UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRARIA, ALIMENTARIA Y DE BIOSISTEMAS

Departamento de Biotecnología-Biología Vegetal

Centro de Biotecnología y Genómica de Plantas (UPM-INIA)

Analysis of fitness trade-offs limiting host range expansion in pepper-infecting tobamoviruses

Ph.D. Thesis:

SAYANTA BERA

Bachelor in Microbiology, M.Sc. in Life Sciences, BRAVE grantee

2018

UNIVERSIDAD POLITÉCNICA DE MADRID
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRARIA, ALIMENTARIA Y DE BIOSISTEMAS
Departamento de Biotecnología-Biología Vegetal
Centro de Biotecnología y Genómica de Plantas (UPM-INIA)

Ph.D. Thesis:

Analysis of fitness trade-offs limiting host range expansion in pepper-infecting tobamoviruses

Author:

Sayanta Bera
Bachelor in Microbiology, M.Sc. in Life Sciences.

Director:

Fernando García-Arenal Rodríguez
Dr. Ingeniero Agrónomo

Madrid, 2018

Table of Contents

Acknowledgements ..... **ix**

List of Abbreviations ..... **xi**

Summary ..... **xiii**

Resumen ..... **xv**

1 INTRODUCTION **1**

1.1 Importance of plant diseases in crops **3**

1.1.1 **Viral diseases in plants** 4

1.2 Host range expansion **6**

1.3 Fitness trade-offs across host species **7**

1.4 Fitness trade-offs across host genotypes **9**

1.4.1 **Costs of overcoming resistance** 10

1.5 Antagonistic pleiotropy as a cause of across-host fitness trade-offs **14**

1.6 Epistasis as a modulating factor of fitness trade-offs **17**

1.7 Pathosystems studied **20**

1.7.1 **The viruses** 20

1.7.2 **The host plants** 28

2 OBJECTIVES **31**

3 MATERIALS AND METHODS **35**

3.1 Viruses and Plants **37**

3.1.1 **Viral isolates and infectious cDNA clones** 37

3.1.2 **Site-directed mutagenesis and characterization of the resulting mutant genotypes** 38

3.1.3 **Host plants** 43

3.2 Construction of an infectious clone and generation of RNA transcripts **43**

3.2.1 ***In vitro*** **transcription** 44

3.3 Purification of viral particles and nucleic acids **45**

3.3.1 **Purification of viral particles** 45

3.3.2 **Extraction of encapsidated viral RNA** 45

3.4 Inoculation of plants and extraction of total RNA for detection and quantification of viral RNA **46**

3.4.1 **Inoculations** 46

3.4.2 **Extraction of total RNA from plant tissue** 46
3.4.3 **Detection of viral RNA by reverse transcription – polymerase chain reaction (RT- PCR)** 47
3.4.4 **Quantification of viral multiplication** 47
3.5 Calculation of virus fitness, epistasis and virulence **48**
3.5.1 **Virus fitness** 48
3.5.2 **Epistasis** 48
3.5.3 **Virulence** 49
3.6 *In vitro* stability of virus particles **49**
3.7 Determination of nucleotide sequences **49**
3.8 Estimation of genetic diversity **49**
3.9 Statistical analyses **50**
4 RESULTS **53**
4.1 Adaptation of Tobacco mild green mosaic virus isolates to different host species **55**
4.1.1 **Genetic diversity of TMGMV isolates** 55
4.1.2 **Multiplication of TMGMV isolates in pepper and in *N. glauca*** 58
4.1.3 **Analysis of the role of dimorphic loci in the coat protein and 3' untranslated region in the multiplication of TMGMV isolates in pepper and *N. glauca*** 61
4.1.3.1 Effect of the dimorphism at position 8 of the coat protein 61
4.1.3.2 Effect of the dimorphism at position 148 of coat protein 65
4.1.3.3 Effect of the dimorphism at 3' untranslated region 68
4.1.4 **Variation of the 3' untranslated region length during virus multiplication in *N. glauca*** 70
4.1.5 **Accumulation in pepper or *N. glauca* of TMGMV genotypes derived from biologically active cDNA clones** 72
4.1.5.1 Molecular characterization of TMGMV genotypes derived from biologically active cDNA clones 72
4.1.5.2 The effect of CP and 3'UTR polymorphisms on viral accumulation in pepper and N. glauca of TMGMV genotypes derived from biologically active cDNA clones 73
4.2 Pleiotropic effects of resistance- breaking mutations on different fitness components of Pepper mild mottle virus **77**
4.2.1 **Characterization of different mutant genotypes of PMMoV** 77
4.2.2 **Pleiotropic effects of resistance-breaking mutations on the multiplication of PMMoV genotypes in susceptible hosts** 78

4.2.3 **Epistasis between resistance-breaking mutations on the multiplication of PMMoV genotypes in susceptible hosts** 82
4.2.4 **Virulence of different PMMoV genotypes in susceptible hosts** 83
4.2.5 **Effect of resistance-breaking mutations on virus particle stability** 86

5 DISCUSSION 89

5.1 Are the isolates of *Tobacco mild green mosaic virus* adapted to their host of origin? 92

5.2 Role of mutations in the coat protein and 3' untranslated region of *Tobacco mild green mosaic virus* in determining virus multiplicaiton in pepper and *N. glauca* 93

5.3 Does resistance-breaking genotypes of *Pepper mild mottle virus* cause fitness penalties in susceptible host genotypes? 97

5.4 Do resistance-breaking mutations of *Pepper mild mottle virus* affect virulence? 100

5.5 Are there trade-offs between different components of virus fitness? 100

6 CONCLUSIONS 105

7 REFERENCES 109

8 ANNEXES 123

ANNEX I: 125

ANNEX II: 135

Acknowledgements

Foremost, I would like to express sincere gratitude to my advisor Prof. Fernando García-Arenal for the continuous support of my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. Otherwise, this PhD would not have been achievable. The constructive criticism from him allowed me to expand the horizons of my knowledge and nurtured my excitement to study the evolution of pathogens. I could not have imagined having a better advisor and mentor for my PhD study.

I gratefully acknowledge the scholarship received towards my PhD from the Erasmus Mundus program, BRAVE (BReeding for plAnt resistance to VirusEs). Thanks to Prof. Supriya Chakraborty for his encouragement who convinced me to apply for this fellowship and also Prof. Ashwani Pareek for giving me the freedom in master's degree to work independently.

I would like to thank Prof. Aurora Fraile, Dr Tetsuro Okuno, Dr Masanori Kaido and Dr Manuel G. Moreno-Pérez for providing me field isolates and biologically active full-length cDNA clones in a generous quantity for my work. I would like to offer my special thanks to Prof. Fraile and Dr Moreno-Pérez for assisting and teaching me the classical techniques in the field of plant virology. The advices from them during lab meetings or in general discussions have been a great help to develop my critical thinking.

Big thanks to Aayushi Shukla for being a very helpful friend, labmate and flatmate. Her successful endeavour of learning a new language, Spanish, was of immense help and was happily exploited by me to fill up several official forms. With her on my side the experience of doing a research away from home was surprisingly comfortable.

I greatly appreciate the assistance provided by Dr. Michael McLeish for doing statistical analysis related to my work. He also coaxed me to use R which helped me eventually in the workshop held in ETH Zurich on Ecological genomics of coevolutionary interactions where I learnt the use of several modeling techniques through R.

This PhD study would not have been possible without the corporation and support extended by the present and ex lab members so, a heartfelt thank you to Israel, Maria Angeles, Soldedad, Cristina, Adrián, Carlos, Viji, Shirin, Ravi, Nuria, Ricky, Rajesh, Awdhesh, Ayo, Gokhan and Enrique.

A very special thank you to Miguel and Antolin for their technical assistance and planting hundreds of plants in greenhouse. Because of their support the work load decreased by many folds and allowed me to invest time in doing more analysis.

I would like to acknowledge Sanjay Swami for being with me through thick and thin for past 6 years and thus, overcoming various challenges together. His spontaneous attitude of doing any task is what I cherish most. I could rely on him for some much-needed respite when the experiments were not going as per the plans. This goes unsaid that he helped me a lot to assist in some experiments when time was critical. In short, doing research in a foreign lab is a challenge in itself, nevertheless, together with Aayushi and Sanjay, we managed to have lots of fun and some adventure also.

Lastly, I am filled with gratitude for my parents for believing in me that I can take this challenging and exciting life and, at the same time praying constantly for my good progress. I am indebted to my brother, Siddhartha Bera and sister-in-law, Aprajita Bera for their endless love, support and care. I will cease this opportunity to thank my niece, Risha Bera, who was born during my doctoral study which gave me the opportunity to take an impromptu break and to be with my family in India.

List of Abbreviations

aa	amino-acid
AVR	avirulence gene
BPeMV	Bell pepper mottle virus
BYDV	Barley yellow dwarf virus
CI	cylindrical inclusion
CMV	Cucumber mosaic virus
CP	Coat Protein
CP_148	Mutation at position 148 of TMGMV's CP
CP_8	Mutation at position 8 of TMGMV's CP
DMFE	distribution of mutational fitness effects
dpi	Days post inoculation
E.C	Extinction Coefficient
GFG model	gene-for-gene model
gRNA	genomic RNA
GzLM	Generalized Linear Model
HCSV	Hibiscus chlorotic spot virus
LSD	least significant difference
ML	Maximum-Likelihood
Ng	*Nicotiana glauca*
NIa	nuclear inclusion A
NIb	nuclear inclusion B
nll	necrotic local lesions
nt	nucleotides
ObPV	Obuda pepper virus
ORF	open reading frame
P	Pepper
PaMMV	Paprika mild mottle virus
PFBV	Pelargonium flower break virus
PMMoV	Pepper mild mottle virus
PPV	Plum pox virus
PVX	Potato virus X
PVY	Potato virus Y
RRSV	Rasp- berry ringspot virus
RT-qPCR	Reverse transcription – quantitative polymerase chain reaction

sg RNA	Subgenomic RNA
TEV	Tobacco etch virus
TICV	Tomato infectious chlorosis virus
TMGMV	Tobacco mild green mosaic virus
TMV	Tobacco mosaic virus
ToMV	Tomato mosaic virus
tRNA	Transfer RNA
TuMV	Turnip mosaic virus
UTR	Untranslated region
VIDE database	Virus Identification Data Exchange database
WT	Wild Type

Summary

The acquisition of new hosts, or host range expansion, provides a virus with more opportunities for transmission and survival, but may be limited by across-host fitness trade-offs. A major cause of across-host trade-offs in viruses is antagonistic pleiotropy, that is, different phenotypic effects of mutations in different host environments, a Genotype x Environment interaction. In addition, epistasis (a Genotype x Genotype interaction) also may play a role in generating across-host fitness trade-offs. Relevant cases of host range expansion are the emergence of viruses in new host species and resistance-breaking (RB), in which viruses acquire the capacity to infect otherwise resistant plant genotypes. Negative effects of resistance-breaking mutations on within-host virus multiplication have been documented for several plant viruses. However, understanding virus evolution requires analyses of potential trade-offs between different fitness components. In this work we have analysed if there are trade-offs, and what are the causes, associated to the acquisition of new host species or genotypes in tobamoviruses that infect pepper crops.

First we explored if there are trade-offs associated to the adaptation of *Tobacco mild green mosaic virus* (TMGMV) to its two major hosts in SE Spain. Field isolates of TMGMV from pepper and *Nicotiana glauca* were molecularly characterised and inoculated in the homologous and heterologous hosts. Results showed that TMGMV isolates from *N. glauca* were adapted to their host, but that pepper isolates were not adapted to pepper. The role of the amino acid at two dimorphic positions in the coat protein (CP), and of the duplication of a region of the 3'UTR, in the differential multiplication of the virus in pepper and *N. glauca* was analysed in the collection of field isolates and through the introduction of mutations in biologically active cDNA clones derived from two TMGMV isolates. Results showed that the effect of these dimorphisms on virus multiplication depended on higher order interactions involving pleiotropy and epistasis, of the type G x G x E.

Fitness costs associated with RB were studied in *Pepper mild mottle virus* (PMMoV) by introducing in a biologically active cDNA clone all reported mutations in the CP determining RB of alleles L^3 and L^4 in pepper. The parental and mutant

genotypes were assayed in different susceptible pepper genotypes, and the within-host multiplication component of fitness, and virulence, were quantified. Within-host fitness did not depend on the virus capacity to overcome a specific resistance allele, but depended on the specific mutation and the genotype of the susceptible host. In addition there was evidence of epistatic interactions between mutations leading to RB, which again depended on the genotype of the susceptible host. Thus, results showed higher order interactions among RB mutations of the type G x G x E. Similar results were obtained relative to the effects of RB mutations in virulence. We also analysed whether coat protein mutations affect particle stability and, thus, survival in the environment. The various RB mutations had effects on particle stability and, thus, survival. The stability of the particles of the different mutants was not correlated with within-host fitness or virulence, thus indicating no trade-offs among different life history traits of the virus that might constraint its evolution.

Taken together the results of this Thesis, show a major general conclusion: that host range evolution in the analysed tobamoviruses will be constrained by higher order interactions between host-range mutations on virus fitness, but not by trade-offs between different fitness components.

Resumen

La adquisición de nuevos huéspedes, es decir, la expansión de su gama de huéspedes, proporciona a un virus más oportunidades de transmitirse y sobrevivir, pero puede estar limitada por compromisos de eficacia entre los distintos huéspedes. Una causa principal de compromisos de eficacia entre huéspedes es la pleiotropía antagonista, es decir, que las mutaciones tengan efectos fenotípicos contrarios en los distintos huéspedes, una interacción Genotipo x Entorno. La epistasia, es decir una interacción Genotipo x Genotipo, puede ser también causa de compromisos de eficacia entre huéspedes. Como casos relevantes de la expansión de la gama de huéspedes podemos citar la emergencia de virus en nuevas especies de huéspedes, y la superación de la resistencia, por la que el virus adquiere la capacidad de infectar a genotipos de huésped previamente resistentes. Se ha documentado en distintos virus que la superación de la resistencia puede comportar efectos negativos en la multiplicación del virus en un huésped susceptible. Sin embargo, a penas se han explorado sus efectos en otros componentes de la eficacia del virus, y entender la evolución de los virus requiere conocer los posibles compromisos entre distintos componentes de la eficacia. En este trabajo hemos analizado si existen compromisos de adaptación asociados a la adquisición de nuevas especies o genotipos de huéspedes, y cuáles son sus causas. Para ello nos hemos centrado en los tobamovirus que infectan a los cultivos de pimiento.

En primer lugar hemos analizado si existen compromisos de adaptación del virus del mosaico verde atenuado del tabaco (*Tobacco mild green mosaic virus*, TMGMV) a sus dos huéspedes principales en el SE de España. Se han caracterizado molecularmente aislados de campo procedentes de los dos huéspedes, pimiento y *Nicotiana glauca*, que se han inoculado en los huéspedes homólogo y heterólogo. El análisis de la multiplicación de estos aislado muestra que los aislados de *N. glauca* están adaptados a su huésped, lo que no ocurre con los aislados de pimiento. Se analizó el papel en la multiplicación diferencial en ambos huéspedes de los amino ácidos de dos posiciones polimórficas de la proteína de la cápsida (CP) y de una duplicación en la región 3' no codificante (3'UTR). El análisis se realizó en la colección de aislados de campo y por manipulación de clones de cDNA derivados de dos aislados de TMGMV. Los resultados muestran que el efecto de los tres dimorfismos

en la multiplicación viral depende de interacciones de orden superior que implican pleiotropía y epistasia, del tipo G x G x E.

Los costes de eficacia asociados a la superación de la resistencia se han estudiado en el virus del moteado atenuado del pimiento (*Pepper mild mottle virus*, PMMoV) introduciendo en un clon de cDNA activo biológicamente todas las mutaciones conocidas que determinan la superación de la resistencia conferida por los alelos L^3 y L^4 del pimiento. Los genotipos parental y mutantes se ensayaron en distintos huéspedes susceptibles, y se cuantificó el componente de eficacia asociado a multiplicación viral, y la virulencia. La eficacia intrahuésped no dependía de la capacidad de virus de superar los distintos alelos de resistencia sino de la mutación en cuestión, y del genotipo del huésped susceptible. Los resultados señalaron también que existen interacciones epistáticas entre las mutaciones de superación de la resistencia, que varían según el genotipo del huésped susceptible. Por tanto, los resultados muestran de nuevo interacciones de orden superior del tipo G x G x E entre las mutaciones. Se obtuvieron resultados similares en cuanto a la relación de las mutaciones de superación de la resistencia y la virulencia. Se analizó también si las mutaciones de superación de la resistencia, que se localizan en la CP, afectan a la estabilidad de la partícula viral y, por tanto, a su supervivencia en el medio. Las mutaciones afectaron a la estabilidad de la partícula. La estabilidad de las partículas de los distintos mutantes no se correlacionó con la eficacia intrahuésped ni con la virulencia, lo que indica que no hay compromisos entre distintas componentes de la historia vital del virus que puedan comprometer su evolución.

En conjunto los resultados de esta Tesis permiten alcanzar una conclusión general importante: que la evolución de la gama de huéspedes en los tobamovirus estará limitada por interacciones de orden superior entre las mutaciones de gama de huéspedes, pero no por compromisos entre los distintos componentes de la eficacia biológica.

1 INTRODUCTION

1.1 Importance of plant diseases in crops

Plants are unique and essential higher organism on which both animals and humans are dependent for their survival. Pathogens that cause diseases in plants have a high socioeconomic impact on human life directly or indirectly. An analysis done by Oerke and Dehne (2004) reported the amount of crop losses due to diseases, as percentage of the actual production, for eight primary food and cash crops, which were: Wheat (25%), Rice (18%), Maize (14%), Barley (12%), Potatoes (30%), Soybean (12%), Sugar beet (21%) and cotton (11%). Worldwide, it has been estimated that plant diseases led to an annual loss of approximately $220 billion (Agrios, 2005) and alone viral diseases are estimated to cause a loss of approximately $30 billion (Sastry and Zitter, 2014). On a different note, the world's human population is expected to rise from 7.6 to about 9.8 billion by 2050 (UN, 2017). Based on projected population growth rates, by the middle of this century, the world will need nearly twice as much food as it currently does. Therefore, food production has to increase at a time when global climate change is likely to make some areas of the planet far less productive than at present due to declining availability of land, water and nutrient supplies (e.g., Tilman *et al*., 2002). Hence, it would be reasonable to say that we should minimize plant production losses due to plant diseases.

What drives the outbreak of diseases in crop plants? As pointed in Fraile *et al*., (2017), plant pathologists have traditionally considered that three factors influence the high impact of diseases in agroecosystems as compared with wild ecosystems: i) the reduced diversity of species, ii) the higher density of crop plants in fields and, iii) the reduced genetic diversity of crops. Thus, cultivation had been found to be intrinsically linked to a decrease in habitat diversity and an increase in host plant density (Stukenbrock and McDonald, 2008) which paves the way for disease outbreaks. Moreover, the sophistication of modern agriculture where genetically homogeneous varieties of plants are grown, which reduces the genetic diversity of potential plant hosts, has been considered as a key factor of pathogen emergence in crops (Jones *et al*., 2008; Roossinck and García-Arenal, 2015). Also, the global expansion of trade in plant products which cause the introduction of foreign plant pathogens to new areas might quickly change to an epidemic. Thus, anthropogenic changes to the environment, which include but are not limited, to introductions, agricultural techniques, or habitat disturbance, are major driving forces for the emergence of plant diseases. The study of pathogens in wild plant communities and in crops at the interface between indigenous vegetation and cultivated areas has the potential to provide critical information not only on disease threats to biodiversity and cultivated species but also on pathogen evolution (Lovisolo *et al*., 2003; Jeger *et al*., 2006; Webster *et al.*, 2007; Alexander *et al.*, 2014). It is also increasingly necessary to address

questions related to pathogen movement between natural and managed ecosystems and the relative threats posed by introduced and native pathogens. This is because the border between natural and agricultural communities (agro-ecological interface) is surprisingly permeable and biologically interactive, and movement of pathogens can be of direct relevance to crop management. Burdon and Thrall (2008) showed that prevalence and dynamics of diseases caused by pathogens, as well as the evolutionary trajectories of pathogens, are conditioned by the attributes of the agro-ecological interface. Indeed, the review by Alexander *et al.*, (2014) indicates several examples of pathogens moving from wild plants to crop plants causing severe losses in crop production, and also emphasized the dynamicity of agro-ecological interfaces which is influenced by anthropic activities and climate change. The anticipated impacts of climate change on vectors of pathogens, host plants, and pathogens themselves, makes it inevitable to not only aggravate the damage caused by epidemics in known pathosystems but will also accelerate the emergence of new plant diseases (Jones, 2009; Aranda and Freitas-Astúa, 2017). In light of these developments, research on the interface between crop fields and natural environments is particularly timely.

1.1.1 **Viral diseases in plants**

In plants, the emerging infectious diseases posse as a serious threat to crop production (Anderson *et al.*, 2004). According to the World Health Organization emerging infectious diseases are defined as “one that has appeared in a population for the first time, or that may have existed previously but is rapidly increasing in incidence or geographic range” (Hanssen *et al.*, 2010). Anderson *et al*., (2004) had reported that around 47% of emerging infectious diseases in plants are due to viruses (Figure 1.1). Emerging viruses can be newly described viruses that were previously unknown. However, frequently, emerging viruses are known viruses with an increased incidence in a certain niche due to changes in the environment, the vector, the host, or in the viral genome (Hanssen *et al.*, 2010; Elena *et al.*, 2014; Tomlinson *et al.*, 2017). The rationale behind viruses causing approximately 50% of the emerging diseases might be the following: viruses (1) are widely distributed across diverse plant taxa and ecosystems (Roossinck *et al.*, 2010), (2) can have shared hosts between crops and wild plants due to their broad taxonomic host range (Wisler and Norris, 2005), (3) are highly mobile across landscapes for their frequent association with insects for their transmission (Hogenhout *et al*., 2008; Uzest and Blanc, 2016) and, (4) may have significant effects on crop yield and plant fitness, even in the absence of visual symptoms (Cooper and Jones, 2006; Thresh, 2006). Therefore, preventing diseases caused by viruses may have a significant positive effect on crop production.

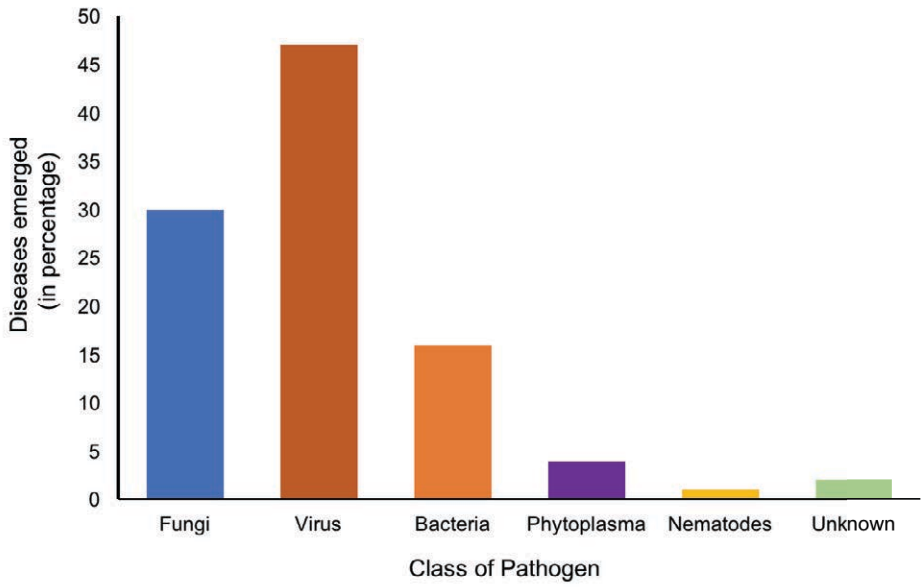


Figure 1.1 Diseases emerged in the last decades caused by different groups of plant pathogens. (Modified from Anderson *et al*., 2004)

A major reason for the importance of viral diseases is the serious limitation in the availability of control strategies. Unlike diseases caused by other micro-organisms (e.g., bacteria, fungi, etc.) which can be effectively treated by the application of bactericides, fungicides; there is no chemical control of viral diseases (Adams and Antoniw, 2006; Sastry and Zitter, 2014). Strategies are mostly based on the elimination of inoculum sources and the prevention of the dispersion of inocula, by reducing the population of virus vectors, by either chemical or biological control of vector populations. Reduction of inoculum sources include the use of virus-free seeds and planting material, prompt removal of virus infected plants, and the elimination of alternative hosts, which require a sound knowledge of host range and inoculum fluxes. Indeed an epidemiological study of a tobamovirus in greenhouses done by Reingold *et al.*, (2016) showed that the use of virus-free compost, seeds combined with early monitoring system reduced the yield losses below the economic threshold. However, all those above-mentioned strategies are not highly effective, are expensive and might cause damage to the environment e.g., insecticides. The only direct strategy for the control of viral diseases is the use of genetic resistance bred into cultivars, which is highly efficient, cheap and environmental friendly. But, the effectiveness of resistance factors is usually short-lived, as the use of resistance exerts a pressure on the virus population that selects for resistance-breaking genotypes (McDonald, 2004). Thus, the study of the factors that determine resistance durability is a main topic in plant pathology research (García-Arenal and McDonald, 2003).

Therefore, to increase the efficiency of the control of virus diseases it is important to understand the evolution of host range in plant viruses, as changes in host range resulting in host switches or host range expansions, are at the root of virus emergence either in new host species or in new host genotypes previously resistant to the virus.

1.2 Host range expansion

The host range of a pathogen is the number of species it can reproduce in. Host range conditions the epidemiology of pathogens, and is predicted to be a major factor in their evolution (Frank, 1996; Woolhouse *et al.,* 2001; Lajeunesse and Forbes, 2002). The acquisition of new hosts, that is, host range expansion, would provide a pathogen (e.g., virus) with more opportunities for transmission and survival. But, the host range of a pathogen cannot be quantified by an absolute number due to ecological factors as indicated in an unpublished review of our group. A pathogen can shift, or change between hosts, which results in a variation of the spectrum of used hosts with the host range being unchanged. Alternatively, pathogen species can expand its host range by adding more host species, or it can contract in response to ecological or other changes (unpublished review). Nevertheless, based on host range, viruses are divided into two groups: specialist viruses only infect, multiply efficiently, and spread, in hosts belonging to one or a few related species, while generalist or multihost viruses are able to do so in hosts from different species, often belonging to quite unrelated taxa (Power *et al*., 2011; Woolhouse *et al.*, 2001; Figure 1.2). It could be considered that the

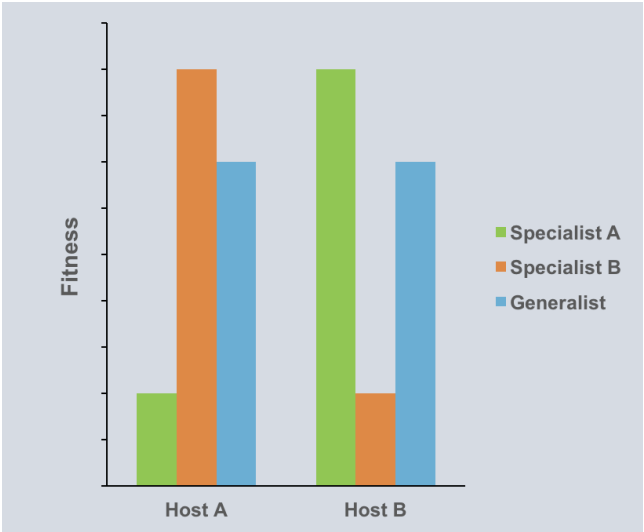


Figure 1.2 Expected fitness for specialist and generalist viruses. Although both A and B specialists perform well in their respective hosts, each one performs poorly in the other host. The generalist virus performs fairly well in both hosts but has lower fitness than either specialist in its preferred host. (Modified from Elena *et al.*, 2014)

higher opportunities for transmission and survival of generalist viruses will result in an advantage over specialist viruses. However, most theoretical analyses predict that evolution should favour narrow host ranges, because the host is the major environment component of the virus and differential host-associated selection would limit the fitness of generalists in any one host, so that they would be outcompeted by specialists (Woolhouse *et al*, 2001; Kirchner and Roy, 2002; Elena and Sanjuan, 2003). Underlying this hypothesis is the concept of trade-offs in the adaptation to different hosts, i.e. that increasing the virus's fitness in one host will result in its decreased fitness in another host (Elena *et al.*, 2009). Despite these considerations, multihost viruses make up a large fraction of viruses of both animals and plants (Woolhouse *et al,* 2001; McDonald and Linde, 2002; Power and Flecker, 2003). Also, although there is ample evidence of fitness trade-offs among hosts (section 1.2), there is also evidence of the opposite (Ebert, 1998; Elena and Lenski, 2003; Elena *et al*., 2009). Hence, identification and quantification of across-host fitness trade-offs is central to understanding virus evolution, as fitness trade-offs will constrain host range expansion (Asplen *et al*., 2012), with important consequences for disease emergence or for developing sustainable control strategies for viral diseases, as pointed out in the previous section.

1.3 **Fitness trade-offs across host species**

It has been proposed that the host range of the vectors is the primary determinant for the host range of the virus, and that plant viruses are generalists for hosts but specialists for vectors (Power and Flecker, 2003). This was proposed based on the data of the VIDE database (Brunt *et al*., 1996) where 474 vector-transmitted viruses were analysed of which only 9.9% had a single host species, while 58.4% had a single vector species. One shortcoming of this study, as pointed by the authors, was that the VIDE database does not differentiate between natural and experimental host ranges, and also there is a scarcity of information on the natural host ranges of plant viruses, and even more so on the vector ranges. García-Arenal and McDonald, (2003) study on the natural host range of 29 virus species indicated that for 17% it was restricted to a single genus, while for 35% it extended over different plant families, in good agreement with the concept that generalism is a common strategy among plant viruses. A recent work by Moury *et al*., (2017) analysed the structure of the plant – virus infectivity matrix derived from the VIDE database, and found that the matrix was significantly nested. This structure means that a large fraction of viruses is generalist able to infect plants belonging to different families.

The high percentage of generalist plant viruses cannot be taken as evidence for the absence of across-host fitness trade-offs. This is because a generalist virus may show differential adaptation to its different hosts. Genetic differentiation of virus populations according to host species or populations may be indicative of host adaptation (Garcia-Arenal *et al*., 2001; Jeger *et al*., 2006; Moury *et al.*, 2006; Elena *et al*., 2011), although most published data usually do not allow deriving this conclusion. An exception is the detailed work on the evolution of *Turnip mosaic virus* (TuMV) from a monocot-infecting virus that acquired the capacity of infecting species of the genus *Brassica* and, later on, of *Raphanus* (Ohshima *et al*., 2002; Tomimura *et al.*, 2003; Nguyen *et al.*, 2013). Differential adaptation of generalist viruses to their different hosts has been explored in a few instances. Power and Mitchell (2004) showed that the prevalence of the PAV strain of *Barley yellow dwarf virus* in seven species of Poaceae was broadly different, suggesting differences in host adaptation. This hypothesis was later confirmed by quantification of within-host multiplication and transmission (e.g. Cronin *et al.*, 2010). Another study of Vassilakos *et al.*, (2017) identified genes in *Potato virus Y* (PVY) involved in a host jump from potato cultivars to pepper plants and had discussed for the presence of severe fitness trade-offs as both pepper and potato are equally susceptible to PVY but there is no single isolate which can infect both hosts. Malpica *et al*., (2006) also analysed the prevalence of five generalist viruses in 21 wild plant species and showed significant associations between viruses and hosts. A recent study with 11 generalist viruses in 47 wild plant species across four habitats also indicated that generalist viruses exhibit host specialization depending on the habitat, behaving as facultative generalists (McLeish *et al.*, 2017). Importantly, from the last three studies it was evident that host selectivity is a successful strategy for generalist viruses which is in good agreement with the hypothesis that considers specialization as advantageous as it allows the optimal use of available resources. These reports also argue that it may not be common for generalist viruses to exploit their different hosts with equally high efficiencies.

Differences in host adaptation are not necessarily linked to fitness trade-offs. For instance, *Cucumber mosaic virus* (CMV) multiplies to lower levels, and is transmitted less efficiently, in melon than in tomato plants, and this was true for three different CMV genotypes. However, there was no evidence of across-host fitness trade-offs, as the three CMV genotypes ranked similarly for multiplication and transmission in tomato and melon (Escriu *et al.*, 2003; Betancourt *et al.,* 2011). Also, trade-offs may occur only among some hosts and/or be evidenced dependent on specific components of the virus fitness. Thus, Sacristán *et al*., (2005), compared the isolates of CMV from different host species in three botanical families and found evidence for a trade-off in infectivity among some of the hosts, but no evidence for trade-offs in within-host multiplication. On the same line, a study done by Bedhomme *et al.*,

(2012) on *Tobacco etch potyvirus* (TEV) under either constant or alternating host environments demonstrated that evolutionary history of the pathogen influences virulence, but found no evidence of trade-offs in within-host multiplication.

More definite evidence for across-host fitness trade-offs had been derived from host passage experiments than from analysis of field isolates (García-Arenal and Fraile, 2013). It has been known for decades that if virus isolates from one host are serially passaged into another host, they become adapted to the new host, which often involves phenotypic changes in the original hosts and/or frequent reversions, suggesting fitness penalties (Yarwood, 1979). The first example of a trade-off across hosts was reported by Matthews in 1949 (cited in Yarwood, 1979): when *Potato virus X* (PVX) from potato was passaged in tobacco, it increased its virulence and infectivity in tobacco, but after 19 passages it lost the capacity to infect potato. More recent experiments have re-examined host adaptation, also provided evidence for across-host fitness trade-offs. Thus, adaptation of *Hibiscus chlorotic spot virus* (HCSV) to *Chenopodium quinoa* W. resulted in a loss of fitness in its original hibiscus host (Liang *et al.*, 2002). Serial passage in peas of *Plum pox virus* (PPV) isolates from peach resulted in adaptation to the new host (increased infectivity and within-host multiplication) at a cost in the original one (decreased transmission efficiency; Wallis *et al.*, 2007). Passaging of TEV in pepper plants resulted in adaptation to the new host, with fitness reduction in the original one (tobacco) (Agudelo-Romero *et al.*, 2008). Passaging of TEV in tobacco failed to increase virus multiplication or virulence, suggesting that the fitness in the original host was already at its optimum (Agudelo-Romero *et al.,* 2008). It is noteworthy that HCSV, PPV and TEV are specialists with narrow natural host ranges, and that adaptation to new hosts in passage experiments resulted in host-range expansion. Interestingly, when six isolates of CMV were passaged in their original three hosts neither fitness nor virulence was improved in any host, suggesting that the fitness landscape of this generalist virus over its host range is at or near its maximum (Sacristán *et al.*, 2005). An exception in this is a study by Moreno-Pérez *et al.*, 2014 where they showed clear evidence of across-host fitness trade-offs in field isolates collected from tomato crops and wild tomato plants. The isolates were adapted to its native host and payed a cost in heterologous host (Moreno-Pérez *et al.,* 2014).

1.4 **Fitness trade-offs across host genotypes**

A particular case of host-range expansion is the capacity to infect and cause disease in host genotypes previously immune or resistant to virus infection, i.e. to acquire increased pathogenicity or infectivity. This process has received much attention because of its relevance for the control of virus diseases in crops. It has been observed that the duration of resistance

factors used against viruses has been, on average, much longer than that deployed against cellular plant pathogens: 12.8 years (average for 25 pathosystems) for viruses as compared to 7.3 years (average for 27 pathosystems) for fungi and oomycetes (Fraile and García-Arenal, 2010). The evolutionary potential of the virus has been identified as a major factor determining resistance durability in different analyses. Thus, the appearance of resistance-breaking genotypes on 10 resistance genes in four plant species was related to the number of amino acid substitutions required (Harrison, 2002). The analysis of the effective life of 50 resistance genes or QTLs in relation to life history traits affecting the virus evolutionary potential indicated a negative relationship between evolutionary potential and resistance durability (García-Arenal and McDonald, 2003). The analysis of 19 resistance genes in 20 pathosystems uncovered a significant association between the evolutionary constraints on the virus genes involved in resistance breaking, estimated as the ratio of nucleotide substitutions at non-synonymous and synonymous sites (dN/dS ratios) and resistance durability (Janzac *et al*., 2009). Hence, analysis of the durability of resistance to viruses suggests that the evolution of increased infectivity may be constrained by fitness penalties that may result in fitness trade-offs across the host genotypes. A detailed theoretical study done by Fabre *et al.*, (2012) explored various combinations of available options such as using resistant cultivar choice, resistance deployment strategy, landscape planning and cultural practices to increase the durability of resistance and they concluded that the optimal strategies consist of resistance deployment range from 'mixture' (where susceptible and resistant cultivars coexist) to 'pure' strategies (with only resistant cultivar) depending on the resistance characteristics and the epidemiological context (epidemic incidence and landscape connectivity).

1.4.1 **Costs of overcoming resistance**

Resistance of plants to viruses can be analysed within the frame of the gene-for-gene (GFG) model: a plant protein encoded by a resistance gene (*R*) recognizes a viral protein encoded by an avirulence gene (*AVR*), this triggers a defence reaction which limits the multiplication of the virus to the infection site (Jones and Dangl, 2006; Brown and Tellier, 2011). This limitation is expressed in most systems as either a typical hypersensitive reaction, or as extreme resistance in which no macroscopic symptoms of virus infection appear (Fraile and García-Arenal, 2010). Any kind of change in the *AVR* genes which impair *R*–*AVR* recognition results in resistance breaking and causes successful infection (Figure 1.3) which eventually leads to universal infectivity, i.e. to the capacity to infect all host genotypes. Thus, the evolution of the capacity to overcome resistance is a case of host range expansion. The GFG model assumes

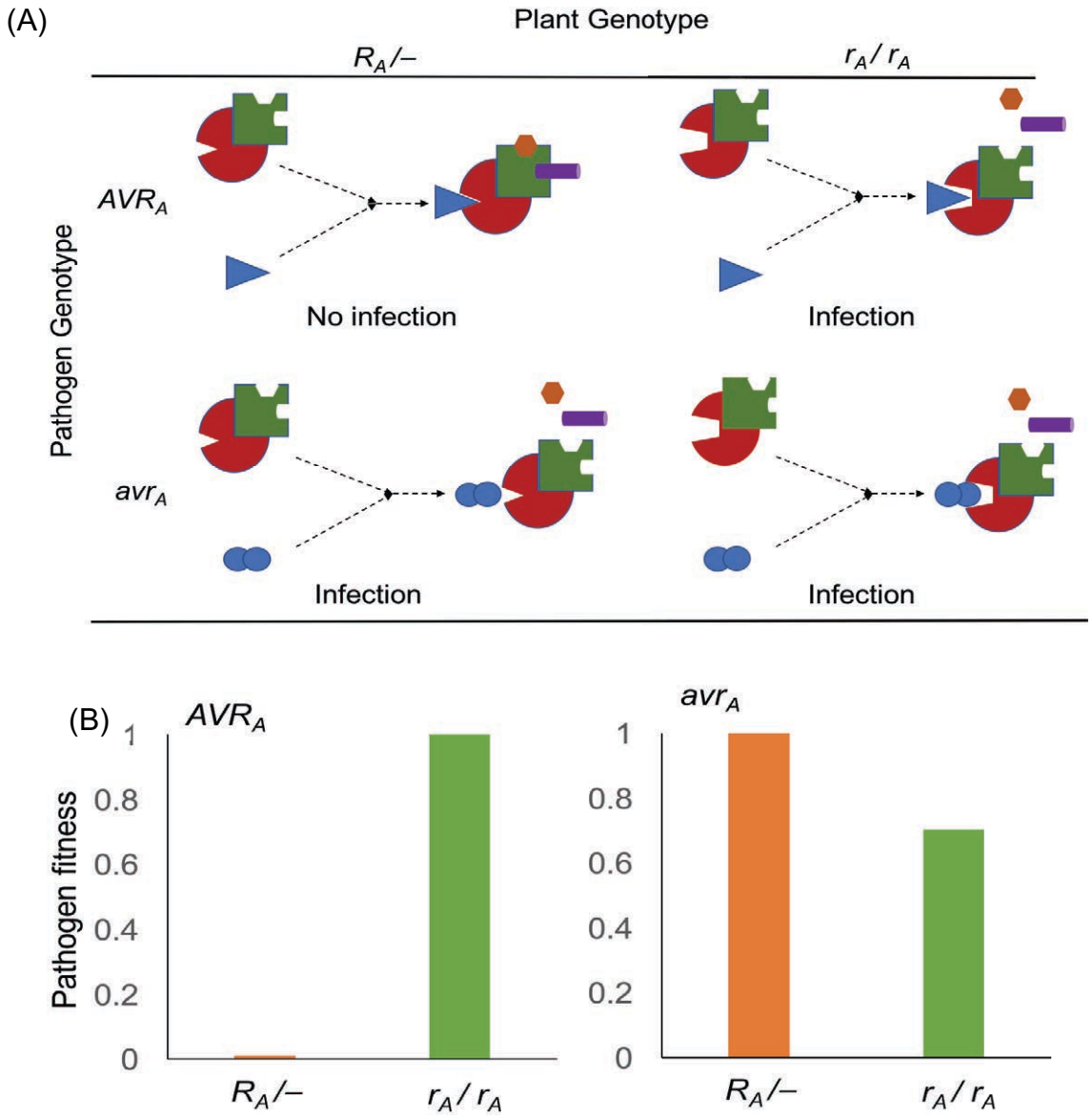


Figure 1.3 Gene-for-gene model. **(A)** The protein encoded by the resistance allele at the host locus A (R_A) ( ) recognizes the protein encoded by the avirulence allele *A* (AVR_A) in the virus ( ), triggering defence reactions by binding to other host proteins (and/or). If the plant genotype is homozygous for the recessive susceptibility allele r_A ( ), or the pathogen genotype has the virulence allele avr_A ( ), the pathogen is not recognized, defences are not triggered and infection occurs. **(B)** In the resistant host genotype (R_A /—), the relative fitness of the avirulent pathogen genotype (AVR_A) is near zero, while that of the virulent one (avr_A) is considered as 1. In the susceptible host genotype (r_A / r_A), the avirulent pathogen genotype has a higher relative fitness than the virulent genotype (cost of infectivity). (Modified from Fraile and García-Arenal, 2010)

there is a cost of infectivity, meaning that a mutant virus which is capable to break the resistance of a plant genotype will be less fit than the wild type in a susceptible plant genotype (Figure 1.3). Thus, resistance-breaking comes with a cost which might limit host range expansion. Costs of infectivity are considered necessary, but not sufficient, for stable

resistance–infectivity polymorphisms (Agrawal and Lively, 2002; Brown and Tellier, 2011; Brown, 2015).

Any virus-encoded protein might be the AVR factor for a particular *R* gene (Maule *et al.,* 2007; Fraile and García-Arenal, 2010). For cellular plant pathogens, the *AVR–R* recognition could be avoided through different mechanisms, which includes point mutations, recombination and even AVR deletion (Sacristán and García-Arenal, 2008; Hartmann *et a*l., 2017). For viruses, evolution of AVR factors is strongly limited by the multifunctionality of viral proteins, and for many *R* genes only one or a few *avr* genotypes have been reported (reviewed in Fraile and García-Arenal, 2010). This limited polymorphism at AVR is also suggestive of fitness penalties associated with increased infectivity. Long-term analyses of the genetic composition of virus populations in relation to the deployment of resistance factors are rare, at odds with similar studies with fungal pathogens such as rusts or powdery mildews. One such analysis examined the relative frequency of tobamovirus pathotypes that were, or were not, able to infect pepper (*Capsicum annuum* L.) cultivars carrying different resistance alleles at the *L* locus, as compared with the relative acreage of the different *L* alleles, for a period of more than 20 years (Fraile *et al.*, 2011). Molecular evolution analysis showed that diversification of the three pathotypes present in the virus population was on a timescale compatible with the dates on which the different resistance alleles were deployed, strongly suggesting that the virus population evolved in response to the selection exerted by the different resistances (Figure 1.4). Moreover, when the use of specific resistance alleles was

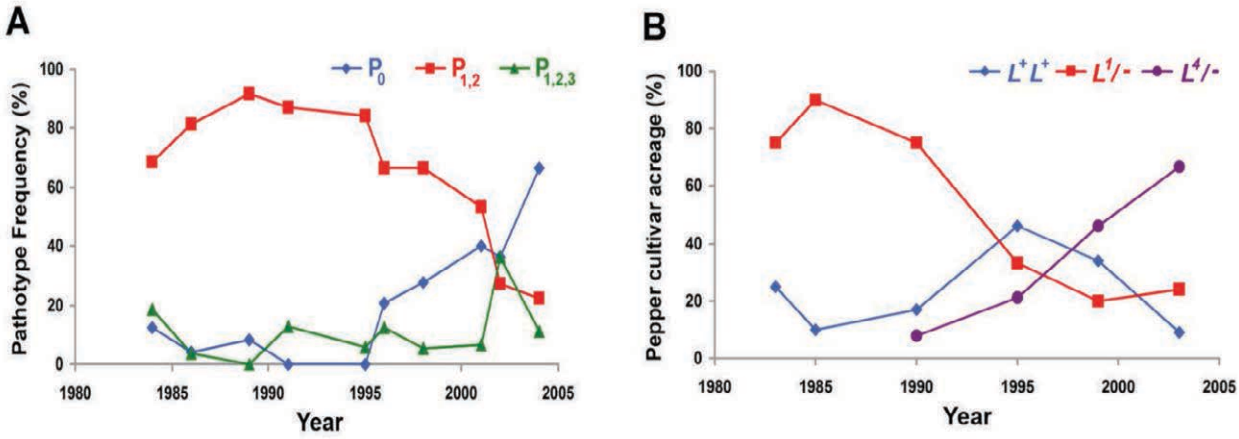


Figure 1.4 Evolution of the pathotype composition of the tobamovirus population and of the genotype composition of pepper crops in Southeast Spain along 21 years (1984–2004). **(A)** Percentage of the total number of sampled isolates identified as pathotypes P_0 (blue diamonds), $P_{1,2}$ (red squares), and $P_{1,2,3}$ (green triangles) in each year. **(B)** Percentage of the pepper acreage corresponding to L^+/L^+ (blue diamonds), $L^1/-$ (red squares), and *L4/-* (round circles) genotypes in each year. (Modified from Fraile *et al*., 2011)

discontinued, the frequency of the virus pathotypes overcoming them decreased, which suggests fitness penalties of unnecessary infectivity.

