View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Sapientia

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências e Tecnologia

APPLICATION OF MOLECULAR TOOLS FOR DETECTION OF PLANT VIRUSES

Submitted by FILIPA ESTEVES

Dissertation for obtaining a Master degree in

Molecular and Microbial Biology

Supervised by Prof. Dr. Filomena Fonseca

Faro, Portugal
2013

UNIVERSIDADE DO ALGARVE

Faculdade de Ciências e Tecnologia

APPLICATION OF MOLECULAR TOOLS FOR DETECTION OF PLANT VIRUSES

Submitted by FILIPA ESTEVES

Dissertation for obtaining a Master degree in

Molecular and Microbial Biology

Supervised by Prof. Dr. Filomena Fonseca

Faro, Portugal
2013

Declaration

I declare this thesis submitted by me for the degree of Magister Scientiae in Molecular and Microbial Biology at the University of Algarve is result of my research work and its contents are my responsibility.

De acordo com a alínea iii) do ponto 4.1 do Anexo II a que se refere o ponto 2 do Artigo 18º do Regulamento nº 287/2021 publicado em Diário da Repúbica, a 24 de Julho de 2012: "Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída."

____

Filipa Esteves

©Copyright by Filipa Alexandra Oleiro Esteves. All Rights Reserved.

Após publicação do seu conteúdo em revistas da especialidade e de acordo com o ponto 4.2 do Anexo II a que se refere o ponto 2 do Artigo 18º do Regulamento nº 287/2021 publicado em Diário da Repúbica, a 24 de Julho de 2012: "A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor."

Acknowledgments

I would like to thank my supervisor Prof. Dr. Filomena Fonseca for allowing me to pursuit my interest in science. I would like to thank you for giving me the opportunity to work with you. I also thank you for the assistance and advice as well for the support and guidance during the development of this project. In addition, I would like to thank the recommendations and corrections to this manuscript.

I would like to thank David for his patience and support during this time. Thank you for your tolerance and love even in those moments when my temper was particularly challenging.

I also would like to thank Gui for making be stop and take a breath of fresh air every day during this process.

Finally, I would like to thank my mother for her unconditional love, support and encouragement through every step of the way. Thank you for always be beside me and share my path. I thank you for being an extraordinary reliable person through my entire life and for always believing in me. Without you I could not have made it here.

Abstract

Grapevine leafroll disease (GLRD) is one of the most important virus diseases of grapevines worldwide, causing major economical impact. The disease has a complex aetiology and currently eleven phloem-limited viruses, termed in general Grapevine leafroll-associated virus (GLRaVs), have been identified. Two of the GLRaVs, GLRaV-1 and GLRaV-3, are included in the European certification scheme of propagation material. However, the flawed notion that GLRaV-3 is more frequent than GLRaV-1 and that all other GLRaVs are possibly not as relevant for GLRD, has until now precluded the development of specific serological and molecular detection assays and limited the scope of molecular characterization of the viruses known to be associated with the disease. Hence, few studies have addressed the phylodynamics of GLRaVs or even characterized the genetic structure of their natural populations. This generalized lack of molecular information, in turn underlie the deficient capacity to detect the viruses.

The phylogenetic analyses were conducted on the basis of the heat shock protein 70 homologue (HSP70h) and the coat protein (CP) genes for GLRaV-1 and the HSP70h, the heat shock protein 90 homologue (HSP90h) and the CP genes for GLRaV-5. The data obtained for GLRaV-1 contributed 83 new CP sequences. This information was combined with previous analysis by other authors and used for the production of new polyclonal IgG, capable of detecting CP variants from all the phylogroups observed. Successful testing of this new tool included tissue print immunoblotting (TPIB) and in situ immunoassay (ISIA).

The data obtained for GLRaV-5, contributed 61 new CP and 28 new HSP90h gene sequences. Eight phylogenetic groups were identified on the basis of the CP. Characterization of the genetic structure of the isolates revealed a higher diversity than previously reported and allowed the identification of dominant virus variants.

For both GLRaV-1 and GLRaV-5, the effect of vegetative propagation on the virus transmission dynamics was addressed.

Keywords: GLRaV-1; GLRaV-5; genetic diversity; diagnosis; transmission dynamics.

Resumo

A Doença do Enrolamento da Videira (GLRD) é uma das mais importantes doenças virais que afectam as videiras a nível mundial, causando um grande impacto económico. A doença tem uma etiologia complexa e, actualmente, onze vírus associados ao floema, designados por *Grapevine leafroll-associated virus* (GLRaVs) foram identificados. Dois dos GLRaVs, os GLRaV-1 and GLRaV-3, estão incluídos na lista de vírus de certificação obrigatória em material de propagação vegetativa indicada pelo esquema de certificação de material de propagação da União Europeia. Contudo, a ideia imprecisa de que o GLRaV-3 é mais frequente que o GLRaV-1 e que, todos os outros GLRaVs, possivelmente, não serão tão revelevantes para a GLRD, tem excluído o desenvolvimento de análises de detecção serológica e molecular e limitado a caracterização molecular dos vírus associados à doença. Deste modo, poucos estudos trataram a filodinâmica dos GLRaVs ou caracterizaram a estrutura genética das suas populações naturais. Por seu turno, a falta generalizada de informação molecular é subjacente à deficiente capacidade de detecção dos vírus.

Para o GLRaV-1, as análises filogenéticas foram efectuadas com base no gene que codifica a *heat shock protein 70 homologue* (HSP70h), bem como no gene que codifica a proteína da cápside (CP). Para o GLRaV-5, as análises foram realizadas com base nos genes que codificam a HSP70h, a *heat shock protein 90 homologue* (HSP90h) e a CP. Os dados obtidos para o GLRaV-1 contribuíram com 83 novas sequências da CP. Esta informação, juntamente com análises anteriores relizadas por outros autores, foi utilizada para produzir um novo IgG policlonal, capaz de detectar as variantes da CP de todos os filogrupos observados. Esta nova ferramenta de detecção foi testada com êxito por *tissue print immunoblotting* (TPIB) e por *in situ immunoassay* (ISIA).

A informação obtida para o GLRaV-5, contibuiu com 61 e 28 novas sequências da CP e HSP90h, respectivamente. Oito grupos filogenéticos foram identificados com base na CP. A caracterização da estrutura genética dos isolados revelou uma maior diversidade que a anteriormente publicada e permitiu a identificação de variantes virais dominantes.

O efeito da propagação vegetativa na dinâmica de transmissão do vírus, foi analisado para o GLRaV-1, bem como para o GLRaV-5.

TABLE OF CONTENTS

Declaration 1

Acknowledgments 3

Abstract 5

Resumo 7

Table of Contents 9

Presentations and Publications 13

List of abbreviations 15

Chapter 1. Introduction and literature review 21

1.1. Viruses in general 21

1.2. Plant Viruses 22

1.3. Economic Importance of Plant Viruses 22

1.4. Virus Transmission in Plants 23

1.5. Virus Diseases of Grapevine 24

1.5.1. Leafroll Disease 25

1.5.1.1. Economic Impact 25

1.5.1.2. Symptomatology 25

1.5.1.3. Implicated Visuses 27

1.5.1.4 Cytopayhology 28

1.5.1.5. Epidemiology 29

1.5.1.6. Certification Scheme of Vegetative Propagation Material 30

1.6. Diagnostic Techniques of Plant Viruses 31

1.6.1. Visual Detection and Biological Indexing 31

1.6.2. Serological Tests 32

1.6.3. Molecular Detection and Typyfycation 34

1.7. Aim of the Work 35

Chapter 2. *Grapevine Leafroll-associated Virus 1* **37**

2.1. Introduction 37

2.2. Material and Methods 39

2.2.1. Sample Collection and Plant Material 39

2.2.2. RNA Extraction 41

2.2.3. cDNA Synthesis and PCR 42

2.2.4. Cloning and Transformation 43

2.2.5. Single-Strand Conformation Polymorphism (SSCP) 44

2.2.6. Plasmid DNA Extraction and Restriction Analysis 44

2.2.7. Sequencing and Sequence Data Analysis 45

2.2.8. Estimates of Genetic Distances and Selection Pressure 46

2.2.9. Recombination Analysis 46

2.2.10. Antibodies Development 46

2.2.11. SDS-PAGE and Western Blot Analysis 47

2.2.12. Tissue Print Immunoblotting (TPIB) 48

2.2.13. In Situ Immunoassay (ISIA) 48

2.2.13.1. ISIA with Polyclonal IgG 48

2.2.13.2. ISIA with MonoIgG-AP 49

2.2.13.3. ISIA with MonoIgG-FITC 49

2.3. Results 50

2.3.1. PCR amplification 50

2.3.2. Genetic Diversity of GLRaV-1 Isolates 50

2.3.3. Phylogenetic Analysis of GLRaV-1 Isolates 53

2.3.4. GLRaV-1 Phylogroups Diversity 56

2.3.5. Recombination and Selection Constraints Analysis 58

2.3.6. Serological Assays 59

2.4. Discussion 65

Chapter 3. *Grapevine Leafroll-associated Virus 5* 69

3.1. Introduction 69

3.2. Material and Methods 71

3.2.1. Sample Collection and Plant Material 71

3.2.2. RNA Extraction and cDNA Synthesis 73

3.2.3. PCR Amplification 73

3.2.4. Cloning and Transformation 74

3.2.5. Single-Strand Conformation Polymorphism (SSCP) 74

3.2.6. Plasmid DNA Extraction and Restriction Analysis 75

3.2.7. Sequencing and Sequence Data Analysis 75

3.2.8. Estimates of Evolutionary Divergence and Selection Pressure 76

3.2.9. Recombination Analysis 76

3.3. Results 77

3.3.1. PCR amplification 77

3.3.2. Diversity of Sequence Variants within GLRaV-5 Isolates 77

3.3.3. Phylogenetic Analysis of GLRaV-5 Isolates and Related Ampeloviruses 80

3.3.4. GLRaV-5 Phylogroups and Related Ampeloviruses Diversity 82

3.3.5. Selection Constraints on CP and HSP90h Genes 85

3.4. Discussion 90

Chapter 4. General Discussion and Conclusions 95

References 99

Presentations and Publications

Part of the results reported in this thesis have been published and presented at scientific meetings.

Publications

Chapter 3:

Esteves, F., Teixeira Santos, M., Eiras-Dias, J. E. and Fonseca F. 2012. Occurrence of grapevine leafroll-associated virus 5 in Portugal: genetic variability and population structure in field-grown grapevines. Arch. Virol. 157(9): 1747-1745. DOI:10.1007/s00705-012-1371-2; IF 2.11.

Poster Presentations

Chapter 2:

Esteves, F., Teixeira Santos, M., Rocha, M. L., Eiras-Dias, J. E. and Fonseca F. 2009. Genetic diversity of portuguese isolates of Grapevine Leafroll-associated Virus 1 (GLRaV-1) based on the capsid protein. In: Proceedings of the 16th Congress of ICVG, Dijon, France, pp. 310-311.

Esteves, F., Teixeira Santos, M., Rocha, M. L., Eiras-Dias, J. E. and Fonseca F. 2012. Novel antibodies for ISIA detection of Grapevine leafroll-associated virus 1 (GLRaV-1) based on the variability of the capsid protein gene. In: Proceedings of the 17th Congress of ICVG, UCDavis, USA, pp. 59-60.

Chapter 3:

Esteves, F., Teixeira Santos, M., Eiras-Dias, J. E. and Fonseca F. 2012. Variability of Grapevine leafroll-associated virus 5 (GLRaV-5) in Portuguese field grown grapevines. In: Proceedings of the 17th Congress of ICVG, UCDavis, USA, pp. 198-199.

List of abbreviations

AA	amino acid
AlkB	alpha-ketoglutarate-dependent deoxygenase
AMV	*Alfafa mosaic virus*
AP	alkaline phosphatase
ArMV	*Arabis mosaic virus*
BCIP	5-bromo-4-chloro-3-indolyl phosphatase
BIC	Bayesian information criterion
bp	base pair
BSA	bovine serum albumin
CAN	Colecção Ampelográfica Nacional
cDNA	complemantary DNA
cm	centimetre
CMV	*Cucumber mosaic virus*
CP	coat protein
CPm	minor coat protein
CTV	*Citrus tristeza virus*
DAS-ELISA	double antibody sandwich enzyme-linked immunosorbent assay
dN	non-synonymous substitutions
dNTP	deoxyribonucleotide triphosphate
ds	double-stranded
dS	synonymous substitutions

DNA	deoxyribonucleic acid
e.g.	*exempli gratia*, for example
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
et al.	*et alia*, with others
FEL	fixed effects likelihood
FITC	fluoresceine isothiocyanate
g	gram
GARD	genetic recombination detection
GFkV	*Grapevine fleck virus*
GFLV	*Grapevine fanleaf virus*
GLRaV	*Grapevine leafroll-associated virus*
GLRD	Grapevine leafroll disease
ha	hectare
HEL	helicase
HSP70h	heat shock protein 70 homologue
HSP90h	heat shock protein 90 homologue
ICTV	International Committee on Taxonomy of Viruses
ICTVdb	International Committee on Taxonomy of Viruses Database
ICVG	International Council for the Study of Virus and Virus-like Diseases of the Grapevine
i.e.	*id est*, that is
IFEL	internal fixed effects likelihood

IgG	Immunoglobulin G
INRB/INIAV	Instituto Nacional de Recursos Biológicos/Instituto Nacional de Investigação Agrária e Veterinária
IPTG	isopropyl β-D-1-thiogalactopyranoside
ISIA	in situ immunoassay
kb	kilobase
kDa	kilodalton
LB	lysogeny broth
L-Pro	papain-like leader protease
M	molar
mA	milliamp
MAb	monoclonal antibody
mg	milligram
min	minute
ml	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
MT	methyltransfrease
NBT	nitroblue tetrazolium
nm	nanometres
nt	nucleotide
OIV	International Organisation of Vine and Wine
ORF	open reading frame

PAb	policlonal antibody
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PVDF	polyvinylidene difluoride
PVP-40	polyvinylpyrrolidone wt 40,000
PVY	*Potato virus Y*
R99	Richter 99
RdRp	RNA-dependent RNA polymerase
REL	random effects likelihood
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcription polymerase chain reaction
SDS-PAGE	sodium dodecyl sulphate polyacrilamide gel electrophoresis
SLAC	single likelihood ancestor counting
SO4	Selektion Oppenheim 4
S.O.C	super optimal broth with catabolite repression
spp.	species
ss	single-stranded
SSCP	single-strand conformation polymorphism
TAE	tris acetate-EDTA
TBE	tris borate-EDTA
TMV	*Tobacco mosaic virus*

TPIB	tissue print immunoblotting
U	units
UTR	untranslated region
V	volt
Vs	versus
v/v	volume/volume
w/v	weight/volume
X-Gal	5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside
μl	microlitre
μm	micrometre
μM	micromollar

Chapter 1. Introduction and literature review

1.1. Viruses in general

Viruses are obligate intracellular parasites. They need the host cellular machinery to reproduce, since they lack the genetic information for conducting biological processes necessary to replicate. Virus particles are designated by virions and contain the genome enclosed by a coat protein known as capsid. Some viruses also have an envelope, a lipid component covering the virion surface. The viruses' genome can be linear or circular and encompass one or more nucleic acid molecules, either deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), which can be single-stranded (ss) or double-stranded (ds). Single-stranded genomes may be positive (+) or negative (-) sense, or ambisense, as they have the same or the complementary nucleotide sequence as the mRNA, or both, respectively. Viruses with negative sense RNA genome need to carry a RNA-dependent RNA-polymerase, since most of their hosts have no mechanism for RNA-dependent RNA polymerization. The majority of virus genomes are constituted by a single molecule of nucleic acid (monopartite) but in some cases the genome is segmented in two (bipartite) or more nucleic acid molecules, which occurs more frequently with RNA viruses. A segmented genome increases the possibility of genes to combine in different ways, thus increasing the evolution rates. Virus genomes can have a size from 3500 nt up to 2400000 nt. In the case of single-stranded RNA, the fragility and high probability of breaking as well as the higher mutation rates, contribute to limit the size of the viral genomes, which are usually found varying from 3.5 kb up to 30 kb (Gelderblom, 1996).

1.2. Plant Viruses

Plant virus species are determined by groups of properties recognized by the International Committee on Taxonomy of Viruses (ICTV). These properties include the plant species infected and the symptoms caused by the viruses, the way viral transmission occurs, the host species that can be used for experimental infection, the shape of virus particles, its nucleic acid and proteins, the genome composition and organization and also their serology. At the present there are 83 genera and 21 families of plant viruses recognized by the ICTV (ICTV, 2011).

Plant viruses' morphology may be rigid or flexuous rod-shaped (e.g. tobacco mosaic virus, TMV and potato virus Y, PVY, respectively), isometric (icosahedral) (e.g. cucumber mosaic virus, CMV) or bacilliform (bullet shape) (e.g. alfafa mosaic virus, AMV).

Most plant viruses have RNA genomes, the majority of which are single stranded and positive sense (Matthews, 1991). Besides the plant viruses with monopartite genomes, like the members of the *Closteroviridae* family, there are also plant viruses with bipartite and tripartite genomes, as the case of the *Comoviridae* and *Bromoviridae* families, respectively.

1.3. Economic Importance of Plant Viruses

Viruses are responsible for many of plant diseases that cause great losses in crops. These diseases can lead to yield and quality decrease in economically important crops, as a result of its symptoms which vary according to the disease. Generally, plant virus symptoms include lesions, necrosis (primary infection), leaf colour alteration and growth abnormalities (secondary or systemic infection). However, most symptoms can be common to several viruses and not all viruses cause visible symptoms.

Exact and up to date estimates related to damages caused by plant viruses to a given crop around the world are usually not available, but there are references to losses of billions of dollars per year in general terms (Gray and Banerjee, 1999). The costs associated with plant viruses, result not only from crop losses but also from

implementation and application of sanitary measures that need to be considered in order to prevent or control outbreaks, such as propagation material certification programs, eradication programs or vector control.

1.4. Virus Transmission in Plants

Plant viruses can be transmitted between different hosts through several routes, including seeds, pollen, propagation material, vectors and mechanical transmission.

Last century research suggested that approximately 20% of plant viruses are transmitted by seed (Matthews, 1991; Mink, 1993). The infection can occur during or after the fertilization of the embryo. In the first case the virus can invade the ovule or be transported by the male gamete and in the second case the virus invade the embryo during its development (Mink, 1993; Maule and Wang, 1996). Transmission by seeds usually leads to unhealthy new crops.

Viruses transmitted by pollen can infect the gametes of the embryo or the mother plant through the fecundated flower (Hull, 2009). This way of transmission seems to be unusual and only a few viruses are transmitted by pollen (Mink, 1993).

Vegetative propagation is a very common horticultural practice, especially with perennial species and therefore constitutes a very effective way of viruses' transmission. Once the parts of a systematically infected plant are used for propagation, normally the virus will be introduced into the new host (Hull, 2009). Grafting is an example of vegetative propagation. In this technique tissues from a plant (scion) are inserted on the roots of another (rootstock) and if one or both is infected and both are susceptible, the grafted plant will also become infected.

Viruses can also be disseminated from plant to plant by living organisms that feed of them and carry virus particles from infected plants to healthy ones, acting as vectors. This is the most common way of spreading of a virus in nature (Andret-Link and Fuchs, 2005). The vectors of plant viruses include arthropods (insects and arachnids), nematodes, fungi and plasmodiophorids (Andret-Link and Fuchs, 2005). Insect vectors of plant viruses include members of the order Hemiptera, such as aphids, leafhoppers, planthoppers, whiteflies, mealybugs, mirids and some treehoppers, and also members from other orders like beetles (Coleoptera), thrips (Thysanoptera),

grasshoppers (Orthoptera) and mites (Acari). Nematodes responsible for plant virus transmission belong to genera Longidorus, Trichodorus, Paralongidorus, Paratrichodorus and Xiphinema. Fungi and plasmodiophorids vectors are included in the orders Plasmodiophorales (Polymyxa spp. and Spongospora spp,) and Chytridales (Olpidium spp) (Agrios, 2005). Because aphids transmit 55% of plant viruses, they are the most important vectors. The remaining virus species are either transmitted by leafhoppers and beetles (11% each), whiteflies (9%), nematodes (7%) fungi and plasmodiophorids (5%), and thrips, mites, mirids and mealybugs (2%) (Andret-Link and Fuchs, 2005), or have no known vector.

According to the way of viruses' transmission by insect vectors, those are classified as non-persistent, semi-persistent and persistent. Non-persistent viruses remain in the stylet of the insect after acquisition and transmission can occur within a few minutes. Semi-persistent viruses are acquired in about fifteen minutes, enter the vector foregut and can be transmitted for up to two days. Persistent transmission involves the passage of the viruses throughout the gut into the haemocoel and then to the salivary glands. Hence, this type of transmission requires a latency or incubation period between acquisition and inoculation of the virus, which can remain in the vector for its lifetime. Non- and semi-persistent viruses, that are not able to replicate in the insect, are known as non-circulative. Persistent viruses are circulative or propagative, since they can circulate and replicate in the body of the insect (Katis *et al.*, 2007).

1.5. Virus Diseases of Grapevine

Grapevine (*Vitis vinifera* L.) is one of the oldest and most economically important cultivated crops. In 2010, the total vineyard cultivated area worldwide was estimated at 7.6 million hectares and the grape production was estimated to be 67950000 tons (OIV, 2012). Grapevine crops can be affected by several diseases, which cause huge economic losses, including those from viruses. Up to date, there are about 70 infectious agents, encompassing virus, viroids and pytoplasmas, recognized by the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG, 2003). However, there are four major virus diseases affecting grapevine: infectious degeneration/decline, rugose wood complex, fleck and leafroll.

1.5.1. Leafroll Disease

Grapevine leafroll disease (GLRD) was first reported in California (Goheen *et al.*, 1958) and nowadays is distributed in all grape-producing regions around the world (Martelli and Boudon-Padieu, 2006).

1.5.1.1. Economic Impact

GLRD economic impact causes average yield losses of 15-20% (Goheen, 1988, Martelli and Boudon-Padieu, 2006) but can achieve values up to 68% (Walter and Martelli, 1997). During a vineyard lifetime, GLRD can cause losses ranging from $25407/ha, considering a 30% yield reduction, to $41000/ha when considered a 50% yield reduction and a 10% penalty for low fruit quality (Atallah *et al.*, 2012). Comparable data is lacking in European countries.

1.5.1.2. Symptomatology

GLRD is latent in American *Vitis* species and their hybrids used as rootstocks but symptomatic in European *V. vinifera* (Uyemoto *et al.*, 2009). The disease affects rooting ability, graft take and plant vigour. The red and white varieties of *V. vinifera* show, respectively, a purple-red or yellow discoloration of the leaf blade with green primary veins that starts to appear at the basal leaves in late spring or early summer and progress through the shoot, not affecting the tip leaves, until leaf drop. In advanced stages of the disease the main veins can also lose their green colour. In some cultivars the disease can cause mesophyll necrosis, noticeable between the veins. Usually, in addition to chlorosis, leaves show downward rolling of the margins and thickening of the blade (Weber *et al.*, 1993, Martelli, 1993). Bunches are smaller and mature irregularly and later than usual, thus affecting fruit ripening. GLRD symptoms also include reduced sugar content of the berries, low Brix in the fruit juice (one degree Brix is 1 gram of sucrose in 100 grams of solution and represents the strength of the solution as percentage by weight (% w/w), reduced wine colour, phloem disruption, reduction of protein content, changes in the pattern of peroxidase and polyphenoloxidase

isoenzymes, potassium depletion and composition as well as aromatic profile of the musts (Martelli and Boudon-Padieu, 2006).

Figure 1.1 *Vitis vinifera* red wine grape variety Monvedro with symptoms of leafroll disease. Photo kindly provided by Dr. Margarida Teixeira Santos.

Figure 1.2 *Vitis vinifera* white wine grape variety Estreito Macio with symptoms of leafroll disease. Photo kindly provided by Dr. Margarida Teixeira Santos.

1.5.1.3. Implicated Viruses

GLRD is one of the most complex viral plant diseases, given the high number of causal agents associated with it. Leafroll is caused by a complex of viruses, named Grapevine leafroll-associated viruses (GLRaVs). All GLRaVs are phloem-restricted filamentous, positive sense single-stranded RNA virus from the family *Closteroviridae*.

The *Closteroviridae* family was established in 1998 (Martelli *et al.*, 2000) and include flexuous filamentous plant virus with either monopartide or bipartide positive sense single-stranded RNA genomes. The genome members of this family have their organized into two modules. The first includes a papain-like leader proteinase (L-Pro) at the 5' end and two ORFs associated with the replication, one comprising a methyltransferase (MET) and a helicase (HEL) and the other a RNA-dependent RNA-polymerase (POL) (Dolja *et al.*, 2006). The second module is a quintuple closterovirus-specific gene block associated with the assembly and movement that codes for a ~6 kDa hydrophobic protein (p6), a 70 kDa heat shock protein homologue (HSP70h), a HSP90-like protein of 50-60 kDa, a major coat protein (CP) and a minor CP (CPm) (Dolja *et al.*, 1994; Dolja *et al.*, 2006; Karasev, 2000).

