Resistance to virus infection mediated by artificial microRNAs: estimating the likelilhood of escape mutants

Guillaume C. G. LAFFORGUE

Programa Oficial de Postgrado en Biotecnología

Directores

Prof. Santiago F. Elena Fito
Dr. José-Antonio Daròs Arnau

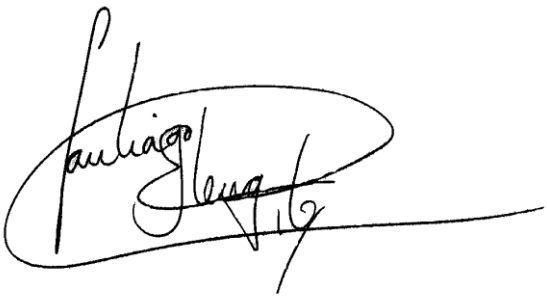
IBMCP
Instituto de Biología Molecular y Celular de Plantas

Santiago F. Elena Fito y José Antonio Darós Arnau, Doctores en Ciencias Biológicas y Profesor de Investigación y Científico Titular, respectivamente, del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Biología Molecular y Celular de Plantas (IBMCP), centro mixto del CSIC y de la Universidad Politécnica de Valencia.

CERTIFICAN

Que D. Guillaume C. G. Lafforgue, Master en Biología Celular y Fisiología y en Biología de Poblaciones y de Ecosistemas, ha realizado bajo nuestra codirección la tesis doctoral titulada "*Resistance to virus infection mediated by artificial microRNAs: estimating the likelilhood of escape mutants*".

Y para que así conste, firman la presente en Valencia, a 14 de Mayo de 2013



Prof. Santiago F. Elena, PhD
santiago.elena@csic.es
http://bioxeon.ibmcp.upv.es/EvolSysVir

46022 València
Tel: 963 877 895
Fax: 963 877 859
Ingeniero Fausto Elío, s/n

Acknowledgments

I spent during my PhD four wonderful years in the city of Valencia, due to the excellent climate and the opportunities given both by the Santiago's lab and Spanish's life. Working in an international lab has been a wonderful and often overwhelming experience. In any case, I am indebted to many people for making the time working on my PhD an unforgettable experience.

First of all, I am deeply grateful to my advisor Santiago F. Elena for insightful comments both on my work and on this thesis. I would like to thank him for leaving me the freedom to start new projects and my free will.

Furthermore, I am very grateful to my second advisor J.A. Daròs, for his comprehension, for his guidance, support and encouragement during my PhD.

In addition, I have been very privileged to get to know and to collaborate with many other great people who became friends over the last several years, as for example the "French group" people:

- Stéphanie, who provides me precious advices and a wonderful project to collaborate with. I enjoyed her philosophy of science and her great histories during endless and repetitive sessions of mashing or measuring plants.

- Nicolas, my lab desk neighbor. Many thanks for his high sense of friendship, for his intense well-argued discussions, for leading me in the wonderful world of protoplast (so pleasant that we stayed several nights) and great moments in the lab.

- Alex, who staid only one year but was precious for her advice in my first steps in bioinformatics. Many thanks for the scientific discussions and for her generosity.

- Romain. We only spent three month together in the lab but what a personality!

I was also very pleased to work with Fernando, my co-PhD, and his molecular biology skills, and Josep who collaborated in this project.

I am grateful to Nacho who continued my formation in phylogeny & Co, and gave me the opportunity to work outside my lab with Beatriz.

I couldn't forget Javi and our conversations leading to laugh for surrealistic incomprehension.

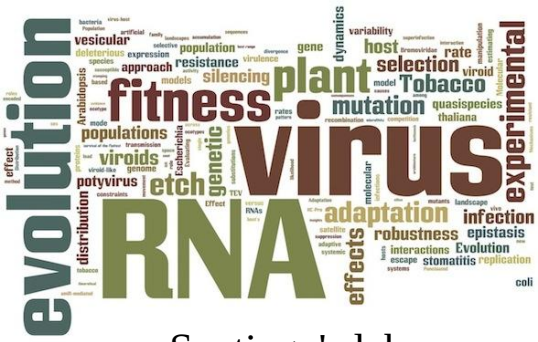
Many people have still to be thanked as Jasna and her direct nature, Susana and her personal vision of life with a touch of madness, Mark with his motivating and passionate discussions.

I would like remember me of the newer as Julia and Anouk, I know that you are doing great job.

There is a special group I would like to honor which was of a great help: the "technicians group". During my stay, I had the pleasure to work first with Ana, then a majority of the time with Àngels and finally with Paula.

To finish I would like to give a special big up to Paqui who is the central nerve system of the lab. I couldn't have produced all these data without her dynamism and love of good work accomplished.

A last thanks for Santi which let me go to NYC for a 3 month stay.



Santiago's lab

Valencia

Thanks,

Table of Contents

ABSTRACT

Plants suffer from the infections caused by many pathogens such as fungi, bacteria, viruses, and nematodes. Virus infections are persistent and frequently are not cleared up by the immune system as in animals. It exist a large variety of plant viruses, around 450 well-characterized species, which cause a high range of diseases in plants. Facing this situation, plants are not merely passive subjects, but they had developed elaborate and effective defense mechanisms to prevent, or limit, damages due to viral infections. Among these defense mechanism, plants have genes that confer resistance to various pathogens, including viruses. These defense systems include single, major resistance genes that induce hypersensitive response, one or more genes that prevents virus replication, cell-to-cell or systemic movement, and more general resistance pathway as RNA silencing. Nowadays, biotechnology can be used to produce virus-resistant crops. Indeed, it is possible to confer transgenic resistance to crops by expressing a resistance gene naturally found in a different plant species. One first biotechnological method to induce resistance to a given virus was by transforming a plant with the gene encoding the virus coat protein (CP). Another natural way plants have to fight virus infection is RNA silencing. This defense mechanism is an important cellular pathway for defense against foreign nucleic acids, including viruses.

In this work we have researched several questions around plants resistant to *Turnip mosaic virus* (TuMV) as a consequence of the transgenic expression of artificial microRNAs (amiRs). MicroRNAs (miRs) are short RNAs (21 nt) found in eukaryotic cells and act as post-transcriptional regulators of gene expression. Their role is to guide the RNA-induced silencing complex (RISC) to cleave the corresponding complementary sequence. It is possible to redesign the miR sequence to target different transcripts using different pre-miRs as backbones. In our case, the pre-miR used as backbone has

been the pre-miR159a precursor from *Arabidopsis thaliana*, and it was engineered to target a 21 nt specific sequence of TuMV HC-Pro cistron. Transgenic expression of these amiRs in *A. thaliana* plants resulted in a complete and specific resistance against TuMV.

Similarly, the gene silencing mechanism (RNAi) has been used in *in vitro* assays antiviral strategies to inhibit the replication of human viruses, such as *Human immunodeficiency virus* type 1 (HIV-1), *Hepatitis C virus* (HCV) and *Influenza A virus*. In all these experiments, a single amiR was expressed, and thus, resistance strictly depended on the match between this amiR and the corresponding viral sequence. A major issue confronting these amiR-based antiviral strategies has been the emergence of resistant virus variants. These variants differ from the wild-type (WT) virus by at least one point mutation in the 21 nt target, leading to imperfect matching with the corresponding amiR and hence to inefficient or ineffective processing by RISC. Although the RNAi machinery tolerates changes in some positions of the 21 nt target, it is sensitive to changes in some others, particularly at the center of the target site. Taken together, these results suggested that some changes in the 21 nt target sequence generate virus escape variants. Until now, apparition of escape variant was observed in cell culture. Such homogeneous environment favors multiplication, colonization and so evolution of the virus.

Transgenic *A. thaliana* plants resistant to TuMV have been chosen to test the durability of the resistance conferred by the amiRs *in vivo*. Indeed, inoculating and following virus infection in whole plants represent a major difference to *in vitro* cell cultures, thus likely having a large impact in virus evolution. This work thus aimed at tracking the likelihood of TuMV to evolve and generate escape mutants able to breakdown the amiR-mediated resistance. In other words, to evaluate the viability of antiviral therapies based on the transgenic expression of amiRs in plants. It is essential to understand how likely are viral populations to contain escape variants, which may be subsequently transmitted to immunized plants. To do so, we used two lines of *A. thaliana* transgenic plants: the 12-4 line which shows

total resistance to the virus and the 10-4 line, with partial resistance to TuMV infection.

In a first experiment, we evolved TuMV in different contexts, wild-type plants or the 10-4 line. Regularly, fully resistant 12-4 plants were challenged with the evolved viral populations on the. At the end of the experiment, all independent viral evolved linages were able to produce escape mutant. We notably evaluate the selective force when the virus evolved in a subinhibitory context and we found acceleration in the emergence of escape mutant. The major characteristic of these variants is always at least one mutation in the 21-nt amiRs target region, mainly at two critical positions at the center. At face value, these results rest interest to the amiR-based resistance technology. Then, how can we improve it to produce a more durable resistance?

In a second experiment, the natural variability of TuMV was overviewed among 100 field isolates. The phylogenetic analysis of these sequences revealed two major points. First, the HC-Pro region targeted by the amiR used in the first experiment has a high genetic variability along its 21 nt, whereas other regions at the CP show no variability at all among the 100 isolates. With this information, a second generation of transgenic plants was produced expressing two amiRs specific to the CP conserved region. In short, the new double transgenic *A. thaliana* plants showed high and more durable resistance to TuMV. Results promise a considerable durability in field.

In our last experiment, we evaluate the potential for evolution of viral suppressors of RNA silencing (VSR) in TuMV genome. With the addition of a new cistron coding for *Cucumber mosaic virus* (CMV) VSR into TuMV genome, we created a TuMV carrying functional redundancy. After evolution by serial passages, the modification encountered in the HC-Pro and the second VSR were only punctual mutations. The general message of these results is that recombination between TuMV and another virus can result in stable chimeras. The addition of a second VSR, which can also be a possible multifunctional protein, may give the chimeric virus several advantages to colonize the host plant.

There are several advantages associated to the use of the amiR strategy. Unlike in other biotechnologies, only one stable small RNA (21 nt) is required, whose sequence can be chosen to reduce off-target effects in plants genome. The amiR strategy also minimizes potential risks for bio-safety concern, reducing any possible negative environmental impact. Finally, broad-spectrum resistance to several viruses can also be achieved by co-expression of appropriately designed multiple amiRs.

This work brings an evolutionary biology vision to the amiRs, indicating several critical choices that have to be made for designing durable resistances.

RESUMEN

Las plantas sufren las infecciones causadas por muchos patógenos como hongos, bacterias, virus y nematodos. Las infecciones virales son persistentes y frecuentemente no se curan por el sistema inmune como en los animales. Existe una gran variedad de virus de plantas, alrededor de 450 especies bien caracterizadas, que cusan una alta gama de enfermedades en plantas. Enfrentando esta situación, las planas no son meros sujetos pasivos, sino que han desarrollado mecanismos defensivos elaborados y efectivos para prevenir, o limitar, los daños debidos a las infecciones virales. Entre estos mecanismos de defensa, las plantas tienen genes que confieren resistencia a virios patógenos, incluyendo los virus. Estos sistemas defensivos incluyen importantes genes únicos de resistencia que inducen la respuesta hipersensible, uno o más genes que previenen la replicación viral, el movimiento célula a célula o sistémico, y rutas de resistencia más generales como el silenciamiento por RNA. Actualmente, la biotecnología puede usarse para producir cultivos resistentes a virus. De hecho, es posible conferir resistencia transgénica a los cultivos mediante la expresión de un gen de resistencia encontrado de forma natural en una especie de planta alternativa. Un primer método biotecnológico para inducir resistencia a un virus determinado fue la transformación de la planta con el gen que codifica la CP del virus. Otra forma natural que las plantas tienen para luchar contra las infecciones virales es el silenciamiento por RNA. Este mecanismo defensivo es una importante ruta celular para la defensa contra ácidos nucleicos exógenos, incluyendo los virus.

En este trabajo hemos investigado algunos temas relacionados con las plantas resistentes al virus del mosaico del nabo (TuMV) como consecuencia de la expresión transgénica de microRNAs (amiRs) artificiales. Los microRNAs (miRs) son pequeños RNAs (21 nt) que se encuentran en las células eucariotas y actúan como reguladores postranscripcionales de la expresión génica. Su papel es guiar al complejo de silenciamiento inducido por RNA (RISC) para

cortar la secuencia complementaria correspondiente. Es posible rediseñar la secuencia del miR para reconocer un transcrito diferente utilizando distintos precursores de miRs (pre-miRs). En nuestro caso, el pre-miR utilizado como precursor ha sido el pre-miR159a de *Arabidopsis thaliana*, y se ingenierizó para reconocer una secuencia específica de 21 nt del cistrón HC-Pro del TuMV. La expresión transgénica de estos amiRs en plantas de *A. thaliana* resultó en una resistencia completa y específica contra el TuMV.

De forma similar, el mecanismo de silenciamiento génico (RNAi) se ha utilizado en ensayos *in vitro* como estrategia antiviral para inhibir la replicación de virus de humanos, tales como el virus de la inmunodeficiencia humana tipo 1 (HIV-1), el virus C de la hepatitis (HCV) o el virus A de la gripe. En todos estos experimentos se expresó un solo amiR, y así, la resistencia dependía exclusivamente del apareamiento entre este amiR y la secuencia viral correspondiente. Una cuestión importante contra estas estrategias antivirales basadas en amiRs ha sido la emergencia de variantes virales resistentes. Estas variantes se diferencian del virus silvestre (WT) en al menos una mutación puntual en la diana de 21 nt, provocando el reconocimiento imperfecto con el correspondiente amiR y, consecuentemente, el procesamiento ineficiente o inefectivo por parte de RISC. Aunque la maquinaria de RNAi tolera cambios en algunas de las posiciones de la diana de 21 nt, es sensible a cambios en otras, particularmente en el centro de la diana. Tomados en conjunto, estos resultados sugieren que algunos cambios en la secuencia diana de 21 nt generan variantes virales de escape. Hasta ahora, la aparición de variantes de escape se observó en cultivos celulares. Un ambiente tan homogéneo favorece la multiplicación, colonización y, en definitiva, la evolución viral.

Las plantas de *A. thaliana* transgénicas que son resistentes al TuMV se seleccionaron para ensayar la durabilidad de la resistencia conferida por los amiRs *in vivo*. De hecho, la inoculación y seguimiento de la infección viral en plantas enteras representa una diferencia importante con los cultivos celulares *in vitro*, lo que probablemente tendrá un gran imparto en la evolución viral. Así, este

trabajo tuvo como objetivo el seguimiento de la posibilidad que el TuMV evolucione y genere mutantes de escape capaces de romper la resistencia mediada por el amiR. En otras palabras, evaluar la viabilidad de las terapias antivirales basadas en la expresión transgénica de amiRs en plantas. Es esencial entender como de probables son las poblaciones virales que contienen variantes de escape, que pueden ser seguidamente transmitidas a las plantas inmunes. Para conseguirlo, usamos dos líneas de plantas de *A. thaliana* transgénicas: la línea 12-4 que muestra total resistencia al virus y la línea 10-4, con resistencia parcial a la infección por TuMV.

En un primer experimento, evolucionamos el TuMV en diferentes contextos, plantas silvestres o la línea 10-4. Regularmente, las plantas 12-4 completamente resistentes se desafiaron con las poblaciones virales evolucionadas. Al final del experimento, todos los linajes virales evolucionados independientemente fueron capaces de producir mutantes de escape. De forma remarcable, evaluamos la fuerza selectiva cuando el virus evoluciono en un contexto subinhibitorio y encontramos una aceleración en la emergencia de mutantes de escape. La principal característica de estas variantes es siempre al menos una mutación en la región diana de 21 nt del amiR, principalmente en dos posiciones críticas en el centro. En principio, estos resultados restan interés a la tecnología de resistencia basada en amiRs. Entonces, ¿cómo podemos mejorarla para producir una resistencia más perdurable?

En un segundo experimento, la variabilidad natural del TuMV se contrastó entre 100 aislados de campo. El análisis filogenético de estas secuencias reveló dos puntos importantes. Primero, la región de HC-Pro reconocida por el amiR utilizado en el primer experimento tiene una gran variabilidad genética a lo largo de los 21 nt, mientras que otras regiones en la CP no muestran ninguna variabilidad entre los 100 aislados. Con esta información, se produjo una segunda generación de plantas transgénicas que expresaban dos amiRs específicos contra la región conservada de CP. En resumen, las nuevas plantas doblemente transgénicas de A. thaliana mostraron una

resistencia mayor y más duradera al TuMV. Estos resultados prometen una durabilidad considerable en el campo.

En nuestro último experimento, evaluamos el potencial para evolucionar de supresores de silenciamiento virales (VSR) en el genoma del TuMV. Con la inserción de un nuevo cistrón que codifica la VSR del virus del mosaico del pepino (CMV) en el genoma del TuMV, creamos un TuMV que lleva una redundancia funcional. Tras la evolución mediante pases seriados, las modificaciones que se encontraron en HC-Pro y en el segundo VSR fueron solo mutaciones puntuales. El mensaje general de estos resultados es que la recombinación entre el TuMV y otro virus puede resultar en quimeras estables. La adición de un segundo VSR, que puede también ser una proteína multifuncional, puede conferir al virus quimérico varias ventajas para colonizar la planta.

Existen varias ventajas asociadas al uso de la estrategia de los amiRs. En contraste con otras biotecnologías, solo se requiere un pequeño RNA (21 nt), cuya secuencia se puede elegir para reducir efectos colaterales en el genoma de la planta. La estrategia de los amiRs también minimiza los riesgos potenciales en bioseguridad, reduciendo cualquier posible impacto negativo sobre el medio ambiente. Finalmente, se puede conseguir una resistencia de amplio espectro a varios virus mediante la coexpresión de múltiples amiRs adecuadamente diseñados.

Este trabajo trae la visión de la biología evolutiva a los amiRs, remarcando algunas decisiones críticas que se tienen que hacer para diseñar resistencias perdurables.

I. INTRODUCTION

A major constraint in agriculture is the negative effect of pathogens on crops. It is an economic challenge to reduce the impact of pest on loss of productivity. Indeed annual worldwide losses due to plant pathogens alone are estimated to be 12% of total production (Cook, 2006). Since the beginning of agriculture, humanity improved its cultures in aspects such as tolerance to drought, production, taste, and naturally, resistance to pathogens. Before the biotechnology era, breeders used existing natural resistance, resulting from the arm races between plants and pests, to enhance protection against phytopathogens. Unfortunately pathogens have the capacity to evolve and adapt themselves to escape resistances after some time (Garcia-Arenal and McDonald, 2003). The challenge is now to extend this period of crop full protection against pathogens using a new generation of tools designed to predict evolution of pests.

In this work we focus on virus protection mediated by artificial microRNAs (amiRs). If the amiR biotechnological tool is well understood and can be used in plants, it remains the question of durability of this new technology against organisms that have the ability to evolve and adapt themselves rapidly. We will first describe the different known resistances in plants against viruses and how viruses can evolve to avoid these resistances. Then this work will directly address the question of the durability of amiR-mediated resistance in different contexts. The results will drive us to ask more questions about the dynamic evolution of viruses and finally with this knowledge in hand, we will propose ways to improve the amiR technology that take into account viral evolution. The second part of this thesis will address the question of the evolution of functional redundancy for the gene silencing suppressor activity of a plant RNA virus.

1. Plant Mechanisms of Resistance

Interactions between virus and host are determined by the effect that the pathogen has on the host and the effect that the host responses have on the pathogen. In their race against the establishment and multiplication of viruses, plants detect pathogens (elicitors and whilst detection of non-specific elicitors) and then develop a wide variety of complexes responses to stop or contain viral infections (Fig. 1). Detection of specific elicitors is required to resist pathogens that have evolved to overcome all the non-specific barriers and detection mechanisms. The model presented in Figure 1 illustrates non-host and host-specific resistance that result in the activation of similar responses and the possible integration of non-host and host-specific resistance signaling pathways sharing same signaling components. In this introductory chapter, we examine the different plant defense mechanisms against viruses as *R* gene mediated resistance against specific pathogens and recessive gene-mediated resistances. In a special section we describe the host RNA silencing as key determinants of the outcome of many plant-virus interactions.

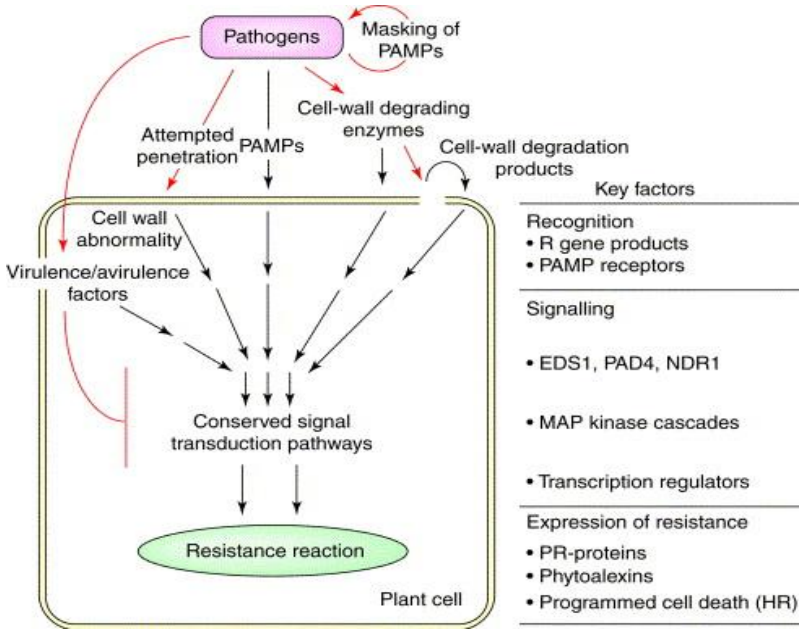


Figure 1. Schematic representation of the signaling responses induced by various plant pathogen elicitors. Elicitors produced by pathogens can be non-specific (that may be loosely defined as plant pathogen-associated molecular patterns or PAMPs), or highly specific (virulence effector/avirulence factors). These highly specific elicitors may have the role of suppressing plant mechanisms reaction. Red arrows indicate pathogen strategies for infection and black arrows indicate plant signaling for resistance. If detection of PAMPs is shown at the cell surface, and the action of virulence effector proteins and their detection as avirulence factors is shown in the cytosol, we have to say that these locations are not mutually exclusive (Jones and Takemoto, 2004).

1.1. Natural Resistance

1.1.1 Non-host resistance

Most plants are resistant to most viruses. This type of resistance when all members of a species are completely resistant to a particular pathogen, is referred as non-host resistance (Staskawicz et al., 1995). Passive resistance might result from a lack of one or more necessary host factors for virus replication or movement. It can be due also to physical barriers to infection, usually resulting from the combined influence of many plants genes (Maule et al., 2007). Non-host resistance shown by an entire plant species to a specific pathogen is the most common and durable form of plant resistance (Heath, 2000). Physical barriers, antimicrobial compounds production, hypersensitive response (HR) are different defensives strategies against, in majority, to fungal and bacteria pathogens (Mysore and Ryu, 2004).

1.1.2 Monogenic dominant resistance

This resistance against a specific pathogen is governed by only one dominant gene able to recognize it. Most resistance genes in plants act against bacterial or fungal pathogens (Hammond-Kosak and Parker, 2003). Until now ca. 13 monogenic dominant resistance genes have been discovered that are specific against virus and all, except at least three RTM genes (for Restricted TEV Movement) *RTM1, RTM2* and *RTM3* identified *in A. thaliana* (Cosson et al., 2012), belong to the nucleotide binding site-leucine rich repeat class (NBS-LRR) of resistance genes. As fungi and bacteria NBS-LRR resistance, antiviral genes lead to complete resistance but not always associated with cell death and/or tissue necrosis. These genes operate through a 'gene-for-gene' recognition of pathogen avirulence factors (Flor, 1971), which for viruses relate to a diverse spectrum of viral gene products (Table 1). The non-NBS-LRR genes *RTM1, RTM2* and *RTM3* effective in *A. thaliana* against *Tobacco etch virus* (TEV), were phenotypically identified as genes inhibiting TEV long-distance movement.

Table 1. Dominant virus resistance genes (Maule et al., 2007).

Gene	Virus	avr determinant	Plant sp	References
HRT	*Turnip crinkle virus* (TCV) Carmovirus	Coat protein	*A.thaliana*	(Cooley et al., 2000), (Ren et al., 2000)
N	*Tobacco mosaic virus* (TMV) Tobamovirus	Replicase/helicase	Tobacco	(Whitham et al., 1996)
RCY1	*Cucumber mosaic virus* (CMV) Cucumovirus	Coat protein	*A.thaliana*	(Takahashi et al., 2001)
Rsv1	*Soybean mosaic virus* (SMV) Potyvirus	Unknown	Soybean	(Hayes et al., 2004)
Rt4-4	CMV Cucumovirus	2a gene	*Phaseolus vulgaris*	(Seo et al., 2006)
RTM1	*Tobacco etch virus* (TEV) Potyvirus	Unknown	*A.thaliana*	(Chisholm et al., 2000)
RTM2	TEV Potyvirus	Unknown	*A.thaliana*	(Whitham et al., 2000)
Rx1	*Potato virus X* (PVX) Potexvirus	Coat protein	Potato	(Bendahmane et al., 1995)
Rx2	PVX Potexvirus	Coat protein	Potato	(Bendahmane et al., 2000)
Sw5	*Tomato spotted wilt virus* (TSWV) Tospovirus	Movement protein	Tomato	(Brommonschenkel et al., 2000)
Tm2^2	*Tomato mosaic virus* (TMV) Tobamovirus	Movement protein	Tomato	(Weber and Pfitzner, 1998)
Y-1	*Potato virus Y* (PVY) Potyvirus	Unknown	Potato	(Vidal et al., 2002)

RTM1 encodes a protein belonging to the jacalin family, with members also involved in defense against insects and fungi (Chisholm et al., 2000). *RTM2* encodes a protein with similarities to small heat shock proteins (Whitham et al, 2000). Finally *RTM3* encodes for a meprin and TRAF homology domain-containing protein (Cosson et al., 2010).

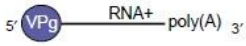
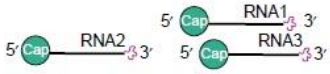
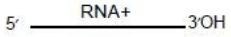
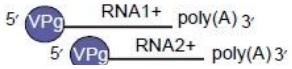
Controlling complete resistance with dominant *R* genes appears to be an attractive option for breeders and have been used when they have been available. Although *R* gene function is based upon precise and specific molecular interactions that lead to the activation of resistance by pathogen avirulence determinants.

1.1.3 Recessive resistance

As obligate parasites, viruses need host factors to complete their infection cycle. Natural mutations of components of the eukaryotic translation initiation complex result in resistance to specific RNA viruses (Diaz et al., 2004; Robaglia and Caranta, 2006). In this case resistance is given by mutated alleles in homozygosis; factually the host is unable to produce the original host factor necessary for the virus‘s cycle replication (Table 2).

A well-documented example is resistance to potyviruses which need an interaction between viral and host protein. Potyviruses and certainly others single-stranded positive-sense RNA viruses posses a covalently linked protein (VPg), instead of a conventional cap structure at the 5’ end of its genomic RNA. To be translated it is necessary that the VPg interact with specific translation factors (eIF4E, eIF(iso)4E or eIF4G depending of the virus). In this case, host resistance gene encode recessive variants of translation factors that are unable to interact properly with the VPg (Leonard et al., 2000; Leonard et al., 2004; Moury et al., 2004; Nicolas et al., 1997) and so break the virus cycle at its early stage. To extend the case of potyviruses, a recessive resistance mediated by eIF(iso)4G gene named *rymv1* confer high resistance to *Rice yellow mottle virus* (RYMV) in rice (Lapidot et al., 1993).

Table 2. Translation initiation factors required for the infection cycle of plant RNA viruses differing in structure and genome expression strategy (Robaglia and Caranta, 2006).

Genus	Virus	Plant	Locus	Gene expression control	Translation factor
Potyvirus	TuMV, TEV	*Arabidopsis*	*lsp1*	Knock-out (EMS-induced)	eIF(iso)4E
	TuMV, LMV	*Arabidopsis*	*lsp1*	Knock-out (T-DNA)	eIF(iso)4E
	CIYVV	*Arabidopsis*	*cum1*	Knock-out (EMS-induced)	eIF4E1
	PVY, TEV	*Capsicum* spp.	*pvr2*	Naturally occurring mutations	eIF4E
	PVMV	*Capsicum* spp.	*pvr6*	Naturally occurring knock-out	eIF(iso)4E
	LMV	*Lactuca* spp.	*mo1*	Naturally occurring mutations	eIF4E
	PSbMV	*Pisum sativum*	*sbm1*	Naturally occurring mutations	eIF4E
	PVY, TEV	*Lycopersicon* spp.	*pot1*	Naturally occurring mutations	eIF4E
Cucumovirus	CMV	*Arabidopsis*	*cum1*	Knock-out (EMS-induced)	eIF4E
	CMV	*Arabidopsis*	*cum2*	EMS-induced mutations	eIF4G
Carmovirus	TCV	*Arabidopsis*	*cum2*	EMS-induced mutations	eIF4G
	MNSV	*Cucumis melo*	*nsv*	Naturally occurring mutations	eIF4E
Bymovirus	BaYMV, BaMMV	*Hordeum vulgare*	*rym4/5*	Naturally occurring mutations	eIF4E
					

1.1.4 Resistance against virus vectors

One possible protection against virus infections is a resistance against theirs vectors, such as aphids or nematodes. This strategy is represented, as far as we know, only by few genes, which belong to the same NBS-LRR group that other plant *R* genes. This indicates the importance of this group on resistance not only against pathogens but also against pests. The first discovered gene is the tomato *Mi-1* gene that confers resistance against the potato aphid (*Macrosiphum euphorbiae*), whitefly (*Bemisia tabaci*) and the root-knock nematode (*Meloidogine spp.*) (Vos et al., 1998). Then the CC-NBS-LRR melon Vatgene controls plant colonization by aphid (*Aphis gossypii*) and transmission of non-persistent viruses (Silberstein et al., 2003). Similarly, the *Nr* gene in lettuce confers resistance to aphid (*Nasonovia ribisnigri*), and presents similarity in functions to *Mi-1* and mapped to a region predicted to contain CC-NBS-LRR *R* genes analogues (Klingler et al., 2005). There is increasing evidence that resistance to vectors may be mediated by single resistance genes of similar structure.

