losses caused by GLR disease are a worldwide problem and can have huge economic impact. Significant yield losses of 30-50% have been recorded (Fuchs et al., 2009b) and even as high as 68% (Walter & Martelli, 1997). In a recent report written for the New Zealand wine industry (Charles et al., 2006), 22 studies were identified that presented data on GLR disease and its effects on yield. There were significant variations between the reports concerning

reduced yield. There is no quantifiable yield loss data available for South African vineyards but we can assume a similar situation exists. We do know that the disease causes serious problems in the South African wine industry due to its rapid spread and infection of certified planting material (Pietersen, 2006). A sector of the local wine industry, Winetech, invests research funds to study aspects of GLR disease, including projects that aim to eradicate virus infected material from vineyards by implementing control strategies for leafroll disease.

GLR disease is transmitted through infected propagation material as well as through mealybug and soft scale insect vectors (Belli et al.,1994; Cabeleiro & Segura, 1997; Douglas & Krüger, 2008; Petersen & Charles, 1997; Sforza et al.,2003; Tsai et al.,2008,).

Recently, several epidemiology studies on GLR disease have been reported from grapevine growing regions worldwide. These studies were mainly done in South Africa (Pietersen, 2006), Spain (Cabeleiro et al., 2006, 2008) and the USA (Golino et al., 2008, Rayapati et al., 2009). A study of the spread of GLR disease in a Napa Valley vineyard in California showed that the disease spread from neighbouring blocks, heavily infected with leafroll, and mapping results of the disease showed a spread rate increase of more than 10% per year in this block (Golino et al.,2008). The possible causes for this sudden rapid spread of GLR in vineyards of California were debated and the authors suggested that something fundamental changed in the vineyards, such as vector epidemiology, grower rootstock preferences and/or new leafroll strains that emerged (Golino et al.,2008). The epidemiological studies reported by Cabeleiro et al. (2006, 2008) described the involvement and spread of GLRaV-3 in GLR disease. The spatial distribution of GLRaV-3 was studied in vineyards from Spain since 1991 (Cabeleiro et al.,2006, 2008) and reported recently (Cabeleiro et al., 2008). From this study it was clear that there was a correlation between mealybug incidence and virus spread (Cabeleiro et al.,2008). Scale insects were implied as vectors of GLRaV-3 in the Meaño vineyard where slow, but constant spread of the virus was observed (Cabeleiro et al.,2008). In two vineyards, in Portomarín and Goian, in the same study, the virus inoculum originated from infected plant material resulting in a random distribution of the disease. A study of vineyards in the Pacific Northwest (PNW) of the U.S.A. documented the presence of genetic variants of GLRaV-1, GLRaV-2, GRSPaV and GFLV in these vineyards (Rayapati et al.,2009).

The spatial distribution and spatial dynamics (changes in distribution patterns) of GLR disease within Mother blocks of the South African Certification Scheme were studied intensively from 2001-2007 (Pietersen, 2004, 2006). Four common distribution patterns of

GLR disease were observed in this study. The most significant distribution pattern identified in local vineyards was secondary spread within vineyards after establishment (Pietersen, 2006). The other means of GLR spread are primary spread of leafroll by infected plant material, GLR spread from a preceding vineyard and gradients of GLR-infected vines associated with proximal leafroll infected vineyards (Pietersen, 2006). An example of the distribution patterns of GLR in vineyards are shown in Figure 3.

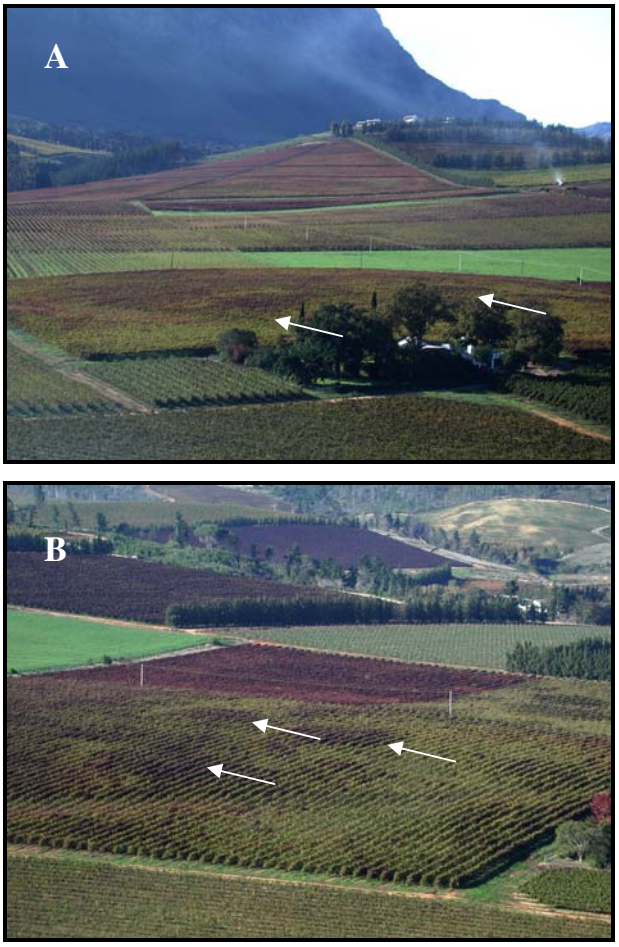


Figure 3. Spatial distribution patterns of GLR infection in vineyards (A,B). GLR infection clusters clearly visible (indicated by white arrows) and a vineyard with 100% leafroll infection in the far background in photoB. (Photos: G. Pietersen).

2.2.4 Control strategies for GLR disease

To combat the spread of GLR disease in vineyards, most wine producing countries recognised the importance of a certification scheme for virus-free propagation material. Published literature on the management options to limit the spread of GLRaV-3 within new vineyards is limited and only a few scientific publications exist on this topic (Charles et al., 2006). Only a few publications exist on the control of GLRaV-3 mainly because leafroll symptoms and

associated damages are not so serious in other parts of the world as in South Africa. In most other countries other diseases causes more serious damages than leafroll and therefore studies and control strategies are focused on such diseases. Leafroll disease control in South Africa is a high priority for the local industry and some extensive work has been done on the control of this disease locally.

The study conducted by Pietersen (2004) in South African vineyards led to the establishment of a management strategy for the disease. It is generally accepted that GLRaV-3 is not infecting hosts other than *Vitis* and the virus cannot be transmitted mechanically but can spread by vegetation propagation and grafting of infected plant material. The virus can be transmitted by mealybugs and scale insects from plant to plant, with mealybugs being the most prevalent vector.

The control measures implemented in the South African vineyards were described in detail in a popular article published in a local Wineland magazine (Pietersen, 2010). Rouging (removal of infected vines), planting strategies for new blocks, the use of certified planting material and the control of mealybugs and ants in vineyards were discussed in the paper.

2.3 GRAPEVINE LEAFROLL-ASSOCIATED VIRUS 3 (GLRaV-3)

2.3.1 Taxonomy

GLRaV-3 is the type member of the *Ampelovirus* genus in the family *Closteroviridae* (Martelliet al., 2002). The family comprise of three genera, namely *Closterovirus, Ampelovirus* and *Crinivirus.* The three genera distinguish between aphid, mealybug and whitefly transmitted viruses. Molecular properties, like genome composition and structure, also differentiate the three genera. Other viruses that belong to the *Ampelovirus* genus are *Grapevine leafroll-associated virus-1, -4, -5, -6, -9* (GLRaV-1, -4, -5, -6, -9), *Pineapple mealybug wilt-associated virus-1, -2* (PMWaV-1, -2), *Little cherry virus-2* (LChV-2) (Martelli et al., 2002) and *Plum bark necrosis stem-pitting-associated virus* (PBNSPaV) (Al-Rwahnih et al., 2007). The ampeloviruseswere recently divided into two subgroups based on the phylogenetic analyses of the Hsp70h, RdRp and HEL domains of viruses in this group (Maliogka et al., 2009). This analysis included two Greek isolates, GLRaV-Pr and GLRaV-De, which represent two newly assigned ampeloviruses(Maliogka et al.,2008). These two isolates, together with GLRaV-4, -5, -6 and -9, PMWaV-1 and PBNSPaV are included in the lineage of subgroup I ampeloviruses and GLRaV-1, -3, PMWaV-2 and LChV-2 included in

the subgroup II lineage. The genome organisation and phylogenetic relationship of *Pineapple mealybug wilt-associated virus-3* (PMWaV-3) with other closterovirusessuggests the addition of another genus within the family *Closteroviridae* (Sether et al., 2009).

2.3.2 Morphology and Genome organization

The virus has flexuous particles of about 1800 nm in length (Figure 4); containing a positive-sense single stranded RNA (ssRNA) genome. The RNA content in closterovirusparticles is about 5% (Dolja et al.,1994).

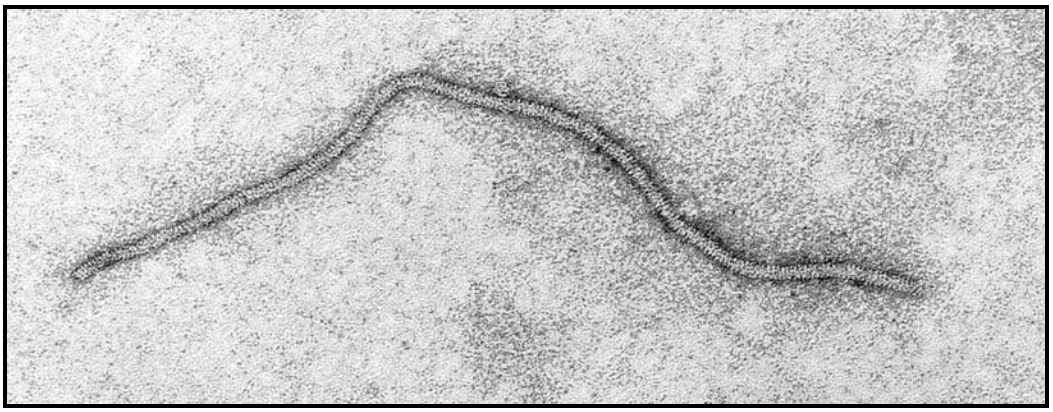


Figure 4. Electron micrograph of a purified GLRaV-3 particle negatively stained with 2% uranyl acetate. (Photo: G.G.F. Kasdorf).

The size of closterovirusgenomes varies from ~15.5 to ~19.5kb with a coding capacity of 10-14 proteins (Dolja et al.,2006).

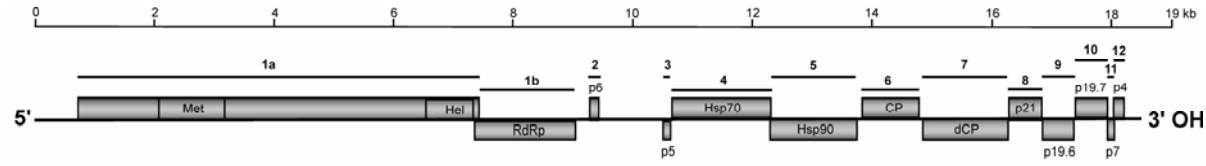


Figure 5. Schematic representation of the GLRaV-3 genome and positions of genes and ORFs.

The first full-length genome sequence of a GLRaV-3 isolate, NY-1, was published by Ling et al. in 2004. The genome organisation of the virus conformed to the genome structure for closterovirusesproposed by Dolja et al.(1994). The relatively large genome of GLRaV-3 is organized into 13 open reading frames (ORFs) (ORF1a, 1b, 2-13) and represent a typical monopartite closterovirus(Ling et al.,2004), the genome organisation seen in Figure 5. In the

Ling study, a comparative study was done on all genes and amino acid sequence similarities between GLRaV-3 and other closteroviruseswere calculated in these regions.

ORF1a encodes a large polyprotein with different domains; leader protease (L-Pro) (Ling et al.,2004), methyltransferase (MET) (Ling et al., 1998), AlkB domain (Engel et al.,2008; Maree et al.,2008) and helicase (HEL) (Ling et al., 1998). L-Pro plays a prominent role in the amplification of the viral genome either activation of the viral replicase or protection of the RNA from degradation by a host defense system (Dolja et al., 2006). The C-terminal portion of ORF1a shared significant similarity with the Superfamily 1 helicase of positive–strand RNA viruses (Ling et al., 2004). Phylogenetic analyses of this region showed that GLRaV-3 grouped in a cluster of its own when compared to other closteroviruses, *Beet yellows virus*(BYV), *Citrus tristeza virus*(CTV) and *Beet yellow stunt virus*(BYSV) (transmitted by aphids), *Lettuce infectious yellows virus*(LIYV) (transmitted by whiteflies) and *Little cherry virus*(LChV) (transmitted by mealybugs). Although LChV is transmitted by mealybugs it showed to be closer related to LIYV. ORF1b encodes for a RNA-dependent RNA polymerase (RdRp) and showed significant similarity to the Supergroup 3 RdRp of positive–strand RNA viruses. Phylogenetic relationships were similar in this region as in ORF1a. An interesting feature of the mealybug-transmitted closteroviruses is a long untranslated intergenic region downstream of ORF1b, which is GC rich and possess extensive RNA secondary structure (Karasev, 2000). The size of this intergenic region is comparable to the size of a protein encoded by ORF2 in the BYSV and CTV genomes (Karasev, 2000). ORF2 encodes a small peptide and for this region no equivalent ORFs were found in BYV and LChV genomes, but in CTV, LIYV and BYSV larger ORFs were found (Karasev et al.,1995, Karasev et al.,1996, Klaasen et al.,1995). The p6 protein resides in the ER and functions in virus movement from cell to cell and can be considered a conventional movement protein (Dolja et al., 2006). ORF 3 encodes a small hydrophobic transmembrane protein. ORF4 encodes the Hsp 70-homologue protein that is the unique hallmark of the closterovirus family (Dolja et al.,1994). Eight conserved motifs (A-H) were identified from the multiple alignments of Hsp 70 homologues of GLRaV-3 and other closteroviruses (Ling et al.,1998). Three of these functionally important motifs (A-C) contain ATPase activity typical of closteroviral Hsp70 chaperone-like proteins. ORF5 encodes a 55K protein but the two conserved regions of the Hsp70-homologue previously delineated in BYV and CTV were not identified in this protein of GLRaV-3 (Ling et al.,1998). ORF 6 and 7 encodes the coat protein (CP) gene and copy of CP (dCP). The duplication of the capsid protein gene is a unique feature of closteroviruses(Boyko et al.,1992).The function of the remaining ORFs 8 to 12 was not determined by Ling et al.

(1998). It is suggested that these ORFs, especially the p21 protein coded for in ORF8, encode for viral silencing suppressors. The p20 proteins, coded for in ORF 9 and 10, are involved as systemic movement proteins (Dolja et al., 2006).

In a recent study the sgRNAs associated with GLRaV-3 infection was investigated (Maree et al., 2010). The production of sgRNA is necessary for the expression of the 3' ORFs (3-12) in positive sense RNA viruses. This study predicted the existence of at least seven 3' co-terminal positive-sense sgRNAs for the expression of ORFs 3-12 (Maree et al., 2010). The gene expression strategy and *cis*- acting elements of GLRaV-3 were reported recently in another study (Jarugula et al., 2010). The study showed that four of the eleven putative 3' co-terminal sgRNAs (specific to ORF6, 8, 9 and 10) were present in higher levels, two sgRNAs (ORF11 and 12) accumulated at intermediate levels and three sgRNAs (ORF7, 5, 3 and 4) were present in very low levels (Jarugula et al., 2010). These results suggest that 3' coterminal sgRNAs accumulate at variable amounts, reflecting differences in their expression levels in infected grapevine tissues. It was suggested that ORF10-12 are likely to be translated from the same sgRNA (Maree et al., 2010).

2.3.3 Full length sequences of GLRaV-3

As mentioned earlier, the first full-length sequence of GLRaV-3, from the NY-1 isolate, was published by Ling et al. (2004). In 2008, a full-length genome sequence of a Chilean GLRaV-3 isolate, Cl-766, was published that showed the same properties as the NY-1 isolate (Engel et al.,2008). In the same year, the complete genome length of a South African GLRaV-3 isolate, GP18 (EU259806), was reported to be 18498 nt (Maree et al.,2008). The extended length of the 5'UTR, consisting of 737 nt, differed from that reported previously by Ling et al. (2004) and Engel et al. (2008) where a 5'UTR of 158 nt for both isolates NY-1 (AF037268) and Cl-766 (EU344893) was described. The length of the 3'UTR of all GLRaV-3 isolates sequenced to date is 277 nucleotides (nt) (Engel et al.,2008; Ling et al.,2004; Maree et al.,2008). Since the report by Maree et al.,three additional GLRaV-3 isolates from South Africa were sequenced, namely isolates 621, 623 and PL-20 (Joosteet al.,2010). The detailed description of the three GLRaV-3 isolates will be discussed in Chapter 3.

2.3.4 Transmission of GLRaV-3

The survival of a plant virus depends on its efficient transmission from plant to plant. Since the association of GLRaV-3 with GLR disease, the vectors responsible for transmitting the virus were studied intensively in combination with the spread of the disease (Cabeleiro et al.,

1997, 2006, 2008; Daane et al., 2006; De Bourbon et al.,2004; Douglas & Krüger, 2008; Fuchs et al., 2009b; Golino et al., 2002, 2008; Mahfoudi et al., 2009; Petersen & Charles, 1997; Pietersen, 2004, 2006; Sforza et al., 2003; Tsai et al., 2008;)

GLRaV-3 is transmitted in a semi-persistent way by its mealybug insect vectors (Martelli et al., 2002), although a recent study suggest a circulative transmission mechanism (Cid et al., 2007).

The first report of GLR disease transmission by the vine mealybug *Planococcus ficus* in South Africa was two decades ago (Engelbrecht & Kasdorf, 1990). The vine mealybug *P. ficus* (Figure 6A) is considered the most important vector of GLRaV-3 in South Africa and the longtailed mealybug, *Planococcus longispinus* (Figure 6B), is far less abundant on grapevine and has a more aggregated distribution in vineyards than *P. ficus* (Walton & Pringle, 2004). Transmission efficiency studies with *P. ficus* and *P. longispinus* showed that the two mealybug species are both efficient vectors for GLRaV-3 in South African vineyards. The study showed for the first time that a single nymph of *P. ficus* or *P. longispinus* is capable of infecting a healthy grapevine plant with GLRaV-3 (Douglas & Krüger, 2008). The age of the mealybug and dispersal of mealybugs play a role in the efficiency of transmission of the virus from plant to plant. In a recent study it was confirmed that the first and second instars of *P. ficus* is more effective (36.7% versus 10%) in transmission of the virus than the adult females (Mahfoudi et al., 2009). These first instar nymphs could be carried by wind over long distances but may not have fed on phloem before dispersing. One can argue that the adult mealybug is more likely to transmit a virus from plant to plant with its less active lifestyle. The dispersal of mealybugs is therefore connected to the transmission ecology and important fact to consider when performing vector-virus studies.

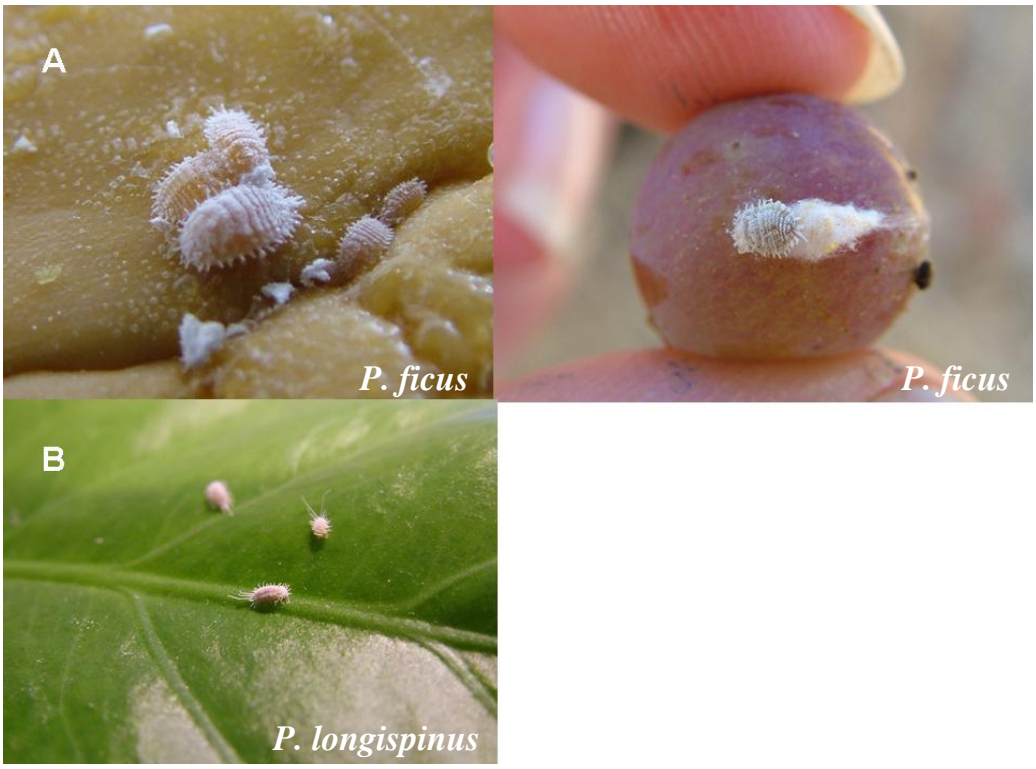


Figure 6. Two mealybug species, *Planococcus ficus* (A) and *Planococcus longispinus* detected and studied in South African vineyards. (Photo A: N. Douglas-Smit; Photo B: D.B. Douglas).

Other species recorded to transmit GLRaV-3 are the soft scales *Pulvinaria innumerabilis*, *Pseudococcus maritimus* (Golino et al., 2002), *Ceroplastes rusci* (Mahfoudi et al., 2009), *Pulvinaria vitis* (Belli et al., 1994) and mealybugs *Heliococcus bohemicus* (Sforza et al., 2003), *Phenacoccus aceris* (Sforza et al., 2003), *Plannococcus citri* (Cabaleiro et al., 1997) and *Pseudococcus calceolariae* (Petersen & Charles, 1997).

2.3.5 Detection techniques for GLRaV-3 and other grapevine infecting viruses

Serological and molecular detection methods for grapevine viruses have been developed during the past years that included the conventional enzyme-linked immunosorbent assay (ELISA), reverse transcription-polymerase chain reaction (RT-PCR), and even more sensitive assays with Real-time PCR.