When it was first established, this family comprised only the genera *Closterovirus* and *Crinivirus* distinguished by having a monopartide and bipartide genome, respectively. In 2000 it was suggested the creation of a new genus as well the classification by type of insect vector instead the number of genomic RNAs (Karasev, 2000). The new genus was first named *Vinivirus* and then *Ampelovirus* by decision of the ICTV Study Group on Closteroviruses and Allied Viruses and the revision of the *Closteroviridae* family was accepted in 2002 (Martelli *et al.*, 2002).

Until 2009, nine serologically different viruses associated with the leafroll disease had been identified and named sequentially GLRaV-1, -2, -3, -4, -5, -6, -7, -8, and -9 according to the order of their discovery. The majority of GLRaVs is classified as an approved or putative species of the genus *Ampelovirus* subgroup I (GLRaV-1 and -3) and subgroup II (GLRaV-4, -5, -6, -9). GLRaV-2 is the only species assigned to the genus *Closterovirus* (Martelli *et al.*, 2002). Since 2009 however, GLRaV-8 is no longer considered a valid species, after its only known sequence (partial CP gene sequence comprising 273 nt) had been found to be part of the grapevine genome (Bertsch *et al.*, 2009). GLRaV-7 has so far remained unassigned to any genus. However, new findings

support the creation of a fourth genus within the *Closteroviridae* family, provisionally designated *Velarivirus*, in which GLRaV-7 would be included (Al Rwahnih *et al.*, 2012, Martelli *et al.*, 2012). The proposed taxonomic alteration is being examined by the various bodies of the ICTV and after it should be ratified by the ICTV plenum. Also in the last years, three other ampeloviruses of subgroup II were reported: GLRaV-10 (or –Pr), GLRaV-11 (or –De), isolated from the greek grapevine varieties Prevezaniko and Debina, respectively (Maliogka *et al.*, 2008, Martelli *et al.*, 2012), and GLRaV-Car, isolated from the *V. vinifera* cultivar Carnelian (Abou Ghanem-Sabanadzovic *et al.*, 2010, Martelli *et al.*, 2012).

In advance to the 17th Congress of the ICVG (October 2012 at UC Davis, California) it was proposed to consider GLRaV-4 as a reference species and GLRaV-5, -6, -9, -Pr, -De and -Car as different strains of GLRaV-4 (Martelli, 2012). However, this taxonomic approach remains yet to be approved by the ICTV.

1.5.1.4. Cytopathology

Since the works of Rowhani and Golino (1995) and Monis and Bestwick (1996) an uneven distribution of GLRaVs in the host tissues has been amply documented. This is related to the fact these viruses are phloem-restricted and spread along the plant passing through the sieve plates of the sieve tubes, mainly during the growing season. Given its location, GLRaVs induce alterations in phloem cellular elements, sieve tubes, companion cells and phloem parenchyma cells, leading to necrotic destruction and hyperplastic proliferation of sieve tubes (Faoro *et al.*, 1992). These viruses are characterized by forming intracellular inclusions in phloem cells. The virus particles aggregate in single or clustered membranous vesicles, which originate from proliferation of peripheral membrane of mitochondria resealed to cytoplasm, as occurs with GLRaV-1, -3 and -5, or from vesiculation of the endoplamastic reticulum, as the case of GLRaV-2 and GLRaV-7 (Martelli and Boudon-Padieu, 2006).

1.5.1.5. Epidemiology

All GLRaVs are transmitted by grafting (from the rootstock to the scion or vice versa) and vegetative propagation, both being responsible for medium and long distance dissemination of the disease (Martelli, 1993). Mechanical transmission of leafroll through infected pruning equipment or harvesters has not been reported. Transmission by seed, although already verified for other grapevine viruses, is considered as not occurring in the case of leafroll associated viruses. GLRaV-2 is the only species with mechanically transmissible variants to herbaceous hosts, albeit inefficiently (Goszczynsky *et al.*, 1996).

Until the 1980s, leafroll disease was thought to be spread only through infected material propagation (Golino, 1993). However, in the 1960s the possibility of leafroll transmission by vectors had already been referred (Hewitt, 1968). In 1983, for the first time, grapevine viruses were shown to be transmissible by mealybugs (Rosciglione *et al.*, 1983). Since then, growers and researchers verified that GLRD spreads within vineyards over the years, forming aggregates in the distribution patterns of infected plants, thus suggesting insect transmission (Engelbrecht and Kasdorf, 1990b; Habili *et al.*, 1995). In 1990, Engelbrecht and Kasdorf, reported the first evidence of GLRaVs transmission by insects. Subsequently, several reports of transmission of GLRaVs on the field (Engelbrecht and Kasdorf, 1990a), in laboratory (Rosciglione and Gugerli, 1989) or greenhouses (Tanne *et al.*, 1989) by mealybugs and soft scales have been shown.

Mealybugs and soft scales are small, phloem-feeding insects from the order Hemiptera and families *Pseudococcidae* and *Coccidae*, respectively. These insects can cause severe damage in plants either by weakening them when feeding of its nutrients, or by excreting honeydew suitable for the proliferation of sooty mould fungi which reduces photosynthesis, or by transmitting viruses while they feed. GLRaVs seem to be transmitted by these insects in a semipersistent manner (Tsai *et al.*, 2008).

Although transmission of GLRaVs by vectors is still poorly understood, there are reports of GLRaV-1, -3, -4, -5 and -9 being vectored by mealybug and soft scale species. GLRaV-1 has been showed to be transmitted by the mealybugs *Heliococcus bohemicus* and *Phenacoccus aceris* and the soft scales *Parthenolecanium corni*, *Neopulvinaria innumerabilis* and *Pulvinaria vitis* (Belli *et al.*, 1994; Fortusini *et al.*, 1997; Sforza *et al.*, 2003). GLRaV-3 have been reported to be spread by the mealybugs

Heliococcus bohemicus, *Phenacoccus aceris*, *Planococcus citri*, *P. ficus*, *Pseudococcus affinis*, *Ps. calceolariae*, *Ps. comstocki*, *Ps. longispinus*, *Ps. maritimus*, *Ps. viburni*, and the soft scales *Neopulvinaria innumerabilis* and *Pulvinaria vitis* (Belli *et al.*, 1994; Cabaleiro and Segura, 1997; Engelbrecht and Kasdorf, 1990a; Garau *et al.*, 1995; Golino *et al.*, 2002; Petersen and Charles, 1997; Rosciglione *et al.*, 1983; Sforza *et al.*, 2003). Up to date, transmission of GLRaV-4, -5 and -9 only has been demonstrated experimentally. GLRaV-4 has been shown to be transmitted by *Planococcus ficus* (Tsai et al., 2010), GLRaV-5 by *Planococcus ficus* (Mahfoudhi *et al.*, 2009; Tsai et al., 2010), *Pseudococcus longispinus* (Golino et al., 2002; Sim et al., 2003) and *Ceroplastes rusci* (Mahfoudhi *et al.*, 2009) and GLRaV-9 by *Planococcus ficus* (Tsai *et al.*, 2010) and *Pseudococcus longispinus* (Sim *et al.*, 2003; Tsai *et al.*, 2010). GLRaV-2 has no known vector, but several members of the genus *Closterovirus* are transmitted by aphids.

1.5.1.6. Certification Scheme of Vegetative Propagation Material

It was already referred that introduction and indiscriminate use of infected propagation material is one of the most important route of grapevine virus dissemination throughout the world. To use certified material is one effective way to prevent and control those agents responsible for the losses in crops yield and quality. With that purpose certification schemes for grapevine material produced with intent to be propagated or sold were established worldwide. To be considered as certified planting material, plants have to be rigorously tested and found free of specified pathogens. Afterwards plants have to be maintained and propagated for one or several stages under strict and different approved conditions to ensure health standards and avoid recontaminations. Because of the uneven distribution of GLRaVs, the standardized protocol recommends each sample should include four to six branches, randomly collected from different sides of the vine to be tested (Rowhani and Golino, 1995). Also, although GLRaVs antigens are usually detected throughout the year (except early in the growing season), samples near the bottom portion of actively growing stems and petioles usually have the highest concentration of virus (Monis and Bestwick, 1996), reinforcing that multiple samples should be taken from this tissues to unambiguously detect GLRaVs.

In Europe, a certification scheme has been implemented by the Council Directive 68/193/CEE (9 April 1968, last amended by Directive 2002/11/EC) on the marketing of material for the vegetative propagation of the vine. According to this directive, propagation material should be free of GFLV and ArMV from the complex of infectious degeneration, GLRaV-1 and GLRaV-3 from the grapevine leafroll disease and GFkV (only for rootstocks). Finland and Sweden were released from the obligation to apply the Council Directive 68/193/EEC as stated in the Commission Decision 2005/931/EC (21 December 2005).

1.6. Diagnostic Techniques of Plant Viruses

Diagnostic tools for detection of plant viruses have evolved through the years. Plant viruses can be detected by several diagnostic methods, ranging from visual observation to molecular detection. Primary detection can be done by symptoms inspection or using electron microscopy to visualize virus particles. Routine diagnosis may be accomplished through serological and molecular detection. Further characterization of isolated virus strains can be accomplished through biological indexing on indicator grapevine cultivars.

1.6.1. Visual Detection and Biological Indexing

A viral infection is possible to identify in the field by recognizing some typical visible symptoms. However, this detection method does not identify the specific infectious agent within a complex aetiology. In fact, the majority of symptoms are not exclusive of one virus and some viruses may not induce visible symptoms. Moreover, plants can express a variety of symptoms, when infected by different strains of a single virus, or similar symptoms caused by different viruses. Thus, posterior identification of the disease viral agent, by other methods, is always required.

Electron microscopy allows direct observation of the virus particles by examining a crude plant extract. The particles morphology may be used to identify the family or genus to which the viral agent belongs to, but not the virus species or strain. In

addition, this technique may allow viruses that are present in low concentrations to pass unnoticed.

Biological indexing is a diagnostic tool used in detection and characterization of causal agents of grapevine virus diseases (Martelli, 1993). It is a technique based in the observation of the presence or absence of symptoms in herbaceous or woody indicator plants after their inoculation (herbaceous) or grafting (woody) with a plant that is being tested (Moore *et al.*, 1948). For each viral disease of the grapevine a suitable indicator plant has been south and identified. Herbaceous indexing is developed in greenhouses and involves inoculation through mechanical transmission by scrubbing a crude extract sample of the plant being tested on the leaves of indicator plants that are sensitive and systematically develop an array of specifically described symptoms. Woody or field indexing involves bud-grafting the indicator plants with material from the grapevine being tested. While herbaceous indexing takes several weeks, woody indexing takes at least two growing seasons. Even though biological indexing requires labour and time it is still very much used, and is the technique applied in certification schemes of propagation material. In the case of grapevine, it is also used for selection of clones to be commercialized. Nevertheless, given that only the disease symptoms are observed, this technique allows the detection of a disease but does not identify the causal agent.

1.6.2. Serological Tests

Serological methods are based on reactions between polyclonal (PAb) or monoclonal (MAb) antibodies and certain antigenic structures of grapevine viruses. Given the high specificity of the antibody-antigen reaction, these can be developed to a highly selective degree.

Enzyme-linked Immunosorbent Assay (ELISA) and its Double Antibody Sandwich form (DAS-ELISA) are the most commonly used serological tests (Engvall and Perlman, 1971). In the later the antibodies are first immobilized onto the microplate wells surface (primary antibody) and then a crude extract of the plant being tested is placed into each well in order to allow the binding of the virus proteins to antibodies. The viral particles can be detected with enzyme-labelled antibodies (second antibody),

that convert an unstained substrate into a coloured product, which absorbance can be measured with a spectrophotometer.

ELISA and DAS-ELISA allow for simultaneously testing of many plants for one virus and are often favoured for being fast, relatively inexpensive and not technically challenging. However, they require the existence of antibodies for each virus of interest, suitable to detect all different strains, without reacting with plant proteins. Often some commercial antisera are raised against a well-known, widespread virus strain, which depending on the diversity of the virus, may lead to false negatives by overlooking plants infected with less common strains. False positives are also a possibility especially when using polyclonal antibodies rose against a plant sample unsuspectedly infected with more than one virus. In addition, because a crude plant extract is used, which may show absorbance at the same peak as the ELISA chromogenic substrates, the technique lacks in sensitivity, and may fail to detect viruses present in low concentrations.

Western Blotting is a technique which first involves the separation of plant proteins from an extract by SDS-PAGE. Next the proteins are transferred to a nylon, nitrocellulose or polyvinylidene difluoride (PVDF) membrane, and the presence of viral proteins is detected with enzyme-labelled antibodies (Towbin *et al.*, 1979).

Tissue Print Immunoblotting (TPIB) is another diagnostic technique that uses anti-virus IgG. It consists in pressing the surfaces of freshly cut plant tissues (leaves, veins, petioles, roots) onto a nylon or nitrocellulose membrane. Virus particles adsorb to the membranes and are afterwards detected with enzyme-labelled antibodies that are able to convert a stainless soluble substrate into a stained insoluble product (Cassab and Varner, 1987). This procedure is simple, fast and inexpensive, allowing testing hundreds of samples per day or even collecting samples in the field to be processed later on.

The In Situ Immunoassay (ISIA) technique, uses extremely thin plant tissue sections (e.g. from veins or petioles), obtained from fresh tissues with a cryostat or with a microtome after resin inclusion, immobilized onto a microscope slide. Antibodies labelled with enzymes or fluorescent dyes identify the virus presence in each section by producing a coloured signal in the precise location of virus particles (Lin *et al.*, 2000), visualized with a microscope.

1.6.3. Molecular Detection and Typification

Molecular testing for plant diseases routine diagnostic has become increasingly popular. These obviously require a priori knowledge of the genome or target gene sequences. When RNA viruses are being tested for, it is necessary to synthesize cDNA by a Reverse Transcriptase prior to the amplification of dsDNA (RT-PCR). Hence RT-PCR and PCR (Saiki *et al.*, 1988) are the most widely used molecular tools for detecting RNA and DNA plant viruses, respectively. These methods involve the selective amplification of a specific virus genome fragment, and enable detection even when the virus is present in very small amounts, with much higher sensitivity that the majority of the serological techniques. Therefore a molecular result tends to be considered as the ultimate evidence of the presence or absence of the virus. Thus, sampling for this purpose is a very critical step in the molecular detection procedure. Representative samples collected at the right phonological stage are vital in determining the validity of the results. Even when an adequate sampling protocol is adopted, however, sample-handling conditions are paramount to obtain reliable results, especially for RNA viruses. In fact the quality of the extraction step is one of the major pitfalls of using a molecular detection method. Subsequent technical aspects of obtaining dsDNA from an RNA or DNA target involve the choice of a DNA polymerase capable of 'proofreading' such as the *Pfu* DNA polymerase, which possesses 3' to 5' exonuclease proofreading activity, and corrects nucleotide-misincorporation errors. *Pfu* DNA polymerase-generated PCR fragments are thus blunt-ended and expected to contain fewer errors than *Taq*-generated PCR inserts. As a result, *Pfu* is more commonly used for molecular cloning of PCR fragments.

Multiplex RT-PCR allows the detection of several viruses in a single PCR reaction, using multiple primer sets to amplify different length amplicons. It requires annealing temperatures adjustment but allows obtaining results in a single run.

RT-PCR or PCR followed by Restriction Fragment Length Polymorphism (RFLP) is a typification method that can be used to identify different variants of the viruses present in a given sample. It requires previous cloning and sequencing of each virus variant to determine the respective restriction profile with a chosen array of endonucleases.

The use of Real-Time PCR (Syvanen *et al.*, 1988) can allow for simultaneous detection and typification of specific sequences. This technique is based in the measurement of fluorescence given by a reporter molecule, which anneals specifically to the target sequence, increasing as the reaction proceeds. Real-Time PCR is more sensitive and rapid than traditional PCR, since it does not require end-point analysis with visualization of final amplification products in agarose gels nor further typification protocols.

Microarray analysis is a method that allows the identification of several different nucleotide sequences at the same time (Schena *et al.*, 1995). These target sequences are labelled with fluorescent dyes and hybridized to probe sequences placed on a solid surface. In the end, specific laser scanners and image analysis software are used to determine the occurrence of probe-target hybridization. Theoretically, this technique permits the detection of several virus sequences in a unique plant; however its application in plant virology is still very limited.

1.7. Aim of the Work

The most effective way to control grapevine diseases is to prevent its appearance by strict observation of certification schemes, i.e. propagation and commercialization of virus-free plant material. To this end, it is extremely important to have efficient detection and diagnostic tools that distinctively identify each virus associated with a given disease. In order to develop these tools, which will increasingly rely on molecular techniques, it is necessary to gather information on way of transmission, genome arrangement and population genetic structure of the viruses involved. In agreement with the European directives, propagation material for viticulture industry must be free of ArMV, GFLV, GFkV, GLRaV-1 and GLRaV-3. The complex aetiology of leafroll disease and its economic impact justify the investment not only in the molecular characterization of GLRaV-1 and -3 but also of all of the known grapevine leafroll-associated viruses. In fact few studies have tackled the phylodynamics of GLRaV-1 and -3, and the other GLRaVs have been considered more as a scientific curiosity in terms of evolution of phylogenetic lineages. Reasons for the later may well be linked to the

fact that, although some molecular evidence for other GLRaVs had been previously reported, their complete genome sequences were obtained only in the last few years. This situation has hitherto precluded the set up of adequate molecular detection assays and the subsequent analysis of the genetic structure of natural populations. A widespread notion that GLRaV-3 is far more frequent than GLRaV-1 and that all other GLRaVs are much less common and of less importance to the GLRD scenario has resulted. Given the present state of knowledge, this work aims to contribute to improve the understanding and diagnosis of leafroll disease and implicated viruses by focusing on two grapevine leafroll-associated viruses - GLRaV-1 and GLRaV-5, in Portuguese grapevine cultivars. Although a quarantine virus, sparse molecular information has been gathered on GLRaV-1 especially regarding the CP gene, in which routine detection methods are based. GLRaV-5 in turn, even though not currently listed in the European certification scheme, has been increasingly reported over the last three years in several grapevine-growing countries. This might indicate that GLRaV-5 is emerging or that it has been overlooked due to lack of implementation of a specific routine detection assay. In this scenario, this work will describe the genetic diversity and population genetic structure of both viruses and, in the case of GLRaV-1, use the collected information will be used to develop antibodies able to detect the array of coat protein variants known for this virus.

Chapter 2. *Grapevine Leafroll-associated Virus 1*

2.1. Introduction

GLRaV-1 has a positive sense single-stranded RNA genome of approximately 19.5 kb organized into 10 ORFs (Fig. 2.1). ORF1a encodes a helicase (HEL) that possibly is part of a larger protein at 5' -end of the genome. ORF1b encodes a RNA-dependent RNA polymerase (RdRp), which overlaps with ORF1a, followed by a non-coding region of 793 nt and ORF2 that encodes for a small endoplasmatic reticulum associated hydrophobic protein (p7). The products of ORF3 and ORF4 are, respectively, a heat shock protein 70 homologue, HSP70h, and a protein with few similarities to the HSP90h of closteroviruses. Next are the ORFs which encode in the 5'-3' direction, the coat protein (CP), two minor copies of CP (CPm1 and CPm2) and two other proteins (p22 and p24) of unknown function (Fazeli and Rezaian, 2000). Unlike other GLRaVs, GLRaV-1 has two diverged copies of the CP. Therefore, the five-gene module characteristic of closteroviruses is present in GLRaV-1.

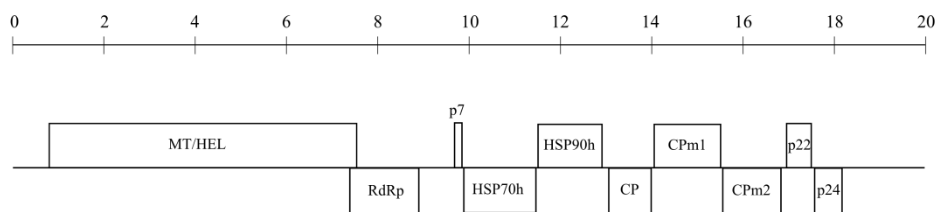

Figure 2.1 Schematic representation of the GLRaV-1 genome organization. Boxes represent the 10 ORFs: MT-methyltransfrease; HEL-helicase; RdRp-RNA-dependent RNA polymerase; p7-7 kDa endoplasmatic reticulum associated hydrophobic protein; HSP70h-heat shock protein 70 homologue; HSP90h-heat shock protein 90 homologue; CP-coat protein; CPm1 and CPm2-minor copies of CP; p22-22 kDa protein; p24-24 kDa protein. The above scale is graduated in kb.

Even thought GLRaV-1 is included in certification schemes of grapevine planting material, very little information is available regarding its molecular variability. Fazeli *et al.* (2000, 2001) were the first to sequence the genome (coding regions) of

GLRaV-1 and to analyze the nucleotide sequences of 10 open reading frames, revealing the presence of higher degree of sequence variation in ORFs 3, 6 and 7 encoding a homologue of heat shock protein 70 and two diverged copies of the coat protein (CPm1 and CPm2), respectively. However, up to date, only two studies on the genetic structure of natural populations of GLRaV-1 were published. One characterized a 540 nt fragment of the HSP70h gene (Komínek *et al.*, 2005) from grape cultivars of Czech Republic and the other analyzed the genetic diversity of partial sequences from the HSP70h, CP, CPm2 and p24 genes (Alabi *et al.*, 2011), in American grapevines from the states of California, Washington and New York. In fact, only 8 complete sequences (969 nt) of the coat protein gene are currently available at GenBank in a total of 41 sequences retrieved from 28 isolates collected from 10 countries. Noticeably, further characterization of the molecular diversity and ecology of the virus from other grapevine growing regions is necessary. In order to obtain data suitable to improve the serological and molecular detection methods, a better description of the diversity of the coat protein, the target for routine serological tests, is badly needed. With that purpose a total of 20 field isolates of GLRaV-1 from 13 grapevine varieties were analyzed in this work. Eighty-three new sequences of the complete CP gene were obtained, and combined with the information at GenBank. A dataset of 120 CP sequences was assembled and used to develop new antibodies for GLRaV-1 diagnostic.

2.2. Material and Methods

2.2.1. Sample Collection and Plant Material

Selection of plants to be analyzed in this study was resultant from an ongoing screening of GLRD symptomatic varieties, using cross information from both serological and molecular detection assays.

Sixteen of the twenty *Grapevine leafroll-associated virus 1* isolates analyzed in this work were collected from *V. vinifera* grapevines maintained in the National Ampelographic Collection (CAN) from the INRB/INIAV situated at Dois Portos, Portugal (Table 2.1). This collection was established in the field 30 years ago (Teixeira Santos *et al.*, 2009b), and a “mother plant” from each grapevine variety was multiplied by grafting onto certified rootstock material (SO4, clone 73). Hence, each variety is at present represented by seven clone plants. The plants from the collection are maintained in phytosanitary controlled conditions, free from nematodes and mealybugs, and are regularly tested for the presence of GLRaV-1, -2, -3, -5, -6 and -7 among other grapevine viruses by DAS-ELISA using commercially available antisera or by RT-PCR. The other four isolates were collected from grapevines established in identical phytosanitary conditions at Quinta do Marquês (INRB/INIAV), in Oeiras, Portugal. Three of these four isolates were grafted onto certified rootstock material (R99) and are maintained in the field. The fourth originated from an ungrafted grapevine cultivar maintained in the greenhouse. In the isolates accession, the numbers in brackets refer to the number of the variety’s clone plant from which the isolate was obtained.

Samples consisted of different plant tissue, either cortical scrapings from dormant canes collected in the fall after harvesting season, or petioles and veins from leaves collected in late spring or early summer, when symptoms are fully visible. To account the possibility of uneven virus distribution throughout the plant, randomly collected samples from a minimum of five dormant canes or leaves from six branches were combined.