1.2. Induced Resistance

Induced resistance depends on factors present only after the contact with the pathogen (Sequeira, 1983). The recent advances in molecular biology and plant transformation have provided evidences that sensitizing a plant to respond more rapidly to infection could confer increased protection against virulent pathogens. The inducing inoculation can involve the same pathogen used in the challenge inoculation, avirulent or incompatible forms of the pathogen or other pathogen or saprophytes (Sequeira, 1983). The two most clearly defined forms of induced resistance are the systemic acquired resistance (SAR) and the induced systemic resistance (ISR) (Fig 2), which can be differentiated on the basis of the nature of the elicitor and the regulatory pathways involved, as demonstrated in model plant systems (Schenk et al., 2000; Uknes et al., 1992).

The fact that strains of viruses vary in their virulence on different crops and even within varieties of the same crop is particularly problematic; indeed a virus used to protect one crop could potentially cause serious diseases on alternate crops or varieties growing nearby.

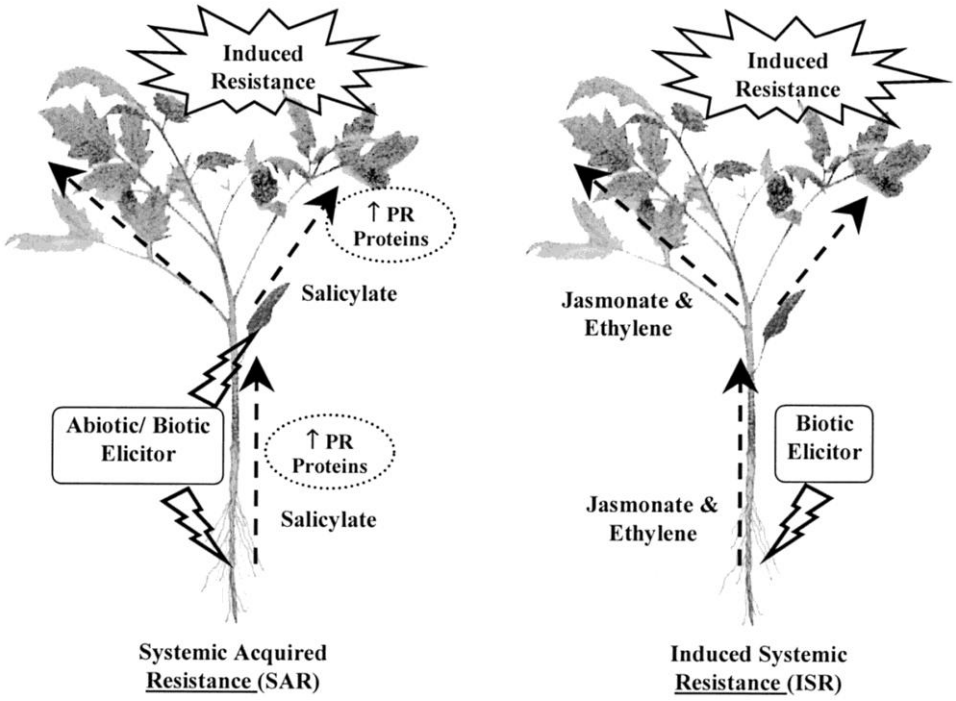


Figure 2. Representation of the two best characterized forms of induced resistance in plants. Systemic acquired resistance (SAR) is dependent on the phytohormone salicylate (salicylic acid) and induced by the exposure of root or foliar tissues to abiotic or biotic elicitors. This form of induced resistance is associated with the accumulation of pathogenesis-related (PR) proteins. Induced systemic resistance (ISR) is dependent of the phytohormones ethylene and jasmonate and induced by the exposure of roots to specific strains of plant growth-promoting rhizobacteria (Vallad and Goodman, 2004).

1.3. **RNA Silencing**

RNA silencing was first observed in experiments dedicated to over-express chalcone synthase (CHS) in petunia by insertion of an extra copy of CHS and that, surprisingly, resulted in the opposite result: suppression of the endogenous CHS expression (Napoli et al., 1990). Indeed petunias flower instead of presenting a more intense coloration appeared to be white, as CHS was not produced. RNA silencing mechanisms act on repression of mRNA transcription by methylation of genomic loci, post-transcriptional degradation of target RNAs and inhibition of mRNAs translation (Aukerman and Sakai, 2003; Llave et al., 2002; Zilberman et al., 2003). This downregulation or complete suppression of gene expression is caused by an antisense RNA molecule. RNA complex long dsRNA transcripts are produced by RNA-dependent RNA polymerase (RdRp) from complementary mRNA-small and small dsRNAs are processed by Dicer (RNase III) in pieces of about 21-24 nt in size (Fig. 3). Then the RNA-induced silencing complex (RISC) incorporates single-stranded small RNAs.

The most common and well-studied example is RNA interference, in which endogenously expressed miRs or exogenously derived small interfering RNA (siRNA) induces the degradation of complementary messenger RNA. Gene silencing can act at two levels, one in the nucleus named transcriptional gene silencing (TGS) and one in the cytosol after transcription, defined as post-transcriptional gene silencing (PTGS). Two of the main roles of RNA silencing are development regulation and defending plants against viruses (Fig 4). dsRNAs generated from aberrant transgenic or viral transcript by a plant RdRp without short siRNA guides are processed to short and long ds siRNAs by RNase Dicer. siRNAs are distinct in size and function, and probably arise from different Dicer activities serving as guide for a range of effector complexes. RISC is then activated by these siRNAs for target cleavage or translational arrest and also guide plant RdRp to amplify dsRNAs, which are diced again to short and long siRNAs. Finally siRNAs are also responsible for short-distance signaling, whereas the long siRNAs are probably involved in long-

distance silencing and transcriptional silencing (Brosnan and Voinnet, 2011).

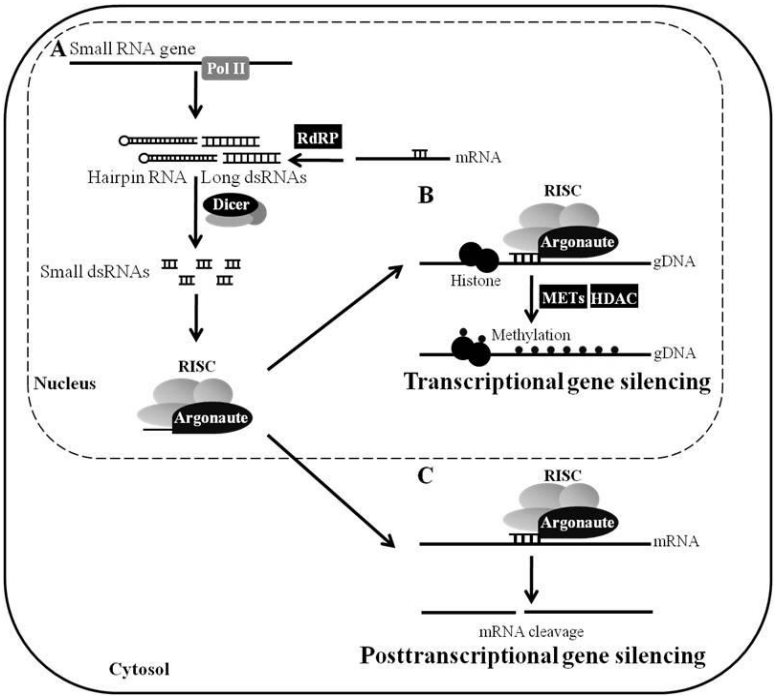


Figure 3. The two gene silencing pathways in plants. (A) Double-stranded (ds) RNAs are formed and used as guides to cleavage the inducing RNAs into short fragments (21-24 nt). (B) Transcriptional gene silencing (TGS). RISC binds onto complementary genomic DNA and then, chromatin modification is promoted by DNA methyltransferase, histone methyltransferase and histone deacetylase (HDAC) activities, resulting in the TGS. (C) Posttranscriptional gene silencing (PTGS). RISC complex moves from nucleus to cytoplasm (Ohrt et al., 2012), and then associates with complementary mRNA to cleave target mRNA by Argonaute activity (from http://bk21.gsnu.ac.kr/jdbahk/research.htm).

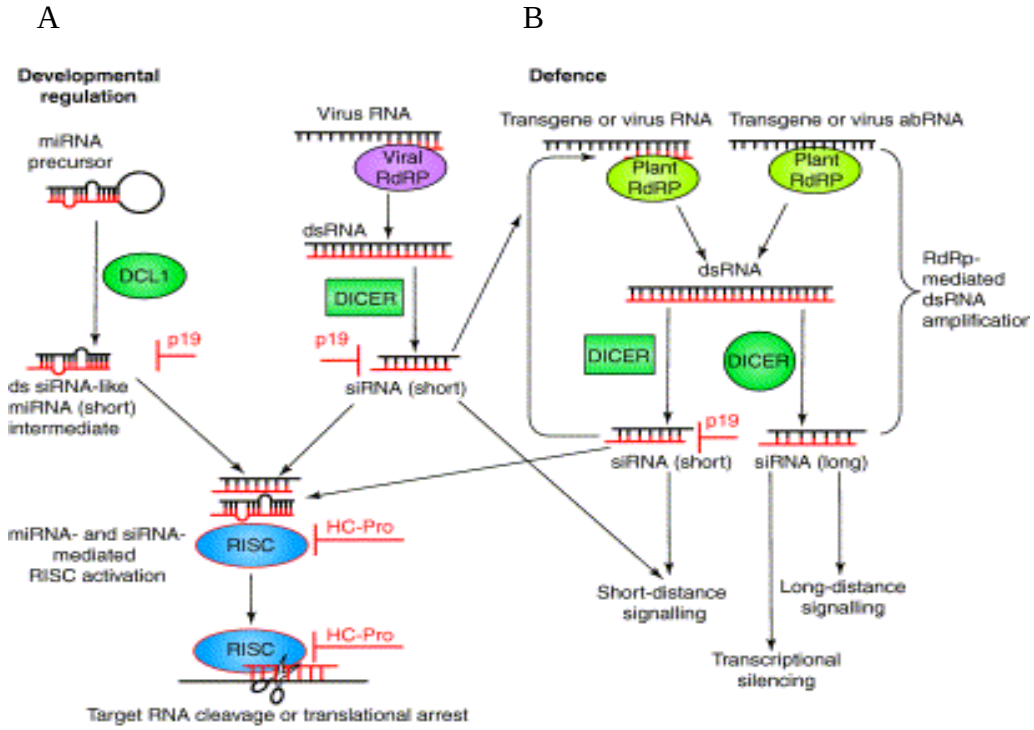


Figure 4. Plant RNA silencing is divided into two main pathways: **(A)** Developmental regulation is initiated with the short, imperfect dsRNA precursors diced by Dicer-like1 (DCL1) into 21-nt ds siRNA-like miRNA intermediates. Then the RISC complex recruits miRs which regulates the endonucleolytic cleavage or translational arrest of target mRNAs by the perfect or near-perfect base pairing between siRNAs and the targeted sequences. **(B)** Defense pathways against parasitic nucleic acids are also initiated by dicing ds or highly structured viral, transgene and transgenic aberrant (ab) RNAs into 21–22 nt (short) and 24–26 nt (long) siRNAs. The viral RdRp produces large amount of ds viral RNAs that are diced to short ds siRNAs. Virus encoded silencing suppressors block different steps of RNA silencing (see 2.2.2 section) such as p19 of tombusviruses and HC-Pro of potyviruses (indicated by blunt-ended red lines). Because p19 efficiently sequesters accessible short siRNAs, it prevents every downstream step of RNA silencing pathways. However, if the unwound miRNAs or siRNAs are already incorporated into RISC, p19 is unable to revert RNA silencing. By contrast, HC-Pro interacts with activated RISC inhibiting target cleavage, which results in accumulation of target RNA and the reversion of RNA silencing (Silhavy and Burgyan, 2004).

1.3.1 Transcriptional gene silencing (TGS)

Transcriptional gene silencing is the result of histone modifications (DNA methylation and changes in the conformation of chromatin), creating an environment of heterochromatin around a gene that makes it inaccessible to transcriptional machinery (RNA polymerase, transcription factors, etc.) (Vaucheret and Fagard, 2001). Thus, gene expression is directly repressed. Such mechanism acts especially during development.

1.3.2 Posttranscriptional gene silencing (PTGS)

Post-transcriptional gene silencing results in specific degradation of translation arrestment of mRNA of a particular gene being degraded or blocked. The degradation of the mRNA prevents translation into an active gene product. The first step for PTGS is the generation of small RNAs (sRNAs). Based on their mode of biogenesis sRNAs can be classified into two categories: siRNAs are processed from long, perfectly double-stranded RNA and miRNAs from single-stranded RNA transcripts that have the ability to fold back themselves to produce imperfectly double-stranded stem loop precursor structures. In the nucleus the primary-miRNA transcript is cleaved by DCL1 with the help of the dsRNA binding protein Hyponastic leaves 1 (HYL1), to produce the shorter precursor-miRNA (pre-miRNA) dsRNA molecule. It is just below the miRNA duplex region of the dsRNA stem loop that the first DCL1-catalyzed cleavage step in the miRNA biogenesis pathway is made (Lu and Fedoroff, 2000). The second cleavage step of the miRNA pathway, which is again directed by the combined action of DCL1 and HYL1, releases from the pre-miRNA stem loop structure the miRNA duplex (Vazquez et al., 2005). Then the second step, that is degradation of mRNA, takes place in the nucleus. The mature single-stranded miRNA is loaded onto AGO1 and is part of an active RNA-induced silencing complex with some other associated proteins (Rana, 2007). RISC uses the siRNA or miRNA as templates for recognizing complementary mRNA and guide the slicer activity of AGO1 to repress the expression

of complementary mRNAs mediated by transcript cleavage (Martinez et al., 2002). RISC cleaves the target mRNA in the middle of the complementary region (Elbashir et al., 2001).

After initiation and degradation of RNA silencing, the last step is propagation of the signal. Systemic silencing, caused by the spread of a silencing signal between cells, occurs in plants and requires RNA-dependent RNA polymerases. RdRps are thought to amplify the silencing effect by dsRNA synthesis of the target mRNAs and/or its cleavage products (Wassenegger and Pelissier, 1998).

1.4. Artificial Resistance

1.4.1 Resistance mediated with R genes

Using natural resistance has been the first approach used by agronomist to control pests. Biotechnological tools allow to introduce specifically and only the resistance gene into crops of interest (Whitham et al., 1996) in only few generations. Strategies to improve resistance for example can involve the overexpression in plants of NBS-LRR proteins (Oldroyd and Staskawicz, 1998).

1.4.2 Resistance mediated by non-viral transgenes

In the use of non-viral proteins to protect plants against viruses, breeders have the choice between the ribosome inactivating proteins (RIP), antibodies and ribonuclease. RIPs are N-glycosidases that inactivate ribosomes by depurination (removing a specific adenine) at the rRNA 60S ribosomal subunits. Successful transgenic resistant plants contain these RIPs or antiviral proteins with similar activities (Hong et al., 1996; Lodge et al., 1993). One side effect of this protein is its toxicity for the plants, inhibiting *in fine* production of proteins. It is possible to limit this negative effect by excretion in intercellular space or under inducible promoter; it is also possible now to separate the antiviral activity from toxic effect of RIPs (Tumer et al., 1997).

Antibodies (Fab, scFV) have been produced successfully in plants (Whitelam and Cockburn, 1996). Resistance acquired by this

technique is assumed to be the viral replication inhibition or during phases of unpacking and/or encapsidation in the first infected cells. Antibodies expressed in the cytoplasm target the CP and confer resistance (Tavladoraki et al., 1993). The low stability of antibodies in the plant cell cytoplasm gives few hopes to use such strategy at large scale in the field.

Finally in the strategy to use non-viral protein, there is some interest about ribonuclease III-like RNase. Found in yeast or bacteria this RNase III specific of dsRNA (dsRNase) could play a role in meiosis inhibition by specific degradation of mRNA (Iino et al., 1991; Rotondo et al., 1997). Using this protein, tobacco and potato plants resistant to ToMV, CMV or PVY have been obtained (Watanabe et al., 1995). However differences in protection have been observed, depending on the virus (Langenberg et al., 1997), that may limit the use of dsRNase.

1.4.3 Resistance mediated by pathogen or pathogen-derived resistance (PDR)

The first demonstration of virus-derived resistance in plants was provided with the coat protein (CP) gene of TMV (Abel et al., 1986). The protection here was explained through the inhibition of virion disassembly in the initially infected cells (Register and Beachy, 1988). The CP-mediated resistance used a fully functional gene of the virus, but in others cases the viral gene used needs to be the dominant negative mutant form. The transgenic expression of viral movement proteins (MP), conferred resistance only when the transgene specified is dysfunctional (Lapidot et al., 1993; Malyshenko et al., 1993). In this case the competition between the natural viral MP and the transgenic mutant prevent a virus spreading. Indeed, protection with one MP of a specific virus can conduct to protection against others viruses (Cooper et al., 1995).

Until now resistance is mediated by viral proteins, however, it exist another level which can also improve the resistance. The RNA or DNA-mediated resistance is due to direct inhibition of the viral

infection cycle by the transgene itself or by its RNA transcript. The idea here is to provide a decoy to the replication machinery that engages to a nonproductive way. There are several types of transgenes that can be used, e.g. satellite RNAs (Gerlach et al., 1987), defective interfering (DI) DNAs (Stanley et al., 1990) or RNAs (Zaccomer et al., 1993). In the case of RNAs protecting plants against virus infection, chimeric genes in which the (+) strand 3'-terminal 100 nt of the noncoding region of the turnip yellow mosaic virus (TYMV) genome were placed downstream from the sense or antisense chloramphenicol acetyltransferase (CAT) coding region. Plants expressing high levels of either chimeric gene showed partial protection against infection by TYMV RNA or virions. This insertion of the non-coding region interfere with viral replication. Protection was overcome when the inoculum concentration was increased. RNA complementary to the initial transcript was detected after infection.

1.4.4 Resistance mediated by gene silencing

In some cases transgenic plants using PDR are actually resistant by the gene silencing process (Lentini et al., 2003; Lindbo et al., 1993). Gene silencing-mediated resistance is now well studied and uses sequences that are shorter than those used in the first applications of the technology, as for example the entire CP sequence used to produce papaya (*Carica papaya*) resistant to *Papaya ringspot virus* (PRSV). This new biotechnological tool provides complete resistance to specific virus and one of first commercial outcome was a transgenic papaya with resistance to PRSV helping to save the papaya industry in Hawaii (Fuchs and Gonsalves, 2007). Indeed the island papaya production had been severely affected. The virus was introduced as early as 1937 and the disease was mild for a number of years until it either mutated or a more aggressive strain was introduced around 1950. Within 12 years, the amount of land under papaya production dropped 94%. Hawaiian papaya production was halved by the end of the decade. Transgenic papaya varieties that are resistant to PRSV entered production in 1998 and resuscitated the industry.

As most plant viruses contain a single-stranded RNA genome, they have been the obvious targets for RNAi technology inspired by nature. Indeed a natural role of the RNA silencing pathway in plants is protection from virus infection (Fusaro et al., 2006). For example transgenic potato have been generated to be resistant to PVY, a potyvirus (Missiou et al., 2004). Recently tomato can be protected from a begomovirus using double and overlapping amiRNA. *Tomato leaf curl virus* (ToLCV) which is generally difficult target for manipulations related to virus resistance, an amiR strategy could be employed to protect plants in an effective manner (Vu et al., 2013).

Controlling virus infection by degrading their RNA within cells can be achieved using amiRs. Niu and collaborators modified a 273 nt sequence of the *A. thaliana* miR159a pre-miRNA transcript expressing amiR against the viral suppressor gene P69 and HC-Pro of *Turnip yellow mosaic virus* (TYMV) and TuMV respectively to provide resistance to these viruses. The application of amiR targeting the gene silencing suppressor of plus-strand RNA viruses can generate specific resistance (Ai et al., 2011; Duan et al., 2008; Niu et al., 2006; Qu et al., 2007). In engineering resistance against plant viruses, the amiR strategy has several advantages as minimal off-target effects, short sequences (21-24 nt) of the amiR. Moreover, the environmental biosafety concerns of viral sequences that might complement or recombine with non-target viruses do not apply to amiRs (Garcia et al., 2006).

2. Fundaments of virology and virus evolution

2.1. Generalities about virus evolution

Knowledge on virus evolution is crucial to the development of efficient and stable control strategies. It is often assumed that viruses evolve rapidly because they have small genomes, short generation times, large population sizes and their RdRps lack of correction activities (Elena and Sanjuan, 2005).

2.1.1 Generation of diversity

The first step for evolution in all organisms is the generation of heritable variation for fitness. The sources of diversity in plant viruses are mutation, recombination, segment reassortment, and migration. The first and initial source of variation, and one defining characteristic of RNA viruses, is their high mutation rate (Drake and Holland, 1999; Sanjuan, 2010) on which natural selection and genetic drift operate. Point mutation is the process that results in differences between the nucleotides incorporated into the daughter strand during nucleic acid replication and those in the template.

Due to their compacted genome, the genetic information has been optimized by natural selection for a specific environment and most mutations are deleterious. Estimation of mutation rates in RNA viruses can vary between 10^{-4} and 10^{-6} mutations per base per generation (Drake and Holland, 1999; Tromas and Elena, 2010). RNA viruses exhibit mutation rates higher than those typical for DNA viruses (Fig. 5) and in the range of 0.03 to 2 per genome and replication round (Chao et al., 2002; Drake and Holland, 1999; Sanjuan, 2010; Tromas and Elena, 2010).

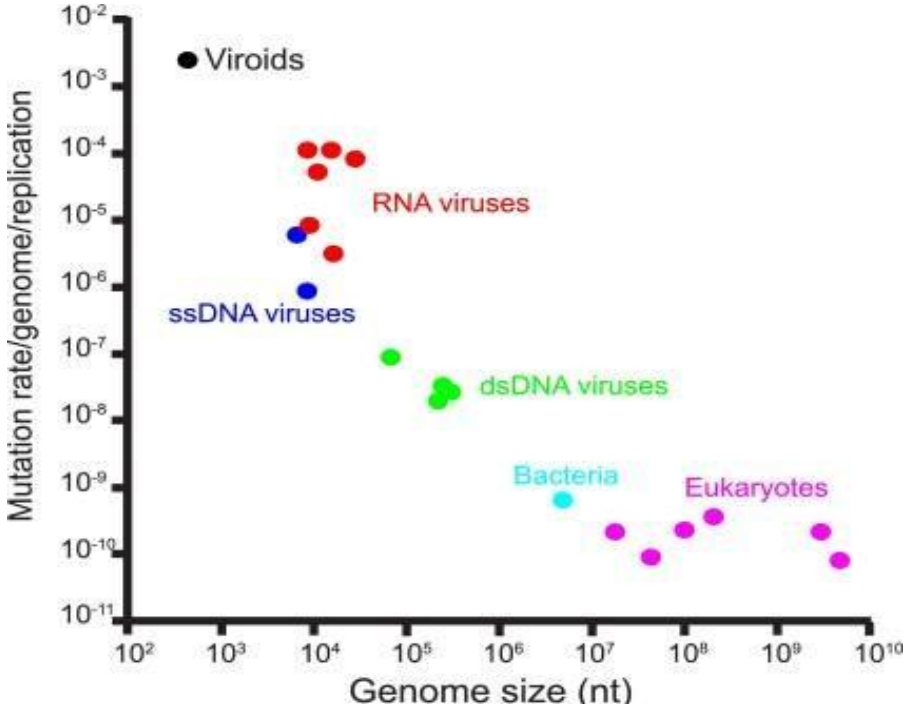


Figure 5. Relationship between mutation rate per nucleotide site and genome size for different genomic systems including viruses (Gago et al., 2009).

Finally the highest mutation rate among all viruses is produced by the positive-sense single-stranded RNA viruses (Fig. 6). Such elevated mutation rate is attributable to the low replication accuracy of viral RdRps that lack proof-reading activity (Barr and Fearns, 2010; Steinhauer et al., 1992). RNA viruses must encode an RdRp, because host cells are not able to replicate long RNA molecules. It is usually thought that low-fidelity of RdRp is not an adaptive trait, but the result of a tradeoff with replication speed (Belsham and Normann, 2008; Elena and Sanjuan, 2005; Furio et al., 2005).

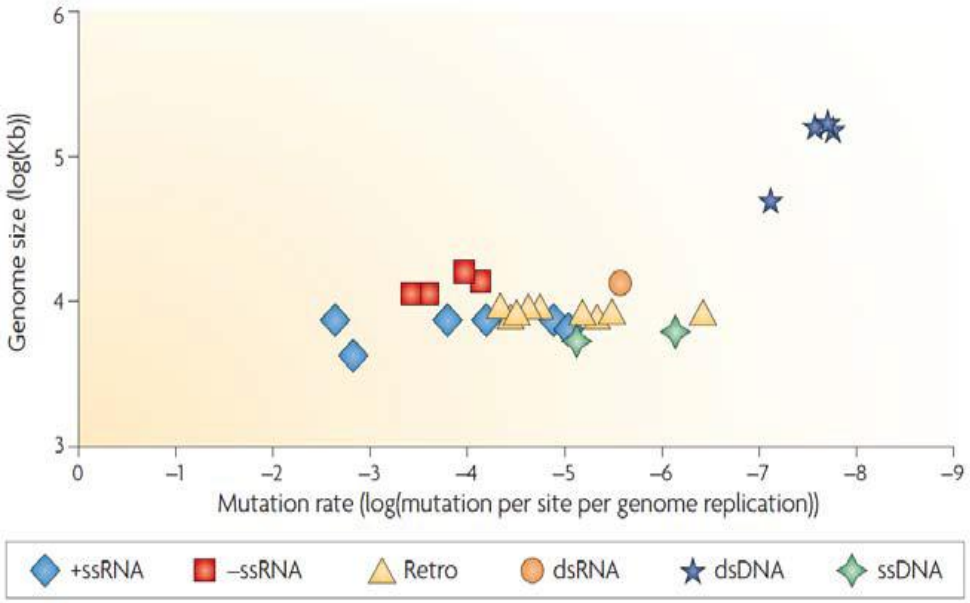


Figure 6. Average mutation rates for different types of viruses. Viruses are denoted by different symbols with respect to their genome type (Duffy et al., 2008).

The actual mechanism of viral replication could also affect mutation frequencies. The two extreme modes of replication are the "stamping-machine" and the "geometric replication" (Duffy et al., 2008). In stamping-machine replication a single virus acts as the template for all progeny genomes, so that mutations accumulate linearly. By contrast, in geometric replication some of the early progeny genomes are themselves used as templates for further progeny, so a mutated template propagates the given error to all its replicate copies (Chao et al., 2002), thereby increasing the rate of mutation accumulation (Drake et al., 1998). Recently it has been demonstrated that the replication mode of TuMV is a mixture of stamping machine and geometric replication, with a much larger contribution of the former (Martinez et al., 2011)

Additionally, some other factors contributing to increased mutation rate have been reported. One of such is the presence of RNA

secondary structures and host or environmental factors (Pita et al., 2007; Schneider and Roossinck, 2001).

The first estimation of mutation rate reported for a plant virus was for TMV at 1.8×10^{-5} substitutions per site and per replication (s/n/r) (Malpica et al., 2002). More recently, estimates have been provided for TEV in the range 2.960×10^{-5} s/n/r (Sanjuan et al., 2009) and 4.754×10^{-6} s/n/r (Tromas and Elena, 2010). In the case of TuMV, mutation rate has been estimated to be 2.640×10^{-6} s/n/r (de la Iglesia et al., 2012).

Using a bayesian approach, the evolutionary rate of a plant virus from the genus *Sobemovirus*: the *Rice yellow mottle virus* (RYMV) have been determined to be between 4 $x10^{-4}$ and 8 $x10^{-4}$ nt/site/year (Fargette et al., 2008).

Additionally, RNA and DNA viruses exhibit recombination which takes place in large virus population and during mixed infections (van der Walt et al., 2009). Recombination is the process by which segments of genetic sequence are switched between the polynucleotide strands of different genetic background during replication (Fig. 7). Experimental studies and phylogenetic analysis on databases of viral gene sequences have revealed that recombination occurs in many families of RNA viruses and it can sometimes have a major impact on their evolution, emergence and epidemiology (Chare and Holmes, 2006). Recombination facilitates major changes in viral genotype and can be a major source of genetic variation for certain viral groups (Garcia-Arenal et al., 2001). Indeed, recombination has been associated with the expansion of viral host range (Brown, 1997; Gibbs and Weiller, 1999), increases in virulence (Khatchikian et al., 1989), the evasion of host immunity (Malim and Emerman, 2001) and the evolution of resistance to antivirals (Nora et al., 2007). Dramatic changes in the biological proprieties at the population level may result from recombination with major epidemiological consequences, including the appearance of resistance-breaking strains or the acquisition of broader host range (Legg and Thresh, 2000; Monci et al., 2002).