Strategies for the detection of multiple grapevine viruses have been developed and tested and these approaches proved to reduce the cost of virus diagnostics dramatically. Examples of the simultaneous detection of viruses associated with GLR disease and other viruses infecting grapevine were described in several papers. These studies described a variety of techniques: one-tube RT-PCR assays (Nassuth et al., 2000), a spot-PCR technique (La Notte et al., 1997), the use of RT-PCR with degenerate primers for simultaneous detection of some members of

the *Clostero-, Viti-,* and *Trichovirus* genera (Salderelli et al., 1998), a spot multiplex nested RT-PCR for detection of viruses involved in the aetiology of GLR disease and RW of grapevine (Dovas & Katis, 2003) and a multiplex RT-PCR developed for the simultaneous detection of nine viruses (ArMV, GFLV, GVA, GVB, GRSPaV, GFkV, GLRaV-1, -2 and -3) (Gambino & Gribaudo, 2006). A more sensitive detection technique, TaqMan RT-PCR, was developed for the sensitive and quantitative detection of GLRaV-1 to GLRaV-5 and GLRaV-9 (Osman et al., 2007). An improvement on this technique was described a year later by the same authors. Low-density arrays have been designed based on real-time RT-PCR (TaqMan) assays for the specific detection of 13 viruses that infect grapevines (Osman et al., 2008). In a recent study, a diagnostic oligonucleotide microarray for the simultaneous detection of a wide range of grapevine viruses was developed (Engel et al., 2010). The microarray developed in this study contained probes designed against species-specific regions, to discriminate between closely related genus members, and against highly conserved regions at the family level, to enable the detection of highly divergent viruses or even previously unidentified viruses (Engel et al., 2010).

ELISA and RT-PCR are basic tools used in grapevine virus diagnostics. Recently the use of deep sequencing of an individual plant (Al Rwahnih et al., 2009) or pooled vines from a diseased South African vineyard (Coetzee et al., 2010) resulted in the identification of newly described viruses as well as determining the frequency in which viruses occurred in a vineyard. The use of the next generation high-throughput sequencing technologies proved to be a powerful tool to identify new viruses in disease complexes and to determine dominant variants of a specific virus. The deep sequencing analyses of 44 pooled vines from the South African study detected GLRaV-3 as dominant virus in the plants from this study (Coetzee et al., 2010).

Real-time RT-PCR has some advantages over conventional PCR and has been used in plant virus diagnostic methods in the last years (Osman et al., 2007). The high cost and expensive equipment needed to use advanced techniques is not feasible for routine diagnostic tests. In the South African context, the polyclonal antisera prepared against GLRaV-1, -2, and -3 (Goszczynski et al., 1995, 1997) are used widely by industry and growers to test with ELISA for these viruses.

2.4 GENETIC VARIABILITY OF PLANT VIRUSES

2.4.1 Definition of a virus variant

It is generally accepted that the genetic structure in a virus population may change with time. Most viruses continue to evolve through genetic exchanges and accumulation of mutations (Seo et al., 2009). Recombination plays a significant role in the evolutionary changes of RNA viruses (Worobey & Holmes, 1999; Chare & Holmes, 2006) and will be discussed in more detail. RNA viruses have genetically diverse populations due to an error-prone replication mechanism with high mutation rates, which causes these viruses to consist of many sequence variants around a consensus sequence (Komínek et al., 2005). This mixture of variants is usually termed quasispecies. A diverse quasispecies ensures better population fitness.

Methods developed to analyse nucleic acids in the 1970s had a big impact on understanding the evolution of plant viruses (García-Arenal & Fraile, 2008). These methods included ribonuclease T1 fingerprinting, restriction fragment length polymorphisms (RFLPs), ribonuclease protein assay (RPA) of a labelled complementary RNA probe, single strand conformation polymorphism (SSCP) analysis and nucleotide sequence determination of genes and entire genomes. In this study SSCP analysis and sequence determination of genes were used to study the variability of GLRaV-3 in South African vineyards.

2.4.2 The use of SSCP analysis in genetic variability studies

2.4.2.1 SSCP analysis as detection method of virus variants

SSCP analysis is one of the methods generally used to identify virus variants (García-Arenal et al., 2001). The analysis of a targeted genomic region with the SSCP technique was first established by Orita and colleagues (1989). SSCP analysis is a simple, reliable method for the detection of sequence variations in genomic loci. Another advantage of the technique is that PCR products from many isolates can be screened simultaneously to determine whether or not DNA fragments are identical in sequence. Pre-screening of isolates with SSCP analysis therefore reduces the amount of sequencing necessary (Sunnucks et al., 2000).

The SSCP technique was introduced soon after the introduction of PCR technology, and relied on the fact that relatively short DNA fragments can migrate in a nondenaturing gel not only as a function of their size but also their sequence (Garinis et al., 2005). In other words, amplified DNA fragments are denatured by heat or chemical agents, cooled down and the

Single-stranded DNA fragments are then electrophoresed through a nondenaturing polyacrylamide gel. Single-stranded DNA fragments adopt a specific three dimensional shape according to their nucleotide sequence with unique conformation. Even a single base difference will result in a different conformation and then migrate as different position during electrophoresis (Figure 7).

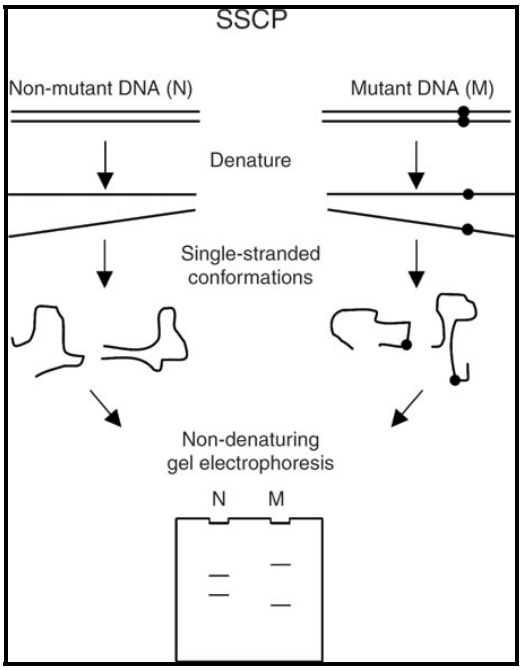


Figure 7. Schematic representation of the SSCP technique. A point mutation (represented by a dot on a DNA strand) leads to the formation of different single-strand conformations of the mutant DNA (M) compared with the non-mutant molecule (N), resulting in differential mobilities in a non-denaturing gel matrix (figure taken from Gasser et al., 2007).

2.4.2.2 Application of SSCP analysis in virus variability studies

SSCP analyses have been used in several genetic variability studies of viruses in the family *Closteroviridae*.

Sequence variability of the coat protein gene of 17 CTV isolates, a closterovirus, was studied and results showed that 1 to 59 nucleotide differences in their CP gene could be distinguished by SSCP analysis (Rubio et al., 1996). In a similar study on the CP gene of CTV isolates introduced into Morocco, SSCP analysis showed that each isolate consisted of several related genomic variants, typical of a quasispecies (Lbida et al., 2004). SSCP analysis was also applied in a study of the p27 gene (dCP) of CTV (Gago-Zachert et al., 1999) and could successfully distinguish between biologically mild and severe CTV isolates in this region.

The population structure and genetic diversity within Californian CTV isolates was studied with SSCP analysis of four genomic regions (Kong et al., 2000). In this study most CTV isolates were composed of a population of genetically related variants, one being predominant. The Kong et al (2000) study suggested that some CTV isolates could have arisen as result of a mixed infection of two divergent isolates. More recently, the population structure of CTV isolates from field Argentinean isolates was studied in three genomic regions of the virus (Iglesias et al., 2008). SSCP analysis showed that most isolates contained high intra-isolate variability. The SSCP technique was also applied to study the different genomic variants in clones from CTV variants (Černi et al., 2008).

The genetic variation of a crinivirus, *Cucurbit yellow stunting disorder virus*(CYSDV), was studied with the use of SSCP analysis (Rubio et al., 1999; Rubio et al., 2001). Genetic variation within individual CYSDV isolates and between CYSDV isolates collected in different years from different locations worldwide were studied. The molecular variability of the whitefly-transmitted *Beet pseudo-yellows virus*(BPYV), a closterovirus, and CYSDV were studied in cucurbits (Rubio et al., 1999). Based on SSCP profiles, CYSDV could be divided into three divergent groups and BPYV into two groups (Rubio et al., 1999).

The identification of two major sequence variants of GLRaV-3 infected vines from South Africa and world wide samples were initially done with SSCP analysis of a region in ORF5 (Jooste & Goszczynski, 2005). In this study it was possible to assign a specific SSCP profile to each of the variant groups. Sequence data confirmed these findings. In the same year a study on the genetic variability and population structure of GLRaV-3 isolates was investigated by SSCP analysis and sequence analysis of three genomic regions of the virus (Turturo et al., 2005). The authors came to the conclusion that GLRaV-3 that they have studied consist of a single undifferentiated population. The sequence variation in GLRaV-3 infected plants from New Zealand was studied with the SSCP technique and sequencing (Chooi et al., 2009). A third molecular variant of GLRaV-3 was identified from South African vineyards with SSCP analysis of individual clones from twelve isolates (Jooste et al., 2010).

The SSCP technique is a very useful tool for the rapid determination of the number of dominant sequence variants of GVA in virus-infected herbaceous host plants as well as in grapevines (Goszczynski & Jooste, 2002). The technique could also be used in an initial screening to discriminate between isolates of different origin and to analyse the genomic structure of each isolate (Lbida et al., 2004).

2.4.3 The role of recombination in shaping diversity

Recombination is one of the main factors in the evolution of positive-strand RNA viruses (Karasev, 2000). Homologous recombination, where the donor sequence replaces a homologous region of the acceptor sequence leaving its structure unchanged, and nonhomologous recombination, recombination between unrelated RNA sequences, are commonly observed (Lai, 1992). The recombination events in closteroviruseshave been studied rather well. The most direct evidence of recombination, in studies on closteroviruses, is the presence of defective RNAs (D-RNA) in infected cells or the exchange of viral genes with sgRNA in the process of replication (Yang et al., 1997) and secondly, the findings of chimeric genomes (Karasev, 2000).

2.4.3.1 Recombination studies in CTV, a Closterovirus

D-RNA was first isolated from a citrus plant infected with CTV (Mawassi et al., 1995). Although D-RNA has been mainly studied in CTV, they are probably characteristic of all closteroviruse*s* (Karasev, 2000). The possibility of chimeric genomes were suggested when two CTV isolates, VT and T36, were 90% identical on nucleotide sequence level in the 3' terminal fragment, compared to the 72% identity in the 5' terminal part (Mawassi et al., 1996). It was even suggested that the two CTV isolates were perhaps too dissimilar to remain the same virus. Two theories were discussed: the one suggesting an uneven evolution rate for the two halves of the genome and the other suggesting possible recombination between the CTV isolate and an unknown CTV genome (Mawassi et al., 1996). In a later study the diverse nature of the 5' terminal of isolate T36 was confirmed when a probe was developed in this region (Hilf et al., 1999). The T36-probe bound only to isolate T36. It was suggested T36 has arisen upon recombination between a normal CTV genome that provided all the 3' terminal genes and an unknown closterovirus that provided most of the ORF1a to form a chimeric genome.

Recombination has also been studied extensively in two CTV isolates, SY568 (Vives et al., 2005) and FS627 (Roy & Brlansky, 2009). The RNA population of isolate SY568 was found to be composed of two diverged sequence variants and different recombinants of them and this report showed the multiple recombination events within a natural virus isolate (Vives et al., 2005). In the study by Roy & Brlansky (2009), the generation of virus recombinants after aphid transmission was proved. Different dominant genotypes were detected in the parent and

aphid-transmitted (AT) subisolates and even intermediate genotypes were detected that differed from the parental or AT subisolates (Roy & Brlansky, 2009).

Advantageous genotypes can be created more rapidly by recombination than in clonal populations and harmful mutations can be removed by recombination with error–free parts of co-infecting genomes (Chare & Holmes, 2006). Recombination has proved to repair defective genes and generate beneficial new variation.

2.4.4 Molecular variability in 5'- and 3'- terminal regions

Information on the variability between CTV genomes are currently the best studied in the closteroviruse*s*. Similarities between genome comparisons of CTV isolates and GLRaV-3 were observed in literature and therefore recorded here. Nucleotide variability between two CTV isolates, T36 and VT, showed that the 5' ends of these isolates have a less than 70% nucleotide identity while the 3' end was relatively conserved (López et al, 1998). The length of the 5'UTR differed between the isolates, 107nt and 105nt, respectively. A feature of the 5'UTR of CTV is the high content of A (27-35%) and C (28-34%) in combination of the low content of G (14-18%). A deletion of seven nucleotides was observed in the T36 sequence and a proposed secondary structure with two stem-looped structures was identified (López et al, 1998). Some of the CTV isolates in this study contained sequences belonging to more than one variant group. Polymorphisms of the 5' terminal region of CTV confirmed the three molecular groups of CTV (Ayllon et al., 2001) although a recent phylogenetic analysis of complete CTV genome sequences showed the existence of more than three groups, with the addition of two New Zealand isolates (Harper et al., 2009). The length of the 5' and 3'-terminal regions of CTV is short (105nt) in comparison to the longer 5'UTR reported for GLRaV-3, isolate GP18 (Maree et al., 2008). The function of the extended 5'UTR of GLRaV-3 is not yet known.

2.4.5 Genetic variation between GLRaV-3 isolates

A focus area of this study is the genetic variation between GLRaV-3 variants.

Research world wide showed the existence of several molecular variants of GLRaV-3. Turturo et al. (2005) investigated the population structure and genetic variability of 45 GLRaV-3 isolates, from different grapevine varieties and 14 different countries, by single stranded conformation polymorphism (SSCP) and sequence analysis of three genomic regions, RdRp, Hsp70h and coat protein (CP). Their results for the RdRp and Hsp70h regions

showed that 10% of the isolates analysed had mixed variant infections, whilst 15% of the isolates had mixed infections when the CP region was analysed (Turturo et al., 2005). Multiple alignment of sequences deposited in Genbank® revealed that the sequences used in the Turturo study had nucleotide identities of above 90% between isolates in the regions studied. High diversity was noted in other studies, such as the divergent strain of GLRaV-3 (GLRaV-3-Tempr) (Genbank accession no. DQ314610), found in a grapevine accession in the cultivar Tempranillo from a Spanish vineyard (Angelini et al., 2006). The GLRaV-3-Tempr isolate was almost 20% divergent to the NY-1 isolate on nucleotide level in the sequenced 3' end of ORF1 (Angelini et al., 2006). GLRaV-3 infected juice grapes (*Vitis labruscana* 'Concord' and *Vitis labruscana* 'Niagara') from Washington State revealed nucleotide identities of 94 to 98% and amino acid identities of 97 to 98% in the Hsp70h gene of the NY-1 isolate (Soule et al.,2006). A survey of leafroll disease-associated viruses showed a 74.1-100% identity at the nucleotide level and 85.9-100% identity at the amino acid level between five GLRaV-3 isolates from New York and 25 isolates from other geographic regions (Fuchs et al.,2009a). Phylogenetic analysis of the HSP70h gene showed at least five possible variant groups in their study (Fuchs et al.,2009a). A study on the viral variants in the 'Waltham Cross' table grape variety, revealed at least two GLRaV-3 variants; one clone (WC-HSP-2) shared a 93.2% nucleotide identity with NY-1 (Ling et al.,2004) and two other clones (WC-HSP-10 and WC-HSP-28) were only 72.3% identical to NY-1 (Prosser et al.,2007). A nucleotide identity of 97.6% was reported between the Chilean isolate Cl-766 and NY-1 (Engel et al.,2008). Another study reported significant variability between New Zealand isolates where, to date, four genetic variants have been identified (Chooi et al.,2009). A study on Portuguese grapevine varieties, infected with GLRaV-3, identified five GLRaV-3 variant groups based on coat protein gene sequences (Gouveia et al.,2009).

From this summary it is clear that the variation in the GLRaV-3 genome is far greater than reported on in the earlier studies.

2.5 CONCLUSION

Grapevine is an important crop globally and is known to be the host of many diseases, especially virus diseases. The main virus diseases known to occur on grapevine includes devastating diseases like GLR disease, the focus of this study, shiraz disease, syrah decline, rugose wood (RW) complex including rupestris stem pitting (RSP) disease, corky bark, kober stem grooving and LN33 stem grooving syndrome. These diseases cause economic losses in worldwide grapevine production areas.

GLRaV-3, ampelovirus, was identified as being one of the main viruses in the GLR disease complex. A number of factors contribute to the spread of GLR disease, *i.e.* efficient virus vectors, infected planting material, virus variants and lack of control strategies. Studies on the mealybug vectors *P. ficus* and *P. longispinus* in South Africa showed that these species are very efficient in transmitting GLRaV-3. The importance of using virus-free planting material is supported by the certification scheme in the South African grapevine industry. Several epidemiology studies on GLR disease showed the importance of this disease worldwide and although control measures are stipulated for GLR disease, the disease still causes major infections in local vineyards.

The genetic variation of GLRaV-3 have been studied more intensively in the past years, since Ling et al. published the first full length sequence of isolate NY-1. New molecular evidence, based on the full length sequence of the South African isolate GP18, showed that the 5'UTR of this isolate is 579 nt longer that the sequence data reported for the NY-1 and Chilean isolate Cl-766. Phylogenetic studies on GLRaV-3 variants showed at least five to six groups of variants. There are a number of sequences that cluster outside these variant groups and further research may find that these isolates are part of a wider range of variation that exists around the world. It is important to understand the interactions between GLRaV-3 variants and insect vectors for control of the disease.

The literature review includes different aspects of the GLR disease complex: GLRaV-3, insect vectors, genetic variability and epidemiology studies and serves as background for further discussions in this dissertation.

2.6 LITERATURE

Al Rwahnih M, Uyemoto JK, Falk BW, Rowhani A (2007) Molecular characterisation and detection of plum bark necrosis stempitting-associated virus. Arch Virol 152:2197-2206

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

Angelini E, Bertazzon N, Bazzo I, Borgo M (2006)Molecular characterization of a divergent strain of grapevine leafroll-associated virus 3. In: Extended abstracts 15th Meeting of ICVG,Stellenbosch, South Africa, 3-7 April 2006, pp 148-150

Ayllón MA, López C, Navas-Castillo J, Garnsey SM, Guerri J, Flores R, Moreno P (2001) Polymorphisms of the 5' terminal region of *Citrus tristeza virus* (CTV) RNA: Incidence of three sequence types in isolates of different origin and pathogenicity. Arch Virol 146:27-40

Belli G, Fortusini A, Casati P, Belli L, Bianco PA, Prati S (1994) Transmission of a grapevine leafroll associated closterovirus by the scale insect *Pulvinaria vitis* L. Rivista di Patologia Vegetale4:105-108

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

Boyko VP, Karasev AV, Agranovsky AA, Koonin EV, Dolja VV (1992) Capsid protein gene duplication in a filamentous virus of plants. Proc. Natl. Acad.Sci. USA. 89:9156-9160

Cabeleiro C, Segura A (1997) Field transmission of Grapevine leafroll Associated virus 3 (GLRaV-3) by the mealybug *Planococcus citri*. Plant Dis81:283-287

Cabeleiro C, Segura A (2006) Temporal analysis of Grapevine leafroll associated virus3 epidemics. Eur J Plant Pathol 114:441-446

Cabeleiro C, Couciero C, Pereira S, Cid M, Barrasa M, Segura A (2008) Spatial analysis of epidemics of *Grapevine leafroll associated virus-3*. Eur J Plant Pathol 121:121-130

Carstens R (2002) Vineyard viruses: Leafroll overview.Wynboer - a technical guide for wine producers. January 2002. URL: http://www.winetech.co.za/news_leafroll.php3

Černi S, Ruščić J, Nolasco G, Gatin Z, Krajačić M, Škorić D (2008) Stem pitting and seedling yellows symptoms of *Citrus tristeza virus* infection may be determined by minor sequence variants. Virus Genes 36:241-249

Chare ER, Holmes EC (2006) A phylogenetic survey of recombination frequency in plant RNA viruses. Arch Virol 152:933-946

Cid M, Cabaleiro S, Faoro C, Segura A (2007) Presence of *Grapevine leafroll-associated virus-3* in primary salivary glands of the mealybug vector *Planococcus citri* suggests a circulative transmission mechanism. Eur J Plant Pathol 118:23-30

Charles JG, Cohen D, Walker JTS, Forgie SA, Bell VA, Breen KC (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

Chooi KM, Pearson MN, Cohen D, Pong JCH (2009) Sequence variation in *Grapevine leafroll-associated virus-3* (GLRaV-3) New Zealand isolates. In: Extended abstracts of the 16th Meeting of ICVG, Dijon, France, 31 Aug-4 Sept 2009, pp 290-291

Coetzee B, Freeborough MJ, Maree HJ, Celton JM, Rees DJG, Burger JT (2010) Deep sequencing analysis of viruses infecting grapevines: Virome of a vineyard. Virology 400:157-163

Daane KM, Bentley WJ, Walton VM, Malakar-Kuenen R, Millar JG, Ingels CA, Weber EA, Gispert C (2006) New controls investigated for vine mealybug. California Agriculture 60, (1):31-38

De Borbón CM, Gracia O, Gómez Talquenca GS (2004) Mealybug and grapevine leafroll-associated virus 3 in vineyards of Mendoza, Argentina. Am. J. Enol. Vitic 55:283-285

Dolja VV, Karasev AV, Koonin EV (1994) Molecular biology and evolution of Closteroviruses: Sophisticated build-up of large RNA genomes. Annu. Rev. Phytopathol. 32:261-285

Dolja VV, Kreuze JF, Valkonen JPT (2006) Comparative and functional genomics of closteroviruses. Virus Res 117:38-51

Douglas N, Krüger K (2008) Transmission efficiency of *Grapevine leafroll-associated virus 3* (GLRaV-3) by the mealybug *Planococcus ficus* and *Pseudococcus longispinus* (Hemiptera:Pseudococcidae). Eur J Plant Pathol 122:207-212

Dovas CI, Katis NI (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(2):217-226

Engel EA, Girardi C, Escobar PF, Arredondo V, Dominguez C, Perez-Acle T, Valenzuela PDT (2008)Genome analysis and detection of a Chilean isolate of *Grapevine leafroll associated virus-3*. Virus Genes 37:110-118

Engel EA, Escobar PF, Rojas LA, Rivera PA, Fiore N (2010) A diagnostic oligonucleotide microarray for simultaneous detection of grapevine viruses. J Virol Methods 163:445-451

Engelbrecht DJ, Kasdorf GGF (1990) Transmission of grapevine leafroll disease and associated closteroviruses by the vine mealybug *Planococcus ficus.*Phytophylactica22:341-346

Fuchs M, Martinson TE, Loeb GM, Hoch HC (2009a)Survey for the three major leafroll disease-associated viruses in Finger Lakes vineyards in New York. Plant Dis93:395-401

Fuchs M, Masella-Herrick P, Loeb GM, Martinson TE, Hoch HC (2009b) Diversity of Ampeloviruses in mealybug and soft scale vectors and in grapevine hosts from leafroll-affected vineyards. Phytopathol 99:1177-1184

Gago-Zachert SP, Costa NB, Semorile L, Grau O (1999) Sequence variability in p27 gene of *Cirus tristeza virus* (CTV) revealed by SSCP analysis. Electronic Journal of Biotechnology, Vol2 No1, Issue of April 15

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

García-Arenal, F, Fraile A, Malpica JM (2001) Variability and genetic structure of plant virus populations.Annu. Rev. Phytopathol 39:157-186