Table 2.1 Details of *Grapevine leafroll-associated virus 1* isolates obtained in this study

Source / Location	Isolate ID	Other GLRaVs detected[a]	Genomic region
cv. Alfrocheiro, red wine grape; Fresh petioles and veins	A0305	GLRaV-3	CP
Oeiras (INRB/INIAV) - Quinta do Marquês (PRT010)			
cv. Brancelho, red wine grape; Phloem scrapings	B3	GLRaV-2, -3	CP
Dois Portos (INRB/INIAV) - CAN (PRT051)			
cv. Coração de Galo, red table grape; Fresh petioles and veins	CdG	GLRaV-3	
Oeiras (INRB/INIAV) - Quinta do Marquês (PRT010)			
cv. Estreito Macio, white wine grape; Phloem scrapings	EM3	GLRaV-3	HSP70h
Dois Portos (INRB/INIAV) - CAN (PRT051)			CP
cv. Gouveio Real, white wine grape; Phloem scrapings	GR3	GLRaV-3, -5	HSP70h
Dois Portos (INRB/INIAV) - CAN (PRT051)			CP
cv. Jacquez, red wine grape; Phloem scrapings	J2	GLRaV-3	HSP70h
Dois Portos (INRB/INIAV) - CAN (PRT051)			CP
	J3	GLRaV-3	CP
	J5	GLRaV-3	CP
	J6	GLRaV-3	CP
cv. Monvedro, red wine grape; Clone plants; Phloem scrapings	M2	GLRaV-3	CP
Dois Portos (INRB/INIAV) - CAN (PRT051)	M4	GLRaV-3	CP
	M7	GLRaV-3	CP
cv. Olho de Pargo, red table grape; Clone plants; Fresh petioles and veins	OdP2	GLRaV-3	CP
Oeiras (INRB/INIAV) - Quinta do Marquês (PRT010)	OdP5	GLRaV-3	CP

Source / Location	Isolate ID	Other GLRaVs detected[a]	Genomic region
cv. Pical-Polho, red wine grape; Phloem scrapings Dois Portos (INRB/INIAV) - CAN (PRT051)	PP1	GLRaV-3	CP
cv. São Saúl, red wine grape; Phloem scrapings Dois Portos (INRB/INIAV) - CAN (PRT051)	SS3	GLRaV-2, -3	HSP70h CP
cv. Sousão Vinhos Verdes, red wine grape; Phloem scrapings Dois Portos (INRB/INIAV) - CAN (PRT051)	SVV1	GLRaV-3	HSP70h CP
cv. Tinta Lameira, red wine grape; Phloem scrapings Dois Portos (INRB/INIAV) - CAN (PRT051)	TL3	GLRaV-3	HSP70h CP
cv. Vinhão Douro, red wine grape; Clone plants; Phloem scrapings Dois Portos (INRB/INIAV) - CAN (PRT051)	VD4 VD7	GLRaV-3 GLRaV-3	CP CP

[a] Molecular detection with CP gene virus-specific primers after cDNA synthesis (Esteves *et al.*, 2009a, Esteves *et al.*, 2009b, , Esteves *et al.*, 2012, Teixeira Santos *et al.*, 2009a).

2.2.2. RNA Extraction

Total RNA was extracted from samples with the E.Z.N.A.™ Plant RNA Mini Kit (Omega Bio-Tek, USA) according to manufacturer's Protocol II for difficult samples with a few modifications as described by MacKenzie *et al.*, 1997. Plant material (0.25 g in the case of cortical scrapings and 0.7 g in the case of petioles and veins) was macerated in a cooled sterile mortar with liquid nitrogen. The powder was homogenised with RNA extraction buffer [4 M guanidine thiocyanate, 0.2 M sodium acetate, pH 5.0, 25 mM EDTA, 2.5% (w/v) PVP-40 and 1% (v/v) β-mercaptoethanol (added just before use)], which was added in a 1:10 or 1:3 proportion (w/v), when using cortical scrapings or petioles/veins, respectively. The sample homogenate (circa 1.4 ml) was transferred to a 1.5 ml microcentrifuge tube and mixed with 100 µl of 20% (w/v) sodium lauryl sarkosyl. After incubation at 70 ºC for 10 min with sporadic agitation, samples were centrifuged at 10,000 x g for 10 min at room temperature. From this point on, manufacturer's instructions were followed.

All RNA samples were treated with DNase with the TURBO DNA-free™ kit (Ambion, USA) according to the manufacturer's protocol.

2.2.3. cDNA Synthesis and PCR

The first-strand cDNA was synthesized in a reaction final volume of 20 µl with *iScript* cDNA First Strand Synthesis Kit™ (Bio-Rad, USA). A mixture of 5 µl of total RNA, 2 µl of random hexamer primers and 8 µl of nuclease-free water was incubated at 65 ºC for 5 min for sample denaturation, cooled on ice and mixed with 4 µl of 5X Reaction Buffer and 1 µl of *iScript* Reverse-Transcriptase (1U/µl). The reverse transcription was carried out at 25 ºC for 5 min followed by 42 ºC for 30 min. The enzyme was inactivated at 85 ºC for 5 min and synthesized cDNA was conserved on ice before direct use or stored at -20 ºC.

Synthesized cDNA was used for PCR amplification of three different GLRaV-1 genome regions. A fragment of 540 bp of the HSP70h encoding gene was amplified using a primer pair reported by Komínek *et al.* (2005). One set of primers reported previously (Esteves *et al.*, 2009a) and one set of primers designed in this study with Oligo – Primer Analysis Software, version 7.37, were used to amplify fragments of 1044 bp and 1021 bp, respectively, each encompassing the entire CP encoding gene (Table 2.2).

Table 2.2 Gene specific primers for *Grapevine leafroll-associated virus 1* used in this work.

Target gene	Primer name	Orientation	Sequence 5'→3'	Genome Position[a]	Product size (bp)	Reference
HSP70h	plazleft	sense	CAGGCGTCGTTTGTACTGTG	416 - 435	540	Komínek *et al.* (2005)
	plazright	antisense	TCGGACAGCGTTTAAGTTCC	936 - 955		
CP	CPLR1F	sense	TCAATAATACTGCGTGCTT	6853 – 6871	1044	Esteves *et al.* (2009a)
	CPLR1R	antisense	CTAACGCAGTCGCCATTGT	7878 - 7896		
	CP1-1	sense	CGTGCTTTTTACAGTATCGGW	6865 - 6885	1021	This work
	CP1-2	antisense	CCATTGTAAATCACTGCTGTC	7865 - 7885		

[a]Primers positions according to GenBank accession AF195822. W = A + T.

PCR amplifications were conducted in a total volume of 50 µl with sterile-filtered water (Sigma-Aldrich, USA), 1X *Pfu* Buffer with $MgSO_4$ [200 mM Tris-HCl (pH 8.8 at 25 ºC), 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 1 mg/ml BSA, 0.1% (v/v) Triton-X 100, 20 mM $MgSO_4$] (Thermo Scientific, USA), 200 µM dNTPs (Thermo Scientific, USA), 200 nM of each primer, 2.5 U of *Pfu* DNA Polymerase (Thermo Scientific, USA) and 5 µl of cDNA. PCR was performed in a Biometra thermocycler and cycling conditions consisted of an initial denaturation step at 95 ºC for 3 min, followed by 30 cycles of denaturation at 95 ºC for 30 sec, annealing at 48 ºC for 30 sec and extension at 72 ºC for 90 sec or 3 min, for HSP70 and CP fragments, respectively, and a final extension step at 72 ºC for 10 min.

The PCR products were analyzed under UV light after electrophoresis in 1.5% agarose gel in 1X TAE buffer (pH 8.3) stained with Greensafe (NZYTech, Portugal). The GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific, USA) was used as size marker.

2.2.4. Cloning and Transformation

PCR products were purified using NZYGelpure purification kit (NZYTech, Portugal) following the manufacturer's instructions and then ligated using CloneJET^{TM} PCR Cloning Kit (Thermo Scientific, USA) according to fabricant's protocol for blunt-end cloning. The products of ligation reactions were used directly to transform XL1-Blue E. *coli* competent cells (Stratagene, USA). For each transformation 10 µl of ligation reaction were added to 100 µl of competent cells and incubated on ice for 30 min. The cells were heat shocked for 45 sec in a water bath at 42 ºC and immediately placed on ice during 2 min. Afterwards 250 µl of S.O.C medium (LB supplemented with 20 mM glucose) were added to the cells and tubes were incubated at 37 ºC for 1 hour with shaking at 225-250 rpm. Five µl of 1 M IPTG were added to the cells, and 150 µl of the suspension were plated onto LB agar plates containing ampicillin (0.05 mg/ml) and 40 µl of X-Gal (20 mg/ml) previously spread. The plates were incubated at 37 ºC overnight.

2.2.5. Single-Strand Conformation Polymorphism (SSCP)

This analysis was conducted in order to select recombinant clones of interest for sequencing. For each isolate, recombinant clones were amplified by PCR with the specific primers pair (see 2.2.3), followed by electrophoresis to verify the clones of interest. Of these, a minimum of 16 recombinant clones per isolate were subsequently analyzed by SSCP (Orital *et al.*, 1989) in order to identify different conformational patterns. For the SSCP analysis, 1 µl of the PCR product was added to 9 µl of denaturing buffer (95% formamide, 20 mM EDTA, pH 8.0, bromophenol blue), denatured at 90 ºC for 5 min and placed on ice. Denatured DNA was analyzed by electrophoresis in 8% polyacrylamide gel in 1X TBE buffer conducted at 4 ºC and 200 V, during 2 or 3 hours, depending on the fragment size being analyzed.

After electrophoresis, polyacrylamide gels were stained with silver nitrate. The gels were placed in a solution of 10% glacial acetic acid during 20 min and washed 3 times for 3 min with distilled water. Next, the gels were passed into a 1 % nitric acid solution and washed again 3 times for 3 min, followed by staining in a silver nitrate solution (1g/L $AgNO_3$, 1.5 ml/L formaldehyde) during 30 min. The gels were then washed with distilled water for 3 min and revealed with a sodium carbonate solution (30 g/L Na_2CO_3, 1.5 ml/L formaldehyde, 2 mg/ml $NaS_2O_3.5H_2O$). The reaction was stopped placing the gels in the initial acetic acid solution.

The SSCP data were used to determine the heterozygosity level of each gene within each isolate with Nei's *h* coefficient (Nei, 1978):

$$h = n(1 - \sum x_i^2)\,/(n-1)$$

where *n* is the total number of clones analyzed for each isolate and *x* the percentage of clones with the same SSCP pattern. *h* values range from 0 to 1.

2.2.6. Plasmid DNA Extraction and Restriction Analysis

For each isolate, at least two clones per SSCP pattern observed were selected to purify the plasmid DNA. Five ml of LB medium with ampicillin (0.05 mg/ml) were inoculated with a single colony and incubated overnight at 37 ºC with shaking at 200

rpm. Plasmid DNA extraction from recombinant E. *coli* was then carried out with NZYMiniprep kit (NZYTech, Portugal) according to the manufacturer's protocol.

The presence of the insert of interest was determined by double restriction of the extracted plasmid DNA, performed with FastDigest® XbaI and XhoI (Thermo Scientific, USA). The restriction reaction was carried out in a final volume of 20 µl, including 14 µl of nuclease-free water (Sigma-Aldrich, USA), 2 µl of 10X FastDigest® Green buffer (Thermo Scientific, USA), 2 µl of plasmid DNA and 1 µl of each FastDigest® endonuclease. The mixture was incubated at 37 ºC during 10 min in a heat block.

The restriction products were analyzed under UV light after electrophoresis in 1.5% agarose gel in 1X TAE buffer (pH 8.3) stained with Greensafe (NZYTech, Portugal). The GeneRuler 100 bp Plus DNA Ladder (Thermo Scientific, USA) was used as size marker.

2.2.7. Sequencing and Sequence Data Analysis

The selected clones were commercially sequenced (CCMAR, UAlg, Portugal). Sequences were processed using the BioEdit Sequence Alignment Editor (Hall, 1999), screened in order to exclude repeated sequences, and aligned with ClustalW (Thompson *et al.*, 1994). A dataset of 106 sequences of the HSP70h gene fragment with 540 bp was constructed. This included 11 sequences obtained in this work from 6 isolates and 95 sequences retrieved from GenBank. Regarding the CP gene, a set of 124 sequences was assembled, with 83 complete sequences (969 bp) obtained in this work from 20 isolates and 41 sequences available at GenBank from which only 8 were complete. The AA multiple alignments graphic views and Hoop and Woods hydrophilicity (Hoop and Woods, 1981) profiles for each gene were carried out using BioEdit. Phylogenetic analysis was performed with MEGA5 (Tamura *et al.*, 2011) with the maximum likelihood method. The suitable substitution models were determined with the Bayesian Information Criterion (BIC) and bootstrap values were estimated with 1000 replicates.

2.2.8. Estimates of Genetic Distances and Selection Pressure

Analysis of genetic diversity among HSP70h and CP sequences variants of GLRaV-1 were determined with MEGA5 with specific substitution models.

Selection pressure analysis was conducted for both genes by estimating the rates of synonymous (dS) and non-synonymous (dN) substitutions using the algorithms for single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL) and internal fixed effects likelihood (IFEL) (Kosakovsky Pond and Frost, 2005, Kosakovsky *et al.*, 2006a) with statistical significance (p-value) of 0.1 and using the HKY85 nucleotide substitution bias model (Hasegawa *et al.*, 1985) implemented in the Datamonkey web server (Delport *et al.*, 2010). The ratio of non-synonymous to synonymous substitutions per site (dN/dS) was also determined. Values indicated the selection pressures acting on each gene: $dN/dS < 1$ indicate negative or purifying selection and $dN/dS > 1$ indicate positive selection.

2.2.9. Recombination Analysis

Potential recombination events in the two sequences datasets were screened with the genetic recombination detection method (GARD) (Kosakovsy Pond *et al*., 2006b) from Datamonkey web server and with the set of methods for recombination detection included in RDP3 Alpha44 software (Martin *et al.*, 2010).

2.2.10. Antibodies Development

The hydrophilicity profile of the coat protein was obtained from the alignment of the deduced amino acid sequences and subsequently used to identify conserved motifs within all the sequences compared. With this information, polyclonal antibodies were commercially obtained (Biogenes, Germany). A 15 AA residue was selected to synthesize a peptide, with which a goat was immunized. Afterwards, the monospecific IgG (MonoIgG) was commercially purified and conjugated with either alkaline phosphatase (MonoIgG-AP) or FITC (MonoIgG-FITC) (Biogenes, Germany).

2.2.11. SDS-PAGE and Western Blot Analysis

Two hundred mg of plant tissue from cortical scrapings or petioles and veins, were macerated with 5 volumes of loading buffer [100 mM Tris-HCl, pH 6.8, 2% (v/v) β-mercaptoethanol, 4% SDS, 20% glycerol, 0.2% bromophenol blue]. The homogenate was transferred into a 2 ml microcentrifuge tube. The extracts were boiled at 100 ºC for 5 min, centrifuged at 13,000 x g for 5 min at room temperature and the supernatant was collected into a new tube. Aliquots of 20 µl of each sample were loaded into a 5-12% gradient SDS-polyacrylamide gel and proteins were separated by electrophoresis under denaturing conditions in 1X Tris/Glycine/SDS buffer [25 mM Tris, pH 8.3, 192 mM glycine, 0.1% (w/v) SDS] at 35 mA per gel and room temperature for 1 hour. The PageRuler Prestained Protein Ladder (Thermo Scientific, USA) was used as size marker.

Gels were stained with Coomassie Brilliant Blue (Bio-Rad, USA) or electroblotted onto PVDF membranes (Millipore, USA) using Trans-Blot® Turbo™ Transfer System (Bio-Rad, USA) according with the fabricant specifications for transfer using traditional semi-dry consumables. After transfer, the membranes were blocked with TBST [20 mM Tris, pH 7.6, 137 mM NaCl, 0.1% (v/v) Tween-20] with 5% non-fat dry milk at 4 ºC overnight. The membranes were washed with TBST for 2 times during 2 min and incubated for 1 hour with polyclonal IgG primary antibody diluted 1:30 in TBST with 1% BSA. After 2 washes of 7 min each with TBST, membranes were incubated with alkaline phosphatase-labelled anti-goat secondary antibody (Sigma, USA) diluted 1:7500 in TBST with 1% BSA. The membranes were subsequently rinsed 3 times each for 7 min with TBST and 1 time for 5 min with alkaline phosphatase substrate buffer (0.1 M Tris-HCl, pH 9.5, 100 mM NaCl, 5 mM $MgCl_2$). Membranes were incubated with freshly prepared NBT-BCIP (Roche, Switzerland) substrate mixture (0.5 mg/ml nitroblue tetrazolium, 0.19 mg/ml 5-bromo-4-chloro-3-indolyl phosphatase in alkaline phosphatase substrate buffer) during 10 min and the reaction was stopped by rinsing with water.

2.2.12. Tissue Print Immunoblotting (TPIB)

Transversal sections of petioles, veins and tendrils 1 mm thick were cut and placed onto nitrocellulose membranes (Millipore, USA) with drawn squared grids (1 x 1 cm). The sections were covered with a layer of tissue paper, pressed for 30 sec by hand and then removed from the membrane. Subsequently, immunodetection was carried out as described in 2.2.9 for the western blot membranes.

2.2.13. In Situ Immunoassay (ISIA)

GLRaV-1 molecularly positive and negative plants were used to prepare transversal plant tissue sections from tendrils and petioles. Sections 40-60 µm thick were obtained with a cryostat (Leica CM1850, Leica Microsystems, Germany) and transferred to glass microscope slides freshly coated with Merckoglas (Merck, Germany). Sections were blocked in PBS (pH 7.4) with 3% BSA during 30 min at room temperature and washed twice with PBS for 5 min. Then, in order to test the IgG, monoIgG-AP and monoIgG-FITC, three groups of slides containing sections from both positive and negative plants were processed.

2.2.13.1. ISIA with Polyclonal IgG

Blocked sections were incubated during 2 hours with polyclonal primary antibody diluted 1:1500 in PBS. After 2 washes each for 5 min with PBST (0.05%), sections were incubated for 2 hours with alkaline phosphatase-labelled anti-goat secondary antibody (Sigma, USA) diluted 1:2500 in PBS and rinsed for 5 min with TBST (0.1%) before incubation with NBT-BCIP (Roche, Switzerland) substrate mixture (0.5 mg/ml nitroblue tetrazolium, 0.19 mg/ml 5-bromo-4-chloro-3-indolyl phosphatase in alkaline phosphatase substrate buffer) during 10 min. The reaction was stopped by rinsing with water.

2.2.13.2. ISIA with MonoIgG-AP

After blocking, slides were directly incubated with monospecific IgG conjugated with alkaline phosphatase (diluted 1:2500 in PBS with 1% BSA) for 2 hours, washed 2 times with TBST (0.1%) for 5 min and incubated with NBT-BCIP substrate mixture for 10 min. The reaction was stopped washing the sections with water.

2.2.13.3. ISIA with MonoIgG-FITC

Other previously blocked sections were incubated with FITC conjugated monospecific IgG (diluted 1:250 in PBS with 1% BSA and 0.1% Triton-X) for 90 min followed by 3 washes with PBS-Trinton-X for 3 min each.

All ISIA results were observed and recorded under an epifluorescence microscope (Axio Imager.Z2, Carl Zeiss, Germany) using either bright field or an excitation filter for FITC.

With NBT-BCIP staining samples that developed a purple colour in the phloem tissue cells were considered positive for the presence of GLRaV-1 and samples that did not were found negative. In the case of FITC labelling, positive samples presented a bright green-fluorescence in the phloem tissue cells.

The captured images were further analyzed with ImageJ 1.43U from imageJ.nih.gov/ij.

2.3. Results

2.3.1. PCR amplification

A total of 20 GLRaV-1 isolates of which 16 were from Dois Portos (CAN) and 4 from Quinta do Marquês (Oeiras) were analyzed in this study. Mixed infections with other viruses associated with leafroll disease were detected in most of the analyzed isolates (Table 2.1).

The assessment of GLRaV-1 genetic variability was carried out for the CP gene. Amplicons encompassing the entire CP gene (969 bp) were amplified with either one of the previously described primers pairs (Table 2.2), cloned and sequenced for all the 20 isolates analyzed. Subsequently, for comparison with the previously published information, a 540 bp fragment of the HSP70h gene was amplified, cloned and sequenced for 6 of the isolates, all originating from Dois Portos.

Two of the isolates (B3 and SS3) had been previously found negative for GLRaV-1 by DAS-ELISA with commercial antibodies. However, evidence of infection by the virus was retrieved in the form of CP (for both isolates) or HSP70h (for SS3) gene sequence variants.

2.3.2. Genetic Diversity of GLRaV-1 Isolates

The SSCP analysis of the CP gene evidenced the existence of high levels of heterozigosity within isolates (Table 2.3). Half of the isolates presented two or more similarly frequent sequence variants, whereas nine (A0305, B3, GR3, J3, J5, M2, OdP2, OdP5 and SS3) presented a dominant variant over a cloud of several others at low frequency. For the isolate CdG a single sequence variant was identified. In the case of the HSP70h fragment less sequence variants were detected per isolate when considered the SSCP patterns and, with the exception of isolate J2 lower heterozigosity levels were obtained. Isolates EM3, SS3 and SVV1 showed an HSP70h dominant variant, GR3 and TL3 presented a unique variant and J2 present two equally frequent variants. For both

genomic regions, the distribution of SSCP patterns showed variability in most isolates, demonstrating that combinations of sequence variants occur within them. But, overall, no systematic pattern of within isolate genetic structure was apparent.

Estimated values of within isolate genetic distance found in this work were considerably low for the CP gene (Table 2.3), ranging from 0.001 ± 0.001 (isolate M7) to 0.011 ± 0.003 (isolate TL3). The corresponding values determined on the basis of the HSP70h 540 nt fragment were similar for the 6 isolates tested, with the lowest being for J2 (0.002 ± 0.001) and the highest within the isolate SVV1 (0.008 ± 0.003).

The average genetic distance between isolates (data not shown on table) ranged from 0.003 ± 0.001 (between M2 and M7) to 0.195 ± 0.018 (between M4 and CdG) for the CP gene, with the lower values corresponding to isolates originating from clone plants, and from 0.025 ± 0.007 (between SS3 and TL3) to 0.062 ± 0.011 (between SVV1 and EM3) for the HSP70h gene. For the later no clone plants from the same grapevine variety were compared.

Regarding overall divergence for each genomic region under study, values varied according to the sequence groups considered (Table 2.3). Estimates of overall diversity were similar when the 83 CP gene sequences obtained in this work were treated as an individual group (0.072 ± 0.005) or when grouped with the 8 complete sequences retrieved from GenBank (0.076 ± 0.005; data not shown on table). However that value was significantly lower than the overall diversity within the group of sequences previously available at GenBank (41 sequences: 143 ± 0.010) and within the group of 29 incomplete CP gene sequences (686 nt) originating from the USA (0.110 ± 0.008; data not shown on table).

For the HSP70h gene, overall mean nucleotide divergence was also found significantly lower when determined for the 11 sequences from this work in comparison with the ones available at GenBank (Table 2.3).

For both the CP and the HSP90h genes, divergence values between the respective group of sequences obtained in this work and the homologous group available at GenBank was lower than the latter's within group divergence value.

Table 2.3 Number of variants, estimates of Nei's heterozygosity (*h*) and average evolutionary divergence (d) within isolates of GLRaV-1 for the HSP70h and CP genes

Isolate ID[a]	CP			HSP70h		
	# of variants	*h*	d	# of variants	*h*	d
A0305	3	0.42	0.003 ± 0.002	-	-	-
B3	13	0.78	0.008 ± 0.001	-	-	-
CdG	1	0	-	-	-	-
EM3	4	1	0.003 ± 0.001	2	0.48	0.004 ± 0.003
GR3	7	0.79	0.004 ± 0.001	1	0	-
J2	4	1	0.005 ± 0.002	2	1	0.002 ± 0.001
J3	5	0.73	0.006 ± 0.002	-	-	-
J5	3	0.52	0.006 ± 0.002	-	-	-
J6	6	0.89	0.007 ± 0.002	-	-	-
M2	4	0.53	0.003 ± 0.001	-	-	-
M4	7	0.91	0.007 ± 0.002	-	-	-
M7	2	0.67	0.001 ± 0.001	-	-	-
OdP2	3	0.52	0.006 ± 0.002	-	-	-
OdP5	3	0.67	0.004 ± 0.002	-	-	-
PP1	2	1	0.005 ± 0.002	-	-	-
SS3	6	0.67	0.002 ± 0.001	1	0	-
SVV1	2	0.67	0.003 ± 0.002	4	0.45	0.008 ± 0.003
TL3	2	1	0.011 ± 0.003	1	0	-
VD4	2	0.50	0.004 ± 0.002	-	-	
VD7	4	1	0.005 ± 0.002	-	-	
All[b]	83	-	0.072 ± 0.005	11	-	0.017 ± 0.009
All[c]	41	-	0.143 ± 0.010	95	-	0.054 ± 0.002
All[d]	124	-	0.106 ± 0.007	106	-	0.052 ± 0.021

Results are based on the alignment of 124 CP gene sequences and 106 HSP70h gene sequences. Standard error estimates were obtained by a bootstrap procedure (1000 replicates). Average evolutionary divergence analyses were conducted in MEGA5 using gene specific substitution models and codon positions included were $1^{st}+2^{nd}+3^{rd}$+Noncoding. Positions containing gaps and missing data were treated as parwise deletion.

[a] GenBank accession numbers and other details of isolates are provided in Table 2.1.

[b] Analysis including all GLRaV-1 gene sequences from this work .

[c] Analysis including all GLRaV-1 gene sequences available at GenBank.

[d] Analysis including all GLRaV-1 gene sequences from this work and available at GenBank.