Experimental evidence of infectivity costs derives from several systems. The mutations responsible for resistance breaking may be lethal for the virus, as shown for PVY. No field isolate of PVY has been reported to overcome *Ry* in potato. Experimental mutations in the protease domain of the viral NIa protein resulted in a failure to elicit *Ry* resistance, but none of these mutants retained the protease function of NIa which is required for the viability of the virus (Mestre *et al.*, 2003). Resistance-breaking mutations may also result in a diminished ability for within-host multiplication. Thus, there is evidence for selection against PVX coat protein (CP) mutants overcoming *Rx* resistance in potato (Goulden *et al*., 1993). TuMV mutants at the cylindrical inclusion helicase protein that overcame *TuRB01* resistance in oilseed rape were out-competed by the wildtype (wt) virus in passage experiments in susceptible hosts (Jenner *et al*., 2002a). Similarly, TuMV mutants at the P3 and CI proteins that overcame a second resistance gene, *TuRB04,* also had a diminished competitive ability in susceptible hosts, and showed a high rate of reversion (Jenner *et al.*, 2002). Janzac *et al*., (2010) showed that a single non-synonymous mutation in the NIb cistron of PVY was sufficient to overcome *Pvr4* resistance in pepper, but that this mutation also resulted in a loss of competitive ability against the wild-type virus in *pvr4* pepper plants. *Tomato mosaic virus* (ToMV) mutants at the helicase domain of the 130 K viral protein, that were able to infect *Tm1*-resistant tomato plants, replicated less efficiently in susceptible tobacco protoplasts and were out-competed in *tm1* tomato plants by the wt ToMV-L strain (Ishibashi *et al.*, 2012). A study by Hillung et al., (2014) evaluated the cost of infectivity conforming to an extended GFG model. For this, they evolved TEV populations in five susceptible genetically heterogeneous ecotypes of *Arabidopsis thaliana*; they found that the specialists infected only the high permissive host, the generalists infected the less permissive hosts and showed a cost of infectivity in high permissive hosts by multiplying less than the specialists. Competition experiments in susceptible pepper lines among field isolates of *Pepper mild mottle virus* (PMMoV) virulent or avirulent on L^3-resistant plants also showed infectivity penalties in the within-host multiplication component of the virus fitness. Estimation of competitive ability showed that these penalties were high, the fitness of resistance-breaking isolates being on average 0.47 times that of isolates that did not overcome the resistance (Fraile *et al*., 2011). Although Fraile *et al.* (2011) provided the only quantitative estimates of infectivity costs in a plant virus, the data in some of the studies discussed above allow the calculation or inference of these costs (García-Arenal and Fraile, 2013). In most cases, costs were similar or higher to those reported for PMMoV, indicating that in plant viruses' high penalties to within-host multiplication are generally associated with the breaking of dominant resistance. Costs of resistance-breaking may be

modulated by various ecological and epidemiological factors. Thus, Moreno-Pérez *et al.*, (2016) showed that fitness penalties of resistance-breaking mutations depended not only on the specific mutation, but also of the specific genotype of the susceptible hosts. Moreover, fitness penalties depended on the type of infection, as they were different in single- and in mixed-infected hosts. These results underscore the complexity of predicting the evolution of resistance-breaking.

Costs of resistance breaking may also affect components of the virus fitness other than within-host multiplication, such as transmission or survival. Thus, it was shown long ago that isolates of *Rasp- berry ringspot virus* (RRSV) overcoming *Irr* resistance in raspberry were less efficiently transmitted by the nematode vector, and showed decreased seed transmission rates in alternate hosts (Murant *et al.*, 1968; Hanada and Harrison, 1977), both traits affecting virus survival and inoculum potential for the focal host, raspberry. It was also shown, through the analysis of a large collection of mutants in the CP of *Tobacco mosaic virus* (TMV), that the maintenance of the CP three-dimensional structure is essential for the elicitation of resistance determined by the *N′* gene of *Nicotiana sylvestris* (Culver, 2002). This result suggests a link between resistance breaking, particle stability and infectivity or survival. Similarly, field isolates of PMMoV that overcome L^3 resistance in pepper differ from the AVR genotype by a few amino acid changes in the CP, and it had been suggested that these changes may alter the stability of virus particles (Berzal-Herranz *et al.*, 1995; Tsuda *et al.*, 1998; Hamada *et al.*, 2007; Hamada *et al.*, 2002). Indeed it had been proved by Fraile *et al.*, 2014 that resistance-breaking mutations of PMMoV overcoming the L^2 and L^3 mediated resistance affected the stability of their virus particles. Interestingly, they also showed that the particle stability was positively correlated with the capacity to survive in the soil (Fraile *et al.*, 2014), which is often the primary inoculum to start an epidemic in the next crop season. However, authors failed to establish trade-off between multiplication and survival as reported for viruses infecting bacteria and animals (De Paepe and Taddei, 2006; Brown *et al.*, 2009; Heineman and Brown, 2012; Handel *et al.*, 2013), which shows a presence of a much more complex relationship between plant viruses and its host.

1.5 **Antagonistic pleiotropy as a cause of across-host fitness trade-offs**

The generation of across-host fitness trade-offs in pathogens has been explained by two different, non-exclusive, mechanisms, related to the differential effects of mutations in different hosts. The first mechanism pertains to pleiotropy, the different phenotypic effects of mutations

in different environments (a Genotype x Environment, G X E interaction). Across host fitness trade-offs may result from antagonistic pleiotropy (Fry, 1990; Whitlock, 1996), i.e., opposite phenotypic effects of mutations across hosts: mutations that are beneficial in one host are deleterious in another one. The second mechanism, mutation accumulation, results from the accumulation by drift of mutations that are neutral in one host, but deleterious in another one, for instance because mutations accumulate in genes unnecessary in one host but necessary in a second one (Kawecki, 1994). For viruses, antagonistic pleiotropy seems to be the major cause of across-host fitness trade-offs which also leads to host specialization (Elena, 2017). There is ample evidence, from different systems, for antagonistic pleiotropy in plant viruses, and evidence derives both from the analysis of adaptation to different host species, and of overcoming resistance genes. For example, Liang *et al.*, (2002) showed that five different HCSV lineages adapted *to C. quinoa* had fixed eight amino acid mutations in the coat protein (CP), three of which were enough, when introduced by site-directed mutagenesis in the original genotype, to cause its fitness loss in hibiscus. A similar study was conducted with a second carmovirus, *Pelargonium flower break virus* (PFBV) which was adapted *to C. quinoa* (Rico *et al.*, 2006). PFBV isolates adapted to *C. quinoa* leaves had five specific non-contiguous amino acid substitutions in the CP. When a wildtype isolate from geranium was inoculated onto *C. quinoa* leaves, after the first passage the viral populations had already fixed two of the *C. quinoa*-specific changes and after four serial passages all changes were introduced (Rico *et al,* 2006). The amino acid substitution N25I in the CP of PVY resulted in increased virus multiplication in tobacco, but in decreased multiplication in potato (Moury and Simon, 2011). A single nucleotide substitution engineered in the CI cistron of TuMV that resulted in overcoming *TuRB01* resistance in oilseed rape also resulted in a decreased competitive ability with the wild type in passage experiments in the susceptible host (Jenner *et a*l., 2002). A study of Calvo *et al.*, (2014) showed that a single amino acid replacements in the 6K1-CI proteolytic site targeted by the NIa-Pro viral protease, or nearby, were responsible for virus adaptation to two different hosts and was enough for negative fitness impact in non-adapted hosts. In support of second mechanism, mutation accumulation, there is not much experimental evidence but a recent study by Hillung *et al.*, (2015) provides support for this. The results showed that mutations fixed during an evolution experiment were so by drift as these mutations were mostly deleterious or neutral in their local host and only a very reduced number had a host-specific beneficial effect when introduced separately in the viral infectious clone (Hillung *et al.*, 2015). Note that most of these studies were based on the analysis of engineered mutants in infectious cDNA clones of the virus, thus precluding phenotypic effects of uncontrolled mutations in other genomic regions.

Probably the most detailed analysis of the distribution of mutational effects (DMFE) across hosts for plant viruses has been reported with TEV (Carrasco *et al.*, 2007; Lalic *et al.*, 2011), which showed that the dependence of DMFE on environment (host) would clearly impact the probability of adaptation to new hosts. For example, if the change of host provides new opportunities, the fraction of beneficial mutations would increase, displacing the mean of the distribution toward higher values or making the variance larger (Figure 1.5). The study of Carrasco *et al.*, (2007) first analysed DMFE in the original host of the virus, tobacco, by estimating the fitness of 66 single nucleotide substitution mutations randomly distributed over

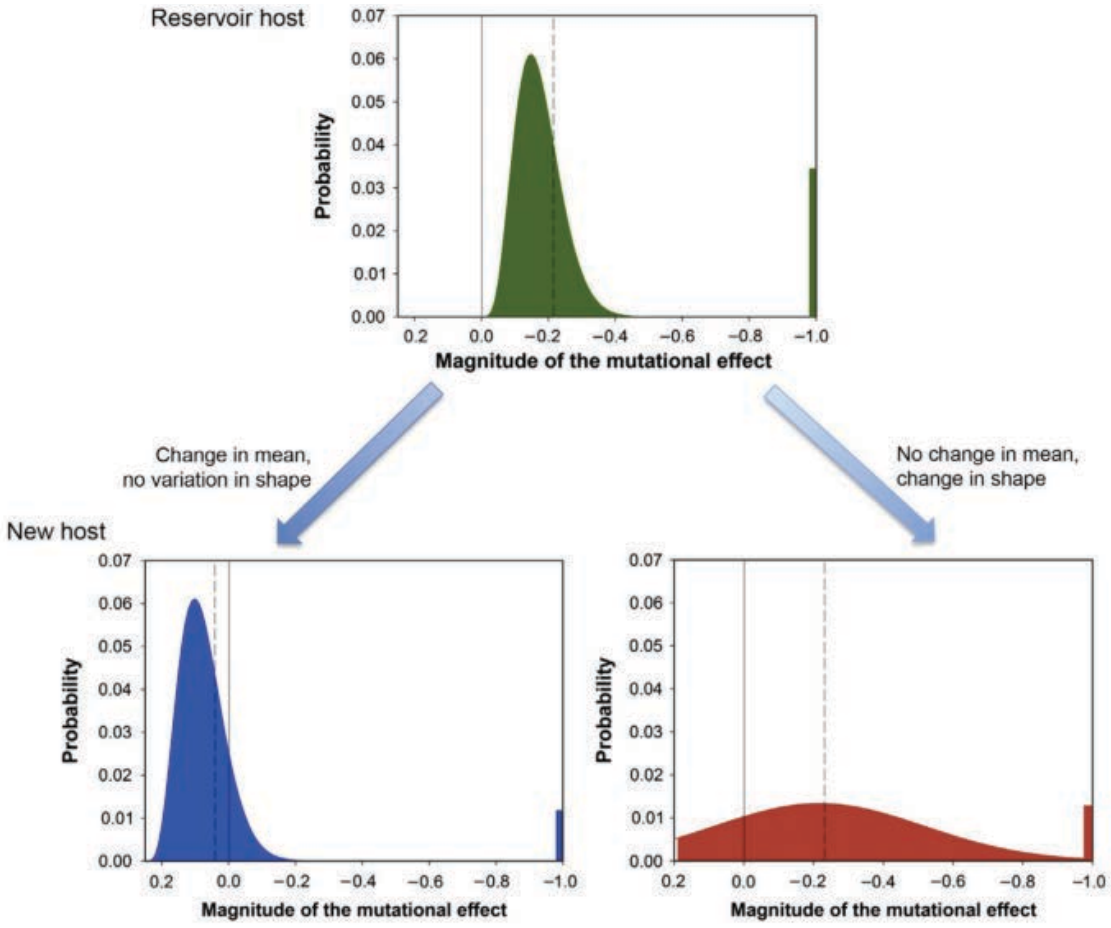


Figure 1.5 Possible effects of host switching in the distribution of mutational fitness effects. In all cases, the vertical dashed line indicates the average effect, whereas the vertical solid line indicates the null (neutral) effect. The surface under the curve that is at the left of the continuous line represents the fraction of beneficial mutations, whereas the surface at the right side of this line represents the fraction of deleterious mutations. The upper diagram shows the distribution for the primary host. The lower diagrams show two possible host effects: the left panel represents a change in the magnitude of the average effect without affecting the shape of the distribution; the right panel illustrates a change in the shape while retaining the same average effect. In both cases, the fraction of mutations with beneficial effect increases. (From Elena, 2011).

the genome: 41% of mutations were lethal, 36% strongly deleterious (decreasing fitness by 40% on average) and 23% apparently neutral, with no beneficial mutations identified (Carrasco *et al.,* 2007). In the second study by Lalic *et al.*, (2011) the DMFE was determined for a subset of 20 mutants from the previous study in a panel of eight different experimental host species of the virus, which belonged to the same genus as the original host (i.e. *Nicotiana*), to different genera within the same family, Solanaceae (*Capsicum, Datura* and *Solanum*) or to genera in unrelated families (*Helianthus, Gomphrena* and *Spinacia*). Results showed a host effect on the DMFE, with a lower mean (i.e. stronger deleterious effects) the closer the taxonomic relatedness with the original host, and a longer right tail (i.e. a larger fraction of beneficial mutations) in the non-solanaceous hosts. Results showed evidence for a strong G X E interaction, which could be explained in part by antagonistic pleiotropy, as the number of mutations whose fitness effects changed signs across hosts was significantly larger than expected by chance.

There is also ample evidence for antagonistic pleiotropy as a generator of across-host fitness trade-offs from RNA and small ssDNA viruses infecting animals and bacteria (Elena *et al.*, 2009). The importance of antagonistic pleiotropy may be a consequence of the nature of the genomes of RNA and ssDNA viruses, which are small, compact, encode multifunctional proteins and, have little neutrality.

1.6 **Epistasis as a modulating factor of fitness trade-offs**

Evolutionary biologists had been using genotype-fitness maps to understand the evolutionary trajectory of an organism fitness or a protein function which prompted Sewall Wright in 1932 to introduce the concept of the fitness landscape (also known as the adaptive landscape), which is a visualization of a high-dimensional map. The iconic graphical rendering is a three-dimensional 'mountainous' landscape in which genotypes are organized in the x–y plane and fitness is plotted on the z axis (Figure 1.6). In such a landscape, evolution can be seen as 'walks' and adaptation as 'climbs' to higher positions (peaks) on the fitness surface. But the path towards adaptation is often hindered by the presence of valleys in between fitness peaks, due to which the fitness landscape is usually "rugged". Thus, to go from peak C to peak A in the fitness landscape represented in Figure 1.6, fitness has to decrease first and then climb the fitness peak A, which will be difficult for the pathogen. The ruggedness of fitness landscapes is owed to the effects of epistasis, i.e., of any kind of genetic interaction that leads

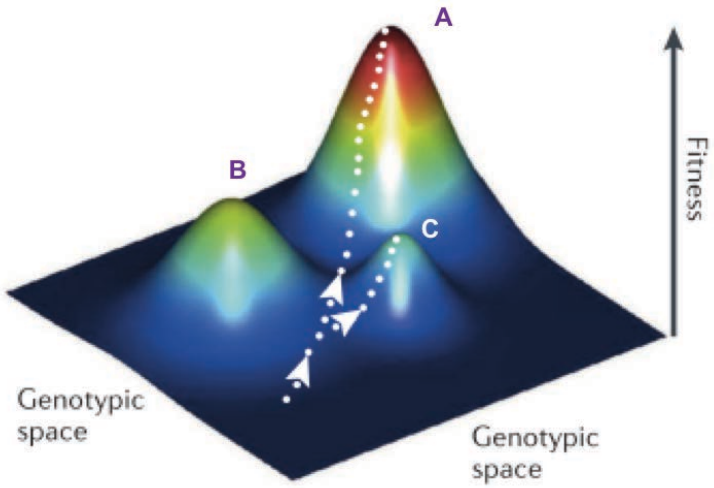


Figure 1.6 A three-dimensional fitness landscape. Genotypes are represented in the x – y plane and fitness on the z axis. The landscape is visualized as 'mountainous' or "rugged", with three fitness peaks (A, B and, C) separated by fitness 'valleys'. Two evolutionary trajectories are shown by white dots and arrows. (Source: de Visser and Krug, 2014).

to a dependence of mutational effects on the genetic background (de Visser & Krug, 2014). As pleiotropy means a genotype x environment interaction, epistasis means a genotype x genotype (G x G) interaction.

A measure of epistasis can be derived from experimental fitness measures of single and double mutants (Kouyos et al, 2007). Thus, epistasis (ε_{xy}) between mutations *x* and *y*, can be calculated as $\varepsilon_{xy} = W_{00}W_{xy} - W_{x0}W_{0y}$ (Kouyos et al., 2007), where W_{00}, W_{xy}, W_{x0}, and W_{0y} correspond to the absolute fitness of the wild-type, the double mutant and each single mutant, respectively. Epistasis can be divided in various types depending on the actual effects of the interaction (Figure 1.7): magnitude epistasis refers to cases where the magnitude effect of a mutation depends on the background while its sign is constant. Magnitude epistasis is positive when the double mutant is fitter than expected from the multiplicative effect of the individual mutant and negative in the opposite case. Sign epistasis refers to cases where the background affects the sign of the effect of a mutation. Reciprocal sign epistasis is a particular case where the sign of the effect of a mutation depends on the allele present at another locus and reciprocally. Reciprocal sign epistasis is a necessary condition for an adaptive landscape to be rugged (Poelwijk et al, 2011). Recent studies in plant viruses (Montarry et al., 2011; Poulicard et al., 2012; Lalić and Elena, 2012; Lalić and Elena, 2015; Lalic and Elena, 2012; Cervera *et al.*, 2016), bacteriophages (Cabanillas *et al.*, 2013) and human viruses (Da Silva *et al.*, 2010) highlighted that sign epistasis, and in particular reciprocal sign epistasis, are more

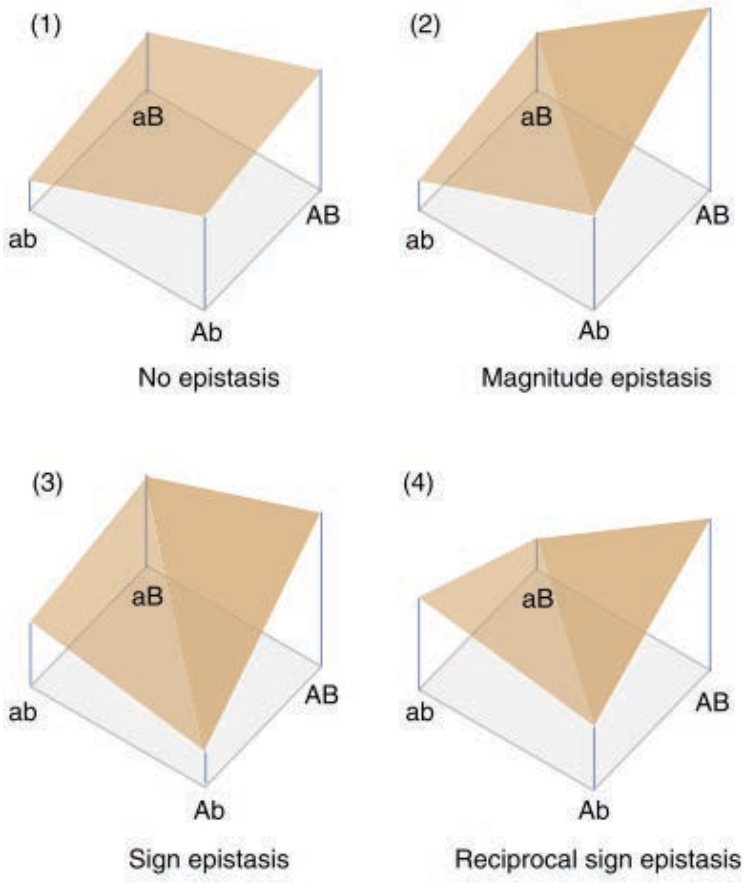


Figure 1.7 Different types of epistasis between two loci. The two loci (a/A and b/B) determine the fitness of a genotype. Small letters are considered as wild type and capital letters are considered as mutants. (1) If there is no epistasis the fitness of the double mutant AB results from multiplying the fitness effects of single mutations on the wild type genetic background (i.e. the fitnesses of genotypes Ab and aB). (2) If magnitude epistasis is present, the fitness of the double mutant AB is different from the multiplicative expectation. In the example, the observed fitness of AB is larger than expected (positive epistasis). Both in the cases of no epistasis or of magnitude epistasis, the effects of mutations A and B are unconditionally beneficial. If the effect of one of the mutations is conditionally beneficial (i.e. beneficial in one genetic background but deleterious in another), then it is called sign epistasis (3). Finally, if both mutations A and B are deleterious by themselves, but beneficial when combined, it is called reciprocal sign epistasis (4). (Source: Bedhomme *et al*., 2015)

frequent than it was previously thought. This suggests that the existence of ruggedness of the adaptive landscapes diminishes the ability of viral populations to escape from specialism to a situation of no-cost generalism. On the other hand, Remold, (2012) presented a model which explains the evolution of specialists and no-cost generalists: epistatic pleiotropy. Epistatic pleiotropy occurs when viral genetic backgrounds differ in how the effect of an allele depends on the host. Epistatic pleiotropy, unlike either antagonistic or magnitude pleiotropy in the absence of epistasis, allows for the evolution of either specialist or no-cost generalist viruses, depending on the virus population's host. However, when epistasis is present with pleiotropy,

it can create rugged fitness landscapes, which may allow either specialists or no-cost generalists, and also may prevent escape from specialism to no-cost generalism depending on the host environment (Remold, 2012). Following this model Hillung *et al.*, (2015) pointed two pairs of mutations which depending on the environment where they evolved might achieve either specialism or no-cost generalism upon a host jump. An interesting review by Bedhomme *et al.*, (2015) further gives emphasis to epistatic pleiotropy model and in simple way explained it as higher order interactions which means that the type and magnitude of epistasis might depend on the host species. This higher order interaction is termed as genotype x genotype x environment (G x G x E) interactions.

1.7 **Pathosystems studied**

In this thesis, I will analyse the role that across-host fitness trade-offs play in virus adaptation to hosts, either to new host species or to new host genotypes. The study focusses in tobamoviruses that infect pepper crops, specifically two virus species *Tobacco mild green mosaic virus* (TMGMV) and *Pepper mild mottle virus* (PMMoV).

1.7.1 **The viruses**

Species of genus *Tobamovirus* (family *Virgaviridae*) are characterized by stable viral particles in the form of rigid cylinders (300 nm long and 18 nm diameter) of helical symmetry (Figure 1.8), with a central cavity 2.3 nm diameter (Caspar, 1964; Fraenkel-Conrat, 1986). The coat protein structure of *Tobacco mosaic virus* (TMV) had been studied in detail and shall be used here to explain the structure of viruses belonging to *Tobamovirus* genus.

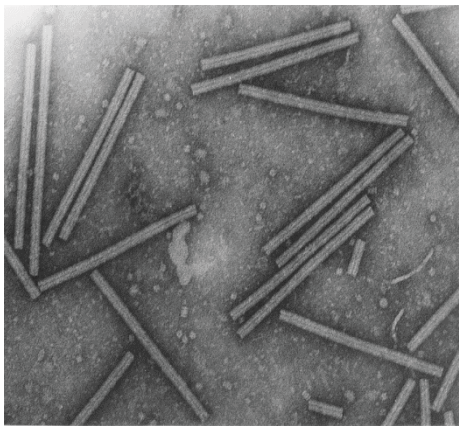


Figure 1.8 A virus belonging to *Tobamovirus* genus under an electron microscope.
(Source: Adams and Antoniw, 2006)

The central core of the coat protein consists of a right-handed antiparallel helical bundle composed of four alpha-helices, the left and right slewed (LS and RS) and left and right radial (LR and RR). A short inner loop connects the LS and RS helices, and a longer loop connects the LR and RR helices. The RNA binding site is composed of residues from both loops and the LR helix. Both the N and C termini are located on the outer surface of the virion (Figure 1.9, Culver, 2002). Depending upon solution conditions, TMV CP forms three general classes

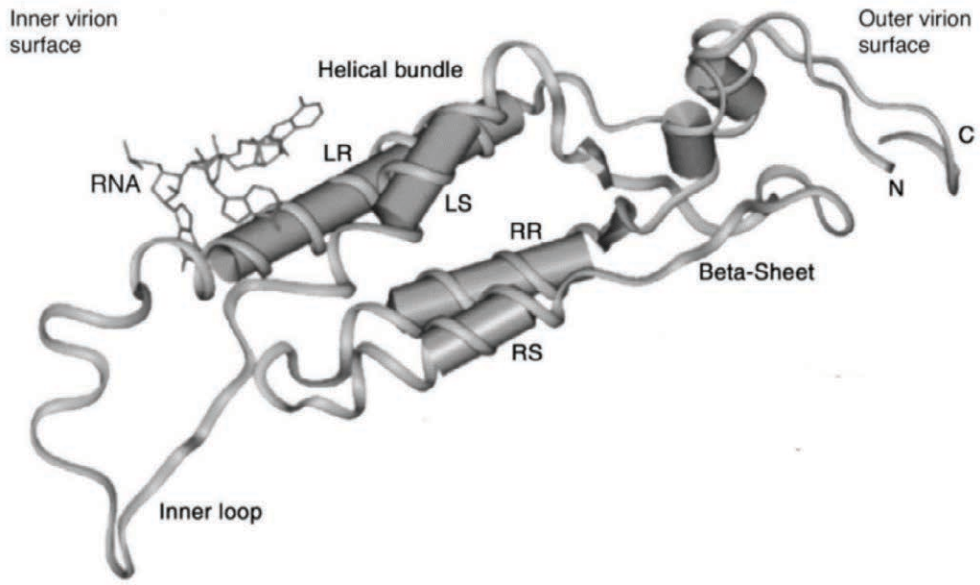


Figure 1.9 Representation of a single subunit of CP of tobacco mosaic virus.
(Source: modified from Culver, 2002)

of aggregates, the 4S or A protein composed of a mixture of low-order aggregates (in high pH and low ionic strength), the 20S disk or helix composed of approximately 38 subunits (in neutral pH), and an extended virion-like rod (in low pH and high iconic strength) (Figure 1.10, Durham *et al.*, 1971; Bloomer and Butler, 1986). Virus particles are metastable structures that must protect the viral genome while in the environment and deliver it for infection upon entry in the infected cell after disassembling. For TMV, virion disassembly has been much analysed both *in vitro* and *in vivo* (Stubbs, 1999; Culver, 2002). Caspar, (1964) proposed that stability switching for infection depends on the protonation state of carboxyl-carboxylate groups from adjacent CP subunits, identifying the pairs E50-D77 (axial interaction) and E95-E106 (side

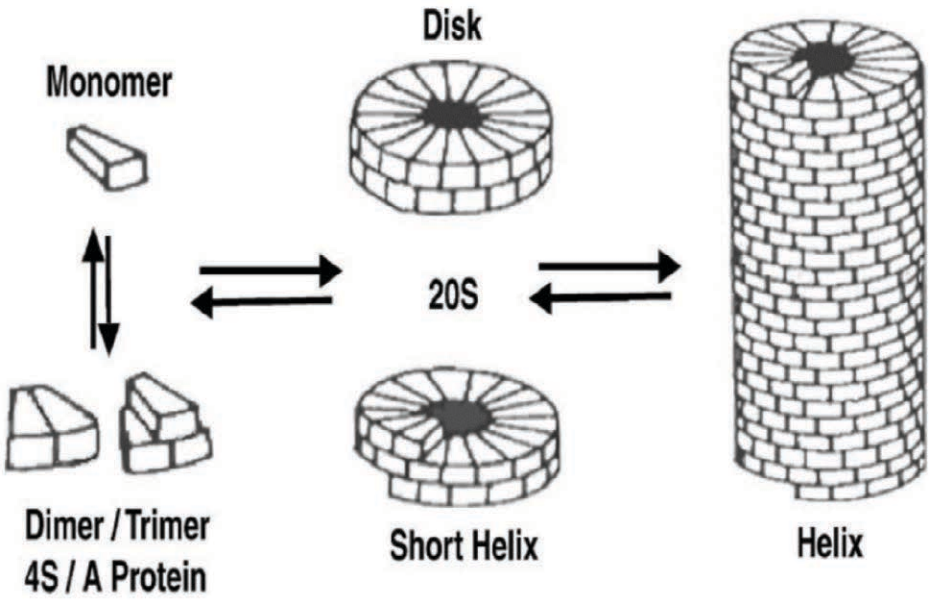


Figure 1.10 Different forms of TMV CP aggregates.
(Source: Culver, 2002)

interaction) in the virion structure (Namba and Stubbs, 1986). These carboxylate groups have since been called “Caspar carboxylates,” and later mutational analyses confirmed their role in disassembly, with the axial interaction E50-D77 having the highest effect (Culver *et al.*, 1995; Lu *et al.*, 1996). In the environment, interactions between these Caspar carboxylates will be stabilized by Ca^{2+} ions or protons, and the higher pH and lower calcium concentration in the cell will result in repulsion between carboxylates and destabilization of the particle structure, which will make the RNA accessible for cotranslational disassembly (Wilson, 1984). Mutations of amino acids involved in the stability switch showed that increased stability resulted in decreased infectivity in a necrotic local lesion assay and less efficient translation of virion particles in rabbit reticulocyte lysate (Culver *et al.*, 1995; Lu *et al.*, 1996).

The genome of tobamoviruses consists of a positive-sense single-stranded RNA molecule of about 6.4 Kb, with a Cap structure (m7G5'pppG) at the 5'-end and a structure similar to the transfer RNA (tRNA) at the 3'-end. (Figure 1.11) (Klug, 1999; King *et al*., 2012). The genome of the tobamoviruses encodes four proteins. The 5’ most open reading frame (ORF 1) is translated from the genomic RNA (gRNA), and encodes a protein that has a molecular weight between 124 and 132 KDa, according to virus species, and contains helicase and methyltransferase motifs. This protein also acts as a suppressor of RNA silencing (Ding *et al*., 2004; Kubota *et al*., 2003). ORF 2, in phase with ORF 1, is translated by read-though

of the ORF 1 stop codon, which results in a protein of between 181 and 189 KDa. The read-though portion of this protein, encoded by ORF 2, is the virus RNA-dependent RNA polymerase. ORF 3 is in a second reading frame, and is translated from a subgenomic RNA (sg RNA) 3' co-terminal with the gRNA. It encodes a protein of between 28 – 31 KDa that is the virus movement protein, with RNA-binding and plasmodesmata-gating functions, necessary for the cell-to-cell movement of the virus. ORF 4, in a different frame than ORF 3, is translated from a second sg RNA, and encodes the 17.5 KDa coat protein (CP).

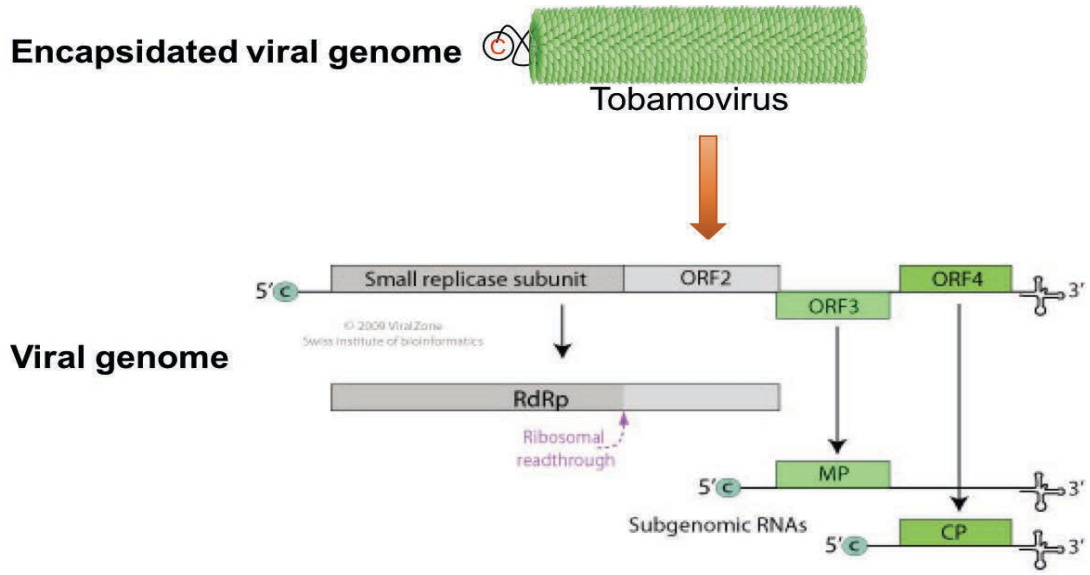


Figure 1.11 Representation of genome organization of viruses belonging to genus *Tobamovirus.* (modified from viralzone.expasy.org)

TMGMV, was first reported in *Nicotiana glauca* Grah. from the Canary Islands, in co-infection with TMV (McKinney, 1929). Several strains of this virus species have been isolated and characterized, including the U2 strain (previously U2-TMV) isolated from tobacco co-infected by TMV or strain U5 (previously U5-TMV, Siegel and Wildman., 1954). TMGMV is economically important in tobacco crops in Germany and the United States of America, and since its first report in Italy in the 1980s, in pepper crops (Wetter, 1984). TMGMV has also been reported naturally infecting ornamental plants such as *Eryngium* spp., *Torenia* spp., *Calibrachoa* spp., *Petunia* spp., *Solanum* spp., and, *Capsicum* spp. (Zeidan *et al.*, 2008).

A study by Fraile *et al.*, (1996) analysed the genetic structure of TMGMV infecting *N. glauca* plants from a number of regions in the world and found that the TMGMV populations were closely similar, although Californian and Cretan populations were twice as variable as

those in Australia and Spain. The Californian and Cretan populations were not separated in a cluster analysis, implying that they were from an older single population. By contrast, the Australian and Spanish populations formed distinct tight subclusters, and thus probably were more recently established from the older population. However, the total diversity of all the populations sampled was no greater than that of the two most diverse ones. An interesting study done by Bodaghi *et al.*, (2000) reported the presence of two types of TMGMV isolates based on its 3' untranslated region (UTR) which are Small type and Large type. Both types were isolated from *N. glauca* and also showed differences in distinctive symptoms in 3 hosts out of 16 assayed hosts. However, Bodaghi *et al*., (2000) could not establish what might be the advantage or disadvantage of the duplication in 3' UTR. Later on, Bodaghi *et al.*, (2004) did a study on cross-protection between two type isolates of TMGMV in *N. tabacum* L. 'Xanthi' and *N. glauca*, and found that irrespective of the assayed host the TMGMV large type showed 100 % cross-protection but the TMGMV small type did not show cross-protection. It was concluded that the large type was more competitive than the small type (Bodaghi *et al*., 2004).

Upon analysing the sequence of Small type TMGMV it was found to be similar to the complete genome sequence reported by Solis and Garcia-Arenal, (1990) (Bodaghi *et al*, 2000). The large type TMGMV had a duplication of 147 bp at the 3' UTR. The repeated sequence contained that part of the tRNA-like structure corresponding to the anti-codon domain upstream of the aminoacyl acceptor domain. In addition, the repeated sequence contained the sequence for three pseudoknotted structures predicted to form in the 3'- UTR of TMGMV between the 5' end of the tRNA-like structure and the 3' end of the CP gene (Figure 1.12). The repeat sequence started at position 39 and ended at position 185 (numbered according to the Figure 1.12) (Bodaghi *et al,* 2000).

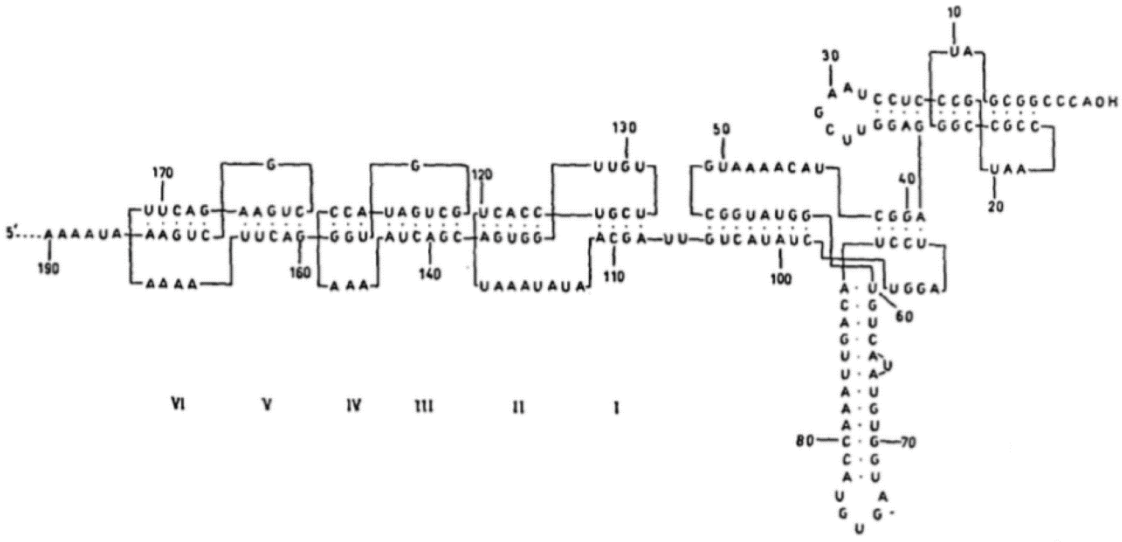


Figure 1.12 The 3' Untranslated region end of TMGMV.(Source: García-Arenal, 1988)

PMMoV, was first reported in the United States of America as a latent strain (Mckinney, 1952), later named the Samsun latent strain of TMV (Greenleaf *et al*., 1964). It was first recognised in 1984 as causing a pepper disease in 1984 in Europe by Wetter *et al.*, (1984). PMMoV causes considerable yield losses in field-grown and protected pepper crops as well as significant loss of quality of fruits. Only peppers (*Capsicum* spp.) are considered as natural hosts of PMMoV. Nevertheless, experimentally PMMoV can infect at least 24 species of six genera of the *Solanaceae* and five species of four genera in four other families, *Chenopodiaceae*, *Cucurbitaceae*, *Labiatae* and *Plantaginaceae* (Avgelis, 1986; Wetter, 1984).

Several isolates of PMMoV were being identified in different parts of the world which had the capacity to overcome the resistance conferred by alleles at the *L* locus present in several *Capsicum* species. Sources of resistance to tobamovirus had been sought in wild relatives and varieties of *C. annuum*. One of the first well-characterized locus of resistance to viruses in *Capsicum* spp. is the *L* locus, located in chromosome 11. The alleles of this locus have been found in different species of cultivated and wild peppers (*C. annuum*, *C. frutescens*, *C. chinense* and *C. chacoense*) (Boukema, 1980; Rast, 1988). The alleles of the *L* locus confer specific resistance to the different species and/or genotypes of viruses of the *Tobamovirus* genus, and these species and/or genotypes, in turn, are classified into pathotypes according to their capacity to overcome these resistances (Table 1.1). Plants that have the allele L^1, are

Table 1.1. Reaction of different tobamovirus species and pathotypes toward *Capsicum* spp. genotypes harbouring different alleles at the *L* locus

		Pepper species and genotype				
		C. annuum	*C. annuum*	*C. frutescens*	*C. chinense*	*C. chacoense*
Virus	Pathotype	L^+/L^+	$L^1/-$	$L^2/-$	$L^3/-$	$L^4/-$
TMV, ToMV, TMGMV, BPeMV	P_0	S	R	R	R	R
PaMMV, ObPV	P_1	S	S	R	R	R
PMMoV	$P_{1,2}$	S	S	S	R	R
PMMoV	$P_{1,2,3}$	S	S	S	S	R
PMMoV	$P_{1,2,3,4}$	S	S	S	S	S

S: susceptibility, that is, systemic infection; R: resistance, that is, necrotic local lesions without systemic infection.
TMV: *Tobacco mosaic virus*; ToMV: *Tomato mosaic virus*; TMGMV: *Tobacco mild green mosaic virus*; BPeMV: *Bell pepper mottle virus*; PaMMV: *Paprika mild mottle virus*; ObPV: *Obuda pepper virus*; PMMoV: *Pepper mild mottle virus* (Souce: Moury & Verdin, 2012)

resistant to the P_0 pathotype, but are susceptible to pathotypes P_1, $P_{1,2}$, $P_{1,2,3}$, and $P_{1,2,3,4}$; plants with L^2 allele are resistant to pathotypes P_0 and P_1, but are susceptible to pathotypes $P_{1,2}$, $P_{1,2,3}$, and $P_{1,2,3,4}$ and so on (Table 1.1) (Boukema, 1980; Rast, 1988). Plants with locus *L* alleles in heterozygous show incomplete dominance that confers less effective resistance. For this reason, the first isolates that overcame resistance were observed in heterozygous pepper cultivars for these *L* alleles. L^i resistance alleles encode resistance proteins of type CC-NB-LRR (Tomita *et al.*, 2011). These proteins (R) detect the presence of the capsid protein (AVR) of the different pathotypes according to a GFG model and trigger a hypersensitive reaction which limits the infection to the point of entry. This recognition is thermosensitive, so that it does not take place at temperatures equal to or higher than 28 ºC (Berzal-Herranz *et al*., 1995, Boukema, 1984, Rast, 1988, Tomita *et al.*, 2011). Specific amino acid changes in the CP prevent its recognition by the resistance protein, which results in the overcoming of resistance and the emergence of resistance-breaking virus genotypes (Berzal-Herranz *et al.,* 1995; de la Cruz *et al*., 1997; Gilardi *et al.*, 2004). Isolates of tomato mosaic virus (*Tomato mosaic virus*, ToMV), TMV and TMGMV are of the pathotype P_0, which can only infect varieties of pepper without alleles of resistance in the locus *L* (L^+ / L^+). Isolates of paprika mild mottle virus (*Paprika mild mottle virus*, PaMMV) and *Obuda pepper virus* (ObPV) are of pathotype P_1 (Table 1.1) (Berzal-Herranz *et al*., 1995). Pathotypes $P_{1,2}$, $P_{1,2,3}$ and $P_{1,2,3,4}$ are isolates of PMMoV that differ in the sequence of their CP. Most of the described isolates of this virus are of pathotype $P_{1,2}$ that only can overcome the resistance alleles L^1 and L^2. In addition, isolates had been described that overcome the resistance of the L^3 allele and subsequently L^4 (Table 1.1 and Figure 1.13) (Antignus *et al.*, 2008; Genda *et al.*, 2007; Hamada *et al.*, 2007; Hamada *et al*., 2002; Tsuda *et al.*, 1998). PMMoV phylogenies cluster isolates into three groups, of which the first and second groups are formed mostly by isolates of pathotype $P_{1,2}$ and the third group is formed by isolates belonging to pathotype $P_{1,2,3,}$ which contains the M138N substitution (Figure 1.13). In the first group two isolates of pathotype $P_{1,2,3}$ had been identified, one isolate had the substitution (S81A) and the other had substitutions (L13F + G66V) (Hamada *et al.*, 2008; Hamada *et al*., 2007). In the second group, an isolate of pathotype $P_{1,2,3}$ had appeared due to the substitutions T43K + D50G, and another one of pathotype $P_{1,2,3,4}$ due to the substitutions Q46R + G85K (Hamda *et al.,* 2008; Genda *et al.,* 2007). Interestingly all those pathotypes of PMMoV had been reported from Japan. In the third group an isolate from Israel of pathotype $P_{1,2,3,4}$ occurs appeared, which contains a substitution A86G in the CP (Figure 1.13) (Antignus *et al*. al., 2008; Genda *et al.,* 2007; Moury and Verdin, 2012). All the amino acid substitutions described are necessary and sufficient for the conversion of the $P_{1,2}$ pathotype into $P_{1,2,3}$ and $P_{1,2,3,4}$ (Tomita *et al.,* 2011).