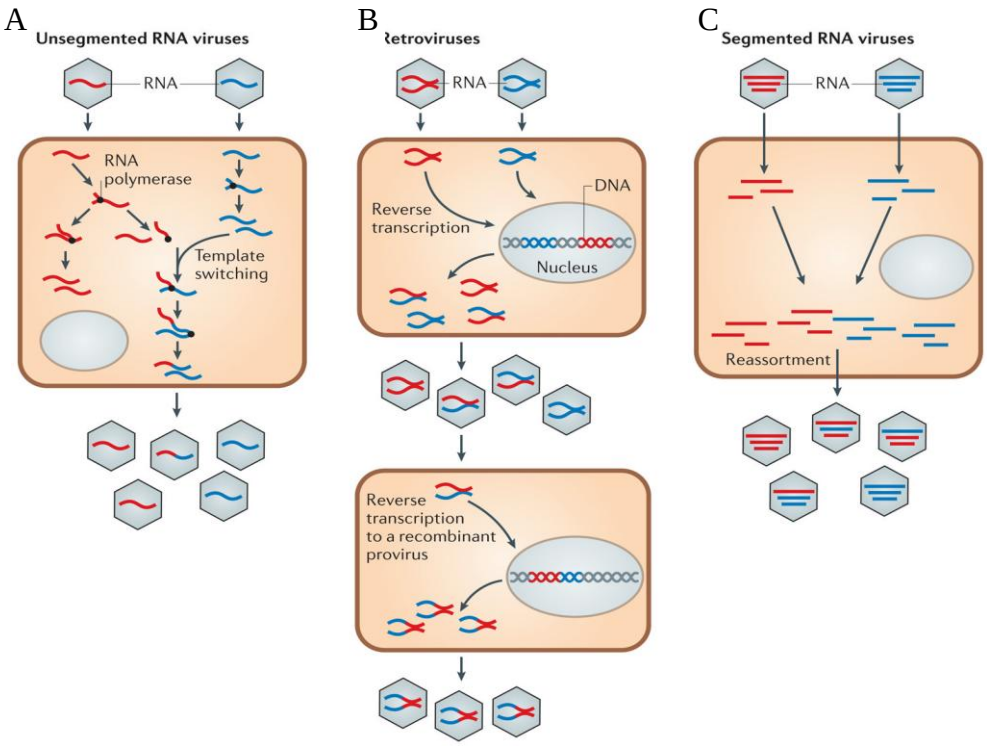


Figure 7. Generation of recombinant and reassortant RNA viruses. **(A)** Generation of recombinant viruses by co-infection of a cell by genetically distinct viral strains. This process can occur in both non-segmented viruses (as shown here) or within a segment of a segmented virus. (**B)** Co-infection of a cell by genetically distinct strains of a retrovirus can lead to the generation of 'heterozygous' virus particles, after which a template-switching event can lead to a recombinant provirus. (**C)** Co-infection of a cell by genetically distinct strains of a segmented virus can generate different combinations of reassortant progeny (Simon-Loriere and Holmes, 2011).

But as mutation rate, RNA recombination occurs at highly variable frequencies among different RNA viruses. For example, in some retroviruses (Mansky, 1998) recombination appears to occur frequently; e.g. in HIV-1, it has been estimated in the range between 1.38×10^{-4} and 1.4×10^{-5} per site per generation (Neher and Leitner,

2010; Shriner et al., 2004). For *Murine leukemia virus* (MLV), the recombination rate was calculated to be close to HIV-1 at 4,6 $\times 10^{-4}$ rate per base per replication cycle (Zhuang et al., 2006). In some (+)ssRNA viruses we can detect also high recombination events, such as enteroviruses (of the family *Picornaviridae*) with 3×10^{-2} mutations/synonymous site/year in the gene encoding viral protein 1 (Savolainen-Kopra and Blomqvist, 2010). In a different category, pararetroviruses which are DNA virus of the family *Caulimoviridae* reach also a high recombination rate on the order of 2×10^{-5} to 4×10^{-5} per base and replication cycle (Froissart et al., 2005). It seems that recombination appears to be far less frequent in other families of (+)ssRNA viruses, including the *Flaviviridae,* (Taucher et al., 2010) only 4×10^{-8} per site per generation were reported during co-infection experiments for HCV.

Noteworthy each study has its own nomenclature as rate/base/replication site (Zhuang et al., 2006), site/generation (Neher and Leitner, 2010) or even by percentage (Urbanowicz et al., 2005). The recombination rate can be calculated experimentally (Taucher et al., 2010) or by phylogenetic studies (Gibbs and Ohshima, 2010; Tomimura et al., 2004) and finally studies can occur only on one small part (Tang et al., 1997) or on the entire part of the genome (Padidam et al., 1999). So it is in general difficult to compare precisely each recombination rate depending on the virus, the method used and which part of the genome is implicated.

These differences in recombination rates among viruses seem to be the result of ecological processes, such as vectors and hosts meet frequently enough for co-infection to occur and differences between viruses in their fundamental biological cycle life.

Certain genome structures could also facilitate recombination. For example, the genomic organization of retroviruses favors the occurrence of genetic exchange. Thus recombination could occur during reverse transcription, by RNA template switching, or after reverse transcription, by breakage and reunion of DNA (Goodrich and Duesberg, 1990). If recombination occurs and results in a successful

genotype, which survives the phase of strong genetic drift until reaching a noticeable frequency in the population and then overcomes other competing genotypes, then we will see the two claimed advantages of recombination over asexual evolution: (i) recombination accelerates the rate at which advantageous genetic combinations are produced (van der Walt et al., 2009) and (ii) allows a more efficient removal of deleterious mutations (FIG. 8) (Kouyos et al., 2007). Experimental studies on *Maize streak virus* (MSV) shows that adaptative recombination is efficient and can yield complex progeny genomes (van der Walt et al., 2009).

The importance of RNA recombination to virus-resistant transgenic plants is that the sequence of viral RNA expressed in a transgenic plant may be available for recombination into a challenging virus. Recovery of viable *Cowpea chlorotic mottle virus* (CCMV) with a deletion in the CP indicates that viral RNA transcribed by the transgenic plant expressing the CP was available to the replicating virus in quantities adequate to support RNA recombination (Greene and Allison, 1994).

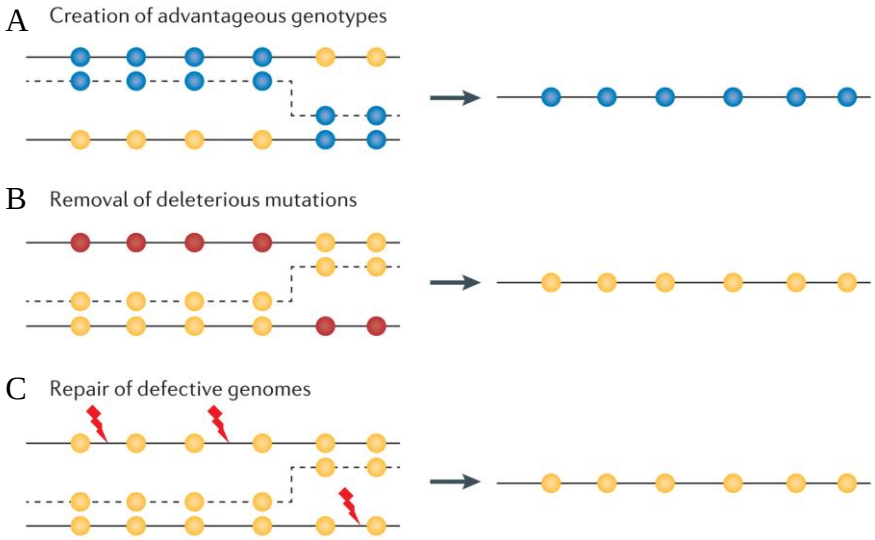


Figure 8. Depending on the acceptor and donor genotypes, and the position of the template switch, recombination can have several positive effects on the genome. Yellow circles indicate wild-type loci. (**A)** Recombination can create advantageous combinations of mutations (blue circles) that increase the rate of adaptive evolution compared with mutation alone, or it can disassociate advantageous and deleterious mutations, allowing the former to spread. (**B)** Recombination can remove deleterious mutations (red circles) and restore the wild-type (fit) genotype, which can lead to a selective advantage for recombination if deleterious mutations occur frequently enough and interact synergistically. (**C)** Recombination can also generate a functional genome from damaged parental molecules. Genetic damage, such as strand breaks or oxidative base modifications, are represented by red lightning symbols (Simon-Loriere and Holmes, 2011).

As for beneficial mutations, recombination events can be selected during evolution and play a role in the emergence of viruses as for example *Geminivirus* (Padidam et al., 1999) and *Tobravirus* (MacFarlane, 1997). The largest family of plant viruses, the *Potyviridae* share between all species the P1 protein. P1 is the first protein of the synthesized polyprotein and the most divergent with

regard to both length and amino acid sequence (Adams et al., 2005). From all recombination events in *Potyviridae*, some affect the P1 sequence in intra and interspecies (Desbiez and Lecoq, 2004; Larsen et al., 2005; Tan et al., 2004; Tugume et al., 2010). Moreover Valli et al. (2007) reported intergenic recombination within the P1 gene affecting the genera *Ipomovirus* and *Potyvirus*. This example illustrate that recombination is one of main forces driving evolution of plant viruses across different genera. Finally evolutionary pathway was designed for potyviruses (Fig. 9) showing frequent recombination events in this genus. TuMV belonging to this genus presents also multiple recombination events (Tomimura et al., 2004). This observation, as the high mutation rate, has to take account in the context of mixed infection potentially generating new variants and leading to breakdown resistance.

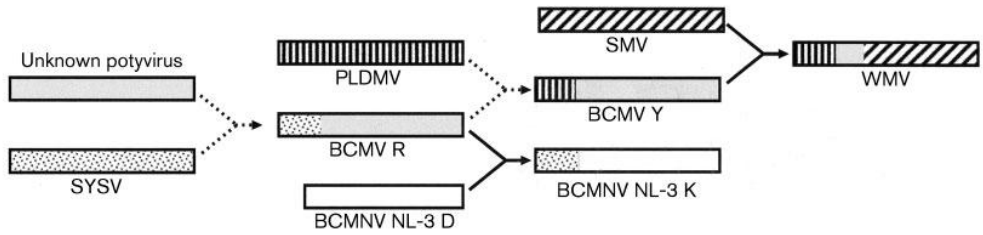


Figure 9. Recombination events at the P1 coding sequence involved in the generation of BCMV-related viruses. Evolutionary pathway proposed for BCMV-related viruses. Boxes represent the P1 proteins and the different shadings indicate the viruses that might supply the different regions. The recombination event involving potyviruses outside the BCMV subgroup is shown with a dashed arrow (Valli et al., 2007).

Multipartite viruses have an additional opportunity for exchanging genetic material during mixed infection: segment reassortment. There is evidence that reassortment occurs in natural population of plants viruses (Fraile et al., 1997) and it may play an important role in virus evolution (White et al., 1995). For example, it is thought that the broad host range and worldwide distribution of

CMV is in part due to frequent reassortment events, contributing to its large evolutionary success (Roossinck, 2002).

Finally migration of vectors and viruses allow a local genetic variation in geographical areas determined. Migration is an evolutionary force which increases the population's genetic variability and can even influence fitness of the population. For example an *in vitro* cell system on *Vesicular stomatitis virus* (VSV) clearly demonstrated a positive correlation between the migration rate and the magnitude of the mean fitness reached by the virus quasispecies populations (Miralles et al., 1999). New individuals from faraway population bring new alleles increasing the genetic pool. Migration implies that single host populations will be invaded by new, different virus genotypes. Competition for the limited number of host cells will occur between the immigrant and the resident virus quasispecies populations. Under this situation, fitness levels increase and the virus strain with the higher basic reproductive rate will overtake the less-virulent strains (Bremermann and Thieme, 1989; Frank, 1992; Mosquera and Adler, 1998; Nowak and May, 1994).

2.1.2 Key evolutionary concepts

Two main evolutionary forces operate upon genetic variability to shape the genetic composition of viral populations: genetic drift and natural selection (Elena et al., 2011). The key parameter driving natural selection in a deterministic manner is differences in Darwinian fitness. By contrast, genetic drift results from random sampling events.

The fitness of a genotype measures its ability to reproduce itself and is defined as the number of viable progeny produced by a genotype that reproductively contributes to the next generation. Indeed from a pure deterministic point of view only genomes with the higher fitness will be represented in the future population. The fittest individual genome in a population is generally set to $W = 1$ and the non-reproductive ones are set to $W = 0$. Fitness represents the combined effects of all other phenotypic properties on the capacity for

the survival and reproduction by a particular genotype in a particular environment (Lenski, 1991). That is why several parameters as the cycle life of the virus (cell entrance, viral reproduction rate (Ro), transcription, encapsidation, movement....), resistance to antiviral responses and transmission or adsorption rates are crucial parameters in many epidemiological models of infectious diseases.

All possible variations of a given genome can be represented in a multidimensional geometric space, where each vertex in the space corresponds to one of these alternative genomic configurations. On this genotypic space is over imposed an additional dimension, the fitness of each of the possible genomes in a given environment (Fig. 10). This corresponds to what is known as the genotypic fitness space.

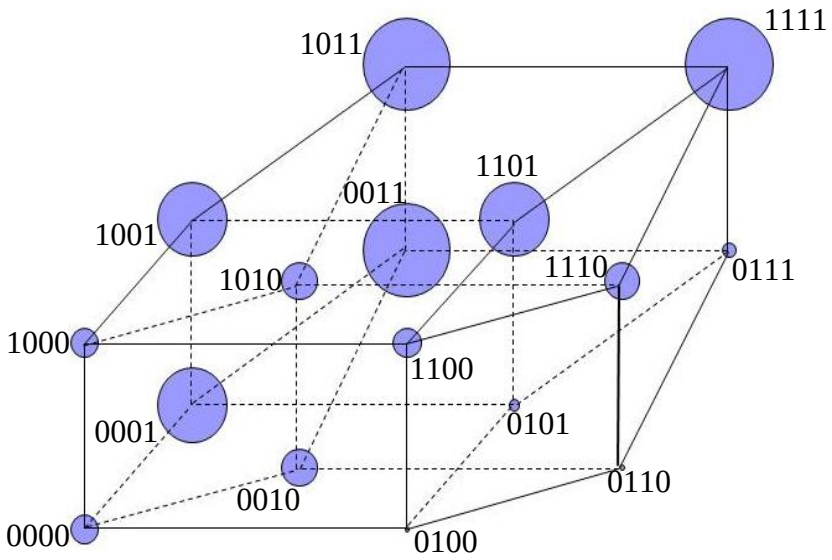


Figure 10. Schematic representation in a hypercube of all possible genomes with 4 genes, each gene has 2 allelic forms (0 or 1). So all possible genome equal to 4^2 are positioned at each vertex of the hypercube. Each genome has its own fitness represented by the size of the blue circle.

It is also common, and perhaps more visually obvious to represent the fitness landscape as a three-dimensional projection of the above multidimensional landscape (Fig. 11). In this alternative representation, the (*x*,*y*) plane contains all possible genotypes whereas the *z* dimension corresponds to the fitness of these genotypes. In such a view, peaks and valleys of fitness appear (Fig. 11).

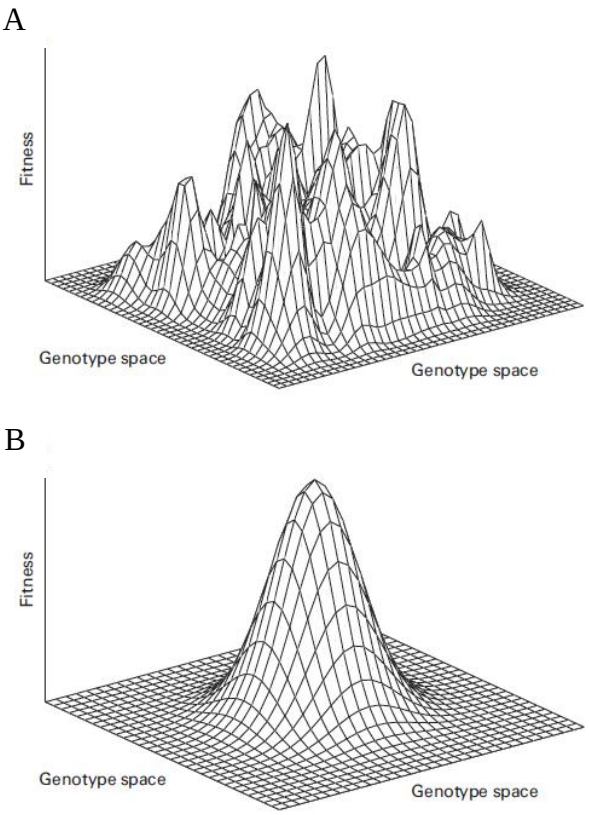


Figure 11. Three-dimensional representation of a fitness landscape in a particular environment as a function of its genetic composition. Two types of classical fitness landscapes are presented: **(A)** a rugged landscape and **(B)** a single peaked landscape (Gavrilets, 2010).

(a) Selection

Selection is the only evolutionary force that produces adaptation, that is, a constant increase in fitness until reaching an evolutionary optimum. It is a deterministic process that entirely depends on the rank of fitness values of competing genotypes. For example if one gene have 2 alleles each with a selective value s and 1-s. Indeed in an infinite panmictic population, if an allele have a fitness $W = 1$ it will invade the population and will be fixed, with a fitness $W = 0,5$ the allele is considered as neutral and do not evolve from its initial frequency and finally $W = 0$ the allele will be removed quickly of the population (Fig. 12).

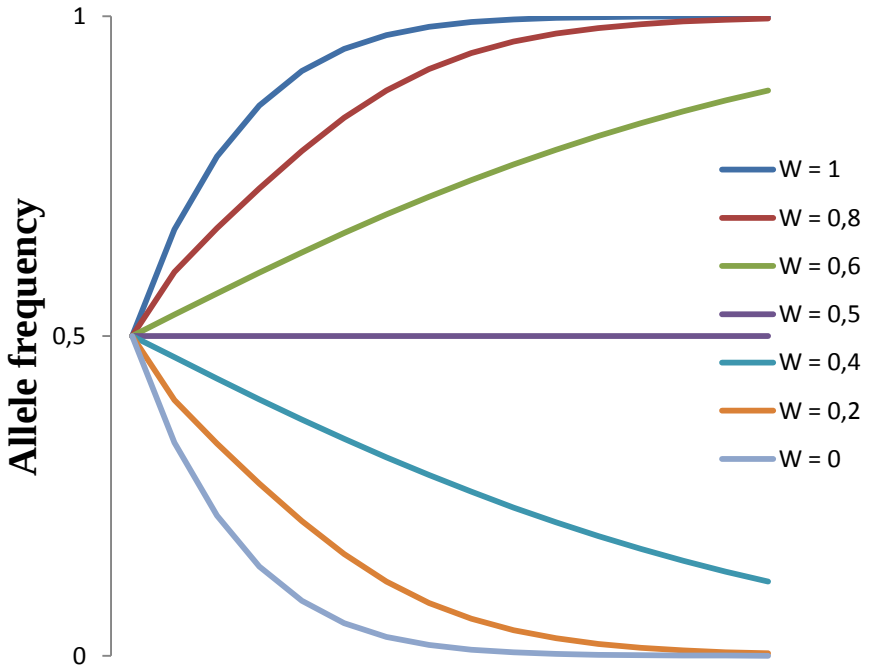


Figure 12. Representation of the evolution of alleles with different fitness values in a infinite panmictic population. At the generation zero each allele starts with frequency equal to 0,5. Each line represent an allele with a fitness *W* (between 0 and 1).

Plant viruses have co-evolved with their wild hosts in communities highly heterogeneous long before any plants were domesticated. This co-evolutionary process shaped both plants hosts and viruses (Cooper and Jones, 2006; Lovisolo et al., 2003). Humans have also shaped virus evolution, indeed since agriculture was invented and spread around the world during the Holocene period, the conditions generated during this era have favored particular viruses (Gibbs et al., 2010; Jones, 2009).

In a constant environment, evolution depends on the distribution of the mutational effects on fitness. A mutation can be beneficial, neutral, deleterious or lethal in a particular environment (Lalic et al., 2011). Most mutations have deleterious impact on organism's fitness (Fisher, 1930). The pervasiveness of spontaneous accumulation of deleterious mutations has been demonstrated in mutation-accumulation experiments under minimal effect of purifying selection and strong bottlenecks (Chao et al., 2002; de la Pena et al., 2000; Duarte et al., 1992; Elena, 1999; Escarmis et al., 1996; Mukai et al., 1964; Muller, 1964) that onset Muller's ratchet (Muller, 1964). As viruses have compact genomes, for a well-adapted virus most of mutations are expected to be deleterious or lethal. Studies using direct mutagenesis approach confirmed the prevalence of deleterious mutations and additionally revealed the large proportion of lethal mutations (Carrasco et al., 2007; Domingo-Calap et al., 2009; Sanjuan et al., 2004). Since neutral mutations have no effect on fitness (or their effect in the experimental assay is too small) (Burch et al., 2007; Kimura, 1991), natural selection does not operate on them, so they can be fixed in a population trough the action of random genetic drift (Carrasco et al., 2007; Coleman et al., 2008; Cuevas et al., 2012; Marsh et al., 2008). It is often thought that synonymous sites are under neutral or weak selection, in many instances, it has been demonstrated that synonymous substitutions are not neutral (Plotkin and Kudla, 2011) and as such may have a major effect on RNA virus fitness (Carrasco et al., 2007; Coleman et al., 2008; Cuevas et al., 2012;

Marsh et al., 2008). Even beneficial mutations are rarely observed, there are also thought to be extremely rare. Hence, in a population, a newly arisen beneficial mutation will have a small probability of being fixed in a finite population only by means of natural selection (Patwa and Wahl, 2008). It has been argued that the existence of beneficial mutations of small effect rather than large is more probable (Orr, 1998). Increases in population size and mutation rate can cause larger-effect beneficial mutations to become fixed (Orr, 2000). A deterministic model using data from experimental evolution studies in microbes was recently described (Sniegowski and Gerrish, 2010), which suggests that beneficial mutations may actually become abundant under periodic selection in combination to high mutation rate. More specifically, virus populations show relative stasis in natural host, but evolve rapidly in a new host (Novella et al., 1999). This assumes that virus population within its reservoir host has reached the global fitness peak characterized by strong purifying selection that operates within host and eliminates unfit mutants in a stable environment (single-host evolution).

Given the high mutation rates, one genome can have several points mutations (Malpica et al., 2002; Tromas and Elena, 2010). Existence of numerous interactions among genes or mutations in determining phenotypes, *i.e.*, epistasis, has vastly been demonstrated (Phillips, 2008). Moreover these interactions can be host dependent (Lalic and Elena, 2012).

(b) Genetic Drift

Genetic drift is a stochastic phenomenon which can lead to unpredicted result (Fig. 13). Indeed drift is stronger in small populations and during bottlenecks, e.g. when only few individuals colonize a new territory/infect a new host; a process known as founder effect.

Populations of plant viruses can reach very large sizes during colonization of the host (Harrison, 1956). In contrast the number of founders able to initiate a new cycle of infection during horizontal

transmission can be really low (Ali et al., 2006; Hall et al., 2001; Sacristan et al., 2003), and the mean number of founders per infection can under some circumstances be as low as 1 (Moury et al., 2004; Zwart et al., 2011).

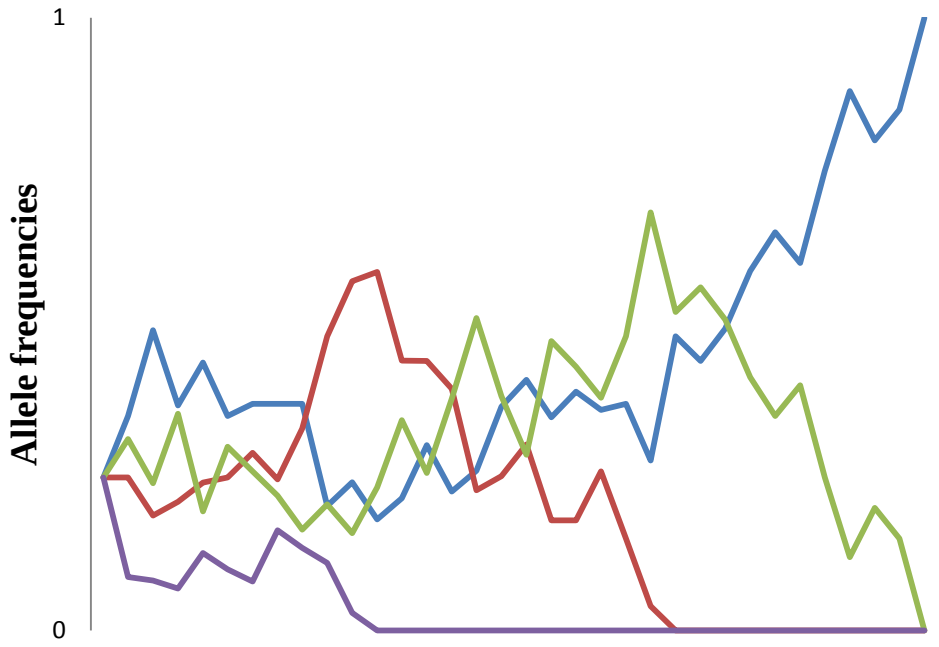


Figure 13. Representation of evolution of neutral allele frequencies. At the generation 0 each allele of a same gene start with an equal frequency of 0,25. Without selection to drive the allele fate, each frequency is determined by one side the frequency at the precedent generation and by random events. Even if all allele are equals during generations, some extinct (as the purple, red and the green allele) and other are fixed in the population (as here the blue one).

In Population Genetics, the strength of genetic drift is measured by effective population size (N_e), which for a population of fluctuating size can be approximated by the smallest size in de fluctuation (Wright, 1931). Hence, for plant viruses, $N_e = 1$ may be quite common and thus the effect of drift of extreme importance. This transition from a large population to few individuals is named genetic bottleneck. Genetic bottlenecks play an important role; the sampling of a small number of individuals from a parental population in which many genetic variants may be present can lead to a profoundly different genetic composition in the new population (Betancourt et al., 2008).

To summarize, viral populations have the potential to evolve rapidly and this evolution can be influenced by diverse factors (selection, migration, genetic drift…) some exerting a directional force during evolution and others being more stochastic (Fig. 14). This capacity to a single genome to evolve in distinct population have been illustrated in a perennial host. Indeed a single event of inoculation in *Prunus* tree of *Plum pox virus* (PPV) gives distinct viral populations which differentiate and evolve independently in the differents parts of the tree (Jridi et al., 2006).

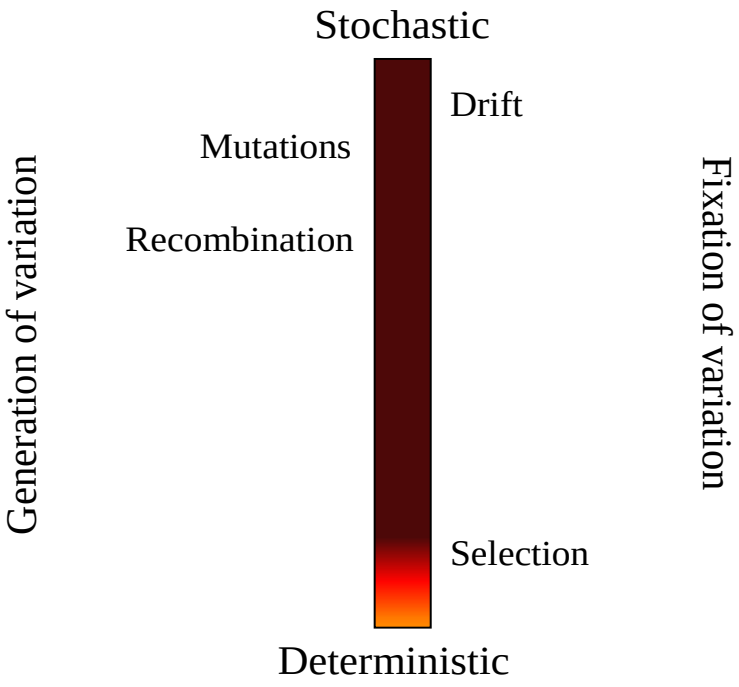


Figure 14. Evolutionary processes continuum, from stochasticity to determinism inspired from (Koonin and Wolf, 2012).

2.1.3 Fitness distribution

The distribution of fitness effects is central to many questions in evolutionary biology, including the rate of genomic decay due to Muller's ratchet (Loewe, 2006), the maintenance of genetic variation at the molecular level (Charlesworth et al., 1995), and the evolution of sex and recombination (Peck et al., 1997). Different mathematical distributions have been used to model the distribution of fitness effects. The most common among these are distributions belonging to the heavy-tail family, for example the gamma or the Weibull distribution. This is a flexible distribution that has two parameters: a shape parameter, *k*, and a parameter that governs the mean of the distribution (Fig. 15).

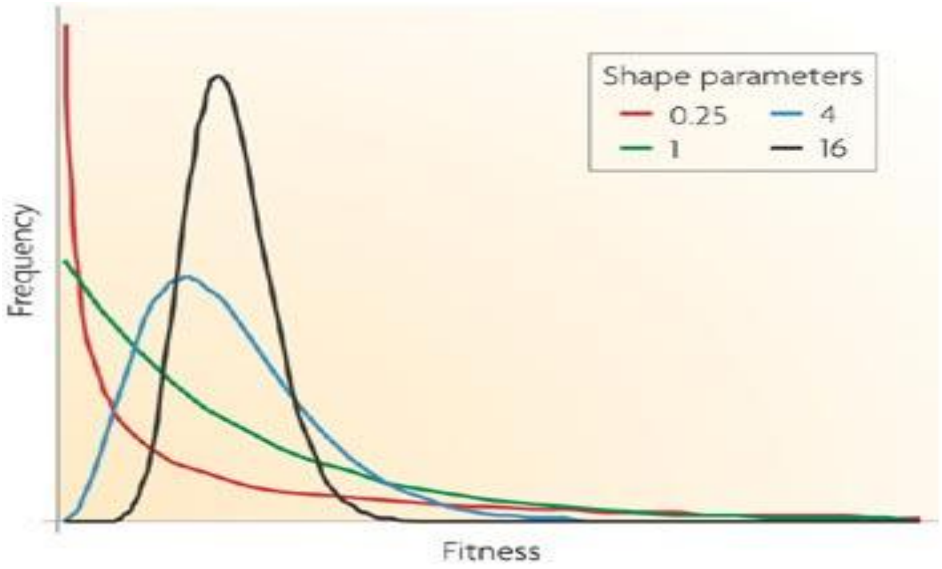


Figure 15. Different gamma distribution. The figure shows the probability density of the gamma distribution with a mean of 1 and varying shape parameters (Eyre-Walker and Keightley, 2007).