2.3.3. Phylogenetic Analysis of GLRaV-1 Isolates

Phylogenetic analysis of global isolates for both HSP70h and CP genes was performed using the sequences obtained in this study and all the correspondent sequences available at GenBank. Maximum likelihood trees were constructed with MEGA5 using the models with the lowest BIC scores (Figs. 2.2 and 2.3). The HSP70h and CP datasets included, respectively, a total of 106 and 124 sequences of GLRaV-1, to which a homologous sequence of GLRaV-3 was added as outgroup. For both genes, the phylogroups obtained from GLRaV-1 global isolates disclosed no evidence of sequences grouping by geographical origin. Interestingly cases were found where the same sequence variant was present in different isolates, either in this work (Fig. 2.2: J2-3 and J3-19) or in isolates from the USA (Figs. 2.2 and 2.3).

The known global variants of CP gene segregated into 8 major phylogroups strongly supported by bootstrap values ≥ 79%. Phylogroups 1 was composed of 82 sequences contributed by Portugal, Poland, USA, Brazil, South Africa and India and harboured six well-resolved sub-clusters (bootstrap values ≥ 84%) named group 1a to 1f. Groups 1a to 1e were exclusively composed of variants from isolates characterized in this study, while subgroup 1f included sequences from India, Poland and USA. Sequences retrieved from the three clonal isolates from variety Monvedro grouped with sequences from China and Iran in phylogroup 4. Phylogroups 2 and 3 presented variants from two (Canada and USA) and four (Australia, Chile, China and USA) countries, respectively. Phylogroups 5, 6, 7 and 8 were exclusively composed of variants from California.

Sequence variants from Portuguese isolates with the same clonal origin always grouped together. None of the Portuguese isolates presented variants assigned to more than one phylogroup. Contrary to this pattern, some isolates from California harboured variants from different phylogroups: CA6 (groups 2 and 5), CA11 (groups 3 and 6), CA16 (groups 3 and 7), CA17 (groups 3 and 7) and CA20 (groups 6 and 8). Only one isolate from California, which position in the dendrogram was unresolved, was not included in any of the considered phylogroups.

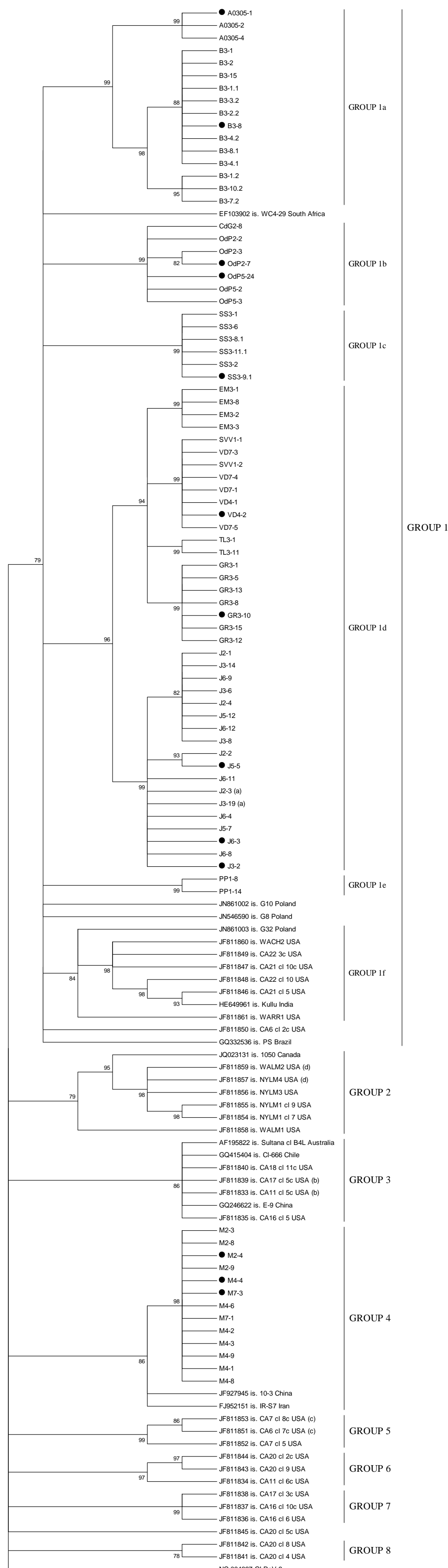

Figure 2.2 Phylogenetic tree of GLRaV-1 global isolates constructed from 125 nucleotide sequences of the CP. Black circles indicate dominant variants within each isolate and letters in brackets indicate duplicate sequences. Branches reproduced in less than 75% of bootstrap replicates are collapsed. Sequences retrieved from GenBank are indicated by accession number and country of origin and sequences obtained in this work are indicated by their isolate name. The GLRaV-3 isolate was used as outgroup.

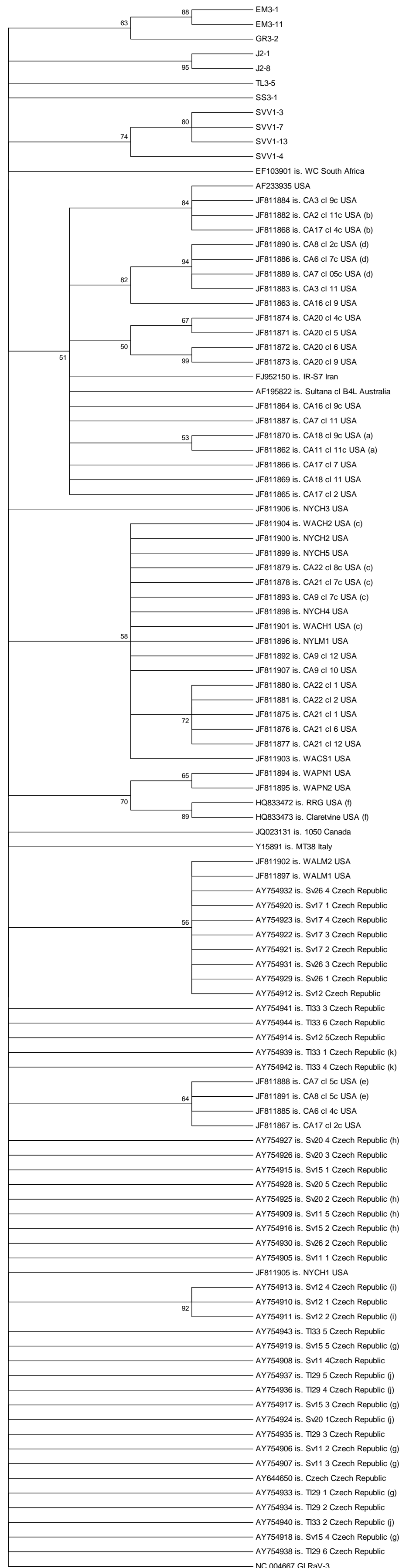

Figure 2.3 Phylogenetic tree of GLRaV-1 global isolates constructed from 107 nucleotide sequences of the HSP70h fragment. Letters in brackets indicate duplicate sequences. Branches reproduced in less than 50% of bootstrap replicates are collapsed. Sequences retrieved from GenBank are indicated by accession number and country of origin and sequences obtained in this work are indicated by their isolate name. The GLRaV-3 isolate was used as outgroup.

The phylogenetic analysis of the HSP70h fragment (Fig. 2.3) lead to a much less resolved tree than the one obtained with the CP sequences. In fact, thirty-five sequence variants remained unassigned to any phylogroup, even when considering a cut-off value of 50%. The HSP70h variants from Portuguese isolates formed groups more closely related between them than between groups of isolates from other countries, but for all the other sequences, the dendrogram did not provide evidence of variants clustering together by geographical origin.

2.3.4. GLRaV-1 Phylogroups Diversity

The analysis of evolutionary divergence was only conducted for the phylogroups identified in the CP phylogenetic analysis. The poorer resolution observed in the HSP70h ML tree precluded the recognition of putative phylogroups.

The lowest value of mean genetic divergence within the phylogroups depicted in Fig. 2.2 was found for group 1e, composed of only two variants from a single Portuguese isolate (0.005 ± 0.002) (Table 2.4).

Phylogroup 3, the cluster of variants with the most diversified origins (four countries) showed a lower mean genetic distance than group 1f (variants from India, Poland and USA) and 2 (variants from Canada and USA) or even groups 6, 7 and 8, exclusively with USA variants. In fact, the highest value (0.105 ± 0.014), similar to the one considering the 124 sequences under analysis, was found for phylogroup 8, with two sequence variants from the California isolate CA20.

The average evolutionary divergence over sequence pairs between the major phylogroups identified for the CP gene ranged from 0.104 ± 0.008 (between groups 5 and 6) to 0.372 ± 0.035 (between groups 2 and 8) as reported in Table 2.5. Hence, genetic diversity within major phylogroups was lower than the evolutionary divergence between them, supporting the phylogenetic inference from this study. The values of genetic distance between groups 1a to 1f were lower than the previous (data not shown on table) ranging from 0.063 ± 0.008 to 0.093 ± 0.011.

Table 2.4 Estimates of average evolutionary divergence (d) within each phylogroup of the CP gene sequences of GLRaV-1

Phylogroup[a]	Number of variants	d
1	82	0.056 ± 0.005
1a	16	0.017 ± 0.002
1b	7	0.018 ± 0.002
1c	6	0.002 ± 0.001
1d	39	0.028 ± 0.003
1e	2	0.005 ± 0.002
1f	8	0.042 ± 0.005
2	7	0.065 ± 0.006
3	7	0.030 ± 0.005
4	15	0.021 ± 0.003
5	3	0.024 ± 0.005
6	3	0.055 ± 0.008
7	3	0.044 ± 0.007
8	2	0.105 ± 0.014
All	124	0.106 ± 0.007

[a] Phylogroups are as defined in Fig. 2.2.

Results are based in the alignment of 124 complete CP gene sequences. Standard error estimates were obtained with 1000 bootstrap replicates.

Average evolutionary divergence analysis was conducted in MEGA5 using gene-specific substitution models and codon positions included 1st + 2nd + 3rd + noncoding. Positions containing gaps and missing data were treated as a pairwise deletion.

Table 2.5 Estimates of average evolutionary divergence (d) between GLRaV-1 CP major phylogroups

Phylogroup[a]	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Group 1		0.017	0.011	0.013	0.014	0.013	0.015	0.025
Group 2	0.175		0.017	0.021	0.021	0.019	0.022	0.035
Group 3	0.112	0.169		0.015	0.014	0.014	0.016	0.027
Group 4	0.139	0.203	0.136		0.014	0.014	0.016	0.028
Group 5	0.129	0.202	0.121	0.108		0.012	0.013	0.024
Group 6	0.130	0.190	0.123	0.120	0.104		0.014	0.023
Group 7	0.147	0.221	0.138	0.133	0.115	0.116		0.023
Group 8	0.273	0.372	0.277	0.272	0.245	0.238	0.252	

[a] Phylogroups are as defined in Fig. 2.2.

Results are based in the alignment of 124 complete CP gene sequences. Standard error estimates were obtained with 1000 bootstrap replicates. Average evolutionary divergence analysis was conducted in MEGA5 using gene-specific substitution models and codon positions included 1st + 2nd + 3rd + noncoding. Positions containing gaps and missing data were treated as a pairwise deletion.

2.3.5. Recombination and Selection Constraints Analysis

Putative recombination events in the HSP70h and CP gene sequences obtained in this work and all available sequences at GenBank were evaluated with GARD method from Datamonkey web server. No evidence of recombination was found. A subsequent analysis with RDP3 software including the CP sequences from this work also did not indicate recombination events in this gene.

Analysis of selection pressures acting on HSP70h and CP genes was carried out on Datamonkey web server. For the CP analysis, the four shorter sequences were not included. Global mean dN/dS values were determined for both datasets based on the trees constructed by the Datamonkey software and estimated values indicated that both genes are under purifying selection (Table 2.6), with the CP gene showing a higher number of negatively selected sites. Moreover, given that the dN/dS obtained for the CP gene is lower than for the partial HSP70h, functional constraints appear to be superior in the CP. The datasets including only the sequences retrieved from GenBank showed slightly lower values, but the same pattern, for the HSP70h and CP genes.

Table 2.6 Selection pressures acting on the HSP70h and CP genes of GLRaV-1

Gene	Global dN/dS[e]		Positively selected sites[f]		Negatively selected sites[f]	
	Log(L)	Mean	N	(%)	N	(%)
HSP70h[a]	-1141.35	0.329	1	1	11	6
HSP70h[b]	-5603.61	0.325	5	3	84	46
CP[c]	-4374.72	0.268	4	1	82	26
CP[d]	-8919.72	0.233	4	1	175	54

[a] Dataset is represented by 11 partial HSP70h gene sequences (181 codons) obtained in this work.
[b] Dataset is represented by 106 partial HSP70h gene sequences (181 codons) obtained in this work and available at GenBank.
[c] Dataset is represented by 83 complete CP gene sequences (322 codons) obtained in this work.
[d] Dataset is represented by 120 CP gene sequences (322 codons) obtained in this work and available at GenBank.
[e] Mean dN/dS < 1 indicates negative or purifying selection, dN/dS = 1 suggests neutral selection, and dN/dS > 1 indicates positive selection.
[f] Positively and negatively selected sites are identified by at least one of the three selection detection methods: single likelihood counting (SLAC), fixed effects likelihood (FEL) and internal fixed effects likelihood (IFEL). SLAC is a counting method while FEL and IFEL are likelihood methods.

The deduced amino acid sequences alignment of the CP showed that variants of each phylogroup share residue substitutions and furthermore, within each phylogroup, the variants from the same isolate show an even more distinctive substitution pattern (Fig. 2.4). The majority of the substitutions are concentrated in the first 154 AA residues, mainly between residues 10 and 100. The hydrophilicity profile of the CP based on the amino acid alignment including a total of 124 sequences (Fig. 2.5) showed that some conserved regions have high antigenic potential, suitable for antibodies production.

2.3.6. Serological Assays

The specificity of the IgG produced against the selected partial sequence of the GLRaV-1 coat protein was first confirmed by western blot analysis after separation by SDS-PAGE of grapevine proteins present in crude extracts. The expected 35 kDa viral protein (Fazeli and Rezaian, 2000) was identified in infected plants (Fig. 2.6) and no identical bands were observed in grapevine extracts free from GLRaV-1 (i.e. molecularly negative).

Samples that previously tested negative or positive for the presence of GLRaV-1 were also tested with the IgG produced in this study, either by tissue print immunoblotting (TPIB) or by in situ immunoassay (ISIA). For TPIB, monoIgG-AP was used and the protein of interest was detected by the formation of a purple insoluble product in the imprints made with tissue sections onto a nitrocellulose membrane. When the imprints developed a purple colour in the phloem tissue area, samples were considered positive for the presence of GLRaV-1. Samples with unstained corresponding imprints were found negative (Fig. 2.7). For ISIA testing, monoIgG-AP or -FITC was used, and in these cases, the phloem tissue cells of positive samples developed, respectively, a dark purple or bright green-fluorescent colour (Fig. 2.8).

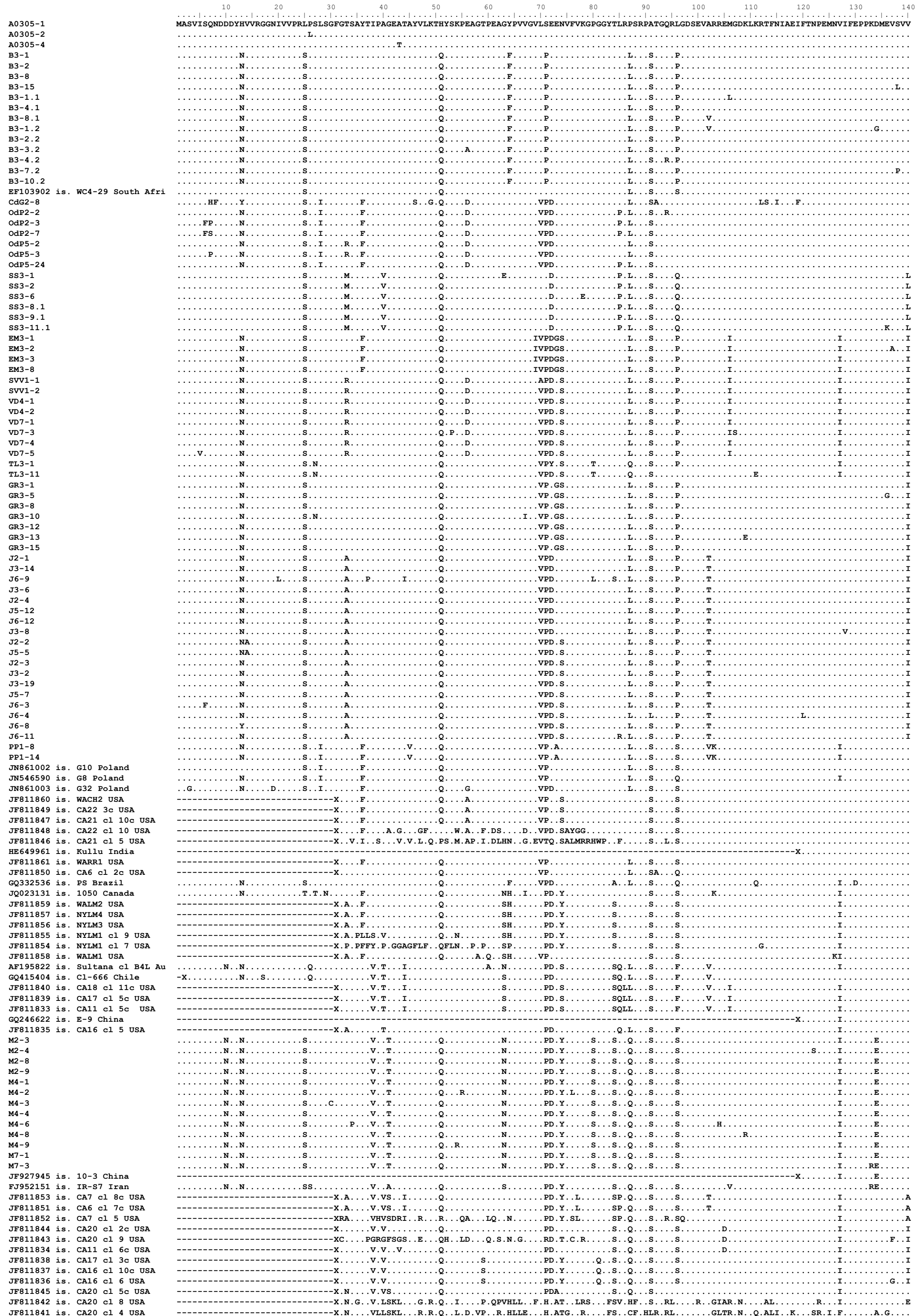

Figure 2.4 Alignment of deduced amino acid sequences of the GLRaV-1 coat protein (322 AA). Sequences retrieved from GenBank are indicated by accession number and country of origin and sequences obtained in this work are indicated by their isolate name. All amino acid substitutions are shown.

```
                                       150       160       170       180       190       200       210       220       230       240       250       260       270       280
                                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
A0305-1                         VPTGPGLVTPAVATAISTELKNLCAGVMGNTDQKSLTDFFLAMLQLMLTFSTSPDTESKEEYFVNLYSNGERKLTYEKVKGAVVKGAEGSTFENPMRQYARLFSATAVHLILNGKLKPNEKVAMQHGVPKRFLPYTFDFC
A0305-2                         .........................E..................................................................................................................
A0305-4                         .........................E..................................................................................................................
B3-1                            .........................E..................................................................................................................
B3-2                            .........................E..................................................................................................................
B3-8                            .........................E....................T................................................L............................................
B3-15                           .........................E..................................................................................................................
B3-1.1                          .........................E..................................................................................................................
B3-4.1                          .........................E..................................................................................................................
B3-8.1                          .........................E.......................................................D................N..............E.....Q..........K........R
B3-1.2                          .........................E...............................C...................................D........................................C.....
B3-2.2                          .........................E..............................................................................................................S...
B3-3.2                          .........................E..................................................................................................................
B3-4.2                          .........................E.......................................................T..........................................................
B3-7.2                          .........................E...............................C..................................................................................
B3-10.2                         .........................E...............................C..................................................................................
EF103902 is. WC4-29 South Afri  .........................E................I.................................................................................................
CdG2-8                          .S.....F...........F.....E......HT................................SR....H...A..E....Q..D....DS..................H...R......LH...............
OdP2-2                          .........................E..........................................................................................R.......................
OdP2-3                          .........................E..........................................................................................R.......................
OdP2-7                          .........................E...................................................................S......................R.......................
OdP5-2                          .........................E..........................................................................................R.......................
OdP5-3                          .........................E..........................................................................................R.......................
OdP5-24                         .........................E..........................................................................................R.......................
SS3-1                           .........................E..................................................................................................................
SS3-2                           .........................E..................................................................................................................
SS3-6                           .........................E..................................................................................................................
SS3-8.1                         .........................E..................................................................................................................
SS3-9.1                         .........................E..........................................................................................R.......................
SS3-11.1                        .........................E..................................................................................................................
EM3-1                           .............A...........E........................................................I.........................................................
EM3-2                           .............A...........E........................................................I.........................................................
EM3-3                           .............A...........E........................................................I.........................................R...............
EM3-8                           .............A..P........E........................................................I................................M........................
SVV1-1                          .............A...........E.....................................................................L............................................
SVV1-2                          .............A...........E.....................................................................L............................................
VD4-1                           .............A...........E.....................................................................L............................................
VD4-2                           .............A...........E.....................................................................L............................................
VD7-1                           .............A...........E.....................................................................L............................................
VD7-3                           .............A...........E.....................................................................L............................................
VD7-4                           .............A...........E..............S.........G............................................L............................................
VD7-5                           .............A...........E.....................................................................L............................................
TL3-1                           .............AT..........E........................................................IF........................................................
TL3-11                          .............AT..........E........................................................I.........................................................
GR3-1                           .............A...........E....................................................................................T.............................
GR3-5                           .............A...........E..................................................................................................................
GR3-8                           .............A...........E..................................................................................................................
GR3-10                          .............A...........E..................................................................................................................
GR3-12                          .............A...........E..................................................................................................................
GR3-13                          .............A...........E...........G......................................................................................................
GR3-15                          .............A...........E..................................................................................................................
J2-1                            .........................E..................................................................................................................
J3-14                           .........................E..................................................................................................................
J6-9                            .........................E..................................................................................................................
J3-6                            .........................E..................................................................................................................
J2-4                            .........................E..................................................................................................................
J5-12                           .........................E............................................................................................................C.....
J6-12                           .........................E..................................................................................................................
J3-8                            .........................E..................................................................................................................
J2-2                            .........................E.........................................................................................S........................
J5-5                            .........................E....................................................................................K.............................
J2-3                            .........................E..................................................................................................................
J3-2                            .........................E..................................................................................................................
J3-19                           .........................E..................................................................................................................
J5-7                            .........................E..................................................................................................................
J6-3                            .........................E..................................................................................................................
J6-4                            .........................E..................................................................................................................
J6-8                            .........................E..................................................................................................................
J6-11                           .........................E..................................................................................................................
PP1-8                           .........................E..................................................................................................................
PP1-14                          .........................E..................................V..............................Y................................................
JN861002 is. G10 Poland         .........................E..................................................................................................................
JN546590 is. G8 Poland          .........................E..................................................................................................................
JN861003 is. G32 Poland         .........................E..................................................................................................................
JF811860 is. WACH2 USA          .........................E.....................................M.......................................................---------------------
JF811849 is. CA22 3c USA        .........................E.............................................................................................---------------------
JF811847 is. CA21 cl 10c USA    .........................E.............................................................................................---------------------
JF811848 is. CA22 cl 10 USA     .........................E..........................................................................................E..---------------------
JF811846 is. CA21 cl 5 USA      .........................E.............................................................................................---------------------
HE649961 is. Kullu India        .........................E.............................X------------------------------------------------------------------------------------
JF811861 is. WARR1 USA          .........................E.............................................................................................---------------------
JF811850 is. CA6 cl 2c USA      .........................EL............................................................................................---------------------
GQ332536 is. PS Brazil          .........................E........................................................I.........................................................
JQ023131 is. 1050 Canada        .........................E..........................................................................................R.......................
JF811859 is. WALM2 USA          .........................E.............................................................................................---------------------
JF811857 is. NYLM4 USA          .........................E.............................................................................................---------------------
JF811856 is. NYLM3 USA          .........................E.............................................................................................---------------------
JF811855 is. NYLM1 cl 9 USA     .........................E.............................................................................................---------------------
JF811854 is. NYLM1 cl 7 USA     .........................E...................................................................I......................N..---------------------
JF811858 is. WALM1 USA          .........................E............................................................................S................---------------------
AF195822 is. Sultana cl B4L Au  L........................E.........................................C........................................................................
GQ415404 is. Cl-666 Chile       .........................E..X---------------------------------------------------------------------------------------------------------------
JF811840 is. CA18 cl 11c USA    .........................E........................................................I....................................---------------------
JF811839 is. CA17 cl 5c USA     .........................E........................................................I....................................---------------------
JF811833 is. CA11 cl 5c  USA    .........................E........................................................I....................................---------------------
GQ246622 is. E-9 China          .........................E.............................X------------------------------------------------------------------------------------
JF811835 is. CA16 cl 5 USA      .........................E........................................................I....................................---------------------
M2-3                            ..........V..............EA.........................T.............................I.R.......................................................
M2-4                            ..........V..............E........................................................I.........................................................
M2-8                            ..........V..............E........................................................I.........................................................
M2-9                            ..........V..............E........................................................I.........................................................
M4-1                            ..........V..........K...E........................................................I.........................................................
M4-2                            ..........V..............E........................................................I.........................................................
M4-3                            ..........V..............E........................................................I.........................................................
M4-4                            ..........V..............E........................................................I.........................................................
M4-6                            ..........V..............E....................................S...................T.........................................................
M4-8                            ..........V..............E........................................................I.........................................................
M4-9                            ..........V..............E........................................................I.........................................................
M7-1                            ..........V..............E........................................................I.........................................................
M7-3                            ..........V..............E........................................................I.........................................................
JF927945 is. 10-3 China         ..S......................E.............................X------------------------------------------------------------------------------------
FJ952151 is. IR-S7 Iran         ..............V..........E........................................................I.........................................................
JF811853 is. CA7 cl 8c USA      I........................E.............................................................................................---------------------
JF811851 is. CA6 cl 7c USA      I........................E.............................................................................................---------------------
JF811852 is. CA7 cl 5 USA       I........................E.............................................................................................---------------------
JF811844 is. CA20 cl 2c USA     .........................E.............................................................................................---------------------
JF811843 is. CA20 cl 9 USA      ...............F.........E.......R.....................................................................................---------------------
JF811834 is. CA11 cl 6c USA     .........................E.............................................................................................---------------------
JF811838 is. CA17 cl 3c USA     .........................E........................................................I....................................---------------------
JF811837 is. CA16 cl 10c USA    .........................E........................................................I....................................---------------------
JF811836 is. CA16 cl 6 USA      .........................EMV...G......LL...VP..V..N.............K..C.R....D...G...IG............P....I......S..........---------------------
JF811845 is. CA20 cl 5c USA     .........................E........................................................I....................................---------------------
JF811842 is. CA20 cl 8 USA      ...AA..T...GG............E.................V.................................R....I....................................---------------------
JF811841 is. CA20 cl 4 USA      ...AV.FR...GE...........PE..............F.............................Q...........I...........L........................---------------------
```