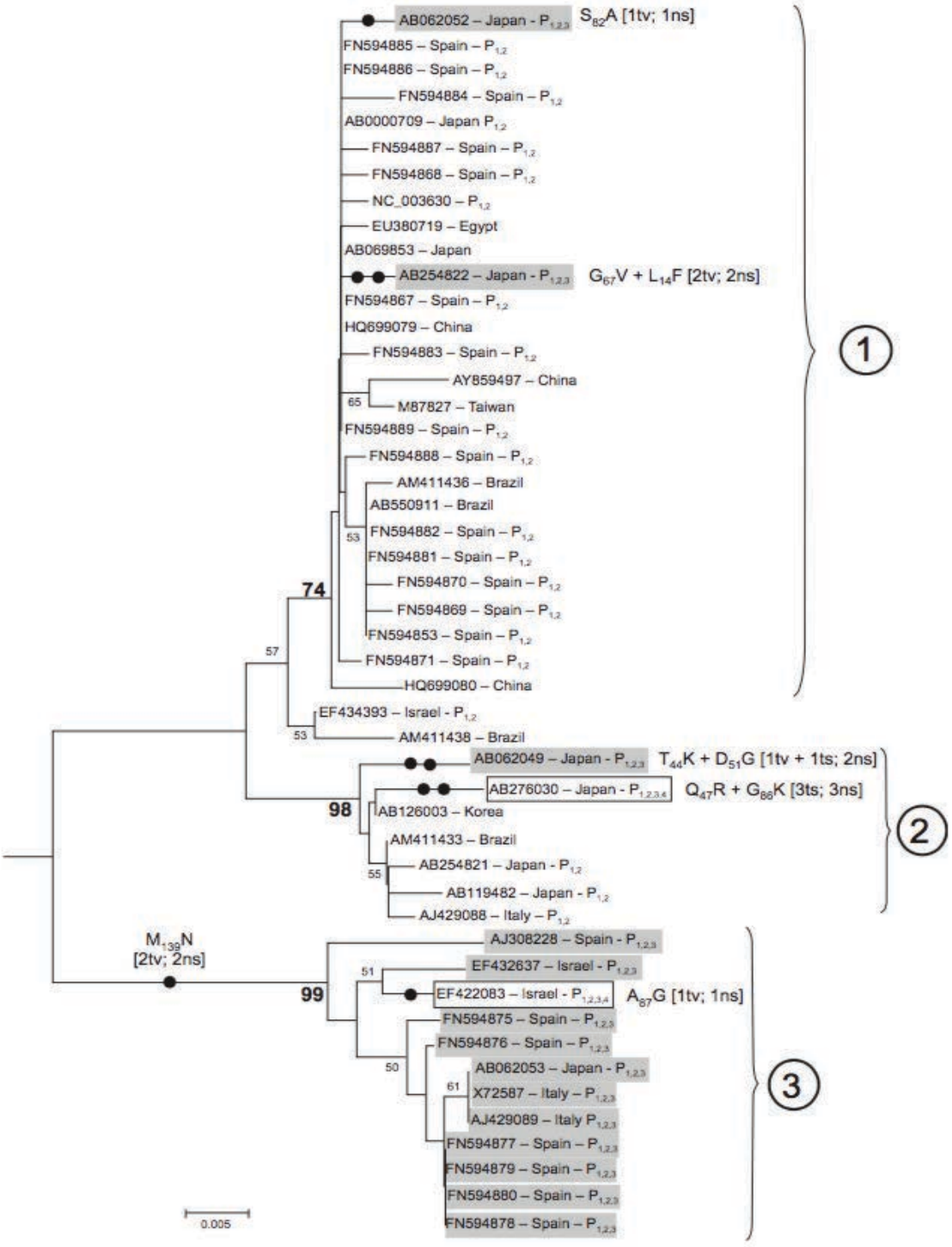


Figure 1.13 Rooted neighbour joining phylogenetic tree of the coat protein (CP) gene of Pepper mild mottle virus (PMMoV). PMMoV pathotypes defined according to the capacity to overcome resistance alleles at the *L* locus are indicated when known. Pathotype $P_{1,2,3}$ PMMoV isolates are shaded in gray and pathotype $P_{1,2,3,4}$ PMMoV isolates are boxed. The number of amino acid changes in PMMoV CP involved in the breakdown of the L^3 and L^4 resistance genes are indicated (black circles), together with the corresponding number of transitions (ts), transversions (tv), and nonsynonymous (ns) nucleotide substitutions. Bootstrap percentages above 50% are shown. The scale bar indicates branch lengths in substitutions per nucleotide. (Source: Moury & Verdin, 2012)

1.7.2 **The host plants**

N. glauca belongs to the Solanaceae family. It is a species of wild tobacco known by the common name of tree tobacco. It is an evergreen shrub native to South America, growing rampantly in Mediterranean climate on other continents as an introduced species (Holmes, 1951; Ollerton *et al.*, 2012; Gray *et al*., 2005) and can produce more than 1.3 million seeds per individual (Fabricante *et al*., 2015). In Australia, *N. glauca* is listed as a naturalised weed of the natural environment (Randall, 2007). It is listed in the Global Compendium of Weeds as an agricultural weed, cultivation escape, environmental weed and noxious weed (GCW, 2013) and is also listed on the Global Invasive Species Database. Historically, *N. glauca* has been transported worldwide as an ornamental species (Ollerton *et al*., 2012), which has been the principal means of its escape and colonization. As it is still used as an ornamental species there is a risk of introduction from future escapes.

N. glauca had been studied as a potential reservoir for virus infection in crops. In Mexico, *N. glauca* was shown to be a potential reservoir of *CMV*, TMV and PVY (Avina-Padilla *et al.*, 2008). In California, *N. glauca* had been reported as a host plant of *Tomato infectious chlorosis virus* (TICV) which causes economic losses in commercial tomato production (Jones, 2001). The potential reservoir role of *N. glauca* is relevant to the objectives of this thesis.

As indicated in section 1.6.1, TMGMV was originally isolated from *N. glauca*, first in the Canary Islands and then in California. Bald and Goodchild, (1960) reported that *N. glauca* was a better host for TMGMV than for TMV. Fraile *et al.,* (1997) analysed *N. glauca* herbarium specimens from Australia for the period 1899 – 1993 for TMV and TMGMV infection, and found that both TMV and TMGMV were isolated from the Australian *N. glauca* herbarium specimens collected before 1950; three of the pre-1950 samples yielded TMV alone, one TMGMV alone, four a mixture, and one a recombinant. Only TMGMV was isolated from more recent specimens. Fraile *et al*., (1997) also showed that although TMGMV virions accumulate to the same concentration in both doubly and singly infected plants, TMV virions attain only one-tenth the concentration in doubly infected plants than in singly infected plants. These reports, in addition to the early detection of TMGMV infection in *N. glauca*, are suggestive of TMGMV adaptation to this host plant.

Capsicum annuum L., is the most cultivated species of the genus *Capsicum* and belongs to the Solanaceae family. The world production of fresh pepper is 31 million tons, with China being the largest producer with 16 million tons per year (50% of production), followed

by Mexico (2.38 million tons), Turkey (2.08 million tons). Indonesia (1.66 million tons), USA (1.06 million tons) and Spain (1.02 million tons) (Liu et al., 2013). After its relatives, the potato (*Solanum tuberosum* L.) and the tomato (*Solanum lycopersicum* L.), pepper is the third most important crop of this family produced in Spain, which is the largest exporter in Europe, showing the importance of this crop in Spain (Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente, 2015).

Viral diseases cause significant losses in pepper production throughout the world (Moury and Verdin, 2012). Hanssen *et al.*, (2010) reported 49 virus species which were able to infect and cause damage in pepper production. Mechanically transmitted viruses like tobamoviruses are predominant in protected crops, whereas insect-transmitted viruses like potyviruses, cucumoviruses, and tospoviruses are more frequent and severe in open fields.

The seven species of *Tobamovirus* known to be able to infect *Capsicum* species are TMV, *Tomato mosaic virus* (ToMV), TMGMV, *Bell pepper mottle virus* (BPeMV), PaMMV, ObPV and PMMoV. Symptom severity varies with virus strain and *Capsicum* genotype and usually includes leaf distortion with chlorotic mottle or mosaic, and small, misshapen, and discoloured fruits. Often there are necrotic patches on fruits and leaves. These symptoms generally affect both quantity and quality of production (Kenyon *et al.*, 2014).

2 OBJECTIVES

With the relevant concepts and research background related to my thesis being presented in the Introduction, now I will point out the objectives of this thesis.

The general objective of the thesis is to study the evolution of host range expansion in plant RNA viruses, by acquiring either new host species or new host genotypes within a host species. I address the following specific objectives:

1) If adaptation of a virus to a new host species or genotype involves fitness penalties in the original host (i.e., across-host fitness trade-offs).
2) What are the mechanisms resulting in across-host fitness trade-offs, specifically which is the role of pleiotropy of and epistasis between, host range mutations.
3) What fitness components are affected by across-host fitness trade-offs, and
4) If trade-offs between different fitness components condition the evolution of host range expansion.

To achieve these objectives, I analysed host range expansion in two tobamovirus species infecting pepper crops, *Tobacco mild green mosaic virus* and *Pepper mild mottle virus*.

3 MATERIALS AND METHODS

3.1 **Viruses and Plants**

3.1.1 **Viral isolates and infectious cDNA clones**

Research in this thesis involved work with two virus species from genus *Tobamovirus*, family *Virgaviridae*:

Tobacco mild green mosaic virus (TMGMV). Twenty-six field isolates of TMGMV from two different hosts, pepper (*C. annuum*, 12 isolates) and *N. glauca* (14 isolates), both belonging to Solanaceae family, were collected in southeastern Spain between 1984 and 2004. Those isolates were then biologically cloned through two steps of necrotic local lesions passage in *Nicotiana glutinosa* L. For the present work, these isolates were multiplied again in their original host and virus particles were purified (Section 3.3). Table 3.1 shows the list of isolates. Isolates were identified with letters indicating the host of origin (P for pepper and Ng for *N. glauca*), two digits corresponding to the year of isolation and an ordinal.

Table 3.1. List of TMGMV field isolates[(a)]

Host of Origin							
Pepper Host				Nicotiana Host			
1	P 00/10	7	P 94/29	1	Ng 89/5	8	Ng 96/16
2	P 00/6	8	P 96/49	2	Ng 89/8	9	Ng 96/19
3	P 01/16	9	P 97/10	3	Ng 90/5	10	Ng 96/5
4	P 04/17	10	P 98/11	4	Ng 90/8	11	Ng 99/11
5	P 83/4	11	P 98/12	5	Ng 92/73	12	Ng 99/15
6	P 92/10	12	P 98/5	6	Ng 94/6	13	Ng 99/16
				7	Ng 96/11	14	Ng 99/20

[(a)] Isolates were identified with letters indicating the host of origin (P for pepper and Ng for *N. glauca*), two digits corresponding to the year of isolation and an ordinal.

Virus genotypes derived from two biologically active cDNA clones were also assayed here. Plasmid pTMGMV-Jap, containing a biologically active full-length cDNA clone derived from a pepper isolate of TMGMV (acc no. AB078435), described elsewhere (Morishima *et al.*, 2003) was generously provided by Dr. Tetsuro Okuno and Dr. Masanori Kaido, Laboratory of Plant Pathology, Kyoto University, Japan. A biologically active full-length cDNA clone, pTMGMV-FA, was derived from a field isolate, Ng 92/73 (see section 3.2). RNA transcripts from these clones were used to inoculate *Nicotiana clevelandii* A. Gray plants, and the resulting viruses were identified as TMGMV-Jap and TMGMV-FA.

Pepper mild mottle virus (PMMoV). Plasmid pPMMoV-MG, containing a biologically active full-length cDNA clone derived from field isolate P 84/8 (Fraile *et al*, 2011) was obtained by Manuel G. Moreno-Pérez, and is described elsewhere (Moreno-Pérez *et al*., 2016). RNA transcripts from pPMMoV-MG were called PMMoV genotype MG. This genotype is of $P_{1,2}$ pathotype (Moreno-Pérez *et al*, 2016), and is considered as the wild type (WT) in this study. On this clone, site-directed mutagenesis was used by Manuel G. Moreno-Pérez to introduce mutations in the CP gene that determine a change of pathotype from $P_{1,2}$ to $P_{1,2,3}$ and $P_{1,2,3,4}$ (Moreno-Pérez *et al*., 2016). The resulting viral genotypes (Table 3.2) were also used in this study.

Table 3.2. Description of PMMoV genotypes used in this work

Genotype	aa substitutions[(a)]	nt substitutions[(b)]	Pathotype
PMMoV-MG-WT	-	-	$P_{1,2}$ [(c)]
PMMoV-MG-(T43K)	T43K	C131A	N.I
PMMoV-MG-(D50G)	D50G	A152G	N.I
PMMoV-MG-(T43K+D50G)	T43K+D50G	C131A + A152G	$P_{1,2,3}$ [(c)]
PMMoV-MG-(L13F)	L13F	A42T	N.I
PMMoV-MG-(G66V)	G66V	G200T	N.I
PMMoV-MG-(L13F+G66V)	L13F + G66V	A42T + G200T	N.I
PMMoV-MG-(Q46L)	Q46L	A140T	N.I
PMMoV-MG-(M138N)	M138N	T433A + G434T	$P_{1,2,3}$ [(c)]
PMMoV-MG-(Q46L+M138N)	Q46L + M138N	A140T + T433A + G434T	N.I
PMMoV-MG-(Q46R)	Q46R	A140G	N.I
PMMoV-MG-(G85K)	G85K	G256A + G257A	N.I
PMMoV-MG-(Q46R+G85K)	Q46R + G85K	A140G + G256A + G257A	N.I
PMMoV-MG-(A86G)	A86G	G260C	$P_{1,2,3,4}$ [(c)]

[(a)] The two digits correspond to the position of amino-acid substituted in the coat protein (CP) from parental genotype PMMoV – MG-WT.
[(b)] The three digits correspond to the nucleotides substituted in the CP gene from parental genotype PMMoV – MG-WT.
[(c)] Described in Moreno-Pérez *et al*., 2016. N.I, Not Identified.

3.1.2 **Site-directed mutagenesis and characterization of the resulting mutant genotypes**

Mutations were introduced in the biologically active cDNA clones of TMGMV and PMMoV using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent), by following the instructions of the manufacturer with minor changes and using the primers designed especially for this purpose (see Tables 3.3 & 3.4). Reaction was set up in a total volume of 25 µl

containing 30 ng of plasmid, 1x reaction buffer, 0.2 µM of each primer, dNTP and QuickSolution reagent as recommended, and water. The cycling parameters were as follows - initial denaturation at 95°C / 2 min; 18 times amplification at 95 °C / 20 s; 60 °C / 10 s; 68 °C / 4 min and final elongation at 68 °C / 5 min. Followed by Dpn I digestion as recommended.

Table 3.3. Primers used for mutagenesis of TMGMV

Primer	Sequence[(a)]	Region	Position[(b)]	Substitution
A148T_F	5' CAGGCTTTGAGACTACT AGTGGACTTGTTTGG 3'	CP	6094-6126	Alanine for Threonine in TMGMV-Jap
A148T_R	5' CCAAACAAGTCCACTAG TAGTCTCAAAGCCTG 3'	CP	6094-6126	
T148A_F	5' CAGGCTTTGAGACTGCT AGTGGACTTGTTTGG 3'	CP	6094-6126	Threonine for Alanine in TMGMV-FA
T148A_R	5' CCAAACAAGTCCACTAG CAGTCTCAAAGCCTG 3'	CP	6094-6126	
J_118_F	5' GCGGGTAGCGGCCCA*G TCGACC* 3'	3' UTR + *Sal I* pUC118	6341-6356	To amplify pUC118
J_118_R	5' AATCTCACAACAATAGCT AAGTAGCTGG 3'	3' UTR	6135-6163	
5'FA _ rep	5' CTTAGCTATTGTTGTGAG ATTTCC 3'	3' UTR	6142-6165	Small 3' UTR of TMGMV-Jap for long 3' UTR of TMGMV-FA
3'FA _ rep	5' AG*GTCGAC*TGGGCCGC TACCCG 3'	pUC118 *Sal I* + 3' UTR	6342-6356	
FA_118_ F	5' GCGGGTAGCGGCCCA*G AGCTC*G 3'	3' UTR + *Sac I* pUC118	6342-6356	Long 3' UTR of TMGMV-FA for small 3' UTR of TMGMV-Jap
3'Jap _ rep	5' C*GAGCTC*TGGGCCGCTA CCCG 3'	pUC118 *SacI* + 3' UTR	6343-6356	

[(a)] The mutated nucleotide is underlined.
[(b)] The position in nucleotides according to the sequence of TMGMV-Jap (accession no. AB078435.1).

In TMGMV, the following mutations were introduced in the inserts of plasmids pTMGMV-Jap and pTMGMV-FA .

- TMGMV-Jap-(A148T). Substitution of guanine to adenine at position 442 of CP gene in pTMGMV-Jap resulted in amino-acid change from Alanine (Ala) to Threonine (Thr) at position 148 of CP. Primer pair, A148T_F and A148T_R (Table 3.3) was used to introduce this mutation in clone, pTMGMV-Jap. This mutant genotype denoted as TMGMV-Jap-C in Table 3.5.

- TMGMV-Jap-(Short 3'UTR Long). Short 3' untranslated region (UTR, 210 nucleotides) of genotype TMGMV-Jap was replaced with long 3' UTR (357 nucleotides) of genotype TMGMV-FA. Primers, 5'FA_rep, 3'FA_rep, J_118_F and J_118_R (Table 3.3) were used to mutate the 3' UTR of TMGMV-Jap. This mutant genotype denoted as TMGMV-Jap-U in Table 3.5.

Table 3.4. Primers used for mutagenesis in the coat protein gene of PMMoV

Primer	Sequence[(a)]	Position[(b)]	Substitution
T43K-F	5' CACAACAGGCTAGAACTAAGGTTC AACAGCAGTTCTCTG 3'	113-151	Threonine for Lysine
T43K-R	5' CAGAGAACTGCTGTTGAACCTTAG TTCTAGCCTGTTGTG 3'	113-151	
D50G-F	5' GGTTCAACAGCAGTTCTCTGGTGT GTGGAAGACTATTCCG 3'	132-171	Aspartic acid for Glycine
D50G-R	5' CGGAATAGTCTTCCACACACCAGA GAACTGCTGTTGAACC 3'	132-171	
L13F-F	5' CCAATCAATTAGTGTATTTTGGTTC TGTATGGGCTGATCC 3'	23-62	Leucine for Phenylalanine
L13F-R	5' GGATCAGCCCATACAGAACCAAAA TACACTAATTGATTGG 3'	23-62	
G66V-F	5' CGACCGCTACAGTTAGATTTCCTG CTACTGTTTTCAAAGTTTTCC 3'	170-214	Glycine for Valine
G66V-R	5' GGAAAACTTTGAAAACAGTAGCAG GAAATCTAACTGTAGCGGTCG 3'	170-214	
Q46L-F	5' GAACTACGGTTCAACTGCAGTTCT CTGATGTG 3'	125-156	Glutamine for Leucine
Q46L-R	5' CACATCAGAGAACTGCAGTTGAAC CGTAGTTC 3'	125-156	
Q46R-F	5' GGCTAGAACTACGGTTCAACGGCA GTTCTCTGATGTGTGG 3'	120-159	Glutamine for Arginine
Q46R-R	5' CCACACATCAGAGAACTGCCGTTG AACCGTAGTTCTAGCC 3'	120-159	
G85K-F	5' GTGTCGGCACTTCTCAAAGCCTTT GATACTAGGAACAGG 3'	241-279	Glycine for Lysine
G85K-R	5' CCTGTTCCTAGTATCAAAGGCTTT GAGAAGTGCCGACAC 3'	241-279	

[(a)] The mutated nucleotide is underlined.
[(b)] The position in nucleotides according to the sequence of PMMoV-MG-WT (accession no. KX063611).

- TMGMV-FA-(T148A). Substitution of adenine to guanine at position 442 of CP gene in pTMGMV-FA resulted in amino-acid change from Thr to Ala at position 148 of CP. Primer pair, T148A_F and T148A_R (Table 3.3) was used to introduce this mutation in clone, pTMGMV-FA. This mutant genotype denoted as TMGMV-FA-C in Table 3.5.

- TMGMV-FA-(Long 3'UTR Short). Long 3' UTR of genotype TMGMV-FA was replaced with Short 3' UTR of genotype TMGMV-Jap. Primers, 5'FA_rep, 3'Jap_rep, FA_118_F and J_118_R (Table 3.3) were used to mutate the 3' UTR of TMGMV-FA. This mutant genotype denoted as TMGMV-FA-U in Table 3.5.

Table 3.5. Description of TMGMV genotypes used in this work

Genotype[(a)]	aa[(b)]	3' UTR
TMGMV-Jap	Ala	Short
TMGMV-Jap-CP	Thr	Short
TMGMV-Jap-UTR	Ala	Long
TMGMV-FA	Thr	Long
TMGMV-FA-CP	Ala	Long
TMGMV-FA-UTR	Thr	Short

[(a)] Mutant genotypes are identified by a letter at the end of parental genotype's name indicating the site of mutation (CP for coat protein and UTR for 3' untranslated region).
[(b)] Amino-acid at position148 of coat protein.

In PMMoV, mutations were introduced in the plasmid pPMMoV-MG whose combinations were the determinants of overcoming the resistance of alleles L^3 and L^4. Therefore, the following mutants were obtained (Table 3.2):

- PMMoV-MG-(T43K).The mutation was introduced at nucleotide 131 of CP gene in pPMMoV-MG by mutating cytosine to adenine, leading to a change of amino-acid, from threonine (T) to lysine (K) at position 43 of CP. Primer pair, T43K-F and T43K-R (Table 3.4) was used for introducing T43K mutation.

- PMMoV-MG-(D50G). The mutation was introduced at nucleotide 152 of CP gene in pPMMoV-MG by mutating adenine to guanine, leading to a change of amino-acid, from aspartic acid (D) to glycine (G) at position 50 of CP. Primer pair, D50G-F and D50G-R (Table 3.4) was used for introducing D50G mutation.

- PMMoV-MG-(L13F). The mutation was introduced at nucleotide 42 of CP gene in pPMMoV-MG by mutating adenine to threonine, leading to a change of amino-acid, from leucine (L) to phenylalanine (F) at position 13 of CP. Primer pair, L13F-F and L13F-R (Table 3.4) was used for introducing L13F mutation.

- PMMoV-MG-(G66V). The mutation was introduced at nucleotide 200 of CP gene in pPMMoV-MG by mutating guanine to threonine, leading to a change of amino-acid, from glycine (G) to valine (V) at position 66 of CP. Primer pair, G66V-F and G66V-R (Table 3.4) was used for introducing G66V mutation.

- PMMoV-MG-(L13F + G66V). The mutation was introduced at nucleotide 200 of CP gene in pPMMoV-MG-(L13F) by mutating guanine to threonine, leading to a change of amino-acid, from glycine (G) to valine (V) at position 66 of CP. Primer pair, G66V-F and G66V-R (Table 3.4) was used for introducing G66V mutation.

- PMMoV-MG-(Q46R). The mutation was introduced at nucleotide 140 of CP gene in pPMMoV-MG by mutating adenine to guanine, leading to a change of amino-acid, from glutamine (Q) to arginine (R) at position 46 of CP. Primer pair, Q46R-F and Q46R-R (Table 3.4) was used for introducing Q46R mutation.

- PMMoV-MG-(G85K). The mutation was introduced at nucleotides 256 and 257 of CP gene in pPMMoV-MG by mutating guanines to adenines, leading to a change of amino-acid, from glycine (G) to lysine (K) at position 85 of CP. Primer pair, G85K-F and G85K-R (Table 3.4) was used for introducing G85K mutation.

- PMMoV-MG-(Q46R + G85K). The mutation was introduced at nucleotide 140 of CP gene in pPMMoV-MG-(G85K) by mutating adenine to guanine, leading to a change of amino-acid, from glutamine (Q) to arginine (R) at position 46 of CP. Primer pair, Q46R-F and Q46R-R (Table 3.4) was used for introducing Q46R mutation.

- PMMoV-MG-(Q46L). The mutation was introduced at nucleotide 140 of CP gene in pPMMoV-MG by mutating adenine to thymine, leading to a change of amino-acid, from glutamine (Q) to leucine (L) at position 46 of CP. Primer pair, Q46L-F and Q46L-R (Table 3.4) was used for introducing Q46L mutation.

- PMMoV-MG-(Q46L + M138N). The mutation was introduced at nucleotide 140 of CP gene in pPMMoV-MG-(M138N) by mutating adenine to thymine, leading to a change

of amino-acid, from glutamine (Q) to leucine (L) at position 46 of CP. Primer pair, Q46L-F and Q46L-R (Table 3.4) was used for introducing Q46L mutation.

3.1.3 **Host plants**

Experiments were performed in plants from the following species: *N. glauca*; *C. annuum cv.* Doux des Landes (L^+/L^+); and *C. annuum cv*. Dulce Italiano (L^+/L^+), used as hosts for TMGMV; and *C. annuum cv.* Dulce Italiano (L^+/L^+); *C. annuum cv.* Yolo Wonder (L^1/L^1); *C. frutescens cv.* Tabasco (L^2/L^2); *C. chinense* PI 159236 (L^3/L^3); *C. chacoense* PI260429 (L^4/L^4), used as hosts for PMMoV. For multiplication of viruses, *N. clevelandii,* and for diagnostic, *Nicotiana tabacum* L. *cv* Xanthi were used. Plants were grown in 15 cm diameter, 1.5 L pots at 23 to 25 °C and with a 16 h light photoperiod in a P1 / P2-level biological containment greenhouse depending on the virus treatment.

3.2 **Construction of an infectious clone and generation of RNA transcripts**

The complete genome of a TMGMV isolate from *N. glauca*, Ng 92/73, was amplified by reverse transcription – polymerase chain reaction (RT-PCR) using the SuperScript III high-yield reverse transcriptase (Life Technologies) and a SuperMix high-fidelity DNA polymerase (Life Technologies) with primers TMGMV_T7_F and TMGMV_SacI_R (Table 3.6). The 5' end of the forward primer (TMGMV_T7_F) had the sequence for the T7 promoter and the reverse pr-

Table 3.6. Primers used for TMGMV cloning

Primer	Sequence	Region[(a)]	Position[(b)]	Use
TMGMV_T7_ F	5' CCA*GTCGAC***TAATACGACT CACTATAG**GATGTTTTAATAG TTTTCGACAACAAC 3'	*Sal I* + **T7 promoter** + 5' UTR	1-25	Full genome clone
TMGMV_SacI_ R	5' ATA*GAGCTC*TGGGCCGCTA CCCGCGGTT 3'	*Sac I* + 3' UTR	6337-6356	
pUC118_R	5' **TAGTGAGTCGTATTA***GTCG AC*CTGCAGGCATGC 3'	part of **T7 promoter** + *Sal I* + pUC118	1-18	To amplify full plasmid
pUC118_F	5' GCGGGTAGCGGCCCA*GAG CTC*GAATTCGTAATCATGG 3'	3' UTR + *Sac I* + pUC118	3140-3162	

[(a)] which part of the genome.
[(b)] The position in nucleotides according to the sequence of TMGMV-Jap (accession no. AB078435.1) .

imer (TMGMV_SacI_R) had a Sac I restriction site. Using Gibson assembly (Gibson *et al.*, 2009) strategy the resulting amplicon was cloned into the pUC118 plasmid implemented by the NEBuilder HiFi DNA Assembly Cloning kit (New England Biolabs). Cloning was performed as mentioned in the manufacturer's protocol. This biologically active full length cDNA clone of TMGMV was dedicated to Fernando-Aurora (FA) for their work with TMGMV for over three decades. Thus, plasmid pTMGMV-FA was obtained and the virus derived from transcripts was called TMGMV-FA.

3.2.1 *In vitro* transcription

To initiate transcription, all the clones were digested at the restriction site flanking the 3' end of the viral genome. For this, at least 10 µg of each plasmid was linearized with 20 U of the corresponding enzyme. The enzyme, Sac I (New England Biolabs) linearizes pTMGMV-FA and its derived mutants, Sal I (New England Biolabs) linearizes pTMGMV-Jap and its derived mutants and Not I (New England Biolabs) releases the PMMoV insert from pPMMoV-MG and its derived mutants (Figure 3.1). After a phenol-chloroform extraction, the linearized plasmids were incubated at 37 °C for 3 hours with a mixture of 5 units of T7 RNA polymerase (New England Biolabs), a 4: 1 ratio of the CAP analog m^7G (5') ppp (5') G (New England Biolabs)

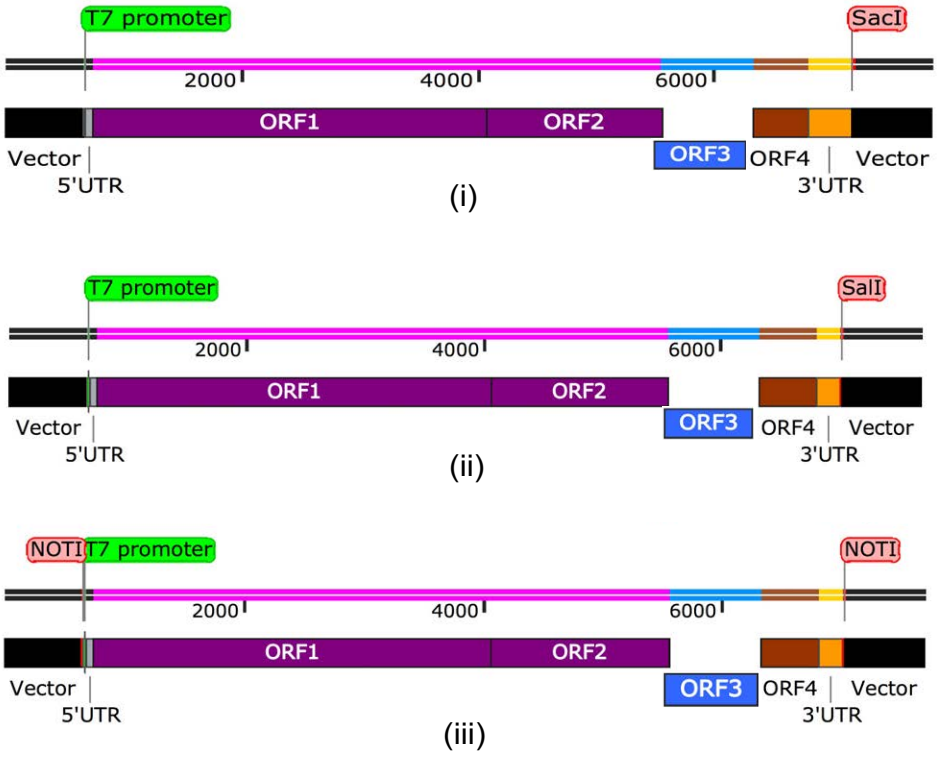


Figure 3.1 Linearized map of plasmids pTMGMV – FA (i), pTMGMV – Jap (ii) and pPMMoV – MG (iii), showing different restriction sites flanking each biologically active full length cDNA clone.

and 2.5 µl of rNTPs (10 mM) (New England Biolabs), followed by DNase (5 U) digestion at 37 °C for 15 min. The obtained RNA was purified and precipitated following the protocol described in section 3.4.2.

3.3 **Purification of viral particles and nucleic acids**

3.3.1 **Purification of viral particles**

TMGMV and PMMoV particles were purified using a modified protocol from that described by Bruening *et al*, 1976. Around 30 – 50 g of infected leaves were ground in 40 mM EDTA, 36 mM NaOH and 0.4% of 2-mercaptoethanol in a ratio volume/weight of 6. The homogenate was centrifuged at 3,000 rpm for 10 min and filtered through four layers of gauze. The filtrate was centrifuged at 13,000 rpm for 20 min and the supernatant was filtered through Whatman # 1 paper (GE Healthcare Life Sciences). The filtrate was incubated on ice for 30 min after adding 0.75 ml of 20% Triton X-100 per 100 ml of filtrate, and then it was centrifuged for one hour at 40,000 rpm, 4 ^{o}C. The pellet obtained was re-suspended overnight by stirring in 1 mM EDTA. The re-suspended pellet was applied to a 6 ml cushion of 30% sucrose in 1mM EDTA, and centrifuged for one hour at 40,000 rpm, 4 ^{o}C. The supernatant was discarded and the pellet containing virus particles was carefully washed with distilled water. Finally, it was re-suspended in 1 mM EDTA by stirring overnight.

The concentration of viral particles was quantified by measuring the absorbance of the suspension at 260 nm. For TMGMV, the extinction coefficient (EC) of 3.16 and for PMMoV, a EC of 3.18 was used (Adams and Antoniw, 2006).

3.3.2 **Extraction of encapsidated viral RNA**

Viral RNA was extracted from particles using a modified protocol from that described by Mandeles, 1968. Viral particles (20-40 µg) were incubated for 10 min in 0.1 M Tris-HCl (pH 8.3), 1% SDS and water at 60 °C. To this, two volumes of phenol-chloroform mixture in ratio (1: 1) was added and homogenized. The homogenate was centrifuged at 10,000 rpm for 10 min. The aqueous phase was re-extracted with phenol-chloroform. Afterwards, RNA was precipitated by 16 hours incubation in 0.08 M sodium acetate with 2.5 volumes of absolute ethanol at -20 °C, followed by centrifugation at 13,000 rpm for 20 min at 4 °C. The sedimented RNA was air dried at room temperature for 10 min and re-suspended with 50 µl of sterile, ribonuclease-free water.

The concentration of viral RNA preparation was determined by measuring the absorbance at 260 nm on a NanoDrop® ND-1000 spectrophotometer (Thermo Scientific). For ssRNA, a E.C of 40 was used.

3.4 Inoculation of plants and extraction of total RNA for detection and quantification of viral RNA

3.4.1 Inoculations

Plants were inoculated at the fully-expanded first two true leaves with either RNA transcripts from biologically active cDNA clones (1 μg per plant) or purified viral particles (400 ng per plant), in 0.1 M phosphate buffer of pH 7.2.

RNA transcripts were inoculated in *Nicotiana benthamiana* D. or *N. clevelandii* and purified viral particles were inoculated in *C. annuum* and *N. glauca*, for all the experiments related to this thesis. Carborundum (400 mesh) was also sprayed over leaves before inoculation to facilitate infection.

3.4.2 Extraction of total RNA from plant tissue

Total RNA was extracted from plant tissue using TRIzol® Reagent (Life Technologies). To this end, 0.1 g of leaf tissue was crushed in liquid nitrogen in an Eppendorf tube (1.5 ml) using a sterilized pipette tip. TRIzol was added in a ratio volume/weight of 10, vortexed for 15 s and incubated for 10 min at room temperature. Chloroform was then added in a ratio volume/weight of 2, homogenized and incubated for 5 min at room temperature. The homogenate was centrifuged at 13,000 rpm for 20 min at 4 °C and 2/3 of the aqueous phase was collected. Isopropanol was added to the aqueous phase in a ratio volume/weight of 5, vortexed and incubated at 4°C for 15 min, followed by centrifugation at 13,000 rpm for 20 min at 4 °C. The obtained RNA pellet was re-suspended in 50 μl of sterile, ribonuclease-free water. Afterwards, RNA was re-precipitated by overnight incubation in 0.08 M sodium acetate with 2.5 volumes of absolute ethanol at -20 °C and centrifugation at 13, 000 rpm for 20 min at 4 °C. The sedimented RNA was washed with 400 μl of 75% ethanol and was allowed to dry for 15 min. The RNA was re-suspended in 30 μl of sterile, ribonuclease-free water.

The total RNA concentration was determined by measuring the absorbance at 260 nm on a NanoDrop® ND-1000 spectrophotometer (Thermo Scientific).

3.4.3 **Detection of viral RNA by reverse transcription – polymerase chain reaction (RT- PCR)**

The CP gene of TMGMV and PMMoV was amplified by RT-PCR. The cDNA was synthesized in a mixture containing 10 pg - 500 ng of viral RNA, 200 units (U) of the enzyme RT Super Script III (Invitrogen), 0.5 mM dNTPs, 5 mM DTT, 40 U from RNAse OUT (Invitrogen) and 0.2 mM of the viral RNA specific primer (see Table 3.7). This mixture was incubated at 55 °C for 30 min for cDNA fragments less than 1,000 nucleotides, or for 1 h for larger fragments (approx. 6 kb). Subsequently the mixture was incubated at 70 °C for 15 min. Afterwards, the PCR was performed in a reaction containing 5 μl of cDNA, 1X reaction buffer (Biotools), 2 mM $MgCl_2$, 200 μM dNTP, 500 μM of each primer and 0.5 U of DNA polymerase (Biotools). The thermal parameters for PCR were: initial denaturation at 94 °C / 5 min, 30 cycles of 94 °C / 30 s, 57 °C / 30 s and 72 °C/ 50 s and final elongation at 72 °C / 5 min.

3.4.4 **Quantification of viral multiplication**

Virus multiplication was quantified as viral RNA accumulation via real-time reverse transcriptase-quantitative polymerase chain reaction (RT-qPCR). For each plant, three leaf discs, 5mm in diameter, were randomly excised from systemically infected leaves. Within-host virus multiplication was estimated at 10 days post inoculation (dpi) for TMGMV in *C. annuum*, 21 dpi for TMGMV in *N. glauca* and at 21 dpi for PMMoV.

For each sample, 0.5–3 ng of total RNA were utilized with Brilliant III Ultra-Fast SYBR Green QRT-PCR Master Mix (Agilent Technologies) according to the manufacturer's recommendations in a final volume of 10 µl. Assays were performed in triplicate on a LightCycler 480 II real-time PCR system (Roche). Primers, POb-Q fwd and PO-Q rev (Table 3.7) for TMGMV and P12 / P123f-Q fwd and P12 / P123f2-Q rev (Table 3.7) for PMMoV were used to quantify the viral accumulation. Reactions including no template and no reverse transcriptase were included in each trial. A standard curve was generated with the help of serial dilutions of purified virion RNA in the range of 10 to 10^{-6} with tenfold dilution factor. Levels of viral RNA was deduced from comparison with the standard curve. Thermal parameter for real-time quantification were 50 °C for 10 min, 95 °C for 3 min, and 40 cycles of 95 °C for 5 s

and 60 °C for 10 s. Dissociation curves were generated to ascertain that only a single product was produced in each case.

Table 3.7. Primers used for detection and quantification of viruses

Primers	Sequence	Position[(a)]	Purpose
TMGMVRNA3	5' AARTAAATAAYAGTGG TAAGAAGGG 3'	5590-5565	Detection of TMGMV
CGM3	5' TGGGCCSCWACCSGS GG 3'	6339-6355	Detection of both TMGMV and PMMoV
PMMoVRNA3	5' GGGTTTGAATAAGGA AGGGAAGC 3'	5617-5595	Detection of PMMoV
P0b-Q fwd	5' GCACCGAATACTACT GAAATCG 3'	5970-5991	Quantification of TMGMV
P0-Q rev	5' GCCTGCTTGATTGAA CATGCCAGT 3'	6075-6098	
P12 / P123f-Q fwd	5' AACTGCCGAGACGCT T 3'	5996-6011	Quantification of PMMoV
P12 / P123f2-Q rev	5' GAGTTGTAGCCCAGG TG 3'	6137-6153	

[(a)] The position in nucleotides according to the sequence of TMGMV-Jap (accession no. AB078435.1) and PMMoV-MG-WT (KX063611).

3.5 Calculation of virus fitness, epistasis and virulence

3.5.1 Virus fitness

Virus fitness was estimated from viral RNA concentration as described in Lalic *et al*., (2011) using the Malthusian parameter (*m*), which represents the population exponential growth rate at a time *t* after inoculation, calculated as $m = (1/t) \ln(Q)$, where Q is virus accumulation. Fitness, *W*, is then computed as $W = e^m$ (Lalic *et al*., 2011).

3.5.2 Epistasis

Epistasis between pairs of mutations *x* and *y*, ε_{xy}, was calculated as $\varepsilon_{xy} = W_{00}W_{xy} - W_{x0}W_{0y}$ (Kouyos *et al*., 2007), where W_{00}, W_{xy}, W_{x0}, and W_{0y} correspond to the absolute fitness of the wild-type, the double mutant and each single mutant, respectively.

3.5.3 **Virulence**

Virulence (*V*) was quantified as the effect of infection on the above-ground host plant biomass, calculated as $V = 1 - (P_i / P_m)$, where P_i is the dry weight of each infected plant and P_m is the mean dry weight of mock-inoculated plants (Pagán *et al*, 2007). Dry weights were determined after incubation of above-ground host plant at 65°C for 2 weeks.

3.6 ***In vitro* stability of virus particles**

The stability of the PMMoV particles was evaluated by *in vitro* disassembly assay under conditions of basic pH and / or high ionic strength as described in Bera *et al*., 2017. Briefly, the kinetics of disassembly of the viral particles was analyzed in three conditions:- 0.1 M Tris-Cl pH 8.75, 6 M Urea at pH 7.4 and pH 10.0 and was incubated for different durations 5 min, 15 min, 30 min, 45 min, 60 min and 90 min on ice. Then, the degree of disassembly was determined by the densitometry analysis of the electrophoresis gel.

3.7 **Determination of nucleotide sequences**

The complete nucleotide sequence of the genomic RNA of TMGMV-FA was determined using primers (Table 3.8) that yield cDNA amplicons of around 800 nucleotides, each amplified region overlapping 100 to 200 nucleotides with contiguous ones, covering the entire genome of the virus. The sequencing was carried out by StabVida (Portugal). All the contigs were assembled using CAP3 Sequence Assembly Program (Huang and Madan, 1999).

The nucleotide sequence of the CP gene plus 3' UTR of 26 field isolates of TMGMV was also determined from virion RNA using the primer pair TMGMVRNA3 and CGM3 (Table 3.7).

3.8 **Estimation of genetic diversity**

The sequences of the CP gene plus 3'UTR of 26 TMGMV isolates were aligned using MUSCLE 3.7 (Edgar, 2004) and adjusted manually. The population diversity of TMGMV isolates was calculated in terms of nucleotide substitutions per site and was estimated by using the Tamura 3-parameter model for nucleotide substitution as implemented in Mega 6.06. This was the best-fitted nucleotide substitution model selected by the corrected Akaike

information criterion as determined by jModelTest 0.1.1 (Posada, 2008). Standard errors of each genetic diversity measure were based on 1,000 replicates bootstrap using MEGA 6.06.

Table 3.8. Primers for sequencing the whole genome of TMGMV

Primer	Sequence	Position [(a)]
TMGMV 5a-FW	5' GATGTTTTAATAGTTTTCGACAACAACAA 3'	1 – 29 nt
801 REV	5' CGTAACCTCCGTCTGGTCTAG 3'	801 - 781 nt
637 FW	5' CATCCRCCAGAGAATAGTGG 3'	637 - 656 nt
1460 REV	5' GCAGCCAGCTTAGTCTGC 3'	1460 - 1443 nt
1290 FW	5' CAGRACATATCAAGCCAAAGCG 3'	1290 - 1311 nt
2127 REV	5' CACRCTATCCACGCAAGC 3'	2127 - 2110 nt
1900 FW	5' GATAAGCCAACCGAGGAGA 3'	1900 - 1918 nt
2723 REV	5' TCTACCGTTCTCACATTGTCC 3'	2723 - 2703 nt
2531 FW	5' ACGAACCTACTGCAAAGATGG 3'	2531 - 2551 nt
3331 REV	5' CTAAAGGATCTAACACTACGG 3'	3331 - 3311 nt
3164 FW	5' TGAACACCGTTCATGAGATCC 3'	3164 - 3184 nt
3976 REV	5' GTTGAGCCTTGATCATGTGC 3'	3976 - 3957 nt
3753 FW	5' CGATTGACATTGAGAGCACYGC 3'	3752 - 3773 nt
4647 REV	5' CACTATTGCTCCCTTATCATGG 3'	4647 - 4626 nt
4443 FW	5' GGGYTCAATGTTACCKATGG 3'	4443 - 4462 nt
5294 REV	5' GCTTYTCAGCATCCTCGG 3'	5294 - 5277 nt
5108 FW	5' GTAGTGTCTGGGGAGTGG 3'	5108 - 5125 nt
5909 REV	5' ATCAACGGATCAAGCGTCG 3'	5909 - 5891 nt

[(a)] The position in nucleotides according to the sequence of TMGMV-Jap (accession no. AB078435).

3.9 **Statistical analyses**

Previous to statistical analyses, outliers in data distributions were detected by calculating the residual value of each data according to a Student's *t*, divided by its standard deviation. Values outside the 95% confidence interval of a Student *t* distribution, calculated by using all residual values, were considered as outliers and were eliminated (Sokal and Rohlf, 1994).

The distribution of the values of all variables analyzed was tested for normality using the Shapiro-Wilk test (Sokal and Rohlf, 1995) and were also tested for homogeneity of variances using the Levene test (Sokal and Rohlf, 1995). It was concluded that most of the

distributions were not normal nor homoscedastic, therefore, appropriate non-parametric tests (for example, Mann-Whitney Test, Kruskal-Wallis Test) were used (Sokal and Rohlf, 1995).

Differences in viral accumulation, viral fitness and virulence were analyzed using Generalized Linear Models (GzLM). The GzLM analysis was selected because it is a robust method with respect to the distribution of the data and allows contrasting both balanced and non-balanced models. To determine if the observed differences between classes of the same factor were significant, least significant difference (LSD) analyses were performed.

All statistical analyses were performed using the SPSS v.24.0 program (SPSS Inc., IL, USA).

4 RESULTS

4.1 **Adaptation of Tobacco mild green mosaic virus isolates to different host species**

From the southeastern part of Spain, our lab had collected field isolates of TMGMV from two different hosts, pepper (*C. annuum*, 12 isolates) and *N. glauca* (14 isolates) growing in proximity. Both hosts belong to the Solanaceae and were highly susceptible to TMGMV infection, which prompted me to investigate the possibility that this virus has adapted to these two different hosts. For this, the isolates were molecularly characterized and a reciprocal inoculation experiment was performed.

4.1.1 **Genetic diversity of TMGMV isolates**

The CP gene plus the 3' UTR of 26 TMGMV field isolates were successfully RT-PCR amplified, and the nucleotide sequences of the amplicons were determined (Annex I). On the basis of this information the population diversity of the TMGMV isolate collection, estimated as nucleotide substitutions per site, was found to be 0.011 ± 0.002. Forty-four nucleotide positions in the CP gene and 16 in the 3' UTR were polymorphic. Most polymorphisms occurred in less than four out of the 26 isolates, except for two dimorphic positions in the CP gene and one in the 3' UTR that showed a higher frequency of each allele (Tables 4.1.1 and 4.1.2): The nucleotide position 22 in the CP gene showed a C/T dimorphism (C in 17 and T in

Table 4.1.1. Polymorphisms in the coat protein gene of TMGMV

Positions[a]	22	48	51	52	53	55	56	57	59	60	63	64	72	79	84	88	105	108	123	129	155	174	176	180	195	201	204	207	210	213	219	253	255	256	264	297	309	319	343	344	387	442	463	464	471
Consensus Sequence	C	C	T	T	A	G	C	A	A	T	T	G	A	C	C	A	G	T	T	A	C	T	T	A	A	T	C	T	G	T	T	T	A	C	T	G	G	G	T	A	G	A	C	G
P 00/10	.	.	.	.	.	.	.	.	.	.	.	.	G	.	T	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	C	.	.	.	.	G	.	.	.	.
P 00/6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P 04/17	.	A	.	.	.	.	.	.	.	.	.	.	G	.	T	.	.	.	C	.	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	.	G	.	C	A	A
P 96/49	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P 98/11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P 98/12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P 01/16	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
P 83/4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
P 92/10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
P 94/29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	A	.	.	.	.
P 97/10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	C	.	.	.	.	A	.	.	.	.
P 98/5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 89/8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 90/5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	G	A	.	.	.	.
Ng 94/6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 96/19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 96/5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 92/73	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 96/11	T	.	.	.	A	G	.	.	.	A	A	.	.	.	T	.	.	.	.	.	C	.	.	G	.	.	C	.	.	.	.	.	.	.	.	A	A	T	A	A	.	.	.	.
Ng 99/20	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.
Ng 89/15	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 90/8	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 96/16	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 99/11	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.
Ng 99/15	T	.	G	G	G	A	G	G	G	C	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.
Ng 99/16	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

[a] The 480 nucleotide positions in the coat protein (CP) gene are numbered starting from the first nucleotide of the CP's starting codon.