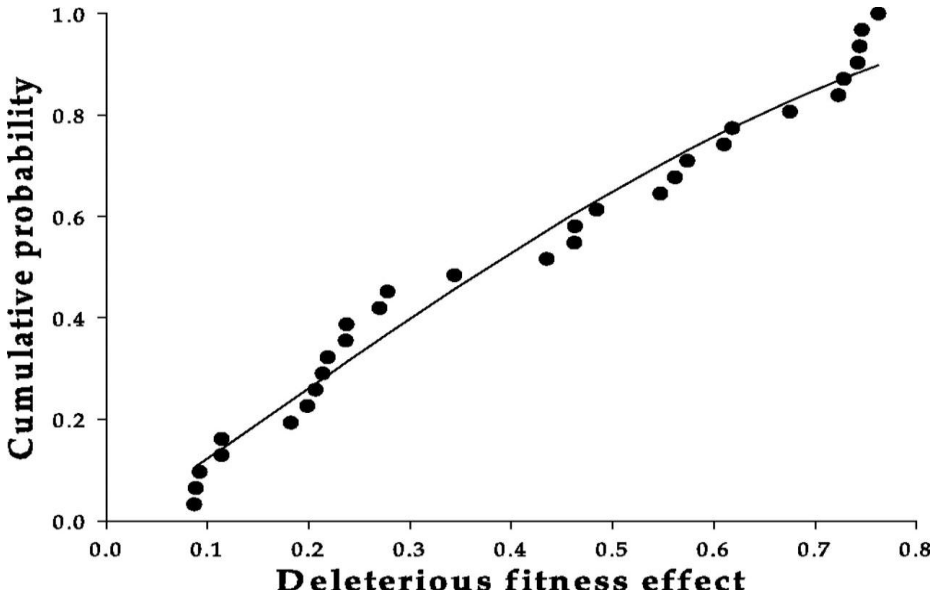


Figure 16. Fit of the beta distribution to the observed distribution function for mutational fitness effects (Carrasco et al., 2007).

Mutations appear randomly in the genome but theirs single effect is determined in a specific environment. In a study on the effect of single point mutations in TEV, Carrasco et al. (2007) found a majority of mutations give a lesser fitter genome than the original clone (Fig. 16). Indeed a collection of 66 clones of TEV, each carrying a different, randomly chosen, single-nucleotide substitution shows that among all mutations, 40.9% were lethal, and among the viable ones, 36.4% were significantly deleterious. Not a single case of beneficial effects was observed and 22.7% are estimated neutral in the range of the median and the level of resolution of measures.

2.1.4 Gene duplication

Finally one possible way by which viral genomes may evolve is by gene duplication. It can be defined as any duplication of a region of DNA. The newly acquired second copy of the gene is often free from selective pressure. Indeed mutation can be deleterious in a single copy gene context but maybe neutral when the gene is redundant. Thus over generations it accumulates mutations faster than a functional single-copy gene. Another view is that both copies are equally free to accumulate degenerative mutations, so long as any defects are complemented by the other copy. This leads to a neutral "subfunctionalization" or duplication-degeneration-complementation (DDC) model (Force et al., 1999; Stoltzfus, 1999), in which the functionality of the original gene is distributed between the two copies (Fig. 17). The two genes that exist after a gene duplication event are called paralogs and usually code for proteins with a similar function and/or structure. By contrast, orthologous genes are the ones that code for proteins with similar functions but exist in different species, and are created from a speciation event. Duplication events can in theory allow the new gene to acquire new functions (Zhang, 2003). One of the initial copies may accumulate mutations that finally conduct to encode a new protein. This copy becomes a new gene in the process called "neofunctionalization" (Fig. 17).

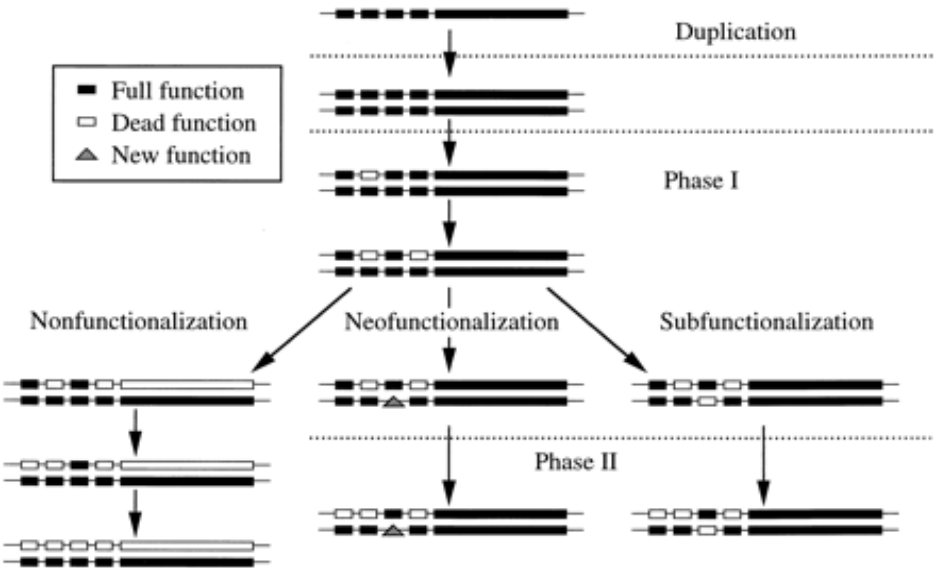


Figure 17. Three potential evolution of duplicate gene with multiple regulatory regions. Regulatory elements with unique functions are noted as small boxes, and the large boxes represents transcribed regions. Solid boxes are intact regions of a gene, while open boxes present null mutations, and triangles show the evolution of a new function. Because the model focuses on mutations fixed in populations, the diagram shows the state of a single gamete. In the first two steps, one of the copies acquires null mutations in each of two regulatory regions. On the left, the next fixed mutation results in the absence of a functional protein product from the upper copy. Because this gene is now a nonfunctional pseudogene, the remaining regulatory regions associated with this copy eventually accumulate degenerative mutations. On the right, the lower copy acquires a null mutation in a regulatory region that is intact in the upper copy. Because both copies are now essential for complete gene expression, this third mutational event permanently preserves both members of the gene pair from future nonfunctionalization. The fourth regulatory region, however, may still eventually acquire a null mutation in one copy or the other. In the center, a regulatory region acquires a new function that preserves that copy. If the beneficial mutation occurs at the expense of an otherwise essential function, then the duplicate copy is preserved because it retains the original function (Force et al., 1999).

The copy can also be preserved and without accumulating deleterious mutations, conserving the original function. This conservation can be explained when increasing the number of copies (that is, gene dosage) gives a certain advantage to the organism. This scenario can be placed in the context of resistance. For example parasites Plasmodium falciparum is treated with mefloquine. *pfmdr1*, a gene encoding a parasite transport protein, is the best overall predictor of treatment failure with mefloquine. Moreover increased copy number of *pfmdr1* was the most important determinant resistance to mefloquine (Price et al., 2004).

2.2. **Mechanisms of resistance-breaking**

A durable resistance is defined as the one that remains effective in a cultivar for a long period of time during its widespread cultivation (Johnson, 1979).

The probability of apparition of resistance-breaking mutants and their spread among otherwise resistant hosts is a complex process that depends, among others, on the fitness of the escape-mutants in the absence of host resistance, the probability of generating a large diversity, the type of resistance encountered, and the number of resistance genes present (Harrison, 2002). In short, the evolutionary potential of plant viruses may be an important determinant of the durability of resistance (Garcia-Arenal and McDonald, 2003).

For plant viruses the cost associated with resistance breakdown can be high for several virus (Carrasco et al., 2007; Fraile et al., 2011; Jenner et al., 2002; Sacristan and Garcia-Arenal, 2008; Sanjuan, 2010). TuMV for example have natural variants (CZE1, CDN1) able to overcome the resistant gene *TuRB01*. In coinfection in susceptible plants, these variants are dominated by the non-resistant UK1 isolate (Jenner et al., 2002).

It is also known that as few as one or two mutations are often enough to break a resistance down (Harrison, 2002; Kang et al., 2005; Lecoq et al., 2004). Fitness costs and the number of mutations determine the equilibrium frequency of resistance-breaking alleles in a

virus population and corresponds to the selection-mutation balance (Ribeiro et al., 1998). The evolutionary dynamics of virus population is also shaped by the presence of susceptible crops and wild plant species which act as a 'reservoir' by providing a 'green bridge' between them and resistant crops (Burdon and Thrall, 2008).

The choice of the resistance gene is the most important factor for durability, and optimal strategies of resistance deployment depend of the resistance characteristics and the epidemiological context (Fabre et al., 2012). Indeed the fitness penalty associated with virulence acquisition can be involved in the durability of resistance. This was demonstrated experimentally for the *Xanthomonas*-rice pathosystem (Vera Cruz et al., 2000). However, this state can be transitory due to compensatory mutation which can increase the fitness of the escape variant (Garcia-Arenal et al., 2001). Another parameter related to resistance durability is the number of mutations required for virulence acquisition, the greater of mutations needed for virulence the more durable was the resistance (Harrison, 2002). So no universal strategy exists. Overall, two broad categories can be used: 'mixture' and 'purely resistant' strategies.

When resistance breakdown occurs rapidly, strategies that mix susceptible and resistant cultivars are optimal. Such strategies actually make a compromise between maximizing yield; and minimizing the probability of resistance breakdown. First, for any epidemic intensity, mixture strategies are optimal for low fitness costs of resistance breakdown, when the equilibrium frequency of the escape variant in susceptible hosts is high. Second, whatever the resistance gene considered, mixture strategies are optimal in landscapes with high epidemic intensities. With no fitness costs, modeling gene-for-gene interactions between crop plants and a fungi-like pathogen during a single season intermediate cropping ratios are optimal (Ohtsuki and Sasaki, 2006). Pure strategies with up to 100% of resistant cultivar can also be optimal. This arises when the pathogen population is unlikely to be invaded by the breaking down resistance variant, typically when two mutations with average fitness costs (for RNA viruses) are

required for resistance breakdown in landscapes with intermediate, or lower, epidemic intensity. This strategy can also be deployed if the reservoir initially hosts very few escape variants, causing few infections of resistant plants and so the viral dynamics in the reservoir respond slowly to the selection pressure exerted by the resistant cultivar (Fabre et al., 2012). With different hypotheses (e.g. durability measured by the time until invasion of the resistance breaking pathogen, without immigration), other mathematical models also demonstrated the value of high cropping ratios (van den Bosch and Gilligan, 2003).

2.2.1 Escape mutants

In many cases, only few point mutations are sufficient to avoid animal or plant resistance. RNA interference (RNAi), a gene-silencing mechanism in animal systems similar to that of miRNA, has been used in clinical trials as antiviral therapeutics to inhibit replication of several human pathogenic viruses (Gitlin et al., 2005; Haasnoot et al., 2007). A major problem of RNAi-mediated antiviral therapies is the emergence of resistant virus variants, which differ from the wild type virus by having fixed point mutations in the target sequence leading to imperfect matching as demonstrated for HIV-1. These escape mutants are not properly processed by the enzymatic silencing machinery (Boden et al., 2003; Westerhout et al., 2005). The RNAi machinery is still efficient for some mismatches within the target sequence whereas other mismatches, such as those in the central region (position 9 to 11) of the target sequence, compromise RNAi-guided antiviral therapies (Elbashir et al., 2001; Westerhout and Berkhout, 2007). To explore the potential of each position to avoid plant resistance Lin et al. (2009) established a heterologous-virus resistance system using a TuMV-GFP viral vector to carry a non-essential 21-nt sequence. This heterologous-virus system allowed to modify any nucleotide within the 21-nt target site without altering virus coding sequences and thus without affecting replication and activity. In other words, this system allowed separating the selective pressure imposed by protein

functionality from the selective pressure imposed at the sequence level by RNA silencing. Using an *in vivo* assay critical positions on the 21-nt target sequence for RISC-amiRNA-mediated cleavage were identified. Scanning mutations on the 21-nt target site of the challenging chimeric virus showed different degree of resistance breakdown on amiR159-P69 transgenic plants (Fig. 18).

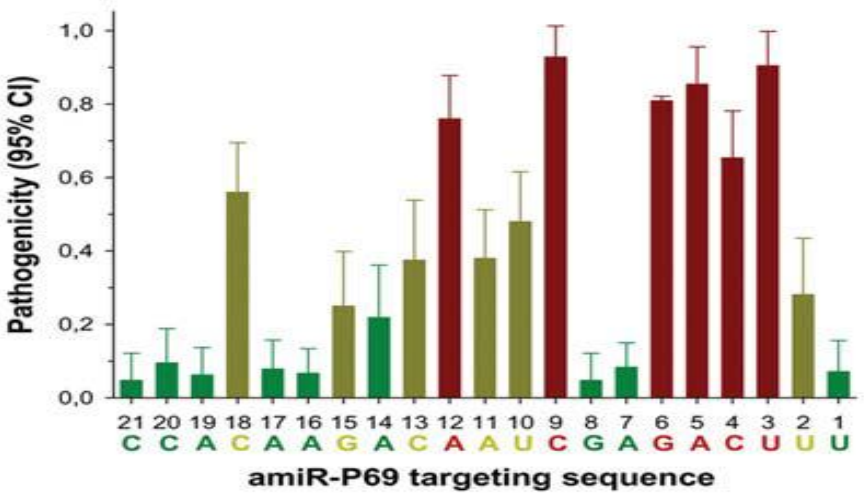


Figure 18. Critical positions within the amiR target site. 21-nt RNA sequence is shown on the x-axis. The degree of resistance breakdown was represented as the ratio of inoculated plants with viral disease symptoms. Red bars represent critical positions for resistance; yellow bars represent positions of moderate importance; green bars represent positions of minimal influence in resistance-breakdown (Lin et al., 2009).

A major result of this study was to determine the most critical positions for resistance-breaking. These critical positions were localized on sequences complementary to the 5' portion of the amiR whereas positions of moderate effect were mainly localized on the central region of the target site (Fig. 18). Imperfect pairing with central mismatches at position 11 and 12 in small RNA-target hybrids promotes translational repression because it excludes slicing. Although the RNAi machinery tolerates changes in some positions of

the 21-nt target, it is sensitive to changes in some others, particularly at the center of the target site (Elbashir et al., 2001; Lin et al., 2009; Westerhout and Berkhout, 2007).

Translation initiation factors, and particularly the eIF4E and eIF4G protein families, are essential determinants in the outcome of RNA virus infections. For most of the potyviruses and for RYMV, the gene implicated in the recognition of the virus by the plant has been identified as the virus genome-linked protein (VPg) (Robaglia and Caranta, 2006). In most plant–potyvirus pathosystems, amino acid changes in the VPg have been shown to be responsible for the ability of the virus to overcome eIF4E-mediated resistance (Ayme et al., 2006; Charron et al., 2008; Kang et al., 2005).

Mutation stepwise pathway of RYMV which need to pass through defined mutation to access to the resistance breakdown (Pinel-Galzi et al., 2007) and the residual multiplication of the avirulent isolates on *rymv1-2* resistant plants (Poulicard et al., 2010) suggest that resistance-breaking mutations are generated *de novo* in resistant plant. This result underlines the importance of selection pressure on new mutations in a resistance context and the potential of evolution of viruses.

2.2.2 Suppressor of silencing, viral-encoded suppressors of RNA silencing (VSRs)

Viruses have also evolved to evade gene silencing by a variety of strategies. To counteract RNA silencing, viruses express suppressors that interfere with siRNA as well as miRNA pathways (Lakatos et al., 2006). Currently, more than 40 VSRs have been characterized (Table 3).

Table 3. Viral suppressors of RNA silencing (VSRs) of plant viruses.

Genome	Genus	Virus	VSR	Suppressed RNA silencing mechanism inhibition	Other biological function(s)	References
DNA	Begomovirus	ACMV	AC2	siRNA binding to RISC	Transcriptional activator protein (TraP)	(Voinnet et al., 1999)
			AC4	Adenylate kinase (ADK)	TraP	(Vanitharani et al., 2004)
		TGMV	AL2	ADK	TraP	(Wang et al., 2003)
		TYLCV-C	C2	ADK2	TraP	(Dong et al., 2003)
(+) RNA	Curtovirus	BCTV	L2	ADK	TraP	(Wang et al., 2005)
	Carmovirus	TCV	p38	Dicer cleavage	Encapsidation, movement	(Thomas et al., 2003)
	Closterovirus	BYV	p21	Bind siRNA	RNA accumulation, Encapsidation	(Reed et al., 2003)
		CTV	p20	siRNA-RISC formation	RNA accumulation, Encapsidation	(Lu et al., 2004)
	Cucumovirus	CMV	2b	AGO1 slicer, siRNA signaling	Movement, anti-SA defense response	(Brigneti et al., 1998)
	Hordeivirus	BSMV	γb	-	Movement, seed transmission	(Yelina et al., 2002)
	Polerovirus	BWYV	P0	Degrade slicer	Pathogenesis related	(Pfeffer et al., 2002)
	Potexvirus	PVX	P25	Signaling	Movement, RNA helicase	(Voinnet et al., 2000)
	Potyvirus	PVY	HC-Pro	Stability of siRNA, siRNA-RISC formation	Aphid transmission, proteinase, movement	(Brigneti et al., 1998)

	Sobemovirus	RYMV	P1	Accumulation miRNA	Movement, accumulation	(Sire et al., 2008)
	Tobamovirus	TMV	126 K	-	RdRp, movement	(Kubota et al., 2003)
	Tombusvirus	TBSV	p19	Signaling, siRNA-RISC formation	Movement	(Lakatos et al., 2004)
(-) RNA	Tenuivirus	RHBV	NS3	-	-	(Bucher et al., 2003)
	Tospovirus	TSWV	NSs	-	Movement	(Takeda et al., 2002)

Indeed many of these suppressors have been identified as pathogenicity factors or as viral cell-to-cell or long-distance movement proteins. They comprise a variety of viral-encoded proteins that interact with various silencing components at different stages in the silencing pathways. VSRs proteins encoded by unrelated RNA and DNA viruses have no similarity to each other in either coding sequence or protein structure, suggesting separate origins and variable functional mechanisms for each suppressor type. Indeed most VSRs are multifunctional proteins, besides being RNA silencing suppressor they can operate as coat protein, replicase, movement protein, helper-component for viral transmission, protease or transcriptional regulators. This is particularly true for the HC-Pro proteins of potyviruses which have others proprieties such as aphid transmission or cell-to-cell movement as well as proteolytic processing of the viral polyprotein (Plisson et al., 2003). The p19 of *Tomato bushy stunt virus* (TBSV) and the 2b proteins of CMV may also have dual functions as VSRs and determinants of virus movement. If each VSR seems to have a particular origin, they also act on different step of gene silencing. Viral suppressors are able to target all effectors of the silencing pathway, such as viral RNA recognition, dicing, RISC assembly, RNA targeting and amplification (Fig 19).

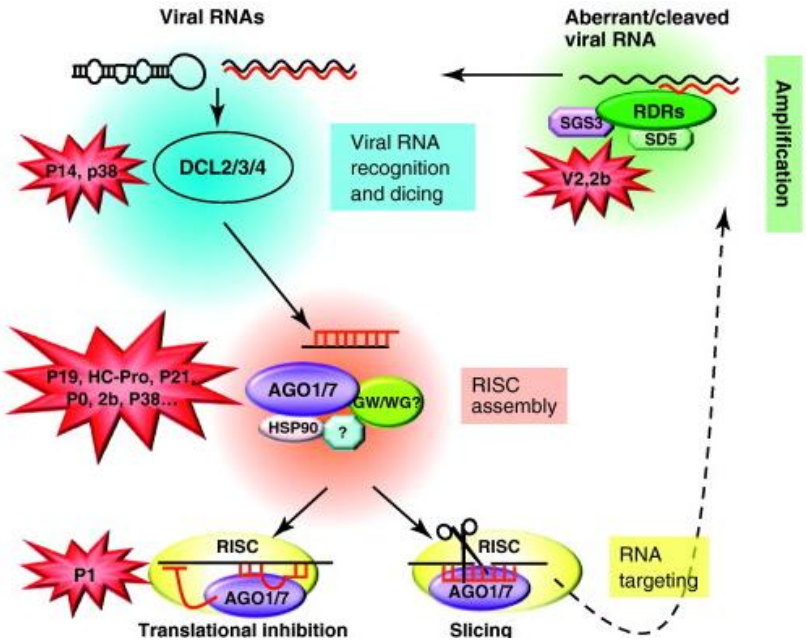


Figure 19. Current model of antiviral RNA silencing in plants and its suppression by virus-encoded silencing suppressors. Recognition of viral dsRNAs or partially double-stranded hairpin RNAs initiate RNA silencing. dsRNA-specific RNases called DCLs generate virus-derived siRNAs (vsiRNA). In the next step, the RISC complex is formed (HSP90-activated AGO1/7 loaded with vsiRNA and probably also incorporate other unidentified proteins). Then RISC target viral RNA by slicing or translational arrest. Secondary vsiRNAs are produced in an amplification loop through the actions of RDRs. Viral-silencing suppressors (i.e., P14, P38, V2, 2b, P19, HC-Pro, P21, P0 and P1) can disrupt many steps of RNA silencing by preventing the assembly of different effectors or inhibiting their actions (Burgyan and Havelda, 2011).

Several VSRs inhibiting viral RNA recognition have been identified; for example, the P14 of *Pothos latent aureus virus* (PoLV) and the P38 of *Turnip crinkle virus* (TCV) which inhibit the processing of dsRNA to siRNAs (Merai et al., 2006; Merai et al., 2005), the P6 of *Cauliflower mosaic virus* (CaMV) interacts with dsRNA-binding protein 4 which is required for the activity of DCL4 (Haas et al., 2008). Preventing RISC assembly is in general done by targeting one of its essential components (siRNAs and miRNAs or AGO proteins) in different ways. The best documented way is dsRNA or siRNA sequestration, which prevents the assembly of the RISC effector (Ding and Voinnet, 2007; Lakatos et al., 2006; Merai et al., 2006; Wu et al., 2010) as it is the case for the P19 protein of tombusviruses which prevents RNA silencing by siRNA sequestration through binding double-stranded siRNA with high affinity (Silhavy et al., 2002). In the case of CMV 2b protein, it prevents the spread of the long-range silencing signal facilitating virus infection (Brigneti et al., 1998). Thus, 2b protein have a dual mode of silencing inhibition, either by sequestering siRNAs or by interacting with AGO1 and preventing RISC assembly, inhibiting the production of viral secondary siRNAs (Diaz-Pendon et al., 2007).

If the VSR do not prevent RISC assembly, there is always the possibility to repress the expression of its components. For example miR168 which is ubiquitous in plant virus infection (Bortolamiol et al., 2007) controls the level of AGO1. P19 of the *Cymbidium ringspot virus* mediates the induction of miR168, which in turn down-regulates the endogenous level of AGO1 (Varallyay et al., 2010).

Some viruses encode more than one VSR. For example *Citrus tristeza virus* (CTV; family *Closteroviridae*) encodes three VSRs. These three VSRs must therefore act in different but overlapping ways to inhibit silencing and promote CTV infection (Lu et al., 2004). Indeed these three VSRs exhibit distinct features in silencing suppression. p23 seems similar to potyvirus HC-Pro and p20 shares similarities in silencing suppression with CMV 2b. Finally the CP activity as VSR is a novelty, because its inhibition of the intercellular

spread of the silencing signal is not associated with suppression of intracellular silencing. Combination of three VSRs in a same genome could suppress the siRNA-mediated antiviral pathway (Carrington et al., 2001; Silhavy and Burgyan, 2004) and possibly also interfere with the host development pathway controlled by miRNAs (Chen et al., 2004; Dunoyer et al., 2004; Mallory et al., 2003), at multiple points in a manner similar to virus synergy in mixed infections with viruses carrying distinct suppressors (Mlotshwa et al., 2002; Pruss et al., 1997).

In numerous studies VSRs are described as pathogenic determinants and responsible for virus-induced symptoms. Indeed when a VSR binds to siRNA and interacts, in general, with siRNA and miRNA biogenesis (Csorba et al., 2007; Kasschau et al., 2003; Lozsa et al., 2008), it compromises plant gene expression. Furthermore, the expression of VSRs in transgenic plants mimics virus symptoms at the phenotypical level (Dunoyer et al., 2004; Kasschau et al., 2003).

2.2.3 Coinfection/complementation/synergism among viruses

Plants can be infected by several viruses at the time. In this case, different kind of interaction can occur between them. A first effect of coinfection, which has already been discussed above, is recombination that may create new viral genomes. In addition to this effect, coinfection with different viruses may affect the accumulation and symptoms development. For example, in mixed tomato infections of PVY and CMV, the symptoms and RNA accumulation of CMV are exacerbated (Mascia et al., 2010). Many viruses have been found to act synergistically with CMV in legumes, cucurbits and solanaceous crops (Palukaitis and Garcia-Arenal, 2003; Wege and Siegmund, 2007). Moreover, the CMV 2b protein was able to neutralize a defense response activated by HC-Pro in transgenic tobacco in combination with TRV (Shams-Bakhsh et al., 2007). In the case of transgenic expression of the RNA silencing suppressor P1/HC-Pro from TEV, enhanced resistance was observed against tobacco mosaic virus (TMV) and tomato black ring virus (TBRV) (Pruss et al., 2004),

whereas in the case of PPV resistance was reduced (Alamillo et al., 2006).

In these mixed infections, interaction between viruses produce an increase in symptoms severity , an increase of virus titer and complementation of movement defects due to host restriction, giving an opportunity to multiply in non-host plants and resistant varieties (Palukaitis and Garcia-Arenal, 2003; Palukaitis and Zaitlin, 1997; Ryabov et al., 2001; Saenz et al., 2002; Wege and Siegmund, 2007).In this context, some recombination between different viruses can generate new functions or duplicate functions, for example the silencing suppression.

3. Potyviruses

Potyvirus is the main genus of the *Potyviridae*, the largest family of plant viruses, containing over 100 different species in eight genera (Lopez-Moya and Garcia, 1999). The *Potyviridae* belongs to the picornavirus supergroup of positive-sense single-strand RNA viruses (Fig 20). The others genera within the family are *Brambyvirus*, *Ipomovirus*, *Macluravirus*, *Poaceavirus*, *Rymovirus*, *Tritimovirus*, and *Bymovirus*. All contain monopartite viruses, but *Bymovirus* gathering viruses with bipartite genomes.

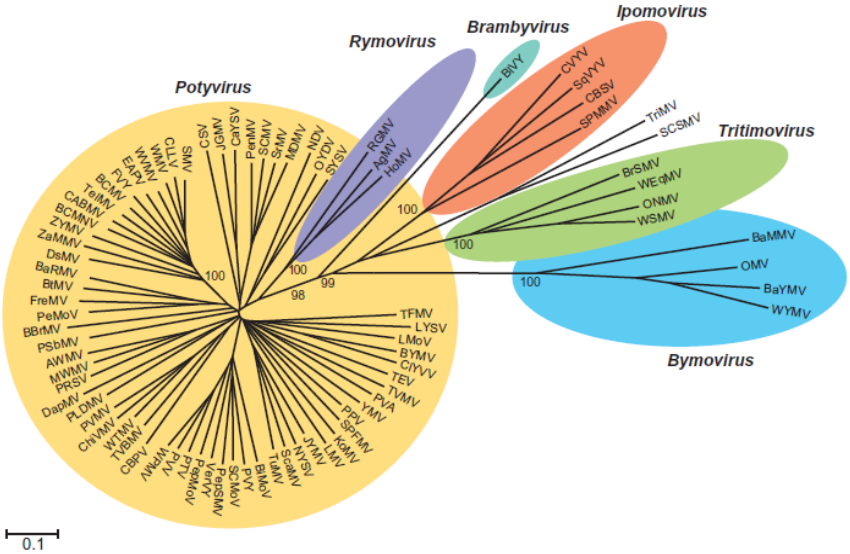


Figure 20. The family *Potyviridae*. Unrooted phylogenetic tree of fully-sequenced members of the family *Potyviridae* based on the codon-aligned nucleotide sequences of the polyproteins from 6K1 to CP (King et al., 2012).

The potyviruses represent ca. 30% of all plant viral pathogens and produce severe crop losses worldwide (Lopez-Moya and Garcia, 1999; Shukla et al., 1994), infecting a large spectrum of plant hosts, including mono and dicotyledonous plants around the world. Overall similarity of the polyproteins of viruses of the family *Potyviridae* is

rather high, with levels of amino acid identity ranging from 25 to 33% in viruses from different genera and from 42 to 56% among different species of the same genus (Adams et al., 2005). More interesting is that for all potyviruses, the gene order is conserved throughout the family (King et al., 2012). Potyviruses share also the same type of capsid and are flexuous particles between 680 and 900 nm in length with 11 to 15 nm wide (Fig. 21).

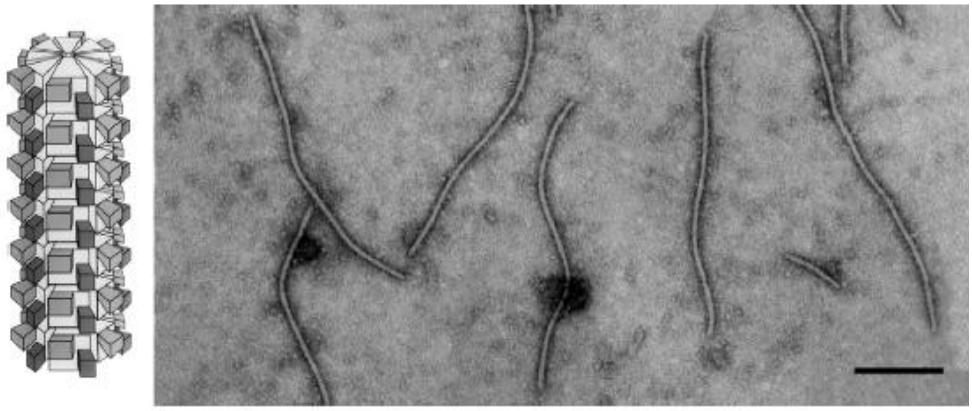


Figure 21. TuMV. Schematic diagram of potyvirus virions (left). Negative contrast electron micrograph of particules of an isolated potyvirus, stained with 1% PTA, pH 6.0 (right) (King et al., 2012). The bar represents 200 nm.