García-Arenal, F, Fraile A (2008) Questions and concepts in plant virus evolution: a historical perspective. In: Plant Virus Evolution, Chapter 1, Roosinck, MJ (Ed), Springer-Verlag Berlin Heidelberg, DOI 10.1007/978-3-540-75763-4

Garinis GA, Panayiotis G, Menounos G, Patrinos GP (2005) Mutation detection by SSCP and heteroduplex analyses. In: Molecular diagnostics, Patrinos GP, Ansorge W (Ed), Elsevier Academic press, pp55

Gasser RB, Hu M, Chilton NB, Campbell BE, Jex JA, Otranto D, Cafarchia C, Beveridge I, Zhu X (2007) Single strand conformation polymorphism for the analysis of genetic variation. Nature protocols 1:3121-3128

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

Golino DA, Weber E, Sim S, Rowhani, A (2008) Leafroll disease is spreading rapidly in a Napa Valley vineyard. California Agriculture 62(4):156-160. DOI:10.3733/ca.v062n04p156

Goszczynski D, Kasdorf GGF, Pietersen G (1995) Production and use of antisera to electrophoretically separated and transferred to nitrocellulose coat proteins of viruses associated with grapevine leafroll disease. African Plant Protection 1:1-8

Goszczynski D, Kasdorf GGF, Pietersen G (1997) Production and use of antisera specific to grapevine leafroll-associated viruses following electrophoretic separation of their proteins and transfer to nitrocellulose. In. "Filamentous viruses of woody plants". Ed. P. Monette. Research Signpost, Trivandrum, India, pp49-58

Goszczynski DE, Jooste AEC (2002) The application of single-strand conformation polymorphism (SSCP) technique for the analysis of molecular heterogeneity of grapevine virus A. Vitis 41:77-82

Gouveia P, Esteves F, Teixeira Santos M, Fonseca F, Eiras-Dias JE, Nolasco G (2009)Assessment of GLRaV-3 variants occuring in Portuguese grapevine varieties according to the coat protein gene. In: Extended abstracts of the 16th Meeting of ICVG,Dijon, France, 31 Aug-4 Sept 2009,pp 316-317

Harper SJ, Dawson TE, Pearson MN (2009) Complete genome sequences of two distinct and diverse *Citrus tristeza virus* isolates from New Zealand. Arch Virol 154:1505-1510

Hilf ME, Karasev AV, Albiach-Marti MR, Dawson WO, Garnsey SM (1999) Two paths of sequence divergence in the citrus tristeza virus complex. Phytopathol 89:336-342

Iglesias NG, Gago-Zachert SP, Robledo G, Costa N, Plata MI, Vera O, Grau O, Semorile L (2008) Population structure of *Citrus tristeza virus* from field Argentinean isolates. Virus Genes 26:199-207

Jarugula S, Gowda S, Dawson WO, Rayapati AN (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 Journal 7:180, DOI:10.1186/1743-422X-7-180

Jiang HE, Zhang YB, Li X, Yi-Feng Y, Ferguson DK, Lu EG, Li CS (2009) Evidence for early viticulture in China: proof of a grapevine (*Vitis vinifera* L. Vitaceae) in the Yanghai tombs, Xinjiang. J Archaeological Science 36:1458-1465

Johnson, H (2005) The story of wine – New Illustrated Edition. Octopus publishing, London

Jooste AEC, Goszczynski DE (2005) Single-strand conformation polymorphism (SSCP), cloning and sequencing reveal two major groups of divergent molecular variants of grapevine leafroll-associated virus 3 (GLRaV-3). Vitis 44(1):39-43

Jooste AEC, Maree HJ, Bellstedt D, Goszczynski DE, Pietersen G, Burger JT (2010) Three genetic grapevine leafroll-associated virus-3(GLRaV-3) variants identified from South African vineyards show high variability in their 5'UTR. DOI: 10.1007/s00705-010-0793-y. Arch Virol 155(12):1997-2006

Karasev AV, Boyko VP, Gowda S, Nikolaeva OV, Hilf ME (1995) Complete sequence of the citrus tristeza virus RNA genome. Virology 208:511-522

Karasev AV, Nikolaeva OV, Mushegian AR, Lee RF, Dawson WO (1996) Organization of the 3'-terminal half of beet yellow stunt virus genome and implications for the evolution of closteroviruses. Virology 221:199-207

Karasev AV (2000) Genetic diversity and evolution of Closteroviruses. Ann. Rev. Phytopathol 38:293-324

Komínek P, Glasa M, Bryxiova M (2005) Analysis of the molecular variability of *Grapevine leafroll-associated virus 1* reveals the presence of two distinct virus groups and their mixed occurrence in grapevines. Virus Genes 31:3, 247-255

Kong P, Rubio L, Polek M, Falk BW (2000) Population structure and genetic diversity within Californian *Citrus tristeza cirus* (CTV) isolates. Virus Genes 21:139-145

Klaasen VA, Boeshore M, Koonin EV, Tian T, Falk BW (1995) Genome structure and phylogenetic analysis of lettuce infectious yellows virus, a whitefly-transmitted, bipartite closterovirus. Virology 208:99-110

La Notte P, Minafra A, Salderelli P (1997) A spot-PCR technique for the detection of phloem-limited grapevine viruses. Journal of Virological Methods 66:103-108

Lai MMC (1992) RNA recombination in animal and plant viruses. Microbiological Reviews 56:61-79

Lbida B, Fonseca F, Santos C, Zemzami M, Bennani A, Nolasco G (2004) Genomic variability of *Citrus tristeza virus* (CTV) isolates introduced into Morocco. Phytopathol.Mediterr. 43:205-210

Ling KS, Zhu HY, Drong RF, Slightom JL, McFerson JR, 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 KS, Zhu HY, 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

López C, Ayllon MA, Nava-Castillo J, Guerri J, Moreno P, Flores R (1998) Molecular variability of the 5'-and 3'-terminal regions of *Citrus tristeza virus*. Phytopathol 88:685-691

Mahfoudi N, Digiaro M, Dhouibi MH (2009) Transmission of Grapevine Leafroll Viruses by *Planococcus ficus* (Hemiptera: Pseudococcidae) and *Ceroplastes rusci* (Hemiptera: Coccidae). Plant Dis. 93:999-1002

Maree HJ, Freeborough M-J, Burger JT (2008) Complete nucleotide sequence of a South African isolate of *grapevine leafroll-associated virus* 3 reveals a 5'UTR of 737 nucleotides.Arch Virol153:755-757

Maree HJ, Gardner HFJ, Freeborough M-J, Burger JT (2010) Mapping of the 5' terminal nucleotides of *grapevine leafroll-associated virus* 3 sgRNAs. Virus Research 151:252-255

Maliogka VI, Dovas CI, Katis NI (2008) Generic and species-specific detection of viruses belonging to an evolutionary distinct lineage within the *Ampelovirus* genus. J Virol Methods 154:41-47

Maliogka VI, Dovas CI, Lotos L, Efthimiou K, Katis NI (2009) Complete genome analysis and immunodetection of a member of a novel virus species belonging to the genus *Ampelovirus*. Arch Virol 154:209-218

Martelli GP, Agranovsky AA, Bar-Joseph M, Boscia D, Candresse T, Coutts RHA, Dolja VV, Falk BW, Gonsalves D, et al. (2002) The family *Closteroviridae* revised. Arch Virol 147/10

Martelli GP, Boudon-Padieu E (2006) Directory of infectious diseases of grapevines viruses and virus –like diseases of the grapevine: Bibliographic report 1998-2004. Options Méditerranéennes B55

Mawassi N, Mietkiewska E, Gifman R, Yang G, Bar-Joseph M (1996) Unusual sequence relationships between two isolates of *Citrus tristeza virus*. J Gen Virol 77:2359-2364

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

Nassuth A, Pollari E, Helmeczy K, Stewart S, Kofalvi SA (2000) Improved RNA extraction and one-tube RT-PCR assay for simultaneous detection of control plant RNA plus several viruses in plant extracts. J Virol Methods 90:37-49

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5:874-879

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, 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(2):292-299

Petersen CL, Charles JG (1997) Transmission of grapevine leafroll-associated closterovirus 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 –A technical guide for wine producers. June 2004.
URL:http://www.wynboer.co.za/recentarticles/articles.php3.

Pietersen G (2006) Spatio-temporal distribution dynamics of grapevine leafroll disease in Western Cape vineyards. In: Extended abstracts 15th Meeting of ICVG,Stellenbosch, South Africa, 3-7 April 2006, pp 126-127

Pietersen G (2010) Grapevine leafroll disease control in South Africa. Wineland magazine, October (2010) URL: http://www.wineland.co.za/index.php?option=com_zine&view=article&id=548:grapevine-leafroll-disease-control-in-south-africa&q=leafroll

Prossser SW, Goszczynski DE, Meng B (2007)Molecular analysis of double-stranded RNAs reveals complex infection of grapevines with multiple viruses. Virus Res124:151-159

Rayapati N, Tefera MA, Olufemi AJ, Sridhar J, Gandhi K, Gutha LR, Martin RR (2009) Current status of grapevine viruses in the Pacific Northwest vineyards of the United States. In: Extended Abstracts, 16th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine, 2009.Dijon, France, pp 108-109.

Roy A, Brlansky RH (2009) Population dynamics of a Florida *Citrus tristeza virus* isolate and aphid-transmitted subisolates: Identification of three genotypic groups and recombinants after aphid transmission. Phytopathol 99:1297-1306

Rubio L, Ayllon MA, Guerri J, Pappu H, Niblett C, Moreno P (1996) Differentiation of citrus tristeza closterovirus (CTV) isolates by single-strand conformation polymorphism analysis of the coat protein gene. Ann. Appl. Biol. 129:479-489

Rubio L, Soong J, Kao J, Falk BW (1999) Geographic distribution and molecular variation of isolates of three whitefly-borne closteroviruses of Cucurbits: *Lettuce Infectious Yellows Virus, Cucurbit Yellow Stunting Disorder virus,* and *Beet Pseudo-Yellows Virus*. Phytopathol 89:707-711

Rubio L, Abou-Jawdah Y, Lin HX, Falk BW (2001) Geographically distant isolates of the crinivirus *Cucurbit yellow stunting disorder virus* show very low genetic diversity in the coat protein gene. J Gen Virol 82(4):929-933

Salderelli P, Rowhani A, Routh G, Minafra A, Digiaro M (1998) Use of degenerate primers in a RT-PCR assay for the identification and analysis of some filamentous viruses, with special reference to clostero- and vitiviruses of the grapevine. Eur J Plant Pathol 104:945-950

Seo JK, Ohshima K, Lee HG, Son M, Choi HS, Lee SH, Sohn SH, Kim KH (2009) Molecular variability and genetic structure of the population of *Soybean mosaic virus* based on the analysis of complete genome sequences. Virology 393:91-103

Sether DM, Melzer MJ, Borth WB, Hu JS (2009) Genome organization and phylogenetic relationship of *Pineapple mealybug wilt-associated virus* 3 with family *Closteroviridae*. Virus Genes 38:414-420

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

Soule MJ, Eastwell KC, Naidu RA (2006) First report of *Grapevine leafroll associated virus-3* in American *Vitis* spp. Grapevines in Washington State. Plant Dis 90:1461

Sunnucks P, Wilson ACC, Beheregaray LB, Zenger K, French J, Taylor AC (2000) SSCP is not so difficult: the application and utility of single stranded conformation polymorphism in evolutionary biology and molecular ecology. Molecular Ecology 9:1699-1710

Swart E, Smit I (2006) The essential guide to South African wines. Chiviot publishing, Green Point, South Africa.

Tsai CW, Chau J, Fernandez L, Bosco D, Daane KM, Almeida RPP (2008) Transmission of *Grapevine leafroll-associated virus* 3 by the vine mealybug (*Planococcus ficus)*. Phytopathol98:1093-1098

Turturo C, Salderelli P, Yafeng D, Digiaro M, Minafra A, Savino V, Martelli GP (2005) Genetic variability and population structure of *grapevine leafroll-associated virus* 3 isolates. J Gen Virol86:217-224

Vives M, Rubio L, Sambade A, Mirkov TE, Moreno P, Guerri J (2004) Evidence of multiple recombination events between two RNA sequence variants within a *Citrus tristeza virus* isolate. Virology 331:232-237

Walter B, Martelli GP (1997) Clonal and sanitary selection of the grapevine, In: B. Walter (ed) Sanitary selection of the grapevine. Protocols for detection of viruses and virus-like diseases, pp 43-95. Les Colleoques, No. 86. INRA Editions, Paris, France

Walton VM, Pringle KL (2004) A survey of mealybugs and associated natural enemies in vineyards in the Western Cape Province, South Africa. South African Journal of Enology and Viticulture 25:23-25

Worobey M, Holmes EC (1999) Evolutionary aspects of recombination in RNA viruses. J Gen Virol 80:2535-2543

Yang ZN, Mawassi M, Gofman R, Gafny R, Bar-Joseph M (1997) Involvement of a subgenomic mRNA in the generation of a variable population of defective citrus tristeza virus molecules. J Virol 71:9800-9802

CHAPTER 3

Three genetic grapevine leafroll-associated virus 3 (GLRaV-3) variants identified from South African vineyards show high variability in their 5'UTR

Arch Virol 155(12):1997-2006 (DOI: 10.1007/s00705-010-0793-y)

A.E.C. Jooste[1], H.J. Maree[2], D.U. Bellstedt[3], D.E. Goszczynski[1], G. Pietersen[1], J.T. Burger[2]

[1]ARC-Plant Protection Research Institute, Private Bag X134, Queenswood, Pretoria, 0121, South Africa
[2]Department of Genetics, Stellenbosch University, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa
[3]Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa

3.1 ABSTRACT

Three genetic variants of grapevine leafroll-associated virus 3(GLRaV-3) were identified in vineyards of the Western Cape, South Africa. The GLRaV-3 variants were identified by single strand conformation polymorphism (SSCP) profiles generated from a region amplified in ORF5. ORF5 sequence data confirmed the three genetic variant groups and a specific SSCP profile was assigned to each variant group. The results of the SSCP analysis on the region in ORF5 showed that this method gives a fast and reliable indication of the GLRaV-3 variant status of a plant, which in many instances showed mixed infections. The full genome sequence of one representative of each variant group *i.e.* isolates 621 (group I), 623 (group II) and PL-20 (group III), was determined by sequencing overlapping cloned fragments of these isolates. The sequences of genomic 5' ends of these isolates were determined by RLM-RACE. Sequence alignment of the 5'UTRs indicated significant sequence and length variation in this region between the three South African variant groups. Alignment of the Hsp70h and CP gene regions of these isolates with those of isolates from elsewhere in the world, followed by phylogenetic analysis, further supported the presence of three variants of GLRaV-3 in South Africa and that two or three additional variant groups occurred elsewhere in the world.

3.2 INTRODUCTION

Grapevine leafroll-associated virus 3 (GLRaV-3) is the main causative agent of grapevine leafroll disease (leafroll) in South Africa and spreads rapidly in local vineyards (Pietersen, 2004). As GLRaV-3 is consistently associated with leafroll it is viewed to be an important etiological agent in this economically important disease (Boscia et al., 1995). GLRaV-3 is the type species of the genus *Ampelovirus* in the family *Closteroviridae* (Martelli et al, 2002). Virions are flexuous, filamentous particles, about 1800 nm long with a positive-sense single stranded RNA genome organised into 13 open reading frames (Ling et al., 2004).

Research world wide showed the existence of several molecular variants of GLRaV-3. The population structure and genetic variability of 45 GLRaV-3 isolates, from different grapevine varieties and 14 different countries, by single stranded conformation polymorphism (SSCP) and sequence analysis of three genomic regions, RdRp, Hsp70h and coat protein (CP) was investigated (Turturo et al., 2005). The results for the RdRp and Hsp70h regions showed that 10% of the isolates analysed had mixed variant infections, whilst 15% of the isolates had mixed infections when the CP region was analysed (Turturo et al., 2005). Multiple alignment of sequences deposited in Genbank® revealed that the sequences used in the Italian study had nucleotide identities of above 90% between isolates in the regions studied. High diversity was

noted in other studies, such as the divergent strain of GLRaV-3 (GLRaV-3-Tempr) (Genbank accession no. DQ314610), found in a grapevine accession in the cultivar Tempranillo from a Spanish vineyard (Angelini et al., 2006). The GLRaV-3-Tempr isolate was almost 20% divergent to the NY-1 isolate on nucleotide level in the sequenced 3' end of ORF1 (Angelini et al., 2006). GLRaV-3 infected juice grapes (*Vitis labruscana* 'Concord' and *Vitis labruscana* 'Niagara') from Washington State revealed nucleotide identities of 94 to 98% and amino acid identities of 97 to 98% in the Hsp70h gene of the NY-1 isolate (Soule et al., 2006). A survey of leafroll-associated viruses showed a 74.1-100% identity at the nucleotide level and 85.9-100% identity at the amino acid level between five GLRaV-3 isolates from New York and 25 isolates from other geographic regions (Fuchs et al., 2009). Phylogenetic analysis of the HSP70h gene showed at least five possible variant groups in their study (Fuchs et al., 2009). A study on the viral variants in the 'Waltham Cross' table grape variety, revealed at least two GLRaV-3 variants; one clone (WC-HSP-2) shared a 93.2% nucleotide identity with NY-1 (Ling et al., 2004) and two other clones (WC-HSP-10 and WC-HSP-28) were only 72.3% identical to NY-1 (Prosser et al., 2007). A nucleotide identity of 97.6% was reported between the Chilean isolate Cl-766 and NY-1 (Engel et al., 2008). Another study reported significant variability between New Zealand isolates where, to date, four genetic variants have been identified (Chooi et al., 2009). A study on Portuguese grapevine varieties, infected with GLRaV-3, identified five GLRaV-3 variant groups based on coat protein gene sequences (Gouveia et al., 2009).

In a previous study of South African isolates of GLRaV-3, single strand conformation polymorphism (SSCP), restriction enzyme (RE) SSCP, cloning and sequencing techniques were used to identify two clearly divergent molecular groups of the virus (Jooste & Goszczynski, 2005). The first molecular variant, represented by isolate 621, was very similar to the NY-1 isolate of Ling et al. (2004) while sequence data of the second molecular variant, represented by isolate 623, was very similar to the complete genome sequence of the South African isolate GP18 (Maree et al., 2008). The molecular divergence between these two variant groups was especially high in the 5' terminal part (partial sequences of the 5'UTR and ORF1a of the virus where nucleotide sequences differed by 35%. Sequence data of the remaining coding regions showed nucleotide similarities above 90% between the variant groups. The two variant groups could be distinguished by unique SSCP profiles generated from an amplified region in ORF 5 (Jooste & Goszczynski, 2005).

The complete genome length of a South African GLRaV-3 isolate, GP18 (EU259806), was reported to be 18498 nt (Maree et al., 2008). The extended length of the 5'UTR, consisting of 737 nt, differed from that reported previously by Ling et al. (2004) and Engel et al. (2008) where a 5'UTR of 158 nt for both isolates NY-1 (AF037268) and Cl-766 (EU344893) was described. There was a 93% sequence identity between the genome sequences of isolate GP18 and NY-1 (Maree et al., 2008).

In this study, the variability of GLRaV-3 in infected plants in South African vineyards was further explored. The aim of the study was to use the SSCP technique to identify GLRaV-3 variants from infected vineyards and to obtain full genome sequences of three representatives of these GLRaV-3 variants, namely isolates 621, 623 and PL-20, including their 5'UTRs. The phylogenetic relationships of these three isolates were determined using Hsp70h and CP gene sequences which were aligned with those of isolates obtained from elsewhere in the world.

3.3 MATERIALS AND METHODS

3.3.1 Virus sources

A field survey was conducted in five mother blocks in different grape producing areas of the Western Cape, South Africa, namely; Stellenbosch, Paarl, Wellington, Rawsonville and Worcester. GLRaV-3 infected plants were selected based on different spatial distribution patterns of leafroll recorded in a survey done in mother blocks from 2001-2007 (Pietersen, 2004). In numerous disease foci (also referred to as disease clusters) the infection point or starting point of disease spread began from a single plant. From this single plant mealybugs transmitted the disease to adjacent plants in rows and across rows forming clusters of infection. The plants collected in this survey all represent the initial infected plant of any given foci of leafroll infection.

Based on SSCP profiles, twelve of these infected plants (3,4,5,7,12,15,16,17,20,32,48, and 50) were selected for further investigation (Table 1). The existence of an extended 5'UTR was studied in these twelve plants and in 57 additional field-collected GLRaV-3 isolates.

Table 1.A description of the twelve plants selected for further investigation, in which cultivar, year planted and collection region are listed

Plant number	Position description (Block/Row/Plant no)	Cultivar	Region	Year planted
3	52/28/95	Cabernet Sauvignon	Paarl	1997
4	52/30/64	Cabernet Sauvignon	Paarl	1997
5	52/53/26	Cabernet Sauvignon	Paarl	1998
7	52/63/89	Cabernet Sauvignon	Paarl	1998
12	10/9/87	Cabernet Sauvignon	Worcester	1996
15	10/23/10	Cabernet Sauvignon	Worcester	1997
16	10/29/2	Cabernet Sauvignon	Worcester	1997
17	10/31/35	Cabernet Sauvignon	Worcester	1997
20	10/39/99	Cabernet Sauvignon	Worcester	1992
32	19b/6/15	Merlot	Rawsonville	1994
48	1/23/63	Cabernet Sauvignon	Stellenbosch	1991
50	1/29/83	Cabernet Sauvignon	Stellenbosch	1991

3.3.2 dsRNA isolation, RT-PCR, cloning and SSCP analysis

Isolation of dsRNA, SSCP analysis and cloning were carried out as described earlier (Goszczynski & Jooste, 2002). SSCP profiles were generated in 15% polyacrylamide gels from GLRaV-3 amplified products in ORF5, genome position 12592-12801 (primer pair H420/C629 in Jooste & Goszczynski (2005). Additionally, the variability in eleven regions of the genomes of the twelve plants listed in Table 1 was further investigated with SSCP analysis and sequencing. The primers used in this study are listed in the supplementary data as Table 1.

3.3.3 Cloning of fragments of isolates 621, 623 and PL-20 and assembly into whole-genome sequences

Based on the SSCP results from the field survey, the complete nucleotide sequences of three GLRaV-3 isolates were determined: GLRaV-3 isolate PL-20, from a *Vitis vinifera* cv. Cabernet Sauvignon plant (Table 1), and GLRaV-3 isolates 621 and 623, from *Vitis vinifera* cv. C. Sauvignon and *Vitis vinifera* cv. Ruby Cabernet, respectively (Jooste & Goszczynski, 2005).

dsRNA and total RNA were isolated from these samples and used for RT-PCR to clone fragments of isolates 621, 623 and PL-20 using standard molecular techniques.