9 isolates), and position 442, showed a G/A dimorphism (G in 12 and A in 14 isolates). The nucleotide dimorphisms at these positions translated into the amino-acid dimorphisms proline/serine at position 8 of the CP and threonine/alanine at position 148. Table 4.1.3 indicates that the amino-acid at position 8 of the CP was a proline in all pepper isolates (12) and in 5 / 14 *N. glauca* isolates, and a serine in 9 / 14 *N. glauca* isolates, the amino acid at this position being significantly associated with the isolate's host of origin ($\chi^2_{(1,26)} = 11.79$, $P = 0.001$). The amino acid at position 148 of the CP was a threonine in 6 / 12 pepper and 8 / 14 *N. glauca* isolates, and an alanine in 6 / 12 pepper and 6 / 14 *N. glauca* isolates (Table 4.1.3), the amino acid at position 148 not being significantly associated with the isolate's host of origin ($\chi^2_{(1,26)} = 0.13$, $P = 0.716$). In the 3' UTR, nine isolates (Table 4.1.2) had a 147 nt duplication and one isolate, Ng 96/5, had a 165 nt duplication, all after position 64. As a result, 3 / 12 pepper and 7 / 14 *N. glauca* isolates had a long 3' UTR (due to duplications) and 9 / 12 pepper and 7 / 14 *N. glauca* isolates had short 3' UTR (Table 4.1.3). The short / long 3' UTR was not significantly associated with the isolate's host of origin ($\chi^2_{(1,26)} = 2.73$, $P = 0.098$).

Table 4.1.2. Polymorphisms in the 3' untranslated region of TMGMV

Positions [a]	2	4	13	21	43	56	57	65	238	287	298	304	307	320	333	356
Consensus Sequence	T	T	G	A	T	G	C	Dup Yes [b]	C	G	A	T	T	A	A	T
P 00/10	.	.	.	.	A	.	.	−	T	A				G	T	.
P 00/6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A
P 04/17	.	.	.	.	.	A	.	−	T	A	.	.	.	.	T	.
P 96/49	.	C	.	.	.	.	.	−	.	.	.	.	.	.	.	.
P 98/11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P 98/12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
P 01/16	.	.	.	.	.	.	.	−	.	.	G	C	.	.	.	.
P 83/4	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.
P 92/10	.	.	.	.	.	.	.	−	T	.	.	.	.	.	.	.
P 94/29	.	.	.	.	A	.	T	−	.	.	.	.	C	.	.	.
P 97/10	.	.	.	.	A	.	.	−	.	.	.	.	C	.	.	.
P 98/5	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.
Ng 89/8	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	A
Ng 90/5	.	.	.	.	.	.	.	−	.	.	.	.	C	.	.	.
Ng 94/6	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
Ng 96/19	.	.	.	.	.	.	.	−	.	.	.	.	C	.	.	.
Ng 96/5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 92/73	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 96/11	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 99/20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 89/15	.	.	.	.	.	.	.	−	T	.	.	.	.	.	.	.
Ng 90/8	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.
Ng 96/16	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.
Ng 99/11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Ng 99/15	C	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Ng 99/16	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.

[a] Positions in the 3' untranslated region (UTR) are numbered considering as position 1 the first nucleotide after the coat protein's stop codon.
[b] The presence of a duplicated fragment (147 nt / 165 nt) after position 64 of the 3' UTR is indicated by a dot; − indicates no duplication.

Table 4.1.3. Genotype of TMGMV isolates according to three dimorphic loci in the coat protein gene and 3' untranslated region

Isolates	aa at 8[(a)]	aa at 148[(a)]	3' UTR[(b)]
P 00/10	Proline	Alanine	Short
P 00/6	Proline	Alanine	Long
P 04/17	Proline	Alanine	Short
P 96/49	Proline	Alanine	Short
P 98/11	Proline	Alanine	Long
P 98/12	Proline	Alanine	Long
P 01/16	Proline	Threonine	Short
P 83/4	Proline	Threonine	Short
P 92/10	Proline	Threonine	Short
P 94/29	Proline	Threonine	Short
P 97/10	Proline	Threonine	Short
P 98/5	Proline	Threonine	Short
Ng 89/8	Proline	Threonine	Short
Ng 90/5	Proline	Threonine	Short
Ng 94/6	Proline	Threonine	Long
Ng 96/19	Proline	Threonine	Short
Ng 96/5	Proline	Threonine	Long
Ng 92/73	Serine	Threonine	Long
Ng 96/11	Serine	Threonine	Long
Ng 99/20	Serine	Threonine	Long
Ng 89/15	Serine	Alanine	Short
Ng 90/8	Serine	Alanine	Short
Ng 96/16	Serine	Alanine	Short
Ng 99/11	Serine	Alanine	Long
Ng 99/15	Serine	Alanine	Long
Ng 99/16	Serine	Alanine	Short

[(a)] Position of amino-acid in coat protein.
[(b)] The 3' untranslated region had been designated as long if it has a duplication after position 64, and as short if it has no duplication.

A maximum-likelihood (ML) phylogenetic tree was inferred from the concatenated sequences of the CP gene and the 3'UTR, which showed that all isolates belonged to a single cluster, with the exception of isolates P 00/10 and P 04/17 (Fig 4.1). Removing the duplication of the 3'UTR did not affect this result (not shown).

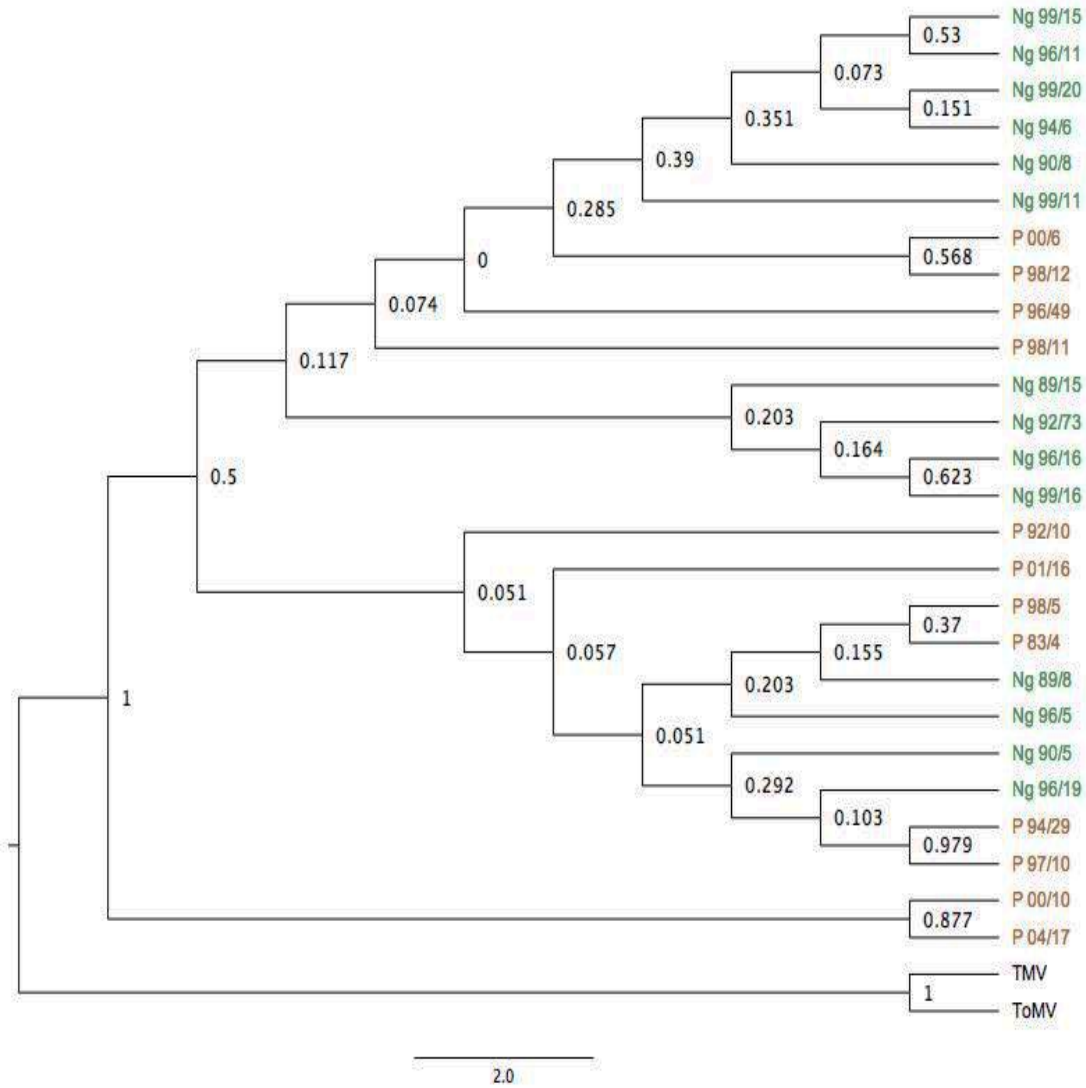


Figure 4.1 Maximum Likelihood phylogeny of TMGMV isolates based on the concatenated nucleotide sequence of the coat protein gene and the 3' untranslated region. Significance of nodes in a bootstrap analysis with 1,000 replicates is indicated. The tree was rooted using the homologous sequences of *Tobacco mosaic virus* (TMV) and *Tomato mosaic virus* (ToMV) as outgroups.

4.1.2 **Multiplication of TMGMV isolates in pepper and in *N. glauca***

An experiment was performed to investigate the possibility of adaptation of TMGMV isolates to the two hosts of origin: in a reciprocal inoculation experiment, the TMGMV field isolates from pepper or from *N. glauca* were assayed in the original and the heterologous host. For pepper, two cultivars with no qualitative resistance to tobamoviruses were assayed, cv. Doux des Landes and cv. Dulce Italiano. The assay followed a random block design with 78 treatments (26 field isolates per three assayed hosts), plus the three mock-inoculated controls,

with ten replicated plants per treatment/control. Plants were harvested at 21 dpi for *N. glauca* and pepper cv. Dulce Italiano and at 10 dpi for pepper cv. Doux des Landes, as the systemic necrosis induced by TMGMV infection in pepper occurs earlier in this cultivar than in Dulce Italiano. Viral multiplication was estimated from viral RNA accumulation in systemically infected leaves (Table 4.1.4) quantified through RT-qPCR.

Table 4.1.4. Accumulation of TMGMV isolates in different assayed hosts[(a)]

Pepper Isolates	Assayed Hosts			*N. glauca* Isolates	Assayed Hosts		
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	*N. glauca*		Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	*N. glauca*
P 00/10	21.9 ± 8.2	4.50 ± 3.89	3.89 ± 2.82	Ng 89/15	317 ± 52	1.63 ± 1.1	1.21 ± 0.8
P 00/6	53 ± 10.4	5.15 ± 2.6	0.79 ± 0.21	Ng 90/8	681 ± 143	9.23 ± 5.4	21.7 ± 3.3
P 04/17	15.9 ± 2.4	1.93 ± 0.76	0.00 ± 0.00	Ng 96/16	48.1 ± 7.5	0.22 ± 0.14	2.77 ± 1.1
P 96/49	155 ± 49	4.25 ± 1.73	0.37 ± 0.16	Ng 99/11	48 ± 5.3	8.89 ± 6.2	1.74 ± 0.61
P 98/11	211 ± 79	2.33 ± 1.50	0.50 ± 0.13	Ng 99/15	49 ± 4.35	0.68 ± 0.45	2.37 ± 0.8
P 98/12	90.7 ± 38	0.34 ± 0.3	0.56 ± 0.16	Ng 99/16	95.9 ± 26.3	0.59 ± 0.47	1.19 ± 0.34
P 01/16	673 ± 99	0.37 ± 0.37	4.44 ± 1.73	Ng 92/73	822 ± 70	12 ± 7.03	23.7 ± 4.8
P 83/4	460 ± 79	0.73 ± 0.37	11.46 ± 3.49	Ng 96/11	388 ± 90	0.01 ± 0.01	26.9 ± 8.8
P 92/10	488 ± 141	1.49 ± 1.01	8.95 ± 2.33	Ng 99/20	247 ± 47	2.48 ± 2.18	37 ± 13.7
P 94/29	348 ± 54	14.24 ± 5.41	14.02 ± 1.73	Ng 89/8	349 ± 63	11.9 ± 6.36	10 ± 3.3
P 97/10	574 ± 83.5	3.27 ± 1.74	16.60 ± 4.07	Ng 90/5	229 ± 45	4.6 ± 4.12	27.8 ± 6.71
P 98/5	777 ± 119	22.72 ± 8.52	12.88 ± 2.69	Ng 94/6	178 ± 50	4.53 ± 4.45	33.3 ± 7.9
				Ng 96/19	404 ± 163	5.76 ± 4.4	81.4 ± 23.6
				Ng 96/5	712 ± 111	12.9 ± 5.3	44.8 ± 10.6
Average	322.53 ± 20.61	4.19 ± 0.74	6.84 ± 0.69	Average	326.23 ± 19.35	4.64 ± 0.95	16.79 ± 1.02

[(a)] Viral accumulation was quantified as RNA accumulation, expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 infected plants.

To determine if virus accumulation depended on the host of origin and/or the assayed host, data were analysed by GzLMs considering host of origin and assayed host as fixed factors, and virus isolate nested to host of origin as a random factor, in a full factorial model. A GzLM considering as assayed hosts *N. glauca* and pepper cv. Doux des Landes, indicated that virus accumulation did not depend on host of origin (Wald $\chi^2_{1,377} = 0.23$, $P = 0.631$) but depended on assayed host, virus isolate nested to host of origin and on the interaction virus isolate nested to host of origin per assayed host (Wald $\chi^2_{1,377} = 488.32$, $P < 0.001$; Wald $\chi^2_{24,377} = 665.63$, $P < 0.001$; Wald $\chi^2_{24,377} = 578.71$, $P < 0.001$, respectively). The accumulation of TMGMV was higher in pepper cv. Doux des Landes than in *N. glauca*. Since virus accumulation depended on virus isolate nested to host of origin and on the interaction virus isolate nested to host of origin per assayed host, pairwise comparisons were done to analyse the performance of the virus isolates in the different hosts. Pairwise comparisons showed that the accumulation of *N. glauca* and pepper isolates was similar in pepper cv. Doux des Landes

(Wald $\chi^2_{1,189}$ = 0.02, *P* = 0.896), but differed in *N. glauca* (Wald $\chi^2_{1,236}$ = 65.18, *P* < 0.001), the accumulation of *N. glauca* isolates in this host being on average 2.45 times higher than that of pepper isolates.

A GzLM considering as assayed hosts *N. glauca* and pepper cv. Dulce Italiano, indicated that virus accumulation depended on host of origin, assayed host, and virus isolate nested to host of origin (Wald $\chi^2_{1,388}$ = 36.43, *P* < 0.001; Wald $\chi^2_{1,388}$ =74.09, *P* < 0.001; Wald $\chi^2_{24,388}$ = 302.83, *P* < 0.001, respectively), and on the interactions host of origin per assayed host and virus isolate nested to host of origin per assayed host (Wald $\chi^2_{1,388}$ = 30.46, *P* < 0.001; Wald $\chi^2_{24,388}$ = 157.83, *P* < 0.001, respectively). Thus, the accumulation of TMGMV was higher in *N. glauca* than in pepper cv. Dulce Italiano. Pairwise comparisons showed that the accumulation of *N. glauca* and pepper isolates was similar in pepper cv. Dulce Italiano (Wald $\chi^2_{1,202}$ = 0.14, *P* = 0.711), but differed in *N. glauca* (Wald $\chi^2_{1,236}$ = 65.18, *P* < 0.001).

Last, a GzLM considering as assayed hosts *N. glauca* and both pepper cultivars combined, indicated that virus accumulation depended on assayed host, virus isolate nested to host of origin and on the interaction virus isolate nested to host of origin per assayed host (Wald $\chi^2_{1,578}$ = 167.29, *P* < 0.001; Wald $\chi^2_{24,578}$ =130.46, *P* < 0.001; Wald $\chi^2_{24,578}$ = 121.28, *P* < 0.001, respectively). The accumulation of TMGMV was higher in pepper than in *N. glauca*; and, again the accumulation of *N. glauca* and pepper isolates did not differ in pepper (Wald $\chi^2_{1,392}$ = 0.29, *P* = 0.584) but differed in *N. glauca* (Wald $\chi^2_{1,236}$ = 65.18, *P* < 0.001).

In summary, irrespective of the isolate's host of origin, pepper cv. Doux des Landes was the best and pepper cv. Dulce Italiano was the worse host (Table 4.1.5). Pepper and *N.*-

Table 4.1.5. Mean Accumulation of TMGMV isolates in different assayed hosts [(a)]

Isolate's host of origin	Assayed hosts			
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cvs combined	*N. glauca*
Pepper	322.53 ± 20.61	4.19 ± 0.74	162 ± 18.5	6.83 ± 0.69
N. glauca	326.23 ± 19.35	4.64 ± 0.95	148 ± 16.9	16.8 ± 1.02
P	0.870	0.711	0.584	< 0.001

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of pepper and *N. glauca* isolates in each assayed host.

glauca isolates accumulated to similar levels in pepper, and *N. glauca* isolates accumulated to higher levels than pepper isolates in *N. glauca*. Thus, results support the hypothesis of *N. glauca* isolates of TMGMV being adapted to their host of origin.

4.1.3 **Analysis of the role of dimorphic loci in the coat protein and 3' untranslated region in the multiplication of TMGMV isolates in pepper and *N. glauca***

The possible role of the three dimorphic loci described in section 4.1.1, position 8 and 148 of CP and the 3' UTR of TMGMV, in the differential accumulation of pepper and *N. glauca* isolates in their original hosts, was analysed, despite that only the allele at position 8 of the CP was significantly associated to the host of origin of the TMGMV isolate. To this aim, data of virus accumulation from the experiment described in section 4.1.2 (Table 4.1.4) were re-analysed according to the alleles at these dimorphic positions.

4.1.3.1 Effect of the dimorphism at position 8 of the coat protein

To determine if the amino acid at position 8 of the CP (CP_8) was associated with the level of virus accumulation in pepper and *N. glauca*, data on virus accumulation in Table 4.1.4 were analysed by GzLMs considering CP_8 and assayed host as fixed factors, and virus isolate nested to CP_8 as a random factor, in a full factorial model.

A GzLM considering as assayed hosts *N. glauca* and pepper cv. Doux des Landes, indicated that virus accumulation didn't depend on CP_8 (Wald $\chi^2_{1,377} = 2.13$, $P = 0.144$) but depended on assayed host, virus isolate nested to CP_8 and on the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,377} = 495.51$, $P < 0.001$; Wald $\chi^2_{24,377} = 692.62$, $P < 0.001$; Wald $\chi^2_{24,377} = 598.22$, $P < 0.001$, respectively). Thus, the accumulation in both pepper cv. Doux des Landes and *N. glauca* of isolates having a proline at CP position 8 was similar than that of isolates having a serine (Wald $\chi^2_{1,189} = 2.01$, $P = 0.156$ for cv. Doux des Landes and Wald $\chi^2_{1,236} = 0.22$, $P = 0.655$ for *N. glauca*). A GzLM considering as assayed hosts *N. glauca* and pepper cv. Dulce Italiano, indicated that virus accumulation depended on CP_8, assayed host, virus isolate nested to CP_8 and the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,388} = 6.67$, $P = 1 \text{ X } 10^{-2}$; Wald $\chi^2_{1,388} = 80.05$, $P < 0.001$; Wald $\chi^2_{24,388} = 311.96$, $P < 0.001$; Wald $\chi^2_{24,388} = 164.13$, $P < 0.001$, respectively).

Accumulation in pepper cv. Dulce Italiano was higher (1.9 times, Wald $\chi^2_{1,178} = 4.16$, $P = 0.041$) for isolates with a proline at CP position 8 than for those with a serine, the amino acid at position 8 of the CP not affecting the accumulation in *N. glauca* (Wald $\chi^2_{1,236} = 0.22$, $P = 0.655$). A GzLM considering as assayed hosts *N. glauca* and both pepper cultivars combined, indicated that virus accumulation didn't depend on CP_8 (Wald $\chi^2_{1,578} = 1.57$, $P = 0.212$) but depended on assayed host, virus isolate nested to CP_8 and the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,578} = 116.44$, $P < 0.001$; Wald $\chi^2_{24,578} = 179.86$, $P < 0.001$; Wald $\chi^2_{24,578} = 127.35$, $P < 0.001$, respectively). Thus, the accumulation in both pepper and *N. glauca* of isolates having a proline at CP position 8 was similar than that of isolates having a serine (Wald $\chi^2_{1,369} = 2.01$, $P = 0.239$ for cv. Doux des Landes and Wald $\chi^2_{1,236} = 0.22$, $P = 0.655$ for *N. glauca*).

Therefore, results show that despite a tendency for higher accumulation in pepper of TMGMV isolates having a proline at position 8 of the CP, accumulation was significantly different only in some pepper cultivars (i.e., Dulce Italiano) (Table 4.1.6).

Table 4.1.6. Accumulation of TMGMV isolates in pepper and *N. glauca* according to the dimorphism at position 8 of coat protein (CP_8) [(a)]

CP_8	Assayed hosts			
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cvs combined	*N. glauca*
Serine	299.44 ± 20.31	2.81 ± 0.91	135.11 ± 20.47	10.86 ± 0.95
Proline	329.76 ± 27.71	5.29 ± 0.81	165.52 ± 1.77	12.90 ± 0.83
P	0.156	0.041	0.239	0.655

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates with a serine or a proline at position 8 of coat protein in each assayed host.

Earlier in section 4.1.1, it was shown that the position 8 of the CP was significantly associated to the host of origin, so to check the effect on viral multiplication of the three factors CP_8, host of origin, and assayed host, we wanted to analyse data on virus accumulation in Table 4.1.4 by a GzLM considering CP_8, host of origin and assayed host as fixed factors, and virus isolate nested to CP_8 as a random factor, in a full factorial model. This model could not be run due to multicollinearity between different factors. So, to overcome this problem, the analysis was done after removing from the data set the five isolates from *N. glauca* with a proline in CP position 8, so that in this new data set of 21 isolates, all isolates of a pepper

origin had a proline and all isolates from *N. glauca* had a serine at position 8 of the CP. This analysis would allow to study the effect on virus multiplication of a proline at position 8 of CP in isolates from pepper in comparison to the effect of a serine at position 8 of CP in isolates from *N. glauca* without introducing "host of origin" as a fixed factor in the GzLM. The GzLMs considered CP_8 and assayed host as fixed factors, and virus isolate nested to CP_8 as a random factor, in a full factorial model. A GzLM considering as assayed hosts *N. glauca* and pepper cv. Doux des Landes, indicated that virus accumulation didn't depend on CP_8 (Wald $\chi^2_{1,315} = 0.43$, $P = 0.511$) but depended on assayed host, virus isolate nested to CP_8 and the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,315} = 435.46$, $P < 0.001$; Wald $\chi^2_{19,315} = 554.25$, $P < 0.001$; Wald $\chi^2_{19,315} = 495.96$, $P < 0.001$, respectively). Thus, in pepper cv. Doux des Landes the accumulation of isolates having a proline at CP position 8 was similar than that of isolates having a serine (Wald $\chi^2_{1,137} = 0.63$, $P = 0.425$) but isolates with a proline accumulated 0.63 times less in *N. glauca* (Wald $\chi^2_{1,180} = 11.73$, $P = 0.001$). A GzLM considering as assayed hosts *N. glauca* and pepper cv. Dulce Italiano, indicated that virus accumulation didn't depend on CP_8 (Wald $\chi^2_{1,322} = 2.52$, $P = 0.113$) but depended on assayed host, virus isolate nested to CP_8, the interaction CP_8 per assayed hosts and the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,322} = 41.23$, $P < 0.001$; Wald $\chi^2_{19,322} = 187.39$, $P < 0.001$; Wald $\chi^2_{1,322} = 10.59$, $P = 1 \text{ X } 10^{-3}$; Wald $\chi^2_{19,322} = 110.02$, $P < 0.001$ respectively). Thus, in pepper cv. Dulce Italiano the accumulation of isolates having a proline at CP position 8 was similar than that of isolates having a serine (Wald $\chi^2_{1,180} = 1.39$, $P = 0.238$) but isolates with a proline accumulated 0.63 times less in *N. glauca* (Wald $\chi^2_{1,180} = 11.73$, $P = 0.001$). A GzLM considering as assayed hosts *N. glauca* and both pepper cultivars combined, indicated that virus accumulation didn't depend on CP_8 (Wald $\chi^2_{1,478} = 0.71$, $P = 0.511$) but depended on assayed host, virus isolate nested to CP_8 and the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,478} = 102.72$, $P < 0.001$; Wald $\chi^2_{19,478} = 128.58$, $P < 0.001$; Wald $\chi^2_{19,478} = 102.28$, $P < 0.001$, respectively). Again, in pepper the accumulation of isolates having a proline at CP position 8 was similar than that of isolates having a serine (Wald $\chi^2_{1,300} = 0.97$, $P = 0.324$) but isolates with a proline accumulated 0.63 times less in *N. glauca* (Wald $\chi^2_{1,180} = 11.73$, $P = 0.001$).

Therefore, results show that excluding the five *N. glauca* isolates having a proline at position 8 of the CP from the analysis, unveils an effect of this proline in the accumulation of TMGMV isolates in *N. glauca* but not in pepper (Table 4.1.7). In *N. glauca*, *N. glauca* isolates with a serine at position 8 of the CP accumulated to higher levels than pepper isolates

with a proline. Again, in pepper, there is a non-significant trend for higher accumulation of isolates from pepper with a proline at position 8 of the CP. Note that when all isolates were considered together there was no difference in *N. glauca* between the accumulation of isolates with a serine or a proline at CP position 8 in Table 4.1.6.

Table 4.1.7. Accumulation of TMGMV isolates in pepper and *N. glauca* according to the dimorphism at position 8 of coat protein after excluding *N. glauca* isolates with a proline at this position[(a)]

CP_8	Assayed hosts			
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cvs combined	*N. glauca*
Serine	299.44 ± 20.31	2.81 ± 0.91	135.11 ± 20.47	10.86 ± 0.94
Proline	322.52 ± 20.61	4.19 ± 0.73	162.32 ± 18.59	6.83 ± 0.68
P	0.425	0.238	0.324	0.001

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates with a serine or a proline at position 8 of coat protein in each assayed host.

Next, to study if a proline or a serine at position 8 of the CP affected the multiplication of *N. glauca* isolates, data on virus accumulation in Table 4.1.4 were analysed in a GzLM after excluding the 12 pepper isolates. The GzLMs considered CP_8 and assayed host as fixed factors, and virus isolate nested to CP_8 as a random factor, in a full factorial model. A GzLM considering as assayed hosts *N. glauca* and pepper cv. Doux des Landes, indicated that virus accumulation depended on CP_8, assayed host, and virus isolate nested to CP_8 (Wald $\chi^2_{1,196}$ = 4.16, P = 0.041; Wald $\chi^2_{1,196}$ = 208.05, P < 0.001; Wald $\chi^2_{12,196}$ = 294.56, P < 0.001, respectively) and the interaction of virus isolate nested to CP_8 per assayed host was also highly significant (Wald $\chi^2_{12,196}$ = 262.64, P < 0.001). Accumulation was higher for isolates with a proline at CP position 8 than for those with a serine, which was due to their higher (2.5 times, Wald $\chi^2_{1,110}$ = 44.73, P < 0.001) accumulation in *N. glauca,* accumulation in pepper cv. Doux des Landes not differing (Wald $\chi^2_{1,87}$ = 2.79, P = 0.094). A GzLM considering as assayed hosts *N. glauca* and pepper cv. Dulce Italiano, indicated that virus accumulation depended on CP_8, assayed host, and virus isolate nested to CP_8 (Wald $\chi^2_{1,207}$ = 41.58, P < 0.001; Wald $\chi^2_{1,207}$ = 68.41, P < 0.001; Wald $\chi^2_{12,207}$ = 92.01, P < 0.001, respectively) and the interaction of CP_8 per assayed host and the interaction of virus isolate nested to CP_8 per assayed host were also highly significant (Wald $\chi^2_{1,207}$ = 11.59, P < 0.001; Wald $\chi^2_{12,207}$ = 67.77, P < 0.001, respectively). Accumulation was higher for isolates with a proline at CP position 8 than for

those with a serine, which was higher in both hosts (2.8 times, Wald $\chi^2_{1,98}$ = 5.06, *P* =0.024 in pepper cv. Dulce Italiano and 2.5 times, Wald $\chi^2_{1,110}$ = 44.73, *P* < 0.001 in *N. glauca*). A GzLM considering as assayed hosts *N. glauca* and both pepper cultivars combined, indicated that virus accumulation didn't depend on CP_8 (Wald $\chi^2_{1,306}$ = 2.27, *P* = 0.132) but depended on assayed host, virus isolate nested to CP_8 and the interaction virus isolate nested to CP_8 per assayed host (Wald $\chi^2_{1,306}$ = 56.71, *P* < 0.001; Wald $\chi^2_{12,306}$ = 48.19, *P* < 0.001; Wald $\chi^2_{12,306}$ = 37.21, *P* < 0.001, respectively). Thus, in pepper the accumulation of isolates having a proline at CP position 8 was similar than that of isolates having a serine (Wald $\chi^2_{1,197}$ = 1.11, *P* = 0.294) but isolates with a proline accumulated 2.5 times higher in *N. glauca* (Wald $\chi^2_{1,110}$ = 44.73, *P* < 0.001).

Therefore, results show that when analyses are restricted to the *N. glauca* isolates, the proline at position 8 of CP was associated with a trend for higher virus accumulation, which was significant in some pepper cultivars (i.e, Dulce Italiano) and in *N. glauca* (Table 4.1.8).

Table 4.1.8. Accumulation of TMGMV isolates from *N. glauca* in pepper and *N. glauca* according to the dimorphism at position 8 of coat protein [(a)]

CP_8	Assayed hosts			
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cvs combined	*N. glauca*
Serine	299.4 ± 20.30	2.81 ± 0.91	135.11 ± 20.42	10.86 ± 0.95
Proline	374.43 ± 39.90	7.9 ± 2.10	173.00 ± 29.81	27.46 ± 2.30
P	0.094	0.024	0.294	< 0.001

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates with a serine or a proline at position 8 of coat protein in each assayed host.

In summary, a proline at position 8 of the CP is associated with a trend for higher virus accumulation, but the effect of the amino acid at position 8 of the CP on virus accumulation varies according to the isolate's host of origin and to the assayed host.

4.1.3.2 Effect of the dimorphism at position 148 of coat protein

To determine if the amino acid at position 148 of CP (CP_148) was associated with the level of virus accumulation in pepper and *N. glauca*, data on virus accumulation in Table 4.1.4 were

analysed by GzLMs considering CP_148 and assayed host as fixed factors, and virus isolate nested to CP_148 as a random factor, in a full factorial model.

A GzLM considering as assayed hosts *N. glauca* and pepper cv. Doux des Landes, indicated that virus accumulation depended on CP_148, assayed host, and virus isolate nested to CP_148 (Wald $\chi^2_{1,377}$ = 158.36, P < 0.001; Wald $\chi^2_{1,377}$ = 486.46, P < 0.001; Wald $\chi^2_{24,377}$ = 319.12, P < 0.001, respectively) and the interactions of CP_148 per assayed host and of virus isolate nested to CP_148 per assayed host were also highly significant (Wald $\chi^2_{1,377}$ = 128.67, P < 0.001; Wald $\chi^2_{24,377}$ = 300.68, P < 0.001, respectively). Accumulation was higher for isolates with a threonine at CP position 148 than for those with an alanine, which was due to their higher accumulation in both *N. glauca* (6.5 times, Wald $\chi^2_{1,212}$ = 200.26, P < 0.001) and in pepper cv. Doux des Landes (3.2 times, Wald $\chi^2_{1,165}$ = 143.41, P < 0.001). A GzLM considering as assayed hosts *N. glauca* and pepper cv. Dulce Italiano, indicated that virus accumulation depended on CP_148, assayed host, and virus isolate nested to CP_148 (Wald $\chi^2_{1,388}$ = 138.12, P < 0.001; Wald $\chi^2_{1,388}$ = 73.71, P < 0.001; Wald $\chi^2_{24,388}$ = 158.29, P < 0.001, respectively), the interactions of CP_148 per assayed host and of virus isolate nested to CP_148 per assayed host were also significant (Wald $\chi^2_{1,388}$ = 68.99, P < 0.001; Wald $\chi^2_{24,388}$ = 91.78, P < 0.001, respectively). Accumulation was higher for isolates with a threonine at CP position 148 than for those with an alanine, which was due to their higher accumulation in both *N. glauca* (6.5 times, Wald $\chi^2_{1,212}$ = 200.26, P < 0.001) and in pepper cv. Dulce Italiano (2 times, Wald $\chi^2_{1,176}$ = 5.96, P = 0.015). A GzLM considering as assayed hosts *N. glauca* and both pepper cultivars combined, indicated that virus accumulation depended on CP_148, assayed host, and virus isolate nested to CP_148 (Wald $\chi^2_{1,578}$ = 49.76, P < 0.001; Wald $\chi^2_{1,578}$ =131.88, P < 0.001; Wald $\chi^2_{24,578}$ = 79.63, P < 0.001, respectively), and the interaction of CP_148 per assayed host and the interaction virus isolate nested to CP_148 per assayed host were also significant (Wald $\chi^2_{1,578}$ = 31.83, P < 0.001; Wald $\chi^2_{1,578}$ = 74.59, P < 0.001, respectively). Again, accumulation was higher for isolates with a threonine at CP position 148 than for those with an alanine, which was due to their higher accumulation in both *N. glauca* (6.5 times, Wald $\chi^2_{1,212}$ = 200.26, P < 0.001) and in pepper cv. Dulce Italiano (3.1 times, Wald $\chi^2_{1,368}$ = 40.34, P = 0.015).

Therefore, results show that a threonine at position 148 of the CP was associated with higher virus accumulation in all assayed hosts, the effect being higher in *N. glauca* than in pepper (Table 4.1.9).

Table 4.1.9. Accumulation of TMGMV isolates in pepper and *N. glauca* according to the dimorphism at position 148 of coat protein (CP_148) [(a)]

CP_148	Assayed hosts			
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cvs combined	*N. glauca*
Threonine	475.00 ± 23.12	5.77 ± 0.99	225.17 ± 21.23	20.00 ± 1.13
Alanine	148.95 ± 14.3	2.87 ± 0.65	73.12 ± 11.03	3.08 ± 0.37
P	< 0.001	< 0.001	< 0.001	< 0.001

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates with a threonine or an alanine at position 8 of coat protein in each assayed host.

Next, to analyse if the effect on virus accumulation of the amino acid at position 148 of the CP depended on the isolate's host of origin, pairwise comparisons between the accumulation of isolates with a threonine or an alanine at position 148 of CP were done for each assayed host (Table 4.1.10) which showed that irrespective of the host of origin, a threonine at position 148 of CP was associated with higher virus accumulation (all P < 0.001) in all hosts except pepper cv. Dulce Italiano.

Table 4.1.10. Accumulation of TMGMV isolates in pepper and *N. glauca* according to the dimorphism at position 148 of coat protein and the host of origin [(a)]

Isolate's host of origin	CP_148	Assayed hosts			
		Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cvs combined	*N.glauca*
Pepper	Threonine	553.72 ± 38.1	5.77 ± 0.99	277.23 ± 35.8	12.62 ± 1.3
	Alanine	91.43 ± 15.8	2.87 ± 0.65	47.24 ± 9.96	1.01 ± 0.44
P		< 0.001	0.244	< 0.001	< 0.001
N. glauca	Threonine	416.00 ± 28.7	6.13 ± 1.47	185.82 ± 25.6	25.52 ± 1.73
	Alanine	206.52 ± 23.9	2.65 ± 1.02	99.06 ± 19.7	5.16 ± 0.58
P		< 0.001	0.068	< 0.001	< 0.001

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates with a threonine or an alanine at position 8 of coat protein in each assayed host.

In summary, the amino-acid at position 148 of CP does affect viral accumulation, which was higher in all assayed hosts when a threonine was present at this position of CP, irrespective of the isolates host of origin.

4.1.3.3 Effect of the dimorphism at 3' untranslated region

To determine if the 3' UTR was associated with the level of virus accumulation in pepper and *N. glauca*, data on virus accumulation in Table 4.1.4 were analysed by GzLM considering 3' UTR and assayed host as fixed factors, and virus isolate nested to 3' UTR as a random factor, in a full factorial model.

A GzLM considering as assayed hosts *N. glauca* and pepper cv. Doux des Landes, indicated that virus accumulation depended on 3' UTR, assayed host, and virus isolate nested to 3' UTR (Wald $\chi^2_{1,377}$ = 6.88, P = 0.009; Wald $\chi^2_{1,377}$ = 511.22, P < 0.001; Wald $\chi^2_{24,377}$ = 699.45, P < 0.001, respectively) and on the interactions of 3' UTR per assayed host and of virus isolate nested to 3' UTR per assayed host (Wald $\chi^2_{1,377}$ = 7.71, P = 0.009; Wald $\chi^2_{24,377}$ = 602.02, P < 0.001, respectively). Accumulation was higher for isolates having a short 3' UTR end, which was due to their higher (1.4 times, Wald $\chi^2_{1,166}$ = 7.31, P = 0.007) accumulation in pepper cv. Doux des Landes, accumulation in *N. glauca* not being different for isolates with a short or a long 3'UTR (Wald $\chi^2_{1,212}$ = 2.44, P = 0.118). A GzLM considering as assayed hosts *N. glauca* and pepper cv. Dulce Italiano, indicated that virus accumulation did not depend on 3' UTR (Wald $\chi^2_{1,388}$ = 0.409, P = 0.522) but depended on assayed host, virus isolate nested to 3' UTR and the interaction virus isolate nested to 3' UTR per assayed host (Wald $\chi^2_{1,388}$ = 80.47, P < 0.001; Wald $\chi^2_{24,388}$ = 303.56, P < 0.001; Wald $\chi^2_{24,388}$ = 171.88, P < 0.001, respectively). Thus, the accumulation in both *N. glauca* and pepper cv. Dulce Italiano of isolates containing a short 3' UTR end was similar than that of isolates containing a long (Wald $\chi^2_{1,212}$ = 2.44, P = 0.118 for *N. glauca* and Wald $\chi^2_{1,178}$ = 0.51, P = 0.475 for cv. Doux des Landes). A GzLM considering as assayed hosts *N. glauca* and both pepper cultivars combined, indicated that virus accumulation didn't depend on 3' UTR (Wald $\chi^2_{1,388}$ = 5.58, P = 0.059) but depended on assayed host, and virus isolate nested to 3' UTR (Wald $\chi^2_{1,578}$ = 121.45, P < 0.001; Wald $\chi^2_{24,578}$ =175.98, P < 0.001, respectively); and the interactions of 3' UTR per assayed host and of virus isolate nested to 3' UTR per assayed host were also highly significant (Wald $\chi^2_{1,578}$ = 4.22, P = 0.04; Wald $\chi^2_{24,578}$ = 123.54, P < 0.001, respectively). Thus, accumulation was higher for isolates containing a short 3' UTR end than for those with a long, which was due to their higher (1.45 times, Wald $\chi^2_{1,369}$ = 3.91, P = 0.048) accumulation in pepper, accumulation in *N. glauca* not being different (Wald $\chi^2_{1,212}$ = 2.44, P = 0.118).

Therefore, results show that a short 3' UTR end was associated with a trend for higher virus accumulation in pepper, which depended on the cultivar, but the length of the 3' UTR did not affect virus accumulation in *N. glauca* (Table 4.1.11).

Table 4.1.11. Accumulation of TMGMV isolates in pepper and *N. glauca* according to the dimorphism at 3' untranslated region [(a)]

3' UTR	Assayed hosts			
	Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper *cv.*s combined	*N. glauca*
Short	347.36 ± 29.07	4.66 ± 0.881	171.12 ± 17.86	10.21 ± 1.09
Long	240.99 ± 32.89	3.87 ± 1.04	117.63 ±18.52	12.79 ± 1.64
P	0.007	0.475	0.048	0.118

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates having a short or a long 3' untranslated region in each assayed host

Next, to analyse if the effect of the dimorphism in the length of the 3'UTR on virus accumulation depended on the isolate's host of origin, pairwise comparisons between short and long 3' UTR end were done for each assayed host (Table 4.1.12). The results showed that pepper isolates with a short 3' UTR and *N. glauca* isolates with a long 3' UTR showed a trend

Table 4.1.12. Accumulation of TMGMV isolates in pepper and *N. glauca* according to the dimorphism at 3' UTR end and the host of origin [(a)]

Isolate's host of origin	3' UTR	Assayed hosts			
		Pepper cv. Doux des Landes	Pepper cv. Dulce Italiano	Pepper cv.s combined	*N. glauca*
Pepper	Short	390.52 ± 25.56	4.72 ± 0.93	196.33 ± 24.12	8.91 ± 0.92
	Long	118.53 ± 27.47	2.6 ± 0.94	60.56 ± 17.15	0.61 ± 0.09
P		< 0.001	0.195	< 0.001	< 0.001
N. glauca	Short	240.51 ± 29.13	4.11 ± 1.41	143.99 ± 23.88	12.31 ± 2.04
	Long	390.48 ± 25.87	5.03 ± 1.32	144.27 ± 25.12	18.52 ± 2.05
P		< 0.039	0.630	< 0.001	0.004

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 plants. *P* values refer to the comparison of isolates nested to virus containing a short and long 3' untranslated region end in each assayed host.

for higher accumulation in all hosts, accumulation being significantly different ($P < 0.001$ for pepper and $P \leq 0.039$ for *N. glauca* host of origin) except in pepper cv. Dulce Italiano ($P = 0.195$ for pepper and $P = 0.630$ for *N. glauca* host of origin).

In summary, the length dimorphism at the 3' UTR does affect virus accumulation differentially according to host of origin and to assayed hosts.

4.1.4 **Variation of the 3' untranslated region length during virus multiplication in *N. glauca***

Prior to the experiments reported in section 4.1.2, all TMGMV field isolates were multiplied in their respective host of origin (Pepper isolates in pepper cv. Dulce Italiano and *N. glauca* isolates in *N. glauca*, Material and Methods section 3.1.1). The 3' UTR length of each isolate was checked by RT-PCR on the encapsidated RNA using primers TMGMVRNA3 and CGM3 (Table 3.7) before and after multiplication in the original host. The length of the 3' UTR did not change for most isolates upon multiplication in their original host, except for two of them, Ng 94/6 and Ng 90/8, (Figure 4.2. lanes 2 and 4). These isolates had changed from a short 3' UTR to a mixed (long and short) 3' UTR after one passage of multiplication in *N. glauca*. Next, to quantify the frequency of long and short 3'UTR in the passaged populations of these two isolates, twenty biological clones for each isolate were generated from random local-lesions in

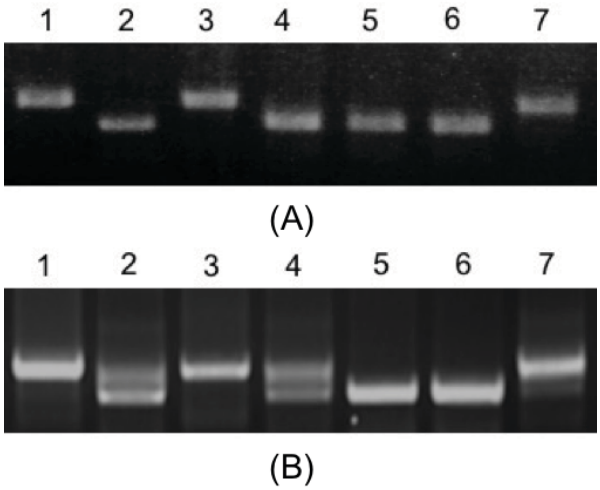


Figure 4.2 Length polymorphisms at the 3' untranslated region in TMGMV isolates from *N. glauca* before (A) and after (B) multiplication in *N. glauca*. Each lane represents the same isolate in panels A and B. Primers TMGMVRNA3 and CGM3 were used for RT-PCR on the encapsidated RNA, and the resulting amplicons were of 780 and/or, 900 base pairs. Data shown are for isolates Ng 96/5, Ng 94/6, Ng 92/73, Ng 90/8, Ng 89/15, Ng 90/5, Ng 96/11 in lanes 1, 2, 3, 4, 5, 6 and, 7, respectively.

N. tabacum cv. Xanthi-nc, which were multiplied in *N. clevelandii* (Figure 4.3). RT-PCR of encapsidated viral RNA showed that for Ng 94/6 5 clones had a short and 15 clones had long 3' UTR end, and for Ng 90/8, 14 clones had a short and 6 clones had a long 3' UTR end.