TuMV is usually spread by 40-50 species of aphids in a non-circulative and non-replicative manner, most notably by *Myzus persicae* and *Brevicoryne brassicae* (Fig. 22) (Kennedy et al., 1962). The virus is adsorbed on the surface of the aphid stylet and released from the bucal apparatus when the insect punches another plant in order to feed (Pirone and Blanc, 1996).

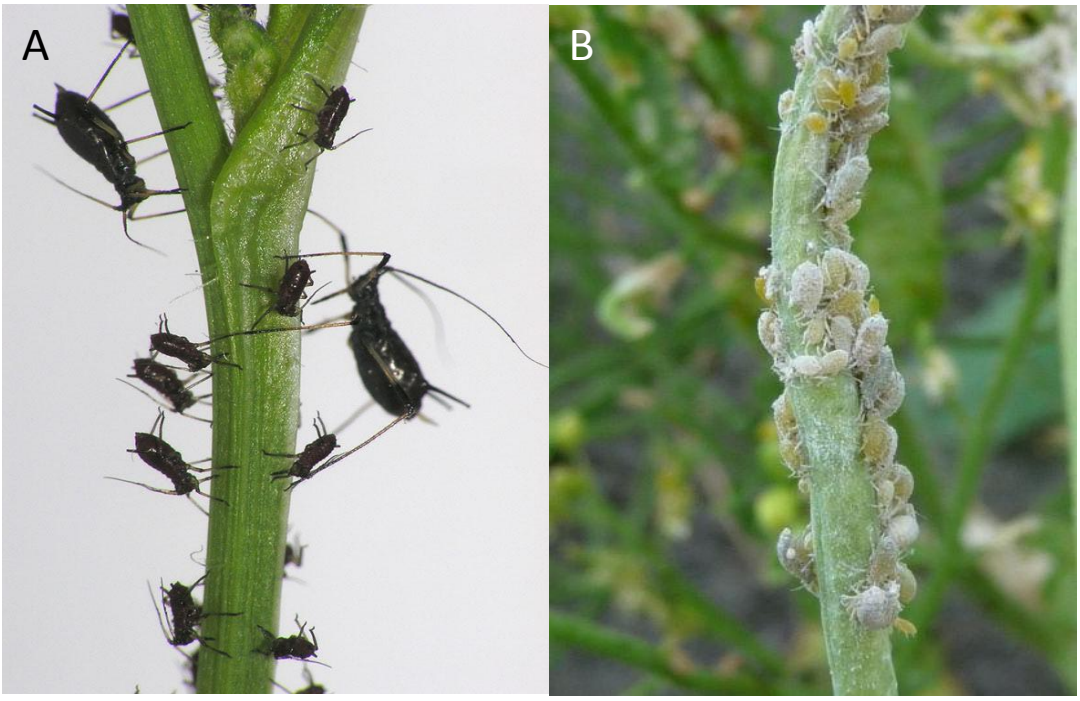


Figure 22. Aphids commonly transmiting TuMV in a non-persistent manner. (**A**) *Myzus persicae* and (**B**) *Brevicoryne brassicae*.

TuMV has a worldwide widespread host range, including many popular weeds and horticultural crops. The majority of its hosts belong to the family *Brassicaceae*. TuMV is ranked among the major pathogens infecting brassica crops, inducing a disease known as the mosaic of *Cruciferae*. Infected plants, especially the natural hosts, show symptoms such as chlorotic local lesions, mosaic, mottling, puckering or rugosity (Fig. 23).

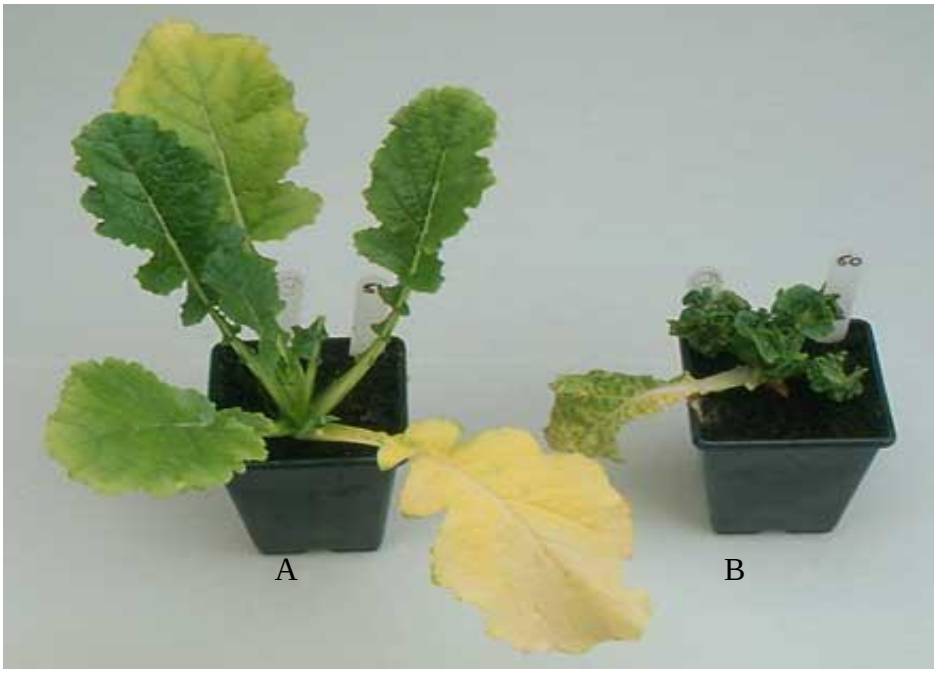


Figure 23. A non-infected *Brassica rapa* plant (**A**) and one infected with TuMV (**B**).

TuMV genome is 9,5 Kb long and is composed by a single strand positive-sense RNA, which encode for eleven mature proteins (Fig. 24).

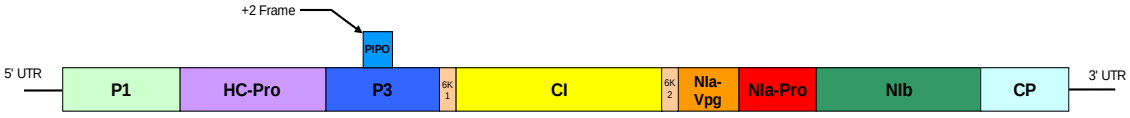


Figure 24. TuMV genome organization. Cleavage of the polyprotein gives 11 proteins. P1 and HC-Pro (helper component proteinase) cleave the bonds between P1 and HC-Pro and HC-Pro and P3 proteins, respectively while the NIa-Pro cleaves in the rest of the polyprotein cleavage sites.

Table 4. Properties of potyviral proteins (Adams et al., 2005; King et al., 2012; Urcuqui-Inchima et al., 2001).

Protein	Properties
P1	P1 is the least conserved in sequence and the most variable in size. It plays a significant role in virus replication probably due to the stimulation of the gene silencing suppressor HC-Pro. A serine protease domain towards the C-terminus cleaves the P1 from the polyprotein, typically at Tyr/Phe-Ser. It is involved in viral movement and symptomatology.
HC-Pro	The HC-Pro (Helper Component-Protease) protein has roles in suppression of gene silencing and in vector transmission. A cysteine protease domain towards the C-terminus cleaves it from the remainder of the downstream polyprotein, typically at Gly-Gly. It plays roles in systemic movement, interaction with 20S proteasome and metilation of siRNAs.
P3	Involved in virus replication and appears to be significant in host range and symptoms.
P3N-PIPO	Cell-to-cell movement
6K1	The function of this small protein is related to P3 and unknown.
CI	CI (Cylindrical Inclusion protein) has helicase activity and accumulates in inclusion bodies in the cytoplasm. It binds to RNA and is involved in cell-to-cell movement.
6K2	A small transmembrane protein anchoring the replication complex to the ER.
NIa-Vpg	VPg (Viral Protein genome-linked) is attached to the 5' terminus of the genome. It plays multiple roles in the viral infection cycle. It is essential for virus replication and translation, interacting with one or several isoforms of the eIF4E translation initiation factor.
NIa-Pro	Serine-like cysteine protease responsible for cleavage of most sites in the polyprotein, typically at Gln/Glu-(Ser/Gly/Ala), recognizing a seven amino-acid motif.
NIb	RNA-dependent RNA polymerase.
CP	Viral coat protein that also has roles in virus movement, genome amplification, vector transmission and cell-to-cell movement.

The polyprotein is cleaved by 3 different proteases (all encoded by the virus) into 11 different mature proteins. Seven of the nine cleavage sites are cut by the viral NIa-Pro acting both in *cis* and in *trans*, while P1 and HC-Pro act only in *cis* separating themselves from the polyprotein during translation (Adams et al., 2005). TuMV replicates in the cytoplasm and packs into a filamentous, flexuous nucleocapsids. Virions consist of genomic RNA with a 3' poly(A) tail and 5' covalently attached genome-linked protein (VPg).

Many of these viral proteins have multiple functions (Table 4). Recently a new protein have been discovered by phylogenetic analysis, the protein PIPO (Chung et al., 2008). It is an overlapping gene resulting from a +2 frame shift in the P3 cistron resulting in the P3N-PIPO protein. This novel protein is involved in cell-to-cell movement (Vijayapalani et al., 2012) along with viral proteins CI and CP (Carrington et al., 1998; Dolja et al., 1995).

II. MATERIALS AND METHODS

1. Biological material

1.1. Infectious material

The TuMV strain used in this study named YC5 (GenBank accession number: AF530055.2) was isolated in Taiwan from calla lily (*Zantedeschia* spp.) cv. Black magic displaying yellow spotting and striping on leaves. This isolate can be mechanically transmitted to various hybrids of *Zantedeschia* and induce systemic symptoms similar to those observed on diseased Black magic. In addition to *Zantedeschia* spp., YC5 also infects several cruciferous species and induce mosaic symptoms (Chen et al., 2003).

1.2. Plant material and growth conditions

1.2.1 *Arabidopsis thaliana* (Fig. 25A)

A. thaliana belongs to *Brassicaceae* family. It is a small flowering plant related to cabbage and mustard. *A. thaliana* is the favorite model organisms used in plant biology and the first plant to have its entire genome sequenced (Arabidopsis, 2000). Two transgenic *A. thaliana* Col-0 lines expressing amiR159-HCPro were used in this study: 10-4 and 12-4 (Niu et al., 2006). Seeds used in all the experiments corresponded to an homozygous T4 generation. Plants were maintained in a growth chamber under conditions of 16 h of light at 25 °C followed by 8 h of darkness at 22 °C. 12-4 plants were fully resistant to infection with the ancestral TuMV clone, whereas 10-4 plants showed incomplete penetrance and variable expressivity of the resistance character.

1.2.2 *Nicotiana benthamiana* (Fig. 25B)

The genus *Nicotiana* belongs to the family *Solanaceae* and includes 76 species of mainly tropical and subtropical distribution from four continents, with the majority occurring in South America and Australia. *N. benthamiana* is a close relative of tobacco and species of *Nicotiana* indigenous from Australia (Knapp et al., 2004). It

is also considered as a model organism for plant research. Indeed *N. benthamiana* is the most widely used experimental host in plant virology, mainly due to the large number of plant viruses that can successfully infect it. *N. benthamiana* genome is 20-fold larger than that of *A. thaliana* (nearly 157 Mbp) (Bennett and Leitch, 2005).

Plants used in this study (section III.2) are wild-type and were maintained in a greenhouse at 25 °C and with 16 h of light at all times.

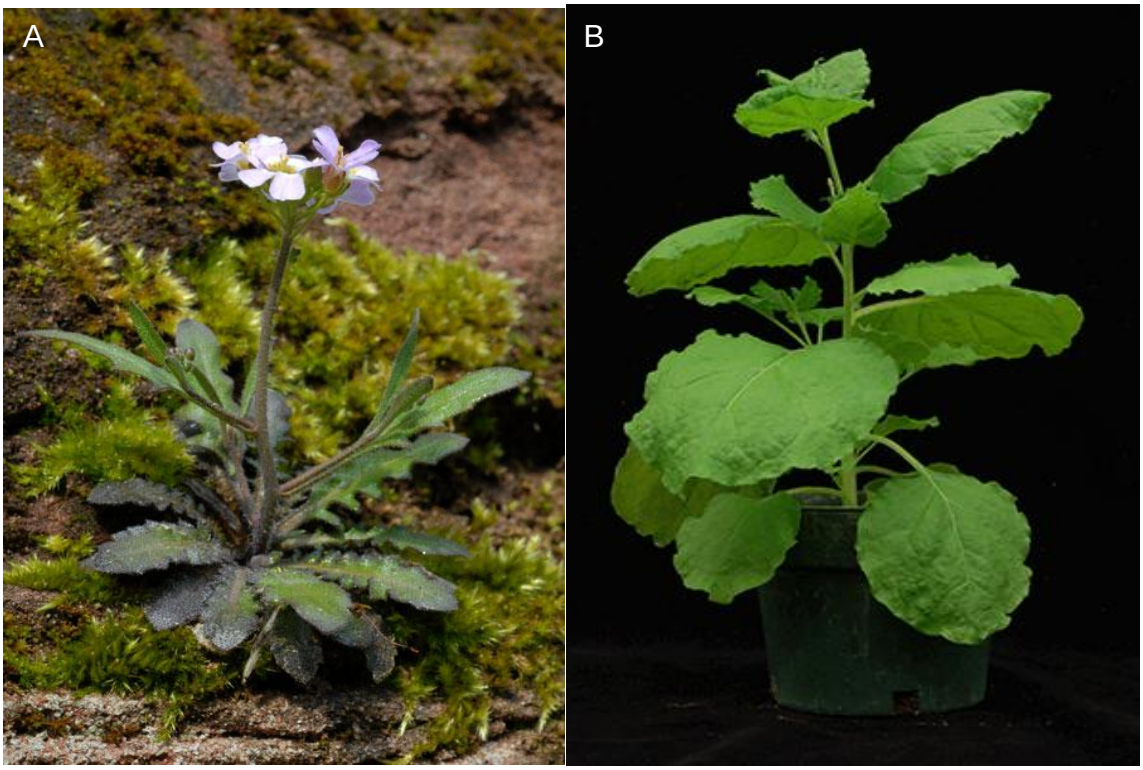


Figure 25. Pictures of the host plants used during this study: *A. thaliana* (A) *and N. benthamiana* (B).

2. Methods

2.1. Quantification of amiR159-HCPro expression

Total RNA was extracted and purified from *A. thaliana* tissue using the TRIzol reagent (Invitrogen). RNA was precipitated with isopropanol, resuspended in H_2O, and quantified by spectrophotometry. The quantification of amiR159-HCPro in RNA preparations was performed by reverse transcriptase (RT) quantitative PCR (qPCR) in triplicate (Varkonyi-Gasic et al., 2007). Standards were prepared by the addition of known amounts of the synthetic oligoribonucleotide 5'-(ACUUGCUCACGCACUCGACUG)-3', corresponding in sequence to amiR159-HCPro, to a non-transgenic *A. thaliana* total RNA preparation. RT reactions were done with a 10 µl mixture containing 100 ng total RNA, 1 pmol primer I (PI) (5'-GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATAC GAC**CAGTCG**-3' [the sequence complementary to amiR159-HCPro is in boldface type]), and 30 U *Moloney murine leukemia virus* (M-MuLV) RT (Fermentas), with incubation for 10 min at 25 °C, 45 min at 42 °C, 10 min at 50°C, 5 min at 60 °C, and finally, 15 min at 70 °C. qPCR was next performed in a 20 µl mixture containing 2 µl of RT reaction product and 10 pmol each primer PII (5'-CGGCGG**ACTTGCTCACGCACT**-3' [the sequence complementary to amiR159-HCPro is in boldface type]) and PIII (5'-GTGCAGGGTCCGAGG T-3' [homologous to the sequence underlined in PI]) with Maxima SYBR green master mix (Fermentas) and incubation for 10 min at 95 °C, followed of 40 cycles of 15 s at 95 °C, 30 s at 60 °C, and 30 s at 72 °C.

2.2. Population passages and evaluation of pathogenicity in *A. thaliana* 12-4 plants

As a source of the TuMV inoculum for all our experiments, we used a large stock of infectious sap obtained from TuMV-infected *N. benthamiana* plants inoculated with a plasmid containing TuMV cDNA under the control of the CaMV 35S promoter(Chen et al.,

2003). Plant infectious saps were obtained by grinding the infected tissue in a mortar with 20 volumes of grinding buffer (50 mM potassium phosphate [pH 7.0], 3% polyethylene glycol 6000 [PEG 6000]). Fig. 26 summarizes the experimental design for the evolution experiments. Aliquots of 5 µl of 10% Carborundum in grinding buffer were applied onto three different *A. thaliana* leaves, and inoculation was done mechanically by gentle rubbing with a cotton swab soaked with infectious sap. Twenty-five wild-type *A. thaliana* and twenty-five 10-4 transgenic plants were initially inoculated. Each plant represented the starting point for an independent evolution lineage. At 14 days postinfection (dpi), symptomatic tissue was collected for each lineage and homogenized in grinding buffer. A portion of the resulting saps was used to inoculate the next set of plants. A second portion of the homogenized sap was frozen at -80°C for further characterization. The remaining portion was used for the challenge experiments designed to estimate pathogenicity in 12-4 plants. This procedure was repeated until all 50 evolutionary lineages overcame the resistance of the 12-4 line. Once a lineage was able of breaking resistance, it was removed from the passaging experiment. For the pathogenicity test experiments, 20 plants of the 12-4 line were inoculated as described above. Plants were visually checked for the presence of symptoms at 14 dpi, and the frequency of infected plants, that is, pathogenicity, was recorded. These challenge experiments were performed after every evolutionary passage for each one of the 50 evolving lineages. A pilot experiment showed that infection always concurred with symptom development. A lineage was considered to be capable of breaking resistance if at least one 12-4 plant showed symptoms.

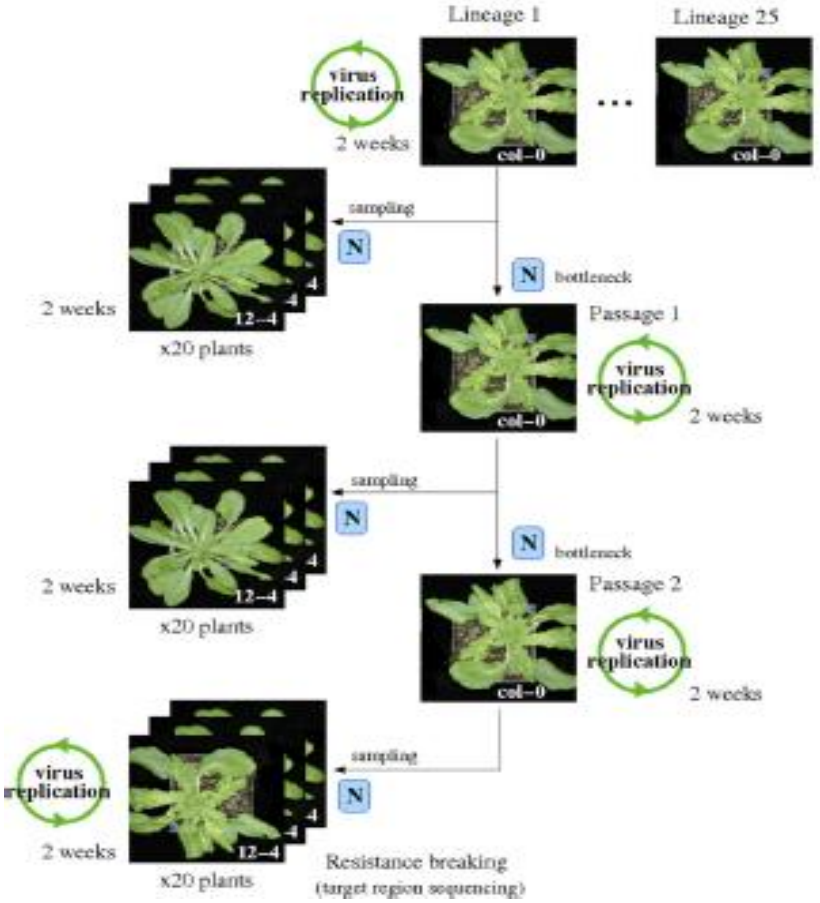


Figure 26. Schematic representation of the experimental design. For illustrative purposes we show only one of the wild-type *A. thaliana*-evolved lineages. The same protocol was repeated for the 10-4 lineages, with the exception that serial passages were performed on *A. thaliana* 10-4 transgenic plants. In the illustrated example, resistance breaking occurred at passage 2 (note symptoms in the corresponding 12-4 plants).

2.3. **Sequence analysis of the 21-nt target region**

The region around the 21-nt target of amiR159-HCPro was sequenced in virus populations breaking resistance. Total RNA from infected *A. thaliana* 12-4 transgenic plants was purified using silicagel columns (Zymo Research), and a viral cDNA was amplified by RT-PCR. RT reactions were carried out in a 10 µl volume containing 50 U M-MuLV RT and 5 pmol primer IV (5'-CCTGGTGACAGTAAAGCATATAATGG-3') for 45 min at 42 °C, 5 min at 50 °C, and 5 min at 60 °C. One µl of the RT reaction was used for PCR amplification in a 20 µl mixture with 0.4 U Phusion DNA polymerase (Finnzymes) and 10 pmol each primer PV (5'-GAC AATGAGTCACAAGATTGTGCACTTT-3') and PVI (5'-CATGAGTGTCCTCCCATTCTGTCCC-3'), with incubation for 30 s at 98 °C; 30 cycles of 10 s at 98 °C, 30 s at 55 °C, and 30 s at 72 °C; and a final extension step for 10 min at 72 °C. Amplification products were separated by electrophoresis in a 1% agarose gel, and the TuMV cDNAs matching the expected 1,427 bp were eluted and sequenced with primer PVII (5'-AAACGATTCTTCAGCAACTACTTTG-3').

2.4. **Simulation algorithm.**

The experiments were simulated by using a bit-string Monte Carlo model (Elena et al., 2010; Marin and Solé, 1999), in which digital genomes were represented by binary strings, ***S***, of a length, *L*, of 31 bits. The digital genomes explicitly considered the 21 nt of the amiR159-HCPro target and added 10 more bits, each corresponding to one of the 10 viral cistrons (Fig. 27). We made this distinction to disentangle the effects due to a mutation in the target (evaluated at the challenge step of the experiment) from those associated with changes in other viral genes and that determine the overall fitness of the virus.

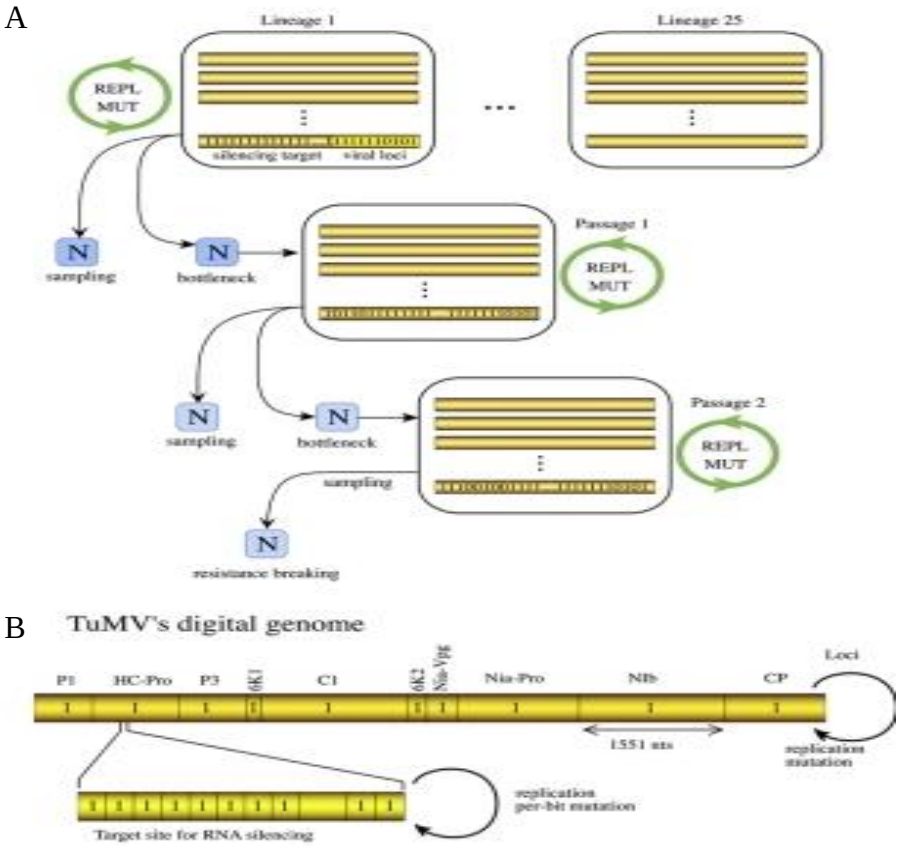


Figure. 27. Schematic diagram of the *in silico* simulation model. (**A**) Each lineage was considered a population of bit-strings containing the 21 bits of the target region plus 10 loci, each corresponding to the different cistrons in the TuMV genome. The model simulated within-host viral replication with mutation and bottleneck transmission between passages. For the simulations of virus evolution in wild-type *A. thaliana* plants, we did not consider target-specific degradation of strings, while for simulating the evolution in 10-4 plants, we included a degradation probability, ε, for strings with a WT target sequence. (**B**) Digital genome of TuMV where the target sequence has been explicitly considered.

The maximum string population size was set to an N_{max} of 5,000 genomes. As in the experiments, the simulation model considered 25 independent lineages (Fig. 27). Each lineage started with a sample of size $N<N_{max}$ of WT genomes. For each lineage we let the population experience τ replication events. At each event, two locations in the population were randomly chosen. If location *i* already contained a string, it was copied to site *j* with probability $P_{ij} = \frac{1}{1+\exp(-\Delta f_{ij}/T)}$ that depends on the fitness difference $\Delta f_{ij} = f_i - f_j$ between strings $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$ (if site *j* is empty, $f_j = 0$). T is the Boltzman temperature, which is a measure of the noise tied to replication events, and it was fixed to $T = 0.2$. The fitness of a given string, $\boldsymbol{S}_k$, was obtained from the binary composition of the 10 loci. We considered four types of deleterious fitness landscapes: the standard additive, antagonistic, and synergistic ones plus one in which mutations in the bits representing the 10 viral cistrons were considered lethal. For the three deleterious landscapes we computed the fitness as $f_k = 1-1-d_H^{\xi}/10$, where d_H is the Hamming distance (i.e., how many different bits we had) between sequence *k* and the corresponding loci of the wild-type genome. ξ measures the sign and strength of epistasis, where $\xi = 1$ if additive, $\xi < 1$ if antagonistic, and $\xi > 1$ if synergistic (Sardanyes et al., 2009). During replication, each bit of the amiR159-HCPro target can mutate with the probability μ. The other 10 loci of the strings mutate with the probability $\mu_{li} = 3\mu v/2l_i$, where l_i is the length of locus *i* and the 2/3 is introduced to consider, as a first approximation, that mutations at the third codon positions are neutral. This correction was done to ensure that all loci mutated proportionally to their length. In order to differentiate between the experiments carried out with wild-type plants and those carried out with 10-4 *A. thaliana* plants, we considered that if the string chosen for replication was the wild-type genome, it would be degraded with probability $\varepsilon = 0$ for simulations in wild-type plants and $\varepsilon > 0$ for simulations in 10-4 plants. As mentioned above, for each lineage we let the population

evolve over τ replication events according to the above-described rules. We then took two random samples of size N (Fig. 27). The first sample was used to initiate the next population (simulating the next passage in the experimental evolutionary lineages) until resistance was broken. The second sample was used to evaluate the likelihood of resistance breaking as follows. For each string $\boldsymbol{S}_i$ in the second sample, we evaluated its pathogenicity as $\theta(\mathbf{S}_i) = 1 - \prod_{k=1}^{21}[1 - \lambda(S_{ik})]$, with $\lambda(S_{ik})$ being the empirical probability that a change in position k of the 21-nt target will be an escape mutation (frequency data from Fig. 35 were corrected by using the Laplace estimator). Next, we evaluated the likelihood of resistance breaking for this second sample after 20 trials (the number of plants inoculated during the challenging experiments) as P_b=1-1(1-<p>)^{20}, where $\langle p \rangle = \frac{1}{N}\sum_{i=1}^{N}\theta(\mathbf{S}_i)$ is the average pathogenicity of all the strains contained in the sample. If $P_b \geq 0.05$, we assumed that resistance was broken. For a sample of 20 plants, this threshold means that at least one plant became symptomatic.

2.5. **Data fitting and parameter inference**

To fit the experimental data to the simulation model and to infer relevant population parameters, we used an optimization algorithm (OA) (Malpica et al., 2002) that systematically searched the parameter space defined by $C = \{\tau, \mu, N, \xi, \varepsilon\}$ we as follows. First, we defined a starting population of 150 parameter sets, $C_1(0)$, $C_2(0)$, …, $C_{150}(0)$. The parameter values for each one of these $C_h(0)$ parameter sets was randomly assigned within the following ranges: $1 \leq \tau \leq 10^{-5}$, $1 \leq N \leq N_{\max}$, $10^{-7} \leq \mu \leq 10^{-3}$, $0.2 \leq \xi \leq 1.8$, and $0.1 \leq \varepsilon \leq 0.5$. For each one of these parameter sets, we ran the simulation algorithm described in the two previous paragraphs. At the end of each simulation we compared the observed cumulative frequencies at passage j shown in Fig. 34, $\rho_{obs}(j)$, with those obtained in the

simulation, $\rho_i(j)$, using the equation $d_h = \sum_{j=1}^{28} |\rho_h(j) - \rho_{obs}(j)|$, which represents a distance value between the empirical and the simulated data. This procedure generated a vector of 150 d_h values between the observed and simulated data. Then we computed then the average distance, $D = \frac{1}{150} \sum_{h=1}^{150} d_h$, from all the 150 parameter sets and chose those sets with a distance smaller than D as a starting point for the next iteration of the OA, $C_i(1)$. Since fewer than 150 parameter sets were left for the next iteration, the rest of the sets were generated by the addition of small perturbations to the retained parameter sets. The whole process was repeated until no change was observed for $\langle d(t) \rangle$ after t iterations of the OA. Notice that for wild-type plants, $C = \{ \tau, \mu, N, \xi$, since amiR-mediated degradation was fixed to $\varepsilon = 0$.