Figure 2.4 continued

```
                                           290       300       310       320
                                   ....|....|....|....|....|....|....|....|..
A0305-1                            RPSYSQFSNDAIRAWQLAAESAFGRKSNVTSSVLRNTSELKV
A0305-2                            ........D.................................
A0305-4                            ..........................................
B3-1                               ..........................................
B3-2                               ..........................................
B3-8                               ..........................................
B3-15                              ..........................................
B3-1.1                             ..........................................
B3-4.1                             ..........................................
B3-8.1                             ..........................................
B3-1.2                             ..........................................
B3-2.2                             ..........................................
B3-3.2                             ..........................................
B3-4.2                             ...............P..........G...............
B3-7.2                             .......I....M.............................
B3-10.2                            ..........................................
EF103902 is. WC4-29 South Afri     .............................-------------
CdG2-8                             ..................................ST......
OdP2-2                             ..........................................
OdP2-3                             ..........................................
OdP2-7                             ..........................................
OdP5-2                             ..........................................
OdP5-3                             ..................................G.......
OdP5-24                            ..........................................
SS3-1                              ..........................................
SS3-2                              ..........................................
SS3-6                              ..........................................
SS3-8.1                            ..........................................
SS3-9.1                            ..........................................
SS3-11.1                           ..........................................
EM3-1                              ..........................................
EM3-2                              ..........................................
EM3-3                              ..........................................
EM3-8                              ..........................................
SVV1-1                             ..........................................
SVV1-2                             ..........................................
VD4-1                              ..........................................
VD4-2                              ...............................P..........
VD7-1                              ...............................L..........
VD7-3                              ..........................................
VD7-4                              ..........................................
VD7-5                              ..........................................
TL3-1                              ..........V......V........................
TL3-11                             ..........................................
GR3-1                              ..........................................
GR3-5                              ..........................................
GR3-8                              ..........................................
GR3-10                             ......................................G...
GR3-12                             ...........................S..............
GR3-13                             ..........................................
GR3-15                             ......................................G...
J2-1                               ..........................................
J3-14                              ..........................................
J6-9                               ..........................................
J3-6                               ...................G......................
J2-4                               ..........................................
J5-12                              ..........................................
J6-12                              ..........................................
J3-8                               ..........................................
J2-2                               ..........................................
J5-5                               ..........................................
J2-3                               ..........................................
J3-2                               ..........................................
J3-19                              ..........................................
J5-7                               ..........................................
J6-3                               ..........................................
J6-4                               ..........................................
J6-8                               ..........................................
J6-11                              ..........................................
PP1-8                              G.........................................
PP1-14                             ................................G.........
JN861002 is. G10 Poland            ..........................................
JN546590 is. G8 Poland             ......................SA.........V........
JN861003 is. G32 Poland            ..........................................
JF811860 is. WACH2 USA             ------------------------------------------
JF811849 is. CA22 3c USA           ------------------------------------------
JF811847 is. CA21 cl 10c USA       ------------------------------------------
JF811848 is. CA22 cl 10 USA        ------------------------------------------
JF811846 is. CA21 cl 5 USA         ------------------------------------------
HE649961 is. Kullu India           ------------------------------------------
JF811861 is. WARR1 USA             ------------------------------------------
JF811850 is. CA6 cl 2c USA         ------------------------------------------
GQ332536 is. PS Brazil             ..........................................
JQ023131 is. 1050 Canada           ..........................................
JF811859 is. WALM2 USA             ------------------------------------------
JF811857 is. NYLM4 USA             ------------------------------------------
JF811856 is. NYLM3 USA             ------------------------------------------
JF811855 is. NYLM1 cl 9 USA        ------------------------------------------
JF811854 is. NYLM1 cl 7 USA        ------------------------------------------
JF811858 is. WALM1 USA             ------------------------------------------
AF195822 is. Sultana cl B4L Au     ..........................................
GQ415404 is. Cl-666 Chile          ------------------------------------------
JF811840 is. CA18 cl 11c USA       ------------------------------------------
JF811839 is. CA17 cl 5c USA        ------------------------------------------
JF811833 is. CA11 cl 5c  USA       ------------------------------------------
GQ246622 is. E-9 China             ------------------------------------------
JF811835 is. CA16 cl 5 USA         ------------------------------------------
M2-3                               ..........................................
M2-4                               ..........................................
M2-8                               ..........................................
M2-9                               ..........................................
M4-1                               .................V........................
M4-2                               ........................................Q.
M4-3                               ..........................................
M4-4                               ..........................................
M4-6                               ..........................................
M4-8                               ..................................G.......
M4-9                               ..........................................
M7-1                               ..........................................
M7-3                               ..........................................
JF927945 is. 10-3 China            ------------------------------------------
FJ952151 is. IR-S7 Iran            ..........................................
JF811853 is. CA7 cl 8c USA         ------------------------------------------
JF811851 is. CA6 cl 7c USA         ------------------------------------------
JF811852 is. CA7 cl 5 USA          ------------------------------------------
JF811844 is. CA20 cl 2c USA        ------------------------------------------
JF811843 is. CA20 cl 9 USA         ------------------------------------------
JF811834 is. CA11 cl 6c USA        ------------------------------------------
JF811838 is. CA17 cl 3c USA        ------------------------------------------
JF811837 is. CA16 cl 10c USA       ------------------------------------------
JF811836 is. CA16 cl 6 USA         ------------------------------------------
JF811845 is. CA20 cl 5c USA        ------------------------------------------
JF811842 is. CA20 cl 8 USA         ------------------------------------------
JF811841 is. CA20 cl 4 USA         ------------------------------------------
```

Figure 2.4 concluded.

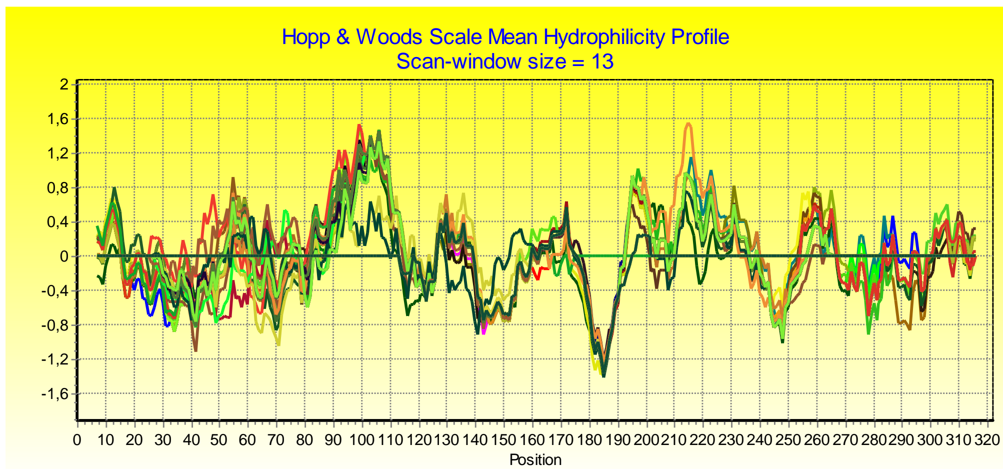

Figure 2.5 Hydrophilicity profile for the CP of GLRaV-1. All the 124 variants in Fig. 2.2 are shown. The average antigenicity values are plotted versus position along the amino acid sequence. The x-axis contains 322 increments, each representing an amino acid. The y-axis represents the range of hydrophilicity values.

From the total of the 20 isolates analyzed in this study, isolates B3 and SS3 had previously tested negative by DAS-ELISA performed with commercial antibodies (Bioreba, Switzerland). However, the infection with GLRaV-1 was identified in those isolates when serological tests were carried out with the antibodies developed in this work (Figs. 2.7 ad 2.8). Hence, the results from the different serological assays performed with the newly developed antibodies were in total agreement with the molecular data. These results suggest that some of the GLRaV-1 variants may pass unnoticed by serological assays using the commercial antibodies currently available.

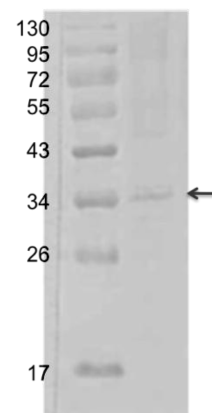

Figure 2.6 Western blot analysis of the IgG produced against the selected partial sequence of the GLRaV-1 coat protein. The arrow indicates the specific detection of the expected 35 kDa band.

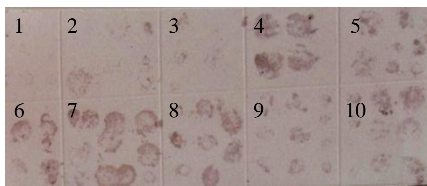

Figure 2.7 Evidence of tissue print immunoblotting with MonoIgG-AP obtained in this work for one GLRaV-1 negative sample and nine positive samples: 1-negative sample; 2-J2; 3-J3; 4-J5; 5-J6; 6-PP1; 7-B3 ; 8-EM3; 9-SS3; 10-VD7.

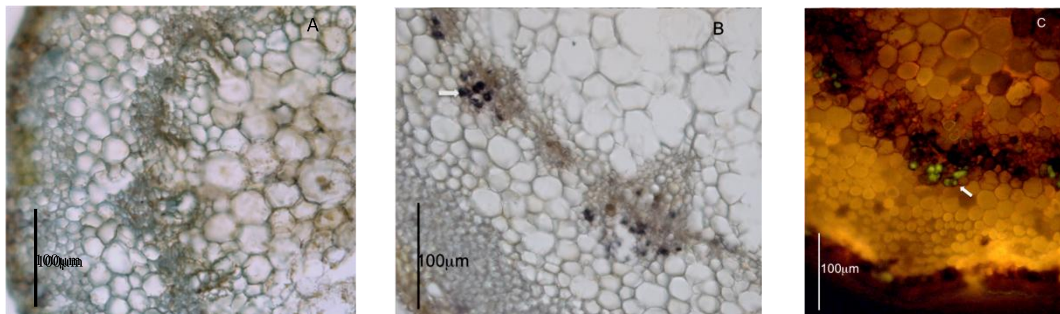

Figure 2.8 Example of the ISIA results for a negative (A) and a positive sample (B, C; grapevine variety Brancelho). The white arrows indicate capsid protein clusters stained with NBT-BCIP (B, bright field) and with FITC (C, fluorescence microscopy with GFP filter) in different slides.

2.4. Discussion

Development of efficient diagnostic tools requires genetic information on a representative array of virus variants in order to allow their detection. Thus, viral population genetic studies are essential to implement proficient routine detection methods. Because of the complex aetiology of the leafroll disease, knowledge of the genetic structure and variability of its causal agents has been developed in a progressive manner. The information obtained in this study on the genetic diversity of 20 Portuguese GLRaV-1 isolates, contributes to a better understanding of the existent variants of this virus. It also allows comparison with previously reported results from variants of different geographical origins (Komínek *et al.*, 2005, Alabi *et al.*, 2011) providing an update on the global assessment of molecular variability in two genomic regions – the HSP70h, which is the taxonomic marker, and the gene encoding the coat protein, the target molecule for serological testing. Furthermore, the use of clone plants, subjected to sanitary control, provide insight into the drift and dynamics of surfacing of new variants caused by vegetative propagation vs. dispersion by vectors.

The first reported study on the GLRaV-1 molecular variability based on a 540 bp fragment of the HSP70h gene, also analyzed in this work separated the GLRaV-1 variants into two distinct groups according to their geographical origin, tentatively named A and E, for American-Australian and European, respectively (Komínek *et al.*, 2005). The other published work on the same subject comprised the analysis of partial sequences of four genes, including the HSP70h and the CP (Alabi *et al.*, 2011). However, in that case, following the HSP70h and CP phylogenetic analysis conducted, three phylogroups were proposed and no segregation according to variant's origin was observed. Findings of the present work are concurrent with the latter, since groups of variants from countries in different continents clustered together. However, the molecular evidence now available for the CP gene supports the existence of 8 well-supported major phylogroups, while the phylogenetic information provided by the HSP70h is clearly not as robust.

Mean genetic diversity values of the HSP70h fragment within Portuguese isolates were similar to those reported for the Czech Republic isolates (Komínek *et al.*, 2005) but substantially lower than the ones determined within USA isolates (Alabi *et*

al., 2011). The divergence value between sequences obtained in this work and sequences retrieved from GenBank was lower than the value of divergence found within the latter group. The same situation was verified when genetic diversity and phylogeny of the CP sequences dataset were analyzed. In this case, albeit mean genetic distance within Portuguese isolates was about half of the distance found within the American isolates, the average divergence between the set of sequences from this work and the one at GenBank was lower than within the latter. These findings suggest that the global GLRaV-1 isolates so far analyzed are part of a single population. This situation may very well be the result of exchange of grafted-propagated infected material between grapevine growing countries, as often seen with other perennial crops (Rubio *et al.*, 2001).

In this work, all the variants present within a single isolate were found in the same phylogroup. In contrast, four of the USA isolates presented variants segregating into different phylogroups. These contrasting patterns suggest two different field scenarios, clearly depicted in Fig. 2.2. The Portuguese isolates, subjected to sanitary control, represent a situation where virus dissemination occurred through vegetative propagation of infected plant material, at the time of establishment of the collection, from mother plant to the respective varietal clone. On the other hand, the USA isolates illustrate a condition in the absence of sanitary control, allowing for vector-mediated dissemination of the virus, and consequent production of mixed infections, by variants from different groups, in a single host.

Even though clustering of CP variants from different countries was observed, 5 out of the 6 subclusters from phylogroup 1 included exclusively Portuguese variants. The close relation between these isolates may result from the fact that these plants were established 30 years ago, during which time sanitary control precluded introduction of foreign variants and clonality (i.e. vegetative propagation). From the isolates characterized in this study, only clonal isolates from variety Monvedro (M2, M4 and M7), originally from the Algarve, are included in a phylogroup with variants of distinct and distant origin. While speculative at this point, one is tempted to suggest that the knowledge of the history of this variety could shed light into its relatedness to isolates from China and Iran.

Although closely related genetically, isolates from clone plants did not show the same dominant variant, which most probably reflect the already mentioned uneven distribution of virus variants in the plant material used for propagation. Hence, although the controlled sanitary conditions constitute a barrier to the introduction of foreign variants, from different phylogroups, it cannot prevent the reshuffling of variants caused by vegetative propagation, to the point of surfacing of different dominant variants between clones.

Results from this study enhance not only the significance of maintaining grapevine crops under phytossanitary control, but also the need to ensure that plant material used for propagation is, at inception, free of viral infections. The lack of correlation between GLRaV-1 variants and geographical origin of the isolates is consistent with the assumption that transmission by propagative material has an imperative role in the dissemination of the virus.

Analysis of the selection pressures acting on the HSP70h and CP genes showed that both regions are under purifying selection and stronger constraints were found for the CP. Evidence that this gene is subjected to stronger constraints than other genomic regions has been previously described for GLRaV-1 (Alabi *et al.*, 2011) as well for GLRaV-3 (Wang et al., 2011). Since both viruses are vector transmitted (Sforza *et al.*, 2003) it is probable that those functional constraints are associated with the CP role in virus acquisition by the vector.

Results from this work suggested that currently available commercial antibodies did not detect all the GLRaV-1 variants disclosed by this work, driving attention to the need of improving this routine detection tool. The information gathered in this study expanded the knowledge of global genetic diversity of GLRaV-1 CP and allowed a comprehensive *in silico* analysis aimed at obtaining a highly specific antibody, capable of detecting all GLRaV-1 variants segregating into groups, with distinct AA profiles. In fact, the IgG produced in this work was tested and recognizes all the groups of variants present in the Portuguese isolates. Although a polyclonal IgG was initially obtained, both that and the purified monospecific IgG have increased the sensitivity of serological assays, showing consistency with the results from molecular detection. Conjugation with FITC was intended to permit co-localization with other GLRaVs, namely GLRaV-3. Given the verified suitability of the obtained IgG for TPIB and ISIA, it also allows

testing a large number of samples in a quicker and less expensive way than ELISA. It overcomes the need for plant extracts, thus eliminating the need for homogenizers, tubes or containers prior to testing. Since a rapid sampling and testing from several sections of the plant can be performed, it allows the detection of the virus even when in a low titre as a single infected cell group is needed to give a clear signal that may go undetected by other methods. TPIB provides yet a convenient method for field survey and samples transportation for testing, minimizing contaminations. The sensitivity of the IgG produced in this work also makes it adequate to follow the virus distribution within the host by TPIB or ISIA in different seasons or at different phonological stages of the host. This represents an advance in the way of obtaining and in the quality of new tools to unravel the phylodinamics of GLRD.

Further characterization of GLRaV-1 genetic diversity, preferably in isolates of diverse origin, and knowledge of the biological and epidemiological repercussions of the different variants will aid to clarify the complexity of GLRaV-1 infections in vineyards and the relative importance of each variant to the infections. This may have important implications for future cross-protection and pathogen-derived resistance studies.

Chapter 3. *Grapevine Leafroll-associated Virus 5*

3.1. Introduction

The complete genome of GLRaV-5 was published in 2012 (Thompson *et al.*, 2012) and deposited at GenBank with 13,384 nt. Since then, two other unpublished variants were sequenced and made available, which differ from the previous in the length of the 5' and 3' untranslated regions (UTRs). The evidence shows that GLRaV-5 has positive sense single-stranded genome organized into 7 open reading frames (ORFs 1a, 1b, 2-6; Fig. 3.1), preceded and followed by relatively short 5' and 3' UTRs. As with other Ampelovirus, ORF1a and ORF1b encode a methyltransferase/helicase (MT/HEL) and a RNA-dependent RNA polymerase (RdRp), respectively. The MT/HEL encompasses 4 domains: papain-like leader protease (L-Pro), MET, AlkB and HEL. ORF2, or P5, encodes a protein of 5262 Da (p5). The heat shock protein 70 homologue, characteristic of all *Closteroviridae*, is the product of ORF3. ORF4 encodes a p60 protein, which is a putative heat shock protein 90 homologue (HSP90h), followed by the genes for the CP (coat protein, ORF 5) and p23 (ORF 6) (see Thompson *et al.*, 2012 for further details).

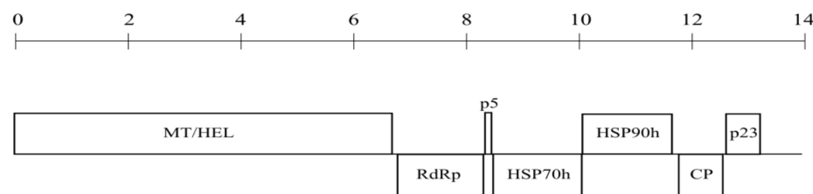

Figure 3.1 Schematic representation of the GLRaV-5 genome organization. Boxes represent the 7 ORFs: MT-methyltransfrease; HEL-helicase; RdRp-RNA-dependent RNA polymerase; p5- 5 kDa protein; HSP70h-heat shock protein 70 homologue; HSP90h-heat shock protein 90 homologue; CP-coat protein; p23-23 kDa protein. The above scale is graduated in kb.

Together with GLRaV-4, -6, -9, -Pr and -De, GLRaV-5, has been considered as forming a distinct evolutionary lineage within ampeloviruses (Maliogka *et al.*, 2008).

However, from these genetically related grapevine viruses, only GLRaV-4 and -6 have also been completely sequenced (Abou Ghanem-Sabanadzovic, et al., 2012).

On the whole this lineage has been considered a scientific curiosity and antibodies for each of the viruses were not made commercially available. Nevertheless, in the last four years the number of reports on the molecular detection of GLRaV-5 as increased. Grapevines with leafroll symptoms, molecularly negative for GLRaV-1 and GLRaV-3 have been found positive for GLRaV-5 when further analyzed. This way, GLRaV-5 has been detected in Argentina (Talquenca *et al.*, 2009), Chile (Engel *et al.*, 2010), China (Pei *et al.*, 2010) and Spain (Padilla *et al.*, 2010). These reports give evidence that GLRaV-5 might be being overlooked for lack of virus-specific routine detection tools and illustrate the possibility for a scenario of a much more widespread and expanding virus than the one considered until now. With the purpose of understanding the implications of GLRaV-5 in leafroll disease, the aim of this work is to study its molecular variability and populational genetic structure in field grown grapevines. The contributed data will allow the identification of genomic variants and transmission dynamics, as well as give new input to clarify the level of GLRaV-5 dissemination worldwide. It will also provide background information crucial for development and improvement of diagnostic tools.

3.2. Material and Methods

3.2.1. Sample Collection and Plant Material

Of the 15 isolates analyzed in this work, 4 were collected at the National Collection of Grapevine Varieties (CAN) from INRB/INIAV in Dois Portos, Portugal, already described in section 2.2.1. Hence, isolate GR3 was obtained from clone plant number 3 of the Gouveio Real grape variety, isolates TB3 and TB4 were obtained from Tempranilla Blanca variety's clone plants 3 and 4, respectively, and isolate TdB from Tinta da Bairrada grape variety. The other 11 isolates were collected from ungrafted field-grown grapevines sampled in several private small-scale vineyards from 5 locations in Algarve, situated in a radius of up to 38 km. Isolates DM and V were collected from Vitoria and Dona Maria grape varieties in the same vineyard. Isolates designated CG, L, ML and MA were collected at Cerro do Guilhim, Lagoa, Mata-Lobos and Moncarapacho, respectively, from different grapevine varieties with vegetatively propagated clones. Isolates detailed information is presented in Table 3.1. The numbers following the isolate designation correspond to the numbers of the clone plants. Mixed infections with GLRaV-5 were detected by PCR after cDNA synthesis using virus-specific primers for GLRaV-1 (Esteves *et al.*, 2009a), GLRaV-2 (Esteves *et al.*, 2009b) and GLRaV-3 (Teixeira Santos *et al.*, 2009a). GenBank accession numbers indicate the sequences obtained in this work for different genomic regions.

To account for the possibility of uneven virus distribution throughout the plant, randomly collected samples from a minimum of five dormant canes or leaves from six branches were combined.