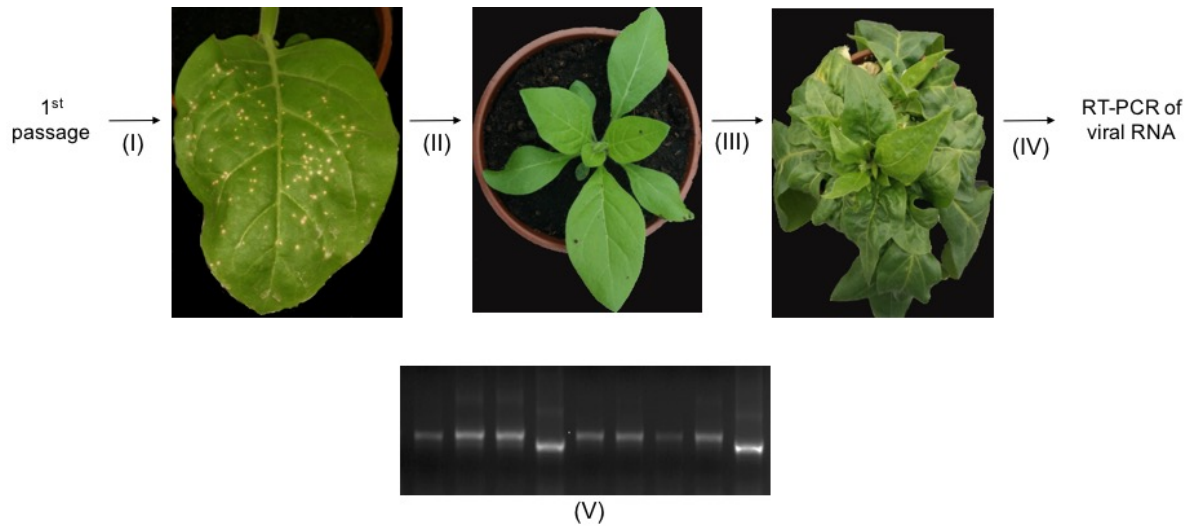


Figure 4.3 Flow diagram showing the steps to estimate the frequency of short and long 3' untranslated region (UTR) in TMGMV isolates Ng 94/6 and Ng 90/8 after multiplication in *N. glauca*. The multiplied virus was inoculated in *N. tabacum* cv. Xanthi-nc to obtain biological clones (I), 20 nll (clones) were transferred to *N. clevelandii* for multiplication (II), At 30 dpi, viral particles were purified from systemically-infected leaves, RNA was extracted from particles (III) & (IV), and the length of the 3'UTR was determined by agarose gel electrophoresis of RT-PCR products.

To further study if host had an effect on 3' UTR length, isolates Ng 94/6 and Ng 90/8 were passaged for a second time in *N. glauca* (Figure 4.4). After this second passage, Ng 94/6, which had a short 3' UTR initially, showed a higher proportion of long 3' UTR end, and Ng 90/8, which had also a short 3' UTR initially, maintained a higher amount of short 3' UTR end, after a transient change for both isolates after first passage.

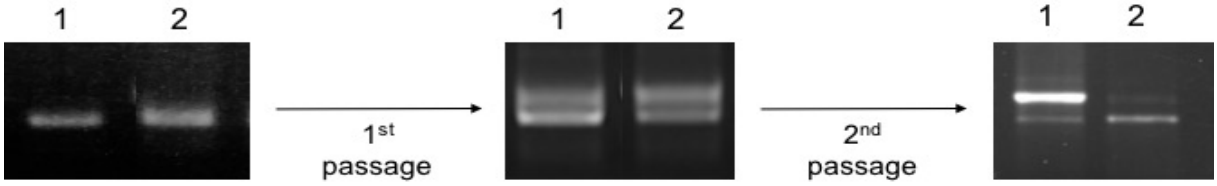


Figure 4.4 Length polymorphisms at the 3' untranslated regions (UTR) of TMGMV isolates Ng 94/6 and Ng 90/8. Agarose gel electrophoresis of 3'UTR RT-PCR 780 and/or 900 base pairs amplicons in the initial virus stock and after passaging in *N. glauca*. Lanes 1 show isolate Ng 94/6 and lanes 2 isolate Ng 90/8.

So, virus multiplication in *N. glauca* had an effect on the length of the 3'UTR that was dependent of the virus isolate.

4.1.5 **Accumulation in pepper or *N. glauca* of TMGMV genotypes derived from biologically active cDNA clones**

As indicated in Material and Methods (see section 3.1.1), pTMGMV-Jap was provided by Dr. Tetsuro Okuno and Dr. Masanori Kaido, Kyoto University, Japan and pTMGMV-FA was constructed by me from field isolate Ng 92/73 as part of this thesis work (see section 3.2). In these plasmids, the dimorphisms at positions 148 of the CP and at the 3'UTR, were introduced to analyse their possible effect on virus multiplication in a known genetic background.

4.1.5.1 Molecular characterization of TMGMV genotypes derived from biologically active cDNA clones

The complete nucleic acid sequence of the TMGMV insert present in the infectious cDNA clone pTMGMV-FA was determined and found to consist of 6,501 nt (Annex 2). Comparison with the corresponding sequence in pTMGMV-Jap showed that the sequences differed at 339 positions: two in the 5' UTR, 114 in the gene encoding the 126-kDa protein open reading frame 1, (ORF 1), 36 in that encoding the read-through portion of the 183-kDa protein (ORF 2), 20 in the movement protein gene (ORF 3), 15 in the CP gene (ORF 4), and 152 in the 3' UTR which includes a 147 nt duplication in TMGMV-FA. Nucleotide substitutions within ORFs resulted in 40 amino-acids substitutions: 28 in the 126 kDa protein, five in the 183 kDa protein read-through portion, four in movement protein and three in the CP. Two out of three substitutions (position 8 and 148) in the CP (Table 4.1.13) were the polymorphisms found in field isolates and analysed in sections 4.1.1 and 4.1.2. The third substitution, a methionine/ isoleucine polymorphism at position 59 of the CP was not considered in further analyses, since all field isolates except P 96/49 had a methionine at this position.

Table 4.1.13. Comparison of the coat protein's amino-acid sequence of TMGMV-FA with TMGMV-Jap

Genotype	Amino-acid		
	8	59	148
TMGMV-FA	Serine	Methionine	Threonine
TMGMV-Jap	Proline	Isoleucine	Alanine

4.1.5.2 The effect of CP and 3'UTR polymorphisms on viral accumulation in pepper and N. glauca of TMGMV genotypes derived from biologically active cDNA clones

Mutations were introduced in plasmids pTMGMV-Jap and pTMGMV-FA as described in Material and Methods (see section 3.1.2). Briefly, in pTMGMV-Jap, the alanine at position 148 of CP was substituted by a threonine (TMGMV–Jap-CP) and the short 3' UTR was substituted by a long 3' UTR (TMGMV–Jap-UTR); in pTMGMV-FA, the threonine at position 148 of CP was substituted by an alanine (TMGMV–FA-CP) and the long 3' UTR was substituted by a short 3' UTR (TMGMV–FA-UTR), generating the set of mutants in Table 4.1.14. No mutations

Table 4.1.14. Description of TMGMV genotypes derived from cDNA clones

Genotype[(a)]	aa[(b)]	3' UTR
TMGMV - Jap	Ala	Short
TMGMV - Jap-CP	Thr	Short
TMGMV - Jap-UTR	Ala	Long
TMGMV - FA	Thr	Long
TMGMV - FA-CP	Ala	Long
TMGMV - FA-UTR	Thr	Short

[(a)] Mutant genotypes are identified by a letter at the end of parental genotype's name indicating the site of mutation (CP for Coat Protein and UTR for 3' Untranslated Region).
[(b)] Amino-acid at position148 of coat protein.

were introduced to construct the proline/serine dimorphism at position 8 of CP, as the results of section 4.1.3 did not show a clear effect of the amino acid at this position on virus accumulation in the different hosts.

To study the effect of amino acid substitutions at position 148 of CP and the duplication in the 3' UTR on viral multiplication in pepper and *N. glauca*, TMGMV-Jap, TMGMV-FA and its derived mutants were assayed in pepper cv. Doux des Landes and *N. glauca*. Initially, genotype TMGMV–FA-CP could not be obtained so, a first assay was performed with genotypes TMGMV-Jap, TMGMV-Jap-UTR, TMGMV-Jap-CP, TMGMV-FA and TMGMV-FA-UTR following a random block design with 10 treatments (5 viral genotype x 2 assayed hosts), plus the two mock-inoculated controls, with ten replicated plants per treatment/control. Plants were harvested at 10 dpi for pepper cv. Doux des Landes and at 21 dpi for *N. glauca*. Viral

multiplication was estimated from viral RNA accumulation in systemically infected leaves (Table 4.1.15) by RT-qPCR.

Table 4.1.15. Accumulation of TMGMV of different genotypes in different hosts[(a)]

Genotype[(b)]	Characteristics		Assayed Host	
	CP_148	3' UTR	Pepper	*N. glauca*
TMGMV-Jap	A	Short	5.62 ± 1.32	34.92 ± 4.35
TMGMV-Jap-CP	T	Short	58.56 ± 20.06	0.008 ± 0.003
TMGMV-Jap-UTR	A	Long	10.47 ± 2.12	0.004 ± 0.001
TMGMV-FA	T	Long	113.02 ± 24.21	53.44 ± 8.01
TMGMV-FA-UTR	T	Short	15.83 ± 3.82	0.002 ± 0.001

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 9 plants.
[(b)] Mutant genotypes are identified by letters at the end of parental genotype's name indicating the site of mutation (CP for coat protein and UTR for 3' untranslated region).

To determine if the virus genotype was associated with the level of virus accumulation in assayed hosts, data on virus accumulation in Table 4.1.15 were analysed by a GzLM considering virus genotype and assayed host as fixed factors, in a full factorial model. Virus accumulation depended on virus genotype, assayed host and on the interaction virus genotype per assayed host (Wald $\chi^2_{1,87}$ = 13.60, P < 0.001; Wald $\chi^2_{4,87}$ = 84.78, P < 0.001; Wald $\chi^2_{4,87}$ = 94.01, P < 0.001, respectively). Pairwise comparisons of TMGMV–Jap with each of its derived mutants showed that both the substitution of a threonine for an alanine at position 148 of the CP, and the introduction of a duplication in the 3'UTR, resulted in higher accumulation in pepper (Wald $\chi^2_{1,19}$ = 4.48, P < 0.035 for TMGMV-Jap-UTR and Wald $\chi^2_{1,18}$ = 7.81, P < 0.005 for TMGMV-Jap-CP) but a lower accumulation in *N. glauca* (Wald $\chi^2_{1,19}$ = 73.02, P < 0.001 for TMGMV-Jap-UTR and Wald $\chi^2_{1,18}$ = 72.91, P < 0.001 for TMGMV-Jap-CP). The pairwise comparison of TMGMV–FA with its derived mutant TMGMV–FA-UTR showed that the long 3'UTR was associated with higher virus accumulation in both hosts (Wald $\chi^2_{1,18}$ = 17.61, P < 0.001 in pepper and Wald $\chi^2_{1,19}$ = 48.96, P < 0.001 in *N. glauca*).

Thus, the amino acid at CP position 148 and the 3'UTR length, had a host-differential effect on virus multiplication in the TMGMV–Jap genetic background, while the 3'UTR length had a host-independent effect on virus multiplication in the TMGMV–FA genetic background.

Later, after some difficulty, genotype TMGMV–FA-CP was obtained and another assay was performed to analyse the effect of the threonine/alanine polymorphism in the TMGMV-FA

context. The experiment included genotypes TMGMV-FA, TMGMV-FA-CP and TMGMV-Jap-UTR as a control, and followed a random block design with 6 treatments (3 viral genotype x 2 assayed hosts), plus the two mock-inoculated controls, with 8 replicated plants per treatment/control. Plants were harvested at 10 dpi for pepper cv. Doux des Landes and at 21 dpi for *N. glauca*. Viral multiplication was estimated from viral RNA accumulation in systemically infected leaves (Tables 4.1.16) by RT-qPCR.

Table 4.1.16. Accumulation of TMGMV of different genotypes in different hosts [(a)]

Genotype[(b)]	Characteristics		Assayed Host	
	CP_148	3' UTR	Pepper	*N. glauca*
TMGMV-FA	T	Long	396.00 ± 78.36	10.94 ± 5.02
TMGMV-FA-CP	A	Long	48.62 ± 10.11	2.65 ± 1.17
TMGMV-Jap-UTR	A	Long	23.13 ± 3.54	0.009 ± 0.002

[(a)] Viral RNA accumulation expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 6 plants.
[(b)] Mutant genotypes are identified by a letter at the end of parental genotype's name indicating the site of mutation (CP for coat protein and UTR for 3' untranslated region).

To determine if the virus genotype was associated with the level of virus accumulation in assayed hosts, data on virus accumulation in Table 4.1.16 were analysed by a GzLM considering virus genotype and assayed host as fixed factors, in a full factorial model. Virus accumulation depended on virus genotype, assayed host and on the interaction virus genotype per assayed host (Wald $\chi^2_{1,35} = 39.49$, $P < 0.001$; Wald $\chi^2_{2,35} = 36.38$, $P < 0.001$; Wald $\chi^2_{2,35} = 30.62$, $P < 0.001$, respectively). Accumulation of TMGMV was higher in pepper cv. Doux des Landes than in *N. glauca*. Pairwise comparisons showed that a threonine at position 148 of the CP in the TMGMV-FA genetic background was associated with higher accumulation in both hosts, differences being significant in pepper (Wald $\chi^2_{1,12} = 23.25$, $P < 0.001$) and marginally significant in *N. glauca* (Wald $\chi^2_{1,13} = 3.09$, $P < 0.079$).

In summary, the results indicated that the amino acid at position 148 of CP and the duplication in the 3' UTR did affect the multiplication of virus. These effects were host-dependent in the TMGMV-Jap genetic background, and host independent in the TMGMV-FA genetic background.

4.2 Pleiotropic effects of resistance- breaking mutations on different fitness components of Pepper mild mottle virus

Under the GFG model, resistance-breaking mutations may result in resistance-breaking costs, so that the resistance-breaking genotypes would be less fit than the non-resistance-breaking ones in susceptible hosts. Resistance-breaking costs are often derived from the within-host multiplication component of the virus fitness. In order to fully understand the evolution of resistance breaking, it is necessary to consider potential costs on other fitness components, as the evolution of viruses may be constrained by conflicting trade-offs between different fitness components. According to life history evolution theory one such trade-off would be between within-host multiplication and survival outside the host, which is related to the stability of the virus particle. In order to test if resistance-breaking mutations are associated with penalties on different fitness components, and if there are trade-offs between those components, a set of PMMoV mutants were constructed (see Material and Methods section 3.1.2) which in addition to four more genotypes from Moreno-Pérez et al., (2016) (see section 3.1.1) were assayed for multiplication in susceptible cultivars of *C. annuum* and for the *in vitro* stability of virus particles. The set of PMMoV genotypes includes all reported single- and double-mutants that result in overcoming resistance alleles at the *L* locus of *Capsicum*, plus the single mutants that are in the evolutionary pathway towards the resistance-breaking double mutants.

4.2.1 Characterization of different mutant genotypes of PMMoV

To determine the pathotype of all the CP mutants derived from the parental genotype PMMoV-MG-WT, they were inoculated in a set of indicator plants consisting of *C. annuum* cv. Dulce Italiano (L^+/L^+), *C. annuum* cv. Yolo Wonder (L^1/L^1), *C. frutescens* cv. Tabasco (L^2/L^2), *C. chinense* PI 159236 (L^3/L^3) and *C. chacoense* PI260429 (L^4/L^4). For convenience, the different indicator hosts will be identified by their genotype at the *L* locus rather than by the full species or cultivar name. Note that the pathotype of PMMoV–MG-WT, PMMoV–MG-(T43K+D50G), PMMoV–MG-(A86G) and PMMoV–MG-(M138N) was already determined in Moreno-Pérez et al., (2016). Results (Table 4.2.1) showed that the single mutants PMMoV-MG-(T43K), PMMoV-MG-(D50G), PMMoV-MG-(L13F), PMMoV-MG-(G66V), PMMoV-MG-(Q46L) and PMMoV-MG-(Q46R) had a $P_{1,2}$ pathotype; the double mutants PMMoV-MG-(L13F+G66V) and PMMoV-MG-(Q46L+M138N) had a $P_{1,2,3}$ pathotype, and PMMoV-MG-(G85K) and PMMoV-MG-(Q46R+G85K) had a $P_{1,2,3,4}$ pathotype.

Table 4.2.1 Capacity of different pepper mild mottle virus genotypes to infect the indicated host [a]

Genotype	L^+/L^+	L^1/L^1	L^2/L^2	L^3/L^3	L^4/L^4	Pathotype
PMMoV-MG-WT	+	+	+	nll	nll	$P_{1,2}$ [b]
PMMoV-MG-(T43K)	+	+	+	nll	nll	$P_{1,2}$
PMMoV-MG-(D50G)	+	+	+	nll	nll	$P_{1,2}$
PMMoV-MG-(T43K+D50G) [b]	+	+	+	+	nll	$P_{1,2,3}$ [b]
PMMoV-MG-(L13F)	+	+	+	nll	nll	$P_{1,2}$
PMMoV-MG-(G66V)	+	+	+	nll	nll	$P_{1,2}$
PMMoV-MG-(L13F+G66V)	+	+	+	+	nll	$P_{1,2,3}$
PMMoV-MG-(Q46L)	+	+	+	nll	nll	$P_{1,2}$
PMMoV-MG-(M138N) [b]	+	+	+	+	nll	$P_{1,2,3}$ [b]
PMMoV-MG-(Q46L+M138N)	+	+	+	+	nll	$P_{1,2,3}$
PMMoV-MG-(Q46R)	+	+	+	nll	nll	$P_{1,2}$
PMMoV-MG-(G85K)	+	+	+	+	+	$P_{1,2,3,4}$
PMMoV-MG-(Q46R+G85K)	+	+	+	+	+	$P_{1,2,3,4}$
PMMoV-MG-(A86G) [b]	+	+	+	+	+	$P_{1,2,3,4}$ [b]

[a] + = systemic infection; nll=necrotic local lesions.
[b] Described in Moreno-Pérez *et al*., 2016.

4.2.2 **Pleiotropic effects of resistance-breaking mutations on the multiplication of PMMoV genotypes in susceptible hosts**

To analyse whether CP mutations (13; ten from my work and three from Moreno-Pérez et al., 2016) in the PMMoV single and double mutants derived from PMMoV–MG-WT were associated with within-host fitness penalties in susceptible hosts, assays were done in host genotypes L^+/L^+, L^1/L^1 and, L^2/L^2. The experiment followed a random block design with 42 treatments (14 virus genotypes per 3 host genotypes), plus three buffer inoculated controls, with eight replicated plants per treatment/control. Plants were harvested at 21 dpi and virus RNA was quantified in systemically infected leaves (Table 4.2.2). Multiplication of the various virus genotypes was compared by means of their within-host fitness, and fitness was estimated as in Materials and Methods (section 3.5.1). In this experiment, due to an error in the processing of the samples, within-host fitness could not be estimated for genotypes PMMoV-MG-(A86G) and PMMoV-MG-(M138N), and so an additional experiment was done in which these two mutant genotypes and PMMoV-MG-WT as a control were assayed in host genotypes L^+/L^+, L^1/L^1 and, L^2/L^2. The experiment followed a similar design as the earlier one with 9 treatments (3 virus genotypes per 3 host genotypes), plus three buffer inoculated controls, with eight replicated plants per treatment/control. For convenience, in the following subsections the PMMoV genotypes will be identified by their mutations (A86G, T43K, D50G,

T43K + D50G, L13F, G66V, L13F + G66V, Q46L, M138N, Q46L + M138N, Q46R, G85K and Q46R + G85K) or as WT.

Table 4.2.2 Accumulation of different PMMoV genotypes in three susceptible hosts [(a)]

Virus genotype	Host genotype		
	L^+/L^+	L^1/L^1	L^2/L^2
WT	61.11 ± 20.00	76.91 ± 16.35	8.94 ± 5.52
M138N	46.76 ± 6.19	49.53 ± 13.94	16.46 ± 2.07
A86G	22.49 ± 5.59	11.57 ± 3.76	18.80 ± 7.03
WT	1.97 ± 0.57	4.77 ± 1.67	12.99 ± 4.43
L13F + G66V	1.35 ± 0.23	0.33 ± 0.13	21.71 ± 5.14
L13F	2.77 ± 0.95	22.12 ± 4.99	12.87 ± 2.81
G66V	0.53 ± 0.17	0.77 ± 0.37	17.86 ± 2.34
T43K + D50G	1.71 ± 0.52	0.41 ± 0.21	4.09 ± 1.62
T43K	0.53 ± 0.16	2.13 ± 0.96	3.94 ± 1.61
D50G	3.23 ± 0.83	4.12 ± 2.03	34.65 ± 13.15
Q46L + M138N	4.15 ± 0.98	0.86 ± 0.38	29.43 ± 5.39
Q46L	3.77 ± 0.77	4.84 ± 2.54	0.96 ± 0.16
Q46R + G85K	3.14 ± 0.72	13.31 ± 3.66	12.19 ± 3.64
Q46R	2.41 ± 0.76	6.83 ± 3.66	11.88 ± 7.13
G85K	1.22 ± 0.19	2.82 ± 1.05	4.16 ± 0.87

[(a)] Viral accumulation was quantified as RNA accumulation, expressed in ng of viral RNA / µg of total RNA. Data are mean ± standard error for at least 8 infected plants.

Table 4.2.2 shows the accumulation of the different PMMoV genotypes in the three assayed hosts, from which the within-host absolute fitness was computed (Table 4.2.3). To perform statistical analyses, since fitness estimates derived from two experiments, fitness were normalised to the fitness of the WT, taken as 1.0.

To determine if fitness depended on the pathotype, normalised data was analysed by a GzLM considering virus pathotype and host genotype as fixed factors and virus genotype nested to pathotype as a random factor in a full factorial model. Virus fitness did not depend on virus pathotype (Wald $\chi^2_{2,258} = 3.97$, $P = 0.138$) but, depended on host genotype and virus genotype nested to pathotype (Wald $\chi^2_{2,258} = 19.34$, $P < 0.001$; Wald $\chi^2_{10,258} = 127.43$, $P = 0.001$, respectively), and on the interactions of virus pathotype per host genotype and on virus genotype nested to pathotype per host genotype (Wald $\chi^2_{4,258} = 41.81$, $P < 0.001$; Wald $\chi^2_{20,258} = 161.59$, $P < 0.001$, respectively). In the L^+/L^+ host genotype, virus fitness was similar for the three pathotypes (Wald $\chi^2 < 0.91$, $P > 0.286$). In L^1/L^1 host genotype, pathotype $P_{1,2,3}$ was sig-

Table 4.2.3 Fitness of different PMMoV mutants in three susceptible hosts [(a)]

Virus genotype	Host genotype		
	L^+/L^+	L^1/L^1	L^2/L^2
WT	1.21 ± 0.02	1.22 ± 0.02	1.07 ± 0.03
M138N	1.20 ± 0.01	1.19 ± 0.02	1.14 ± 0.01
A86G	1.14 ± 0.02	1.07 ± 0.04	1.12 ± 0.02
WT	1.02 ± 0.01	1.05 ± 0.03	1.15 ± 0.03
T43K	0.92 ± 0.02	0.94 ± 0.05	1.04 ± 0.03
D50G	1.06 ± 0.02	1.00 ± 0.04	1.22 ± 0.03
T43K + D50G	0.98 ± 0.03	0.88 ± 0.03	1.05 ± 0.02
L13F	1.04 ± 0.02	1.20 ± 0.02	1.16 ± 0.02
G66V	0.92 ± 0.02	0.88 ± 0.04	1.20 ± 0.01
L13F + G66V	1.01 ± 0.01	0.90 ± 0.02	1.20 ± 0.02
Q46L	1.06 ± 0.03	1.04 ± 0.04	0.98 ± 0.02
Q46L + M138N	1.07 ± 0.02	0.90 ± 0.04	1.24 ± 0.01
Q46R	1.02 ± 0.02	1.06 ± 0.04	0.95 ± 0.10
G85K	1.00 ± 0.01	1.03 ± 0.02	1.08 ± 0.02
Q46R + G85K	1.05 ± 0.02	1.15 ± 0.03	1.15 0.02

[(a)] Absolute fitness was computed as $W = e^m$

nificantly less fit (Wald $\chi^2 > 7.85$, $P < 0.001$) than pathotypes $P_{1,2}$ and $P_{1,2,3,4}$ that did not differ in fitness (Wald $\chi^2_{1,15} = 2.75$, $P = 0.605$). In L^2/L^2 host genotype, the three pathotypes showed significantly different fitness, which ranked $P_{1,2} < P_{1,2,3,4} < P_{1,2,3}$ (Wald $\chi^2 > 6.25$, $P < 0.038$).

Since pathotype was not a factor on the fitness of the analysed PMMoV genotypes, data were reanalysed by a GzLM considering virus genotype and host genotype as fixed factors in a full factorial model. Virus fitness depended on virus genotype and host genotype (Wald $\chi^2_{12,258} = 131.86$, $P < 0.001$; Wald $\chi^2_{2,258} = 14.43$, $P = 0.001$, respectively), and on the interaction of virus genotype per host genotype (Wald $\chi^2_{24,258} = 185.67$, $P < 0.001$). Since virus fitness depended on the interaction of virus genotype per host genotype, pairwise comparisons (Table 4.2.4) were done to analyse the magnitude and sense of fitness differences between mutant and parental genotypes for each host. In the L^+/L^+ host, mutants A86G, G66V, and T43K were significantly less fit than the parental WT (Wald $\chi^2 > 5.66$, $P \leq 0.017$), while the rest of mutants were as fit as the WT (Wald $\chi^2 < 2.63$, $P > 0.104$, Table 4.2.4). In addition, pairwise comparisons of the mutant fitness relative to that of WT showed that the fitness of mutants A86G, T43K, and G66V was similar among them (Wald $\chi^2 < 3.51$, $P > 0.109$) and lower than the fitness of mutants D50G, T43K+D50G, L13F, L13F+G66V, Q46L, M138N, Q46L+M138N, Q46R, G85K and Q46R+G85K mutants (Wald $\chi^2 > 4.32$, $P <$

Table 4.2.4 Magnitude and sense of fitness differences between mutant and parental genotypes in three susceptible hosts

Virus genotype	Host Genotype											
	L^+/L^+				L^1/L^1				L^2/L^2			
	Wm[(a)]	Wp[(b)]	Wm - Wp	P[(c)]	Wm[(a)]	Wp[(b)]	Wm - Wp	P[(c)]	Wm[(a)]	Wp[(b)]	Wm - Wp	P[(c)]
M138N	1.20	1.21	-0.01	0.650	1.19	1.22	-0.03	0.245	1.14	1.07	0.07	0.009
A86G	1.14	1.21	-0.07	0.017	1.07	1.22	-0.15	<0.001	1.12	1.07	0.05	0.158
L13F + G66V	1.01	1.03	-0.02	0.477	0.90	1.06	-0.15	<0.001	1.21	1.15	0.05	0.147
L13F	1.04	1.03	0.01	0.658	1.21	1.06	0.15	0.001	1.17	1.15	0.01	0.687
G66V	0.92	1.03	-0.10	0.001	0.89	1.06	-0.17	0.005	1.20	1.15	0.05	0.130
T43K + D50G	0.99	1.03	-0.04	0.312	0.88	1.06	-0.17	0.001	1.06	1.15	-0.09	0.020
T43K	0.93	1.03	-0.10	0.001	0.95	1.06	-0.11	0.121	1.04	1.15	-0.11	0.015
D50G	1.06	1.03	0.03	0.259	1.00	1.06	-0.05	0.372	1.22	1.15	0.07	0.162
Q46L + M138N	1.08	1.03	0.05	0.105	0.91	1.06	-0.15	0.014	1.24	1.15	0.09	0.009
Q46L	1.07	1.03	0.04	0.317	1.04	1.06	-0.01	0.855	0.98	1.15	-0.17	<0.001
Q46R + G85K	1.05	1.03	0.03	0.420	1.15	1.06	0.10	0.051	1.15	1.15	0.00	0.969
Q46R	1.03	1.03	0.00	0.984	1.06	1.06	0.01	0.881	0.96	1.15	-0.20	0.067
G85K	1.01	1.03	-0.02	0.293	1.04	1.06	-0.02	0.670	1.08	1.15	-0.07	0.056

(a) W_m, average fitness of the mutant in each host at 21 dpi.
(b) W_p, average fitness of the parental genotype WT.
(c) Fitness differences of mutant and parental genotype significant at a P value of ≤ 0.05 level according to a LSD analysis are underlined

0.045), with no differences among the fitness of this second group of mutants (Wald $\chi^2 < 1.84$, $P > 0.080$). In the L^1/L^1 host, mutant L13F was significantly more (Wald $\chi^2_{1,14} = 10.47$, $P = 0.001$), and mutants A86G, T43K+D50G, G66V, L13F+G66V and, Q46L+M138N were significantly less fit than the parental WT (Wald $\chi^2 > 6.09$, $P \leq 0.014$), while the rest of the mutants were as fit as the WT (Wald $\chi^2 < 3.80$, $P > 0.051$, Table 4.2.4). The pairwise comparisons of mutant fitness relative to that of the WT showed that the fitness of mutants were significantly different and was ranked A86G, T43K+D50G, G66V, L13F+G66V, Q46L+M138N **<** T43K, D50G, G86K, M138N, Q46L, Q46R **<** Q46R+G85K, L13F ($P \leq 0.05$ for the differences). In the L^2/L^2 host, mutants M138N and Q46L+M138N were significantly more (Wald $\chi^2 > 6.82$, $P = 0.009$) and mutants T43K, T43K+D50G and Q46L were significantly less fit than the parental WT (Wald $\chi^2 > 5.37$, $P \leq 0.020$), while the rest of the mutants were as fit as the WT (Wald $\chi^2 < 3.35$, $P > 0.067$, Table 4.2.4). The pairwise comparisons of mutant fitness relative to that of the WT showed that the fitness of mutants Q46R+G86K, L13F, A86G, G66V, L13F+G66V, D50G, M138N and, Q46L+M138N was similar among them (Wald $\chi^2 < 1.27$, $P > 0.354$) and higher than the fitness of Q46R, Q46L, T43K, T43K+D50G and G86K (Wald $\chi^2 > 4.23$, $P < 0.043$), with no significant differences among the fitness of this second group of mutants (Wald $\chi^2 < 2.21$, $P > 0.115$).

In summary, results showed that the overcoming *L*-gene resistance is not linked to fitness penalties in susceptible hosts, as the fitness of the analysed PMMoV genotypes did

not depend on the pathotype. However, resistance-breaking mutations, and the single mutations necessary to generate the resistance-breaking double mutants, may have pleiotropic effects on the within-host fitness of the virus. The occurrence of pleiotropy, and its magnitude and sense, depended on the specific mutation and on the genotype of the susceptible host, being in most cases antagonistic.

4.2.3 **Epistasis between resistance-breaking mutations on the multiplication of PMMoV genotypes in susceptible hosts**

Next, for the double mutants *x*+*y*, it was analysed if there was epistasis between each of the single mutations *x* and *y* that might affect within-host fitness. Epistasis, ε_{xy}, was computed as described in Material and Methods 3.5.2. For each double mutant, *t* tests were done to assess if the value of epistasis, ε_{xy}, was significantly different from 0. Epistasis between mutations Q46L and M138N could not be estimated as estimates for the absolute fitness of the corresponding mutants derived from two different experiments. Table 4.2.5 shows in which cases epistasis was detected, and the different types of epistatic interactions in all susceptible

Table 4.2.5 Epistatic interactions between pairs of mutations in three susceptible hosts

Host	WT	*x* [(a)]	Fitness of *x* [(b)]	*y* [(a)]	Fitness of *y* [(b)]	Fitness of *x* + *y* [(c)]	*t* test	df [(c)]	Epistasis ε_{xy} [(d)]	*P* for epistasis [(e)]	Type of Epistasis
L^+/L^+	1.03	T43K	0.93	D50G	1.06	0.99	1.285	30	0.034 ± 0.026	0.300	–
L^+/L^+	1.03	L13F	1.04	G66V	0.93	1.01	3.764	29	0.079 ± 0.021	< 0.001	Sign
L^+/L^+	1.03	Q46R	1.03	G85K	1.01	1.05	2.164	29	0.047 ± 0.022	0.040	Magnitude
L^1/L^1	1.06	T43K	0.95	D50G	1.00	0.88	0.360	29	-0.016 ± 0.044	0.700	–
L^1/L^1	1.06	L13F	1.21	G66V	0.89	0.90	3.307	29	-0.118 ± 0.036	0.005	Sign
L^1/L^1	1.06	Q46R	1.06	G85K	1.04	1.15	3.123	30	0.114 ± 0.037	0.005	Magnitude
L^2/L^2	1.15	T43K	1.04	D50G	1.22	1.06	1.538	29	-0.052 ± 0.033	0.200	–
L^2/L^2	1.15	L13F	1.17	G66V	1.20	1.21	0.560	30	-0.014 ± 0.024	0.600	–
L^2/L^2	1.15	Q46R	0.96	G85K	1.08	1.15	5.410	28	0.293 ± 0.054	< 0.001	Reciprocal Sign

[(a)] *x* and *y* indicate the single mutation.
[(b)] Fitness differences between the single mutants and the WT significant at a *P* value of ≤ 0.05 according to a LSD analysis are underlined.
[(c)] df, degree of freedom.
[(d)] ε_{xy}, epistasis among pair of mutations *x* and *y*.
[(e)] Epistasis significant at a *P* value of ≤ 0.05 level according to a *t* test are underlined. '–' indicates no epistasis.

hosts L^+/L^+, L^1/L^1 and L^2/L^2. In total, four cases showed significant positive epistasis, whereas one case had negative epistasis (all $P \leq 0.04$). There was significant sign epistasis between mutations L13F and G66V in hosts L^+/L^+ and L^1/L^1, magnitude epistasis between mutations Q46R and G85K in hosts L^+/L^+ and L^1/L^1, and reciprocal sign epistasis between mutations Q46R and G85K in host in L^2/L^2. Interestingly no epistatic interaction was detected between the pair of mutations T43K and D50G in any host.

In summary, the occurrence and type of epistasis depended on the different pairs of mutations, and on the host genotype, showing a genotype x genotype x environment interaction.

4.2.4 **Virulence of different PMMoV genotypes in susceptible hosts**

To analyse the possible effect of the set of resistance breaking mutations on virulence, virulence was quantified as the effect of infection on plant biomass, and the derived virulence values are presented in Table 4.2.6.

Table 4.2.6 Virulence of different PMMoV mutants in three susceptible hosts

Viral genotype	Host genotype		
	L^+/L^+	L^1/L^1	L^2/L^2
WT	0.15 ± 0.02	0.17 ± 0.05	0.69 ± 0.03
M138N	0.10 ± 0.03	0.20 ± 0.05	0.49 ± 0.02
A86G	0.50 ± 0.03	0.33 ± 0.05	0.89 ± 0.00
L13F + G66V	0.13 ± 0.05	0.22 ± 0.05	0.79 ± 0.03
L13F	0.22 ± 0.05	0.17 ± 0.03	0.80 ± 0.03
G66V	0.10 ± 0.03	0.22 ± 0.06	0.82 ± 0.02
T43K + D50G	0.06 ± 0.08	0.22 ± 0.04	0.54 ± 0.05
T43K	0.05 ± 0.07	0.34 ± 0.07	0.67 ± 0.02
D50G	0.05 ± 0.05	0.30 ± 0.06	0.77 ± 0.03
Q46L + M138N	-0.01 ± 0.03	0.25 ± 0.02	0.38 ± 0.06
Q46L	0.19 ± 0.05	0.47 ± 0.06	0.73 ± 0.03
Q46R + G85K	0.34 ± 0.04	0.11 ± 0.03	0.85 ± 0.02
Q46R	0.04 ± 0.03	0.20 ± 0.05	0.32 ± 0.05
G85K	0.56 ± 0.04	0.62 ± 0.03	0.83 ± 0.01

[(a)] Virulence was computed as $V = 1 - (P_i/P_m)$, where P_i is the dry weight of the total above-ground biomass of each infected plant and P_m the mean dry weight of mock-inoculated plants. Data are means ± standard errors from eight replicated plants.

To determine the effect of pathotype on virulence, virulence data (Table 4.2.6) were analysed by a GzLM considering virus pathotype and host genotype as fixed factors and virus genotype nested to pathotype as a random factor in a full factorial model. Virulence depended

on virus pathotype, host genotype and virus genotype nested to pathotype (Wald $\chi^2_{2,285}$ = 300.01, P < 0.001; Wald $\chi^2_{2,285}$ = 1191.91, P < 0.001; Wald $\chi^2_{11,285}$ = 168.31, P = 0.001, respectively), and on the interactions of virus pathotype per host genotype and on virus genotype nested to pathotype per host genotype (Wald $\chi^2_{4,285}$ = 62.63, P < 0.001; Wald $\chi^2_{22,285}$ = 226.17, P < 0.001, respectively). In the L^+/L^+ host genotype, pathotype $P_{1,2,3,4}$ was significantly more virulent than pathotypes $P_{1,2}$ and $P_{1,2,3}$ (Wald χ^2 > 6.73, P < 0.001), with no significant difference in virulence between them (Wald $\chi^2_{1,13}$ = 0.85, P = 0.111). In L^1/L^1 host genotype, pathotype $P_{1,2,3,4}$ was significantly more virulent (Wald χ^2 > 5.23, P < 0.016) than pathotypes $P_{1,2}$ and $P_{1,2,3}$ with marginally significant difference in virulence between them (Wald $\chi^2_{1,14}$ = 3.45, P = 0.069). In L^2/L^2 host genotype, all pathotypes differed in virulence, ranking $P_{1,2,3}$ < $P_{1,2}$ < $P_{1,2,3,4}$ (Wald χ^2 > 7.60, P < 0.001).

Then, the virulence data were analysed by a GzLM considering virus and host genotypes fixed factors in a full factorial model. Virulence depended on virus genotype, host genotype and on their interaction (Wald $\chi^2_{13,289}$ = 465.2, P < 0.001, Wald $\chi^2_{2,289}$ = 1251.2, P < 0.001, Wald $\chi^2_{26,289}$ = 303.96, P < 0.001, respectively). Since virulence depended on the interaction of virus genotype per host genotype, pairwise comparisons (Table 4.2.7) were do-

Table 4.2.7 Magnitude and sense of virulence differences between mutant and parental genotypes in three susceptible hosts

Virus genotype	Host Genotype											
	L^+/L^+				L^1/L^1				L^2/L^2			
	V_m[(a)]	V_p[(b)]	V_m - V_p	P[(c)]	V_m[(a)]	V_p[(b)]	V_m - V_p	P[(c)]	V_m[(a)]	V_p[(b)]	V_m - V_p	P[(c)]
M138N	0.10	0.15	-0.05	0.461	0.20	0.17	0.03	0.547	0.49	0.69	-0.19	0.001
A86G	0.50	0.15	0.34	0.001	0.33	0.17	0.16	0.053	0.89	0.69	0.20	0.001
L13F + G66V	0.13	0.15	-0.02	0.695	0.22	0.17	0.05	0.712	0.79	0.69	0.10	0.023
L13F	0.22	0.15	0.06	0.345	0.17	0.17	0.00	0.675	0.80	0.69	0.11	0.01
G66V	0.10	0.15	-0.05	0.478	0.22	0.17	0.04	0.5	0.82	0.69	0.13	0.001
T43K + D50G	0.06	0.15	-0.09	0.177	0.22	0.17	0.04	0.768	0.54	0.69	-0.14	0.017
T43K	0.05	0.15	-0.10	0.157	0.34	0.17	0.17	0.082	0.67	0.69	-0.02	0.571
D50G	0.05	0.15	-0.10	0.152	0.30	0.17	0.13	0.162	0.77	0.69	0.08	0.045
Q46L + M138N	0.01	0.15	-0.17	0.015	0.25	0.17	0.08	0.344	0.38	0.69	-0.30	0.001
Q46L	0.19	0.15	0.03	0.61	0.47	0.17	0.29	0.001	0.73	0.69	0.04	0.34
Q46R + G85K	0.34	0.15	0.18	0.011	0.11	0.17	-0.06	0.179	0.85	0.69	0.16	0.001
Q46R	0.04	0.15	-0.11	0.005	0.20	0.17	0.03	0.589	0.32	0.69	-0.36	0.001
G85K	0.56	0.15	0.41	0.001	0.62	0.17	0.45	0.001	0.83	0.69	0.14	0.001

(a) V_m, average virulence of the mutant in each host at 21 dpi.
(b) V_p, average virulence of the parental genotype WT.
(c) Virulence differences of mutant and parental genotype significant at a P value of $\leq$ 0.05 level according to a LSD analysis are underlined

ne to analyse the magnitude and sense of virulence differences between mutant and parental genotypes for each host. In the L^+/L^+ host genotype, genotypes Q46R+G85K, G85K and, A86G were significantly more (Wald $\chi^2 > 15.65$, $P \leq 0.011$) and Q46L+M138N and Q46R were significantly less virulent than the parental WT (Wald $\chi^2 > 14.21$, $P \leq 0.015$), while the rest of the genotypes were as virulent as WT (Wald $\chi^2 < 3.81$, $P > 0.152$, Table 4.2.7). In addition, pairwise comparisons showed that the genotypes A86G and G85K were similar between them (Wald $\chi^2_{1,15} = 2.87$, $P = 0.243$) and more virulent than the other 12 genotypes, WT, T43K, D50G, T43K+D50G, L13F, G66V, L13F+G66V, Q46L, M138N, Q46L+M138N, Q46R and, Q46R+G85K (Wald $\chi^2 > 22.31$, $P < 0.001$) with no significant difference in virulence among them (Wald $\chi^2 < 4.65$, $P > 0.064$). In the L^1/L^1 host genotype, genotypes G85K and, Q46L were significantly more virulent than the parental WT (Wald $\chi^2 > 13.05$, $P < 0.001$), while the rest of the genotypes were as virulent as the WT (Wald $\chi^2 < 6.25$, $P > 0.052$, Table 4.2.7). The pairwise comparisons showed that the genotype G85K was more virulent than the other 13 genotypes, WT, A86G, T43K, D50G, T43K+D50G, L13F, G66V, L13F+G66V, Q46L, M138N, Q46L+M138N, Q46R and, Q46R+G85K (Wald $\chi^2 > 10.12$, $P < 0.017$) with no significant difference in virulence among them (Wald $\chi^2 < 3.80$, $P > 0.171$). In the L^2/L^2 host genotype, genotypes A86G, D50G, L13F, G66V, L13F+G66V, G85K and, Q46R+G85K were significantly more (Wald $\chi^2 > 4.02$, $P \leq 0.045$) and T43K+D50G, M138N, Q46L+M138N and, Q46R were significantly less virulent than the parental WT (Wald $\chi^2 > 6.74$, $P \leq 0.015$), while the genotypes T43K and Q46L were as virulent as the WT (Wald $\chi^2 < 2.74$, $P > 0.340$, Table 4.2.7). The pairwise comparisons showed that the genotypes T43K+D50G, M138N, Q46L+M138N and, Q46R had similar virulence (Wald $\chi^2 < 1.48$, $P > 0.106$), which was lesser than the other ten genotypes, WT, A86G, T43K, D50G, L13F, G66V, L13F+G66V, Q46L, G85K and, Q46R+G85K (Wald $\chi^2 > 8.46$, $P < 0.026$) with no significant difference in virulence among them ($P > 0.065$).

In summary, results showed that the resistance-breaking mutations, and the single mutations necessary to generate the resistance-breaking double mutants, may have pleiotropic effects on virulence. The magnitude and sense of the pleiotropy depended on the pathotype, on the specific mutation and on the genotype of the susceptible host.

4.2.5 **Effect of resistance-breaking mutations on virus particle stability**

The stability of virus particles of genotype WT (pathotype $P_{1,2}$) and its derived CP mutants of $P_{1,2,3}$ pathotype (M138N, T43K+D50G and L13F+G66V) and of $P_{1,2,3,4}$ pathotype A86G was compared through the kinetics of particle disassembly were analysed in both 0.1 M Tris-HCl pH 8.75 and in 6 M urea (Figure 4.5). All the treatments were repeated at least three times. The shown electrophoreses illustrate how RNA mobility increases from that of RNA encapsid-

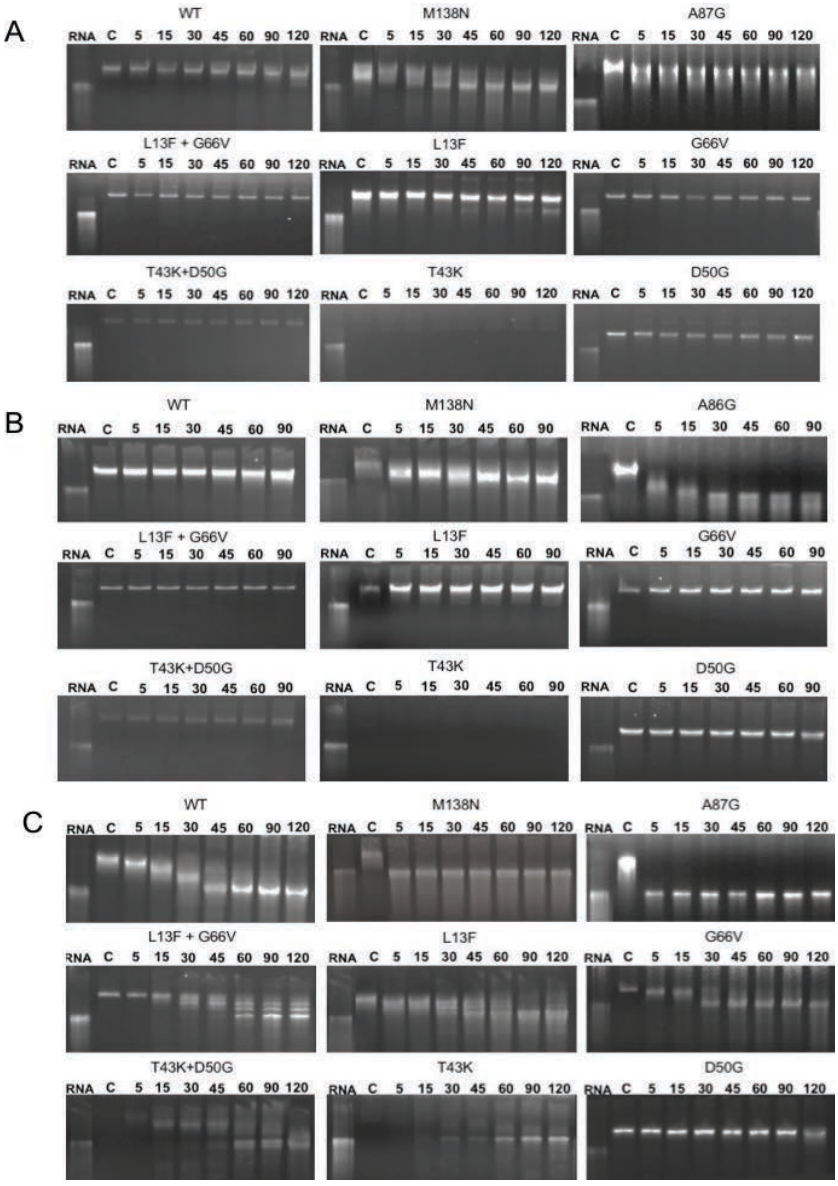


Figure 4.5 Disassembly of virus particles of MG-WT and coat protein mutants in 0.1 M Tris – HCl pH 8.75 (A), 6 M urea (pH 7.4) (B) and, 6M Urea (pH 10.0) (C). Each panel indicates the incubation time in minutes. RNA, electrophoretic mobility of genomic RNA, C, electrophoretic mobility of virus particles at incubation time zero (negative control).