Primer design for cloning and sequencing of the genome fragments of isolates 621 and 623 was done with the Genefisher program (Giegerich et al., 1996) using the published sequences of isolate NY-1 (Ling et al., 2004) as a template. The sequencing strategy for fragments of isolates 621 and 623 was based on the amplification, cloning and sequencing of 18

overlapping clones in the genomes of isolates 621 and 623, with product sizes ranging from 283 to 986 bp (primer sequences not shown). Cloning of PCR-amplified products was carried out using a pGEM®-T Easy cloning system (Promega). At least three clones of each amplicon were sequenced in both directions. Consensus sequences of isolates 621 and 623, similar to the nucleotide length of NY-1 (Ling et al., 2004), were assembled with DNAMAN version 6 sequence analyses software (Lynnon Biosoft, 1996) from the overlapping sequences generated from the cloned fragments.

The sequence of isolate PL-20 was compiledfrom the amplification and sequencing of ten overlapping cloned fragments (Figure 3). Primer design was done with Genefisher (Giegerich et al., 1996) and OligoExplorer 1.1.0 Software (http://www.genelink.com/tools/gl-oe.asp) with the GP18 isolate sequence as a template (Maree et al.,2008). Primers designed to amplify the fragments of isolate PL-20 in ten overlapping clones are listed in Table 2. The PCR products (Table 2) were ligated into the pDrive vector, using a Qiagen® PCR Cloning kit according to the manufacturer's instructions and transformed into competent DH5α cells. At least three clones of each amplicon were sequenced in both directions with the SP6 and T7 primers, and with additional primers where required. The additional sequencing primers that were designed to sequence the larger cloned fragments are listed in Table 3.

Table 2. Primers designed to amplify and sequence ten overlapping clones of isolate PL-20. Primer positions are based on the genome sequence of GLRaV-3 isolate GP18 (EU259806)

Clone	Primer	Sequence (5' to 3')	Product size(bp)	Position of amplified products
K	GL3.5	TGCTCTAGTAGGATTCGAAC		
	GL3H.R	CGATCAATAGACCTCCTCTTG	1976	ORF1a (6-1973)
J	GL3J1814F	GTCACTCAATGAAGAGCGCACC		
	GL3J3108R	CAACCTTGCTGTGGACGCC	1294	ORF1a (1814-3108)
I	GL3J.F	GACTTTGTCGACAGGATC		
	GL3.4133R	CACGCTTCGAGGTGAATGG	1763	ORF1a (2967-4730)
G	GL3.3874F	GGGGCTTGCTTAACGACAC		
	GL3.5591R	AACGCCCTGTATGTCCTCTC	1734	ORF1a (4453-6187)
F	GL3.5297F	GTCACCAGGTGTTCCAAACC		
	GL3.7707R	CCTGCTTCATGAGAGCACTC	2429	ORF1a/b (5876-8305)
E	GL3.7466F	CGCCATTGTCGAAGTACG		
	GL3.3RRev	GAATACTTCCACAGCCCTAG	3534	ORF1b/4 (8045-11579)
D	GL3.3RFor	TGCTGGTTCTTACGTTCG		
	GL3.12524R	TGACCAGCTTGAGCGTAG	2547	ORF3/5 (10573-13120)
C	GL3.12212F	CGATCGTGCCGTTAAGAG		
	GL3.15218R	GCCTTTCGAAAGAACGAGTC	3026	ORF5/7 (12791-15817)
B	GL3B15705F	CGGACTCAGTGTTGTCAATCAGTG		
	GL3B17172R	GCTTCCTCCAACGCTGAAACAG	1467	ORF7/9 (15705-17172)
A	GL3.16399F	GGGTGCCGTATTAAGAGACTC		
	3'UTRev	AAGGGCCCGACCTAACTTATTGTCGATAAG	1520	ORF9/3'UTR (16978-18498)

Table 3. Additionalprimers designed for sequencing the gaps in the large clones generated for PL-20. The positions of the primers are indicated in the primer name according to positions on the GP18 (EU259806) sequence

Clone	Primer	Sequence (5' to 3')
K	GL3K790F	TTTACGGGGAATGTGAAGTTG
F	GL3F6544F	CTCCTCCAGGTGGTGGTAAGAC
E	GL3E8837F	CGCGGAAATATAAGCATTCGGGATG
	ORF2+3Rev	GGGGAAAAGCAGATTGTGC
D	GL3D11187F	CCGACAGCCGCAGCCCTCTATTC
	GL3D12353R	GCCTCGTTAGGGTTCAATATCC
C	GL3C13579F	CACAGGAACAATTGGAAGACGCTG
	GL3C15054R	CTGGCTATCGATAACGTAACCG

To determine the sequences of the 5' and 3' ends of isolates 621 and 623, poly (A) tailing was attempted (Ling et al., 1998; Ling et al., 2004). The 5'ends of isolates 621, 623 and PL-20 were amplified using the RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) with the First Choice® RLM-RACE kit (Ambion, USA) according to the manufacturer's instructions, using total RNA extracted from 2 g of phloem tissue by the CTAB method (White et al., 2008). The genome-specific reverse primers used for the 5' RLM-RACE kit to perform the 5' RACE PCRs, were LR3 365Rev, 5' CGTCCGCTTCACCCCTTTGG 3' (used in the outer PCR reaction), and LR3 868, 5' GGGTGTGAAGTCAGATAACTT 3' (in the inner PCR reaction). The PCR products of the 5'UTR region were ligated into the pDrive vector and sequenced as described above. Sequences were analysed and assembled with Vector NTI *v10* (Invitrogen) and BioEdit software (Hall, 1999), and consensus sequences of the 5' and 3' ends of isolates 621, 623 and PL-20 were compiled.

The 3' ends of isolates 621 and 623 were successfully determined with the polyadenylation and reverse transcription of purified dsRNA (Ling et al., 1998). The 3' end of isolate PL-20 was cloned with the extension of polyA cDNA.

The whole-genome sequences of GLRaV-3 isolates 621, 623 and PL-20 were deposited in the GenBank® database and assigned accession numbers GQ352631, GQ352632 and GQ352333. The predicted functions of the ORFs were confirmed with the Conserved Domain Search on the NCBI website (Marchler-Bauer, 2009; Marchler-Bauer & Bryant, 2004). The nucleotide (nt) and amino acid (aa) sequences of the ORFs of these isolates, those of the Chilean isolate C1-766 and the South African isolate GP18 were compared with those of the USA isolate

NY-1, and were expressed as percentage sequence identity to isolate NY-1 and were listed in tabular form (Table 6).

3.3.4 Confirmation of 5'UTR in field-collected samples

Primers were designed to amplify a 362-422nt fragment (depending on the variant group) in the 5'UTR of selected GLRaV-3 isolates by RT-PCR. These primers were designed in conserved regions, identified by multiple sequence alignment of the 5'UTRs of isolates 621, 623, PL-20 and GP18, and were designated GL3.5F (5' TGCTCTAGTAGGATTCGAAC 3') and GL3.342R (5' CCCAACACGATAAAGAGAAC 3'). PCR products were sequenced and sequence data were analysed to identify variant groups.

3.3.5 Phylogenetic analysis

In order to assess the relationship of the three South African isolates, their Hsp70h and CPgene sequences were used in a phylogenetic analysis in which the Hsp70h and CP sequences of a large number of isolates from elsewhere in the world were included. The origins of the Hsp70h and CPsequences and Genbank accession numbers are shown in Table 4 and 5 respectively. These Hsp70h and CP sequences of isolates 621, 623 and PL-20 as representatives of the South African variant groups I, II and III, respectively, were aligned with the Hsp70h and CP sequences downloaded from GenBank (Table 4, 5), respectively utilizing the software package BioEdit (Hall, 1999). These sequences were aligned using the Clustal W (v 1.4) alignment function embedded within the BioEdit package and the alignment was refined manually. Phylogenetic analyses of the aligned Hsp70h and CP sequence matrices were performed using PAUP 4.0b10 (Swofford, 2003). In both analyses theGLRaV-1 Hsp70h and CPsequence (GenBank accession no. AF 195822) was used as the outgroup. A heuristic search (1 000 replicates) using TBR branch swapping with all characters weighted equally was performed to search for the shortest possible trees from both data matrices. A bootstrap analysis (1 000 replicates) using TBR branch swapping was performed to establish clade support. Branches with bootstrap values $\geq$ 75% were considered as well supported, whilst values between 75% and 50% were considered as moderately supported. Values below 50% were considered weakly supported and in line with other phylogenetic analyses were not indicated on phylograms.

Table 4. Virus isolates used in this study to determine the phylogenetic relationships between the different variant groups of GLRaV-3 based on the Hsp70h gene

Isolate	Cultivar	Country	GenBank accession no.
NY-1	Pinot noir	U.S.A	AF037268
GP18	Cabernet Sauvignon	South Africa	EU259806
621	Cabernet Sauvignon	South Africa	GQ352631
623	Ruby Cabernet	South Africa	GQ352632
PL-20	Cabernet Sauvignon	South Africa	GQ352633
Cl-766	Merlot	Chile	EU344893
NZ-1	Unknown	New Zealand	EF508151
MT48-1	LN33	Italy	AJ748518
MT48-2	LN33	Italy	AJ748519
MT48-3	LN33	Italy	AJ748520
MT48-4	LN33	Italy	AJ748521
C1-1	Unknown	U.S.A	DQ780885
C2-1	Unknown	U.S.A	DQ780886
C3-1	Unknown	U.S.A	DQ780887
C4-1	Unknown	U.S.A	DQ780888
C5-1	Unknown	U.S.A	DQ780889
C6-1	Unknown	U.S.A	DQ780890
N1-1	Unknown	U.S.A	DQ780891
TU32	Unknown	Tunisia	AJ748522
AUSG5-2	Unknown	Austria	AJ748510
AUSG5-4	Unknown	Austria	AJ748511
AUSG5-5	Unknown	Austria	AJ748512
AUSG5-6	Unknown	Austria	AJ748513
SY2-2	Homos	Syria	AJ748515
SY2-4	Homos	Syria	AJ748516
SY2-7	Homos	Syria	AJ748517
IL1	Hillmanis	Israel	AJ748524
USA6	Shermann	U.S.A	AJ748523
GLRaV-1	Unknown	Australia	AF195822
C3	Unknown	China	AJ748514

Table 5. Virus isolates used in this study to determine the phylogenetic relationships between the different variant groups of GLRaV-3 based on the coat protein gene

Isolate	Cultivar	Country	GenBank accession no.
NY-1	Pinot noir	U.S.A	AF037268
GP18	Cabernet Sauvignon	South Africa	EU259806
621	Cabernet Sauvignon	South Africa	GQ352631
623	Ruby Cabernet	South Africa	GQ352632
PL-20	Cabernet Sauvignon	South Africa	GQ352633
Cl-766	Merlot	Chile	EU344893
NZ-1	Unknown	New Zealand	EF508151
MT 48	Unknown	Italy	AJ606351
MT 38	Unknown	Italy	AJ606350
MN 18	Unknown	Italy	AJ606349
SS 5.1	Unknown	Italy	AJ606348
SS 5.2	Unknown	Italy	AJ606347
TA 3.3	Unknown	Italy	AJ606346
TA 3.1	Unknown	Italy	AJ606345
TA 3.2	Unknown	Italy	AJ606344
GR 1.2	Unknown	Greece	AJ606343
GR1.1	Unknown	Greece	AJ606342
NIG 3.1	Unknown	Nigeria	AJ606341
NIG 3.2	Unknown	Nigeria	AJ606358
USA	Unknown	Tromelin Island	AJ606340
AUSG 5.2	Unknown	Austria	AJ606339
AUSG 5.2	Unknown	Austria	AJ606338
LN	Venus seedless	China	FJ786016
SL10	Unknown	China	DQ911148
Pet-3	Unknown	Brazil	DQ062152
Pet-4	Unknown	Brazil	AY753208
Pet-1	Unknown	Brazil	DQ680141
Cl-765	Merlot	Chile	EU344896
Cl-644	Merlot	Chile	EU344895
Cl-817	Chardonnay	Chile	EU344894
Dawanhong 2	Unknown	China	DQ119574
CH 5.2	Unknown	China	AJ606357
CH 5.1	Unknown	China	AJ606356
IL 1.1	Unknown	Israel	AJ606355
IL 1.2	Unknown	Israel	AJ606354
TU 16	Unknown	Tunisia	AJ606353
SY 2.3	Unknown	Syria	AJ606352
GLRaV-1		Australia	AF195822

3.4 RESULTS

3.4.1 Field survey, SSCP analysis

SSCP profiles performed on the isolated dsRNA of 46 GLRaV-3 infected plants showed two distinct profiles (Figure 1a, lane 4 and 5) as previously described (Jooste & Goszczynski, 2005), as well as additional (hitherto unreported) profiles (Figure 1a-d). The SSCP profiles consisted of 'simple' and 'complex' patterns, as described previously(Turturo et al., 2005). 'Complex' and atypical profiles were selected for further analyses in this study. The SSCP profiles of the twelve plants selected for further analysis are indicated by arrows in Figure 1a-d (3, 4, 5, 7, 12, 15, 16, 17, 20, 32, 48 and 50). The SSCP profiles of plants 4 and 5 represent the two variants, isolates 621 (group I) and 623 (group II), previously identified (Jooste & Goszczynski, 2005).

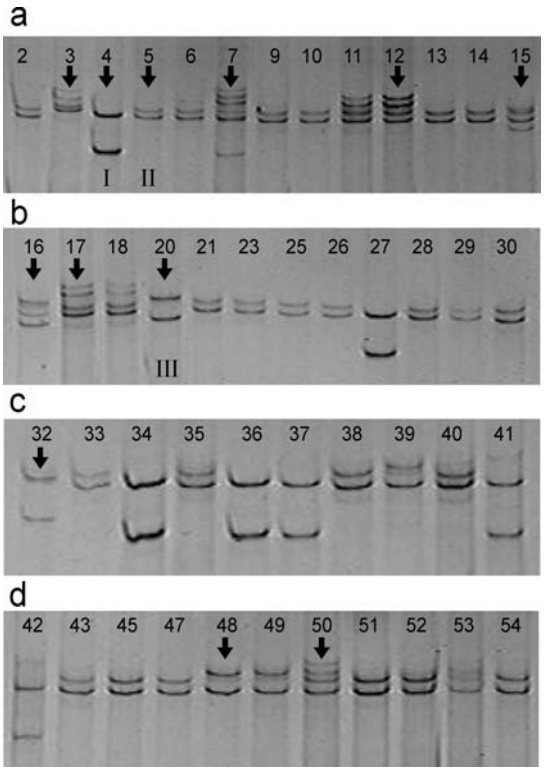


Figure 1a-d. SSCP results of plants collected in vineyards in the (a) Paarl and Worcester regions, lanes 2-10 and 11-15, respectively, (b) Worcester and Wellington regions, lanes 16-20 and 21-30, respectively, (c) Rawsonville region, lanes 32-41 and (d) Stellenbosch region, lanes 42-54.

Ten clones from each of the twelve plants were generated and individually analysed by SSCP. A minimum of four clones per plant were selected for sequencing of this region (indicated by circled numbers in Figure 2). Variability in SSCP profiles between the clones was detected in most of the plants (4, 7, 12, 15, 16, 32, and 48), illustrating that combinations of variants occurred in different plants. In some of the plants (5, 17, 20 and 50) the SSCP profiles of all

10 selected clones were identical. The nucleotide sequences of 55 clones and four reference sequences (Group I: NY-1 and 621, Group II: GP18 and 623) were compared. Sequence alignment showed that the majority of the clones grouped with the variant group II isolates 623 and GP 18 (Supplementary data, Figure 1). A third variant group was identified based on sequence results and distinct SSCP profiles detected in plants 15, 16, 20 and 32. The SSCP profiles shown in Figure 2 of clones 15.2, 15.9, 16.1, 16.5, 16.7, 20.1, 20.4, 20.7, 20.9, 32.5 and 32.8 represent this third group of variants. Based on these results, the viral isolate from plant 20 (referred to as isolate PL-20) was selected as a representative of the third variant group. Sequences of other clones from plant 15, 16 and 32 clustered with sequences of clones in group I. Nucleotide sequences from each clone and the SSCP profiles (Figure 2) of individual clones of the three variant groups, correlated. SSCP profiles of clones from plant 12 were all similar, except for 12.8, which had a single nucleotide change, resulting in a different SSCP profile.

Clones derived from plant 4, *i.e.* 4.1, 4.3, 4.6 and 4.10, showed SSCP profiles typical for variant group I and clustered into the group that was similar to isolates 621 and NY-1. Clones 4.5, 4.7 and 4.9 showed atypical SSCP profiles for group I and sequence results confirmed that these clones were similar to clones from variant group II. The alignment of clones from plant 4 is shown in Figure 2 as supplementary data to this chapter.

The SSCP profiles determined previously for the three variant groups are indicated in the profiles of plants 4, 5 and 20 (Figure 1a, b).

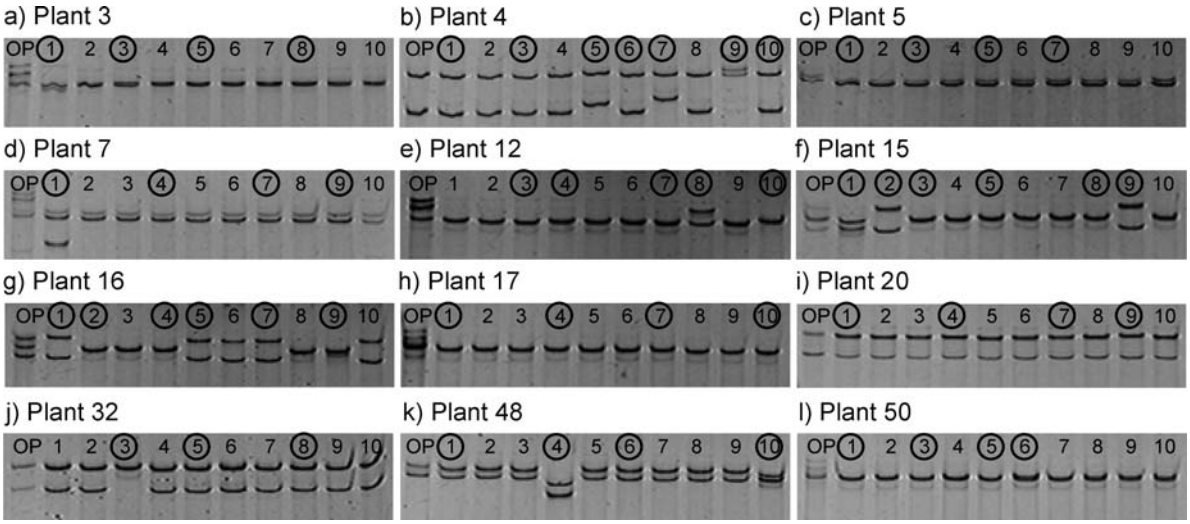


Figure 2. SSCP profiles of ten clones of each of the twelve plants. Clones indicated with circles were sequenced. The first lane of each gel represents the original SSCP profile (OP) of each plant.

The bands seen in the SSCP profiles of the original plants 3,7,12 and 17 were not all found in the 10 cloned samples. This is because the expected SSCP pattern should be composed of two or three bands (one of them being a 'conformomer', i.e a conformational structure of the same sequence as described by Orita et al., 1989) (Turturo et al., 2005). A closer look at the SSCP profiles of these plants show these conformational structures of the same sequence as light bands, some more visible than others.

Additional SSCP and sequence data were analysed to investigate the variability of the viral genomes present in the twelve selected plants. The objective was to investigate how SSCP analysis and sequencing results compared in identifying GLRaV-3 variants. Sequencing results, presented as phylogenetic analyses, and SSCP analysis of primer pairs 2 in ORF1a (259 nt), 6 in ORF5 (312 nt), 8 in ORF7 (326 nt), 10 in ORF9(254 nt) and 11 in ORF10 (212 nt) are shown in Figures 3-7 in the supplementary data. At least four or more different SSCP profiles were detected in the twelve plants analysed. Phylogenetic analysis of the sequence results from the plants in these regions confirmed three clear variant groups in all the regions studied. The additional sequence data confirmed the three variant groups detected in ORF5 with H420, C629 primers.

3.4.2 Whole-genome sequencing results and variability in the 5'UTR

Whole-genome sequences of isolates 621, 623 and PL-20 were successfully generated, each representing one of the variant groups identified in the SSCP studies. The genome organisation of the three GLRaV-3 variants was similar to that described previously for isolates NY-1, Cl-766 and GP18 (Ling et al., 2004; Engel et al., 2008; Maree et al., 2008). Similar to isolate GP18, all three isolates contained an extended 5'UTR. ORF1a of isolates 621 (group I) and 623 (group II) started at nucleotide position 738 on the respective genomes and at position 673 on the PL-20 genome (group III). In this region, the methyltransferase (MET), AlkB (2OG-FeII) and helicase (HEL) domains were found in all three isolates. The other domains were similar to the NY-1, GP18 and Cl-766 isolates (Ling et al., 2004; Maree et al., 2008; Engel et al., 2004) and included the two intergenic regions from positions 9058-9286 and 9443-10508.

The variation between the three variant groups compared to isolate NY-1 is shown in Table 6. The percentage sequence identity in the nt and aa sequences in ORF1a showed no clear differentiation between group I and II variants. In ORF2 the aa prediction for the variant group II isolates, GP18 and 623, differed by almost 20% to group I and III variants. The

genomic regions towards the 3'UTR, namely ORF10, ORF11 and ORF12, showed the most variation between variants in the nt and aa sequences and the group III variant showed 80%, 67% and 78% aa identity, respectively, to the NY-1 isolate in these regions.

Table 6. The nucleotide (nt) and amino acid (aa) percentage sequence identity of the GLRaV-3 isolates, Cl-766 (EU344893), 621 (GQ352631), GP18 (EU259806), 623 (GQ352632) and PL-20 (GQ352633), representing the three variant groups, as compared to isolate NY-1 (AF037268)

GLRaV-3 isolate	**ORF1a Met/Hel nt/aa**	**ORF1b RdRp nt/aa**	**ORF2 p6 nt/aa**	**ORF3 p5 nt/aa**	**ORF4 Hsp70h nt/aa**	**ORF5 Hsp90h nt/aa**	**ORF6 CP nt/aa**	**Variant group**
Cl-766	95.7/95.5	99.0/99.1	98.7/98.0	97.1/100	98.8/98.4	99.2/99.0	99.0/99.0	I
621	96.2/96.5	99.4/100	98.7/98.0	97.1/100	98.8/98.7	99.2/99.2	99.3/99.0	
GP18	94.1/95.3	95.4/97.6	90.4/80.4	93.5/97.8	94.6/96.9	92.8/93.2	92.5/94.6	II
623	93.4/95.6	90.3/95.7	91.0/82.4	93.5/97.8	94.8/97.8	92.9/93.4	92.8/95.2	
PL-20	87.5/90.9	92.3/96.5	87.8/78.4	92.0/91.1	90.0/94.7	90.8/91.3	91.5/96.5	III
	ORF7 dCP nt/aa	**ORF8 p21 nt/aa**	**ORF9 p19.6 nt/aa**	**ORF10 p19.7 nt/aa**	**ORF11 p7 nt/aa**	**ORF12 p4 nt/aa**	**3'UTR nt**	
Cl-766	99.2/98.1	98.9/98.4	99.4/99.4	98.7/97.2	98.2/97.2	93.4/91.7	97.5	I
621	99.1/98.3	99.5/100	99.6/99.4	98.7/96.6	97.3/94.4	93.4/91.7	96.8	
GP18	92.1/89.9	93.5/97.3	91.6/88.7	90.6/86.0	91.0/88.9	97.3/96.7	97.5	II
623	92.3/90.6	93.7/97.3	91.6/88.7	89.8/84.9	91.9/88.9	97.8/95.0	97.1	
PL-20	88.3/89.3	90.1/94.6	90.8/89.3	83.0/79.9	77.5/66.7	84.2/78.3	94.9	III

The 5'UTR sequences of isolates 621, 623 and PL-20 were successfully sequenced after RLM-RACE. The full genome length of isolates 621 and 623 was 18498 nt with a 5'UTR length of 737 nt, identical to that reported for the South African GP18 isolate (Maree et al., 2008). Multiple sequence alignment of the 5'UTR (Figure 3) indicated significant variation amongst the three variant groups with two regions differing substantially amongst them. In the first region, two insertions were observed in the sequence of the group I and III variants (isolates 621 and PL-20) compared to the group II variants, in which a 61 nt insertion between positions 205 and 267 and an insertion of 4 nt between positions 273 and 278 was found. In the second region, deletions between positions 567 and 661 in variant group I (isolate 621) of 65 nt, and in variant group III (isolate PL-20) of 132 nt were found. The sequence data of the 5'UTR was confirmed with clones generated from the RLM-RACE reaction as well as clones from primer set GL3.5 and LR 365Rev that amplified a product of 941 bp from positions 5-946 on the GP18 genome.