2.6. Phylogenetic analysis

2.6.1 Alignment and phylogenetic tree

The genomes of a representative world-wide collection of 32 TuMV isolates with ”minimum recombination” were retained (Tomimura et al., 2003). Only coding sequences were used for this analysis.

The sequences of each TuMV protein were aligned with Muscle (Edgar, 2004) and back-translated into codon-aligned nucleotides sequences with Revtrans (Wernersson and Pedersen, 2003).

Phylogenetic relationships were calculated for each separate gene as well as for the TuMV concatenated genome. Phylogenetic reconstructions were performed by maximum likelihood (ML) using RAXML v7.2.8 (Huson and Bryant, 2006; Stamatakis et al., 2005) at the nucleotide level, using the GTR model, considering nine partitions, one per codon position per gene. Reliability of the resulting trees was evaluated by 500 bootstrap pseudosamples. Furthermore, phylogenetic networks were used to represent incompatibilities within

and between data sets. To tackle this issue we generated a split network computed from trees as a ”supernetwork” using 100 number of runs by the SplitsTree4 software (Huson and Bryant, 2006).

2.6.2 Distances among phylogenetic trees

Two phylogenetic trees can be compared by differences in both topologies and branch lengths. We compared the gene trees with the corresponding tree resulting from the concatenated genome, using the program Ktreedist (Soria-Carrasco et al. 2007), which takes into account both topology and branch length information of a phylogenetic tree. This program computes a K-score that measures overall differences in the relative branch length and topology of two phylogenetic trees after scaling one of the trees to have a global divergence as similar as possible to the other tree. This allows for the comparison among trees differing in evolutionary rates and/or divergence scales. The scale factor defines a value which indicates how many times a tree's total branch length have to be multiplied to be equal to the reference tree. Trees compared with Ktreedist must be all rooted or unrooted. To this end, trees were saved with estimated branch lengths. With Ktreedist gene trees were compared to the concatenated. High K-scores indicate a poor match between the estimated tree and the reference tree. We also compared reference and estimated trees using the symmetric difference (Robinson and Foulds, 1981), which only takes into account the topology of a phylogenetic tree. This method calculates the number of splits that disagree between two trees. If a rooted tree is input, it will be automatically unrooted.

Tree distances were calculated using RAxML software through the option -f x and with partitions.

2.6.3 Modes of evolution

In order to identify selection pressures along a protein the Selecton server was use (Doron-Faigenboim et al., 2005), which calculates the ratio between non-synonymous (Ka) and synonymous

(Ks) substitutions (w) at each site of the protein looking for site-specific positive selection, purifying selection or lack of selection. Positions with a w between 0 and 0.019 are considered under purifying, between 0.2 and 0.99 under neutral and > 1 under positive selection. Then the Hubert M-estimator of all position of each gene was determined, to take account of outliers.

2.6.4 Graphical representation of pairwise nucleotide vs. amino-acids distances

For each tree obtained from each individual gene, pairwise distances of all possible paired terminal taxa were calculated, from nucleotide and amino acid sequences. To summarize the information and being able to represent all genes into the same graph, the Huber M-estimator and the median absolute deviation (MAD) values were calculated for each gene tree and for the tree obtained for the concatenated genome. In order to be able to compare our data between genes of the same genome, the pairwise distances were normalised to the concatenate, and then, nucleotides and amino acids M-estimator was calculated.

Four different areas can be defined in the plain formed by the amino acid and nucleotide distances; each area reflecting different evolutionary constraints (Fig 28).

When the distance of a gene is located along the line parallel to the y-axis, the nucleotide distance of the gene is equal to the nucleotide distance of the reference point. When the distance of a gene is located to the left of this line, the nucleotide distance of the gene is smaller than the concatenated. However, when the distance of a gene is located to the right of the line, the nucleotide distance of the gene is larger than the concatenated. Likewise, when the distance of a gene is located along the line parallel to the x-axis, the amino acid distance of the gene is equal to the amino acid distance of the reference point. When the distance of a gene is located bellow this line, the amino acid distance of the gene is smaller than the concatenated. By contrast,

when the distance of a gene is located above the line, the amino acid distance of the gene is larger than the concatenated.

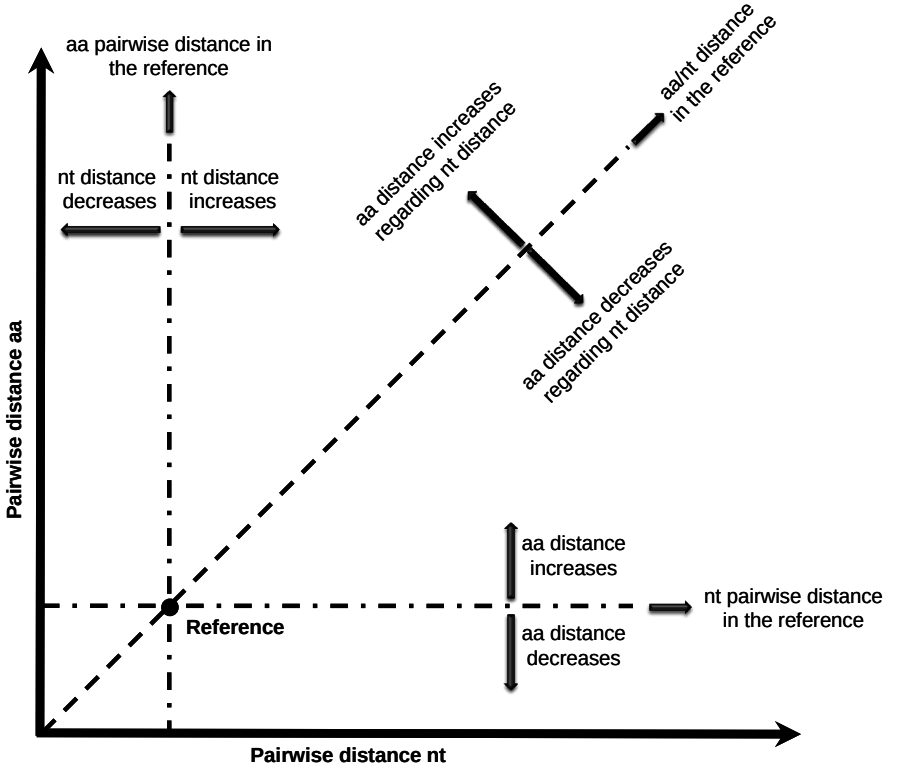


Figure 28. Schematic graphical representation between pairwise distances between nucleotide and amino acid sequences in which each gene is compared to the reference defined by the concatenated genes (whole genome).

When the distance of the gene is located along the straight line passing through the origin and point of reference, the individual gene behaves like the entire genome. However, when the gene is located above the straight line the amino acid distance increases regarding the nucleotide distance with respect the concatenated and when the gene is located below the straight line the amino acid distance decreases regarding the nucleotide distance with respect the concatenated.

2.7. **Generation of new transgenic lines expressing two amiRs**

To take advantage of the precedent 12-4 resistant plant a second amiR targeting the CP region (TuHC-CP) was inserted (Fig 29A). We also decided to generate an overlap double amiR in the CP region (TuCPa-b) (Fig 29B).

Transgenic plants expressing two amiRs from the same promoter are obtained with the same protocol as Niu et al. (2006) using the *A .thaliana* miR159a as backbone.

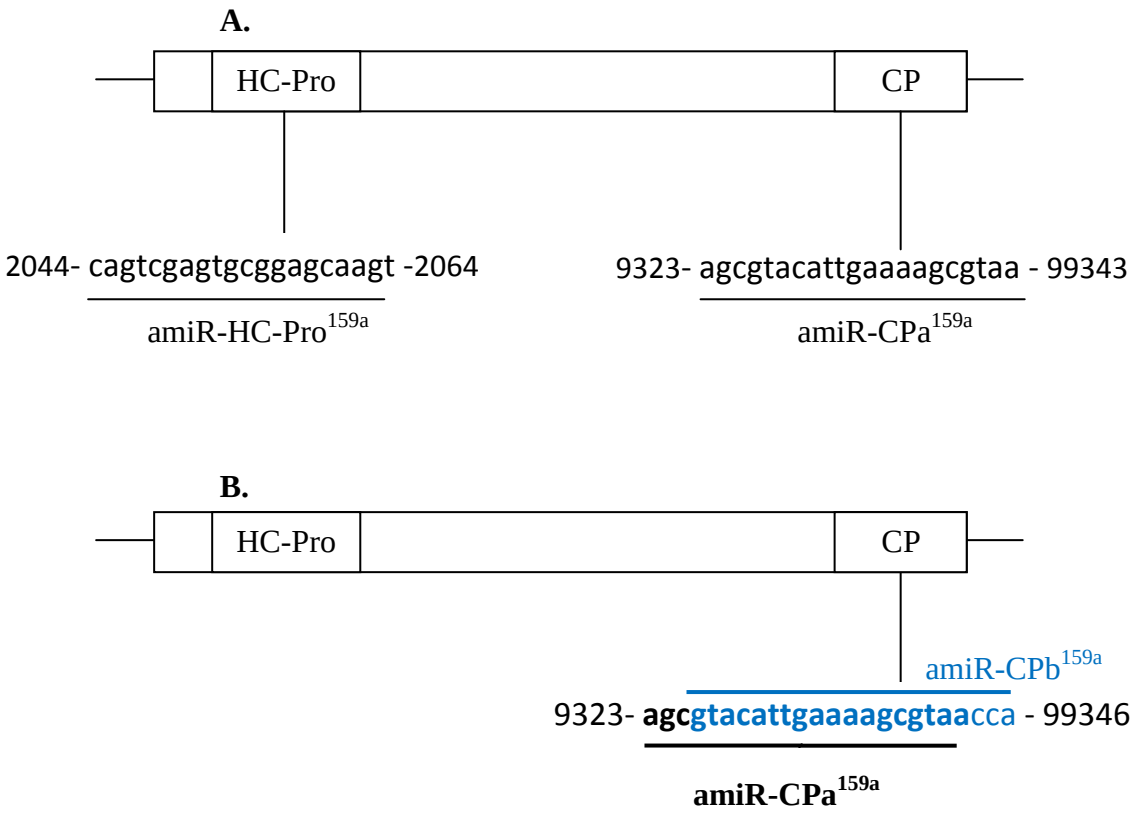


Figure 29. Position of targets of amiR in the TuMV genome. (**A**). Double amiR targeting HC-Pro and CP sequence. (**B**). Two amiR targeting overlapping sequences in the CP region.

The amiR sequence of amiR159a-HC-Pro (5'-ACTTGCTCACGCACTCGACTG-3'), amiR159a-CPa (5'-TTACGCTTTTCAATGTACGCT-3') and amiR159b-CPb (5'-TGGTTACGCTTTTCAATGTAC-3') target respectively HC-Pro (2044-2064 nt) and CP (9323-9343 nt; 9326-9346 nt) were inserted in the miR159a backbone.

2.8. **Challenge of double amiR transgenic plants**

Fig. 30 shows a scheme of the experimental design used to test the resistance of the transgenic lines carrying two amiR. Infectious sap was obtained from *N. benthamiana* plants that were previously inoculated with a plasmid containing TuMV cDNA (Chen et al., 2003) and that had developed symptoms of infection. From this primary infectious sap, wild-type or 10-4 plants were inoculated to give wild-type or 10-4 sap. Five hundred wild-type, 12-4 fully resistant, TuHC-CP and TuCPa-b *A. thaliana* plants were inoculated each with sap from wild-type or 10-4 plants. Infected plants were recorded 15 dpi by symptom inspection.

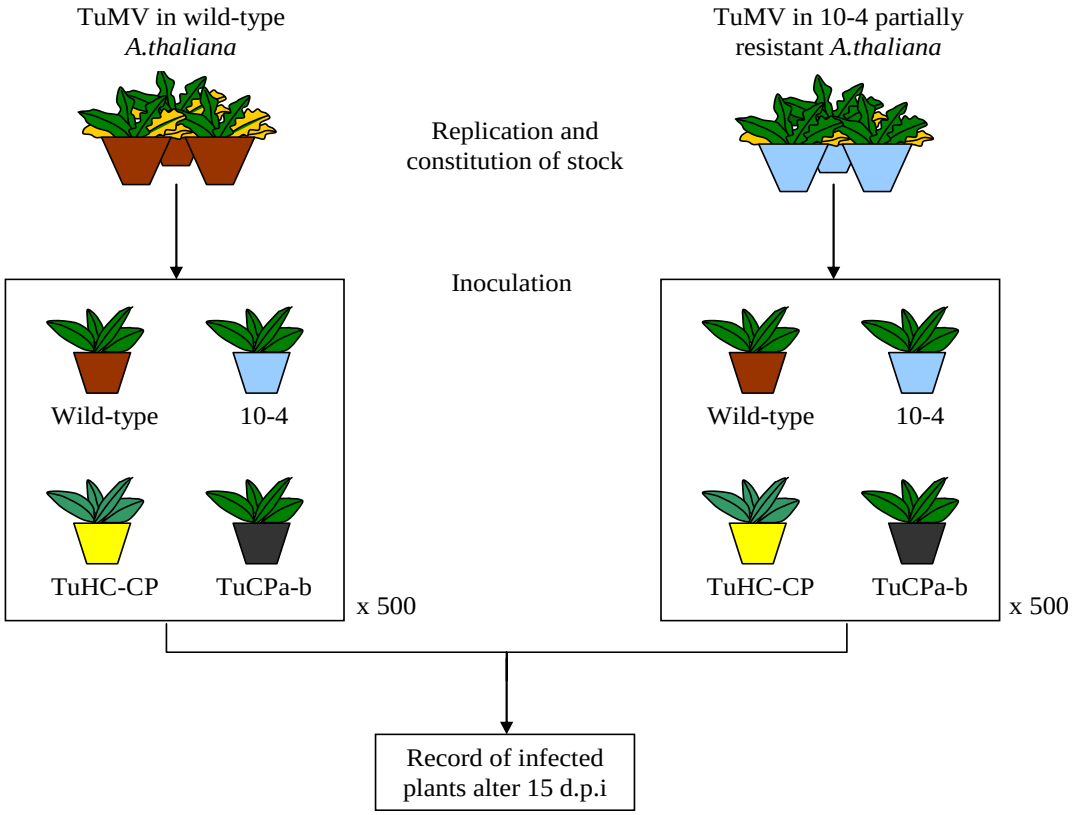


Figure 30. Schematic representation of the experimental design for challenge experiments. After a multiplication step in *N. benthamiana*, infectious saps were inoculated on wild-type or 10-4 partially resistant plants. Then more than 500 wild-type, 12-4, TuHC-CP, and TuCPa-b plants were inoculated with wild-type or 10-4 saps. Symptoms were recorded 15 dpi.

2.9. **Construction of viral genomes containing a second VSR**

To generate all constructs we use as starting point an infectious plasmid containing the TuMV genome (GenBank accession number AF530055.2) previously engineered to express a green fluorescent protein (GFP) inserted between NIb and CP cistrons. The TuMV-K clone carrying mutations CGC to AAG in the FRNK highly conserved

domain of HC-Pro, rendering a FKNK motif, was generously gifted by Dr. S.S. Lin (NTU, Taiwan).

Plasmids containing CMV 2b (GenBank accession number NP_619631.1) or TBSV p19 (GenBank accession number NP_062901.1) were used as sources to amplify the two VSR genes by PCR using the following primers: 5'-GGCGCCATGGAATTGAACGTAGGTGCAATGACAAACGTC-3' (P2b-F) and 5'-CCGCCGCTAGCGAAAGCACCTTCCGCCCA-3' (P2b-R) for CMV 2b, and 5'-GGCGCCATGGAACGAGCTATACAAGGAAACGACGCTAG-3' (P19-F) and 5'-CCGCCGCTAGCCTCGCTTTCTTTTTCGAAGGTCTCAGT-3' (P19-R) for TBSV p19. Restriction enzymes *Nco*I and *Nhe*I (Fermentas) were used to cut the PCR products and to linearize the p35S-TuMV-GFP or p35S-TuMV-K-GFP plasmids (Fig. 31). Then ligation between PCR products and linearized plasmids gave the following four constructions: TuMV-2b, TuMV-p19, TuMV-K-2b and TuMV-K-p19. Ligations were performed with a relation insert:vector 1:5, in a final volume of 10 µl. In total 50 ng of plasmid and 15-30 ng of insert were mixed with 1% of PEG 6000, DNA ligase 1X buffer (ligase 10X buffer : 660 mM Tris-HCl, 50 mM $MgCl_2$, 50 mM DDT, 10 mM ATP, pH 7.5) and 1 U of T4 DNA ligase (Roche. USA). The reactions were maintained at 22°C for 2 hours.

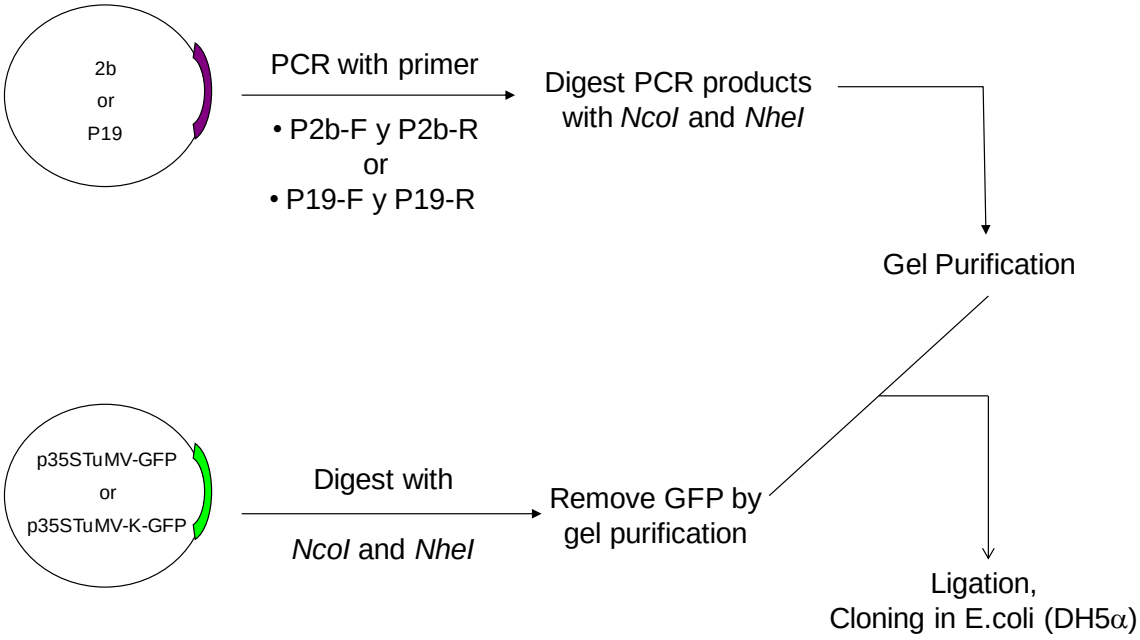


Figure 31. Generation of the TuMV with CMV 2b or TBSV p19 proteins.

2.10. **Experimental evolution of viruses carrying two VSR**

Ten µl of each plasmid (200 ng/µl) were gently rubbed on the third true leaf of *N. benthamiana* plants except for the TuMV-K-GFP which was inoculated with an *Agrobacterium tumefaciens* clone gifted by Dr. S.S. Lin (Taiwan National University, Taiwan). Once symptoms appeared, primary infectious tissue was prepared and inoculated to *N. benthamiana* plants.

Two different evolution protocols were set up, varying on the frequency at which serial passages were done. These two modes of evolution are based on the biological cycle life of the pathosystem TuMV-*N. benthamiana*. Indeed symptoms appeared 10 dpi, time at which the viral load was maximal. Therefore, a first set of plants were serial passaged every 14 dpi. To minimize founder effects and give

more time for selection to operate, a second set of plants were serially passaged every 28 dpi. For each genome and passage scheme, ten independent evolution lineages were generated and maintained for two months in *N. benthamiana* plants (Fig. 32). Henceforth, 10 passages in the case of 14 dpi evolution or 5 passages in the case of 28 dpi evolution. At the end of the evolution phase, the consensus sequences for HC-Pro and the corresponding additional VSR gene were determined.

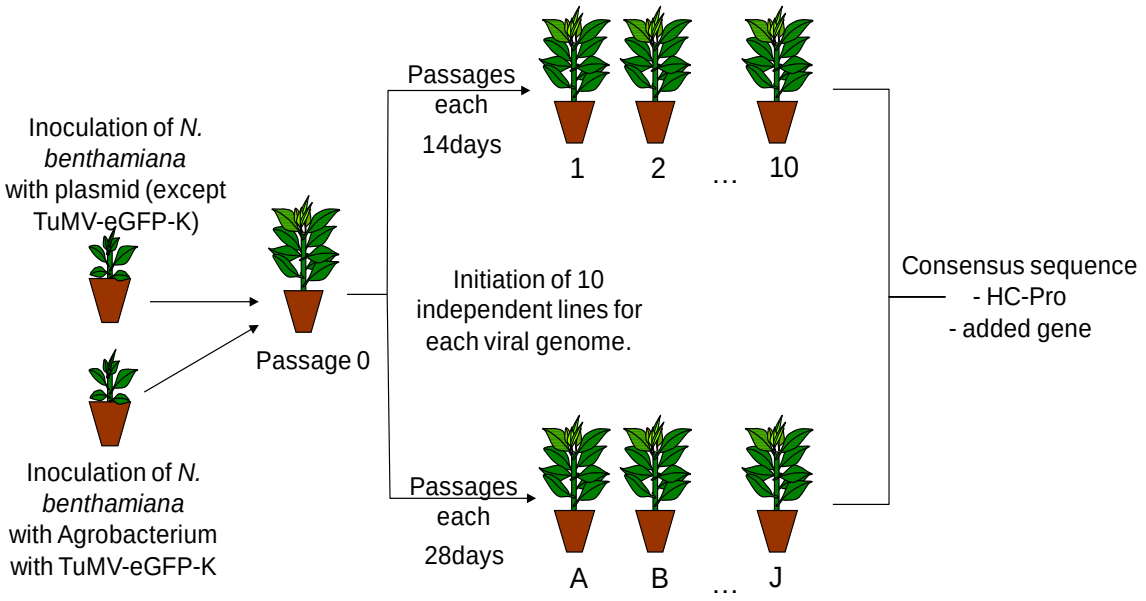


Figure 32. Schematic representation of the experimental design. After a multiplication step in *N. benthamiana* of the six different TuMV genomes, infectious saps were inoculated on *N. benthamiana* wild-type for five month. Ten independents lineages were generated for each genome during experimental evolution with two different time passages: 14 or 28 dpi. At the end of evolution, consensus sequences of HC-Pro and of the additional VSR were determined.

We used the primer TuCP-R 5'-CCGGCGTTCACTTCCCTGTCGCGTT-3’ (at the position 9645 nt of the TuMV-GFP genome) for RT. To perform the RT in 10 µl, we mixed 1 µl of RNA with 3.5 µl of H_2O, 1 µl of dNTPs (40 mM), 0.8 µl DTT (100 mM), 1,5 µl of TuCP-R primer (10 µM), 0.2 µl RNase inhibitor (40 U/µl), 1 µl of buffer (10x) and 1 µl of Accuscript enzyme.

Then to amplify the HC-Pro, we used the following primers: TuHC-F 5’-ATTCTAACCCAGAAACGAAGCAGAA-3’ and TuHC-R 5’-TATCTATGTTGCTGGCTCCATGA-3’ located at nucleotides 1114 nt and 2610 nt of TuMV genome, respectively; generating an amplicon of 1496 nt in lenght. To amplify the inserted sequence between NIb and CP and the primer TuNIb-F 5’-TGATTGAATCGTGGGGATACGACAA-3’ (position 8540 nt) and TuCP-R were used. The length of this amplicon depends of the gene added; it is 1105 nt for GFP, 901 nt for P19 and 715 nt for 2b. For a total of 25 µl of PCR mix we pulled 1 µl of buffer (10X), 1 µl of dNTPs (40 mM), 0,8 µl of DTT (100 mM), 1,5 µl of primers (10 µM), 0,2 µl RNase inhibitor (40 U/µl), 1 µl of Accuscript and 1,5 µl of DNA.

III. RESULTS

III.1 BIOTECHNOLOGICAL APPLICATION OF RNA SILENCING

It has been shown that changes within the miRNA 21-nucleotide (nt) sequence do not affect its biogenesis and maturation (Guo et al., 2005; Vaucheret et al., 2004), and this finding opened up the possibility for a redesigning of the miRNA sequence to target different transcripts using different pre-miRNAs as backbones (Niu et al., 2006; Qu et al., 2007; Warthmann et al., 2008). One of the many applications of this technology is to produce amiRs targeting viral genomes, thus generating transgenic plants that are resistant to viral infection (Niu et al., 2006; Qu et al., 2007). Niu et al. (2006) previously used the pre-miRNA159a precursor to express two amiR159s with sequences complementary to the RNA genomes of *Turnip yellow mosaic virus* (TYMV) and of *Turnip mosaic virus* (TuMV), respectively. The amiR159-P69 was designed to target the sequence of the P69 silencing suppressor protein of TYMV. Similarly, amiR159-HCPro was designed to target the sequence of the TuMV silencing suppressor HC-Pro. The transgenic expression of the amiRs conferred high levels of specific resistance to the corresponding virus. Two independent transgenic lines, plants fully resistant to TuMV, named as 12-4, and plants partially resistant, designed as 10-4, have been used in our first experiment. Lin et al. (2009) demonstrated that despite the amiRNA technology gives full and specific resistance, it can be broken with only one point mutation with more impact depending of the position of the mutations.

1. Tempo and Mode of Plant RNA Virus Escape from RNA Interference-Mediated Resistance

1.1. Objectives

Changes in certain sites within the 21-nt target may generate virus escape variants. However, the relevance, if any, of these escape variants in natural viral populations remains to be established. In other words, to evaluate the viability of antiviral therapies based on the transgenic expression of amiRs in crops, it is essential to evaluate the likelihood of viral populations infecting susceptible reservoir host species to contain escape variants that may be subsequently transmitted to the transgenic crops by vectors. Moreover, it is also pivotal to evaluate whether variations in the expression of the amiR transgenes, especially at subinhibitory concentrations, would affect the accumulation and evolution of escape viral mutants. More specifically, we are interested in addressing the following issues. What is the likelihood of escape mutations arising and accumulating in a WT host population? Does partial resistance favor the accumulation of escape mutants? What sites in the 21-nt target are more critical for escape from RNAi surveillance? What are the basic population genetic parameters governing the escape process?

We performed two sets of evolution experiments together with their corresponding *in silico* simulations. In the first set, 25 independent TuMV populations evolving in fully susceptible wild-type *A. thaliana* plants were periodically tested for the presence of escape mutants by challenging fully resistant *A. thaliana* plants 12-4 (see Materials and Methods section). The second set of experiments was similar to the first one, except that 25 independent TuMV populations were evolved in partially susceptible *A. thaliana* plants 10-4 expressing subinhibitory concentrations of amiR159-HCPro. The *in silico* simulation algorithm was used to evaluate population genetic parameters governing the evolutionary dynamics of escape mutants.

1.2. Results

1.2.1 *A. thaliana* lines 10-4 and 12-4 differ in amiR159-HCPro expression and susceptibility to TuMV infection.

First, we evaluated whether TuMV had the same level of pathogenicity, *p*, in both *A. thaliana* transgenic lines 10-4 and 12-4. All plants were inoculated at Boyes' stage 1.03 (i.e., when three rosette leaves are greater than 1 mm in length) (Boyes et al., 2001) and with TuMV infectious sap applied by gentle abrasion onto leaves occupying the same position on the plant. None of the 30 inoculated 12-4 plants developed symptoms of infection at 14 dpi (p = 0.000±0.048 [±95% confidence interval, computed by using the Wald adjusted method]). In sharp contrast, 152 out of 218 inoculated 10-4 plants developed obvious symptoms after the same period of time (p = 0.697±0.063). The difference between the results obtained from 10-4 plants and those obtained from 12-4 plants was highly significant (P < 0.001 by Fisher's exact test). Significantly, TuMV pathogenicity in 10-4 plants was only 11.80% lower than that in the fully susceptible wild-type plants (166 out of 210; p = 0.791±0.057), although this small difference was still statistically significant (P = 0.035 by Fisher's exact test).

To elucidate the difference in pathogenicity between the two transgenic lines, we first evaluated whether there was any difference in the overall accumulation of amiR159-HCPro. To this end, we analyzed the concentration of the amiR accumulated in sets of each transgenic line at Boyes' stage 1.03 (the developmental stage at which the above-described pathogenicity tests were performed) by RT-qPCR. Twelve 12-4 plants and 11 10-4 plants were analyzed; three independent quantifications were obtained for each plant. The data were analyzed by using a general linear model (GLM), using "plant genotype" as the main random factor and "plant replicate" nested within the plant genotype. This analysis showed that significant heterogeneity exists among plants of the same genotype (χ^2 = 264.698; 21 df; P < 0.001). Despite this heterogeneity, the differences among genotypes were highly significant (χ^2 = 389.442; 2 df; P <

0.001). On average, 12-4 plants accumulated 111.367±6.998 pg amiR159-HCPro per mg of plant tissue (here errors represent ±1 standard error of the mean [SEM]), whereas 10-4 plants accumulated 4.961±1.370 pg/mg (i.e., 22.45-fold less than 12-4 plants).

Second, we characterized the temporal pattern of the accumulation of amiR159-HCPro in 10-4 leaves whose developmental stage was equivalent to the developmental stage of those inoculated in the pathogenicity tests (e.g., the zero in the ordinate corresponds to Boyes' stage 1.03). Four independent 10-4 plants were analyzed at each time point, and the estimates were averaged among plants. Fig. 15A shows that the amount of amiR159-HCPro accumulated per nanogram of plant total RNA increased in a nonlinear fashion as a leaf developed. Indeed, during the first days of the experiment the increase in the amiR159-HCPro concentration was minor, but accumulation significantly accelerated 10 days after the beginning of the experiment (i.e., accumulation was not linear but exponential) (Fig. 33A).