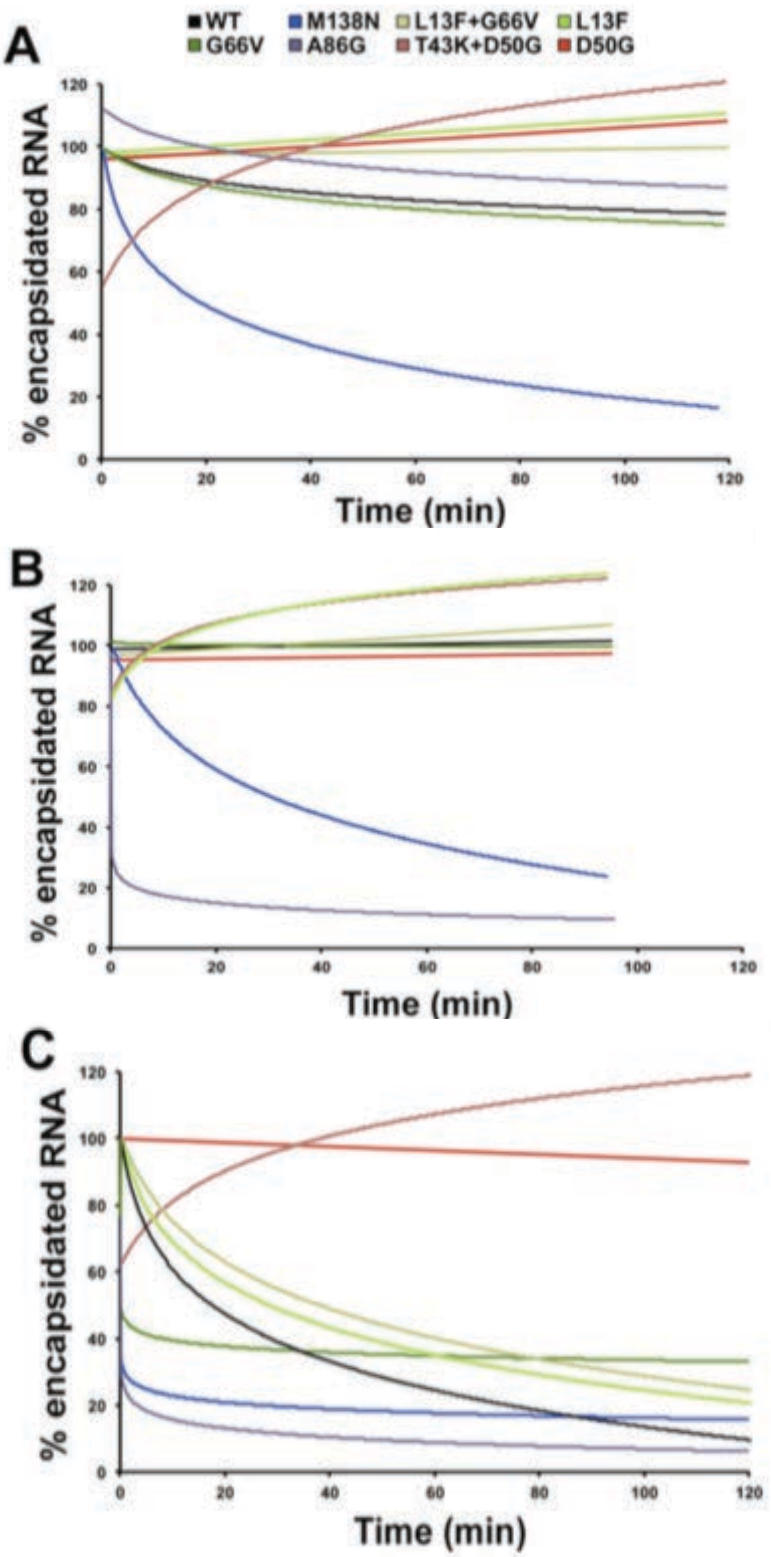


Figure 4.6 Kinetics of disassembly of WT and its derived coat protein mutants in 0.1 M Tris-HCl (pH 8.75) (A), 6 M urea (pH 7.4) (B), and 6 M urea (pH 10.0) (C). Disassembly curves are presented as the percentage of encapsidated RNA over a period of 90 to 120 min. Curves were adjusted using data from at least three replicated assays and according to exponential functions of the form $y = ab^x$.

ated in the assembled virus particle (C, zero time controls) to that of free RNA after full disassembly (purified genomic RNA, indicated as RNA), as particles disassemble during treatment (e.g., genotype A86G in 6 M Urea pH 7.4); or does not change because particle disassembly does not occur (e.g., genotype L13F+G66V in 6 M Urea pH 7.4). Note that the ethidium bromide fails to stain the RNA in particles of genotype T43K, and only weakly stains the RNA in particles of genotype T43K+D50G in all treatments.

The amount of RNA in virus particles was quantified by densitometry and exponential curves describing the disassembly kinetics were adjusted on the data of at least three replicate experiments (Figure 4.6). Note that no densitometry and, hence, no disassembly curves, could be obtained for T43K, as encapsidated RNA did not stain, and that curves for T43K+D50G are partly artefactual, showing an apparent increase in encapsidated RNA with incubation time due to again poor staining of encapsidated RNA at early times, followed by a more efficient staining as particle structure relaxes upon treatment (compare Figures 4.5 and 4.6). Comparison of the slopes of linearized disassembly kinetics curves showed that in Tris-HCl pH 8.75 particle stability ranked T43K+D50G > D50G = G66V = L13F+G66V > L13F = A86G = WT > M138N ($P \leq 0.018$) (Figure 4.6A); in 6M urea pH 7.4 particle stability ranked T43K+D50G = G66V > D50G =L13F= L13F+G66V = WT > M138N > A86G ($P \leq 0.023$) (Figure 4.6B) and in 6M urea pH 10 particle stability ranked T43K+D50G = D50G > G66V = L13F+G66V = WT > L13F = M138N = A86G ($P \leq 0.030$) (Figure 4.6C).

Results from *in vitro* disassembly experiments indicate that resistance-breaking mutations in the CP responsible for the conversion of pathotype $P_{1,2}$ to pathotypes $P_{1,2,3}$ or $P_{1,2,3,4}$ have pleiotropic effects on virus particle stability, but the magnitude and sense of these effects depended on the specific mutations rather than on the pathotype.

5 DISCUSSION

Host range in a key property of parasites, which conditions their epidemiology and evolution (Frank, 1996; Woolhouse *et al.,* 2001; Lajeunesse and Forbes, 2002). Host range is a parasite trait that may evolve, leading to host switches or host range expansions and understanding host range evolution is highly relevant, as it is at the root of processes such as disease emergence, or the efficiency and sustainability of strategies for the control of infectious diseases (García-Arenal and Fraile, 2013; García-Arenal and McDonald, 2003; McDonald, 2004; Adams and Antoniw, 2006). Understanding the evolution of host range in plant viruses may be particularly relevant, as viruses represent about half the number of emerging plant pathogens (Anderson *et al.*, 2004), and as host range evolution compromises the durability of resistance bred in to crop cultivars, the only direct and most efficient strategy for the control of plant virus diseases (García-Arenal and McDonald, 2003; Adams and Antoniw, 2006).

It is often speculated that RNA viruses have a high potential to adapt to new hosts because their high census numbers and the high error rates of RNA-dependent RNA polymerases will translate into a strong potential to generate host range mutants (Elena *et al.*, 2014; Bedhomme *et al.*, 2015; Woolhouse *et al.*, 2005). However, RNA viruses have small, information-compact genomes encoding few multifunctional proteins, which will favour pleiotropic effects of mutations (Bedhomme *et al.*, 2015; Elena and Sanjuán, 2007). A consequence can be the generation of across-host fitness trade-offs so that mutants that perform well in one host will perform poorly in another one, limiting host range expansion and, more generally, virus emergence (Elena and Sanjuán, 2007; Gandon *et al.*, 2013). It is thought that the major cause of across-host fitness trade-offs in viruses is antagonistic pleiotropy, resulting from the interaction of the virus genotype per host environment (G x E), so that there is a negative fitness effect in the original host of mutations that increase fitness in the new one (Whitlock, 1996). Antagonistic pleiotropy limits the range of adaptation and promotes the evolution of ecological specialization (Remold, 2012). Along with antagonistic pleiotropy, recent studies had also shown the role of epistasis (genotype per genotype interactions, G x G) in the host range evolution of RNA viruses (Hillung *et al.*, 2014; Hillung *et al.*, 2015). In this thesis I have analysed if adaptation of plant RNA viruses to new hosts (either new host species or new host genotypes) is limited by across-host fitness trade-offs, and if such trade-offs are generated by antagonistic pleiotropy of, and by epistasis between, host range mutations. Evidence for antagonistic pleiotropy and epistasis as determinants of across-host derives mostly from experiments in which a virus adapts to a new host in the course of serial passages (Hillung *et al.*, 2014; Hillung *et al.*, 2015; Bedhomme *et al.*, 2015). The approach followed here is different, as it is based on the analysis of field isolates of tobamoviruses that infect pepper crops, and of the host range mutations reported in field isolates of these viruses. The analysis of field isolates and of field-occurring mutations may better illustrate on the role of across-host

fitness trade-offs in the evolution of viruses in crops than the analysis of genotypes generated in serial passage experiments.

5.1 **Are the isolates of *Tobacco mild green mosaic virus* adapted to their host of origin?**

To detect host adaptation of a pathogen, Kawecki and Ebert, (2004) had proposed two criteria, 'local versus foreign' and 'home versus away'. The 'local versus foreign' criterium emphasizes comparison between pathogens within hosts: in each host, it is expected that the original pathogen of the host will multiply to higher levels than a foreign pathogen which comes from another host. In contrast, the 'home versus away' criterium emphasizes the comparison of a pathogen's multiplication across hosts: host adaptation would be said to occur if each pathogen showed a higher level of multiplication in its own host (at home) than in other hosts (away). Therefore, to check for the possibility of adaptation of TMGMV to their host of origin and to analyse the role of antagonistic pleiotropy and epistasis in the adaptation (if present), a reciprocal inoculation experiment was performed. Twenty-six TMGMV isolates from *N. glauca* (14) and pepper (12) were used to inoculate both *N. glauca* and pepper plants and the multiplication of each isolate was quantified in the assayed hosts, as a proxy to their within-host fitness. According to the criteria proposed by Kawecki and Ebert, (2004), isolates from *N. glauca* were adapted to its own host as indicated by the fact that they multiplied to higher levels in this hosts than isolates from pepper thus fulfilling the 'local versus foreign' criterion (Tables 4.1.4 and 4.1.5). According to this criterium though, pepper isolates were not adapted to pepper. We could neither conclude that pepper isolates were adapted to pepper according to the 'home versus away' criterion: support for this criterion was ambiguous, as pepper isolates multiplied to higher levels in pepper cv. Doux des Landes than in *N. glauca*, but this was not the case in pepper cv. Dulce Italiano where the pepper isolates multiplied to lower levels (Tables 4.1.4 and 4.1.5). Also, it may be pointed out that the 'home versus away' criterium is not as unequivocal as the "local versus foreign" to detect host adaptation, because it may confound the effects of divergent selection on the virus with intrinsic differences in habitat quality. For instance, in our case, it is clear that pepper cv. Doux des Landes provides a better habitat (is a better host) for all TMGMV isolates than pepper cv. Dulce Italiano or *N. glauca*. Thus, it can be concluded that there is evidence in support of adaptation of *N. glauca* isolates, but not of pepper isolates, to their host or origin. Because pepper isolates mutliplied in *N. glauca* to lower levels than *N. glauca* isolates, there is a cost for the virus isolates from pepper. This cost is not indicative of an adaptation trade-off, as there is no adpatation to

pepper. It could be interpreted as an atypical instance of maladaptation (Bull, 1994; Kaltz *et al.*, 1999; M., Lajeunesse and Forbes, 2002).

The association of TMGMV with *N. glauca* is certainly old, and may have ocurred worldwide since the second half of the XIX century, as the virus was found infecting this plant in the Canary Islands and in Eastern Australina since late in the XIX century or early in the XX century (McKinney, 1929; Fraile *et al.*, 1996). In contrast, TMGMV infection in pepper crops was first reported in the early 1980s (Wetter, 1984), which given the relevance of viral diseases in this crop, and the attention paid by virologists to pepper viruses, strongly suggests that infection of pepper by TMGMV was not much older. These data strongly suggest that *N. glauca* is the original or reservoir host for TMGMV, to which the virus is adapted, and pepper a much more recent host to which adpatation has not occurred. Still data on the the adaptation of viruses to new hosts show it may be fast (Moreno-Pérez *et al.*, 2014; Sacristan *et al.*, 2005). It can be speculated that adaptation of TMGMV to pepper in SE Spain was prevented by frequent rates of inoculum migration between sympatric *N. glauca* and pepper crops, and/or by the extensive use of L^1 gene resistance in pepper crops since the early 1980s for the control to *Tomato mosaic virus* (Fraile *et al*., 2011). L^1 resistance being effective against TMGMV could have prevented the building of high effective population sizes of this virus in pepper crops. Also, the fast turnover of varieties and cultivars in the intensive pepper cropping system in SE Spain (Fraile et al. 2011) would mean a faster rate of genetic change of the host than of the virus, which may prevent adaptation and result in maladaptation (Kaltz *et al*., 1999).

5.2 **Role of mutations in the coat protein and 3' untranslated region of *Tobacco mild green mosaic virus* in determining virus multiplicaiton in pepper and *N. glauca***

In an attempt to identify determinants of the differential multiplication of TMGMV in *N. glauca* and pepper, we focussed on three loci that were dimorphic in the analysed set of 26 TMGMV isolates, at positions 8 and 148 of the CP and in the 3' UTR of TMGMV. Of these three dimorphisms, only that at position 8 of CP was significantly associated with the TMGMV isolate host of origin. Analyses of the multiplication in pepper and *N. glauca* of TMGMV isolates with either a proline or a serine at CP position 8 revealed that isolates with a proline showed a trend for higher multiplication in pepper, significant only in some cultivars, than isolates with a serine at position 8 of CP (Table 4.1.6). Results also indicated that a proline at position 8 of CP in pepper isolates affected negatively multiplication in *N. glauca*, with no effect in pepper,

as compared with the effect of a serine at position 8 of the CP in *N. glauca* isolates (Table 4.1.7). Last, for *N. glauca* isolates, multiplication of those with a proline at position 8 of CP was higher in its original host, and in some pepper cultivars (Table 4.1.8). Taken together, these results indicate that a proline at position 8 of the CP is generally positive for the virus, but the extent of this positive effect depends on the assayed host, indicating a pleiotropic effect of the mutation, and on the genetic background, as it differs in isolate from pepper and from *N. glauca*, indicating epistatic interactions. Moreover, the results also showed higher order interactions involving pleiotropy and epistasis (G x G x E). Evidence of this derives from the ranking of virus multiplication, which in pepper is *N. glauca* isolates with a serine < pepper isolates with a proline < *N. glauca* isolates with a proline and in *N. glauca* is pepper isolates with a proline < *N. glauca* isolates with a serine < *N. glauca* isolates with a proline (Tables 4.1.7 and 4.1.8). From these results the fitness landscapes of TMGMV in each host shown in Figure 5.1 can be derived, by considering as lowest the multiplication of pepper isolates with a proline, as they were not found in the sampled population. The landscape would be different in each host, indicating reciprocal sign epistasis in pepper and magnitude epistasis in *N. glacua* (Figure 5.1).

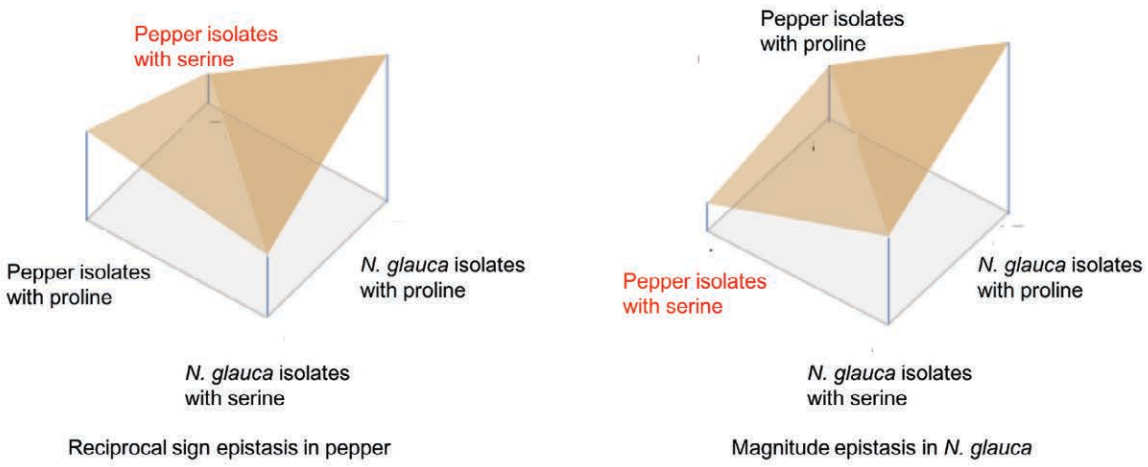


Figure 5.1 Fitness landscape of Tobacco mild green mottle virus isolates from pepper and *N. glauca* in these two hosts.

The amino-acid at position 148 of CP, alanine/threonine, was equally distributed in isolates from both host of origin and this dimorphism of either a hydrophobic (alanine) or a hydrophilic (threonine) amino-acid was intriguing to me. So, I checked for its effect on virus multiplication despite no association with the host of origin. The results indicated that in isolates from both hosts, pepper and *N. glauca*, a threonine at position 148 of CP was

associated with higher multiplication than an alanine in both assayed hosts (Table 4.1.10). So, there is no evidence of antagonistic pleiotropy in virus multiplication associated to this dimorphism. It can be speculated that the position 148 of CP might play other roles, as in the case of TMV, where the substitution of a hydrophilic (serine) for a hydrophobic (phenylalanine) amino-acid resulted in initiating N'-mediated hypersensitive reaction (Culver *et al.*, 1994). The amino acid at CP position 148 is in the outer surface of the viroin structure (James N Culver, 2002; Bera, Manuel G Moreno-Pérez, *et al.*, 2017) and although in this study both hosts were susceptible to TMGMV, this position might play a role in interacting with host proteins paving the way to an efficient infection thus increasing the level of multiplication of isolates with a threonine at this position. The positive effect of a threonine on virus multiplication in all assayed hosts gives rise to the question of why this sequence variant is not fixed in the population, and the alanine polymorphism is maintained. One posibility, of course, is that the virus populations is not at equilibrium, which certainly would be the case for the pepper population, although evidence for equilibrium was reported for the virus population in *N. glauca* (Moya *et al.*, 1993). A second posibility is that TMGMV infects other, non-identified hosts in the same region, in which an alanine at CP position 148 would be advantageous.

The short/long dimorphism of 3' UTR of TMGMV was intriguing to me due to the frequent presence of long 3' UTR in isolates from *N. glauca* as found in other virus populations (Bodaghi *et al.*, 2000; Bodaghi *et al.*, 2004) but not in ours so despite of no association with the host of origin, I decided to study its' effect on virus multiplication. The results indicated that the TMGMV isolates having a short/long 3' UTR, did not have an effect on virus multiplication in *N. glauca* but had an effect in pepper, with isolates having a short 3' UTR multiplying to higher levels than the isolates having long 3' UTR (Table 4.1.11). In addition, in all assayed hosts multiplication was higher for pepper isolates with a short 3'UTR and for *N. glauca* isolates with a long 3'UTR (Table 4.1.12). These results indicate that the effect of the 3'UTR length on virus multiplication depends on the genetic background, i.e., they are evidence of epistasis with non-identified loci in other parts of the genome. The hypothesis of epistatic interactions involving the 3'UTR is also supported by the fact that long or short 3'UTR were selected in *N. glauca* upon passaging in this host of different virus isolates (Figure 4.4). It remains to be elucidate the exact function of the duplication at the 3'UTR but as this region has a role at binding the replication machinery and initiating the replication of the genome, it could affect the efficiency of replication, as shown in other viruses (Villordo *et al.*, 2016; Gritsun *et al.*, 2014). In agreement with this hypothesis, Bodaghi *et al.*, (2004) did showed that TMGMV isolates having a long 3' UTR outcompeted those having a short 3'UTR in mixed infections in *N. glauca*. It is interesting to point that for the *N. glauca* isolates the length dimorphism at TMGMV 3'UTR is associated to the sampling date of the isolate, the long end

being more frequent in summer and the short end in winter ($\chi^2_{(1,14)} = 7.14$, $P = 0.008$). This suggests that the 3'UTR could have a temperature differential effect on the genome replication, a suggestive hypothesis that remains to be tested.

From the results of field isolates of TMGMV, it was clear that the amino-acid at position 148 of CP and the short or long 3' UTR of TMGMV had an effect on virus multiplication which might be dependent on the isolate host of origin and on the assayed host. But, these results might include the effect of other unknown mutations in other parts of the viral genome. So, to analyse the role in virus multiplication of the amino-acid at CP position 148 and the length of the 3'UTR in defined genetic backgrounds substitutions were introduced in the biologically active infectious clones of pTMGMV-FA and pTMGMV-Jap. These two clones were selected because TMGMV-FA derives from a field isolate from *N. glauca*, Ng 92/73, while TMGMV-Jap derives from a pepper isolate from Japan (Morishima *et al.*, 2003). In the case of TMGMV-FA, where the long 3' UTR was substituted for a short 3' UTR and a threonine at position 148 of CP was substituted for an alanine, the ability to multiply decreased by more than 100 times in *N. glauca* and by 10 times in pepper (Tables 4.1.15 and 4.1.16). In the case of TMGMV-Jap, the short 3' UTR was substituted for a long 3' UTR and an alanine at position 148 of CP was substituted for a threonine, the fitness decreased by more than 1000 times in *N. glauca*, and increased by at least two times in pepper. These results showed unequivocal evidence for the epistatic interactions that were deduced from the analyses of the field isolates. In addition, the analysis of mutants derived from clones showed that in TMGMV-Jap, but not in TMGMV-FA, a threonine at position 148 of CP increased multiplication in pepper and decreased multiplication in *N. glauca,* and a short 3'UTR decreased multiplication in pepper but increased multiplication in *N. glauca*. Thus, it can be concluded that in TMGMV-Jap, but not in TMGMV-FA there are pleiotropic effects of these mutations, another instance of higher order interactions of the type G X G X E.

The detailed analysis of the three dimorphic loci present in the CP and 3' UTR of TMGMV genome shows that both antagonistic pleiotropy and epistasis play a role in limiting the host range expansion of a virus, and may partly explain the loss of fitness of pepper isolates in *N. glauca*.

5.3 **Does resistance-breaking genotypes of *Pepper mild mottle virus* cause fitness penalties in susceptible host genotypes?**

Antagonistic pleiotropy and epistasis may play a pivotal role in the generation of across-host fitness trade-offs that may limit host range expansion. A particular case of host range expansion is the overcoming of host resistance in GFG host-pathogen interactions, which compromises the use of resistance for the control of crop viral diseases. In GFG systems there is a hierarchy of infectivity alleles in the pathogen and of resistance alleles in the host so that some pathogen infectivity alleles are intrinsically better than others, conferring the capacity to infect and multiply in a larger set of host genotypes, and, conversely, some host resistance alleles confer the capacity to resist a larger set of pathogen genotypes (Agrawal and Lively, 2002; Dybdahl *et al.*, 2014). Host-pathogen coevolution under the GFG model has been extensively modelled; models assume that resistance-breaking mutations have fitness costs in the susceptible host (Brown and Tellier, 2011; Brown, 2015), thus generating across-host fitness trade-offs. Previous reports refer to different plant-virus systems (Ishibashi *et al.*, 2012; Janzac *et al.*, 2010; Goulden *et al.*, 1993; Jenner., *et al.*, 2002; Jenner., *et al.*, 2002), including the pepper-PMMoV system analysed here (Fraile *et al.*, 2011; Fraile *et al.*, 2014; Moreno-Pérez *et al.*, 2016)

The interaction of PMMoV genotypes with pepper genotypes carrying different resistance alleles at the *L* locus is according to the GFG model (Berzal-Herranz *et al.*, 1995; de la Cruz *et al.*, 1997; Gilardi *et al.*, 2004; Matsumoto *et al.*, 2008). Most PMMoV field isolates are of a $P_{1,2}$ pathotype; i.e., they infect plants of the universally susceptible L^+/L^+ genotype and those carrying resistance alleles L^1 and L^2, but $P_{1,2,3}$ isolates, which also infect L^3 plants, also occur in pepper crops. Analyses of field PMMoV isolates of $P_{1,2}$ and $P_{1,2,3}$ pathotypes have shown fitness costs associated with L^3 resistance breaking (Fraile *et al.*, 2011). To identify the mechanisms that generate such fitness costs, different mutations in the virus CP reported as determinants of resistance breaking of the L^3 or L^4 alleles ($P_{1,2,3}$ or $P_{1,2,3,4}$ pathotype, respectively) (Antignus *et al*., 2008; Berzal-Herranz *et al.*, 1995; Matsumoto *et al.*, 2008; Tsuda *et al.*, 1998) were introduced into a PMMoV $P_{1,2}$ isolate by site-directed mutagenesis of an infectious cDNA clone, pPMMoV-MG obtained by Moreno-Pérez *et al*., (2016). The set of PMMoV genotypes include all reported resistance-breaking single- and double-mutants, plus the single mutants that are in the evolutionary pathway towards the resistance-breaking double mutants. All the resistant-breaking mutants had the expected pathotype, confirming that the introduced CP mutations are responsible for the viral pathotype. Most of the single mutants whose combination causes resistance-breaking were of $P_{1,2}$ pathotype e.g., mutants

PMMoV-MG-(T43K), PMMoV-MG-(D50G), PMMoV-MG-(L13F), PMMoV-MG-(G66V), PMMoV-MG-(Q46L) and PMMoV-MG-(Q46R) had a $P_{1,2}$ pathotype, except PMMoV-MG-(G85K) which had a $P_{1,2,3,4}$ pathotype. The resistance-breaking double mutant, (Q46R+G85K) was first characterised by Genda *et al.*, (2007) as a $P_{1,2,3,4}$ pathotype, and these authors reported that none of the single mutants Q46R and G85K were able to systemically infect *C. chacoense* (L^4/L^4) up to 15 dpi. In our study, PMMoV-MG-(Q46R) did neither systemically infect *C. chacoense*, but plants inoculated with PMMoV-MG-(G85K) showed systemic infection at the end of the observation period (21 dpi). Plants inoculated with the double mutant PMMoV-MG-(Q46R+G85K) showed systemic infection at 10 dpi.

The fitness of the mutants was compared to that of the parental genotype in three susceptible hosts that differed broadly genetically, as they represent different species of the genus *Capsicum*, but that, for the purpose of this work are characterised by different alleles at the *L* locus (Table 4.2.4). The virus fitness depended on the virus genotype and the host genotype, but did not depend on the virus pathotype, thus there was not a consistent cost of resistance-breaking. This analysis also showed that virus fitness depended on the interaction between virus and host genotypes, i.e., on a genotype per environment (G x E) interaction, which is evidence of pleiotropic effects of the resistance-breaking mutations (Bedhomme *et al.*, 2015; Lalic *et al.*, 2011). When the sense and magnitude of the pleiotropy, whether negative (i.e., antagonistic) or positive, were analysed for the various mutants, it was found to depend on both the specific mutation and on the host genotype. Three out of thirteen mutants showed antagonistic pleiotropy in the universally susceptible host L^+/L Five of 13 mutants showed antagonistic pleiotropy in the L^1/L^1 host, and 3/13 mutants showed antagonistic pleiotropy in L^2/L^2. Conversely, no mutant showed positive pleiotropy in L^+/L^+, only 1/13 mutant showed positive pleiotropy in the L^1/L^1 host and 2/13 mutants showed positive pleiotropy in L^2/L^2 (Table 4.2.4). This trend of increasing numbers of positive pleiotropy is comparable to the study of Lalic *et al.*, (2011) where it was shown that single substitutions introduced in a virus were mostly deleterious in its original host but as the genetic distance between the hosts increased the number of beneficial mutation also increased. Here, PMMoV-MG-WT was derived from an isolate (P84/8) from *C. annuum* (Moreno-Pérez *et al.*, 2016), and the L^2/L^2 host belong to *C. frutescens*, so the increase of positive pleiotropy agrees with the study of Lalic *et al.*, (2011).

Antagonistic pleiotropy has been shown to explain resistance-breaking-associated fitness costs in several plant viruses (Ishibashi *et al.*, 2012; Fraile *et al.*, 2011; Goulden *et al.*, 1993; Jenner, Tomimura, *et al.*, 2002), as shown here for some interactions. Our results also provide evidence of resistance breaking without cost, as occasionally reported (Bedhomme *et*

al., 2012; Coffey and Vignuzzi, 2011; Cooper and Scott, 2001; Greene *et al.*, 2005; Novella *et al.*, 1999). Nevertheless, our results agree with the findings of Moreno-Pérez *et al.*, (2016) and could provide the evidence of the host-dependent pleiotropy of resistance-breaking mutations, which may be costly or favourable in different susceptible hosts.

To determine the fitness landscape of different genotypes of PMMoV in each susceptible host genotype, epistasis was computed. In both hosts L^+/L^+ and L^1/L^1, epistasis was found to be either of magnitude or sign epistasis types, which indicates that the fitness landscape in these hosts are mostly smooth with the presence of small valleys (minor ruggedness). In host L^2/L^2, only reciprocal sign epistasis was found which shows that the fitness landscape has deep valleys (major ruggedness). The two mutations T43K and D50G did not show epistatic interactions in any hosts which means that there will be no limitation to the evolution of the double mutant, while there will be such limitations for the other double mutants. However, all $P_{1,2,3}$ and $P_{1,2,3,4}$ pathotype mutants, with the exception of M138N have been reported only once and in a very limited geographical area (Moury and Verdin, 2012) which could be unrelated to evolutionary histories. The pairs of mutations L13F and G66V, Q46R and G85K, show epistasis of different types, or no epistasis, depending on the host, indicating that the ruggedness of the fitness landscape will be host dependent, another instance of higher order G x G x E interactions. Thus understanding the evolution of resistance breaking in PMMoV would require detailed knowledge of several factors, including the genetic composition of the susceptible host population, making predictions difficult.

Our results on the occurrence and type of epistasis are on line with other studies. For example, Hillung *et al*., (2015) and Lalić and Elena, (2015) analysed epistasis between mutations fixed during the evolution of Tobacco etch virus (TEV) in different ecotypes of *A. thaliana*, finding that magnitude and sign epistasis was more common than reciprocal sign epistasis These, and our study, may suggest that magnitude and sign epistasis are the most frequent interactions between mutations generated during virus evolution in different hosts. On the other hand, reciprocal sign epistasis was the most frequent between mutations introduced randomly in the genome of TEV (Lalić and Elena, 2012).

5.4 **Do resistance-breaking mutations of *Pepper mild mottle virus* affect virulence?**

Apart from the within-host multiplication component of virus fitness, resistance-breaking mutations could also affect other components of the virus life history, such as virulence. Therefore, I analysed the virulence of the parental and the mutant genotypes in the three susceptible hosts. Plant biomass correlates with both plant fecundity and survival (Sacristán and García-Arenal, 2008; Escriu *et al.*, 2003; Doumayrou *et al.*, 2013); thus, the effect of infection on biomass is a proxy to virulence. As was the case for fitness, virulence depended on the interaction of G x E, revealing a new pleiotropic effect of resistance-breaking mutations. Pleiotropic effects on virulence were mutation-specific and were unrelated to those on fitness, and, overall, virulence and fitness did not correlate, as is often the case for plant viruses (Alizon *et al.*, 2009; García-Arenal and Fraile, 2013). It is noteworthy that although pathotype has a significant effect on virulence, increased infectivity does not necessarily translate into virulence changes and this agrees with the findings of Moreno-Pérez *et al.*, (2016).

This is an interesting result as the relationship between infectivity and virulence has rarely been analysed either experimentally or theoretically (Van Den Bosch *et al.*, 2006). Our data are consistent with the only other analysis we know off, in which it was shown that *Rice yellow mottle virus* genotypes differing in infectivity on the *Rymv-2* resistance allele did not differ in virulence (Sorho *et al.*, 2005); that is, virulence and infectivity were not linked. Virulence may also be an important determinant of resistance-breaking evolution in the pepper-PMMoV system because growers eliminate infected plants when detected (Moury and Verdin, 2012), and detection will be easier in plants infected by more virulent virus genotypes.

5.5 **Are there trade-offs between different components of virus fitness?**

The mutations that determine the breaking of *L*-gene resistance in *Capsicum* by tobamoviruses occur in the CP (Berzal-Herranz *et al.*, 1995; de la Cruz *et al.*, 1997; Gilardi *et al.*, 2004; Matsumoto *et al.*, 2008), so it could be hypothesized that they have effects on particle stability. It has been shown that tobamovirus field isolates of pathotypes P_0, $P_{1,2}$, and $P_{1,2,3}$ differ in particle stability, which was ranked as $P_{1,2} > P_{1,2,3} > P_0$. Particle stability was positively correlated with survival of infectious particles in the soil for extended time periods of up to 6 months, which is usually more than the duration of the crop cycle and hence more than

the infectious period of infected hosts (Fraile *et al.*, 2014). These results demonstrated that increased resistance-breaking capacity from P_0 to $P_{1,2}$ was not associated with a cost on survival but that there was a cost from $P_{1,2}$ to $P_{1,2,3}$. The field isolates of each pathotype assayed by Fraile *et al.*, (2014) differed in various amino acid positions in the CP, so that different haplotypes were described. About 30% of the amino acid positions of the CPs of P_0 isolates (TMGMV) differ from those of $P_{1,2}$ and $P_{1,2,3}$ isolates (PMMoV). Hence, that study could not determine whether changes in particle stability were due specifically to resistance-breaking mutations. This question was analysed in the present work by comparing the effect of reported CP mutations responsible for the conversion of PMMoV pathotype $P_{1,2}$ to pathotypes $P_{1,2,3}$ and $P_{1,2,3,4}$. The results demonstrate that resistance-breaking mutations have effects on particle stability, as estimated by the *in vitro* kinetics of disassembly under conditions of high pH and high urea concentration. These assay conditions were chosen because they have been used extensively in the past to analyse the disassembly kinetics of TMV, providing useful information on its kinetics and mechanisms (James N Culver, 2002), not because they mimic natural soil conditions. Still, as noted above, stability under these assay conditions positively correlates with tobamovirus particle survival in the soil (Fraile *et al.*, 2014). Considering together the data from the three assayed conditions for disassembly, the stability of the different mutants analysed depended on the specific mutation and was ranked as PMMoV-MG-(T43K+D50G) ($P_{1,2,3}$) > PMMoV-MG-(L13F+G66V) ($P_{1,2,3}$) ≥ PMMoV-MG-WT ($P_{1,2}$) ≥ PMMoV-MG-(A86G) ($P_{1,2,3,4}$) ≥ MG-(M138N) ($P_{1,2,3}$).

The present results show pleiotropic effects of resistance-breaking mutations on PMMoV particle stability, a proxy for survival in the soil. These results confirm our previous finding that deployment of resistance in the host population results in selection for resistance breaking (Fraile *et al.*, 2011) and for differential survival, a trait unrelated to the plant-virus interaction (Fraile *et al.*, 2014). As was the case for virus multiplication in susceptible hosts (our present results and Moreno-Pérez *et al.*, 2016), the sense and magnitude of the pleiotropy depend on the specific mutations that determine resistance breaking and not on the pathotype. Hence, there is no trade-off in this system between increased host range (i.e., resistance breaking) and survival, as has been reported for bacteriophages (De Paepe and Taddei, 2006; Keen, 2014). Also, the analysis of the within-host fitness of these coat protein mutants (Table 4.22) shows that fitness and particle stability rank differently. Moreover, no correlation was found between the values of fitness and those of the parameters of exponential curves describing disassembly kinetics ($R^2 \leq 0.074$ and $P \geq 0.514$ in a Spearman test for any assayed disassembly conditions), so there is no evidence of a trade-off between virus multiplication (i.e., reproduction) and particle stability (i.e., survival). To our knowledge, the relationship

between survival and multiplication has not been studied for other plant viruses, but studies with bacterial or animal-infecting viruses have demonstrated that such a trade-off is not general across systems (De Paepe and Taddei, 2006; Keen, 2014; Dessau *et al.*, 2012; Handel *et al.*, 2014). Trade-offs between survival and reproduction, predicted by the life history theory, need not apply for viruses, because there is no obvious mechanistic reason to expect them due to the differences between the extracellular and intracellular environments where survival and reproduction, respectively, occur, as pointed out by Goldhill and Turner, (2014).

Survival may be particularly relevant for pathogens transmitted through the environment or through both the environment and direct contact (Walther and Ewald, 2004; Roche *et al.*, 2011; Gandon, 1998), as is the case for tobamoviruses. Transmission through the environment will break the trade-off between virulence and transmission in directly transmitted pathogens (Alizon *et al.*, 2009), establishing a positive correlation between survival and virulence and allowing the evolution of highly virulent strains with high survival, the "curse of the pharaoh" hypothesis (Bonhoeffer *et al.*, 1996). Again, our data show that the PMMoV CP mutants do not rank similarly for virulence and particle stability, and there is no correlation between the values of virulence and the parameters of exponential curves describing disassembly kinetics ($R^2 \leq 0.159$ and $P \geq 0.207$ in a Spearman test for any assayed disassembly conditions). Again, this result is consistent with previous ones, as comparison of the virulence on different host genotypes of some of the PMMoV mutants analysed here (Fraile *et al.*, 2014) and their particle stability do not provide evidence of a relationship between survival and virulence. Therefore, no survival-virulence correlation appears to occur in the analysed system, even though it fulfills the conditions predicted by theoretical models for such a positive correlation to occur, i.e., that infection prevalence is not at equilibrium and that the death rate of the infected host is not much higher than that of virions in the environment (Bonhoeffer *et al.*, 1996). In epidemics of PMMoV on susceptible or resistant pepper cultivars, prevalence will not be in equilibrium, and PMMoV infection has little effect on plant survival (our unpublished observations). Similarly, our present and previous (Fraile *et al.*, 2014) data show no trade-off between survival and infectivity.

Our results show that joint consideration of different life history traits is necessary for understanding virus evolution, as shown with experimental populations of phages or wild populations of avian influenza virus (Dessau *et al.*, 2012; Brandon Ogbunugafor *et al.*, 2013; Roche *et al.*, 2014). The consequence of the reported pleiotropic effects of resistance-breaking mutations for the evolution of resistance-breaking is a complex one. Selection on traits that determine different components of the virus fitness may be in the same or opposite

directions, which may favor or limit the durability of the resistance, depending on the specific resistance-breaking mutations. An additional conclusion from the results presented here, relevant to the understanding of virus evolution, is that trade-offs between different life history traits, predicted by theory for parasite evolution, may not apply to the evolution of viruses.

6 CONCLUSIONS

In this thesis I have studied if the host-range expansion of pepper-infecting tobamoviruses by acquiring new species or genotypes or hosts involves fitness penalties, and if so, which are the underlying mechanisms. Result show that:

1. TMGMV isolates from *N. glauca* are adapted to tis host, which may be considered as the native reservoir host. The recent acquisition of pepper as a host has not resulted into adaptation of TMGMV isolates to pepper, but has resulted in a loss of fitness in the reservoir host.

2. The analysis in field isolates of TMGMV of the possible role of the amino acid at two dimorphic positions in the coat protein, and of a duplication in the 3' untranslated region, in the differential multiplication in pepper and N, glauca, shoes that the effect depend on the mutations, the genetic background of the isolate as associated to the host of origin, and the assayed host. These results provide evidence of pleiotropic effects of the analysed mutations and of epistatic interactions between them and unidentified positions in the virus genome.

3. The analysis of the effects of these mutations in a defined genetic background was performed by mutagenesis of biologically active cDNA clones derived from two TMGMV isolates. Results confirmed the relevance of pleiotropy and epistasis in determining host differential virus multiplication, and revealed higher order interactions of the type G x G x E of the analysed mutations.

4. The breaking *L*-gene resistance by PMMoV is not linked to fitness penalties in susceptible hosts, as the fitness of engineered PMMoV mutants did not depend on their pathotype.

5. However, the resistance-breaking mutations, and the single mutations necessary to generate the resistance-breaking double mutants of PMMoV, may have pleiotropic effects on the within-host fitness of the virus, which magnitude and sense depend on the specific mutation and the specific genotype of the susceptible host.

6. The occurrence and type of epistasis between resistance-breaking mutations in PMMoV depends on the specific pair of mutations and on the genotype of the susceptible host, mostly revealing a smoothly rugged fitness landscape.

7. Again, the analysis of the within-host fitness of resistance breaking mutants, showed higher order interactions of the type G x G x E among the rsistance-breaking mutations.

8. Resistance-breaking mutations were shown to have pleiotropic effects on two other life history traits of PMMOV, virulence and survival in the environment, again associated to the specific mutations and not to the pathotype.

9. The analysis of the pleiotropic effects of resistance-breaking mutations on within-host multiplication, virulence and survival, indicates there are no trade-offs among these life history traits that would constraint the evolution of the virus.

10. The major general conclusion of this study is that host range evolution in the analysed tobamoviruses will be constrained by higher order interactions between host-range mutations on virus fitness, but not by trade-offs between fitness components.

7 REFERENCES

Adams, M.J. and Antoniw, J.F. (2006) DPVweb : a comprehensive database of plant and fungal virus genes and genomes., **34**, 382–385.

Agrawal, A. and Lively, C.M. (2002) Infection genetics: gene-for-gene versus matching- alleles models and all points in between. *Evol. Ecol. Res.*, **4**, 79–90.

Agrios, G.N. (2005) *Plant Pathology* 5th ed., Elsevier.

Agudelo-Romero, P., Carbonell, P., Perez-Amador, M.A. and Elena, S.F. (2008) Virus Adaptation by Manipulation of Host's Gene Expression. *PLoS One*, **3**, e2397.

Alexander, H.M., Mauck, K.E., Whitfield, A.E., Garrett, K.A. and Malmstrom, C.M. (2014) Plant-virus interactions and the agro-ecological interface. *Eur. J. Plant Pathol.*, **138**, 529–547.

Alizon, S., Hurford, A., Mideo, N. and Baalen, M. Van (2009) Virulence evolution and the trade-off hypothesis: History, current state of affairs and the future. *J. Evol. Biol.*, **22**, 245–259.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R. and Daszak, P. (2004) Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol. Evol.*, **19**, 535–544.

Antignus O.; Pearlsman, M.; Maslenin, L.; Rosner, A., Y.. L. (2008) A New Pathotype of Pepper mild mottle virus (PMMoV) Overcomes the L4 Resistance Genotype of Pepper Cultivars. *Plant Dis.*, **92**, 1033–1037.

Aranda, M.A. and Freitas-Astúa, J. (2017) Ecology and diversity of plant viruses, and epidemiology of plant virus-induced diseases. *Ann. Appl. Biol.*, **171**, 1–4.

Asplen, M., Bruns, E., David, A. and Al, E. (2012) Do trade-offs have explanatory power for the evolution of organismal interactions? *Evolution.*, **66**, 1297–307.

Avgelis, A.D. (1986) A pepper strain of TMV who is new in Crete (Greece). *Phytopathol. Mediterr.*, **25**, 33–38.

Avina-Padilla, K., Ochoa-Sanchez, J.C. and Martinez-Soriano, J.P. (2008) Nicotiana glauca L. Arvense is a reservoir of plant pathogen viruses. (Nicotiana glauca L. Arvense es Reservorio de Virus Fitopatógenos). *Rev. Mex. Fitopatol.*, **26**, 188–190.

Beachy, R.N. (1976) In Vitro and in Vivo translation of the ribonucleic acids of a Cowpea Strain of Tobacco Mosaic Virus . *Virology*, **517**, 498–517.

Bedhomme, S., Hillung, J. and Elena, S.F. (2015) Emerging viruses : why they are not jacks of all trades ? *Curr. Opin. Virol.*, **10**, 1–6.

Bedhomme, S., Lafforgue, G. and Elena, S.F. (2012) Multihost Experimental Evolution of a Plant RNA Virus Reveals Local Adaptation and Host-Specific Mutations. *Mol. Biol. Evol.*, **29**, 1481–1492.

Bera, S., Moreno-Pérez, M.G., García-figuera, S., Pagán, I., Fraile, A., Pacios, L.F. and García-Arenal, F. (2017) Pleiotropic effects of resistance-breaking mutations on particle stability provide insight on life history evolution in a plant RNA virus. *J. Virol.*, **91,** 1-17.

Berzal-Herranz, A., la Cruz, A. de, Tenllado, F., Díaz-Ruíz, J.R., López, L., Sanz, a I., Vaquero, C., Serra, M.T. and García-Luque, I. (1995) The Capsicum L3 gene-mediated resistance against the tobamoviruses is elicited by the coat protein. *Virology*, **209**, 498–505.