The variation in the 5'UTR sequences presented here correlated with the three variant groups shown by SSCP profiles generated from amplicons of ORF5, and sequence data of the 5'UTRs of the three variants (621, 623 and PL-20) reflected the three variant groups. The 5'UTR is highly variable between the GLRaV-3 molecular variants compared to the rest of

the genome. Intergroup variation between the group I (621) and group II (623) variants in the 5'UTR was as much as 30%, while variant group II (623) and group III (PL-20) differed by 22%. Group I (621) and III (PL-20) varied by 33%.

3.4.3 Confirmation of 5'UTR sequence results

Amplified RT-PCR products obtained from the twelve selected plants, using primer pair GL3.5F and GL3.342R, were of the expected sizes for the different variant groups, *i.e.* 422 bp for isolates of variant group I and III, and a smaller product, 362 bp, for isolates of group II variants. These PCR products were sequenced and nucleotide sequences of 382 nt and 317 nt (excluding primer sequences) were compiled. The multiple sequence alignment of the twelve plants showed the insertion of 65 nts for group I and III isolates (plants 4 and 20).

In addition to these plants, 57 GLRaV-3 infected plants, collected randomly from 10 mother blocks in different grapevine growing regions of the Western Cape, were analysed in the same area of the 5'UTR (not shown). Fifty-two of these plants grouped in the group II variant clade with GP18 and 623, four of these plants grouped with the group I variant, 621, and one plant in the group III clade with PL-20.

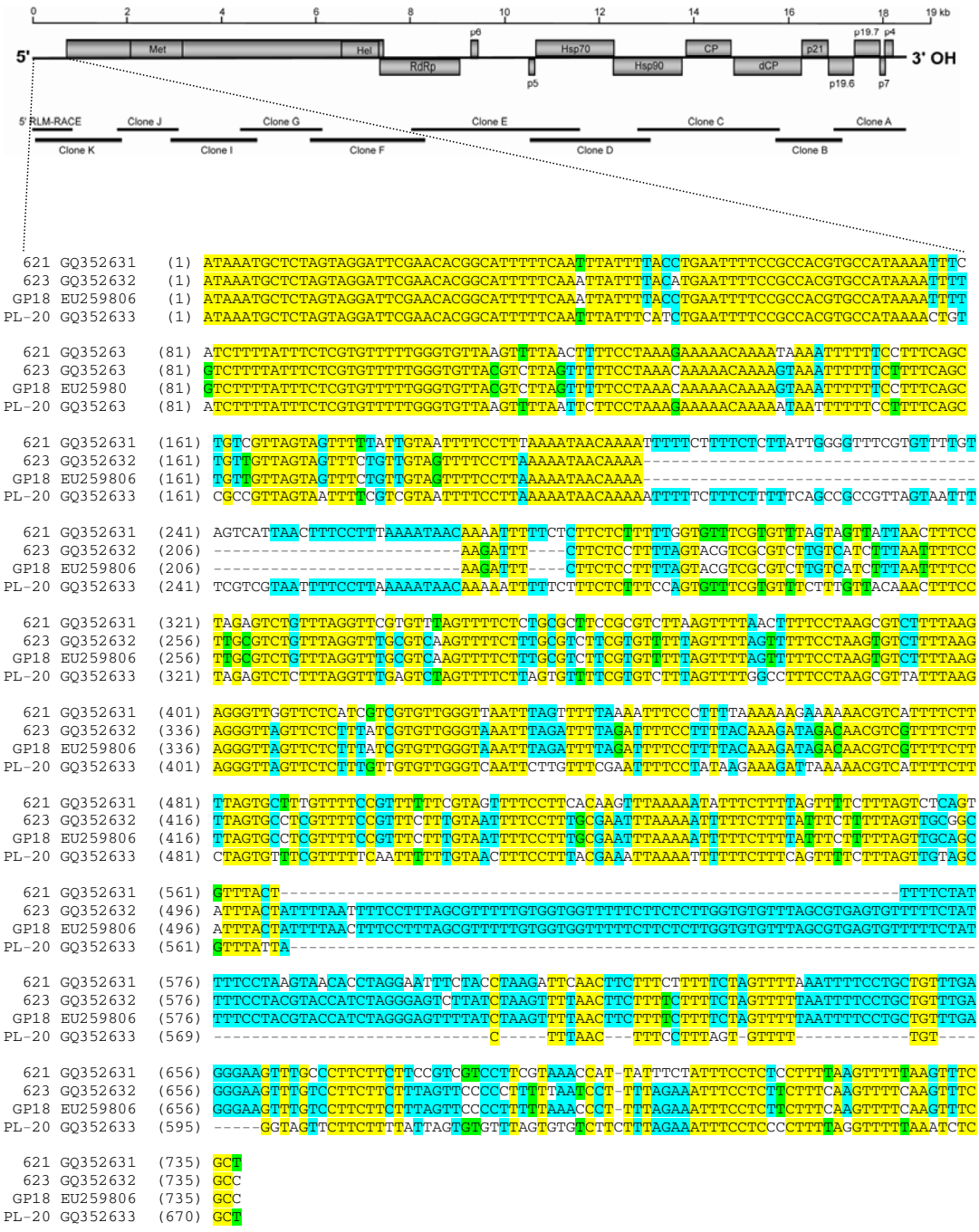


Figure 3.A schematic representation of the genome organisation of GLRaV-3 and the position of the ten overlapping clones of isolate PL-20 illustrating the sequencing strategy (top). The multiple alignments of the complete 5'UTR sequences of isolates 621, 623, PL-20 and GP18 is shown below this. This reveals a first deletion in isolates 623 and GP18 (group II) of 61 and 4 nt, a second deletion of 65 nt in isolate 621 and PL-20 (group I and III) and a further deletion of 51 nt in the isolate PL-20 sequence (group III).

3.4.4 Phylogenetic analysis of the Hsp70h and CP gene sequences

The alignment of the Hsp70h gene sequences of the isolates 621, 623 and PL-20 with the sequences from Genbank® resulted in a matrix which was 933 bp in length. The phylogenetic analysis revealed that 213 of the 933 characters were constant, 545 (58%) characters were found to be parsimony uninformative while 175 (19%) characters were parsimony informative. The heuristic search retrieved 8960 trees with a tree length of 988, with a consistency index (CI) of 0.906 and a retention index (RI) of 0.786. As many of the Genbank® depositions of CP sequences were not complete, the alignment that was used in the phylogenetic analysis was restricted to the first 499 bp of the CP gene. The analysis revealed that 174 of the 499 characters were constant, 228 (45.7%) characters were found to be parsimony uninformative while 97 (19.4%) characters were parsimony informative. The heuristic search retrieved 16 trees with a tree length of 458, with a CI of 0.856 and an RI of 0.876. The trees that were retrieved were used to generate a strict consensus tree. One of the shortest trees retrieved for both regions is shown in Figure 4. This representation of the phylogenetic analysis was chosen as it shows branch lengths, which indicate the actual numbers of differences between the sequences that were included in the analysis. Those nodes that collapsed in the strict consensus tree are indicated with arrows. The phylogeny confirms that, based on strong bootstrap support of key nodes, three distinct variants of GLRaV-3 occur in South Africa.

The phylogenetic position of the three South African variants in the Hsp70h region (Figure 4A), in relation to GLRaV-3 from other geographic regions, showed that the group I variant, isolate 621, grouped with the NY-1 (Ling et al., 2004) and Cl-766 (Engel et al., 2008) isolates with very little variation between these isolates. The variant I clade included sequences from isolates from the USA (NY-1, N1-1, C6-1, C4-1, C3-1, C2-1, C1-1, USA6), China (C3), Chile (Cl-766), Israel (IL1), Italy (MT48-4, MT48-1), Syria (SY2-2, SY2-7), and Austria (AUSG5-5). The group II variant, isolate 623, was sister to isolate GP18 (Maree et al, 2008). Other accessions in the group II clade were from Austria (AUSG5-2, AUSG5-6), Syria (SY2-4) and Tunisia (TU32). Branch lengths were longer in this clade indicating greater sequence heterogeneity, and even the two South African isolates showed some sequence divergence. The South African variant group III, as represented by isolate PL-20, grouped with two accessions from Italy (MT48-2, MT48-3). The branch lengths in this clade indicated that there was very little variation between these isolates. Two isolates, C5-1 and NZ-1, each resolved in isolated positions on their own in the phylogeny. They appear to represent two further

variants of GLRaV-3 in addition to the three variants present in South Africa and elsewhere in the world.

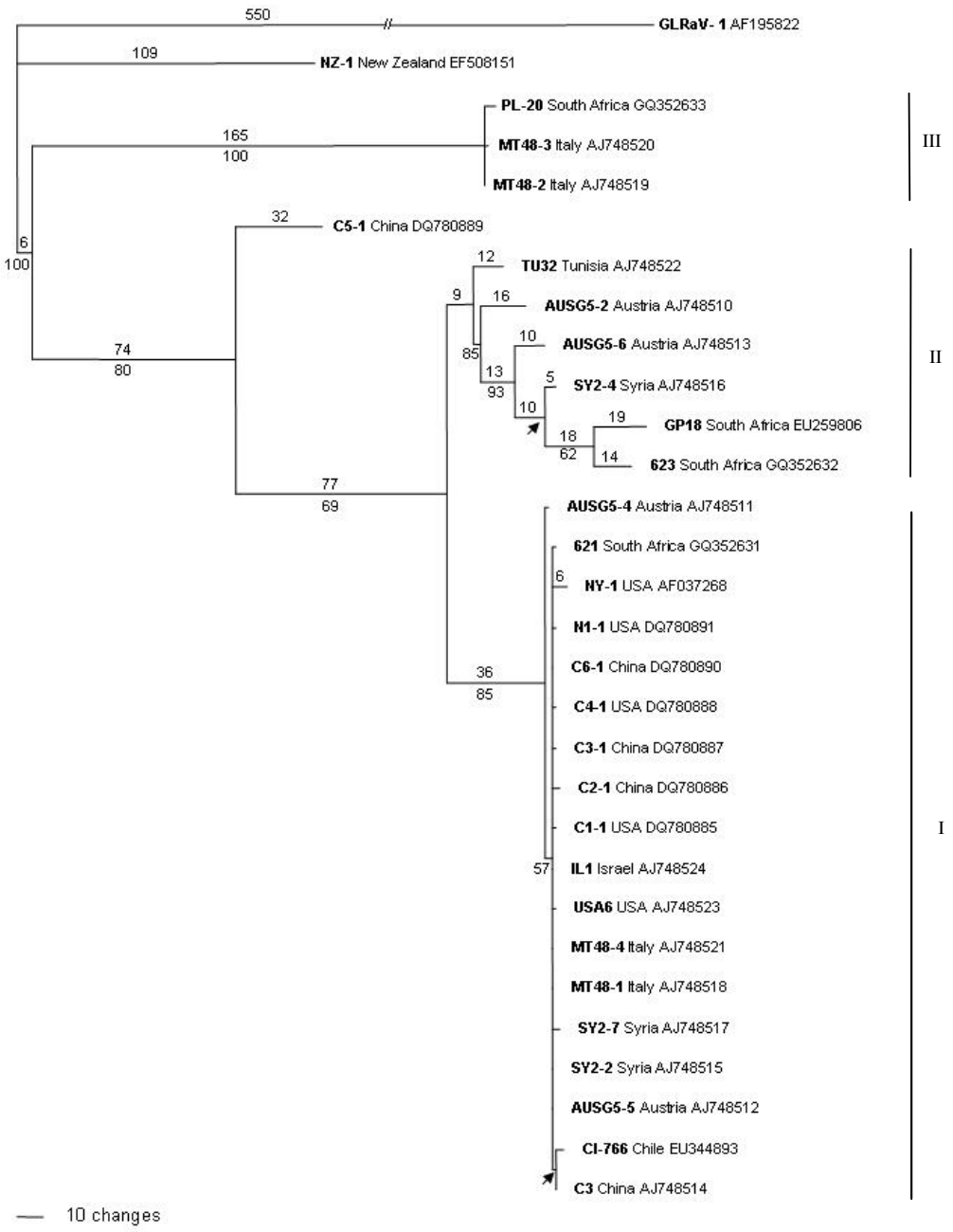


Figure 4A. One of the shortest trees of a heuristic search performed on the aligned Hsp70h sequence matrix. Branch lengths that are longer than 5 are shown above branches and bootstrap values are indicated beneath. Branches that collapsed in the strict consensus tree are shown with arrows. The branch length of the outgroup is not drawn according to scale. The three variant groups studied here are indicated as groups I, II and III.

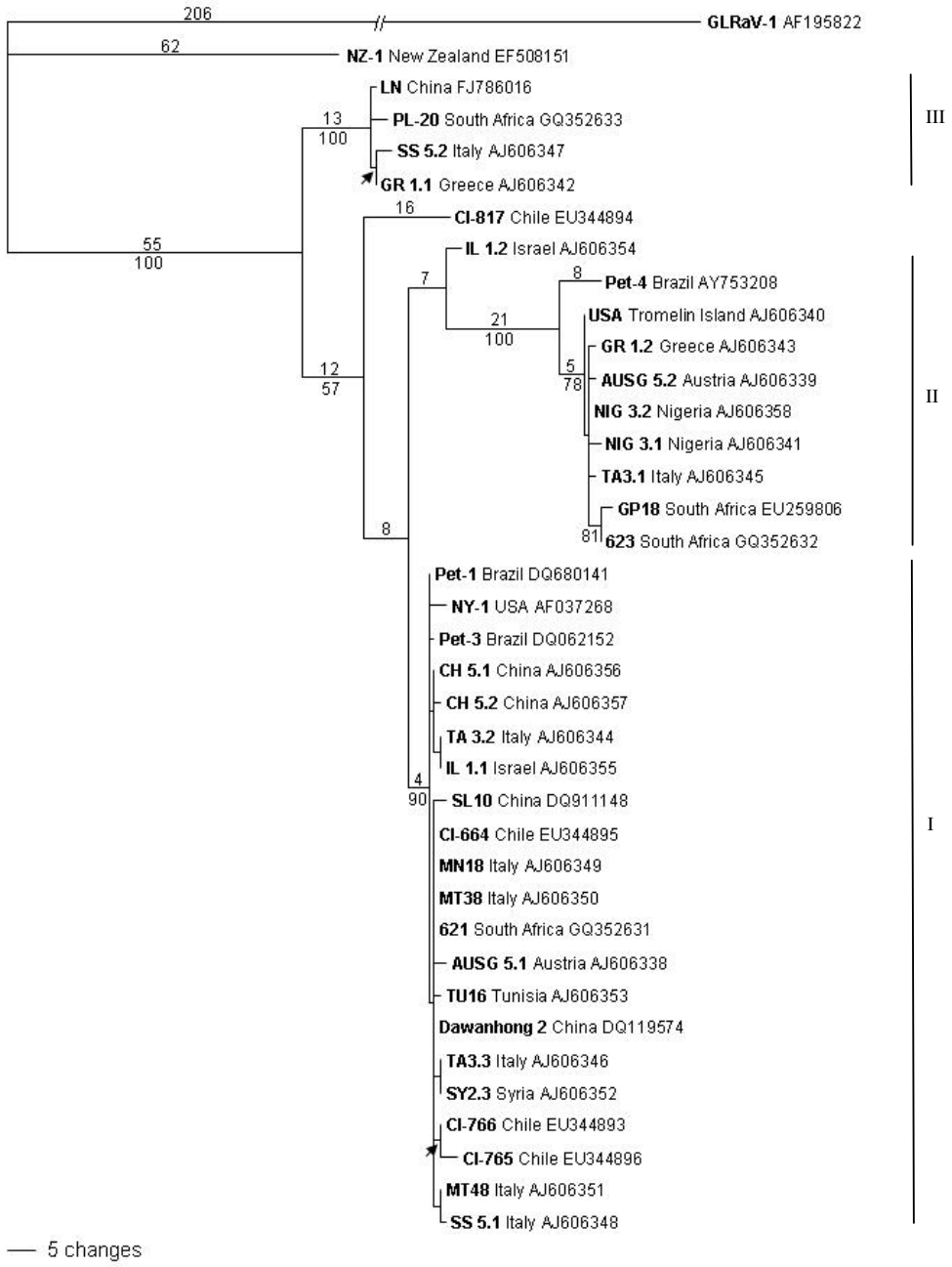


Figure 4B. One of the shortest trees of a heuristic search performed on the aligned coat protein sequence matrix. Branch lengths that are longer than 5 are shown above branches and bootstrap values are indicated beneath. Branches that collapsed in the strict consensus tree are shown with arrows. The branch length of the outgroup is not drawn according to scale. The three variant groups studied here are indicated as groups I, II and III.

The phylogenetic analysis of the CP gene region (Figure 4B) showed that the three South African isolates also grouped into three well-supported monophyletic clades, as in the Hsp70h analysis. In the CP analysis, the majority of isolates grouped in the variant group I clade, together with isolate 621 and NY-1. This group included isolates from Brazil (Pet-1, Pet-3), China (SL10, CH 5.1, CH 5.2, Dawanhong 2), Italy (TA 3.2, TA 3.3, MN18, MT38, MT48, SS 5.1), Israel (IL1.1), Chile (Cl-664, Cl-766, Cl-765), Austria (AUSG 5.1), Tunisia (TU16), and Syria (SY 2.3). In the group II clade, isolates 623 and GP18 grouped together with isolates from Italy (TA 3.1), Nigeria (NIG 3.1, NIG 3.2), Austria (AUSG 5.2), Greece (GR 1.2), Tromelin Island (USA) and Brazil (Pet-4). Isolates that grouped with the third variant, isolate PL-20, were from China (LN), Italy (SS 5.2) and Greece (GR 1.1). The New Zealand isolate NZ-1, the Chilean isolate Cl-817 and the Israeli IL 1.2 isolate resolved in isolated positions but with poor bootstrap support.

3.5 DISCUSSION

As in another study (Turturo et al., 2005), the SSCP technique proved a useful tool for studying the genetic diversity of South African GLRaV-3 isolates. In our study, the results showed that SSCP analysis of a region in ORF5 gave a fast and reliable indication of the GLRaV-3 variant status of a plant. This region in ORF5 was successfully amplified in all GLRaV-3-infected plants (>200 plants) collected during two local surveys in red grapevine cultivars (data not shown). Based on this gene region, three main GLRaV-3 variant groups were identified in South African vineyards. The three variant groups were confirmed by sequencing and distinct SSCP profiles could be assigned to each variant group. SSCP and sequencing results showed complete correlation in all of the genomic regions studied. Additional SSCP analysis and sequence data from other regions of the genomes confirmed the same three variant groups. The SSCP profiles detected for group II variants were heterogeneous, indicating greater sequence complexity within this group. In some of the analysed plants, “complex” SSCP profiles were detected which, in most cases, contained a single dominant GLRaV-3 variant. Plants infected with a mixture of variants were also detected. This information is in agreement with a previous study (Turturo et al., 2005) in which combinations of variants in the same isolate were also detected.

The variation detected by SSCP was confirmed by sequencing the whole genomes of three isolates representative of the three variant groups. The three GLRaV-3 variant groups correlated with previous reports that more than one group occurred in South Africa. Prosser *et*

al. (2007) reported that the WC-HSP-2 isolate- from South Africa, was 93.2% identical to the NY-1 isolate and isolates WC-HSP-10 and WC-HSP-28 were only 72.3% identical to NY-1. Similarly, Maree *et al.* (2008) reported that isolate GP18 was 93% similar to NY-1. Jooste & Goszczynski (2005) found that the percentage identity between isolates 621 and 623 was 91.8-96.2% in ORF4-7 and that the most and least divergent fragments were of ORF5 and ORF7 (91.8-92.3% and 96.0-96.2%).

The phylogenetic analysis of the aligned sequences of the Hsp70h gene region of these South African isolates and those of isolates from elsewhere in the world, further confirmed that three variant groups of GLRaV-3 occurred in South Africa and that these variants also occur elsewhere in the world. The two further isolates, one from New Zealand (NZ-1) and one from China (C5-1) appear to represent two further variants. Five GLRaV-3 variant groups were also identified by Fuchs *et al.* (2009) in their phylogenetic analysis of partial Hsp70h sequences available in GenBank. The five variant groups in the Fuchs study were represented by isolates NY-1, GP18 and MT48-2, found in our groups I, II and III respectively, and also C5-1 and NZ-1. The two isolates, C5-1 and NZ-1, each also resolved in isolated positions on their own in the phylogenetic analysis in the Fuchs study. The phylogenetic analysis of the CP largely corroborated the identification of the three variant groups present in South Africa, in spite of the fact that most of the CP sequences from elsewhere in the world, with the exception of NY-1, Cl-766 and GP18, were not from the same source as the Hsp70h gene sequences. Again the NZ-1 isolate grouped in an isolated position, confirming the unique identity of a further variant in New Zealand, but the isolated position of two further isolates, Cl-817 from Chile, and IL 1.2 from Israel may prove to present two further variant groups of GLRaV-3. Thus, the phylogenetic analysis of the CP region indicates six possible variants instead of the five indicated by the phylogenetic analysis of the Hsp70h gene. Further whole-genome sequencing of viral isolates from other regions of the world will have to be undertaken to establish whether, besides the three isolates that are represented in South Africa, two or three further GLRaV-3 variants occur.

The high divergence detected in the 5'UTR between isolates 621 and 623 supports the previous report where a 417 nt fragment, which included 88 nt of the 5'UTR and adjacent 329 nt of ORF1a, only revealed a 81.8% identity between isolates 621 and 623 (Jooste & Goszczynski, 2005). In the present study, the sequence of the 5'UTR of GLRaV-3 showed three clear variant groups of GLRaV-3 that correlated with the identification of the three variant groups by SSCP profiles, whole-genome sequencing and phylogenetic analysis of

South African isolates. The third variant, represented by PL-20, is especially interesting as it contains a shorter 5'UTR resulting in a genome of 18433 nt, 65 nt shorter than the sequences of isolates 621 (variant group I), 623 (variant group II) and GP18 (variant group II). A similar study on variability of *Citrus tristeza virus* (CTV) isolates showed that the 5'UTR of CTV could be used for the classification of sequences into three groups (López et al., 1998).