Third, we looked for differences in the amounts of amiR159-HCPro on different leaves of the same plants (at Boyes' stage 1.06; i.e., six rosette leaves are greater than 1 mm) to see whether this accumulation pattern was consistent among plants. To do so, we estimated the concentration of the amiR in each of six leaves from four different plants. Fig. 33B shows the observed pattern of amiR159-HCPro accumulation. A GLM model in which "plant" was treated as a random factor and "leaf" was treated as a covariable highlighted several interesting results. First, the amounts of amiR159-HCPro significantly varied among leaves at different developmental stages, significantly increasing as leaves became older ($\chi^2 = 88.713$; 1 df; $P < 0.001$). Second, in agreement with data from our first test, plants were heterogeneous in their average amounts of accumulated amiR159-HCPro ($\chi^2 = 497.603$; 4 df; $P < 0.001$). Third, differences existed among plants in the rate at which the concentration of amiR159-HCPro increased ($\chi^2 = 96.531$; 3 df; $P < 0.001$ [test for homogeneity of slopes]). In other words early stochastic events during

development determined the initial amount of amiR159-HCPro that would characterize a leaf, and these differences were further amplified as leaves expanded and developed.

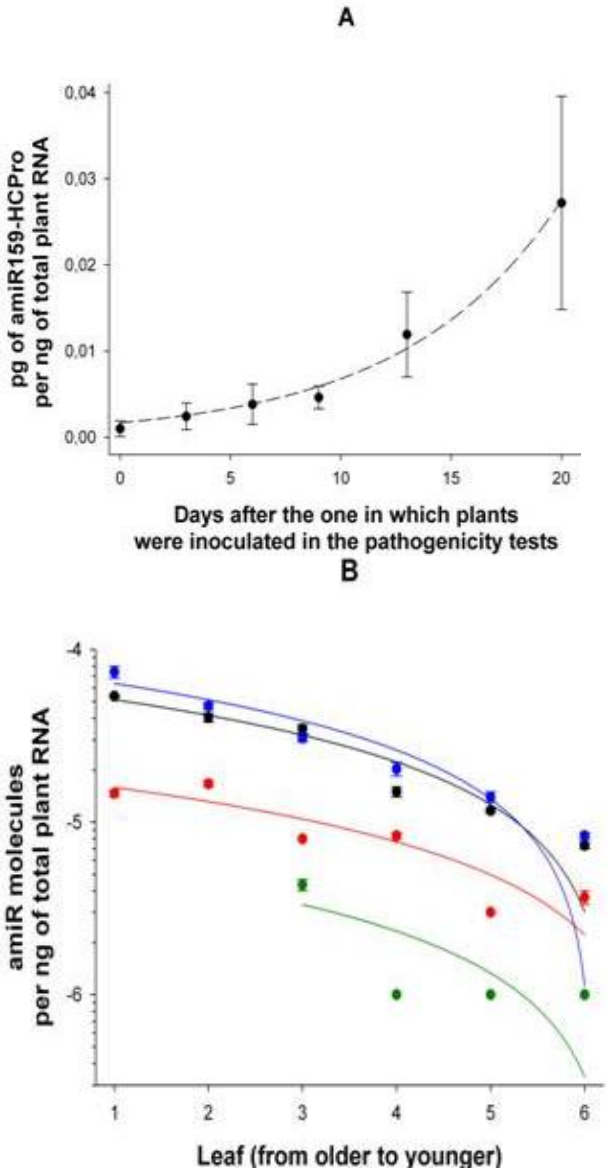


Figure 33. Pattern of amiR159-HCPro accumulation in the partially resistant transgenic lineage 10-4. (A) Curve of amiR159-HCPro accumulation in the

leaf inoculated in the pathogenicity tests (units of pg of amiR159-HCPro per ng of total plant RNA). The dashed line represents the fit to a two-parameter exponential growth model ($R^2 = 0.990$, $F_{1,4} = 405.167$, $P < 0.001$). (B) Pattern of amiR159-HCPro accumulation in six leaves that differ in their developmental stages from four different plants (units of amiR159-HCPro molecules per ng of total plant RNA). Each plant is represented by a different color. In all cases, error bars represent ±1 SEM.

Therefore, all these analyses led to the conclusion that the transgenic line 10-4 shows incomplete genetic penetrance (i.e., not all individual transgenic plants are resistant) and variable gene expressivity for resistance (i.e., not all resistant individuals express amiR159-HCPro at the same level). In contrast, line 12-4 showed a complete genetic penetrance of the resistance trait. These phenotypic differences are due to differences in the amount and timing of expression of amiR159-HCPro. Rather than being an issue, we will take full advantage of the 10-4 peculiarity to evaluate the effect of evolving TuMV populations on the subinhibitory and variable expression of amiR159-HCPro.

1.2.2 Resistance breaking in TuMV populations evolving in wild-type *A. thaliana* plants.

We aimed to evaluate the likelihood that TuMV populations replicating and evolving in fully susceptible wild-type *A. thaliana* hosts contained escape mutants able to overcome the resistance mediated by amiR159-HCPro. To this end, 25 independent evolution lineages were founded by inoculating wild-type *A. thaliana* plants with sap obtained from a pool of *N. benthamiana* plants previously inoculated with an infectious TuMV cDNA genome. Therefore, the amount of genetic variability in the inoculum will not be zero but the lowest amount technically possible. All plants were inoculated with the same amount of this infectious sap. All plants became infected, as confirmed by the presence of symptoms. Every 14 dpi, infected plants were sampled; one portion of the sample was used to inoculate the following set of plants, another portion was stored for future analyses,

and a third portion was used to challenge 20 12-4 transgenic plants per evolving lineage (total, 20 × 25 = 500 plants per challenge experiment) (Fig. 26). The pathogenicity of each evolving lineage at each passage was evaluated by an inspection of symptoms; a lineage was considered capable of breaking resistance if it was able to infect at least one 12-4 plant in the challenge experiments (i.e., pathogenicity of ≥ 0.05). We hypothesized that mutants in the amiR159-HCPro target would arise and stay in the population at the mutation-drift balance and that they would be transferred to the 12-4 plants during challenge in a rather stochastic manner. The black line in Fig. 34 shows the cumulative frequency of lineages that overcame resistance at each passage. The first break out occurred at passage six, and all 25 lineages were capable of breaking resistance after 28 passages. A Kaplan-Meier regression showed that the median time for resistance breaking was 14.000±0.480 passages in wild-type *A. thaliana* plants.

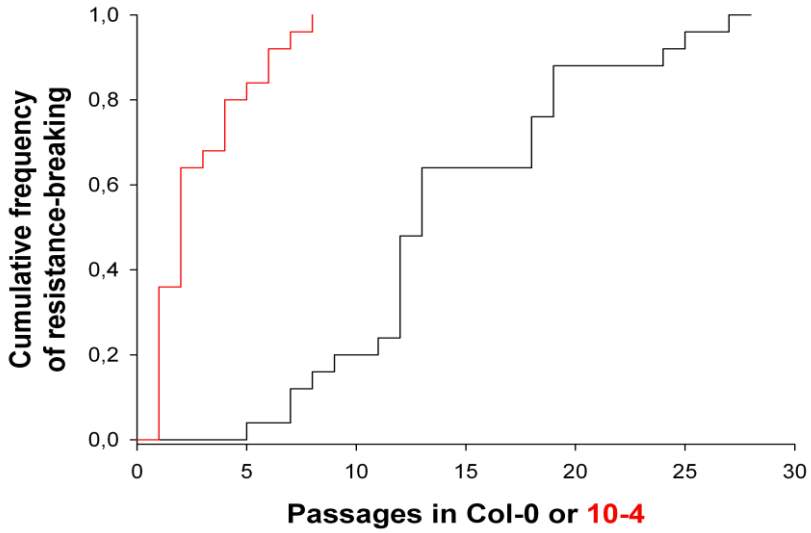


Figure 34. Cumulative frequency of lineages capable of escaping from amiR159-HCPro resistance. The black line corresponds to the lineages evolved in wild-type *A. thaliana* plants. The red line corresponds to the lineages evolved in partially resistant *A. thaliana* 10-4 plants. The ability of evolving TuMV populations to escape from amiR159-HCPro was evaluated

in 12-4 plants that were fully resistant to the ancestral TuMV genotype. A population was considered to be able to escape from resistance when at least one 12-4 plant was infected.

1.2.3 Resistance breaking in TuMV populations evolving in partially susceptible *A. thaliana* 10-4 plants

Next, we sought to evaluate the effect that TuMV replication under subinhibitory concentrations of amiR159-HCPro had on resistance durability. To this end, we repeated the evolution experiment by performing serial passages in partially resistant *A. thaliana* 10-4 plants; all other operations were kept identical. We reasoned that in this case, the TuMV populations infecting plants would be under the selective pressure imposed by the presence of the amiRNA in the cells but at concentrations that may still allow viral replication. We predicted that in such a situation, escape mutations would have a selective advantage and accumulate in the population at the mutation-selection-drift balance, at frequencies higher than those in the previous experiment. This would allow a faster resistance breaking after challenge of the 12-4 plants. The red line in Fig. 16 illustrates the time course accumulation of lineages able to break the resistance. As we predicted, lineages broke resistance faster than in the previous experiment, with many of them already containing escape mutants after the first passage and all 25 being able to do so after only eight passages. A Kaplan-Meier regression showed that in this case, the median time for resistance breaking was 2.000±0.343 passages in 10-4 plants, a value that is significantly lower than that obtained for the 12-4-evolved lineages (χ^2= 54.971; 1 df; $P < 0.001$ [by a Mantel-Cox test]).

1.2.4 Changes in amiR159-HCPro target

After determining that a TuMV lineage was capable of escaping from amiR159-HCPro-mediated resistance, we sought to characterize the genetic changes associated with its new phenotype. Based on results reported previously by Lin et al. with TuMV (Lin et al., 2009), supported by previous accumulated knowledge from HIV-1

(Boden et al., 2003; Das et al., 2004; von Eije et al., 2008; Westerhout et al., 2005) and poliovirus (Gitlin et al., 2005) cell culture experiments, we hypothesized that in all cases, the dominant TuMV genotype in the infected 12-4 plants after challenge would carry at least one mutation in the target sequence. To test this expectation, we obtained the 21-nt target consensus sequence for the viral population replicating in each 12-4 plant. Tables 5 and 6 show the different escape alleles found in TuMV populations evolving in wild-type *A. thaliana* and 10-4 plants, respectively.

Regarding Table 5, a total of 10 different alleles were characterized, although four of them (alleles 1, 2, 3, and 4) were pervasively seen in more than one lineage, a clear example of convergent evolution. The two most common nucleotide substitutions were a synonymous one at target site 11 (in 10 cases) and a nonsynonymous one at position 12 (in 7 instances), which gave rise to a conservative amino acid replacement, V to M, in the HC-Pro protein. Half of the alleles contained a single substitution (alleles 1, 2, 4, 8, and 10), whereas the other half contained two mutations. Four of these substitutions were synonymous, and eight were associated with amino acid replacements. Interestingly, lineages 6, 11 and 23 all showed a polymorphism at position 20 of the target. In all three cases one of the coexisting alleles was a synonymous substitution, whereas the other one involved a conservative amino acid replacement, K to N, in HC-Pro.

Table 5. Escape alleles found in the TuMV populations evolved in fully susceptible wild-type *A. thaliana* plants.
a The 21 nt of the target is underlined. The mutated sites are shown in boldface type.
b The lineage (and passage) in which each allele was observed is indicated.
c Indicates whether the mutations were synonymous or involved an amino acid replacement.

Allele (sequence)[a]	Lineage (passage)[b]	Type of mutation
ACA GUC GAG UGU GUG AGC AAG UUA	12(6), 18(9), 9(13), 7(14), 19(14), 21(14), 20(19)	synonymous
ACA GUC GAG UGC AUG AGC AAG UUA	10(10), 3(14), 8(14), 22(14), 24(14), 4(19), 16(20)	V→M
ACA GUC GAG UGU GUG AGC AAN UUA	6(14), 11(27), 23(27)	synonymous/synonymous or K→N
ACA GUC GAG UGC GUG AGU AAG UUA	14(13), 1(27)	synonymous
ACA GUC AAG UGC GUA AGC AAG UUA	15(8)	E→K/synonymous
ACA GUC GAG UGC GUG GGU AAG UUA	25(8)	S→G
ACA GUC GUA UGC GUG AGC AAG UUA	13(14)	E→V
ACA GUC GAG UGC GUG AGC AGG UUA	5(18)	K→R
ACA AUC AAG UGC GUG AGC AAG UUA	2(20)	V→I/E→K
ACA GUC GAG UGC GUG AGC GAG UUA	17(20)	K→E

Regarding Table 6, seven escape alleles were identified in the TuMV populations evolved in the partially resistant 10-4 plants. Four of them were not observed for the populations evolving in wild-type *A. thaliana* plants (alleles 11, 12, 13, and 14), although only one of the mutations in these alleles was not previously observed (the A-to-C nonsynonymous change at position 19 of allele 11). The two most common alleles in this experiment were also those observed in the first experiment (alleles 1 and 2): the synonymous substitution at position 11 of the target (in 12 cases) and the second most abundant one, the nonsynonymous replacement at site 12 (in 8 cases). Pooling data from both experiments, 52 of the 55 observed mutations were transitions, with G-to-A and C-to-U changes dominating the mutational spectrum. Consistent with the principle that transitions are biochemically more likely to occur than transversions, the maximum composite likelihood estimate of the overall transition-to-transversion rate ratio is 14.176. This excess also occurs when purines (ratio, 20.599) and pyrimidines (ratio, 40.639) are considered separately.

It is well known that viral coding regions show an excess of transitions over transversions (Burch et al., 2007; Haydon et al., 1998; Lin et al., 2009; Tromas and Elena, 2010). Three reasons can account for this bias: (i) the underlying mechanisms of mutation render transitions easier than transversions, (ii) the redundancy of the genetic code is expected to make the average effect of transitions smaller than that of transversions, and (iii) RNA editing by deaminase-like enzymes has been shown to induce transition mutations in single-stranded regions of certain viral genomes (Bishop et al., 2004).

Table 6. Escape alleles found in TuMV populations evolved in partially resistant 10-4 plants. The 21 nt of the target is underlined. The mutated sites are shown in boldface type.

Allele (sequence)[a]	Lineage (passage)	Type of mutation
ACA GUC GAG UGU GUG AGC AAG UUA	3(1), 6(1), 8(1), 13(1), 15(1), 14(2), 21(2), 25(2), 7(4), 11(4), 22(5)	synonymous
ACA GUC GAG UGC AUG AGC AAG UUA	2(1), 9(1), 17(2), 10(3), 4(4), 18(6), 24(7), 20(8)	V→M
ACA GUC GAG UGC GUG AGC ACG UUA	1(2), 19(6)	K→T
ACA GUC AAG UGU GUG AGC AAG UUA	5(1)	E→K/synonymous
ACA AUC GAG UGC GUG AGC AAG UUA	16(1)	V→I
ACA GUC AAG UGC GUG AGC AAG UUA	12(2)	E→K
ACA GUC GAG UGC GUG AGU AAG UUA	23(2)	synonymous

Convergent evolution would imply that the frequency distribution of changes along the 21-nt target should be similar in both experiments. Fig. 35 shows these distributions for both types of TuMV populations. A homogeneity test detected no differences between both pattern distributions ($\chi^2 = 8.388$; 11 df; $P = 0.678$), thus supporting the notion of widespread convergent evolution, likely driven by the selective advantage of mutations at sites 11 and 12 of the target.

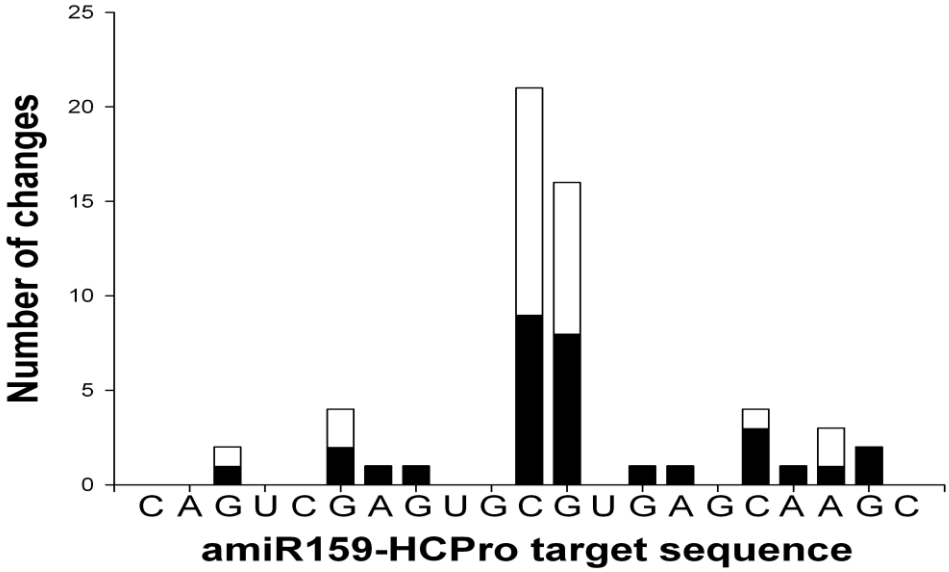


Figure 35. Distribution of mutations in escape viruses along the amiR159-HCPro target sequence. Black bars correspond to the frequency of mutations that arose in wild-type *A. thaliana* plants; white bars correspond to those observed for the 10-4 transgenic line.

1.2.5 Estimates of population genetic parameters by *in silico* simulations

To provide new insights into the above-described results as well as to evaluate the range of population parameters compatible with our observations, we simulated the two evolution experiments by using digital viral genomes replicating, mutating, and being subjected to transmission bottlenecks as in the experiments described above (Fig. 27 and see Materials and Methods). We performed a search of parameter space using an optimization algorithm (OA) to find a set of parameters that minimized the distance between the data shown in Fig. 16 and those simulated. For the simulations of the evolution experiments carried out with wild-type *A. thaliana* plants (i.e., without sequence-specific degradation), we analyzed a total of 393 runs of the

OA: 129 runs assuming that mutations had additive effects, 210 runs assuming that they interacted epistatically, and 54 runs assuming that mutations outside the target and affecting other genes were lethal. Each run of the OA consisted of 400 generations, with a population of parameter sets of 150, resulting in more than 25 million simulations. The parameter set that generated the lowest and more robust distance (d = 0.56; R^2 = 0.976; $F_{1,27}$ =1114.571; P < 0.001) between the experiments and the simulation model was obtained with the additive fitness landscape with the following parameters: τ = 13918.23±75.64 viral replications between passages, μ = (4.11±0.33)×10^{-5} mutations per site and generation, and N_e = 956.89±23.76 digital viruses transmitted per bottleneck event (i.e., ~19% of the total population). Fig. 36A shows the results of the simulation obtained with this set of parameters. The simulated values of the frequency of lineages escaping from amiR159-HCPro are shown with red dots on top of the black line that represents the experimental data.

For the evolution experiments in the partially resistant 10-4 plants (i.e., with sequence-specific degradation), we followed the same procedure, although we restricted the study to only the additive fitness landscape (that gave the best fit for the wild-type plants) and added a degradation rate, ε > 0, to the parameter set. This degradation rate simulated the assumption that 10-4 plants expressed amiR159-HCPro and that; hence, the silencing machinery may still be capable of degrading a fraction of the viral population (i.e., the strings containing a wild-type target sequence are degraded with probability ε [see Materials and Methods]). For this case, we ran 150 replicas of the OA, thus exploring a total of 6 million simulations. Among all these simulations, the parameter combination providing the smallest distance between experimental and simulated data (d = 0.16; R^2 = 0.995; $F_{1,7}$ = 733.253; P < 0.001) was as follows: τ = 5629.51±63.79, μ = (7.69±1.12)×10^{-5}, N_e = 68.61±11.78 (i.e., ~1.4% of the potential maximum population size), and ε = 0.223±0.098 per genome. The best fit to the experimental data is shown in Fig. 36B. As described above,

the red dots represent the simulated values for this parameter set. Not surprisingly, the mutation rates estimated for both experiments are on the same order of magnitude and close to the only experimental value reported previously for potyviruses (Tromas and Elena, 2010).

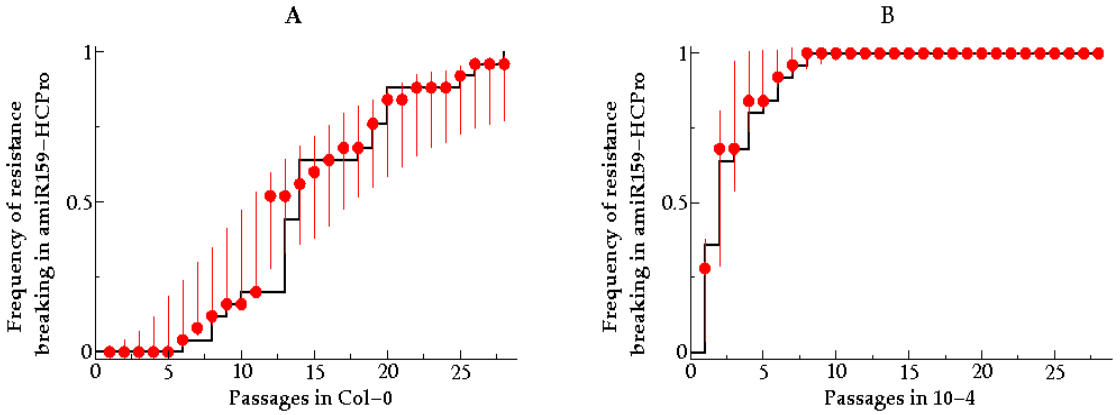


Figure 36. Results of the simulation studies for the set of parameters that showed the best fit to data shown in Fig. 3**.** (**A**) Simulation results for the WT *A. thaliana*-evolved TuMV lineages. (**B**) Simulation results for the TuMV lineages evolved in partially resistant 10-4 plants. The red dots correspond to the best-fitting trajectory obtained from the most optimized parameter set. Red bars indicate the standard deviations among 103 runs of the simulation model using the best-fitting parameters.

The degradation of genomes containing no mutated amiR159-HCPro targets in 10-4 plants has two interlinked effects. First, there was a reduction of 92.83% in N_e: not all genomes contained in the inoculum were capable of replicating in the partially susceptible plants, and a certain fraction was degraded. Second, we expected an apparent reduction in the number of viral replication events supported by the two plant genotypes. In Col-0 plants all the viral progeny produced may eventually contribute to future replications. On the other hand, in 10-4 plants, we expected part of the progeny to be degraded by the amiRs and, hence, not contribute to future replications. The model catches this expectation and shows that 10-4 plants supported ~2.5-times-fewer replication events than the wild-type plants. Consistently, the viral populations replicating in 10-4

plants did not reach carrying capacity, and therefore, the number of genomes transmitted to the next infection cycle was 13.95-fold lower. This reduction in the size of the transmitted population enhances the effect of genetic drift in the 10-4 lineages. That being said, it is important to recall that two different evolutionary regimes are in play for each plant genotype. In the fully susceptible wild-type plants, purifying selection and drift should be the only factors affecting allele frequencies, since mutations in the target would be either deleterious or neutral; deleterious alleles will not reach high frequencies. The time to fixation of a neutral allele whose initial frequency is negligible is $4N_e$ = 3827.56 generations (Hedrick, 2004), which is less than the estimated number of viral replications, τ, thus making it likely that some neutral alleles in the target would drift to high frequencies in the population. In contrast, in partially resistant 10-4 plants, positive selection also enters the picture, since escape alleles will clearly be beneficial in the presence of amiR159-HCPro. Indeed, we estimate that the average selection coefficient for such a beneficial allele to survive drift should be $s > 1/N_e = 0.015$ (Hedrick, 2004), a low value that ensures that many mutations conferring resistance will survive drift.

Finally, the mutation rates estimated for both experiments are in the range of 4×10^{-5} to 8×10^{-5} mutations per site, values that are very close to recent estimates obtained for another potyvirus, TEV (Sanjuan et al., 2009; Tromas and Elena, 2010), and, more generally, for other plant viruses (Malpica et al., 2002; Sanjuan et al., 2009). This excellent agreement gives support to the validity of our modeling approach as well as to the conclusions derived from it.

1.3. **Discussion**

The long-term effectiveness of genetic resistances to plant viruses is constantly being challenged by the evolutionary potential of RNA viruses (Garcia-Arenal and McDonald, 2003), creating the necessity for the development of new resistance strategies. In the early 1990s it was recognized that the transgenic expression of virus-derived sequences resulted in a highly efficient defense against plant viruses (Lindbo and Dougherty, 2005), with this defense being mediated by the posttranscriptional degradation of RNA genomes guided by virus-derived small interfering RNAs (siRNAs) (Hamilton and Baulcombe, 1999). In recent years, plants that are resistant to virus infection have been engineered by use of this approach (Chen et al., 2004; Di Nicola-Negri et al., 2005; Kalantidis et al., 2002; Missiou et al., 2004; Wang et al., 2000). However, the transgenic expression of long viral sequences raises biosafety concerns regarding the possibility of recombination and the generation of new and potentially virulent strains (Turturo et al., 2008). Taking advantage of the functional similarities between siRNA and miRNAs, Niu et al. (2006) modified the backbone of *A. thaliana* pre-miRNA159, replacing it with short 21-nt viral sequences, resulting in highly specific resistant plants. This approach has at least two advantages compared with the expression of long viral sequences. First, it should have fewer off-target effects, as the amiR sequences are shorter than those required for homology-dependent gene silencing. Second, recombination is not a concern anymore, given the shortness of the amiRs. However, this approach may still raise a major concern: the high mutability of RNA viruses makes it likely that resistant virus variants will emerge, as already observed in *in vitro* experiments with mammalian viruses (Boden et al., 2003; Das et al., 2004; Ge et al., 2003; Gitlin et al., 2005; Kronke et al., 2004; Sabariegos et al., 2006). The objective of the present work was to evaluate the likelihood of the emergence of such escape variants in viral populations replicating in fully susceptible reservoir plants as well as in plants expressing resistance at subinhibitory levels. Toward accomplishing this objective, we have

performed two different evolution experiments using the pathosystem TuMV/*A. thaliana*, together with *in silico* computational models simulating both evolution experiments. The first experiment was designed to mimic a situation in which crops of resistant transgenic plants coexisted with crops of fully susceptible ones that acted as virus reservoirs. In this case, we observed an increase in the number of evolving lineages that were capable of successfully infecting the fully resistant host. Such escape mutants should most likely be neutral, or perhaps even slightly deleterious, maintained by complementation, in the evolving population. Our second experiment was aimed at mimicking a situation in which the expression level of the antiviral amiR was variable among plants, with some of them having suboptimal levels that allow virus replication and the selection of escape variants. In the second case we found that these populations accumulated escape mutations at a much higher frequency and, therefore, were able to successfully infect the fully resistant hosts at earlier times in virus evolution. The second result was highly predictable, since it recapitulates the evolution of bacteria at antibiotic concentrations below the MIC (Couce and Blazquez, 2009) and has been solidly established. At subinhibitory concentrations of the antiviral amiR, mutant genotypes gain a fitness advantage, given their ability to replicate despite the presence of the antiviral amiR, whereas wild-type genomes may still suffer from the inhibitory effects. This fitness advantage results in the accumulation of escape alleles above what was expected from the first experiment.

In all 50 cases, the molecular characterization of the escape mutants confirmed the presence of mutations in the amiR159- HCPro target. In agreement with the mutant spectra described previously for other viruses, including TuMV, we have observed an excess of transition mutations (Burch et al., 2007; Haydon et al., 1998; Lin et al., 2009; Tromas and Elena, 2010). Particularly interesting is the fact that G-to-A and C-to-U transitions represented 95% of all mutations observed. These transitions are from the particular type induced by cellular cytidine deaminases involved in innate immune responses to

viral infection (Conticello et al., 2005), a phenomenon particularly well described for HIV-1 and other retroviruses (Cullen, 2006) but hitherto not described for RNA viruses. This observation is in good agreement with those described previously Lin et al. (2009), thus giving additional support to the hypothesis that as an antiviral strategy, plants may have an RNA-editing system that induces hypermutagenesis in viral genomes. We note that *A. thaliana* contains a family of nine paralogous genes that have been annotated cytidine deaminases owing to their homology to the *CDA1* locus (Vincenzetti et al., 1999).

Indeed, mutations were unevenly distributed along the 21-nt target and concentrated mainly at positions 11 and 12, in a clear case of convergent evolution at the molecular level. Convergent evolution is a widespread phenomenon in RNA viruses both in experimental (Bull et al., 1997; Cuevas et al., 2002; Wichman et al., 1999) and in natural (Martinez-Picado et al., 2000) populations. Although these convergences could in principle be explained from a neutralist point of view as resulting from mutational bias, it is more likely that parallel and convergent substitutions are adaptive. This pattern would result from viruses facing identical selective pressures, with few alternative adaptive pathways, as expected for their simple and compacted genomes. In agreement with our observation, Lin et al. (2009) classified position 11 as being moderately crucial and position 12 as being critical for resistance breaking, although other sites qualified as being crucial did not show a high frequency of variation in our experiments. In contrast to the study by Lin et al. (2009), in which the targeted sequence was neutral to the virus, here amiR159-HCPro targeted a coding region of the TuMV HC-Pro cistron, and consequently, mutations in escape variants must result form the balance between avoiding recognition by amiR159-HCPro and retaining biological function. Indeed, this coding effect may explain why Lin et al. observed an excess of critical positions at the 5' end of the amiR. Additionally, a potential explanation for convergence in these two central sites relies on the fact that imperfect pairing with

central mismatches in small RNA-target hybrids promotes translational repression as it excludes slicing (Brodersen et al., 2008). This observation suggests the possibility that imperfect pairing between the amiR and mutated targets might lead to translational repression rather than viral RNA cleavage. In contrast to the catalytic effects of amiR-mediated viral RNA cleavage, translational repression requires stoichiometric amounts of amiRs and therefore is not as efficient. Inefficient translation inhibition might allow residual viral replication, and progeny virus can still escape repression by fixing changes in the target sequence.