Betancourt, M., Fraile, A. and Garcia-Arenal, F. (2011) Cucumber mosaic virus satellite RNAs that

induce similar symptoms in melon plants show large differences in fitness. *J. Gen. Virol.*, **92**, 1930–1938.

Bloomer, A.C. and Butler, P.J.G. (1986) Tobacco Mosaic Virus Structure and Self-Assembly. In M. H. . Van Regenmortel and H. Fraenkel-Conrat, eds. *The Plant Viruses*. Boston, MA: Springer US, pp. 19–57.

Bodaghi, S., Mathews, D.M. and Dodds, J.A. (2004) Natural Incidence of Mixed Infections and Experimental Cross Protection Between Two Genotypes of Tobacco mild green mosaic virus. *Phytopathology*, **94**, 1337–41.

Bodaghi, S., Yassi, M.N.A. and Dodds, J.A. (2000) Heterogeneity in the 3'-terminal untranslated region of tobacco mild green mosaic tobamoviruses from Nicotiana glauca resulting in variants with three or six pseudoknots. *J. Gen. Virol.*, **81**, 577–586.

Bonhoeffer, S., Lenski, R.E. and Ebert, D. (1996) The Curse of the Pharaoh: The Evolution of Virulence in Pathogens with Long Living Propagules. *Proc. R. Soc. B Biol. Sci.*, **263**, 715–721.

Bosch, F. V. D, Akudibilah, G. Seal, S. and Jeger, M. (2006) Host resistance and the evolutionary response of plant viruses. *J. Appl. Ecol.*, **43**, 506–516.

Boukema, I.W. (1980) Allelism of genes controlling resistance to TMV in Capsicum L. *Euphytica*, **29**, 433–439.

Brandon Ogbunugafor, C., Alto, B.W., Overton, T.M., Bhushan, A., Morales, N.M. and Turner, P.E. (2013) Evolution of Increased Survival in RNA Viruses Specialized on Cancer-Derived Cells. *Am. Nat.*, **181**, 585–595.

Brown, J.D., Goekjian, G., Poulson, R., Valeika, S. and Stallknecht, D.E. (2009) Avian influenza virus in water: Infectivity is dependent on pH, salinity and temperature. *Vet. Microbiol.*, **136**, 20–26.

Brown, J.K.M. (2015) Durable Resistance of Crops to Disease: A Darwinian Perspective. *Annu. Rev. Phytopathol.*, **53**, 513–539.

Brown, J.K.M. and Tellier, A. (2011) Plant-Parasite Coevolution: Bridging the Gap between Genetics and Ecology. *Annu. Rev. Phytopathol.*, **49**, 345–367.

Brunt, A.A., Crabtree, K., Dallwitz, M.J., Gibbs, A.J. and Watson, L. (1996) *Viruses of plants. Descriptions and lists from the VIDE database.*, Wallingford, Oxfordshire, UK: CABI Publishing.

Bull, J.. (1994) Virulence. *Evolution.*, **48**, 1423–1437.

Burdon, J.J. and Thrall, P.H. (2008) Pathogen evolution across the agro-ecological interface : implications for disease management. *Evol. Appl.*, **1**, 57–65.

Cabanillas, L., Arribas, M. and Lázaro, E. (2013) Evolution at increased error rate leads to the coexistence of multiple adaptive pathways in an RNA virus. *BMC Evol. Biol.*, **13**, 11.

Calvo, M., Malinowski, T. and García, J.A. (2014) Single Amino Acid Changes in the 6K1-CI Region Can Promote the Alternative Adaptation of *Prunus* - and *Nicotiana* -Propagated *Plum pox virus* C Isolates to Either Host. *Mol. Plant-Microbe Interact.*, **27**, 136–149.

Carrasco, P., la Iglesia, F. de and Elena, S.F. (2007) Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco Etch virus. *J. Virol.*, **81**, 12979–84.

Caspar, D.L.D. (1964) Assembly and Stability of the Tobacco Mosaic Virus Particle. *Adv. Protein*

Chem., **18**, 37–121.

Cervera, H., Lalić, J. and Elena, S.F. (2016) Effect of host species on topography of the fitness landscape for a plant RNA virus. *J. Virol.*, **90**, 10160–10169.

Coffey, L.L. and Vignuzzi, M. (2011) Host alternation of Chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. *J. Virol.*, **85**, 1025–1035.

Cooper, I. and Jones, R.A.C. (2006) Wild Plants and Viruses: Under-investigated ecosystems. *Adv. Virus Res.*, **67**, 1–47.

Cooper, L.A. and Scott, T.W. (2001) Differential evolution of eastern Equine Encephalitis Virus populations in response to host cell type. *Genetics*, **157**, 1403–1412.

Cronin, J.P., Welsh, M.E., Dekkers, M.G., Abercrombie, S.T. and Mitchell, C.E. (2010) Host physiological phenotype explains pathogen reservoir potential. *Ecol. Lett.*, **13**, 1221–1232.

Culver, J.N. (2002) Tobacco mosaic virus assembly and disassembly: determinants in pathogenicity and resistance. *Annu. Rev. Phytopathol.*, **40**, 287–308.

Culver, J.N., Dawson, W.O., Plonk, K. and Stubbs, G. (1995) Site-directed mutagenesis confirms the involvement of carboxylate groups in the disassembly of tobacco mosaic virus. *Virology*, **206**, 724–730.

Culver, J.N., Stubbs, G. and Dawson, W.O. (1994) Structure-function Relationship Between Tobacco Mosaic Virus Coat Protein and Hypersensitivity in Nicotiana sylvestris. *J. Mol. Biol.*, **242**, 130–138.

Dessau, M., Goldhill, D., McBride, R.L., Turner, P.E. and Modis, Y. (2012) Selective Pressure Causes an RNA Virus to Trade Reproductive Fitness for Increased Structural and Thermal Stability of a Viral Enzyme. *PLoS Genet.*, **8**. e1003102.

Ding, X.S., Liu, J., Cheng, N.-H., Folimonov, A., Hou, Y.-M., Bao, Y., Katagi, C., Carter, S.A. and Nelson, R.S. (2004) The *Tobacco mosaic virus* 126-kDa Protein Associated with Virus Replication and Movement Suppresses RNA Silencing. *Mol. Plant-Microbe Interact.*, **17**, 583–592.

Doumayrou, J., Leblaye, S., Froissart, R. and Michalakis, Y. (2013) Reduction of leaf area and symptom severity as proxies of disease-induced plant mortality: The example of the Cauliflower mosaic virus infecting two Brassicaceae hosts. *Virus Res.*, **176**, 91–100.

Durham, A.C., Finch, J.T. and Klug, A. (1971) States of aggregation of tobacco mosaic virus protein. *Nat. New Biol.*, **229**, 37–42.

Dybdahl, M.F., Jenkins, C.E. and Nuismer, S.L. (2014) Identifying the Molecular Basis of Host-Parasite Coevolution: Merging Models and Mechanisms. *Am. Nat.*, **184**, 1–13.

Ebert, D. (1998) Experimental evolution of parasites. *Science.*, **282**, 1432–5.

Edgar, R.C. (2004) MUSCLE : multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.

Elena, S. (2011) Evolutionary constraints on emergence of plant RNA viruses. In C. Caranta, M. A. Aranda, M. Tepfer, & J. J. Lopez-Moya (Eds.), Recent advances in plant virology (pp. 283–300). United Kingdom: Caister Academic Press.

Elena, S. and Lenski, R. (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.*, **4**, 457–69.

Elena, S. and Sanjuan, R. (2003) Climb every mountain? *Science*., **302**, 2074–5.

Elena, S.F. (2017) Local adaptation of plant viruses: lessons from experimental evolution. *Mol. Ecol.*, **26**, 1711–1719.

Elena, S.F., Agudelo-romero, P. and Lali, J. (2009) The Evolution of Viruses in Multi-Host Fitness Landscapes. *Open Virol. J.*, **3**, 1–6.

Elena, S.F., Bedhomme, S., Carrasco, P., et al. (2011) The evolutionary genetics of emerging plant RNA viruses. *Mol. Plant. Microbe. Interact.*, **24**, 287–293.

Elena, S.F., Fraile, A. and García-Arenal, F. (2014) Evolution and emergence of plant viruses. *Adv. Virus Res.*, **88**, 161-191

Elena, S.F. and Sanjuán, R. (2007) Virus Evolution: Insights from an Experimental Approach. *Annu. Rev. Ecol. Evol. Syst.*, **38**, 27–52.

Escriu, F., Fraile, A. and García-Arenal, F. (2003) The evolution of virulence in a plant virus. *Evolution.*, **57**, 755–65.

Fabre, F., Rousseau, E., Mailleret, L. and Moury, B. (2012) Durable strategies to deploy plant resistance in agricultural landscapes. *New Phytol.*, **193**, 1064–1075.

Fabricante, J.R., Castro, R.A. de, Araújo, K.C.T. de and Siqueira-Filho, J.A. de (2015) Ecological attributes of alien Nicotiana glauca Graham (Solanaceae) and assessment of the susceptibility of the species occuring in Brazil. *Ciência Florest.*, **25**, 959–967.

Fraenkel-Conrat, H. (1986) Tobacco Mosaic Virus The History of Tobacco Mosaic Virus and the Evolution of Molecular Biology. In M. H. . Van Regenmortel and H. Fraenkel-Conrat, eds. *The Plant Viruses*. Boston, MA: Springer US, pp. 5–17.

Fraile, A., Escriu, F., Aranda, M.A., Malpica, J.M., Gibbs, A.J. and Garcia-Arenal, F. (1997) A century of tobamovirus evolution in an Australian population of Nicotiana glauca. *J. Virol.*, **71**, 8316–8320.

Fraile, A. and García-Arenal, F. (2010) The coevolution of plants and viruses: resistance and pathogenicity. *Adv. Virus Res.*, **76**, 1–32.

Fraile, A., Hily, J.M., Pagan, I., Pacios, L.F., Garcia-Arenal, F., Pagán, I., Pacios, L.F. and García-Arenal, F. (2014) Host resistance selects for traits unrelated to resistance-breaking that affect fitness in a plant virus. *Mol. Biol. Evol.*, **31**, 928–939.

Fraile, A., Malpica, J.M., Aranda, M.A., Rodríguez-Cerezo, E. and García-Arenal, F. (1996) Genetic diversity in tobacco mild green mosaic tobamovirus infecting the wild plant Nicotiana glauca. *Virology*, **223**, 148–55.

Fraile, A., McLeish, M.J., Pagán, I., González-Jara, P., Piñero, D. and García-Arenal, F. (2017) Environmental heterogeneity and the evolution of plant-virus interactions: Viruses in wild pepper populations. *Virus Res.*, **241**, 68–76.

Fraile, A., Pagan, I., Anastasio, G., Saez, E. and Garcia-Arenal, F. (2011) Rapid genetic diversification and high fitness penalties associated with pathogenicity evolution in a plant virus. *Mol. Biol. Evol.*, **28**, 1425–1437.

Frank, S. (1996) Models of parasite virulence. *Q. Rev. Biol.*, **71**, 37–78.

Fry, J.D. (1990) Trade-Offs in Fitness on Different Hosts: Evidence from a Selection Experiment with a Phytophagous Mite. *Am. Nat.*, **136**, 569–580.

Gandon, S. (1998) The curse of the pharaoh hypothesis. *Proceedings. Biol. Sci.*, **265**, 1545–52.

Gandon, S., Hochberg, M.E., Holt, R.D. and Day, T. (2013) What limits the evolutionary emergence of pathogens? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **368**, 20120086.

García-Arenal, F. (1988) Sequence and structure at the genome 3' end of the U2-strain of tobacco mosaic virus, a histidine-accepting tobamovirus. *Virology*, **167**, 201–206.

Garcia-Arenal, F., Fraile, A. and Malpica, J.M. (2001) Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.*, **39**, 157–186.

García-Arenal, F. and Fraile, a. (2013) Trade-offs in host range evolution of plant viruses. *Plant Pathol.*, **62**, 2–9.

García-Arenal, F. and McDonald, B.A. (2003) An Analysis of the Durability of Resistance to Plant Viruses. *Phytopathology*, **93**, 941–952.

Genda, Y., Kanda, A., Hamada, H., Sato, K., Ohnishi, J. and Tsuda, S. (2007) Two Amino Acid Substitutions in the Coat Protein of Pepper mild mottle virus Are Responsible for Overcoming the L(4) Gene-Mediated Resistance in Capsicum spp. *Phytopathology*, **97**, 787–793.

Gibson, D.G., Young, L., Chuang, R., Venter, J.C., Iii, C.A.H., Smith, H.O. and America, N. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. , **6**, 12–16.

Gilardi, P., García-Luque, I. and Serra, M.T. (2004) The coat protein of tobamovirus acts as elicitor of both L2 and L4 gene-mediated resistance in Capsicum. *J. Gen. Virol.*, **85**, 2077–2085.

Goldhill, D.H. and Turner, P.E. (2014) The evolution of life history trade-offs in viruses. *Curr. Opin. Virol.*, **8**, 79–84.

Goulden, M.G., Köhm, B.A., Santa Cruz, S., Kavanagh, T.A. and Baulcombe, D.C. (1993) A feature of the coat protein of potato virus X affects both induced virus resistance in potato and viral fitness. *Virology*, **197**, 293–302.

Gray, A., Pelembe, T. and Stroud, S. (2005) The conservation of the endemic vascular flora of Ascension Island and threats from alien species. *Oryx*, **39**, 449-453.

Greene, I.P., Wang, E., Deardorff, E.R., Milleron, R., Domingo, E. and Weaver, S.C. (2005) Effect of alternating passage on adaptation of sindbis virus to vertebrate and invertebrate cells. *J. Virol.*, **79**, 14253–14260.

Greenleaf, W.H., Cook, A.A. and Heyn, A.N.J. (1964) Resistance to Tobacco mosaic virus in Capsicum, with reference to the Samsun latent strain. *Phytopathology*, **54**, 1367–1371.

Gritsun, D.J., Jones, I.M., Gould, E.A. and Gritsun, T.S. (2014) Molecular archaeology of Flaviviridae untranslated regions: duplicated RNA structures in the replication enhancer of flaviviruses and pestiviruses emerged via convergent evolution. *PLoS One*, **9**, e92056.

Hamada, H., Matsumura, H., Tomita, R., Terauchi, R., Suzuki, K. and Kobayashi, K. (2008) SuperSAGE revealed different classes of early resistance response genes in Capsicum chinense plants harboring L3-resistance gene infected with Pepper mild mottle virus. *J. Gen. Plant Pathol.*, **74**, 313–321.

Hamada, H., Takeuchi, S., Kiba, A., Tsuda, S., Hikichi, Y. and Okuno, T. (2002) Amino Acid Changes in Pepper mild mottle virus Coat Protein That Affect L 3 Gene-mediated Resistance in Pepper. *J. Gen. Plant Pathol.*, **68**, 155–162.

Hamada, H., Tomita, R., Iwadate, Y., Kobayashi, K., Munemura, I., Takeuchi, S., Hikichi, Y. and Suzuki, K. (2007) Cooperative effect of two amino acid mutations in the coat protein of Pepper mild mottle virus overcomes L 3 -mediated resistance in Capsicum plants. *Virus Genes*, **34**, 205–214.

Hanada, K. and Harrison, B.D. (1977) Effects of virus genotype and temperature on seed transmission of nepoviruses. *Ann. Appl. Biol.*, **85**, 79–92.

Handel, A., Brown, J., Stallknecht, D. and Rohani, P. (2013) A Multi-scale Analysis of Influenza A Virus Fitness Trade-offs due to Temperature-dependent Virus Persistence C. Fraser, ed. *PLoS Comput. Biol.*, **9**, e1002989.

Handel, A., Lebarbenchon, C., Stallknecht, D. and Rohani, P. (2014) Trade-offs between and within scales: environmental persistence and within-host fitness of avian influenza viruses. *Proceedings. Biol. Sci.*, **281**, 20133051

Hanssen, I.M., Lapidot, M. and Thomma, B.P.H.J. (2010) Emerging Viral Diseases of Tomato Crops. *Mol. Plant Microbe Interact.*, **23**, 539–548.

Harrison, B.D. (2002) Virus variation in relation to resistance-breaking in plants. *Euphytica*, **124**, 181–192.

Hartmann, F.E., Sánchez-vallet, A., Mcdonald, B.A. and Croll, D. (2017) A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *Int. Soc. Microb. Ecol.*, 1–16.

Heineman, R.H. and Brown, S.P. (2012) Experimental Evolution of a Bacteriophage Virus Reveals the Trajectory of Adaptation across a Fecundity/Longevity Trade-Off. *PLoS One*, **7**, e46322.

Hillung, J., Cuevas, J.M. and Elena, S.F. (2015) Evaluating the within-host fitness effects of mutations fixed during virus adaptation to different ecotypes of a new host. *Philos. Trans. R. Soc. B Biol. Sci.*, **370**, 20140292

Hillung, J., Cuevas, J.M., Valverde, S. and Elena, S.F. (2014) Experimental evolution of an emerging plant virus in host genotypes that differ in their susceptibility to infection. *Evolution.*, **68**, 2467–2480.

Hogenhout, S.A., Ammar, E.-D., Whitfield, A.E. and Redinbaugh, M.G. (2008) Insect Vector Interactions with Persistently Transmitted Viruses. *Annu. Rev. Phytopathol.*, **46**, 327–359.

Holmes, F.O. (1951) Indications of a New-World origin of tobaccomosaic virus. *Phytopathology*, **41**, 341–349.

Huang, X. and Madan, A. (1999) CAP3 : A DNA Sequence Assembly Program. , 868–877.

Ishibashi, K., Mawatari, N., Miyashita, S., Kishino, H., Meshi, T. and Ishikawa, M. (2012) Coevolution and Hierarchical Interactions of Tomato mosaic virus and the Resistance Gene Tm-1 F. Garcia-Arenal, ed. *PLoS Pathog.*, **8**, e1002975.

Janzac, B., Fabre, F., Palloix, A. and Moury, B. (2009) Constraints on evolution of virus avirulence factors predict the durability of corresponding plant resistances. *Mol. Plant Pathol.*, **10**, 599–610.

Janzac, B., Montarry, J., Palloix, A., Navaud, O. and Moury, B. (2010) A Point Mutation in the Polymerase of *Potato virus Y* Confers Virulence Toward the *Pvr4* Resistance of Pepper and a High Competitiveness Cost in Susceptible Cultivar. *Mol. Plant-Microbe Interact.*, **23**, 823–830.

Jeger, M., Seal, S. and Bosch, F. Van den (2006) Evolutionary epidemiology of plant virus disease. *Adv. Virus Res.*, **67**, 163–203.

Jeger, M.J., Seal, S.E. and Bosch, F. Van Den (2006) Evolutionary epidemiology of plant virus disease. *Adv. Virus Res.*, **67**, 163–203.

Jenner, C.E., Tomimura, K., Ohshima, K., Hughes, S.L. and Walsh, J.A. (2002) Mutations in Turnip mosaic virus P3 and cylindrical inclusion proteins are separately required to overcome two Brassica napus resistance genes. *Virology*, **300**, 50–9.

Jenner, C.E., Wang, X., Ponz, F. and Walsh, J.A. (2002) A fitness cost for Turnip mosaic virus to overcome host resistance. *Virus Res.*, **86**, 1–6.

Jones, D.R. (2001) *Summary pest risk assessment of Tomato infectious chlorosis virus*,.

Jones, J.D.G. and Dangl, J.L. (2006) The plant immune system. *Nature*, **444**, 323–329.

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L. and Daszak, P. (2008) Global trends in emerging infectious diseases. *Nature*, **451**, 990–993.

Jones, R.A.C. (2009) Plant virus emergence and evolution: Origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Res.*, **141**, 113–130.

Kaltz, O., Gandon, S., Michalakis, Y. and Shykoff, J.A. (1999) Local Maladaptation in the Anther-Smut Fungus Microbotryum violaceum to Its Host Plant Silene latifolia: Evidence from a Cross-Inoculation Experiment. *Evolution.*, **53**, 395–407.

Kawecki, T.J. (1994) Accumulation of Deleterious Mutations and the Evolutionary Cost of Being a Generalist. *Am. Nat.*, **144**, 833–838.

Keen, E.C. (2014) Tradeoffs in bacteriophage life histories. *Bacteriophage*, **4**, e28365.

Kenyon, L., Kumar, S., Tsai, W. and Hughes, J.A. (2014) Virus Diseases of Peppers (Capsicum spp .) and Their Control. *Adv. Virus Res.*, **90**, 297–354.

Kirchner, J. and Roy, B. (2002) Evolutionary implications of host–pathogen specificity: fitness consequences of pathogen virulence traits. *Evol. Ecol. Res.*, **4**, 27–48.

Klug, A. (1999) The tobacco mosaic virus particle: structure and assembly. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **354**, 531–5.

Kouyos, R.D., Silander, O.K. and Bonhoeffer, S. (2007) Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol.*, **22**.

Kubota, K., Tsuda, S., Tamai, A. and Meshi, T. (2003) Tomato mosaic virus replication protein suppresses virus-targeted posttranscriptional gene silencing. *J. Virol.*, **77**, 11016–26.

la Cruz, A. de, López, L., Tenllado, F., Díaz-Ruíz, J.R., Sanz, A.I., Vaquero, C., Serra, M.T. and García-Luque, I. (1997) The Coat Protein Is Required for the Elicitation of the *Capsicum* L ^{2} Gene-Mediated Resistance Against the Tobamoviruses. *Mol. Plant-Microbe Interact.*, **10**, 107–113.

Lajeunesse, M. and Forbes, M. (2002) Host range and local parasite adaptation. *Proc. R. Soc.*

London B, **269**, 703–710.

Lalic, J., Cuevas, J.M. and Elena, S.F. (2011) Effect of Host Species on the Distribution of Mutational Fitness Effects for an RNA Virus. *PLoS Genet.*, **7**, e1002378.

Lalic, J. and Elena, S.F. (2012) Epistasis between mutations is host-dependent for an RNA virus. *Biol. Lett.*, **9**, 20120396–20120396.

Lalić, J. and Elena, S.F. (2012) Magnitude and sign epistasis among deleterious mutations in a positive-sense plant RNA virus. *Heredity.*, **109**, 71–77.

Lalić, J. and Elena, S.F. (2015) The impact of high-order epistasis in the within-host fitness of a positive-sense plant RNA virus. *J. Evol. Biol.*, **28**, 2236–2247.

Liang, X.Z., Lee, B.T.K. and Wong, S.M. (2002) Covariation in the capsid protein of hibiscus chlorotic ringspot virus induced by serial passaging in a host that restricts movement leads to avirulence in its systemic host. *J. Virol.*, **76**, 12320–12324.

Liu, W. Y. Kang, W.H. and Kang, B.C. (2013). Basic Information on Pepper. Genetics, Genomics and Breeding of Peppers and Eggplants, 1.

Lu, B., Stubbs, G. and Culver, J.N. (1996) Carboxylate Interactions Involved in the Disassembly of Tobacco Mosaic Tobamovirus. *Virology*, **225**, 11–20.

Malpica, M., Sacrista, S., Fraile, A. and Garcia-Arenal, F. (2006) Association and Host Selectivity in Multi-Host Pathogens. *PLoS One*, **e41**.

Mandeles, S. (1968) Location of Unique Sequences in Tobacco Mosaic Virus Ribonucleic Acid. *J. Biol. Chem.*, **243**, 3671–3674.

Matsumoto, K., Sawada, H., Matsumoto, K., et al. (2008) The coat protein gene of tobamovirus P0 pathotype is a determinant for activation of temperature-insensitive L1a-gene- mediated resistance in Capsicum plants. *Arch. Virol.*, **153**, 645–650.

McDonald, B.A. (2004) Population Genetics of Plant Pathogens. *Plant Heal. Instr.*

McDonald, B. and Linde, C. (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.*, **40**, 349–79.

McKinney, H.H. (1929) Mosaic diseases in the Canary Islands, West Africa and Gibraltar. *J. Agric. Res.*, **39**, 577–578.

McKinney, H.H. (1952) Two strains of Tobacco-mosaic virus, one of which is seed-borne in an etch-immune pungent Pepper. *Plant Dis. Report.*, **36**, 184–187.

McLeish, M., Sacristán, S., Fraile, A. and Garcia-Arenal, F. (2017) Scale dependencies and generalism in host use shape virus prevalence. *Proc. R. Soc. B Biol. Sci.,* **284**, 20172066.

Mestre, P., Brigneti, G., Durrant, M.C. and Baulcombe, D.C. (2003) Potato virus Y NIa protease activity is not sufficient for elicitation of Ry-mediated disease resistance in potato. *Plant J.* , **36**, 755–61.

Montarry, J., Doumayrou, J., Simon, V. and Moury, B. (2011) Genetic background matters: A plant-virus gene-for-gene interaction is strongly influenced by genetic contexts. *Mol. Plant Pathol.*, **12**, 911–920.

Moreno-Pérez, M.G., García-Luque, I., Fraile, A. and García-Arenal, F. (2016) Mutations That Determine Resistance Breaking in a Plant RNA Virus Have Pleiotropic Effects on Its Fitness

That Depend on the Host Environment and on the Type, Single or Mixed, of Infection. *J. Virol.*, **90**, 9128–9137.

Moreno-Pérez, M.G., Pagán, I., Aragón-Caballero, L., Cáceres, F., Fraile, A. and García-Arenal, F. (2014) Ecological and genetic determinants of Pepino Mosaic Virus emergence. *J. Virol.*, **88**, 3359–68.

Morishima, N., Ido, T., Hamada, H., Yoshimoto, E., Mizumoto, H., Takeuchi, S., Kiba, A., Hikichi, Y. and Okuno, T. (2003) Infectious in vitro transcripts from a cDNA clone of Tobacco mild green mosaic tobamovirus and its biological activity in host and nonhost plants and in their protoplasts. *J. Gen. Plant Pathol.*, **69**, 335–338.

Moury, B., Desbiez, C., Jacquemond, M. and Lecoq, H. (2006) Genetic diversity of plant virus populations: towards hypothesis testing in molecular epidemiology. *Adv. Virus Res.*, **67**, 49–87.

Moury, B., Fabre, F., Hebrard, E. and Froissart, R. (2017) Determinants of host species range in plant viruses. *J. Gen. Virol.*, **98**, 862–873.

Moury, B. and Simon, V. (2011) dN/dS-Based Methods Detect Positive Selection Linked to Trade-Offs between Different Fitness Traits in the Coat Protein of Potato virus Y. *Mol. Biol. Evol.*, **28**, 2707–2717.

Moury, B. and Verdin, E. (2012) *Viruses of Pepper Crops in the Mediterranean Basin: A Remarkable Stasis*, p 127–162. In Loebenstein G, Lecoq H (ed), Viruses and virus diseases of vegetables in the Mediterranean basin. Aca- demic Press, San Diego, CA.

Moya, A., Rodríguez-Cerezo, E. and García-Arenal, F. (1993) Genetic Structure of Natural Populations of the Plant RNA Virus Tobacco Mild Green Mosaic Virus. *Mol. Biol. Evol.*, **10**, 449–456.

Murant, A.F., Taylor, C.E. and Chambers, J. (1968) Properties, relationships and transmission of a strain of raspberry ringspot virus infecting raspberry cultivars immune to the common Scottish strain*. *Ann. Appl. Biol.*, **61**, 175–186.

Namba, K. and Stubbs, G. (1986) Structure of tobacco mosaic virus at 3.6 A resolution: implications for assembly. *Science (80-.).*, **231**, 1401–1406.

Nguyen, H.D., Tomitaka, Y., Ho, S.Y.W., Duchêne, S., Vetten, H.-J., Lesemann, D., Walsh, J.A., Gibbs, A.J. and Ohshima, K. (2013) Turnip Mosaic Potyvirus Probably First Spread to Eurasian Brassica Crops from Wild Orchids about 1000 Years Ago A. Rao, ed. *PLoS One*, **8**, e55336.

Novella, I.S., Hershey, C.L., Escarmis, C., Domingo, E. and Holland, J.J. (1999) Lack of evolutionary stasis during alternating replication of an arbovirus in insect and mammalian cells. *J. Mol. Biol.*, **287**, 459–465.

Oerke, E. and Dehne, H. (2004) Safeguarding production — losses in major crops and the role of crop protection. *Crop Prot.*, **23**, 275–285.

Ohshima, K., Yamaguchi, Y., Hirota, R., et al. (2002) Molecular evolution of Turnip mosaic virus : evidence of host adaptation, genetic recombination and geographical spread. *J. Gen. Virol.*, **83**, 1511–1521.

Ollerton, J., Watts, S., Connerty, S., et al. (2012) Pollination Ecology of the Invasive Tree Tobacco

Nicotiana Glauca: Comparisons Across Native and Non-Native Ranges. *J. Pollinat. Ecol.*, **9**, 85–95.

Paepe, M. De and Taddei, F. (2006) Viruses' Life History: Towards a Mechanistic Basis of a Trade-Off between Survival and Reproduction among Phages T. Kirkwood, ed. *PLoS Biol.*, **4**, e193.

Pagán, I., Alonso-blanco, C. and Garcia-Arenal, F. (2007) The Relationship of Within-Host Multiplication and Virulence in a Plant-Virus System. *PLoS One*., **2**, e786.

Poelwijk, F.J., Tănase-Nicola, S., Kiviet, D.J. and Tans, S.J. (2011) Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol.*, **272**, 141–144.

Posada, D. (2008) jModelTest : Phylogenetic Model Averaging. *Mol. Biol. Evol.*, **25**, 1253–1256.

Poulicard, N., Pinel-Galzi, A., Traoré, O., Vignols, F., Ghesquière, A., Konaté, G., Hébrard, E. and Fargette, D. (2012) Historical contingencies modulate the adaptability of rice yellow mottle virus. *PLoS Pathog.*, **8**. e1002482.

Power, A. and Flecker, A. (2003) Virus specificity in disease systems: are species redundant? In P. Kareiva and S. Levin, eds. *The Importance of Species*. Princeton: Princeton University Pres, pp. 330–47.

Power, A.G., Borer, E.T., Hosseini, P., Mitchell, C.E. and Seabloom, E.W. (2011) The community ecology of barley/cereal yellow dwarf viruses in Western US grasslands. *Virus Res.*, **159**, 95–100.

Power, A.G. and Mitchell, C.E. (2004) Pathogen Spillover in Disease Epidemics. *Am. Nat.*, **164**, S79–S89.

Rast, A. T. B. (1988). Pepper tobamoviruses and pathotypes used in resistance breeding (invited paper). Capsicum News Letter, 1(1), 20-23.

Reingold, V., Lachman, O., Belausov, E., Koren, A., Mor, N. and Dombrovsky, A. (2016) Epidemiological study of Cucumber green mottle mosaic virus in greenhouses enables reduction of disease damage in cucurbit production. *Ann. Appl. Biol.*, **168**, 29–40.

Remold, S. (2012) Understanding specialism when the jack of all trades can be the master of all. *Proc. R. Soc. B Biol. Sci.*, **279**, 4861–4869.

Rico, P., Ivars, P., Elena, S.F. and Hernandez, C. (2006) Insights into the Selective Pressures Restricting Pelargonium Flower Break Virus Genome Variability: Evidence for Host Adaptation. *J. Virol.*, **80**, 8124–8132.

Roche, B., Drake, J.M., Brown, J., Stallknecht, D.E., Bedford, T. and Rohani, P. (2014) Adaptive Evolution and Environmental Durability Jointly Structure Phylodynamic Patterns in Avian Influenza Viruses. *PLoS Biol.*, **12**.

Roche, B., Drake, J.M. and Rohani, P. (2011) The curse of the Pharaoh revisited: evolutionary bi-stability in environmentally transmitted pathogens. *Ecol. Lett.*, **14**, 569–575.

Roossinck, M.J. and García-Arenal, F. (2015) Ecosystem simplification, biodiversity loss and plant virus emergence. *Curr. Opin. Virol.*, **10**, 56–62.

Roossinck, M.J., Saha, P., Wiley, G.B., Quan, J., White, J.D., Lai, H., Chavarria, F., Shen, G. and Roe, B.A. (2010) Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Mol. Ecol.*, **19**, 81–88.

Sacristan, S., Fraile, A., Malpica, J.M. and Garcia-Arenal, F. (2005) An analysis of host adaptation and its relationship with virulence in Cucumber mosaic virus. *Phytopathology*, **95**, 827–833.

Sacristán, S. and García-Arenal, F. (2008) The evolution of virulence and pathogenicity in plant pathogen populations. *Mol. Plant Pathol.*, **9**, 369–384.

Sastry, K.S. and Zitter, T.A. (2014) *Management of Virus and Viroid Diseases of Crops in the Tropics*,.

Siegel, A. and Wildman, G.S. (1954) Some natural relationships among strains of tobacco mosaic virus. *Phytopathology*, **44**, 277–282.

Silva, J. Da, Coetzer, M., Nedellec, R., Pastore, C. and Mosier, D.E. (2010) Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. *Genetics*, **185**, 293–303.

Sokal, R.R. and Rohlf, F.J. (1994) Biometry: the principles and practice of statistics in biological research.

Solis, I. and Garcia-Arenal, F. (1990) The complete nucleotide sequence of the genomic RNA of the tobamovirus tobacco mild green mosaic virus. *Virology*, **177**, 553–558.

Sorho, F., Pinel, A., Traoré, O., Bersoult, A., Ghesquière, A., Hébrard, E., Konaté, G., Séré, Y. and Fargette, D. (2005) Durability of natural and transgenic resistances in rice to Rice yellow mottle virus. *Eur. J. Plant Pathol.*, **112**, 349–359.

Stubbs, G. (1999) Tobacco mosiac virus particle structure and the initiation of disassembly. *Philos. Trans. R. Soc. B Biol. Sci.*, **354**, 551–557.

Thresh, J.M. (2006) Plant Virus Epidemiology: The Concept of Host Genetic Vulnerability. In *Advances in virus research*. pp. 89–125.

Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R. and Polasky, S. (2002) Agricultural sustainability and intensive production practices. *Nature*, **418**.

Tomimura, K., Gibbs, A.J., Jenner, C.E., Walsh, J.A. and Ohshima, K. (2003) The phylogeny of Turnip mosaic virus; comparisons of 38 genomic sequences reveal a Eurasian origin and a recent “emergence” in east Asia. *Mol. Ecol.*, **12**, 2099–111.

Tomita, R., Sekine, K.-T., Mizumoto, H., Sakamoto, M., Murai, J., Kiba, A., Hikichi, Y., Suzuki, K. and Kobayashi, K. (2011) Genetic basis for the hierarchical interaction between Tobamovirus spp. and L resistance gene alleles from different pepper species. *Mol. Plant. Microbe. Interact.*, **24**, 108–117.

Tomlinson, K.R., Bailey, A.M., Alicai, T., Seal, S. and Foster, G.D. (2017) Cassava brown streak disease: historical timeline, current knowledge and future prospects. *Mol. Plant Pathol.*, 1–13.

Tsuda, S., Kirita, M. and Watanabe, Y. (1998) Characterization of a pepper mild mottle tobamovirus strain capable of overcoming the L3 gene-mediated resistance, distinct from the resistance-breaking Italian isolate. *Mol. Plant. Microbe. Interact.*, **11**, 327–331.

Vassilakos, N., Simon, V., Tzima, A. and Johansen, E. (2017) Genetic Determinism and Evolutionary Reconstruction of a Host Jump in a Plant Virus. *Mol. Biol. Evol.*, **33**, 541–553.

Villordo, S.M., Carballeda, J.M., Filomatori, C. V and Gamarnik, A. V (2016) RNA Structure Duplications and Flavivirus Host Adaptation. *Trends Microbiol.*, **24**, 270–83.

Visser, J.A.G.M. de and Krug, J. (2014) Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.*, **15**, 480–490.

Wallis, C.M., Stone, A.L., Sherman, D.J., Damsteegt, V.D., Gildow, F.E. and Schneider, W.L. (2007) Adaptation of plum pox virus to a herbaceous host (Pisum sativum) following serial passages. *J. Gen. Virol.*, **88**, 2839–2845.

Walther, B.A. and Ewald, P.W. (2004) Pathogen survival in the external environment and the evolution of virulence. *Biol. Rev. Camb. Philos. Soc.*, **79**, 849–69.

Webster, C.G., Coutts, B.A., Jones, R.A.C., Jones, M.G.K. and Wylie, S.J. (2007) Virus impact at the interface of an ancient ecosystem and a recent agroecosystem : studies on three legume-infecting potyviruses in the southwest Australian floristic region. *Plant Pathol.*, **56**, 729–742.

Wetter, C. (1984) Serological identification of four tobamoviruses infecting pepper. *Plant Dis.*, **68**, 597–599.

Wetter, C., Conti, M., Altschuh, D., Tabillion, R. and Regenmortel, M.H.V. Van (1984) Pepper mild mottle virus, a tobamovirus infecting pepper cultivars in Sicily. *Phytopathology*, **74**, 405–410.

Whitlock, M.C. (1996) The Red Queen Beats the Jack-Of-All-Trades: The Limitations on the Evolution of Phenotypic Plasticity and Niche Breadth. *Am. Nat.*, **148**, S65–S77.

Wilson, T.M.A. (1984) Cotranslational disassembly of tobacco mosaic virus in vitro. *Virology*, **137**, 255–265.

Wisler, G.C. and Norris, R.F. (2005) Interactions between Weeds and Cultivated Plants as Related to Management of Plant Pathogens. *Weed Sci.*, **53**, 914–917.

Woolhouse, M.E., Taylor, L.H. and Haydon, D.T. (2001) Population biology of multihost pathogens. *Science*, **292**, 1109–12.

Woolhouse, M.E.J., Haydon, D.T. and Antia, R. (2005) Emerging pathogens: The epidemiology and evolution of species jumps. *Trends Ecol. Evol.*, **20**, 238–244.

Yarwood, C.E. (1979) Host passage effects with plant viruses. *Adv. Virus Res.*, **25**, 169–90.

Zeidan, M., Beckelman, E. and Holdengreber, P. (2008) Detection of Tobacco mild green mosaic virus in ornamental crops. In: Proceedings of the 12th International Symposium on Virus Diseases of Ornamental Plants,.

8 ANNEXES

ANNEX I:

The nucleotide sequence of the coat protein gene and 3' untranslated region of 26 field isolates of *Tobacco mild green mosaic virus*

A.I.1 Alignment of the nucleotide sequence of the coat protein gene to that of U2-TMGMV (accession no. AB078435)

```
           1                                                                                                   60
TMV-U2     ATGCCTTATACAATCAACTCTCCGAGCCAATTTGTTTACTTAAGTTCCGCTTATGCAGAT
P 00/10    ............................................................
P 00/6     ............................................................
P 04/17    ..........................................A.................
P 96/49    ............................................................
P 98/11    ............................................................
P 98/12    ............................................................
P 01/16    ............................................................
P 83/4     ............................................................
P 92/10    ............................................................
P 94/29    ............................................................
P 97/10    ............................................................
P 98/5     ............................................................
Ng 89/8    ............................................................
Ng 90/5    ............................................................
Ng 94/6    ............................................................
Ng 96/19   ............................................................
Ng 96/5    ............................................................
Ng 92/73   ..................T.........................................
Ng 96/11   ..................T.....................................AG....
Ng 99/20   ..................T.........................................
Ng 89/15   ..................T.........................................
Ng 90/8    ..................T.........................................
Ng 96/16   ..................T.........................................
Ng 99/11   ..................T.........................................
Ng 99/15   ..................T.........................GGG.AGG.GC
Ng 99/16   ..................T.........................................

           61                                                                                                  120
TMV-U2     CCTGTGCAGCTAATCAATCTGTGCACAAATGCATTGGGTAACCAGTTTCAAACGCAACAA
P 00/10    ..........G.........T.....................  ..............
P 00/6     ..........................................A.................
P 04/17    ..........G.........T.....................  ..............
P 96/49    ............................................................
P 98/11    ..........................................A.................
P 98/12    ..........................................A.................
P 01/16    ..................T.........................................
P 83/4     ............................................................
P 92/10    ............................................................
P 94/29    ............................................................
P 97/10    ............................................................
P 98/5     ............................................................
Ng 89/8    ............................................................
Ng 90/5    ............................................................
Ng 94/6    ..........................................C.................
Ng 96/19   ..........................................C.................
Ng 96/5    ............................................................
Ng 92/73   ............................................................
Ng 96/11   ..AA.....................T.....................................
Ng 99/20   ............................................................
Ng 89/15   ............................................................
Ng 90/8    ............................................................
Ng 96/16   ............................................................
Ng 99/11   ............................................................
Ng 99/15   ..A.........................................................
Ng 99/16   ............................................................
```

```
          121                                                          180
TMV-U2    GCTAGGACAACAGTCCAACAGCAATTTGCGGATGCCTGGAAACCTGTGCCTAGTATGACA
P 00/10   ............................................................
P 00/6    ...........................................................G
P 04/17   ..C.........................................................
P 96/49   .......................................................C....
P 98/11   ...........................................................G
P 98/12   ...........................................................G
P 01/16   ..................................T.........................
P 83/4    ............................................................
P 92/10   ............................................................
P 94/29   ........G...................................................
P 97/10   ........G...................................................
P 98/5    ............................................................
Ng 89/8   ............................................................
Ng 90/5   ............................................................
Ng 94/6   ............................................................
Ng 96/19  ............................................................
Ng 96/5   ............................................................
Ng 92/73  ............................................................
Ng 96/11  .....................................................C......
Ng 99/20  ............................................................
Ng 89/15  ............................................................
Ng 90/8   ............................................................
Ng 96/16  ............................................................
Ng 99/11  ............................................................
Ng 99/15  ............................................................
Ng 99/16  ............................................................
```

```
          181                                                          240
TMV-U2    GTGAGATTTCCTGCATCGGATTTCTATGTGTATAGATATAATTCGACGCTTGATCCGTTG
P 00/10   .......................T....................................
P 00/6    ............................................................
P 04/17   .......................T....................................
P 96/49   ............................................................
P 98/11   ............................................................
P 98/12   ............................................................
P 01/16   .......................T....................................
P 83/4    ............................................................
P 92/10   ............................................................
P 94/29   ............................................................
P 97/10   ............................................................
P 98/5    ............................................................
Ng 89/8   ............................................................
Ng 90/5   ....................C.......................................
Ng 94/6   ......................................C.....................
Ng 96/19  ......................................C.....................
Ng 96/5   ................................C...........................
Ng 92/73  ..........................C.................................
Ng 96/11  ..............G...........C.................................
Ng 99/20  ..........................C.................................
Ng 89/15  .............................A..............................
Ng 90/8   ............................................................
Ng 96/16  ............................................................
Ng 99/11  ............................................................
Ng 99/15  ............................................................
Ng 99/16  ............................................................
```

```
          241                                                                                                  300
TMV-U2    ATCACGGCGTTATTAAATAGCTTCGATACTAGAAATAGAATAATAGAGGTTGATAATCAA
P 00/10   . . . . . . . . . . . . . . . . . . . . . . . . T. . . . . . . . . . . . . . . . . . . . . . . . . . . . . C. . .
P 00/6    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 04/17   . . . . . . . . . . . . . . . . . . . . . . . . T. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 96/49   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/11   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/12   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 01/16   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 83/4    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 92/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 94/29   . . . . . . . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . C. . .
P 97/10   . . . . . . . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . C. . .
P 98/5    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 94/6   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/19  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 92/73  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/11  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/20  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/11  . . . . . . . . . . . . . . . G. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/15  . . . . . . . . . . . . . . . G. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

          301                                                                                                  360
TMV-U2    CCCGCACCGAATACTACTGAAATCGTTAACGCGACTCAGAGGGTAGACGATGCGACTGTA
P 00/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 00/6    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 04/17   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 96/49   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/11   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/12   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 01/16   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 83/4    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 92/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 94/29   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 97/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/5    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 94/6   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/19  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 92/73  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/11  . . . . . . . . . A. . . . . . . . . A. . . . . . . . . . . . . . . . . . . . . . . TA. . . . . . . . . . . . . . .
Ng 99/20  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/11  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
```

```
          361                                                         420
TMV-U2    GCTATAAGGGCTTCAATCAATAATTTAGCTAATGAACTGGTTCGTGGAACTGGCATGTTC
P 00/10   ..........................G.................................
P 00/6    ............................................................
P 04/17   ..........................G.................................
P 96/49   ............................................................
P 98/11   ............................................................
P 98/12   ............................................................
P 01/16   ............................................................
P 83/4    ............................................................
P 92/10   ............................................................
P 94/29   ............................................................
P 97/10   ............................................................
P 98/5    ............................................................
Ng 89/8   ............................................................
Ng 90/5   ..........................G.................................
Ng 94/6   ............................................................
Ng 96/19  ............................................................
Ng 96/5   ............................................................
Ng 92/73  ............................................................
Ng 96/11  ............................................................
Ng 99/20  ............................................................
Ng 89/15  ............................................................
Ng 90/8   ............................................................
Ng 96/16  ............................................................
Ng 99/11  ............................................................
Ng 99/15  ............................................................
Ng 99/16  ............................................................
```

```
          421                                                         480
TMV-U2    AATCAAGCAGGCTTTGAGACTGCTAGTGGACTTGTCTGGACCACAACTCCGGCTACTTAG
P 00/10   ............................................................
P 00/6    ............................................................
P 04/17   ..........................................CA......A.........
P 96/49   ............................................................
P 98/11   ............................................................
P 98/12   ............................................................
P 01/16   .....................A......................................
P 83/4    .....................A......................................
P 92/10   .....................A......................................
P 94/29   .....................A......................................
P 97/10   .....................A......................................
P 98/5    .....................A......................................
Ng 89/8   .....................A......................................
Ng 90/5   .....................A......................................
Ng 94/6   .....................A......................................
Ng 96/19  .....................A......................................
Ng 96/5   .....................A......................................
Ng 92/73  .....................A......................................
Ng 96/11  .....................A......................................
Ng 99/20  .....................A......................................
Ng 89/15  ............................................................
Ng 90/8   ............................................................
Ng 96/16  ............................................................
Ng 99/11  ............................................................
Ng 99/15  ............................................................
Ng 99/16  ............................................................
```