Table 3.1 Details of *Grapevine leafroll-associated virus 5* isolates obtained in this study

Source / Location	Isolate ID	Other GLRaVs detected[a]	Genomic region	GenBank accession number
cv. unkown, red wine grape; Fresh leaf petioles and veins	CG2	GLRaV-3	HSP70h	
Cerro de Guilhim, Algarve			HSP90h	JQ619065 – JQ619068
			CP	JQ619077 – JQ619080
	CG4	GLRaV-3	HSP70h	
			HSP90h	JQ619069 – JQ619070
			CP	JQ619081 – JQ619084
	CG6	GLRaV-3	HSP90h	JQ619071 – JQ619072
			CP	JQ619085 – JQ619086
cv. Dona Maria, white wine grape; Fresh leaf petioles and veins	DM	GLRaV-2, -3	CP	JQ619126 – JQ619129
Tavira, Algarve				
cv. Gouveio-Real, white wine grape; Phloem scrapings	GR3	GLRaV-1, -3	HSP70h	
Dois Portos (INIAV) - CAN (PRT051)			HSP90h	JQ619049 – JQ619053
			CP	JQ619132 – JQ619136
cv. unkown , red wine grape; Fresh leaf petioles and veins	L5	-	HSP70h	
Lagoa, Algarve			CP	JQ619099 – JQ619105
	L6	-	CP	JQ619106 – JQ619110
	L10	-	CP	JQ619111 – JQ619116
cv. unkown, white wine grape; Fresh leaf petioles and veins	MA4	GLRaV-3	CP	JQ619117 – JQ619119
Moncarapacho, Algarve				
cv. unkown, white wine grape ; Fresh leaf petioles and veins	ML2	GLRaV-3	HSP70h	
Mata-Lobos, Algarve			HSP90h	JQ619074 – JQ619076
			CP	JQ619090 – JQ619093
	ML5	GLRaV-3	HSP90h	JQ619073
			CP	JQ619087 – JQ619089
cv. Tempranilla Blanca, white wine grape; Phloem scrapings	TB3	GLRaV-1, -2, -3	HSP70h	
Dois Portos (INIAV) - CAN (PRT051)			HSP90h	JQ619054 – JQ619059
			CP	JQ619094 – JQ619096; JQ619120 – JQ619122
	TB4	GLRaV-1, -2, -3	HSP70h	
			HSP90h	JQ619060 – JQ619064
			CP	JQ619097 – JQ619098; JQ619123 – JQ619125
cv. Tinta da Bairrada, red wine grape; Phloem scrapings	TdB	-	CP	JQ619137
Dois Portos (INIAV) - CAN (PRT051)				
cv. Vitória, white wine grape; Fresh leaf petioles and veins	V	GLRaV-2, -3	HSP70h	
Tavira, Algarve			CP	JQ619130 – JQ619131

[a] Molecular detection was carried out with gene specific primers after cDNA synthesis (Esteves *et al.*, 2009a, Esteves *et al.*, 2009b, Teixeira Santos *et al.*, 2009a).

3.2.2. RNA Extraction and cDNA Synthesis

RNA extraction was carried out as described previously in section 2.2.2.

First strand cDNA was synthesized with RevertAid^{TM} First Strand cDNA Synthesis Kit (Thermo Scientific, USA) in a reaction volume of 20 µl. A mixture of 5 µl of total RNA, 1 µl of 100 µM random hexamer primer and 2 µl nuclease-free water, was first incubated at 65 ºC for 5 min and cooled on ice before the addition of 4 µl of 5X Reaction Buffer, 20 U of RiboLock RNase Inhibitor, 2 µl of 10 mM dNTP mix and 200 U of RevertAid M-MuLV Revert Transcriptase. Reverse transcription mixture was incubated for 5 min at 25 ºC followed by 60 min at 42 ºC. The reaction was terminated by heating at 70 ºC for 5 min. Synthesized cDNA was conserved on ice before direct use or stored at -20 ºC.

3.2.3. PCR Amplification

PCR amplifications were conducted in a total volume of 50 µl with sterile-filtered water (Sigma-Aldrich, USA), 1X *Pfu* Buffer with $MgSO_4$ [200 mM Tris-HCl (pH 8.8 at 25 ºC), 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 1 mg/ml BSA, 0.1% (v/v) Triton-X 100, 20 mM $MgSO_4$] (Thermo Scientific, USA), 200 µM dNTPs (Thermo Scientific, USA), 200 nM of each primer, 2.5 U of *Pfu* DNA Polymerase (Thermo Scientific, USA) and 5 µl of cDNA. PCR was performed in a Biometra thermocycler and cycling conditions consisted of an initial denaturation step at 95 ºC for 3 min, followed by 30 cycles of denaturation at 95 ºC for 30 sec, annealing at 48 ºC for 30 sec and extension at 72 ºC for 2 min for HSP70 and CP genes and 4 min for the HSP90h gene. Gene specific primers are presented in table 3.2.

Table 3.2 Gene specific primers for *Grapevine leafroll-associated virus 5* designed in this work.

Target gene	Primer name	Orientation	Sequence 5'→3'	Genome Position[a]	Product size (bp)
HSP70h	HSP70hLR5-1F	sense	GGCGAGAGCGATAATAGAG	879-897	657
	HSP70hLR5-2R	antisense	GTGCTTAAAGATGGCTGGT	1517-1535	
HSP90h	HSP90hLR5-2F	sense	ACCAGCCATCTTTAAGCAC	1517-1535	1779
	HSP90hLR5-4R	antisense	TGTTCCAGACATCTGTTC	3278-3295	
CP	CPLR5-1	sense	CGAGAAAAGGAAATCCAGTAG	3106-3126	1045
	CPLR5-4	antisense	GACAAGCACAAAGTAACCTCC	4130-4150	

[a]Primers positions according to GenBank accession AF233934

3.2.4. Cloning and Transformation

Amplicons obtained for HSP70h, HSP90h and CP genes were purified using NZYGelpure purification kit (NZYTech, Portugal) following the manufacturer's instructions and ligated using CloneJET^{TM} PCR Cloning Kit (Thermo Scientific, USA) according to fabricant's protocol for blunt-end cloning. Ligation products were used directly to transform XL1-Blue E. *coli* competent cells (Stratagene, USA). For each transformation 10 µl of ligation reaction were used and the ensuing procedure was the same as in 2.2.4.

3.2.5. Single-Strand Conformation Polymorphism (SSCP)

The analysis was conducted in order to select recombinant clones of interest for sequencing. In the case of the HSP70h and CP genes, a minimum of 16 recombinant clones per isolate were analyzed by SSCP after PCR amplification with gene specific primers to verify the presence of the expected insert, as described in section 2.2.5.

With regard to HSP90h, given that the fragment length was outside the detection limit of the method, SSCP analysis was not carried out.

3.2.6. Plasmid DNA Extraction and Restriction Analysis

Concerning the HSP70h and CP genes, a minimum of two recombinant clones for each SSCP pattern detected was selected to purify the plasmid DNA for sequencing. In the case of the HSP9h gene, since SSCP analysis was not performed, plasmid DNA was extracted from at least 10 positive clones per isolate. Plasmid DNA extraction was carried out as described in section 2.2.6.

In order to confirm the presence of the DNA insert after plasmid extraction, each clone obtained was subjected to double restriction reaction of the extracted plasmid DNA, performed with FastDigest® XbaI and XhoI (Thermo Scientific, USA) and the repective products were analyzed as described in 2.2.6.

3.2.7. Sequencing and Sequence Data Analysis

Selected clones were commercially sequenced (CCMAR, UAlg, Portugal). Sequences obtained in this work were first processed with BioEdit and repeated sequences within isolates were excluded. All the homologous sequences of HSP70h, HSP90h and CP genes from GLRaV-5 and related ampeloviruses available at GenBank were retrieved and three final datasets including respectively a total of 43 HSP70h, 37 HSP90h and 110 CP gene sequences were assembled. Although incomplete, 9 of the CP gene sequences from GenBank were included in the alignments for comparison of deduced AA sequences but not used for diversity or phylogenetic analysis. Alignments of nucleotide and amino acid sequences were performed with ClustalW. The graphic views of AA alignments and Hoop and Woods hydrophylicity profiles (Hoop and Woods, 1981) from HSP90h and CP genes were made with BioEdit. Phylogenetic trees were obtained with MEGA5 (Tamura *et al.*, 2011) using the maximum likelihood method and a different model to each gene chosen according to software's option "Find Best DNA/Protein Models". Bootstrap values were estimated with 1000 replicates.

3.2.8. Estimates of Evolutionary Divergence and Selection Pressure

Estimates of evolutionary divergence were inferred with MEGA5 with specific substitution models. The rates of synonymous (dS) and non-synonymous (dN) substitutions per site and mean dN/dS values were determined using the algorithms for single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL) and internal fixed effects likelihood (IFEL) (Kosakovsky Pond and Frost, 2005, Kosakovsky *et al.*, 2006a) and using the HKY85 nucleotide substitution bias model (Hasegawa *et al.*, 1985) from Datamonkey web server (Delport *et al.*, 2010).

Evolutionary analysis was carried out with DnaSP software v. 5.10.01 (Librado and Rozas, 2009) and included Tajima's D test of neutrality (Tajima, 1989) and the test of natural selection, G-test statistics, of the McDonald-Kreitman (MK) test (McDonald and Kreitman, 1991). Tajima's D test is one of the most used statistical tests to estimate the correlation between the number of segregating sites (S) and the average number of nucleotide differences. This test detects an excess or deficiency of mutations at intermediate frequencies compared to new mutations segregating at low or high frequencies. The value of D gives information about the evolutionary history of a given sequence: negative D indicates an excess of low frequency polymorphisms, thus the possibility of selective allelic removal and purifying selection; a positive D indicates low levels of both low and high frequency polymorphisms, hence balancing selection. MK test was used to detect the signature of natural selection on protein coding sequences, comparing the ratio of non-synonymous to synonymous fixed substitutions (differences) and the ratio of non-synonymous and synonymous polymorphic differences between isolates.

3.2.9. Recombination Analysis

Putative recombination events in the three datasets were searched with genetic algorithm for recombination detection (GARD) (Kosakovsy Pond *et al.*, 2006b) from Datamonkey web server and with recombination detection methods from RDP3 Alpha44 software (Martin *et al.*, 2010).

3.3. Results

3.3.1. PCR amplification

A total of 15 GLRaV-5 isolates, 4 from Dois Portos (CAN) and 11 from Algarve, were analyzed (Table 3.1).

During an assessment of the infection of grapevine leafroll disease, isolates from CAN and Algarve were found positive for GLRaV-5. Preliminary detection was carried out with PCR amplification of a fragment from the HSP70h gene in which the taxonomy of grapevine leafroll-associated viruses is based on and sequences were obtained for 8 isolates, 3 from CAN and 5 from Algarve. Subsequently and focusing in one of the main purposes of this work which consists in the improvement of diagnostic tools for GLRaVs, PCR amplification, cloning and sequencing of the CP gene was conducted for a broader range of isolates, in a total of 15. After analysis of the data obtained for the CP gene, 8 isolates (3 from CAN and 5 from Algarve), were selected for further testing if the genetic variability and structure found for this gene was present in other regions of the genome. For that purpose the complete sequence of the HSP90h gene, adjacent to the CP gene, was targeted for amplification.

Mixed infections with other grapevine leafroll-associated viruses were detected in the majority of the isolates (Table 3.1).

3.3.2. Diversity of Sequence Variants within GLRaV-5 Isolates

Molecular variability of GLRaV-5 was first assessed considering the CP gene sequences. Initial SSCP analysis of the CP gene profiles, including determination of heterozigosity coefficient *h* (Table 3.3) showed evidence of high levels of heterozigosity within isolates. Variability in SSCP patterns was detected in most cases, showing that several sequence variants occurred in the same isolate. In the case of the CP gene, isolates GR3, TB, V and DM presented one dominant sequence variant and several others with low frequency. In other isolates high frequency levels were found in two or

three sequence variants along with other much less frequent ones and in the case of TdB only one variant was identified. SSCP profiles of the more frequent CP gene sequence variants of each isolate are showed in Fig. 3.2.

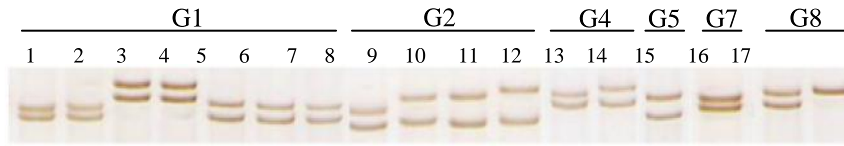

Figure 3.2 CP gene SSCP profiles of the dominant variants of each isolate from the phylogroups defined in Fig. 3.3C. 1 to 17: ML2-10; ML5-4; TB3-6; TB4-6; CG2-5; CG4-1; CG6-2; L5-4; L6-32; L10-30; MA4-28; TB3-3; TB4-1; DM-8; V-12; GR3-2; TdB-24.

Even though the HSP70h gene was not entirely amplified nor sequenced for all the isolates under study, the molecular variability of the genomic fragment, with which GLRaV-5 was first detected, was also analyzed. As verified for the CP gene, the SSCPs profiles revealed high levels of heterozygosity for each isolate, similar to those of the respective CP gene. Isolates DM, GR3, TdB and V showed one dominant sequence variant, while isolate CG4 presented high levels of two sequence variants. For isolate L5 all variants were equally frequent.

In order to understand if the CP genetic variability was extended to other genes and since CP gene data analysis revealed that the highest and lowest heterozigosity levels were found between the isolates obtained from vegetatively propagated clone plants CG2, CG4, CG6, GR3, ML2, ML5, TB3 and TB4, amplification, cloning and sequencing of the HSP90h gene was focused on this group of isolates. With exception of the isolate ML5, sequence variants within each isolate, in the same or lower number that the ones identified in the CP gene, were also found for HSP90h gene.

Genetic distances were estimated for each gene dataset with MEGA5. Diversity within isolates, overall nucleotide divergence and the number of segregating sites are presented in Table 3.3. Values obtained for divergence within isolates, as well as the number of segregating sites varied significantly when the CP, the HSP90h or the partial HSP70h genes were considered. For determination of CP and HSP90h genes overall nucleotide diversity, only complete gene sequences were used. The values obtained for all three genomic regions were similar and did not significantly vary when sequences available at GenBank were included. However, nucleotide divergence between the

group of sequences obtained in this work and those previously reported, was higher for the CP (0.067 ± 0.005) or the HSP70h (0.061 ± 0.007) than within each group. For the HSP90h gene, nucleotide divergence values were also determined considering the sequences from this work and the sequences available at GenBank as one group or two separate groups. Given the fact that only three complete sequences of this gene were retrieved from the database, estimated values were not significantly different.

Table 3.3 Number of variants, estimates of Nei's heterozygosity (*h*), average evolutionary divergence (d) and number of segregating sites (S) within isolates of GLRaV-5 for the HSP70h, HSP90h and CP genes

Isolate ID[a]	HSP70h				HSP90h			CP			
	# of variants	*h*	d	S	# of variants	d	S	# of variants	*h*	d	S
CG2	3	0.57	0.009 ± 0.003	9	4	0.010 ± 0.002	31	4	0.77	0.011 ± 0.002	17
CG4	5	0.69	0.007 ± 0.002	11	2	0.008 ± 0.002	13	4	0.74	0.006 ± 0.001	8
CG6	-	-	-	-	2	0.007 ± 0.002	11	2	0.67	0.011 ± 0.006	9
DM	-	-	-	-	-	-	-	4	0.75	0.006 ± 0.002	10
GR3	6	0.72	0.005 ± 0.001	9	5	0.007 ± 0.001	25	5	0.71	0.003 ± 0.001	7
L5	4	1	0.002 ± 0.001	2	-	-	-	7	0.54	0.006 ± 0.001	14
L6	-	-	-	-	-	-	-	5	0.48	0.006 ± 0.001	13
L10	-	-	-	-	-	-	-	6	0.72	0.005 ± 0.001	12
MA4	-	-	-	-	-	-	-	3	0.62	0.011 ± 0.003	13
ML2	6	0.79	0.005 ± 0.002	10	3	0.013 ± 0.002	31	4	0.73	0.004 ± 0.001	7
ML5	-		-	-	1	-	-	3	0.32	0.007 ± 0.001	8
TB3	5	0.63	0.004 ± 0.001	6	6	0.010 ± 0.002	48	6	0.67	0.037 ± 0.005	54
TB4	1	0	ND	-	5	0.011 ± 0.002	42	5	0.70	0.040 ± 0.005	56
TdB	-	-	-	-	-	-	-	1	0	ND	-
V	4	0.69	0.010 ± 0.003	12	-	-	-	2	0.53	0.007 ± 0.003	6
All[b]	34	-	0.049 ± 0.005	116	28	0.050 ± 0.004	317	61	-	0.055 ± 0.005	253
All[c]	3	-	0.044 ± 0.007	39	3	0.042 ± 0.010	102	18	-	0.053 ± 0.004	147
All[d]	37	-	0.050 ± 0.005	140	31	0.052 ± 0.004	374	79	-	0.059±0.005	294

Results are based on the alignment of 37 HSP70h gene sequences, 31 HSP90h gene sequences and 79 CP gene sequences. Standard error estimates were obtained by a bootstrap procedure (1000 replicates). Average evolutionary divergence analyses were conducted in MEGA5 using gene specific substitution models and codon positions included were 1st+2nd+3rd+Noncoding. Positions containing gaps and missing data were treated as complete deletion.

[a] GenBank accession numbers and other details of isolates are provided in Table 3.1.

[b] Analysis including all GLRaV-5 sequences from this work .

[c] Analysis including all GLRaV-5 sequences available at GenBank.

[d] Analysis including all GLRaV-5 sequences from this work and available at GenBank.

3.3.3. Phylogenetic Analysis of GLRaV-5 Isolates and Related Ampeloviruses

Three datasets of nucleotide sequences encompassing respectively the HSP70h fragment, the HSP90h and CP genes were constructed including the sequences obtained in this work and homologous sequences from GLRaV-4, -5, -6, -9, -10 (-Pr), -11 (-De) and -Carn available at GenBank. For each dataset, maximum-likelihood (ML) trees were constructed using the best evolutionary model determined with MEGA5 (Fig. 3.3A, B and C). A total of 43 sequences were included in the dendrogram obtained for the HSP70h gene fragment (Fig. 3.3A), of which 34 were obtained in this work. HSP90h gene tree (Fig. 3.3B) incorporated 37 sequences, 28 of which retrieved from Portuguese isolates. The CP gene tree (Fig. 3.3C) was constructed with 101 sequences, 61 from this work. In all cases, GLRaV-5 sequence variants clustered forming a resolved major phylogroup strongly supported by bootstrap values of 95% (HSP70h), 100% (HSP90h) or 99% (CP). In all three dendrograms, GLRaV-9 and GLRaV-11 were the ampeloviruses most closely related to GLRaV-5. For all the analyzed genomic regions, GLRaV-6 sequences grouped together with GLRaV-11.

The 79 GLRaV-5 CP gene sequences segregated into eight well-supported major phylogroups (>90%; Fig. 3.3C). No evidence of clustering by geographical location or by white or red grape type was observed. Four out of the eight phylogroups were exclusively composed by sequences from this work (groups 1, 2, 4 and 5) while two other groups (groups 3 and 6) were exclusively composed of sequences from GenBank.

Phylogroups 1, 2 and 4 comprised only sequences from isolates of clone plants, from Dois Portos and Algarve, and within each one of those groups, isolates with the same clonal origin clustered forming highly supported subgroups. Only isolates TB3 and TB4 showed sequence variants distributed in more than one phylogroup (groups 1 and 4). Phylogroup 6 was composed by two isolates from Argentina. Phylogroup 3 comprised isolates originating in three different countries (Argentina, France and USA), while phylogroup 8 included isolates from Argentina and Portugal (GR and TdB). Three isolates, namely Colgadera, Tamares and Pigato, remained unassigned to a phylogroup as a result of insufficient support.

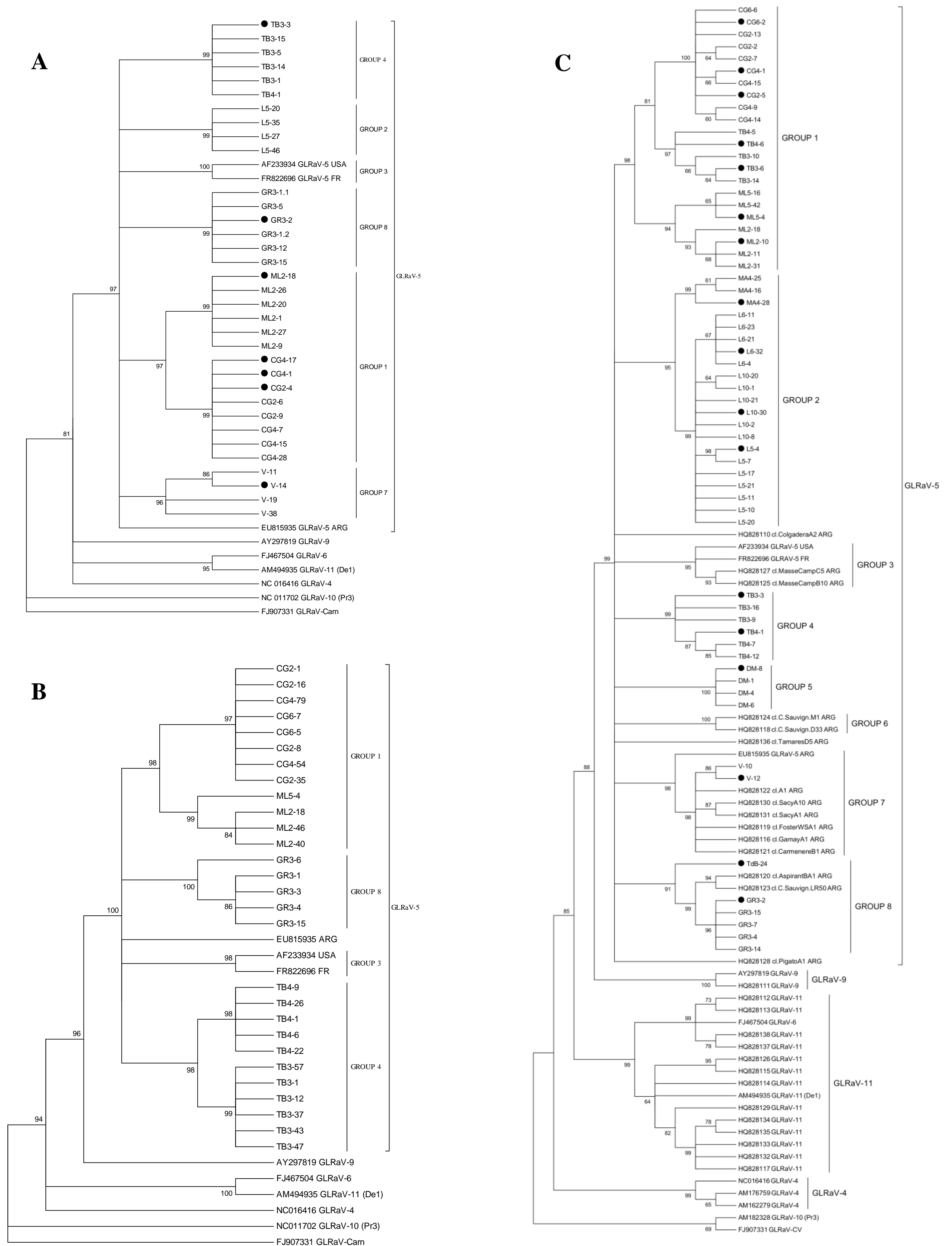

Figure 3.3 Phylogenetic trees of global isolates of GLRaV-5 and related ampeloviruses. Dendrograms are based on the nucleotide sequences of: A) HSP70h partial gene; B) HSP90h gene and; C) CP gene. Black circles indicate the most frequent variant within each isolate analyzed in this work. Sequences retrieved from GenBank are indicated by accession number and sequences obtained in this work are indicated by their isolate name. A tentative correspondence between the phylogroups in each tree is showed.

The 43 partial HSP70h gene sequences and the 37 complete HSP90h gene sequences (Figs. 3.3A and B) clustered respectively, into six and into four phylogroups. Similar to the case of the CP, sequences from isolates with the same clonal origin were found in the same phylogroups, with high bootstrap support.

3.3.4. GLRaV-5 Phylogroups and Related Ampeloviruses Diversity

Considering the CP gene, for which, a higher number of sequences was available, GLRaV-5 presented lower genetic diversity than the related ampeloviruses, GLRaV-4, -9 and -11 (Table 3.4). Two of the highest within-group divergence values and number of segregating sites were found in groups 1 and 2, which included most of sequence variants from clonal plants isolates. The average evolutionary distance between the set of sequences composing these groups (43) and remaining 18 sequences from Portuguese isolates was 0.076 ± 0.008. The value was 0.069 ± 0.007 when the group of sequences from other countries was compared with sequences from groups 1 and 2. The lowest values of average evolutionary divergence and segregating sites were found for phylogroup 5 and 6, which included respectively, the four variants from isolate DM and two Argentinean variants, followed by phylogroup 3 with four sequences from three distinct countries, Argentina, France and USA (Table 3.4).