RT-PCR of all South African field samples tested thus far confirmed the presence of an extended 5'UTR for GLRaV-3, as described previously (Maree et al., 2008). Sequence results of the 5'UTR from 69 of these field collected isolates clearly indicated that variability between isolates of the same variant group is low, but higher between isolates from different variant groups.

The possibility of potential folding and the significance of the 5'UTR were investigated in this study. Bioinformatic analysis of the 5'UTRs, *i.e.* conserved functional motifs, ORF analyses and predicted secondary structure of these three molecular variants were unable to predict a possible function for these 5' UTRs.

Three genetic variants of GLRaV-3 represented by isolates 621, 623 and PL-20, were identified from South African vineyards. We conclude that the 5'UTR of the GLRaV-3 genome is a key region that can be used to study the variation amongst variants of the virus. The full-length sequences of these three isolates and partial 5'UTR sequence data from field isolates reaffirm the existence of an extended 5'UTR (Maree et al., 2008). Currently these four sequences represent the only complete full-length sequences of GLRaV-3. Sequencing of the 5' UTRs of other GLRaV-3 variants occurring elsewhere in the world may therefore be of particular value in future studies of the variation in GLRaV-3 and warrants further investigation.

3.6 LITERATURE

Angelini E, Bertazzon N, Bazzo I, Borgo M (2006)Molecular characterization of a divergent strain of grapevine leafroll-associated virus 3. In: Extended abstracts 15th Meeting of ICVG, Stellenbosch, South Africa, 3-7 April 2006, pp 148-150

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

Chooi KM, Pearson MN, Cohen D, Pong JCH (2009) Sequence variation in *Grapevine leafroll-associated virus-3* (GLRaV-3) New Zealand isolates. In: Extended abstracts of the 16th Meeting of ICVG, Dijon, France, 31 Aug-4 Sept 2009, pp 290-291

Engel EA, Girardi C, Escobar PF, Arredondo V, Dominguez C, Perez-Acle T, Valenzuela PDT (2008)Genome analysis and detection of a Chilean isolate of *Grapevine leafroll associated virus-3*. Virus Genes 37:110-118

Fuchs M, Martinson TE, Loeb GM, Hoch HC (2009)Survey for the three major leafroll disease-associated viruses in Finger Lakes vineyards in New York. Plant Dis93:395-401

Giegerich R, Meyer F, Sleiermacher C (1996) GeneFisher - software support for the detection of postulated genes. Proc Int Conf Intell Syst Mol Biol. 4:68-77

Goszczynski DE, Jooste AEC (2002) The application of single-strand conformation polymorphism (SSCP) technique for the analysis of molecular heterogeneity of grapevine virus A. Vitis 41:77-82

Gouveia P, Esteves F, Teixeira Santos M, Fonseca F, Eiras-Dias JE, Nolasco G (2009)Assessment of GLRaV-3 variants occuring in Portuguese grapevine varieties according to the coat protein gene. In: Extended abstracts of the 16th Meeting of ICVG, Dijon, France, 31 Aug-4 Sept 2009, pp 316-317

Hall TA (1999) BioEdit: a user-friendly biological sequence and genomic alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95-98

Jooste AEC, Goszczynski DE (2005) Single-strand conformation polymorphism (SSCP), cloning and sequencing reveal two major groups of divergent molecular variants of grapevine leafroll-associated virus 3 (GLRaV-3). Vitis 44(1):39-43

Ling KS, Zhu HY, Drong RF, Slightom JL, McFerson JR, 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 KS, Zhu HY, 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

López C, Ayllón MA, Navas-Castillo J, Guerri J, Moreno P, Flores R (1998) Molecular variability of the 5'- and 3'-terminal regions of *Citrus tristeza virus* RNA. Phytopathology88:685-691

Marchler-Bauer A (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res37:205-210

Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res32:327-331

Maree HJ, Freeborough M-J, Burger JT (2008) Complete nucleotide sequence of a South African isolate of *grapevine leafroll-associated virus* 3 reveals a 5'UTR of 737 nucleotides.Arch Virol153:755-757

Martelli GP, Agranovsky AA, Bar-Joseph M, Boscia D, Candresse T, Coutts RHA, Dolja VV, Falk BW, Gonsalves D, et al. (2002) The family *Closteroviridae* revised. Arch. Virol 147/10

Pietersen G (2004) Spread of Grapevine leafroll disease in South Africa - a difficult, but not insurmountable problem.Wynboer –A technical guide for wine producers. June 2004. URL: http://www.wynboer.co.za/recentarticles/articles.php3.

Prosser SW, Goszczynski DE, Meng B (2007)Molecular analysis of double-stranded RNAs reveals complex infection of grapevines with multiple viruses. Virus Res124:151-159

Soule MJ, Eastwell KC, Naidu RA (2006) First report of *Grapevine leafroll associated virus-3* in American *Vitis* spp. Grapevines in Washington State. Plant Dis 90:1461

CHAPTER 4

Distribution of grapevine leafroll-associated virus 3 (GLRaV-3) variants in South African vineyards

Submitted to European Journal of Plant Pathology, accepted 15 November 2010

A.E.C. Jooste[1], G. Pietersen[1], J.T. Burger[2]

[1]ARC-Plant Protection Research Institute, Private Bag X134, Queenswood, Pretoria, 0121, South Africa

[2]Department of Genetics, Stellenbosch University, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa

4.1 ABSTRACT

Three genetic variants of grapevine leafroll-associated virus 3 (GLRaV-3) were identified from vineyards of the Western Cape, South Africa. In a previous study, three full-genome sequences of isolates representing each of the variant groups were determined. The three variant groups were represented by accessions 621, 623 and PL-20, of variant groups I, II and III respectively. A specific single strand conformation polymorphism (SSCP) profile was assigned to each variant which was used as a quick, reliable detection and differentiation method. In this study we analysed the occurrence of these three GLRaV-3 variants in mother blocks in different cultivars and from different vine growing regions using SSCP analysis. The majority of the plants studied were infected with the group II variant, similar to isolates 623 and GP18. The distribution of three GLRaV-3 variants within a spatio-temporally recorded cluster of diseased plants was studied by means of SSCP profile analysis of ORF5 amplified PCR products. We showed that different GLRaV-3 variants are transmitted to adjacent plants in an infection cluster. Results showed that, in some leafroll disease clusters, the variant that was present in the original GLRaV-3-infected plant of a cluster was transmitted to adjacent plants in a row and across rows.

4.2 INTRODUCTION

Grapevine leafroll (GLR) disease is one of the most important diseases of grapevines, occurring in all grape-producing countries worldwide, including South Africa (Pietersen 2004). Several phloem-limited filamentous viruses, identified as grapevine leafroll-associated viruses(GLRaVs), have been characterized from leafroll infected grapevines (Fuchs et al., 2009a). These viruses include species from the genera *Closterovirus* (GLRaV-2) and *Ampelovirus* (GLRaV-1, GLRaV-3, GLRaV-4, GLRaV-5, GLRaV-6 and GLRaV-9) and GLRaV-7, not assigned to a genus yet (Fuchs et al. 2009a).

The disease delays ripening of grapevine berries, decreases the accumulation of sugar and ultimately influences the overall quality of the wines. GLR symptoms vary depending on environmental conditions as well as the cultivars concerned. Symptoms are best observed in the period between harvesting and shedding of leaves (late summer and autumn) and appear as typical downward rolling of leaves with leaf veins that stay green. In red wine cultivars the areas between leaf veins turn red, whereas yellowing of the same leaf areas occur in white wine cultivars (Carstens 2002). GLR disease is transmitted through infected propagation material as well as by mealybug and soft scale insect vectors (Belli et al 1994; Cabeleiro

&Segura 1997a; Cabeleiro & Segura 1997b; Douglas & Krüger 2008; Petersen & Charles 1997; Sforza et al. 2003; Tsai et al 2008).

GLRaV-3 is known to be closely associated with leafroll disease and occurs commonly in South African vineyards (Pietersen 2004; Pietersen 2006). The mealybug *Planococcus ficus* is considered the most important vector of GLRaV-3 in South Africa (Douglas & Krüger 2008). Transmission efficiency studies with *P. ficus* and *P. longispinus* showed that the two mealybug species are both efficient vectors for GLRaV-3 in South African vineyards (Douglas & Krüger 2008).

The first report of natural spread of GLR disease in South Africa was recorded in 1985 (Engelbrecht & Kasdorf 1985). The natural spread of this disease in a vineyard was proven a few years later when 100 healthy LN33 indicators were planted in a leafroll infected vineyard of cv. Tinta Barocca (Engelbrecht & Kasdorf 1990). The first symptoms on the indicator plants appeared 2-3 years after planting and after 7 years 71% of the plants displayed symptoms. GLRaV-3 was detected in all symptomatic plants.

Recently, several epidemiology studies on leafroll disease have been reported from grapevine growing regions world-wide. The spatial distribution of GLRaV-3 was studied in vineyards from Spain (Cabeleiro & Segura 2006; Cabeleiro et al. 2008) and the field spread of GLRaV-3 was monitored in these vineyards since 1991 (Cabeleiro et al. 2008). From this study it was clear that there was a correlation between mealybug incidence and virus spread (Cabeleiro et al. 2008). Scale insects were implied as vectors of GLRaV-3 in the Meaño vineyard where slow, but constant spread of the virus was observed (Cabeleiro et al. 2008). In two vineyards, in Portomarín and Goian, in the same study, the virus inoculum originated from infected plant material resulting in a random distribution of the disease. A study of the spread of leafroll disease in a Napa Valley vineyard in California showed that spread of the disease came from neighbouring blocks, heavily infected with leafroll, and mapping results of the disease showed an increase in spread of more than 10% per year in this block (Golino et al. 2008). The possible causes for the sudden rapid spread of leafroll in vineyards of California were debated and the authors suggested that something fundamental changed in the vineyards, such as vector epidemiology, grower rootstock preferences and/or new leafroll strains that emerged (Golino et al. 2008). A study of vineyards in the Pacific Northwest (PNW) of the U.S.A. documented the presence of several genetic variants of GLRaV-1, GLRaV-2, *Rupestris stempitting-associated virus* and *Grapevine fanleaf virus* in these vineyards (Rayapati et al.

2009). The identification of variants of the viruses is therefore important to understand the spread of a complex disease like leafroll.

Three genetic variants of grapevine leafroll associated virus 3(GLRaV-3) were identified in vineyards of the Western Cape, South Africa (Jooste & Goszczynski 2005; Jooste et al 2010). The GLRaV-3 variants were identified by single strand conformation polymorphism (SSCP) profiles generated from a region amplified in ORF5. SSCP results and sequence data confirmed the three variant groups and a specific SSCP profile was assigned to each variant group. Results showed that SSCP analysis on the region in ORF5 gives a fast and reliable indication of GLRaV-3 variant status in a plant. In many plants, combinations of variants were detected. The full genome sequence of a representative from each variant group, 621 (GQ352631, group I), 623 (GQ352632, group II) and PL-20 (GQ352633, group III), was determined (Jooste et al. 2010). The most variation between the three variants occurred in their 5'UTR.

The spatial distribution and spatial dynamics (changes in distribution patterns) of GLR disease within mother blocks of the South African Certification Scheme were studied intensively from 2001-2007 (Pietersen 2004; Pietersen 2006). Four common distribution patterns of GLR were observed in this study. The most significant distribution pattern identified in local vineyards was secondary spread within vineyards after establishment (Pietersen 2006). The other described means of GLR spread are primary spread by infected plant material, GLR spread from a preceding vineyard and gradients of GLR infected vines associated with proximal leafroll infected vineyards (Pietersen 2006).

No information is currently available on the spread of specific GLRaV-3 variants in vineyards. The recent data on molecular characteristics of three genetic variants of GLRaV-3 together with the occurrence of GLR-infected plants within mother blocks of the certification scheme led to the objectives of this study: firstly, to determine which variant occurs predominantly in mother blocks and secondly, to investigate disease clusters and the spread of individual GLRaV-3 variants within such a disease cluster.

4.3 MATERIALS AND METHODS

4.3.1 Field survey to determine the occurrence of GLRaV-3 variants in mother blocks

Spatio-temporal distribution patterns of leafroll-infected plants for the period 2001 to 2005 were observed from mother blocks from different regions, including the Stellenbosch, Paarl, Wellington, Worcester, and Somerset West grape production areas (Pietersen unpublished results).

Mother blocks are a category of propagation vineyard within the South African Wine Grape Certification Scheme, and generally refer to the second generation of vines derived from nuclear plants. Nuclear plants are plants from which virus was eliminated through heat therapy and meristem tip culture and tested for the presence of any viruses by indexing on indicator vine plants, more specifically for GLRaV-1, -2, and -3 by ELISA and for all leafroll-associated viruses and grapevine virus A by immuno-electron microscopy. These plants are maintained in vector-free gauze houses. Mother blocks represent the second propagation generation planted outside the gauze houses, the first generation being foundation blocks. These vineyards may only be established from a foundation block or from approved other mother block planting material. Mother blocks are inspected for leafroll disease annually and infected vines either rouged or marked with paint and the canes pruned before planting material is collected from the rest of the mother block. Planting material is no longer collected from mother blocks once leafroll infection levels exceed 3%, whereupon these vineyards loose their mother block status. All mother blocks referred to in these studies were at incidences of leafroll below 3% at the initiation of them being monitored for the spatial spread of leafroll.

The relative position of vines showing leafroll symptoms were recorded and plotted in an XY matrix using the row number and vine position as co-ordinates (Pietersen 2004). Leafroll-infected plants were recorded in vineyards on a yearly basis, based on symptom expression. In numerous disease foci (also referred to as disease clusters), the infection point or starting point of disease spread appeared to begin from a single plant, from where mealybugs transmitted the disease to adjacent plants in rows and across rows forming clusters of GLR infection. Plants were recorded to be positive for GLR when the typical symptoms were first visible. Assuming that the actual infection of a plant may have preceded the GLR symptom expression by a constant number of seasons, the appearance of symptoms does reflect the timing of infection.

Eighty one plants were collected from 10 mother blocks (1, 4, 9, 17, 50, 54, 64, 65, M and V) in different grapevine growing regions (Table 1). Eight plants were randomly selected per block, except for block 4 where nine plants were collected and block 65 from which seven plants were collected. The plants were chosen based on the vine position and leafroll distribution data collected previously (Pietersen 2004). Additionally, nine plants (V1-9) were collected randomly from five mother blocks (72, 73, 77, 84, and 108) on the Vergelegen Wine Estate. On this farm, a virus control strategy was implemented (Pietersen, 2010) and the nine plants were isolated sources with GLR infection. Grapevine plants displaying leafroll symptoms, spatially distant from other infected plants, as well as plants that were part of heavily-infected disease clusters were selected. Cultivars included Cabernet Sauvignon, Merlot, Shiraz, Petit Verdot, red cultivars, and Palomino, a white cultivar, grafted on different rootstocks (Table 1).

Table 1. Grapevine plants collected from ten mother blocks from different geographical regions and their GLRaV-3 variant status

Plant no.	Plant position Block/Row/Plant	Cultivar x Rootstock	Year planted	Region	Variant group
1.1	1/4/62	Cabernet Sauvignon 1CxRichter99 179	1991	Stellenbosch	II
1.2	1/10/17	Cabernet Sauvignon 1CxRichter99 179	1991	Stellenbosch	I
1.3	1/15/37	Cabernet Sauvignon 1CxRichter99 179	1991	Stellenbosch	I
1.4	1/18/2	Cabernet Sauvignon 1CxRichter99 13A	1991	Stellenbosch	II
1.5	1/18/49	Cabernet Sauvignon 1CxRichter99 13A	1991	Stellenbosch	II
1.6	1/22/40	Cabernet Sauvignon 1CxRichter99 13A	1991	Stellenbosch	I+II
1.7	1/30/63	Cabernet Sauvignon 1CxRichter99 13A	1991	Stellenbosch	II
1.8	1/30/113	Cabernet Sauvignon 1CxRichter99 13A	1991	Stellenbosch	II
4.1	4/1/16	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	II
4.2	4/3/122	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	II
4.3	4/4/125	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	I+II+III
4.4	4/7/70	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	I+II+III
4.5	4/8/18	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	I+II+III
4.6	4/9/5	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	II
4.7	4/6/32	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	II
4.8	4/9/252	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	I+II+III
4.9	4/10/124	Cabernet Sauvignon 1Cx101-14 219A	1997	Somerset West	II
9.1	9/2/75	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/Stellenbosch	II
9.2	9/3/6	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	unknown
9.3	9/3/36	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	I
9.4	9/8/64	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	I+II+III
9.5	9/11/30	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	I
9.6	9/11/99	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	nt
9.7	9/12/85	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	II
9.8	9/13/21	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl/ Stellenbosch	I
17.1	17/2/6	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	II+III
17.2	17/3/44	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	II+III

17.3	17/4/120	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	I
17.4	17/12/34	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	II
17.5	17/15/58	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	II
17.6	17/18/82	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	I+II+III
17.7	17/28/27	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	I+II+III
17.8	17/31/4	Cabernet Sauvignon 163IxRichter99 179	1997	Paarl	I+II+III
50.1	50/12/22	Merlot 12x101-14 28A	1992	Paarl	I+II+III
50.2	50/12/64	Merlot 12x101-14 28A	1992	Paarl	II+III
50.3	50/17/21	Merlot 12x101-14 28A	1992	Paarl	II
50.4	50/23/58	Merlot 12x101-14 28A	1992	Paarl	II
50.5	50/30/9	Merlot 12x101-14 28A	1992	Paarl	II
50.6	50/32/24	Merlot 12x101-14 28A	1992	Paarl	II
50.7	50/33/56	Merlot 12x101-14 28A	1992	Paarl	I
50.8	50/34/62	Merlot 12x101-14 28A	1992	Paarl	II
54.1	54/6/54	Cabernet Sauvignon 1CxRichter99 179	1997	Franshoek/Paarl	I+II
54.2	54/12/45	Cabernet Sauvignon 1CxRichter99 179	1997	Franshoek/Paarl	II
54.3	54/14/19	Cabernet Sauvignon 1CxRichter99 179	1997	Franshoek/Paarl	II
54.4	54/16/93	Cabernet Sauvignon 1CxRichter99 179	1997	Franshoek/Paarl	II
54.5	54/23/39	Cabernet Sauvignon 1CxRichter99 179	1997	Franshoek/Paarl	II
54.6	54/52/41	Shiraz 22BxRichter99 13	1998	Franshoek/Paarl	II
54.7	54/65/40	Shiraz 22BxRichter99 30B	1997	Franshoek/Paarl	II
54.8	54/66/4	Shiraz 22BxRichter99 30B	1997	Franshoek/Paarl	II
64.1	64/2/77	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	II
64.2	64/8/82	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	II
64.3	64/9/55	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	II
64.4	64/13/12	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	II
64.5	64/23/11	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	II
64.6	64/37/20	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	nt
64.7	64/41/18	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	I
64.8	64/59/22	Petit Verdot 400Dx101-14 219F	2000	Stellenbosch/Belville	unknown
65.1	65/16/46	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	nt
65.2	65/20/19	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	I
65.3	65/38/116	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	I+II
65.4	65/39/18	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	I+II
65.5	65/40/10	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	II
65.6	65/44/62	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	I
65.7	65/50/44	Cabernet Sauvignon 341Bx101-14 219F	2000	Klapmuts/Franshoek	II
M1	38/1/9	Palomino	1980	Worcester/Villiersdorp	II
M2	38/3/4	Palomino	1980	Worcester/Villiersdorp	I+II
M3	38/4/12	Palomino	1980	Worcester/Villiersdorp	I
M4	38/4/55	Palomino	1980	Worcester/Villiersdorp	III
M5	38/16/1	Palomino	1980	Worcester/Villiersdorp	II
M6	38/19/7	Palomino	1980	Worcester/Villiersdorp	II
M7	38/31/23	Palomino	1980	Worcester/Villiersdorp	nt
M8	38/42/25	Palomino	1980	Worcester/Villiersdorp	I+II
V1	72/5/136	Shiraz SA99xRichter 110	1999	Somerset West	II
V2	72/16/84	Shiraz SA99xRichter 110	1999	Somerset West	II
V3	73/27/107	Cabernet Sauvignon 46Cx101-14	1999	Somerset West	II+III
V4	77/2/66	Cabernet Sauvignon 46Cx101-14 219A	2002	Somerset West	II
V5	77/16/32	Cabernet Sauvignon 46Cx101-14 219A	2002	Somerset West	II
V6	72/16/78	Shiraz SA99xRichter 110	1999	Somerset West	I+II+III
V7	84/6/44	Cabernet Sauvignon 46Cx101-14 219A	2002	Somerset West	I

V8	84/27/11	Cabernet Sauvignon 46Cx101-14 219A	2002	Somerset West	I+II+III
V9	108/37/19	Cabernet Sauvignon	1999	Somerset West	II

nt= not determined

4.3.2 Distribution of GLRaV-3 variants in infected disease clusters

The spread of three GLRaV-3 variants, previously identified from South African vineyards (Jooste et al.,2010), was investigated within three leafroll disease clusters. The disease clusters were studied in a Cabernet Sauvignon block in the Worcester region, Western Cape. The identification of GLRaV-3 variants in leafroll-infected plants was described in a previous study (Jooste et al 2010; Chapter 3 of this dissertation). In this study, SSCP results of plants 16, 17 and 20 showed distinct SSCP profiles that indicated the specific variant status of the plants (Jooste et al 2010). The SSCP profiles were correlated with sequence data. SSCP profiles of plant 16 showed a mixed infection with group II (623) and group III (PL-20) variants, the dominant variant in plant 17 was the group II (623) variant and plant 20 (PL-20) was infected with the group III variant. These three plants represented the initial infected plant of the particular cluster and were recorded to be leafroll-infected in 2001 based on symptom expression. The first symptom expression in a plant and the position of the infected plant was used to reflect the relative time of infection of a plant. We assumed that actual infection of a plant may have preceded the symptom expression of leafroll in a plant by a number of seasons. The history of symptom development of infected plants in the infection cluster, which developed around the initial infected leafroll plants, was recorded in consecutive years. Ten plants directly adjacent to each of plants 16, 17 and 20 (A-J) were collected in 2008, as seen in Figure 1, X indicating the originally-infected plant.

Figure 1. A diagram of the positions (A-J) of plants collected in a leafroll infected disease cluster relative to the original GLRaV-3 infected plant (X).

4.3.3 RT-PCR and SSCP analysis

Isolation of dsRNA, SSCP analysis and cloning were carried out as described earlier (Goszczynski & Jooste 2002). Double-stranded RNA was isolated from the plants in both experiments and SSCP analysis was performed on PCR-amplified products from a region in ORF5 with primer set H420 and C629 as described in Jooste et al. (2010). PCR products of the expected size were purified directly from low melting agarose gels using a Wizard PCR Prep DNA Purification System (Promega). SSCP profiles were generated in 15% polyacrylamide gels, run for two hours at 200V and stained with ethidium bromide.