A.I.2 Alignment of the nucleotide sequence of the 3’ untranslated region to that of U2-TMGMV (accession no. AB078435)

```
           1                                                           60
TMV-U2     CTATTGTTGTGAGATTTCCTAAAATAAAGTCGCTGAAGACTTTAAATTCAGGGTGGCTGA
P 00/10    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . . . . . . . . . . . . . . .
P 00/6     . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 04/17    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . .
P 96/49    . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/11    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/12    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 01/16    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 83/4     . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 92/10    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 94/29    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . . . . . . . . . . . T. . .
P 97/10    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . . . . . . . . . . . . . . .
P 98/5     . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/8    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/5    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 94/6    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/19   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/5    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 92/73   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/11   . . . . . . . . . . . . A. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/20   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/15   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/8    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/16   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/11   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/15   . C. . . . . . . . . . . . . . . . . T. . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/16   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

           61                                                          120
TMV-U2     TACC- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 00/10    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 00/6     . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
P 04/17    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 96/49    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 98/11    . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
P 98/12    . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
P 01/16    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 83/4     . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 92/10    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 94/29    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 97/10    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
P 98/5     . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 89/8    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 90/5    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 94/6    . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
Ng 96/19   . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 96/5    . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
Ng 92/73   . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
Ng 96/11   . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
Ng 99/20   . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGAATCC
Ng 89/15   . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 90/8    . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 96/16   . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Ng 99/11   . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
Ng 99/15   . . . . AAAATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCC
Ng 99/16   . . . . - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

```
        121                                                         180
TMV-U2  ------------------------------------------------------------
P 00/10 ------------------------------------------------------------
P 00/6  AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
P 04/17 ------------------------------------------------------------
P 96/49 ------------------------------------------------------------
P 98/11 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
P 98/12 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
P 01/16 ------------------------------------------------------------
P 83/4  ------------------------------------------------------------
P 92/10 ------------------------------------------------------------
P 94/29 ------------------------------------------------------------
P 97/10 ------------------------------------------------------------
P 98/5  ------------------------------------------------------------
Ng 89/8 ------------------------------------------------------------
Ng 90/5 ------------------------------------------------------------
Ng 94/6 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
Ng 96/19 ------------------------------------------------------------
Ng 96/5 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGAGAGGTTCG
Ng 92/73 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
Ng 96/11 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
Ng 99/20 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
Ng 89/15 ------------------------------------------------------------
Ng 90/8 ------------------------------------------------------------
Ng 96/16 ------------------------------------------------------------
Ng 99/11 AGCAGTTAAACCATGTGATGGTATATACTGTGGTATGGCGTAAAACAACGGA--------
Ng 99/15 AGCAGTTAAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGA--------
Ng 99/16 ------------------------------------------------------------
```

```
        181                                                         240
TMV-U2  -------------------------------------------------AAAATCAGCAG
P 00/10 -------------------------------------------------........T..
P 00/6  ----------AAAGTCCCTGAACACTTTAAATTCAGGGTGGCTGATAAC...........
P 04/17 -------------------------------------------------........T..
P 96/49 -------------------------------------------------...........
P 98/11 ----------AAAGTCGCTGAAGACTTTAAATTCAGGGTGGCTGATACC...........
P 98/12 ----------AAAGTCCCTGAACACTTTAAATTCAGGGTGGCTGATAAA...........
P 01/16 -------------------------------------------------...........
P 83/4  -------------------------------------------------...........
P 92/10 -------------------------------------------------........T..
P 94/29 -------------------------------------------------...........
P 97/10 -------------------------------------------------...........
P 98/5  -------------------------------------------------...........
Ng 89/8 -------------------------------------------------...........
Ng 90/5 -------------------------------------------------...........
Ng 94/6 ----------AAAGTCCCAAACCACTTTAAATTCCGGGTGGCTGACAAA...........
Ng 96/19 -------------------------------------------------...........
Ng 96/5 AATCCTCCCCAAAGTCGCTGAAGACTTTAAATTCAGGGTGGCTGATACC...........
Ng 92/73 ----------AAAGTCGCTGAAGACTTTAAATTCAGGGTGGCTGATACC...........
Ng 96/11 ----------AAAGTCCCTGACCACTTTAAATTCCGGGTGGCTGATAAA...........
Ng 99/20 ----------AAAGTCCCTAAACACTTTAAATTCCGGGTGGCTGATAAA...........
Ng 89/15 -------------------------------------------------........T..
Ng 90/8 -------------------------------------------------...........
Ng 96/16 -------------------------------------------------...........
Ng 99/11 ----------AAAGTCCCTGAACACTTTAAATTCAGGGTGGCTGATAAA...........
Ng 99/15 ----------AAAGTCCCTAAACACTTTAAATTCCGGGTGGCTGATAAA...........
Ng 99/16 -------------------------------------------------...........
```

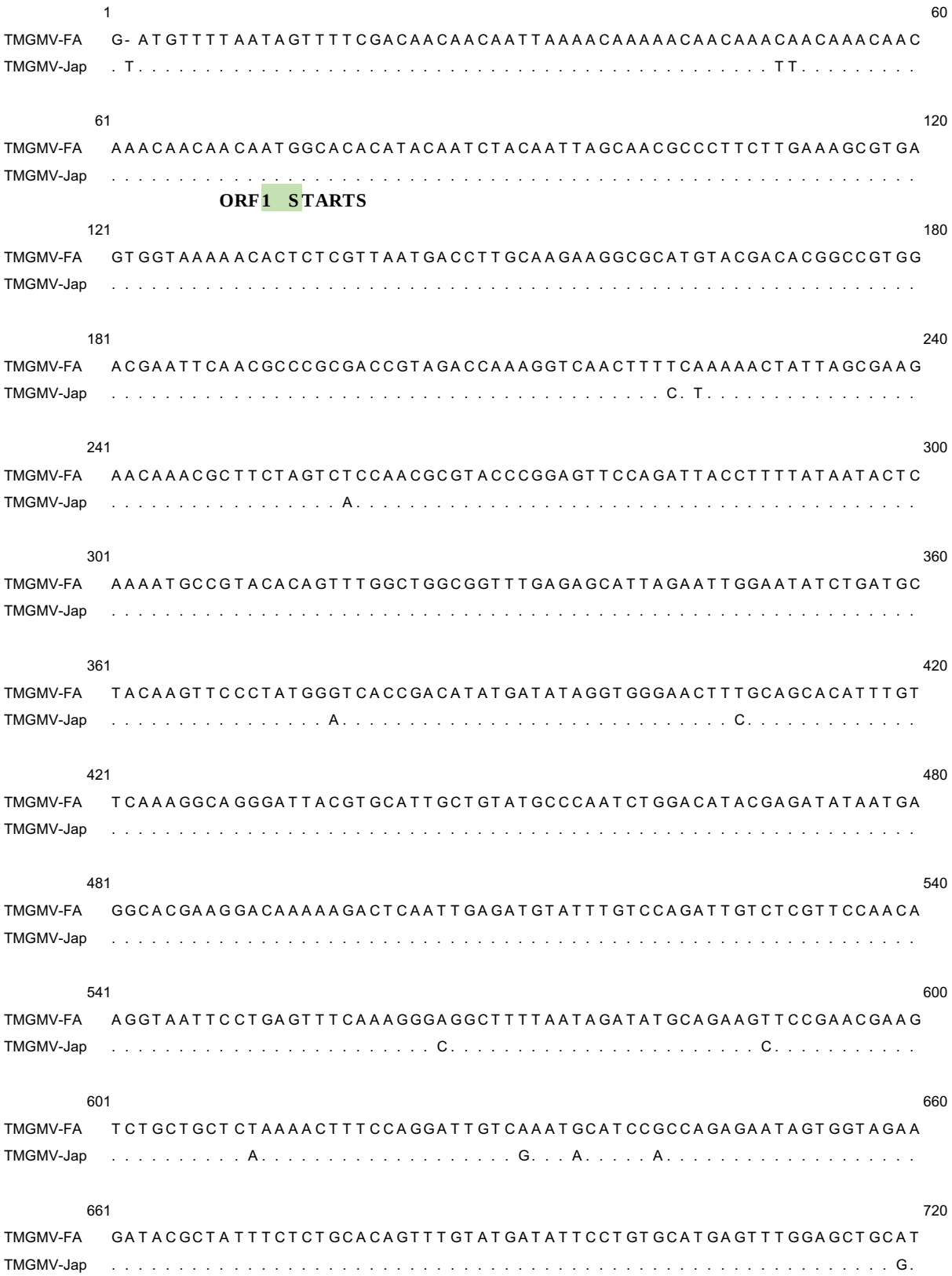
```
          241                                                                                                  300
TMV-U2    TGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCCAGCAGTTAAACCATG
P 00/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . . . . . . . . . .
P 00/6    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 04/17   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . . . . . . . . . .
P 96/49   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/11   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/12   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 01/16   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G. .
P 83/4    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 92/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 94/29   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 97/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/5    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 94/6   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/19  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 92/73  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/11  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/20  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/11  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

          301                                                                                                  360
TMV-U2    TGATGGTGTATACTGTGGTATGGCGTAAAACAACGGAGAGGTTCGAATCCTCCCCTAACC
P 00/10   . . . . . . . . . . . . . . . . . . . G. . . . . . . . . . . T. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 00/6    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . .
P 04/17   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . T. . . . . . . . . . . . .
P 96/49   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/11   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/12   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 01/16   . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 83/4    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 92/10   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 94/29   . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 97/10   . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
P 98/5    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A. . . .
Ng 90/5   . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 94/6   . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/19  . . . . . . C. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/5   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 92/73  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/11  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/20  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 89/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 90/8   . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 96/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/11  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/15  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Ng 99/16  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
```

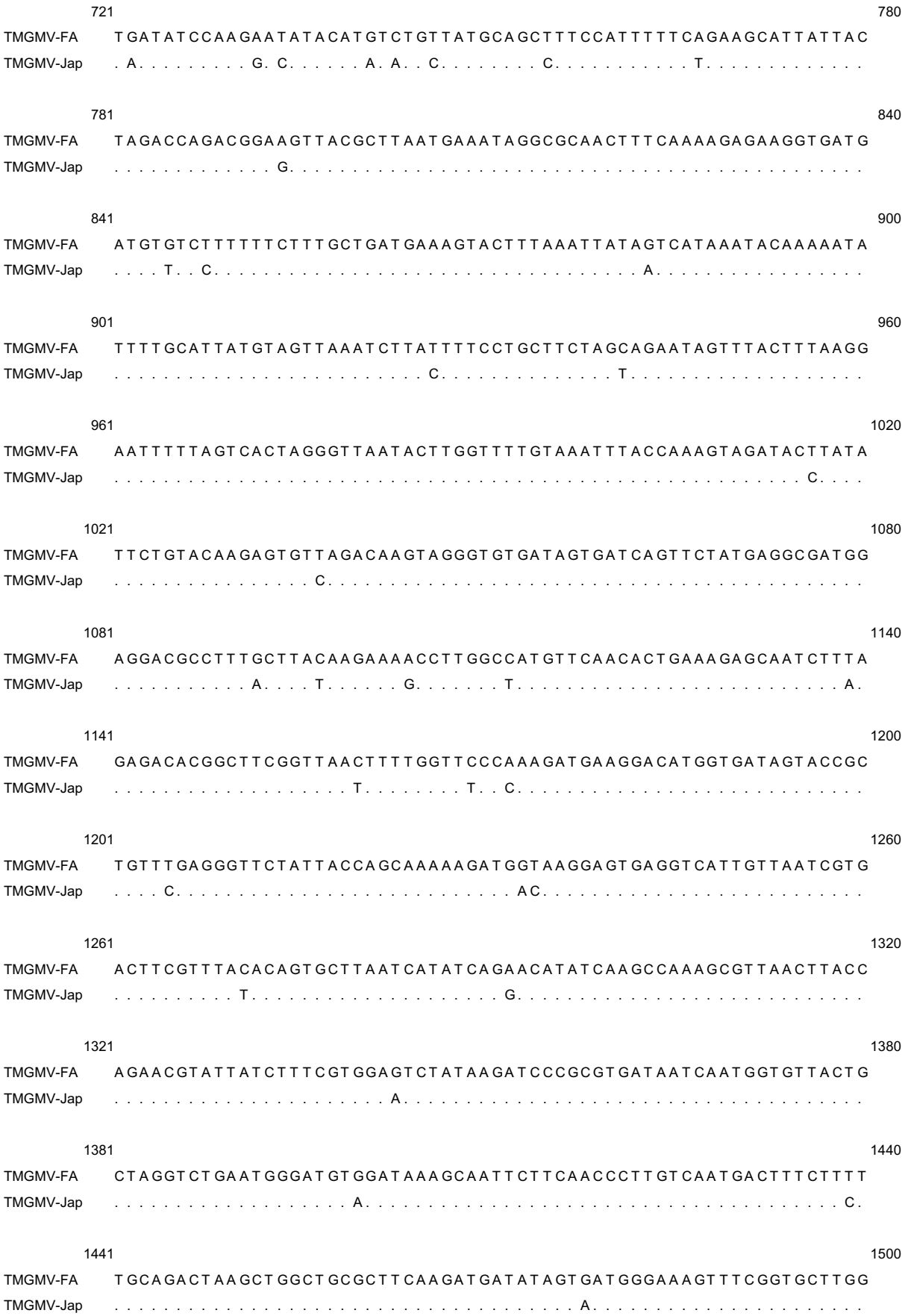
	361	375
TMV-U2	GCGGGTAGCGGCCCA	
P 00/10	. . . . .	
P 00/6	. . . . .	
P 04/17	. . . . .	
P 96/49	. . . . .	
P 98/11	. . . . .	
P 98/12	. . . . .	
P 01/16	. . . . .	
P 83/4	. . . . .	
P 92/10	. . . . .	
P 94/29	. . . . .	
P 97/10	. . . . .	
P 98/5	. . . . .	
Ng 89/8	. . . . .	
Ng 90/5	. . . . .	
Ng 94/6	. . . . .	
Ng 96/19	. . . . .	
Ng 96/5	. . . . .	
Ng 92/73	. . . . .	
Ng 96/11	. . . . .	
Ng 99/20	. . . . .	
Ng 89/15	. . . . .	
Ng 90/8	. . . . .	
Ng 96/16	. . . . .	
Ng 99/11	. . . . .	
Ng 99/15	. . . . .	
Ng 99/16	. . . . .	

ANNEX II:

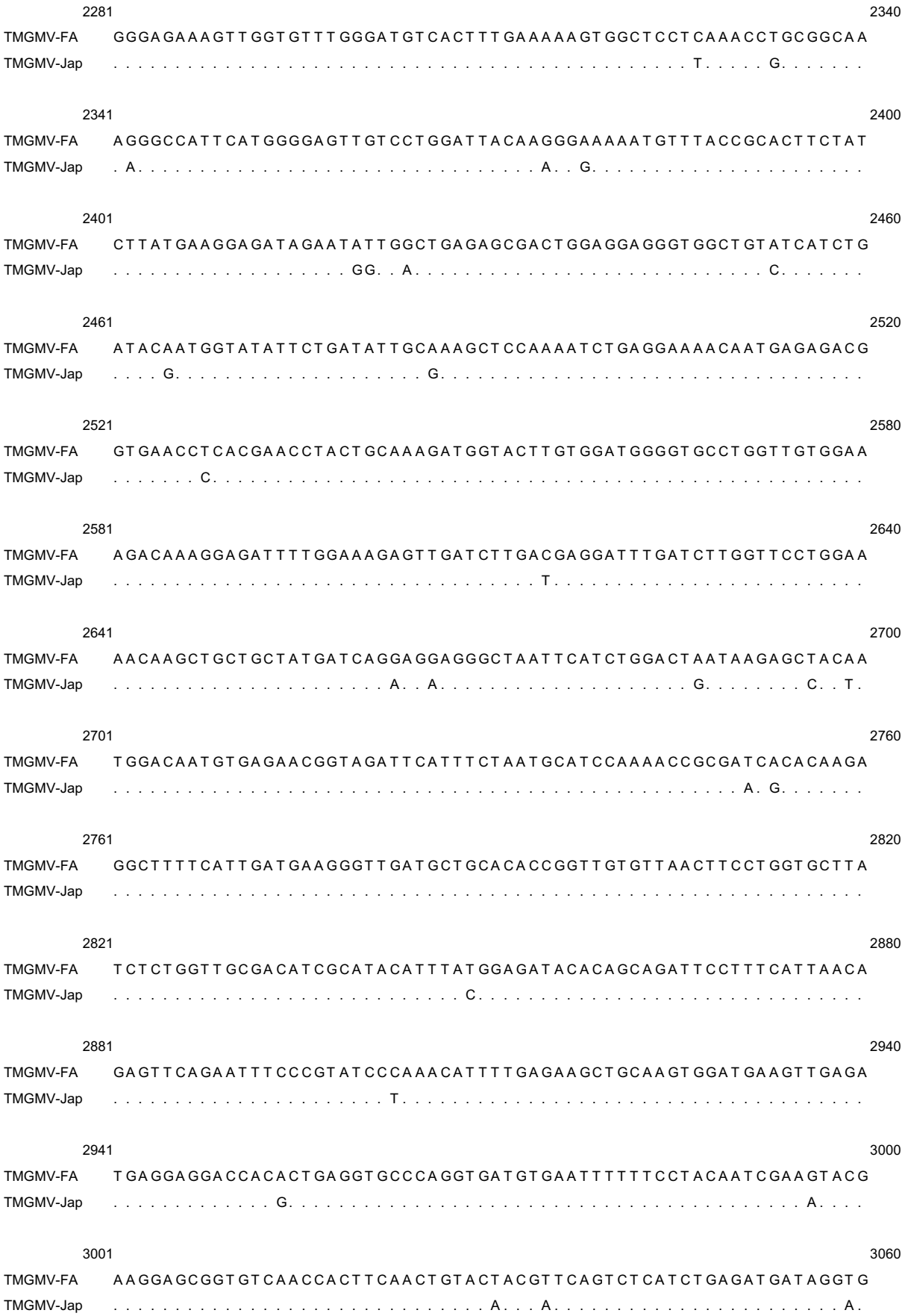
Comparison of the complete nucleotide sequence of the genomic RNA of *Tobacco mild green mosaic virus* from clones: pTMGMV-FA and pTMGMV-Jap

A.II.1 Alignment of the nucleotide sequence of the genomic RNA of TMGMV-FA and TMGMV-Jap




721 780
TMGMV-FA TGATATCCAAGAATATACATGTCTGTTATGCAGCTTTCCATTTTTCAGAAGCATTATTAC
TMGMV-Jap .A.....G.C.....A.A..C.....C.....T.....
781 840
TMGMV-FA TAGACCAGACGGAAGTTACGCTTAATGAAATAGGCGCAACTTTCAAAAGAGAAGGTGATG
TMGMV-Jap .....G.....
841 900
TMGMV-FA ATGTGTCTTTTTTCTTTGCTGATGAAAGTACTTTAAATTATAGTCATAAATACAAAAATA
TMGMV-Jap ....T..C.....A.....
901 960
TMGMV-FA TTTTGCATTATGTAGTTAAATCTTATTTTCCTGCTTCTAGCAGAATAGTTTACTTTAAGG
TMGMV-Jap .....C.....T.....
961 1020
TMGMV-FA AATTTTTAGTCACTAGGGTTAATACTTGGTTTTGTAAATTTACCAAAGTAGATACTTATA
TMGMV-Jap .....C....
1021 1080
TMGMV-FA TTCTGTACAAGAGTGTTAGACAAGTAGGGTGTGATAGTGATCAGTTCTATGAGGCGATGG
TMGMV-Jap .....C.....
1081 1140
TMGMV-FA AGGACGCCTTTGCTTACAAGAAAACCTTGGCCATGTTCAACACTGAAAGAGCAATCTTTA
TMGMV-Jap .....A....T.....G.....T.....A.
1141 1200
TMGMV-FA GAGACACGGCTTCGGTTAACTTTTGGTTCCCAAAGATGAAGGACATGGTGATAGTACCGC
TMGMV-Jap .....T.....T..C.....
1201 1260
TMGMV-FA TGTTTGAGGGTTCTATTACCAGCAAAAAGATGGTAAGGAGTGAGGTCATTGTTAATCGTG
TMGMV-Jap ....C.....AC.....
1261 1320
TMGMV-FA ACTTCGTTTACACAGTGCTTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACC
TMGMV-Jap .....T.....G.....
1321 1380
TMGMV-FA AGAACGTATTATCTTTCGTGGAGTCTATAAGATCCCGCGTGATAATCAATGGTGTTACTG
TMGMV-Jap .....A.....
1381 1440
TMGMV-FA CTAGGTCTGAATGGGATGTGGATAAAGCAATTCTTCAACCCTTGTCAATGACTTTCTTTT
TMGMV-Jap .....A.....C.
1441 1500
TMGMV-FA TGCAGACTAAGCTGGCTGCGCTTCAAGATGATATAGTGATGGGAAAGTTTCGGTGCTTGG
TMGMV-Jap .....A.....

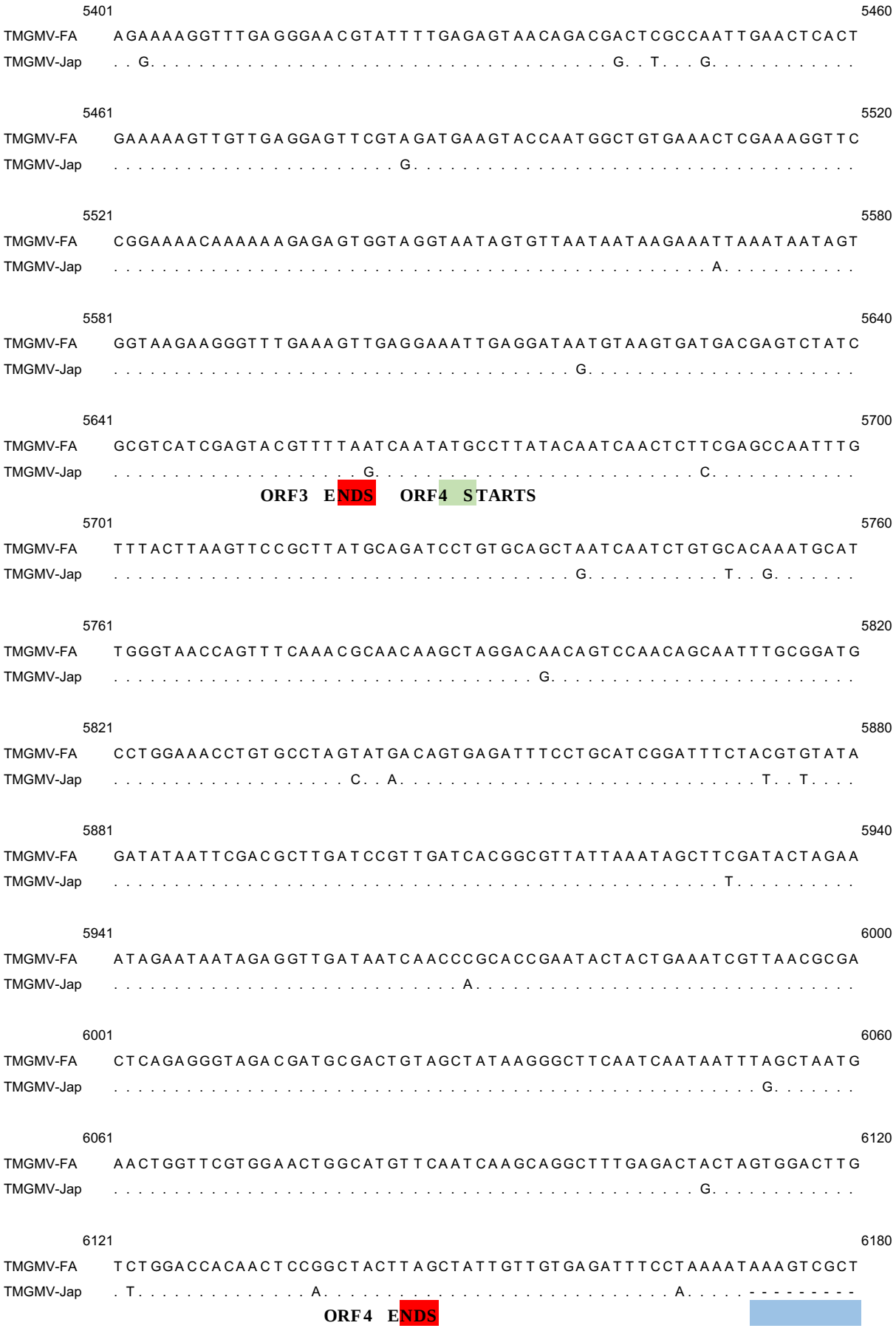
```
          1501                                                                1560
TMGMV-FA  ATAAGACCACTTCTGAACTTATTTGGGATGAGGTGGGCAAATTCTTTGGAAACGTTTTCC
TMGMV-Jap ................................A............................

          1561                                                                1620
TMGMV-FA  CCACTATCAAAGAGAGATTGGTGAGCAGAAAAATTCTGGATGTAAGTGAGAATGCTCTGA
TMGMV-Jap ............................G......T.......G................

          1621                                                                1680
TMGMV-FA  AGATCAAGATCCCGGACCTGTATGTCACATGGAAAGACAGATTCGTAGCTGAGTATACCA
TMGMV-Jap ................T...........G.....G.....G....................

          1681                                                                1740
TMGMV-FA  AGTCTGAGGAGTTACCGCATCTAGATATCAAGAAGGACTTAGAAGAAGCTGAGCACATGT
TMGMV-Jap ......................G...................................A....

          1741                                                                1800
TMGMV-FA  ACGACGCATTGTCGGAATTATCTATCCTTAGGGATGCTGATAATTTCGATATCGCGAAGT
TMGMV-Jap .......G..A.......................................................

          1801                                                                1860
TMGMV-FA  TCAAAGACATGTGTAAAACTTTAGGTGTTAGTCCTGATGTAGCAGCACGAGTAATCGTTG
TMGMV-Jap ............C..GG.....T......A.........G....................

          1861                                                                1920
TMGMV-FA  CAGTGGCTGAGAATAGAAGCGGTTTGACTCTTACTTTTGACAAGCCAACTGAGGAGAATG
TMGMV-Jap .......C.................................T........C..........

          1921                                                                1980
TMGMV-FA  TGGCTAAGGCTCTTAAAAGCACGGCGTCTGAGGCCGTGGTGTGTCTTGAACCGACATCCG
TMGMV-Jap ..............................................C..............T.

          1981                                                                2040
TMGMV-FA  AAGAGGTGAAGGTAAACAAATTTTCTATTGCTGAGAGGGGAAAATTGCCTGTGTGTGCTG
TMGMV-Jap .T.......T......T....................A..G........C..........

          2041                                                                2100
TMGMV-FA  AAAGTCATGGTTTGACGAATGCTAACTTAGAGCACCAGGAGTTGGAGTCCCTCAATGATT
TMGMV-Jap ..........................................................T.....C....

          2101                                                                2160
TMGMV-FA  TCCATAAAGCTTGCGTGGATAGTGTGATTACAAAGCAAATGGCATCGGTTGTCTATACTG
TMGMV-Jap ............................................................................

          2161                                                                2220
TMGMV-FA  GCTCACTCAAAGTTCAACAAATGAAGAACTATGTGGACAGTTTGGCAGCTTCGTTGTCCG
TMGMV-Jap ............................................................................

          2221                                                                2280
TMGMV-FA  CCACTGTATCAAATCTATGCAAGTCATTAAAGGATGTTGTTGGGTATGATTCTGATTCCA
TMGMV-Jap ............................................................................
```

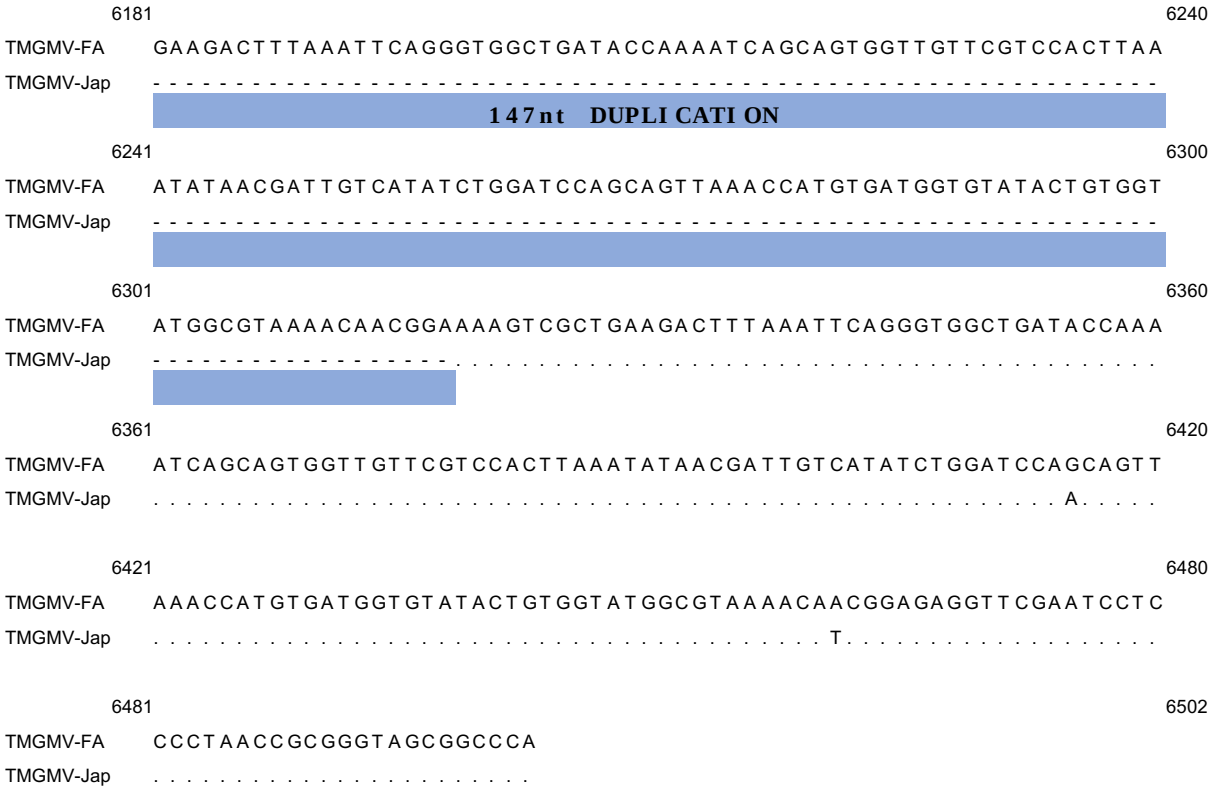

2281 2340
TMGMV-FA GGGAGAAAGTTGGTGTTTGGGATGTCACTTTGAAAAAGTGGCTCCTCAAACCTGCGGCAA
TMGMV-Jap . . . . . T. . . . . G. . . . . .
2341 2400
TMGMV-FA AGGGCCATTCATGGGGAGTTGTCCTGGATTACAAGGGAAAAATGTTTACCGCACTTCTAT
TMGMV-Jap . A. . . . . A. . G. . . . . .
2401 2460
TMGMV-FA CTTATGAAGGAGATAGAATATTGGCTGAGAGCGACTGGAGGAGGGTGGCTGTATCATCTG
TMGMV-Jap . . . . . GG. . A. . . . . C. . . . . .
2461 2520
TMGMV-FA ATACAATGGTATATTCTGATATTGCAAAGCTCCAAAATCTGAGGAAAACAATGAGAGACG
TMGMV-Jap . . . . G. . . . . G. . . . . .
2521 2580
TMGMV-FA GTGAACCTCACGAACCTACTGCAAAGATGGTACTTGTGGATGGGGTGCCTGGTTGTGGAA
TMGMV-Jap . . . . . C. . . . . .
2581 2640
TMGMV-FA AGACAAAGGAGATTTTGGAAAGAGTTGATCTTGACGAGGATTTGATCTTGGTTCCTGGAA
TMGMV-Jap . . . . . T. . . . . .
2641 2700
TMGMV-FA AACAAGCTGCTGCTATGATCAGGAGGAGGGCTAATTCATCTGGACTAATAAGAGCTACAA
TMGMV-Jap . . . . . A. . A. . . . . G. . . . . C. . T.
2701 2760
TMGMV-FA TGGACAATGTGAGAACGGTAGATTCATTTCTAATGCATCCAAAACCGCGATCACACAAGA
TMGMV-Jap . . . . . A. G. . . . . .
2761 2820
TMGMV-FA GGCTTTTCATTGATGAAGGGTTGATGCTGCACACCGGTTGTGTTAACTTCCTGGTGCTTA
TMGMV-Jap . . . . . .
2821 2880
TMGMV-FA TCTCTGGTTGCGACATCGCATACATTTATGGAGATACACAGCAGATTCCTTTCATTAACA
TMGMV-Jap . . . . . C. . . . . .
2881 2940
TMGMV-FA GAGTTCAGAATTTCCCGTATCCCAAACATTTTGAGAAGCTGCAAGTGGATGAAGTTGAGA
TMGMV-Jap . . . . . T. . . . . .
2941 3000
TMGMV-FA TGAGGAGGACCACACTGAGGTGCCCAGGTGATGTGAATTTTTTCCTACAATCGAAGTACG
TMGMV-Jap . . . . . G. . . . . A. . . .
3001 3060
TMGMV-FA AAGGAGCGGTGTCAACCACTTCAACTGTACTACGTTCAGTCTCATCTGAGATGATAGGTG
TMGMV-Jap . . . . . A. . . A. . . . . A.

```
3061                                                                3120
TMGMV-FA    GTAAGGGAGTACTGAACAGTATTTCCAAACCACTTAAAGGGAAAATTGTAACTTTCACTC
TMGMV-Jap   ....................G.............G.........................

3121                                                                3180
TMGMV-FA    AAGCTGATAAATTTGAGTTAGAGGAGAAGGGCTATAAGAATGTGAACACTGTTCATGAGA
TMGMV-Jap   .G..................................................C.........

3181                                                                3240
TMGMV-FA    TCCAAGGAGAAACCTTTGAAGATGTGTCGCTGGTTAGATTGACAGCAACTCCACTGACTC
TMGMV-Jap   ....................................C.........................

3241                                                                3300
TMGMV-FA    TGATTTCCAAGTCTTCCCCGCATGTTCTAGTCGCTCTGACTAGACACACAAAGAGCTTCA
TMGMV-Jap   ....................................C.........................

3301                                                                3360
TMGMV-FA    AATATTACACCGTAGTGTTAGATCCTTTAGTACAGATAGTTAGTGATTTGTCTTCTTTAA
TMGMV-Jap   ....................................................A.........

3361                                                                3420
TMGMV-FA    GCTCCTTCCTTTTAGAAATGTATATGGTGGAAGCAGGTAGTAGATAGCAATTACAGATGG
TMGMV-Jap   ..............................A.........G.....................
                                                  ORF1   ENDS
3421                                                                3480
TMGMV-FA    ATGCAGTGTTCAAAGGTCATAATCTCTTCGTGGCAACACCTAAATCAGGAGACTTTCCAG
TMGMV-Jap   ..............................T.....................................

3481                                                                3540
TMGMV-FA    ATCTGCAGTTCTATTATGATGTATGTCTCCCTGGTAATAGTACTATACTTAACAAGTATG
TMGMV-Jap   ....................................................C.........

3541                                                                3600
TMGMV-FA    ATGCTATTACTATGAGATTACGCGATAATAGTCTTAATGTGAAGGATTGTGTTCTTGATT
TMGMV-Jap   .....G....C.....G.....................................................

3601                                                                3660
TMGMV-FA    TTTCCAAAAGCATTCCGATGCCAAAGGAGGTGGAACCATGTCTAGAGCCAGTTTTGCGTA
TMGMV-Jap   ....................................A.........................

3661                                                                3720
TMGMV-FA    CCGCGGCGGAACCGCCAAGGGCTGCAGGACTACTCGAAAATCTGGTTGCAATGATTAAAA
TMGMV-Jap   ........A.............T.....................................

3721                                                                3780
TMGMV-FA    GAAATTTCAACGCACCAGACCTGACGGGGACGATTGACATCGAGAGCACCGCATCTGTTG
TMGMV-Jap   ....................................................T.........

3781                                                                3840
TMGMV-FA    TAGTAGATAAGTTTTTTGATAGTTATTTTATTAAAAAAGAAAAATACACAAAAAATATTG
TMGMV-Jap   ..............................T.A.........................
```

```
          3841                                                                3900
TMGMV-FA  CTGGAGTGATGACGAAGGATTCAATGATGAGATGGTTGGGAAAACAGGAGAGAAGTACTA
TMGMV-Jap ............................................................

          3901                                                                3960
TMGMV-FA  TTGGACAGTTGGCTAACTACAATTTTGTAGATCTGCCGGCCATCGATCAGTACAAGCACA
TMGMV-Jap .C..........................................................C.....T.......

          3961                                                                4020
TMGMV-FA  TGATCAAGGCTCAACCAAAACAGAAGTTGGATCTTTCAATTCAGAATGAATACCCTGCTC
TMGMV-Jap ........................A.....C.............................

          4021                                                                4080
TMGMV-FA  TGCAAACAATTGTCTACCATTCAAAGCAGATCAACGGTATTTTTGGCCCGGTTTTTTCAG
TMGMV-Jap ............................................................

          4081                                                                4140
TMGMV-FA  AGCTTACAAGGTTGCTGCTCGAGGCAGTTGATTCTCAAAAGTTTCTTTTCTTTACTAGGA
TMGMV-Jap ........................................G...................

          4141                                                                4200
TMGMV-FA  AAACTCCAGAACAGATTCAAGAATTTTTCTCGGATCTCGACTCGCACGTTCCTATGGATG
TMGMV-Jap ............................................................C.

          4201                                                                4260
TMGMV-FA  TGTTAGAACTGGATATTTCTAAGTATGATAAGTCACAGAACGAGTTTCATTGTGCTGTAG
TMGMV-Jap ............C.....................G.........................

          4261                                                                4320
TMGMV-FA  AGTATGAAATATGGAAAAGATTGGGTCTCAATGAGTTTTTGGCCGAAGTGTGGAAACAAG
TMGMV-Jap ....................C.....................................

          4321                                                                4380
TMGMV-FA  GGCATAGGAAAACAACTTTGAAGGATTACATTGCTGGAATCAAGACATGTCTGTGGTACC
TMGMV-Jap .A..C.......................................................T.

          4381                                                                4440
TMGMV-FA  AAAGGAAAAGCGGTGATGTGACTACTTTCATCGGCAATACTGTTATAATAGCAGCTTGCT
TMGMV-Jap ............................................................

          4441                                                                4500
TMGMV-FA  TGGGTTCAATGTTACCGATGGAAAAGGTCATAAAAGGTGCTTTTTGTGGAGATGATTCCG
TMGMV-Jap ....C..........T....................C.............C....

          4501                                                                4560
TMGMV-FA  TTTTGTATTTCCCGAAGGGTTTGGATTTCCCTGACATTCAGTCATGTGCTAATCTCATGT
TMGMV-Jap ..........................................................C.....

          4561                                                                4620
TMGMV-FA  GGAATTTTGAGGCCAAACTGTACAGAAAGAGGTACGGTTACTTTTGTGGCAGGTACATCA
TMGMV-Jap ..............................A.................T.....
```

```
          4621                                                                                                4680
TMGMV-FA  TACACCACGATAAGGGAGCAATAGTGTATTATGATCCTTTGAAGTTGATCTCCAAACTTG
TMGMV-Jap .......T.............................................................G.

          4681                                                                                                4740
TMGMV-FA  GGGCAAAACATATCAAGGATTATGATCACTTGGAAGAGTTGAGGGTGTCTTTGTGTGATG
TMGMV-Jap .............................................................C.

          4741                                                                                                4800
TMGMV-FA  TTGCTTGTTCGCTCGGAAACTGTGCTTACTTTCCGCAGCTGAACGCAGCTATCAAGGAGG
TMGMV-Jap ............................................................

          4801                                                                                                4860
TMGMV-FA  TTCATAAAACCGCAATTGATGGTTCGTTTGCTTTTAATTGTGTTAATAAATTTTTGTGTG
TMGMV-Jap ............G.............................C.............

          4861                                                                                                4920
TMGMV-FA  ATAAATTTTTATTTAGAACTTTGTTTTTAAATGGCTGTTAGTCTCAGAGATACTGTCAAA
TMGMV-Jap ............................................................
                                                ORF3 STARTS     ORF2 ENDS
          4921                                                                                                4980
TMGMV-FA  ATTAGCGAGTTCATTGATCTTTCGAAACAGGATGAGATACTTCCGGCATTCATGACTAAG
TMGMV-Jap ..........T.............................................................A

          4981                                                                                                5040
TMGMV-FA  GTCAAGAGCGTCAGAATATCGACTGTGGATAAGATTATGGCTGTTAAGAATGATAGTCTT
TMGMV-Jap ........T..T................C.............................

          5041                                                                                                5100
TMGMV-FA  TCTGATGTAGATTTACTTAAAGGTGTTAAGTTAGTTAAGAATGGGTATGTGTGCTTAGCT
TMGMV-Jap ..........C.............................................................

          5101                                                                                                5160
TMGMV-FA  GGTTTGGTAGTGTCCGGGGAGTGGAATCTCCCGGACAACTGCCGTGGTGGTGTCAGTGTT
TMGMV-Jap ..............T..................T.........................

          5161                                                                                                5220
TMGMV-FA  TGTATTGTAGATAAGAGAATGAAAAGGAGTAAGGAAGCAACGCTGGGTGCGTATCACGCC
TMGMV-Jap .............................................................A.........

          5221                                                                                                5280
TMGMV-FA  CCTGCTTGCAAAAAGAATTTTTCCTTTAAGCTAATCCCTAATTATTCAATAACGTCCGAG
TMGMV-Jap ......................T.............................A......

          5281                                                                                                5340
TMGMV-FA  GATGCTGAGAAGCACCCATGGCAAGTGTTAGTGAATATCAAAGGAGTGGCTATGGAAGAA
TMGMV-Jap ................G.....G.....................................

          5341                                                                                                5400
TMGMV-FA  GGATACTGTCCTTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATAAAAATAATGTA
TMGMV-Jap ............................................................
```


5401
5460
TMGMV-FA AGAAAAGGTTTGAGGGAACGTATTTTGAGAGTAACAGACGACTCGCCAATTGAACTCACT
TMGMV-Jap ..G.....G..T...G.....
5461
5520
TMGMV-FA GAAAAAGTTGTTGAGGAGTTCGTAGATGAAGTACCAATGGCTGTGAAACTCGAAAGGTTC
TMGMV-Jap .....G.....
5521
5580
TMGMV-FA CGGAAAACAAAAAAGAGAGTGGTAGGTAATAGTGTTAATAATAAGAAATTAAATAATAGT
TMGMV-Jap .....A.....
5581
5640
TMGMV-FA GGTAAGAAGGGTTTGAAAGTTGAGGAAATTGAGGATAATGTAAGTGATGACGAGTCTATC
TMGMV-Jap .....G.....
5641
5700
TMGMV-FA GCGTCATCGAGTACGTTTTAATCAATATGCCTTATACAATCAACTCTTCGAGCCAATTTG
TMGMV-Jap .....G.....C.....
ORF3 ENDS ORF4 STARTS
5701
5760
TMGMV-FA TTTACTTAAGTTCCGCTTATGCAGATCCTGTGCAGCTAATCAATCTGTGCACAAATGCAT
TMGMV-Jap .....G.....T..G.....
5761
5820
TMGMV-FA TGGGTAACCAGTTTCAAACGCAACAAGCTAGGACAACAGTCCAACAGCAATTTGCGGATG
TMGMV-Jap .....G.....
5821
5880
TMGMV-FA CCTGGAAACCTGTGCCTAGTATGACAGTGAGATTTCCTGCATCGGATTTCTACGTGTATA
TMGMV-Jap .....C..A.....T..T....
5881
5940
TMGMV-FA GATATAATTCGACGCTTGATCCGTTGATCACGGCGTTATTAAATAGCTTCGATACTAGAA
TMGMV-Jap .....T.....
5941
6000
TMGMV-FA ATAGAATAATAGAGGTTGATAATCAACCCGCACCGAATACTACTGAAATCGTTAACGCGA
TMGMV-Jap .....A.....
6001
6060
TMGMV-FA CTCAGAGGGTAGACGATGCGACTGTAGCTATAAGGGCTTCAATCAATAATTTAGCTAATG
TMGMV-Jap .....G.....
6061
6120
TMGMV-FA AACTGGTTCGTGGAACTGGCATGTTCAATCAAGCAGGCTTTGAGACTACTAGTGGACTTG
TMGMV-Jap .....G.....
6121
6180
TMGMV-FA TCTGGACCACAACTCCGGCTACTTAGCTATTGTTGTGAGATTTCCTAAAATAAAGTCGCT
TMGMV-Jap .T.....A.....A.....---------
ORF4 ENDS


6181
6240
TMGMV-FA GAAGACTTTAAATTCAGGGTGGCTGATACCAAAATCAGCAGTGGTTGTTCGTCCACTTAA
TMGMV-Jap
147nt DUPLICATION
6241
6300
TMGMV-FA ATATAACGATTGTCATATCTGGATCCAGCAGTTAAACCATGTGATGGTGTATACTGTGGT
TMGMV-Jap
6301
6360
TMGMV-FA ATGGCGTAAAACAACGGAAAAGTCGCTGAAGACTTTAAATTCAGGGTGGCTGATACCAAA
TMGMV-Jap
6361
6420
TMGMV-FA ATCAGCAGTGGTTGTTCGTCCACTTAAATATAACGATTGTCATATCTGGATCCAGCAGTT
TMGMV-Jap A
6421
6480
TMGMV-FA AAACCATGTGATGGTGTATACTGTGGTATGGCGTAAAACAACGGAGAGGTTCGAATCCTC
TMGMV-Jap T
6481
6502
TMGMV-FA CCCTAACCGCGGGTAGCGGCCCA
TMGMV-Jap