When the average evolutionary divergence was analyzed considering all ampeloviruses in each dataset, the obtained values were similar for the three genomic regions: 0.163 ± 0.011 for the HSP70h, 0.199 ± 0.008 for the HSP90h and, 0.178 ± 0.012 for the CP. The average evolutionary divergence between GLRaV-5 and all the related ampeloviruses was 0.474 ± 0.034 for HSP70h, 0.525 ± 0.022 for HSP90h and 0.373 ± 0.027 for CP. The average nucleotide diversity within the group of GLRaV-5 related ampeloviruses presented the highest value for the HSP90h (0.619 ± 0.028), followed by the HSP70h (0.565 ± 0.041) and the CP (0.282 ± 0.018). The analysis of the evolutionary divergence between the different ampeloviruses considering the CP gene, showed the highest values for GLRaV-4 vs. GLRaV-9 (0.437 ± 0.034) and the lowest for GLRaV-5 vs. GLRaV-9 (0.290 ± 0.025). For this analysis, since GLRaV-6 clustered with GLRaV-11 in all the trees, both viruses were included in the same group.

Table 3.4 Estimates of average evolutionary divergence (d), number of segregating sites (S), and Tajima's test of neutrality within each CP gene phylogroup

Phylogroup[a]	Number of variants	d	S	Tajima's D
1	22	0.021 ± 0.001	78	-0.90452 (ns)
2	21	0.016 ± 0.003	71	-1.61705 (ns)
3	4	0.007 ± 0.001	10	0.38910 (ns)
4	6	0.012 ± 0.002	24	-0.62499 (ns)
5	4	0.006 ± 0.002	10	-0.83379 (ns)
6	2	0.004 ± 0.002	3	ND
7	9	0.013 ± 0.003	39	-1.33858 (ns)
8	8	0.019 ± 0.007	49	-1.18521 (ns)
GLRaV-5	79	0.055 ± 0.002	294	-1.25747 (ns)
GLRaV-4	3	0.076 ± 0.024	92	ND
GLRaV-9	2	0.078 ± 0.039	63	ND
GLRaV-11	15	0.096 ± 0.009	235	-0.29690 (ns)
All	101	0.178 ± 0.012	477	-1.20027 (ns)

[a] Phylogroups are as defined in Fig. 3.3C.

Results are based in the alignment of 101 complete CP gene sequences. Standard error estimates were obtained with 1000 bootstrap replicates.

Average evolutionary divergence analysis was conducted in MEGA5 using gene-specific substitution models and codon positions included 1st + 2nd + 3rd + noncoding. Positions containing gaps and missing data were treated as a complete deletion.

Tajima's D statistic as a measure of the departure from neutrality for all mutations in the gene was performed using the DnaSP v. 5.10.11. The D test is based on the differences between the number of segregating sites ant the average number of nucleotide differences. The DnaSP software provides a measure of significance of the D value. Values for neutrality were not significant (ns) in all cases ($P > 0.10$). ND = not determined due to insufficient data.

Evolutionary analysis including Tajima's D test of neutrality and the McDonald-Kreitman (MK) test of natural selection was only conducted for HSP90h and CP genes (Tables 3.4 and 3.5), given that sequences of HSP70h did not encompass the entire gene. After performing Tajima's D test, D negative values were obtained within all the HSP90h phylogroups (data not shown on table). In the case of the CP gene, only phylogroup 3 presented a positive D value (Table 3.4). This may result from the fact that variants originated from different populations, which would be expected with isolates from different countries, as occurs in this group. However the value did not significantly deviate from zero ($P < 0.10$). For the HSP90h dataset, MK test was only

carried out for phylogroups of the Portuguese isolates due to insufficient information on isolates from other countries. The obtained results showed significant differences between the ratios of synonymous to non-synonymous substitutions with good statistical support (P < 0.001; P-values of Fisher's and G-test are not shown), indicating that natural selection, instead of random processes like isolation or drift, could be leading the evolution of HSP90h variants in the isolates analyzed. Regarding the CP dataset analysis, comparisons between phylogroups 1, 2, 3 and 4, consistently presented significant differences (Table 3.5). However, significant values for G-test statistics were not found for most of the comparisons with phylogroups including variants from other countries (groups 6, 7 and 8). These results indicate that the ratio of fixed to polymorphic substitutions is not significantly different between synonymous and non-synonymous differences, and consequently, that random processes together with natural selection might be shaping the evolutionary divergence of those phylogroups.

Table 3.5 Estimates of average evolutionary divergence (d) between phylogroups of the CP gene sequences of GLRaV-5 and McDonald-Kreitman test of natural selection

Phylogroup[a]	1	2	3	4	5	6	7	8
1		8.607 (***)	4.572 (*)	15.790 (***)	9.628 (***)	4.001 (*)	2.892 (ns)	14.401 (***)
2	0.060 ± 0.004		5.361 (*)	15.172 (***)	14.360 (***)	1.475 (ns)	2.849 (ns)	12.128 (***)
3	0.068 ± 0.009	0.063 ± 0.009		9.016(**)	3.793 (ns)	1.004 (ns)	0.922 (ns)	2.372 (ns)
4	0.063 ± 0.011	0.070 ± 0.009	0.068 ± 0.017		24.475 (***)	2.408 (ns)	3.459 (ns)	7.985 (**)
5	0.071 ± 0.009	0.074 ± 0.010	0.066 ± 0.020	0.062 ± 0.016		1.949 (ns)	1.776 (ns)	7.668 (**)
6	0.061 ± 0.009	0.068 ± 0.010	0.061 ± 0.020	0.057 ± 0.016	0.058 ± 0.020		0.944 (ns)	0.447 (ns)
7	0.071 ± 0.007	0.067 ± 0.007	0.072 ± 0.015	0.073 ± 0.013	0.066 ± 0.014	0.061 ± 0.015		1.381 (ns)
8	0.072 ± 0.007	0.078 ± 0.008	0.075 ± 0.016	0.065 ± 0.012	0.079 v 0.018	0.064 ± 0.016	0.083 ± 0.013	

[a] Phylogroups are as defined in Fig. 3.3C.

Lower diagonal: evolutionary divergence between phylogroups. Results are based in the alignment of 101 complete CP gene sequences. Standard error estimates were obtained with 1000 bootstrap replicates. Average evolutionary divergence analysis was conducted in MEGA5 using gene-specific substitution models and codon positions included 1st + 2nd + 3rd + noncoding. Positions containing gaps and missing data were treated as a complete deletion.

Upper diagonal: G statistics of the McDonald-Kreitman test. The G-test (G) statistics of the MK test was used to test for evidence that divergence in GLRaV-5 lineages is driven by natural selection. *0.01 < P < 0.05, ** 0.001 < P < 0.01, *** P < 0.001. ns = not significant values. MK test was performed using DnaSP v. 5.10.11.

3.3.5. Selection Constraints on CP and HSP90h Genes

Analysis of selection constraints was only carried out for the HSP90h and CP sequence datasets, given the fact that only these two comprised the complete corresponding ORF. Analysis performed with Datamonkey web server showed no evidence of recombination in either sequence dataset.

Overall mean dN/dS values, based on the trees constructed by the Datamonkey software for HSP90h and CP genes indicated that both are under purifying selection (Table 3.6). The CP dataset comprising the sequences gathered in this work showed a higher dN/dS value than the dataset composed by those sequences and the ones available at GenBank.

Table 3.6 Selection pressures acting on the HSP90h and CP genes of GLRaV-5

Gene	Global dN/dS[d]		Positively selected sites[f]		Negatively selected sites[f]	
	Log(L)	Mean	N	(%)	N	(%)
HSP90h[a]	-5146.73	0.327	3	1	86	16
CP[b]	-3390.42	0.364	9	3	50	19
CP[c]	-4322.74	0.296	8	3	85	32

[a] Dataset is represented by 31 complete HSP90h gene sequences (539 codons) obtained in this work and available at GenBank.
[b] Dataset is represented by 61 complete CP gene sequences (269 codons) obtained in this work.
[c] Dataset is represented by 79 complete CP gene sequences (269 codons) obtained in this work and available at GenBank.
[d] Mean dN/dS < 1 indicates negative or purifying selection, dN/dS = 1 suggests neutral selection, and dN/dS > 1 indicates positive selection.
[e] Positively and negatively selected sites are identified by at least one of the three selection detection methods: single likelihood counting (SLAC), fixed effects likelihood (FEL) and internal fixed effects likelihood (IFEL). SLAC is a counting method while FEL and IFEL are likelihood methods.

After the deduced amino acid sequences from the HSP90h (data not shown) and the CP (Fig. 3.4) genes were aligned it was possible to verify that variants of each isolate share a characteristic residue substitution pattern. At the same time, variants from the same phylogroup share some of those AA substitutions. In the CP, these substitutions are concentrated at the N-terminal domain, in the first 35 AA residues, where six of the eight sites found to be under positive selection (4, 5, 9, 13, 14 and 15; Table 3.6) are situated. For that reason, the shorter sequences available at GenBank,

which lack the N-terminal domain of the protein, contain little phylogenetic information. This pattern of high frequency of AA substitution in the N-terminal domain is also observed in the sequences of GLRaV-11 isolates. Interestingly, the GLRaV-6 deduced CP sequence has only two AA substitutions relatively to the GLRaV-11 sequences (AA9 and AA175 in the alignment). In the GLRaV-5 HSP90h gene, the residue substitutions observed are also characteristic of each isolate; however their distribution occurs throughout the protein sequence. In both HSP90h and CP sequence variants, the hydrophilicity profiles were not altered by the residues substitution. Crossed information given by the AA alignment and the hydrophilicity profile of the CP gene shows that the region with the highest immunogenic index is also the most variable one within the peptide sequence (Fig. 3.5) either for GLRaV-5 or related ameloviruses. This suggests that antibodies raised against the CP of GLRaV-5 and designed to cover all its variability, may in turn fail to specifically detect the virus in mixed infections with related ampeloviruses. HSP90h hydrophilic profile also reveals several hydrophilic peaks, which majority is shared with the related ampeloviruses (Fig. 3.6), but one of those peaks, localized between AA215 and AA270, coincides with a region conserved among all GLRaV-5 sequences but not between GLRaV-5 and the other ampeloviruses.

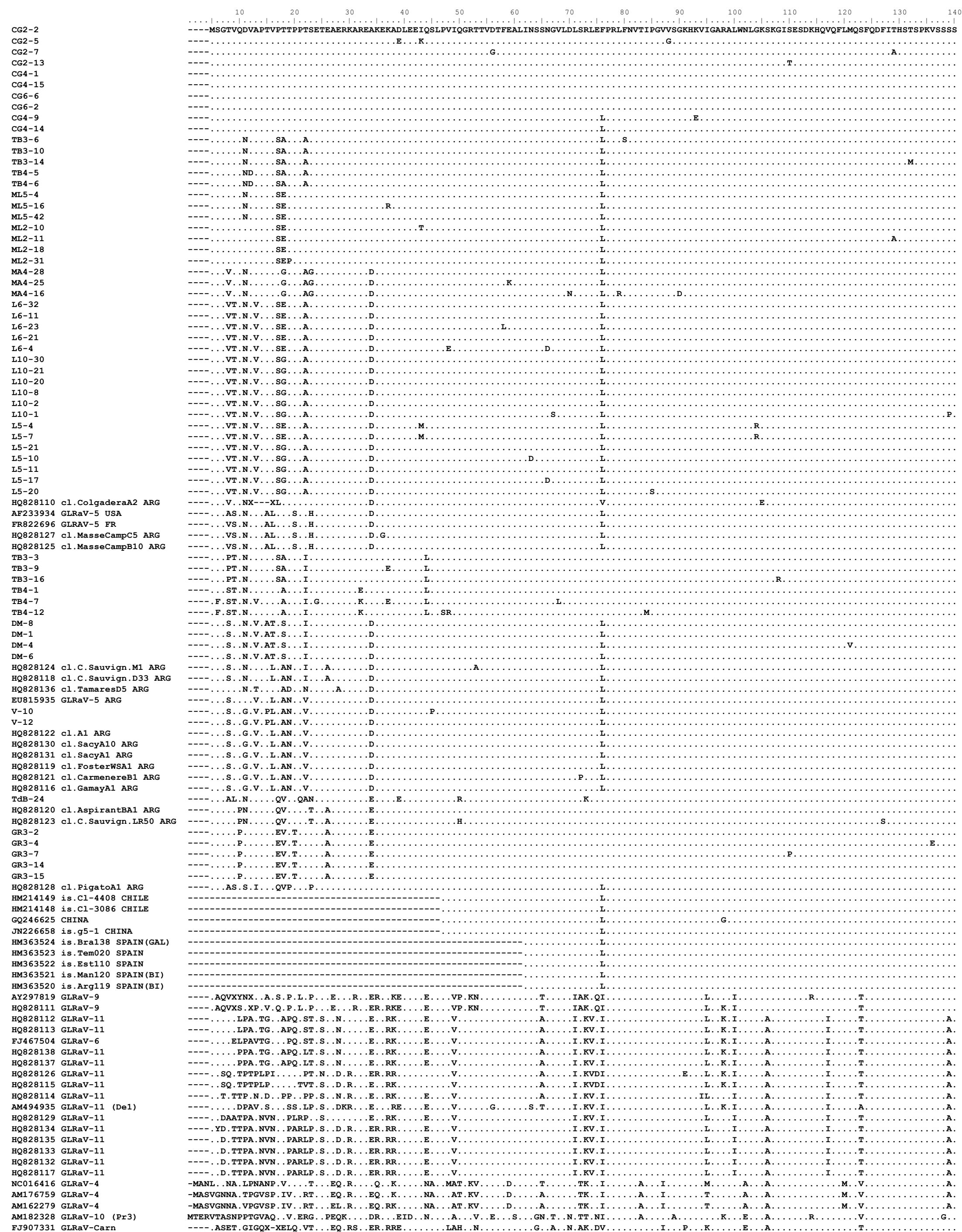

Figure 3.4 Alignment of deduced amino acid sequences of the coat proteins of GLRaV-5 and related ampeloviruses (273 AA). Sequences retrieved from GenBank are indicated by accession number and sequences obtained in this work are indicated by their isolate name. All amino acid substitutions are shown.

Figure 3.4 concluded.

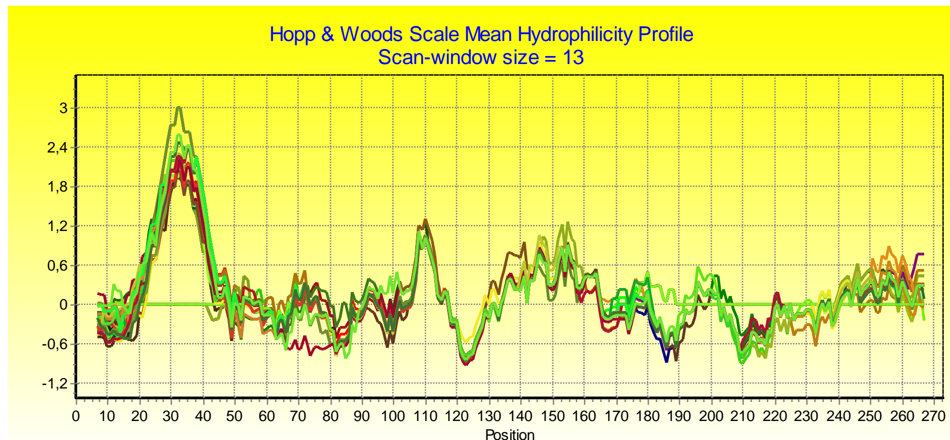

Figure 3.5 Hydrophilicity profile for the CP of GLRaV-5 and related ampeloviruses. All sequences in Fig. 3.3C are shown. The average antigenicity values are plotted versus position along the amino acid sequence. The x-axis contains 273 increments, each representing an amino acid. The y-axis represents the range of hydrophilicity values.

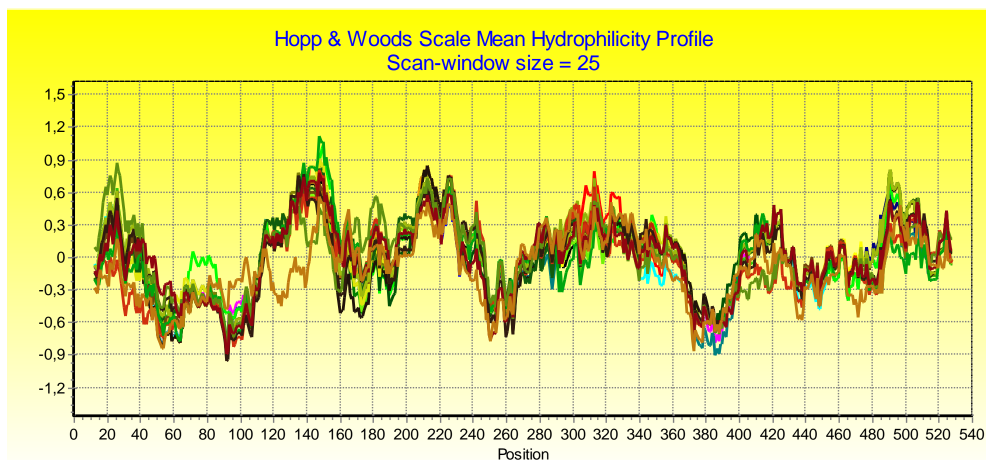

Figure 3.6 Hydrophilicity profile for the HSP90h of GLRaV-5 and related ampeloviruses. All sequences in Fig. 3.3B are shown. The average antigenicity values are plotted versus position along the amino acid sequence. The x-axis contains 539 increments, each representing an amino acid. The y-axis represents the range of hydrophilicity values.

3.4. Discussion

Knowledge of the structure and genetic variability of natural populations of RNA viruses affecting long-lived crop plants is gradually emerging (Kong *et al.*, 2000, Lozano *et al.*, 2009, Rubio *et al.*, 2001, Turturo *et al.*, 2005). These studies showed evidence that intra-host (within isolate) range of sequence variants exist as virus quasispecies (García-Arenal *et al.*, 2001), a genetic structure that appears to play an important role in shaping virus adaptability and pathogenesis (Domingo *et al.*, 2008) and act as a reservoir of emerging viral variants (Domingo *et al.*, 2008, Moya *et al.*, 2004).

This study analyzed the genetic diversity and population structure of 15 GLRaV-5 isolates obtained from field grapevines, representing the first molecular characterization of GLRaV-5 in Portugal. This work was based on sequence determination of molecular clones (partial HSP70h gene and complete HSP90h and CP genes), recognized as a valid approach for characterizing the within-isolate genetic structure of a virus (Rossinck and Schneider, 2006). The gathered data showed that most of the GLRaV-5 isolates had more than one sequence variant, regardless of the genomic region considered. Also, the within-isolate genetic diversity values determined are comparable to those found for other GLRaVs (Alabi *et al.*, 2011, Turturo *et al.*, 2005).

The heterozigosity levels determined for the CP gene revealed that distribution of variant frequency for this gene is highly variable, even when clone isolates were compared. In addition, the genetic distances between isolates from plants of clonal origin, also indicated divergence. Within-isolate sequence divergence was equally observed for the HSP70h and HSP90h, although in a smaller extent.

Regarding their GLRaV-5 variants, clonal isolates analyzed in this study appear to be in different stages of divergence dynamics. This suggests the divergence observed might be related to vegetative propagation of the host. As previously mentioned, it is reasonable to expect that the relative proportion of variants composing the virus load differ between sources of propagative plant material. New frequency combinations between variants transferred into the new isolate can eventually result in the appearance

and establishment of new sequence variants. In most of the clonal isolates (CG, L, MA, ML, TB) there is also an indication of an increase in effective population size by divergence and establishment of variants at high frequency.

Genetic diversity and phylogenetic analysis of GLRaV-5 sequences obtained in this study provide an update on the GLRaV-5 molecular variability, adding new information on virus variants and phylogroups to the one existing from other countries. The fact that nucleotide divergence is higher between the sequences set from Portugal and the set composed by sequences from other countries, than within each one of them, suggests that each represent a different population.

Phylogenetic trees constructed for the three genomic regions encoding proteins with different functions showed the tendency for sequence variants obtained in this work to group according to isolate and clonal origin. Topology of the CP based tree supported segregation of global GLRaV-5 isolates into eight major phylogroups. Estimates of genetic distances within each phylogroup are significantly lower than between phylogroups, supporting a clear separation of variants into eight lineages. Three of these lineages (groups 1, 2 and 4) were composed exclusively by variants from isolates of clone plants. Although close clustering was evidenced by isolates from clone plants in the Algarve, (groups 1 and 2), the close relationship between sequences of isolates obtained from different countries seen in groups 3, 7 and 8 suggest lack of clustering by geographical origin. This scenario in which apparently geographically unrelated variants group together is familiar to long-lived crops vegetatively propagated such as grapevine, where exchanged of infected planting material between wine-producing countries has undoubtedly occurred. Recently, it has been reported for citrus species infected with citrus tristeza virus (CTV) that traffic of infected material can overcome dispersion boundaries established by a biological vector when geographically distant regions are compared (Rubio *et al.*, 2001). The lack of correlation between clustering and geographic origin has been described for grapevine, either in virus with a known biological vector, as shown for GLRaV-1 (Alabi *et al.*, 2011) and GFLV (Vine *et al.*, 2004) or in virus with no known vector as the case of GLRaV-2 (Jarugula *et al.*, 2010).

Up to date, transmission of GLRaV-5 by insect vectors as only been observed in experiments carried out under controlled conditions (Golino *et al.*, 2002, Mahfoudhi *et*

al., 2009, Sim *et al.*, 2003, Tsai *et al.*, 2010) but not in field situations. However, in a situation where a disseminating vector exists, it can be expected to introduce some degree of spatial delineation of variants or to induce the appearance of variants from different phylogroups in a single isolate, due to serial inoculation.

The results from this study do not provide evidence of GLRaV-5 vector-mediated transmission between field-established grapevines. Phylogenetic analysis did not disclose segregation of variants from the same isolate into more than one phylogroup, except in the case of the clonal isolates TB3 and TB4 and only for the CP gene. Reduced or absence of vector transfer is also suggested by the formation of isolate-specific phylogroups for the three genomic regions analyzed in addition to the fact that variants of different isolates from the same vineyard segregate to distinct phylogroups.

Given the fact that the genetic diversity dynamics of a virus is closely related to its way of spreading, understanding a virus transmission mechanism is of great importance to understand its evolution (Roy and Brlansky, 2009) as well as to establish effective control measures. The type of transmission is also fundamental in the selection pressures acting on different viral genes (Chare and Holmes, 2004), and once the presence or absence of a biological vector is established, knowing those selection forces is of great interest in vegetatively propagated perennial crops. It allows the inference of functional constraints, determinants of vector-mediated transmission and virus-host interactions that may modulate variant divergence dynamics.

The global dN/dS value obtained for the CP gene when considering only sequence variants obtained in this study was higher than when sequences from GenBank were included, suggesting that although GLRaV-5 variants are under purifying selection, it is less evident in the Portuguese isolates. However, in either situation, dN/dS value was higher than the ones usually reported for other leafroll-associated virus with known vectors (Alabi *et al.*, 2011, Lozano *et al.*, 2009), adding further support to the suggested absence of a GLRaV-5 field vector effect. Additional evidence that groups showing different selection force signal were being compared was provided by the results of the MacDonald-Kreitman test. However, this difference might be due to the fact that although sequences retrieved from GenBank convey phylogenetic

and lineage-related information, they do not provide comparable information of intra-isolate variant divergence.

The coat protein deduced sequences showed, in the N-terminal domain, amino acid substitutions specific to each isolate, yet producing similar hydrophilic profiles. A comparable situation was observed for the HSP90h. The less evident abundance of diverging variants in HSP90h relatively to the CP suggests different gene mutation rates or different selection constraints, probably resulting from different protein function.

Since the CP potentially takes an active part in RNA structure modulation, particle formation and in virus vector transmission implying conserved amino acid sequences, it is expected a reasonable constraint level. The same happens with HSP90h, which function is accepted to be fundamental in cell-to-cell movement of the viral genome. The verification that the non-synonymous substitutions observed for the CP or the HSP90h resulted in amino acid with side-chains of analogous chemical nature suggests that this feature is crucial for protein viability and consequently a target for selection. Albeit the apparent determinant role of the CP in plant viruses' transmission, no particular conserved amino acid motifs that might imply systematic transmissibility of GLRaV-5 by vectors were found in the hydrophilic N-terminus of the CP. Molecular determinants of vector transmissibility or specificity for ampeloviruses were not identified until now, but it has been suggested (Maliogka *et al.*, 2008), based on information from the CP and partial HSP70h, that GLRaV-5 and related ampeloviruses may have reached an adaptive peak and amino acid changes that reduce suitability are removed by purifying selection. This scenario may have arisen caused by the long-term coexistence with one host (grapevine) and dissemination through grafting and vegetative propagation. The results gathered in this work support this premise and add evidence from a different protein.

The CP hydrophilic profile, with the N-terminal domain of the protein exposed on the surface of the particle is characteristic of GLRaV-5 and related ampeloviruses (Maliogka *et al.*, 2008). That information and the findings from this work on the AA substitutions accumulation in the hydrophilic region, help foreseeing the difficulty to obtain specific GLRaV-5 antibodies based on this protein. On the other hand, in the HSP90h hydrophilic profile it is possible to identify a conserved region for GLRaV-5

variants but not for related ampeloviruses, providing a potentially more suitable target for specific antibody design, leading to virus detection improvement.