4.4 RESULTS

4.4.1 Incidence of GLRaV-3 variants in mother blocks

GLRaV-3-specific RT-PCR amplicons of 209nt from ORF5 were obtained from all 81 plants. SSCP profiles were generated, analysed and each unique profile, representing a specific variant group, reported as a percentage of the total number of SSCP profiles analysed. SSCP results of blocks 1, 4, 9, 38, 54, and 64 are shown in Figure 2. SSCP profiles representing pure sources (single infections) of the group I variant can be seen in plants 1.2, 1.3, 9.3, 9.5, 9.8, 64.7 and M3. Similarly, the profile for pure sources of the group II variant can be seen in the profiles of plants 1.1, 1.4, 1.5, 1.7, 1.8, 9.1, 9.7, 54.2, 54.3, 54.4, 54.5, 54.6, 54.7, 54.8, 64.1, 64.2, 64.3, 64.4, 64.5, M1, M5, M6, 4.1, 4.2, 4.6, 4.7, and 4.9. Profiles of plants infected with combinations of variants can be seen in plants 1.6 and 54.1, which are infected with variant groups 1 and II, while plants 4.3, 4.4, 4.5 and 4.8, showed ‘complex’ profiles, suggesting that they represent mixed infections of group I, II and III variants. Results showing the respective variant groups detected in every plant are summarized in Table 1. Single infection with the group II variant was detected in 54% of the plants followed by the group I variant infecting 16.2% of the plants studied. The group III variant was detected in only one plant as a single infection, representing 1.4%. Six plants (8.1%) were infected with a combination of group I and group II variants, and four plants (5.3%) were infected with variants from groups II and III simultaneously. Eleven plants or 14.9% of all plants were infected with a combination of all three variants. Two plants had unknown SSCP profiles (9.2 and 64.8) and in four plants it was not possible to generate SSCP profiles because of inadequate concentrations of the amplified products.

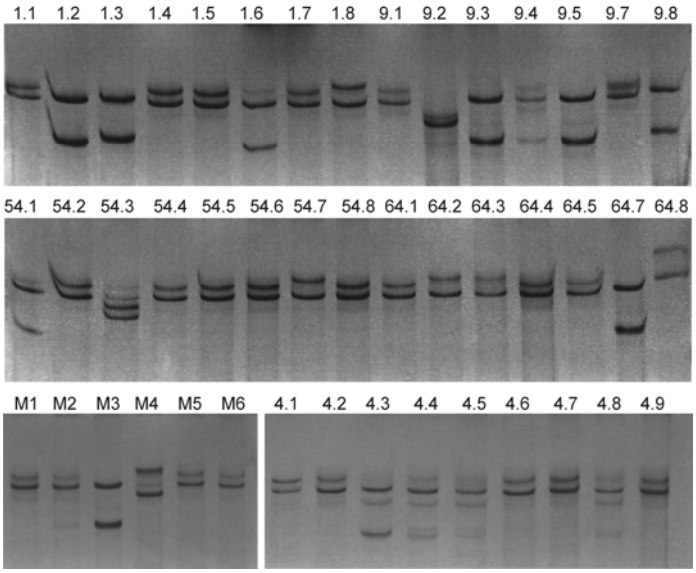


Figure 2. Examples of SSCP results of ORF5 from plants collected in six mother blocks (1, 9, 54, 64, M and 4). Pure sources of each variant group, as well as mixed variant infections were differentiated by unique SSCP profiles.

4.4.2 Distribution of GLRaV-3 variants in infected disease clusters

Three infection clusters, surrounding plants 16, 17 and 20, were collected from a vineyard in the Worcester region. The spread of variants from the putative originally-infected plant (the first one showing symptoms) to neighbouring plants were studied with SSCP analysis (Figure 3), in combination with the plotted data of leafroll infected plants from the survey of 2001 to 2005 (Figure 4). These originally infected plants in a disease cluster were sequenced in a previous study and ten clones of each plant were analysed by SSCP analyses and sequencing (Jooste et al.2010).

Plant 16, showing leafroll symptoms for the first time in 2001, was infected with variants from group II and III. The plants directly adjacent to it, in the same row, i.e. plants A and B, displayed leafroll symptoms three years later in 2004. The combination of these two variants was transmitted to the two plants along the row. Plants H, I and J, plants across the row, on the right side, showed leafroll symptoms in 2005 and both variants were detected in these plants. Plants E, F and G, in the row on the left of plant 16, showed the separation of group II and III variants. Plants E and G were infected with the group II variant and plant F with the group III variant.

A second infection cluster of plants, those surrounding plant 17, was analysed. Plant 17 was infected with the group II variant and was recorded as leafroll symptomatic in 2001. At that

time the adjacent plants did not show any leafroll symptoms. In the following season, 2002, two plants in the row adjacent to plant 17 showed leafroll symptoms and only the group II variant was detected in these plants. Three years after plant 17 was infected with leafroll, in 2004, the plants in the row on the left of the infection focus showed symptoms (plants H, I, and J). Plant H was infected with a combination of variants from groups I and II and plants I and J infected with variant groups III and II, respectively. Plants E, F, and G, on the left side of plant 17, showed symptoms a year later in 2005, and the group II variant was detected in these as well as the group III variant in plant F.

Plant 20, the representative of the group III variant, displayed symptoms for the first time in 2001 and was an isolated focus point of leafroll disease in the vineyard. Four years later in 2005, plant A, directly next to plant 20 in the row, displayed symptoms and was found to be infected with the group III variant. The other plant recorded to be leafroll infected in 2005 was plant J, diagonally left of plant 20. Plant J was infected with the group II variant. It is possible that plant J was infected by a mealybug that fed on a plant infected with the group II variant, from outside the 10 plants in the cluster, and transmitted the variant to plant J. None of the other plants in the disease cluster showed symptoms in 2005, but in 2008, when the plants in the cluster were tested, all plants tested positive for GLRaV-3 and contained the group II variant, according to SSCP profiles, except plant C that was infected with a combination of group II and III variants. The movement of mealybugs from the surrounding area could have transmitted variants in the cluster.

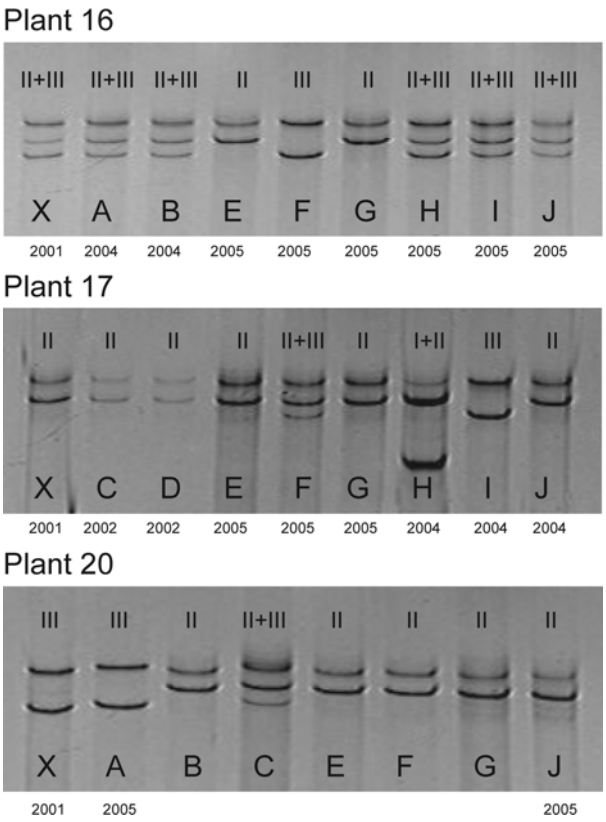


Figure 3.SSCP analyses of three disease clusters in Block 10 in the Worcester vine growing region, Western Cape. SSCP profiles of eight surrounding plants of plants 16, eight plants surrounding plant 17 and seven plants surrounding plant 20 are shown.

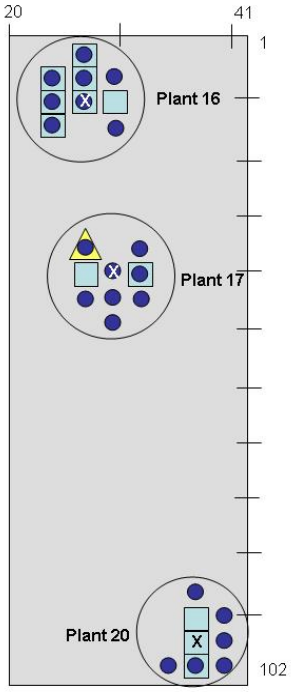


Figure 4.A graphic presentation of the relative position of the infection clusters in the Worcester vineyard. Rows are indicated horizontally and plant number vertically. The original infected plant of an infection cluster is marked as X, and represents plants 16, 17 and 20, respectively. The GLRaV-3 variant status of every plant in the disease cluster is indicated by the following symbols: variant I (△), variant II (○), variant III (□). Mixed infections with more than one variant are indicated with a combination of the symbols.

4.5 DISCUSSION

The distribution of three GLRaV-3 variants in local vineyards was investigated, with the focus on the presence of the three GLRaV-3 variants in leafroll disease clusters. A field survey was done; firstly to determine which of the variants occurred predominantly in the selected mother blocks and secondly, to determine if there were any differences in the spatial distribution patterns of the three GLRaV-3 variants. The distribution of GLRaV-3 variants was compared with the spatial distribution study of leafroll disease within mother blocks of the South African Certification Scheme, from 2001-2005 (Pietersen 2004; Pietersen 2006). Distribution data of newly-infected leafroll plants within vineyards, based on the visual inspection of diseased plants on a yearly basis, was used as basis for this study. Each of the individual monitored vineyards was planted at the same time with the same cultivar, and grafted on the same rootstock, and therefore it validates the monitoring strategy based on symptom expression. The distribution of a specific GLRaV-3 variant in a disease cluster was based on the appearance of GLR symptoms in plants.

In previous studies, phylogenetic analysis of the Hsp70h and coat protein gene sequences deposited in Genbank, showed that the majority of GLRaV-3 isolates grouped into the variant group I clade, represented by the South African accession 621 and American isolate NY-1 (Ling et al. 2004; Fuchs et al. 2009b; Jooste et al. 2010). However, in the current study, we determined that variant group II is the predominant variant in the plants collected from 14 mother blocks, suggesting that variant II is the most widespread in local vineyards. Whether this difference reflects a unique situation in vineyards in South Africa is still unknown and warrants further investigation. Factors like specific virus-vector interactions, weather conditions or the dissemination of group II variant-infected propagation material may individually or collectively contribute to the fact that the group II variant occur predominantly in South Africa. The effectiveness with which the vine mealybug, *P. ficus*, transmits the different variants needs further investigation.

We made the assumption that the dominant variant in a plant would be detected with SSCP analysis. The SSCP profiles characteristic for each variant group, as determined previously (Jooste et al. 2010), could be corroborated by sequence data. Furthermore, we assumed that plants with the same SSCP profile contained the same variant(s), as observed in the Turturo et al. (2005) study. The upper band of the SSCP profiles in plants 1.8, 54.7, 64.2 and 64.3 (Figure 2) are slightly higher than the typical SSCP profile (plants 1.4, 1.5, 1.7) for group II variants. Sequence data previously obtained from plants 5 and 48 described in the Figure 1 in Jooste et al. (2010), page 48 in this thesis, showed that a two-nucleotide change resulted in the SSCP profile with the slightly higher upper band. Similarly, the four-band profile detected in plant 54.3 was also found in the Jooste et al. (2010) study in plants 7, 12, and 17, and, based on sequence results, were shown to be group II members. Sequence results obtained from four clones from each plant with a 'duplicate' four-band profile proved to be identical to that of plant 5, which has a two-band profile, and represent the group II profile (Jooste et al., 2010). Plant 54.3 was therefore identified as a group II variant. Based on SSCP results, the variant status of originally infected plants, plants 16, 17 and 20 in this study, remained the same in two consecutive years, 2007-2008.

In plants 17 and 20 (Figure 3 and 4), single variant infections were detected, but adjacent plants had mixed variant infections. We observed that the single variant infection detected with SSCP analysis in plant 20 (group III variant) only spread to an adjacent plant four years later, suggesting that this variant maybe slower in expressing symptoms or is spread slower

by mealybugs in contrast with the group II variant, plant 17, detected in adjacent plants a year later. This observation still needs to be proved. Plants 20 B, C, E, F and G all tested positive for GLRaV-3 in 2008. The exact year when symptoms became visible were not recorded in these plants. A combination of variants in a plant may be transmitted more effectively to adjacent plants as seen in the transmission of variants in the plant 16 disease cluster. It was not possible to analyse plants in positions 16 C, D, 17 A, B or plants 10 I, H as these positions did not contain vines in the field. In the disease clusters analysed it is clear that the spread of variants was not always from the originally infected plant, marked X. Plausible explanations would be that these plants were new primary infections due to 1) first instar mealybugs being dispersed long distances by wind, machinery or workers clothing, or 2) adult mealybugs moving around from plant to plant.

It is likely that plants with older leafroll symptoms are infected with combinations of variants as seen in certain of the SSCP profiles of Figure 2, for example plants in blocks 4 and 17 (Table 1). This suggests that with time, plants might get infected with a selection of variants, depending on the mealybug's transmission efficiency of different variants. The influence of an insect vector on the change of a viral population was discussed in work done on a well-studied *Closterovirus, Citrus tristeza virus* (CTV) (Ayllón et al., 1999; d'Urso et al., 2000). SSCP analysis of genes p18 and p20 showed that the profiles characteristic of field CTV isolates were frequently altered after aphid transmission (d'Urso et al., 2000). Similarly, it has been shown (Brlansky et al., 2003) that frequencies of genomic variants in a *Citrus tristeza virus* (CTV) populations may alter following aphid-mediated virus transmission to a new host. Another study by Roy & Brlansky (2009), proved the generation of virus recombinants after aphid transmission. Different dominant genotypes were detected in the parent and aphid-transmitted (AT) sub isolates and even intermediate genotypes were detected that differed from the parental or AT sub isolates (Roy & Brlansky, 2009). A study by Broadbent et al. (1996) showed that the influence of aphid transmission sometimes even alters pathogenic characteristics in CTV. Whether mealybug vectors influences the transmission efficiency of specific GLRaV-3 variants or are influencing the viral population in a certain way, needs to be studied.

The classification of GLRaV-3 variants reported in this and previous studies (Jooste & Goszczynski, 2005; Jooste et al. 2010) are similar to the phylogenetic classification described by Fuchs et al. (2009a). The Fuchs study classified GLRaV-3 accessions into groups NY-1, GP-18, C5-1, MT48-2 and NZ-1. Jooste et al. (2005 and 2010) identified three variant groups

and numbered them I, II and III. The clades identified are similar to the Fuchs study. In the Jooste et al. (2010) paper the genome of isolate 621 (group I) was completely sequenced, and could replace NY-1 as representative of group I. GP-18 is the representative of group II variants and 623 joins this variant group (both complete genomes). PL-20 (complete genome), similar to MT48-2 (partial sequence), from the Jooste et al (2010) study is the representative of variant group III. We propose that isolates from the NZ-1 clade belong to variant group IV and isolates from the C5-1 clade to variant group V, with representative isolate sequences to be determined when more data is available. In our opinion the proposed Roman numeral classification system for the different variant groups of GLRaV-3 is a sensible approach that will reduce the risk for confusion.

This is the first study to show that a specific GLRaV-3 variant, or a combination of GLRaV-3 variants, is transmitted to adjacent plants in a leafroll infected cluster. As discussed above, the importance of the interaction between the mealybug vector and a specific GLRaV-3 variant warrants further investigations. Mealybug populations in the vineyards and their role in transmitting the GLRaV-3 variants from plant to plant were not monitored in this study. Results of this study concluded that GLRaV-3 variants from group II occur predominantly in the vineyards studied. It is important to continue to study the biological properties of GLRaV-3 variants, including their possible role in causing mild or severe symptoms in grapevine plants.

4.6 LITERATURE

Ayllón MA, Rubio L, Moya A, Guerri J, Moreno P (1999) The haplotype distribution of two genes of *Citrus tristeza virus* is altered after host change of aphid transmission. Virology255:32-39

Belli G, Fortusini A, Casati P, Belli L, Bianco PA, Prati S (1994) Transmission of a grapevine leafroll associated closterovirus by the scale insect *Pulvinaria vitis* L. Rivista di Patologia Vegetale 4:105-108

Brlansky RH, Damsteegt VD, Howd DS, Roy A (2003) Molecular analysis of *Citrus tristeza virus* subisolates separated by aphid transmission. Plant Dis 87:397-401

Broadbent P, Brlansky RH, Indsto J (1996) Biological characterisation of Australian isolates of citrus tristeza virus and separation of subisolates by single aphid transmissions. Plant Dis 80:329-333

Cabeleiro C, Segura A (1997a) Some characteristics of the transmission of grapevine leafroll associated virus 3 by *Planococcus citri* Risso. Eur J Plant Pathol 103:373-378

Cabeleiro C, Segura A (1997b) Field transmission of Grapevine leafroll Associated virus 3 (GLRaV-3) by the mealybug *Planococcus citri*. Plant Dis 81:283-287

Cabeleiro C, Segura A (2006) Temporal analysis of grapevine leafroll associated virus 3 epidemics. Eur J Plant Pathol 114:441-446

Cabeleiro C, Couciero C, Pereira S, Cid M, Barrasa M, Segura A (2008) Spatial analysis of epidemics of grapevine leafroll associated virus-3. Eur J Plant Pathol 121:121-130

Carstens R (2002, January) Vineyard viruses:Leafroll overview. Wynboer-a technical guide for wine producers.Retrieved March 2010, from http://www.winetech.co.za/news_leafroll.php3.

Douglas N, Krüger K (2008) Transmission efficiency of grapevine leafroll-associated virus 3(GLRaV-3) by the mealybug *Planococcus ficus* and *Pseudococcus longispinus* (Hemiptera:Pseudococcidae). Eur J Plant Pathol122:207-212

d'Urso F, Ayllón MA, Rubio L, Sambade A, de Mendoza H, Guerri J, Moreno P (2000) Contribution of uneven distribution of genomic RNA variants of citrus tristeza virus (CTV) within the plant to changes in the viral population following aphid transmission. Plant Pathol49:288-294

Engelbrecht DJ, Kasdorf GGF (1985) Association of a closterovirus with grapevines indexing positive for grapevine leafroll disease and evidence of its natural spread in grapevine. Phytopathologia Mediterranea 24:101-105

Engelbrecht DJ, Kasdorf GGF (1990) Transmission of grapevine leafroll disease and associated closteroviruses by the vine mealybug *Planococcus ficus*. Phytophylactica 22:341-346

Fuchs M, Masella-Herrick P, Loeb GM, Martinson TE, Hoch HC (2009a) Diversity of Ampeloviruses in mealybug and soft scale vectors and in grapevine hosts from leafroll-affected vineyards. Phytopathol 99:1177-1184

Fuchs M, Martinson TE, Loeb GM, Hoch HC (2009b)Survey for the three major leafroll disease-associated viruses in Finger Lakes vineyards in New York. Plant Dis 93:395-401

Golino DA, Weber E, Sim S, Rowhani A (2008) Leafroll disease is spreading rapidly in a Napa Valley vineyard. California Agriculture 62(4):156-160. DOI:10.3733/ca.v062n04p156

Goszczynski DE, Jooste AEC (2002) The application of single-strand conformation polymorphism (SSCP) technique for the analysis of molecular heterogeneity of grapevine virus A. Vitis 41:77-82

Jooste AEC, Goszczynski DE (2005) Single-strand conformation polymorphism (SSCP), cloning and sequencing reveal two major groups of divergent molecular variants of grapevine leafroll-associated virus 3 (GLRaV-3). Vitis 44(1):39-43

Jooste AEC, Maree HJ, Bellstedt D, Goszczynski DE, Pietersen G, Burger JT (2010) Three genetic grapevine leafroll-associated virus-3 variants identified from South African vineyards show high variability in their 5'UTR, DOI 10.1007/s00705-010-0793-y. Arch Virol, 155(12):1997-2006

Ling KS, Zhu HY, 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

Maree HJ, Freeborough M-J, Burger JT (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

Petersen CL, Charles JG (1997) Transmission of grapevine leafroll-associated closterovirus by *Pseudococcus longispinus* and *P. calceolariae*. Plant Pathol46:509-515

Pietersen G (2004, June) Spread of Grapevine leafroll disease in South Africa - a difficult, but not insurmountable problem. Retrieved from Wynboer –A technical guide for wine producers, June 2004, http://www.wynboer.co.za/recentarticles/articles.php3

Pietersen G (2006) Spatio-temporal dynamics of grapevine leafroll disease in Western Cape vineyards. Extended Abstracts, 15th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine, 2006. Stellenbosch, South Africa, pp 126-127

Rayapati N, Tefera MA, Olufemi AJ, Sridhar J, Gandhi K, Gutha LR, Martin RR (2009) Current status of grapevine viruses in the Pacific Northwest vineyards of the United States. Extended Abstracts, 16th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine, Dijon, France, pp 108-109.

Roy A, Brlansky RH (2009) Population dynamics of a Florida *Citrus tristeza virus* isolate and aphid-transmitted subisolates: Identification of three genotypic groups and recombinants after aphid transmission. Phytopathol99:1297-1306

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

Tsai C-W, Chau J, Fernandez L, Bosco D, Daane KM, Almeida RPP (2008) Transmission of grapevine leafroll-associated virus 3 by the vine mealybug (*Planococcus ficus)*. Phytopathol98:1093-1098

Turturo C, Saldarelli P, Yafeng D, Digiaro M, Minafra A, Savino V, Martelli GP (2005) Genetic variability and population structure of *grapevine leafroll-associated virus* 3 isolates. J Gen Virol 86:217-224

CHAPTER 5

CONCLUSIONS

Grapevine leafroll (GLR) disease is one of the most important diseases of grapevines, and occurs in all grape-producing countries worldwide, including South Africa. The rapid spread of GLR disease and the associated GLRaV-3 in South African vineyards (Pietersen, 2004) is of major concern to the local grapevine industry and makes GLRaV-3 an important virus to study. In this project the molecular variability of GLRaV-3 isolates, and the spread of GLRaV-3 variants within GLR-infected clusters were investigated.

In this study we confirmed the result of an extended 5'UTR of GLRaV-3 described by Maree et al. in 2008, with the full-length nucleotide sequences of representatives of variant groups I and II, namely isolates 621 (GQ352631) and 623 (GQ352632). Field surveys done in mother blocks in 2007 and 2008 resulted in the identification of a third molecular variant and added another objective - to determine the full-genome sequence of this new variant - to the study. The complete nucleotide sequence of this isolate represented by PL-20 (GQ352633), showed that this molecular variant is significantly different, especially in the 5'UTR, from isolates 621 (group I), 623 and GP18 (group II). We therefore propose that isolate PL-20 represents a third molecular variant, group III.