The results from this study substantiate the relation between GLRaV-5 and GLRaV-4, -9, -10, -11 and -Carn. In addition, the relatedness observed between GLRaV-6 and GLRaV-11, which formed a cohesive group in the phylogenetic trees constructed in this work for the three genomic regions, raise questions about the differentiation of those viruses as two distinct species.

The assembled data on the GLRaV-5 genetic diversity from vegetatively propagated grapevines disclosed the coexistence of high frequency variants for the CP gene within a single isolate, drawing attention to the emergence of new variants by reshuffling and redistribution of virus variants sources following vegetative propagation.

The situation observed for the Portuguese isolates could be widespread because it may derive from host propagation and may have contributed to the continual dispersal of this virus. Considering this, routine monitoring of GLRaV-5, in the same way it is done for GLRaV-1 and GLRaV-3 in propagation material and field plants should be implemented.

Chapter 4. General Discussion and Conclusions

Grapevine leafroll disease (GLRD) has been reported in most grapevine-growing countries. It has a great economic impact, constituting a major limiting factor in yield and quality of grapes in crops worldwide, thus restricting the development of the reliant industries.

GLRD is one of the most complex viral plant diseases, given the high number of associated causal agents, generically named Grapevine leafroll-associated viruses (GLRaVs). This complex aetiology and its economic impact justify the investment not only in the molecular characterization of GLRaV-1 and -3, which are believed to be the most frequent GLRaVs, but also of all the putative grapevine leafroll-associated virus species. From these, however, knowledge on the genetic structure and variability of natural populations, i.e. in field situations, has evolved slowly.

Since the introduction and indiscriminate use of infected propagation material is one of the most important routes of grapevine virus dissemination, the most effective way to control GLRD is to prevent its appearance by strict observation of certification schemes, i.e. propagation and commercialization of virus-free plant material throughout the world. In most of the grapevine growing countries, strict legislation has been issued concerning this subject. Portugal has also complied with the EU directive on certification of grapevine propagation material by incorporating it into national legislation. However, to fully observe the legislative demands, it is paramount to have efficient detection and diagnostic tools that specifically identify each virus species in a given sample. Thus, population genetic studies on GLRaVs are essential to implement proficient routine detection methods.

This study gathered and analyzed information on the genetic diversity of 20 Portuguese GLRaV-1 isolates and has expanded the knowledge on the existent variants. This new data allowed comparison with previously reported results on variants from different geographical origins (Komínek *et al.*, 2005, Alabi *et al.*, 2011), and updated

global assessment of molecular variability of two genomic regions, the CP and the HSP70h.

Systematic screening for GLRaV-1 infection with commercially available antibodies, in which grapevine certification schemes rely on to certify propagation material, produced false negatives, detected through corresponding systematic screening with RT-PCR based methods. The need to improve the serological detection called for an alternative way to obtain the anti-GLRaV-1 IgG, one that would guarantee i) the capture of all the coat protein known variants, regardless of their variability, instead of the traditional immunization with a partially purified form of coat protein obtained from one isolate and ii) the absence of cross-reaction with other virus present in mixed infections, which can be frequent when using partially purified extracts for immunization. This in turn required the assembly of a suitable dataset of CP gene sequence variants and respective deduced AA sequences for *in silico* analysis, which was accomplished in this work, contributed from previously reported and newly analyzed isolates.

Results showed that the new polyclonal IgG and respective purified monospecific IgG were able to detect viral variants present in all of the Portuguese isolates from which molecular data had been obtained. These products were also adequate for routine assays with higher throughput and less expensive than ELISA, such as TPIB.

The systematic molecular screening for GLRaVs in the CAN accessions and field grown grapevines revealed an unexpected high incidence of GLRaV-5. Reports on this virus have been appearing around the world. However, the lack of information on its molecular variability, genetic structure and way of transmission in natural populations has precluded any credible attempt at developing specific detection tools and prevention measures. The 15 Portuguese GLRaV-5 isolates analyzed in this work for the HSP70h, the HSP90h and the CP genes have made a solid contribution to advance the characterization of the variability of the virus. Moreover it further documented its relatedness to the ampeloviruses in the subgroup I, adding support to its status as a cohesive taxon, at the species level, within the group. On a side note, it is interesting to observe that the data available on the three genomic regions studied raises questions about the status of GLRaV-6 as a putative species, due to its systematic

grouping with GLRaV-11. This shows how much the taxonomy of Ampelovirus is still a work in progress.

Comparison of the genetic diversity and phylogenetic analysis results for GLRaV-1 and GLRaV-5 disclosed different global situations. The fact that groups of GLRaV-1 CP variants from distinct origins showed higher nucleotide divergence values within than between them, suggests that isolates currently documented are part of the same population, and not divided between an American-Australian and an European population, as once suggested. The situation with GLRaV-5, as insofar documented, is contrasting to this. Since a higher divergence value was obtained between the group of CP gene sequences variants available at GenBank and the corresponding group from the Portuguese isolates, the observation suggests that each group is part of a different population.

Evolutionary analyses of both viruses showed evidence that GLRaV-1 dissemination can be contributed by a transmission-vector. In the case of GLRaV-5 no indication of vector-mediated transmission between field-established grapevines was found. In fact, for GLRaV-1, variants from different phylogroups were found to coexist in a single isolate, probably due to consecutive inoculation by an insect vector, but the same was not observed for GLRaV-5 isolates. In effect, up to date, transmission of GLRaV-5 by insect vectors as only been observed in experiments carried out under controlled conditions (Golino *et al.*, 2002, Mahfoudhi *et al.*, 2009, Sim *et al.*, 2003, Tsai *et al.*, 2010) but not in field situations.

For both viruses, the present study clearly indicated the contribution of vegetative propagation not only in virus dissemination but also in allowing for lesser variants to surface, by altering the frequency of established sequence variants when a cutting is taken from the mother plant.

Future research

Adding the isolates characterized in this study, more positive varieties have been detected for GLRaV-1 and GLRaV-5 in the meantime and are being analyzed,

continuing the description of the whole range of molecular variants present in Portugal. Next, biological indexing for different variants evaluating i) the symptoms developed by the host an also ii) the genetic expression profile under infection, will aid to clarify the impact of both viruses in vineyards and the degree of virulence of each variant. That will allow the identification of severe and mild variants. These in turn can be tested for the possibility of using mild strains in cross-protection, thus reducing the economic impacts caused by severe variants.

References

Abou Ghanem-Sabanadzovic, N., Sabanadzovic, S., Gugerli, P. and Rowhani, A. 2012. Genome organization, serology and phylogeny of *Grapevine leafroll-associated viruses 4* and *6*: Taxonomic implications. Virus Res. 163: 120-128.

Abou Ghanem-Sabanadzovic, N., Sabanadzovic, S., Rowhani, A. and Martelli, G. P. 2003. Multiplex RT-PCR detection of Grapevine fleck virus-like viruses in grapevine with co-amplification of control plant mRNA. In: Proceedings of the 14th Congress of ICVG, Locoronto, Italy, p. 195.

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

Agrios, G. N. 2005. Plant Pathology. 5th Edition, Academic Press, USA.

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

Alabi, O. J., Rwahnih, M. A., Karthikeyan, G., Poojari, S., Fuchs, M., Rowhani, A. and Naidu, R. A. 2011. *Grapevine leafroll-associated virus 1* occurs as genetically diverse populations. Phytopathology 101(12): 1446-1456.

Andret-Link, P. and Fuchs, M. 2005. Transmission specificity of plant viruses by vectors. J. Plant Pathol. 87(3): 153-165.

Atallah, S. S., Gómez, M. I., Fuchs, M. F. and Martinson, T. E. 2012. Economic impact of Grapevine Leafroll Disease on *Vitis vinifera* cv. Cabernet franc in Finger Lakes vineyards of New York. Am. J. Enol. Vitic. 63(1): 73-79.

Belli, G., Fortusini, A., Casati, P., Belli, L., Bianco, P. A., and Prati, S. 1994. Transmission of a grapevine leafroll associated closterovirus by the scale insect Pulvinaria vitis L. Riv. Pathol. Veg. 4: 105-108.

Bertsch, C., Beuve, M., Dolja, V. V., Wirth, M., Pelsy, F., Herrbach, E. and Lemaire O. 2009. Retention of the virus-derived sequences in the nuclear genome of grapevine as a potential pathway to virus resistance. Biol. Direct 4: 21.

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

Cassab, G. I. and Varner, J. E. 1987. Immunocytolocalization of extension in developing soybean seeds coats by immunogold-silver staining and tissue printing on nitrocellulose paper. J. Cell Biol. 105: 2581-2588.

Chare, E. R. and Holmes, E. C. 2004. Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission. J. Gen. Virol. 85: 3149-3157.

Delport, W., Poon, A. F. Y., Frost, S. D. W. and Pond S. L. K. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26(19): 2455-2457.

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

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

Domingo, E., Escarnis, C., Menéndez-Arias, L., Perales, C., Herrera, M., Novella, I. and Holland, J. J. 2008. Viral quasiespecies: dynamics, interactions and pathogenesis. In: Domindo, E., Parrish, C. and Holland, J. (Eds.), Origin and evolution of viruses. Elsevier, Amsterdam, pp. 87-118.

Elbeaino, T., Sabanadzovic, S., Digiaro, M., Abou Ghanem-Sabanadzovic, N., Rowhani, A., Kyriakopoulou, P. E. and Martelli, G. P. 2001. Molecular detection of Grapevine fleck virus-like viruses. Vitis 40: 65-68.

Engel, E. A., Escobar, P. F., Rivera, P. A. and Valenzuela, P. D. T. 2010. First report of *Grapevine leafroll-associated virus 5* in Chilean grapevines. Plant Dis. 94(8): 1067.

Engelbrecht, D. J. and Kasdorf, G. G. F. 1990a. Transmission of grapevine leafroll disease and associated closteroviruses by the vine mealybug *Planococcus ficus*. Phytophylactica 22: 341-346.

Engelbrecht, D. J. and Kasdorf, G. G. F. 1990b. Field spread of corky bark, fleck, leafroll and Shiraz decline diseases and associated viruses in South African grapevines. Phytophylactica 22: 347-354.

Engvall, E. and Perlman, P. 1971. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. Immunochemistry 8(9): 871-874.

Esteves, F., Teixeira Santos, M., Rocha, M. L., Eiras-Dias, J. E. and Fonseca F. 2009a. Genetic diversity of Portuguese isolates of *Grapevine leafroll-associated virus 1* (GLRaV-1) based on the capsid protein. In: Proceedings of the 16th Congress of ICVG, Dijon, France, pp. 310-311.

Esteves, F., Teixeira Santos, M., Rocha, M. L., Eiras-Dias, J. E. and Fonseca F. 2009b. Genetic variability of *Grapevine leafroll-associated virus 2* (GLRaV-2) based in a portuguese vineyard. In: Proceedings of the 16th Congress of ICVG, Dijon, France, pp. 312-313.

Esteves, F., Teixeira Santos, M. L., Eiras-Dias, J. E. and Fonseca F. 2012a. Novel antibodies for ISIA detection of *Grapevine leafroll-associated virus 1* (GLRaV-1) based on the variability of the capsid protein gene. In: Proceedings of the 17th Congress of ICVG, UCDavis, USA, pp. 59-60.

Esteves, F., Teixeira Santos, M., Eiras-Dias, J. E. and Fonseca F. 2012b. Occurrence of *Grapevine leafroll-associated virus 5* in Portugal: genetic variability and population structure in field-grown grapevines. Arch. Virol. 157(9): 1747-1745.

Faoro, F., Tornaghi, S., Cinquanta, G. and Belli, G. 1992. Cytopathology of grapevine leafroll associated virus III (GLRaV-III). Riv. Patol. Veg. 2: 67-83.

Fazeli, F. C. and Rezaian, M. A. 2000. Nucleotide sequence and organization of ten open reading frames of the *Grapevine leafroll-associated virus 1* genome and identification of three subgenomic RNAs. J. Gen. Virol. 81: 605-615.

Fortusini A., Scattini G., Prati S., Cinquanta S. and Belli G. 1997. Transmission of G*rapevine leafroll-associated virus 1* (GLRaV-1) and *Grapevine virus A* (GVA) by scale insects. Proceedings of 12th Meeting of ICVG, Lisbon, Portugal. pp. 121-122.

Garau, R., Prota, V. A., Boscia, D., Fiori, M., and Prota, U. 1995. Pseudococcus affinis Mask., new vector of grapevine trichoviruses A and B. Vitis 34: 67-68.

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

Gelderblom, H. R. 1996. Structure and classification of viruses. In: Baron, S. (Ed.), Medical Microbiology. 4th edition, Galveston (TX): University of Texas Medical Branch at Galveston.

Goheen, A. C., Herman, F. N. and Weinberger, J. H. 1958. Leafroll (White Emperor disease) of grapes in California. Phytopathology 48: 51-54.

Goheen, A. C. 1988. Leafroll. In: Pearson R.C. and A. C. Goheen (Eds.), Compendium of Grape Diseases, pp. 52. American Phytopathological Society Press, St. Paul, Minnesota, USA.

Golino, D. A. 1993. Potential interactions between rootstocks and grapevine latent viruses. Am. J. Enol. Vitic. 44: 148-152.

Golino, D. A., Sim, S. T., Gill, R. and Rowhani, A. 2002. California mealybugs can spread grapevine leafroll disease. Calif. Agric. 56: 196-201.

Goszczynski, D. E., Kasdorf, G. G. F., Pieterson, G. and van Tonder, H. 1996. Detection of two strains of *Grapevine leafroll-associated virus 2*. Vitis 35: 133-135.

Gray, S. M. and Banerjee, N. 1999**.** Mechanisms of arthropod transmission of plant and animal viruses. Microbiol. Mol. Biol. Rev. 63: 128–148.

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

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

Hasegawa, M., Kinshino, H. and Yano, T.A. 1985. Dating of the human-ape splitting by a molecule clock of a mithocondrial DNA. J. Mol. Evol. 22: 160-174.

Hewitt, W. B. 1968. Viruses and virus-like diseases of grapevine. Rev. Appl. Mycol. 47: 433-455.

Hopp, T. P. and Woods, K. R. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. U.S.A. 78: 3824-3828.

Hull, R. 2009. Comparative Plant Virology. 2nd Edition, Academic Press, USA.

ICTV. 2011. http://ictvonline.org/. Quoted in 04/11/2012.

ICVG. 2003. Recommendation for the certification of grapevine propagating material. http://www.icvg.ch/data/recomm.pdf. Quoted in 04/11/2012.

Jarugula, S., Alabi, O. J., Martin, R. R. and Naidu, R. A. 2010. Genetic variability of natural population of *Grapevine leafroll-associated virus 2* in Pacific Northwest Vineyards. Phytopathology 100: 698-707.

Katis, N. I., Tsitsipis, J. A., Stevens, M. and Powell, G. 2007. Transmission of Plant Viruses. In: Van Emden, H. F. and Harrington, R (Eds.), Aphids As Crop Pests, pp. 353-372. CABI, Oxfordshire, UK.

Karasev, A. V. 2000. Genetic diversity and evolution of closteroviruses. Annu. Rev. Phytopathol. 38: 293-324.

Komínek, P., Glasa, M. and Bryxiová, 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. and Falk, B. W. 2000. Population structure and genetic diversity within California *Citrus tristeza virus* (CTV) isolates. Virus Genes 21: 139-145.

Kosakovsky Pond, S. L. and Frost, S. D. W. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22(5): 1208-1222.

Kosakovsky Pond, S. L., Frost, S. D. W., Grossman, Z., Gravenor, M. B., Richman, D. D. and Brown, A. J. L. 2006a. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Computational Biology 2(6): e62.

Kosakovsky Pond, S. L, Posada, D., Gravenor, M. B., Woelk, C. H. and Frost, S. D. 2006b. GARD: a genetic algorithm for recombination detection. Bioinformatics 22: 3096-3098.

Librado, P. and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lin, Y., Rundell, P. A. Xie, L. and Powell, C. A. 2000. In situ immunoassay for detection of *Citrus tristeza virus*. Plant Dis. 84: 937-940.

Lozano, G., Grande-Pérez, A. and Navas-Castillo, J. 2009. Populations of genomic RNAs devoted to replication or spread of a bipartide plant virus differ in genetic structure. J. Virol. 83: 12973-12983.

MacKenzie, D. J., McLean, M. A., Mujerki, S. and Green, M. 1997. Improved RNA extraction from woody plants for the detection of viral pathogens by reverse transcription-polymerase chain reaction. Plant Dis. 81: 222-226.

Mahfoudhi, N., Digiaro, M. and Dhouibi, M. H. 2009. Transmission of grapevine leafroll viruses by *Planococcus ficus* (Hemiptera: *Pseudococcidae*) and *Ceroplastes rusci* (Hemiptera: *Coccidae*). Plant Dis. 93: 999-1002.

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

Martelli G. P. 1993. Leafroll. In: Martelli G. P. (Ed.), Graft-transmissible diseases of grapevines. Handbook for Detection and Diagnosis, pp. 37-44. FAO Publication Division, Rome, Italy.

Martelli, G. P. 2012. Grapevine virology highlights: 2010–2012. In: Proceedings of the 17th Congress of ICVG, Davis, California, USA, pp. 13-31.

Martelli, G. P., Abou Ghanem-Sabanadzovic, N., Agranovsky, A. A., Al Rwahnih, M., Dolja, V. V., Dovas, C., Fuchs, M., Gugerli, P., Hu, J. S., Jelkmann, W., Katis, N. I., Maliogka, V. I., Melzer, M. J., Menzel, W., Minafra, A., Rott, M. E., Rowhani, A., Sabanadzovic, S. and Saldarelli, P. 2012. Taxonomic revision of the family Closteroviridae with special reference to the grapevine leafroll-associated members of

the genus Ampelovirus and the putative species unassigned to the family. Journal of Plant Pathol. 94: 7-19.

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

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

Martelli, G.P. and Boudon-Padieu, E. 2006. Directory of infectious diseases of grapevines. In: Martelli, G.P., Boudon-Padieu, E. (Eds.), Options Mediterraneennes, vol. 55. CIHEAM, Bari.

Martelli, G. P. and Savino, V. 1990. Fanleaf degeneration. In: Pearson R.C. and Goheen A. (Eds.), Compendium of grape diseases. APS Press, St. Paul, MN, USA.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D. and Lefeuvre, P. 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462-2463.

Matsumoto, T. and Ohki, S. T. 1998. A possible new necrotic disease of grapevine associated with small isometric particles and novel membrane bound particles. Ann. Phytopathol. Soc. Japan 64: 560-564.

Matthews, R. E. F. 1991. Plant Virology. 3rd Edition, Academic Press, Inc., San Diego, CA.

Maule, A. J. and Wang, D. 1996. Seed transmission of plant viruses: a lesson in biological complexity. Trends Microbiol. 4: 153–158.

McDonald, J. L. and Kreitman, M. 1991. Adaptative protein evolution at the Adh locus in Drosophila. Nature 351: 652-654.

Mink, G. I. 1993. Pollen- and seed-transmitted viruses and viroids. Annu. Rev. Phytopathol. 31: 375-402.

Monis, J. and Bestwick, R. K. 1996. Detection and localization of grapevine leafroll associated closteroviruses in grennhouse and tissue culture grown plants. Am. J. Enol. Vitic. 47(2): 199-205.

Moore, J. D., Boyle, J. S. and Keit, G. W. 1948. Mechanical transmission of a virus disease to cucumber from sour cherry. Science 168: 623-624.

Moya, A., Holmes, E. C. and González-Candelas, F. 2004. The population genetics and evolutionary epidemiology of RNA viruses. Nat. Rev. Microbiol. 2: 279-288.

Nei, M. 1978. Estimation of average heterozigosity and genetic distance from a small number of individuals. Genetics 89: 583-590.

OIV (International Organisation of Vine and Wine). 2012. In: OIV Statistical Report on World Vitiviniculture 2012.

Orita, M., Iwahana, H., Kanazawa, K. and Sekiya, T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single stranded conformation polymorphisms. Proc. Natl. Acad. Sci. 86: 2766-2770.

Padilla, C.V., Cretazzo, E., Hita, I., López, N., Padilla, V. and Velasco, L. 2010. First Report of *Grapevine leafroll-associated virus 5* in Spain. Plant Dis. 94(12): 1507.

Pei, G. –Q., Dong, Y. -F, Zhang, Z. –P. and Fan, X. –D. 2010. First Report of *Grapevine leafroll-associated virus 4 and 5* in Grapevines in China. Plant Dis. 94(1): 130.

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

Rosciglione, B., Castellano, M. A., Martelli, G. P., Savino, V. and Cannizzaro, G. 1983. Mealybug transmission of *Grapevine virus A*. Vitis 22: 331-347.

Rosciglione, B. and Gugerli, P. 1989. Transmission of grapevine leafroll disease and an associated closterovirus to healthy grapevine by the mealybug Planococcus ficus Signoret. In: Proceedings of the 9th Meeting of ICVG, Kiryat Anavim, Israel, pp. 67-69.

Rossynck, M. J. and Schneider, W. L. 2006. Mutant clouds and occupation of sequence space in plant RNA viruses. Curr. Top. Microbiol. Immunol. 299: 337-348.

Rowhani, A. and Golino, D. A. 1995. ELISA test reveals new information about leafroll disease. Calif. Agr. 49(1): 26-29.

Roy, A. and Brlansky, R. H. 2009. Population dynamics of a Florida *Citrus tristeza virus* isolate and aphid-transmitted subisolates: identification of the three genotypic groups and recombinants after aphid transmission. Phytopathology 99: 1297-1306.

Rubio, L., Ayllón, M. A., Kong, P., Fernández, A., Polek, M., Guerri, J., Moreno, P. and Falk, B. W. 2001. Genetic variation of *Citrus tristeza virus* isolates from California and Spain: evidence for mixed infections and recombination. J. Virol. 75: 8054-8062.

Saiki, R., Gelfand, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., **Mullis, K**. and Erlich, H. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239(4839): 487–491.

Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270(5235): 467–470.

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

Sim, S. T., Rowhani, A. and Golino, D. A. 2003. Experimental transmission of *Grapevine leafroll-associated virus 5* and *9* by longtailed mealybugs. In: Proceedings of the 14th Congress of ICVG, Locorotondo, Italy, pp. 211-212.

Syvanen, A. C., Bengtstrom, M., Tenhunen. J. and Soderlund, H. 1988. Quantification of polymerase chain reaction products by affinity-based hybrid collection. Nucleic Acids Res. 16: 11327-11338.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

Talquenca S.G, Muñoz C., Grau O. and Gracia, O. 2009. First description of *Grapevine leafroll-associated virus 5* in Argentina and partial genome sequence. Virus Genes 38: 184-186.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol. 28(10): 2731-2739.

Tanne, E., Ben-Dov, Y. and Raccah, B. 1989. Transmission of closterolike particles associated with grapevine leafroll by mealybugs (*Pseudococcidae*) in Israel. In: Proceedings of the 9th Congress of ICVG, Kiryat Anavim, Israel, pp. 71-73.

Teixeira Santos, M., Esteves, F., Rocha, M. L., Eiras-Dias, J. E. and Fonseca, F. 2009a. Comparison of phylogenetic inference based on HSP70, HSP90 and CP genes of GLRaV-3: implications for detection and taxonomy. In: Proceedings of the 16th Congress of ICVG, Dijon, France, pp. 276-277.

Teixeira Santos, M., Rocha, M. L., Guedes da Silva A., Esteves, F., Fonseca, F. and Eiras-Dias, J. E. 2009b. Old grapevine varieties vineyards: a window over a pre sanitary selection era and a source of viruses. In: Proceedings of the 16th Congress of ICVG, Dijon, France, pp. 62-63.

Thompson, J. R., Fuchs, M. and Perry, K. L. 2012. Genomic analysis of *Grapevine leafroll-associated virus 5* and related viruses. Virus Res. 163(1): 19.27.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Towbin, H., Staehelin, T. and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Nat. Acad. Sci. USA 76(9): 4350–54.

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

Tsai, C. -W., Rowhani, A., Golino, D. A., Daane, K. M. and Almeida, R. P. P. 2010. Mealybug transmission of grapevine leafroll viruses: An analysis of virus–vector specificity. Phytopathology 100: 830-834.

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

Uyemoto, J. K., Martelli, G. P. and Rowhani, 2009. Grapevine Viruses, Viruslike Diseases and Other Disorders. In: Virus diseases of plants: grape, potato, and wheat image collection and teaching resource CD-Rom. APS Press, St. Paul, MN.

Vigne, E., Bergdoll, M., Guyader, S. and Fuchs, M. 2004. Population structure and genetic variability within isolates of *Grapevine fanleaf virus* from a naturally infected vineyard in France: evidence for mixed infection and recombination. J. Gen. Virol. 85: 2435-2445.

Walter, B. and Martelli, G.P. 1997. In: Walter B. (Eds.), Sanitary Selection of the Grapevine: protocols for detection of virus and virus-like diseases. INRA, Paris, France.

Weber, E., Golino, D. and Rowhani, A. 1993. Leafroll disease of grapevines. PWV Journal 13: 21-25.