The variation in the 5'UTR sequences presented here correlated with the characteristic SSCP profiles generated from amplicons of ORF5 of the three variants (621, 623 and PL-20). Compared to the rest of the genome, the 5'UTR is highly variable between GLRaV-3 molecular variants. The significance of the extended 5'UTR is not yet known and future studies to determine the possible function of this region is important, especially since bioinformatic analyses of the 5'UTRs, *e.g.* conserved functional motifs, ORF analyses and predicted secondary structure of these regions for the three molecular variants were unable to predict a possible function.

The identification of the genetic variants of GLRaV-3 was described in Chapter 3. Results from this study suggest that more than the current three GLRaV-3 variants may exist in South African vineyards. These results were generated from SCCP analysis, and subsequent

corroboration by sequencing of the amplicons of ORF 5. In order to confirm the early suggestion that even more variants may exist, whole-genome sequencing of viral isolates from other regions of the world will have to compared with those identified in the current study as potential additional variants.

Another aspect investigated in this study was the distribution and spread of the genetic variants in selected vineyards in the Western Cape. The results of the field surveys were described in detail in Chapter 4. We concluded that a specific GLRaV-3 variant, or a combination of GLRaV-3 variants, is transmitted to adjacent plants in a leafroll-infected cluster. In the disease clusters analysed, it is clear that the spread of variants was not always from the originally infected plant, marked X. Theoretically it is possible that not all genetic variants in plant X were detected with the H420, C629 primer set, especially since primer H420 was chosen from a region with limited homology. A second more plausible explanation could be that these surrounding plants were new primary infections due to 1) first instar mealybugs being dispersed long distances by wind, machinery or workers' clothing, or 2) adult mealybugs moving around from plant to plant. The importance of the interaction between the mealybug vector and specific GLRaV-3 variants warrants further investigation. Mealybug populations in the vineyards and their role in transmitting the GLRaV-3 variants from plant to plant were not monitored in the current study, but studies to investigate these interactions are underway. Future studies will include transmission efficiency experiments of the three identified GLRaV-3 variants in South African vineyards.

Based on the results of this study we concluded that GLRaV-3 variants from group II occurred predominantly in the vineyards we studied. Whether the dominant occurrence of group II variants reflects a unique situation in vineyards of South Africa is still unknown and warrants further investigation. The abundance of the group II variant may be the result of positive selection, a process by which the variants that are the fittest increase their frequency in a population, or could simply have resulted from the accidental use of infected plant material or infected rootstocks.It is important to continue to study the biological properties of GLRaV-3 variants, including their possible role in causing mild or severe symptoms in grapevine plants. The future challenge will be to associate pathogenicity of GLR with a specific GLRaV-3 variant and to determine if the mealybug vectors transmit specific GLRaV-3 variant(s) more efficiently.

In many of the plants in this study, combinations of GLRaV-3 variants were detected by SSCP analysis. The mealybug vector is probably the main contributor to the spread of the genetic variability that already exists in the vines. It is possible that grapevine plants acquire multiple GLRaV-3 variants (and other grapevine viruses) when mealybugs transmit these viruses from plant to plant. The role of recombination in GLRaV-3-infected plants has not been studied previously. The possible role of recombination in the emergence of new GLRaV-3 variants needs to be investigated.

This study clearly demonstrated that SSCP analysis of a portion of ORF5 provided a fast and reliable indication of GLRaV-3 variant status in a plant. One can argue that the SSCP analysis, without confirmation with sequence data, was not sensitive enough to detect all the genetic variants in vineyards. However, repeated SSCP analyses of this region in ORF5 and corroboration of these SSCP profiles with sequence data strongly support our notion that this is a reliable method to identify individual or complexes of variants infecting individual plants. All the GLRaV-3 infected plants screened to date were amplified with the H420 and C629 primers, positioned in ORF5. Other regions of the genome were also investigated with SSCP analysis and sequence comparisons. The predicted three variant groups were further confirmed with SSCP profiles from amplified products of multiple sets of primers, spanning the entire genome, as well as by sequencing these regions. We therefore propose that this method be used as an initial screening of field material to give a preliminary indication of variant status of plants.

Moreover, the use of the SSCP technique is relatively simple, quick and inexpensive and does not require expensive equipment. The question if we can solely rely on SSCP analysis to identify genetic variants in plants needs attention. The fact that single base changes alter SSCP profiles can result in complicated SSCP profiles that are not easy to interpret. SSCP profiles of the group II variant, described in this study, showed more variable and complicated SSCP profiles in different genomic regions. Variation within this group of variants should therefore be further characterised. Variant group I and III profiles were always distinct in all regions studied. With genome sequence data of several GLRaV-3 variants available now, it is possible to design variant-specific primers to detect variants individually and universal primers to detect all known GLRaV-3 variants in grapevine plants. This could be useful in studies designed for the detection new viruses and variants of viruses. The recent deep sequencing of viruses infecting grapevines provides the ultimate way to detect new viruses

and variants of viruses (Coetzee et al., 2010), but is prohibitively expensive for most laboratories.

The number of full-length GLRaV-3 sequences submitted to GenBank was doubled by this study, with the addition of the 621, 623 and PL-20 sequences (Jooste et al. 2010). The other full-length sequence is the published GP18 sequence (Maree et al., 2008) and the Chilean Cl-766 and NY-1 accessions, the latter two (Engel et al., 2008, Ling et al, 2004) lacking the extended 5'UTR ends.

The results obtained in this study form a basis for future studies to further advance our understanding of GLRaV-3 and the spread and management of the virus. Some future prospects from this study include:

1) The significance of the extended 5'UTR, described by Maree et al., 2008, is not yet known and future studies to determine the function of this region are important.
2) Whole-genome sequencing of GLRaV-3 variants from other regions of the world and South Africa should be undertaken to obtain a complete picture of GLRaV-3 variability.
3) The role of recombination in the emergence of new GLRaV-3 variants needs to be studied.
4) Molecular detection methods for these variants should be optimised.
5) The biological properties of the virus are important to consider and the interaction between the mealybug vector and specific GLRaV-3 variants needs to be studied.
6) A transmission efficiency experiment is necessary to explain the fast spread of certain variants in a vineyard.
7) It will be interesting to continue the search for more GLRaV-3 variants in vineyards, also in white cultivars and the occurrence of variants in table grapes.

Finally, the taxonomical classification of GLRaV-3 variants presented in this thesis and related publications (Jooste et al, 2005, 2010), is based on the Roman numerical classification system, and is similar to the phylogenetic classification described earlier (Fuchs et al., 2009). We support the proposed Roman numeral classification system for the different variant groups of GLRaV-3 as a sensible approach to classify GLRaV-3 variants, which will reduce the risk of confusion.

LITERATURE

Coetzee B, Freeborough MJ, Maree HJ, Celton JM, Rees DJG, Burger JT (2010) Deep sequencing analysis of viruses infecting grapevines: Virome of a vineyard. Virology 400:157-163

Engel EA, Girardi C, Escobar PF, Arredondo V, Dominguez C, Perez-Acle T, Valenzuela PDT (2008)Genome analysis and detection of a Chilean isolate of *Grapevine leafroll associated virus-3*. Virus Genes 37:110-118

Jooste AEC, Goszczynski DE (2005) Single-strand conformation polymorphism (SSCP), cloning and sequencing reveal two major groups of divergent molecular variants of grapevine leafroll-associated virus 3 (GLRaV-3). Vitis 44(1):39-43

Jooste AEC, Maree HJ, Bellstedt D, Goszczynski DE, Pietersen G, Burger JT (2010) Three genetic grapevine leafroll-associated virus-3 variants identified from South African vineyards show high variability in their 5'UTR. Arch Virol 155(12):1997-2006

Ling KS, Zhu HY, 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

Maree HJ, Freeborough M-J. Burger JT (2008) Complete nucleotide sequence of a South African isolate of *grapevine leafroll-associated virus* 3 reveals a 5'UTR of 737 nucleotides.Arch Virol153:755-757

Pietersen G (2004) Spread of Grapevine leafroll disease in South Africa - a difficult, but not insurmountable problem.Wynboer –A technical guide for wine producers. June 2004. URL: http://www.wynboer.co.za/recentarticles/articles.php3

Supplementary data

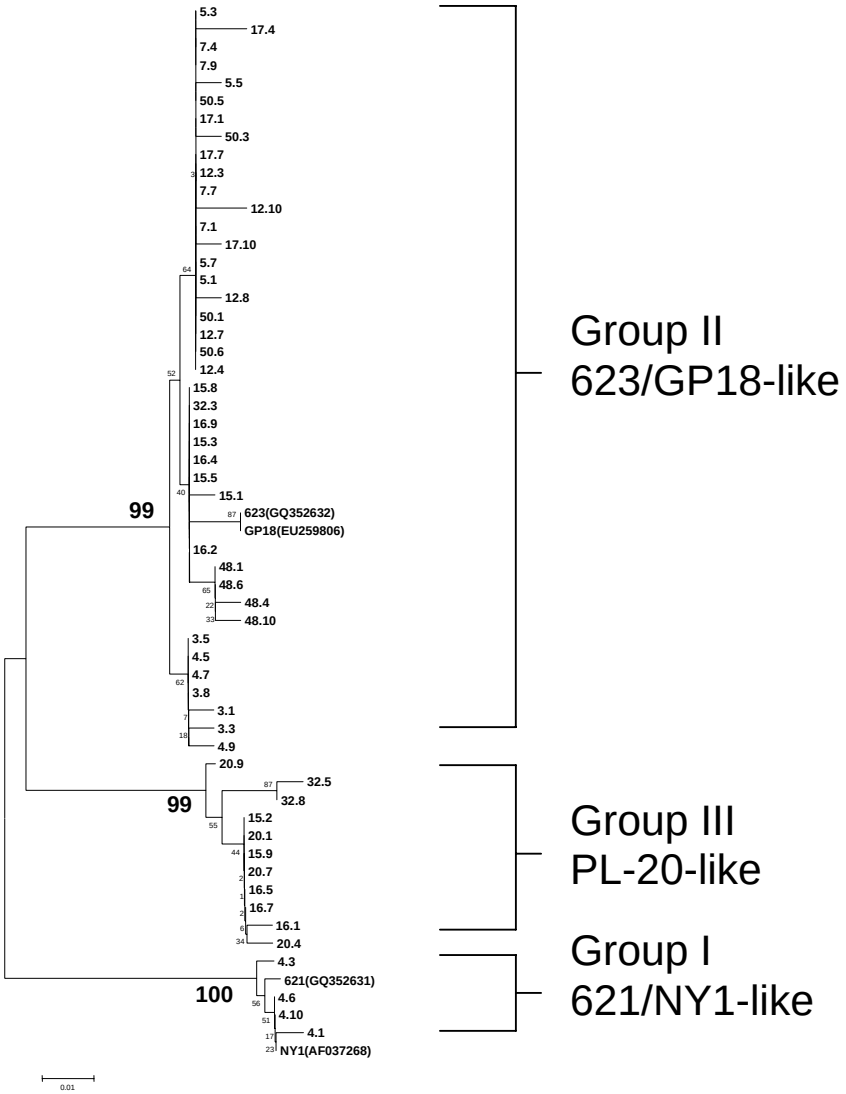


Figure 1. Alignment of 55 clones from twelve plants, selected in Figure 2 in the main text, and four reference sequences (GP18, 623, 621 and NY-1) selected for sequencing to confirm variant groups of clones detected with SSCP analysis.

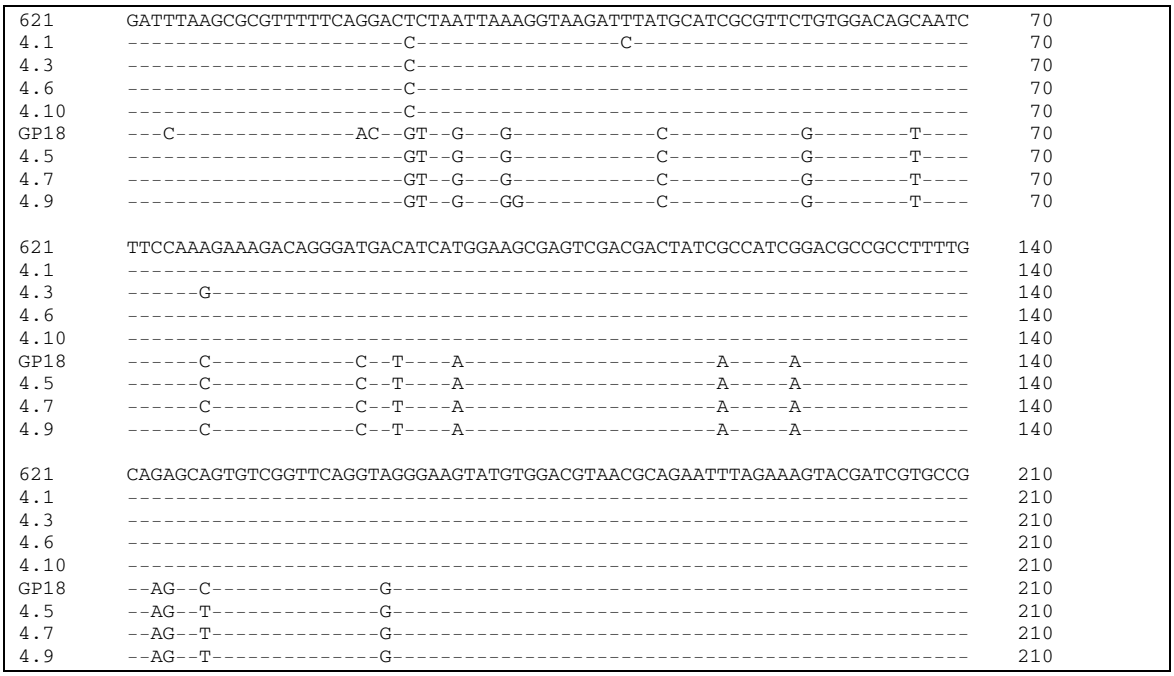


Figure 2. Clones derived from plant 4 aligned with 621 (variant group I) and GP 18 (variant group II) sequences. SSCP profiles of clones 4.5, 4.7, and 4.9 clearly showed atypical profiles of the dominant group I variant in plant 4.

Table 1. Primer sequences to amplify eleven additional region in the genomes of plants 3,4,5,7,12,15,16,17,20,32,48, and 50 to study the variability of their genomes with SSCP and sequencing

Primer pair nb	Position	Forward primer	Primer sequence (5'-3')
1	ORF1a	GL3.253	AGC TTT CCT AAC CAC CAT GAAG
2	ORF1a	GL3.3874	GGG GCT TGC TTA ACG ACA C
3	ORF1a	GL3.5297	GTC ACC AGG TGT TCC AAA CC
4	ORF1b	GL3.7466	CGC CAT TGT CGA AGT ACG
5	ORF4	GL3.10953	ACT GGA CGC AAT CGT TGC
6	ORF5	GL3.12212	CGA TCG TGC CGT TAA GAG
7	ORF6	GL3.13450	CCA CAG CAG CTT TGG CTA C
8	ORF7	GL3.14893	TTC CCC AGT GCG CAT CTT C
9	ORF8	GL3.15736	TAA TTA CAG TTC GCC GTG ATC C
10	ORF9	GL3.16399	GGG TGC CGT ATT AAG AGA CTC
11	ORF10	GL3.17104	CTG GAG GTC ATC AAG TCG ATG

Primer pair nb	Position	Reverse primer	Primer sequence (5'-3')	Product size
1	ORF1a	GL3.539	CGG TGG TAG TGT ACG ACA AG	286
2	ORF1a	GL3.4133	CAC GCT TCG AGG TGA ATG G	259
3	ORF1a	GL3.5591	AAC GCC CTG TAT GTC CTC TC	292
4	ORF1b	GL3.7707	CCT GCT TCA TGA GAG CAC TC	241
5	ORF4	GL3.11250	TTA CCA CCG GCT GAA G	298
6	ORF5	GL3.12524	TGA CCA GCT TGA GCG TAG	312
7	ORF6	GL3.13816	GAA CTC CGT CGA AGA CGA TG	366
8	ORF7	GL3.15218	GCC TTT CGA AAG AAC GAG TC	326
9	ORF8	GL3.16038	TAG GTA ACC GGC GCG TTG	302
10	ORF9	GL3.16653	AAC GTC GGA TCC ACA ATC AC	254
11	ORF10	GL3.17316	CAA CAA AGC GTC AAG AGC AAC	212

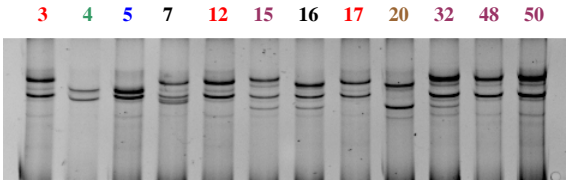


Figure 3a.SSCP analysis from a region in ORF1a show at least five different SSCP profiles.

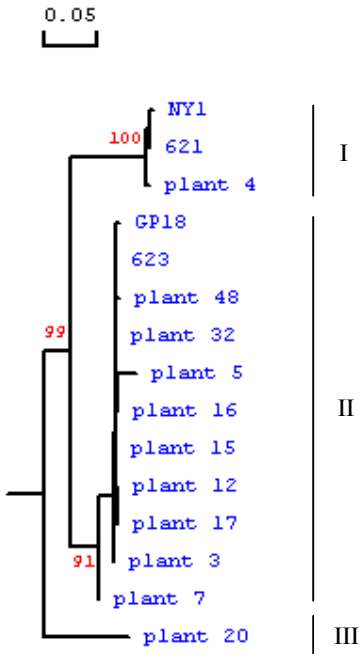


Figure 3b. Phylogenetic analysis of sequence data from a region in ORF1a confirms that there are three main variant groups in the plants analysed.

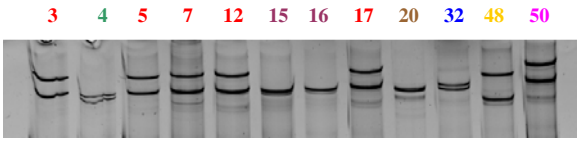


Figure 4a.SSCP analysis from a region in ORF5 show at least four different SSCP profiles.

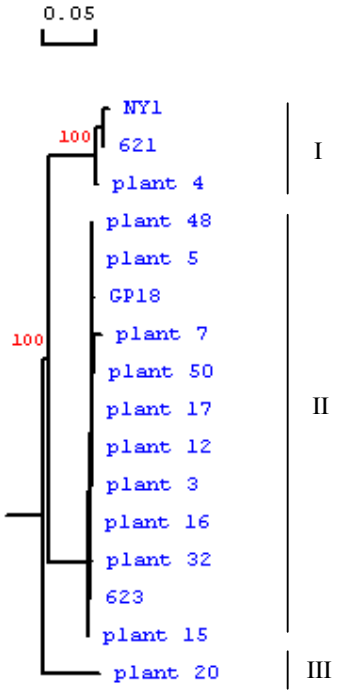


Figure 4b. Phylogenetic analysis of sequence data from a region in ORF5 confirms that there are three main variant groups in the plants analysed.

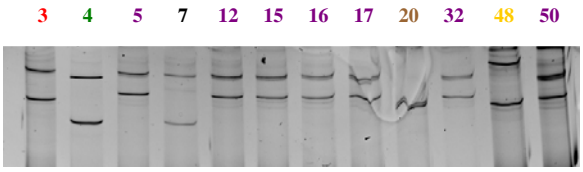


Figure 5a.SSCP analysis from a region in ORF7 show at least six different SSCP profiles.

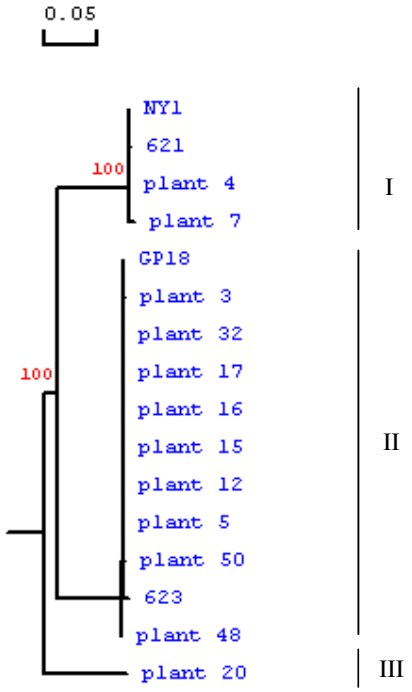


Figure 5b. Phylogenetic analysis of sequence data from a region in ORF7 confirms that there are three main variant groups in the plants analysed.

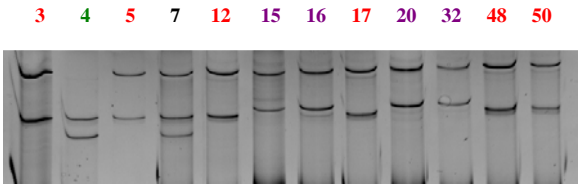


Figure 6a.SSCP analysis from a region in ORF9 show at least four different SSCP profiles.

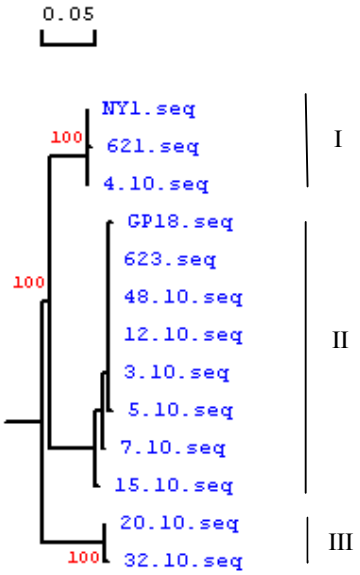


Figure 6b. Phylogenetic analysis of sequence data from a region in ORF9 confirms that there are three main variant groups in the plants analysed.

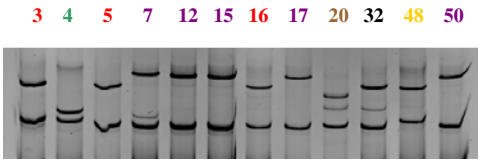


Figure 7a.SSCP analysis from a region in ORF10 show at least four different SSCP profiles.

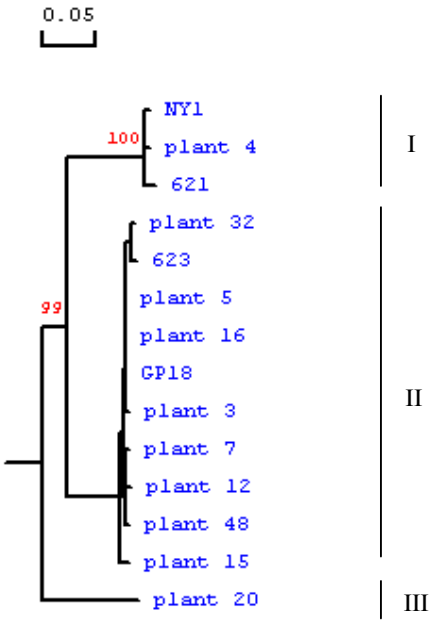


Figure 7b. Phylogenetic analysis of sequence data from a region in ORF10 confirms that there are three main variant groups in the plants analysed.