All in all, our results suggest that the durability of amiR-based resistance may be too short in time to make it a profitable approach. However, this assertion has to be carefully considered in the context that we designed our experiments in such a way that they represent the most favorable possible situation for resistance breaking. For instance, our challenge experiments were done with inocula that represent 1 to 20% of the whole viral population, according to our simulations. In a natural situation in the field, transmission would be mediated by vectors, which impose more dramatic bottlenecks, in the order of units per vector and transmission event (Ali et al., 2006; Betancourt et al., 2008; Moury et al., 2007), thus minimizing the likelihood of the transmission of very-low-frequency escape alleles, although large vector populations will contribute to an increase of the chances of transmission. Furthermore, the way in which we sampled viral populations, homogenizing the whole plant, provided transmission probability for all genomes present in the plant. The spatial structure imposed by the plant architecture limits gene flow among distal parts of the plant, up to the point that each part may be dominated by different viral genotypes (Jridi et al., 2006). This means that variants may not reach a high frequency within the whole metapopulation despite having some local fitness advantage. Therefore, by feeding on particular leaves, vectors would miss loading escape mutants that may be abundant in other parts of the plant. All these factors, plus surely some additional ones, increase the stochasticity of escape alleles

spilling over from their reservoirs to the amiR transgenic crops, thus perhaps increasing resistance durability. Another factor that may affect durability, as suggested by our results, is the level at which the amiR is expressed. We have shown that subinhibitory expression levels would indeed select for resistance alleles, facilitating their spread in transgenic populations. This adds a cautionary note for biotechnologists when selecting their new transgenic plants. Another way of increasing the resistance durability could be to express more than one amiR in a transgenic crop, to target different highly conserved RNA sequences in the viral genome, or to combine amiR-mediated resistance with other genetic resistances. By combining multiple amiRs into a single plant, the likelihood of resistance breaking will drop exponentially. This possibility is explored later on this thesis.

2. How design a better antiviral technology using mutation rate and virus constraint evolution data

2.1. Phylogeographic analysis of TuMV

As we have already discussed in the Introduction, viral genomes can evolve and have their sequence modified by mutation or recombination. As we have shown, despite the target sequence is only 21-nt long, the high mutability of RNA viruses can have a significant impact on breaking down the amiR159HC-Pro-mediated resistance. An important question, is how much natural variation exists that may jeopardize amiR159HC-Pro-mediated resistance against TuMV. It is important to know and integrate the natural variability of TuMV for the design of amiR.

The whole genome of ca. 100 TuMV isolates was available in GenBank at the time we started this study (Table 7). The majority of these isolates come from Asia (Fig. 37).

In the previous study, our evaluation of the likelihood that TuMV escape variants will emerge under different contexts already

provides some clues to enhance the durability of amiR-mediated resistance in transgenic plants. To go further to understand the natural variability and its impact on the durability of the amiR strategy, we have analyzed the relationship between natural isolates and the different rate of evolution of ten TuMV genes.

Table 7. List of the 100 TuMV genomes isolated worldwide used for our phylogenetic study.

Isolate	Original Host	Location	Year of collection	GenBank accession	Group
1J	*Raphanus sativus*	Japan	1977	D83184	
2J	*Brassica pekinensis*	Japan	1994	AB093622	5
59J	*R. sativus*	Japan	1996	AB093620	3
A102/11	*Anemone coronaria*	Italy	1993	AB093597	1
A64	*A. coronaria*	Italy	1991	AB093599	1
AD178J	*R. sativus*	Japan	1998	AB252094	
AD181J	*R. sativus*	Japan	1998	AB252095	
AD853J	*R. sativus*	Japan	2002	AB252096	
AD855J	*R. sativus*	Japan	2002	AB252097	
AD860J	*R. sativus*	Japan	2002	AB252098	
AKD161J	*R. sativus*	Japan	1998	AB252099	
AKD934J	*R. sativus*	Japan	2000	AB252100	
AKH937J	*B. pekinensis*	Japan	2000	AB252101	
Al	*Alliaria officinalis*	Italy	1968	AB093598	1
AT181J	*Eustoma russellianum*	Japan	1998	AB252102	
BZ1	*B. oleracea*	Brazil	1996	AB093611	6
C1	Not known	China	2001	AF394601	
C42J	*B. rapa*	Japan	1993	AB093625	5
Cal1	*Calendula officinalis*	Italy	1979	AB093601	2
CAR37	*Cochlearia armoracia*	Poland	2004	DQ648592	
CAR37A	*C. armoracia*	Poland	2004	DQ648591	
CAR39	*C. armoracia*	Poland	2004	EF374098	
CDN1	*B. napus*	Canada	1989	AB093610	6
CH6	*R. sativus*	China	1999	AB252103	
CHK16	*R. sativus*	China	2000	AB252104	
CHL13	*R. sativus*	China	1999	AB252105	5
CHN1	*Brassica sp.*	Taiwan	1980	AB093626	
CHN12	Not known	China	1990	AY090660	

CHZJ26A	*B. campestris*	China	1999	AB252106	
CP845J	*C. officinalis*	Japan	1997	AB093614	2
CZE1	*B. oleracea*	Czech Republic	1981	AB093608	4
CZE5	*B. rapa*	Czech Republic	1994	AB252107	
DMJ	*R. sativus*	Japan	1996	AB093623	5
DNK2	*B. napus*	Denmark	1993	AB252108	
FD27J	*R. sativus*	Japan	1998	AB093618	
FKD001J	*R. sativus*	Japan	2000	AB252109	
FKD004J	*R. sativus*	Japan	2000	AB252110	
FKH122J	*B. pekinensis*	Japan	1998	AB252111	
FRD1	*B. oleracea*	Germany	1987	AB252112	
GBR36	*B. oleracea*	UK	1999	AB252113	
GBR50	*B. oleracea*	UK	2000	AB252114	
GFD462J	*R. sativus*	Japan	2001	AB252115	
GRC17	*B. oleracea*	Greece	1993	AB252116	
GRC42	*Allium sp*	Greece	1999	AB252117	
H1J	*R. sativus*	Japan	1996	AB252118	
HOD517J	*R. sativus*	Japan	1998	AB093617	
HRD	*R. sativus*	China	1998	AB093627	3
HZ6	*Brassica sp.*	China	1998	AB252119	
IS1	*Allium ampeloprasum*	Israel	1993	AB093602	1
ITA3	*Brassica sp*	Italy	1994	AB252122	
ITA7	*Raphanus raphanistrum*	Italy	1994	AB093600	2
IWD032J	*R. sativus*	Japan	2000	AB252120	
IWD038J	*R. sativus*	Japan	2000	AB252121	
Ka1J	*B. pekinensis*	Japan	1994	AB093624	5
KD32J	*R. sativus*	Japan	1998	AB093621	3
KEN1	*B. oleracea*	Kenya	1994	AB093605	5
KGD54J	*R. sativus*	Japan	1998	AB252123	
KWB778J	*B. oleracea*	Japan	2004	AB252124	
KWB779J	*B. rapa*	Japan	2004	AB252125	
KYD073J	*R. sativus*	Japan	2000	AB252126	
KYD81J	*R. sativus*	Japan	1998	AB093613	2
MED302J	*R. sativus*	Japan	2001	AB252127	
MYD013J	*R. sativus*	Japan	2000	AB252128	
MYD015J	*R. sativus*	Japan	2000	AB252129	
ND10J	*R. sativus*	Japan	1998	AB252130	
NDJ	*R. sativus*	Japan	1997	AB093616	
NID048J	*R. sativus*	Japan	2000	AB252131	
NID119J	*R. sativus*	Japan	1998	AB252132	

NLD1	*B. oleracea*	The Netherlands	1995	AB252133	
NRD350J	*R. sativus*	Japan	2001	AB252134	
NZ290	*B. pekinensis*	New Zealand	1998	AB093612	6
PV0104	*L. sativa*	Germany	1993	AB093603	2
PV376-Br	*B. napus*	Germany	1970	AB093604	6
Q-Ca	*B. rapa*	Canada	1991	D10927	6
RC4	*Zantedeschia sp.*	China	2000	AY134473	
Rn98	*Ranunculus asiaticus*	Italy	1997	AB252135	
RUS1	*Armoracia rusticana*	Russia	1993	AB093606	4
RUS2	*B. napus*	Russia	2001	AB093607	4
SGB088J	*B. rapa*	Japan	2000	AB252136	
SGD311J	*R. sativus*	Japan	1998	AB093619	
SMD060J	*R. sativus*	Japan	2000	AB252137	
St48	*Limonium sinuatum*	Italy	1993	AB093596	1
TANX2	*R. sativus*	China	2007	EU734433	
TD88J	*R. sativus*	Japan	1998	AB093615	
TRD052J	*R. sativus*	Japan	2000	AB252138	
TRD053J	*R. sativus*	Japan	2001	AB252139	
Tu-2R1	*R. sativus*	Japan	2003	AB105135	
Tu-3	*B. oleracea*	Japan	2003	AB105134	
TuR1	Not known	Turkish	2005	AB362512	
TuR9	Not known	Turkish	2005	AB362513	
TW	Not known	China	2001	AF394602	5
UK1	*B. napus*	UK	1975	AF169561	
USA1	*B. oleracea*	United States	1980	AB093609	5
WFLBO6	*R. sativus*	China	2006	EU734434	
YAD020J	*R. sativus*	Japan	2000	AB252140	
YAL018J	*Lactuca sativa*	Japan	2000	AB252141	5
YC5	*Zantedeschia sp.*	China	2000	AF530055	
YMD069J	*R. sativus*	Japan	2000	AB252142	
YMD070J	*R. sativus*	Japan	2000	AB252143	

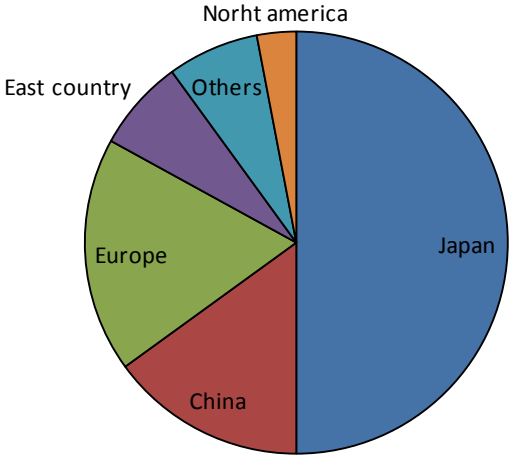


Figure 37. Geographic distribution of the 100 complete sequenced TuMV isolates available in GenBank.

Phylogenetics studies on 30 representative and "non recombinant" (Tomimura et al., 2003) TuMV isolates shows a West European origin and a recent emergence in Eastern Asia. We reconstructed the phylogeny of TuMV according to the concatenated genes. If we represent these isolates in a ML unrooted phylogenetic tree 6 groups can be distinguished (Fig. 38.I).

The first group is formed exclusively by isolates from Europe (Fig. 38). Then all other groups contain Asian isolates, except for the group 4 specific of East Europe and group 6 which is shared between worldwide isolates. To extend our knowledge on the relationship between these isolates, we determine the phylogenetic tree of each gene individually (Fig 38).

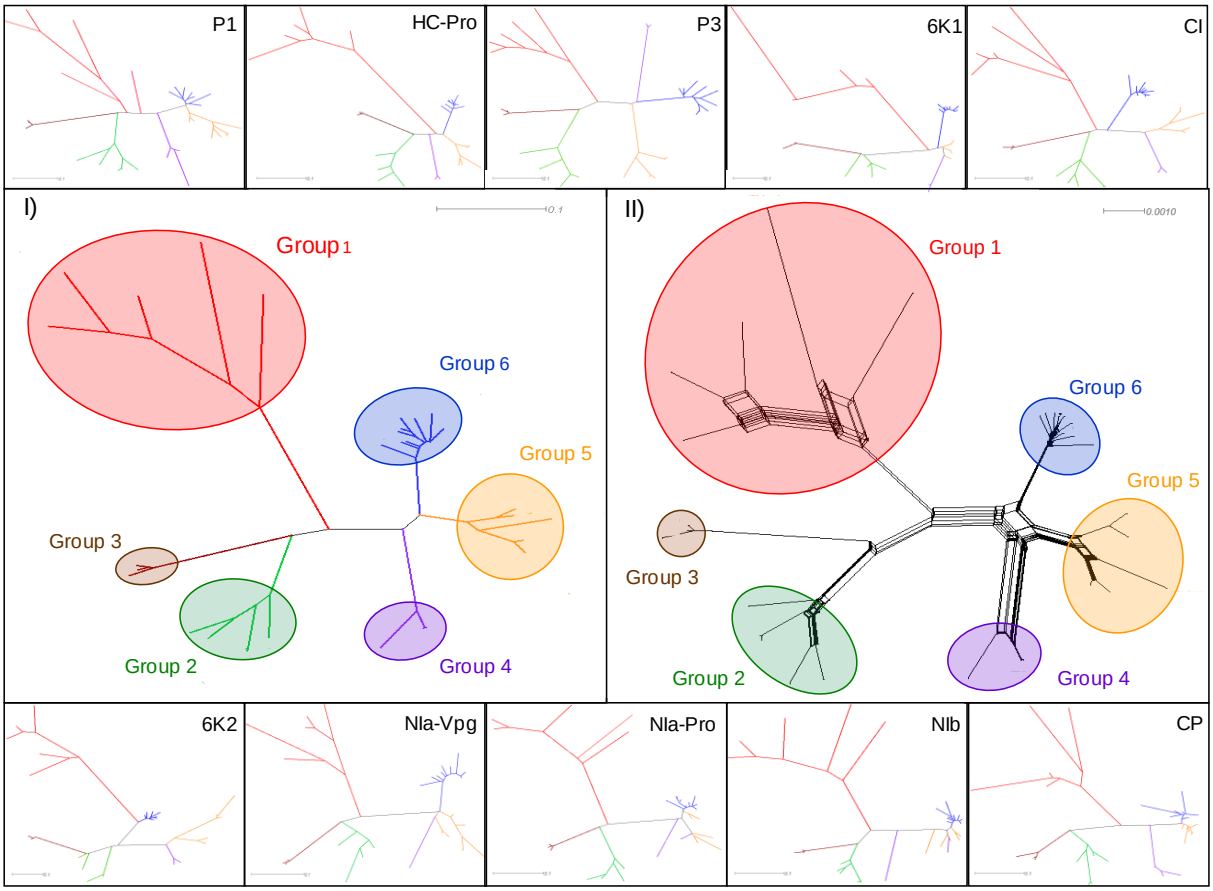


Figure 38. Phylogenetic reconstructions from TuMV genomes at nucleotide level. At the center, the phylogenetic tree of the complete genome (I) and the phylogenetic network (II) obtained by the integration of phylogenetic trees of all genes. The small boxes in the top and bottom of the figure contain the phylogenetic trees computed for each cistron.

Phylogenetic tree of each gene contain these six groups with support values higher than 70 (Table 8). However we can see a major perturbation for the fourth and fifth basal clade for a cluster of genes (from VPg to NIb) in which taxa are mixed between the two clades. The phylogenetic network (Fig. 38.II) shows the six clades defined. The perturbation found in VPg, NIa-Pro and NIb is represented among clades 4, 5 and 6. Moreover, a split network using all gene trees

obtained by the ML-approach is generated (Fig 38.II). This last analysis demonstrates the existence of possible alternative evolutionary pathway between the isolates.

Table 8. Support of groups for each tree gene and for concatenated genes of the 30 TuMV isolates.

TuMV	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
P1	92 (except St48_Italy)	100	100	100	88	-
HC-Pro	100	100	100	100	99	100
P3	100	100	100	100	100	98
6K1	82	79 (except ITA7_Italy)	96	78	-	87
CI	100	100	100	100	86	100
6K2	100	-	99	78	86	74
Nia-Vpg	100	88	100	-	-	98
Nia-Pro	100	100	99	-	-	43
Nib	100	99	100	-	-	89
CP	100	78	97	100	53	53
Concatenated genes	100	100	100	100	92	100

In order to detect incongruence in gene tree topologies, suggested by the split network, different methods have been developed. For example, the TopD (TOPological Distance) software (Puigbo et al., 2007) measures the means of the symmetric difference between two topologies. This distance equals to the minimum number of elementary operations, consisting of merging and splitting nodes necessary to transform one tree into the other. In this algorithm non-shared long branches have a same contribution than small conflicting branches. When two topologies are the same, the algorithm returns a value of zero, however when the topologies are completely different this value ups to one. Ktreedist (Castresana, 2007) is another software, which takes into account both, the topology and the branch length between two trees, scaling one to result in equal global evolutionary rates. In order to determine whether the differences observed in both, topology and branch length between the gene and the concatenate trees –the last one used as reference tree, because it should reflect the average divergence of all genes– from TuMV, the Robinson and

Foulds and Ktreedist algorithms were performed at the nucleotide and amino acid level (Table 9).

Table 9. Comparison of the concatenated tree vs. each cistron's trees of TuMV.

Individual TuMV's genes vs. Concatenated	Split		K-score		Scale-Factor	
	nt	aa	nt	aa	nt	aa
P1	0,37	0,39	0,15	0,06	1,04	0,27
HC-Pro	0,74	0,42	0,18	0,08	1,24	1,36
P3	0,74	0,42	0,24	0,08	1,12	0,41
6K1	0,44	0,91	0,19	0,12	1,36	0,46
CI	0,74	0,46	0,21	0,07	0,97	1,99
6K2	0,44	0,56	0,19	0,11	0,77	0,36
NiaPro	0,56	0,56	0,25	0,09	0,87	0,57
NiaVpg	0,56	0,70	0,18	0,10	0,85	1,72
Nib	0,56	0,56	0,27	0,08	0,79	1,28
CP	0,63	0,67	0,27	0,09	0,48	1,26

The case of TuMV genes is interesting in the asymmetry between the nucleotide and amino acid level found for the split distance. Indeed, HC-Pro, P3 and CI reach a high level at the nucleotide level (0.74) contrasting with the amino acid result (0.45). Except for P1 which have a quite good congruent tree (0.37 and 0.39 at the nucleotide and amino-acids level respectively) the others genes have a Robinson and Foulds value around 0.5 (Table 9).

For all genes the K-score is low, with some cases being close to zero. Indeed if we report to the different TuMV genes trees (Fig 38) few mismatches are encounter between them and the concatenated tree (Fig 38.II).

The scale factor also presents some particularities; with a higher value at the nucleotide level than the amino acid level, P1, P3 and 6K1 indicate a big influence of nucleotide changes on the protein sequence. Contrasting this case, CP shows a robustness of the amino acid sequence change for nucleotide variation with a high scale factor at amino acid level (1.26) comparing the nucleotide level (0.48) (Table 9).

Another parameter to evaluate is the tree length (substitutions per site) (Fig. 39). In the case of trees obtained from nucleotide sequences, this parameter can be estimated for individual codon positions. If we analyze the sum of the tree length of the three codon positions we can distinguish three different rates of nucleotide changes among the ten TuMV genes. The four first genes in the TuMV genome (P1, HC-Pro, P3, and 6K1) are the more variable. The five following genes (CI, 6K2, NIa-Pro, VPg and NIb) have a reduced rate of nucleotide change. Finally the last gene, CP, is the most conserved among all TuMV genes. P1 have more flexibility to accept changes in its protein sequence than others genes, followed by P3 and 6K2. Indeed for P1 the first and second codon positions participate more in its tree length than for the others genes.

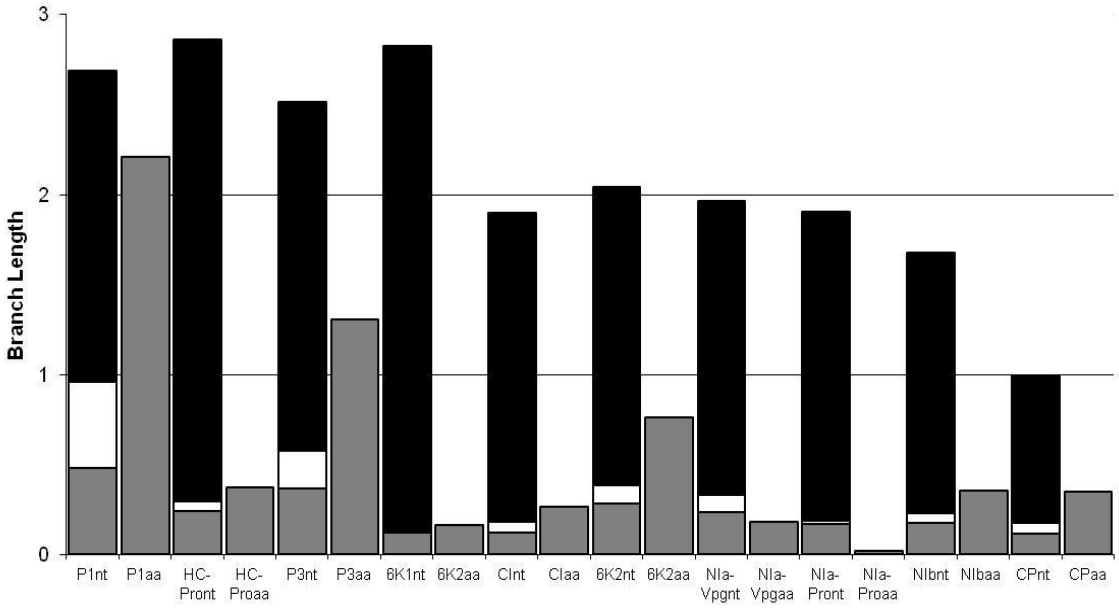


Figure 39. Tree length (number of substitution per site) of each codon position for each TuMV's genes. The nucleotide tree length in light grey, white and black represent the first, second and third codon position. The second column in light grey represents the tree lengths computed from amino acid sequences.

To evaluate if some positions are under selective pressure the Selecton software calculated the Ka/Ks of each position identifying site-specific positive and purifying selection of TuMV genes (Fig 40). With a high Huber M-estimation of Ka/Ks characterised by a 10% of its sequence under positive selection, P1 being the exception. We can observe a majority of cases of purifying selection (80%) for half of the genes (HC-Pro, 6K1, CI, NIa-Pro and NIb). For the other genes, we have a balance between neutral evolution and purifying selection. This result can suggest a global evolution of TuMV by purifying selection, indeed over the 3152 positions across the TuMV genome 1963 (62.3%) are under purifying selection, 1140 (36.1%) are neutral and only 49 (1.6%) are under positive selection.

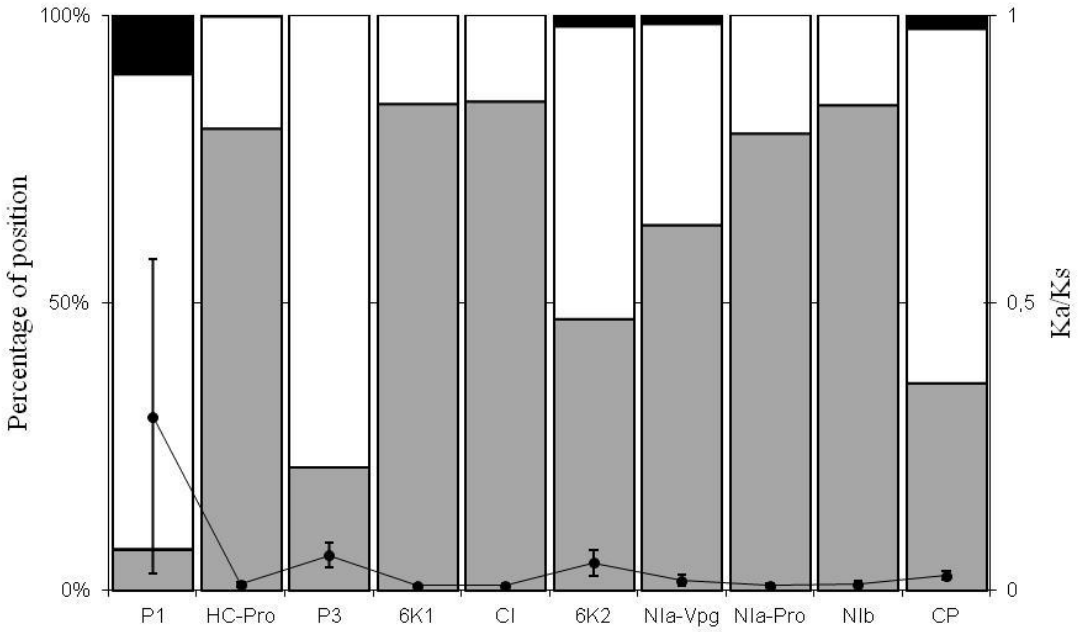


Figure 40. Percentage of sites under positive (black), neutral (white) and purifying selection (light grey) according to the Selecton software for each of TuMV genes (left scale). The solid black line represents the Huber M-estimation of Ka/Ks of all position for each gene (right scale).

For each tree corresponding of each individual gene, pairwise distances of all possible paired terminal taxa were calculated, both for nucleotides and for amino acid sequences. In order to compare our data between each genes of the same genome, the pairwise distances were normalised to the concatenate sequence, and then, M-estimator were calculated for nucleotide and amino acid sequences and plotted on Fig. 41. With this figure we have a global vision on the different rates of evolution of TuMV genes, with some genes being more variable in amino acids, like P1, or both in nucleotide and amino acids, like P3. Contrasting this we can observe conserved genes in amino acid like HC-Pro or 6K1 or conserved in both nucleotide and amino acid as CP.

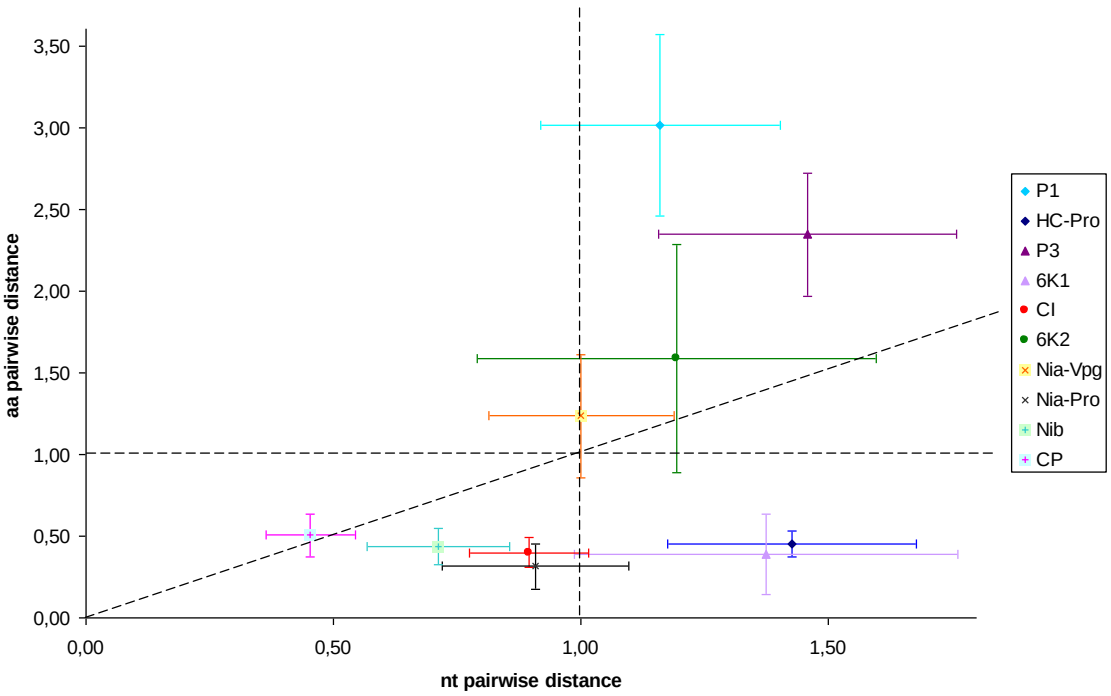


Figure 41. Hubert M-estimator of pairwise distance at the nucleotide and amino acid level for each of the TuMV gene normalized by their respective concatenated genes sequence.

In our previous work, the variability and the potential for evolution of each sequence determined the durability of amiR-mediated resistance. To explore difference in evolution of the TuMV genes we calculated, for each gene tree, the pairwise distances between all possible paired terminal taxa, both from nucleotide and amino acid sequences, normalized to the entire estimates obtained for the concatenated genome. The values obtained are represented in a graph with the mean pairwise in nucleotide distance as x-axis and the amino acids distance as y-axis for each gene (Fig. 41). This representation gives us clues on the different rate of evolution for each part of the TuMV genome. Indeed we can confirm that the P1 gene is the most variable one (Ohshima et al., 2002), while that the most conserved gene at the nucleotide and amino acid levels is the CP. This result confirms a possible region in the TuMV genome with strong evolutionary constrains on its sequence. By contrast the HC-Pro gene, where the amiR159-HCPro target sequence is contained, shows a high nucleotide divergence between isolates though most of mutations are synonymous.

To go deeper in the analysis of natural variability, we observed the abundance of mutations found among all natural isolates for the 21-nt sequence targeted by the amiR159-HCPro. Indeed the distribution of mutations over the 100 field isolates along the amiR159-HCPro target sequence shows natural variability for this region (Fig. 42). Over the 21 nt only nine are strictly conserved. Variable positions are distributed along the sequence and can reach more than 60% of isolates that differ from our sequence. Among these variable positions, one is central (position 11) and is a critical position for the amiR to be sliced by Dicer is variable. Actually, if only one mutation is enough to break the resistance, as shown in our previous study, only three isolates would not be able to infect the amiR159HC-Pro transgenic plants.

Adding to this, result from the ultra-deep sequencing analysis of population dynamics of virus escape mutants in RNAi-mediated resistant plants confirm the presence of this variability along all the

experiment (Martinez et al., 2012). Interestingly, in this case all positions of the amiR159-HCPro target showed variability in evolution in wild-type plants. Under the hypothesis of neutral accumulation of mutations, we should expect all sites to show approximately the same variability. When the virus evolve in 10-4 plants some position revealed to be more variable as site 11 and 12. This result confirms that for breakdown amiR, some positions are more efficient as shown in a neutral gene (Lin et al., 2009).

Taking advantage of knowledge of the natural variability is crucial to define new amiR against non-flexible region of the virus.

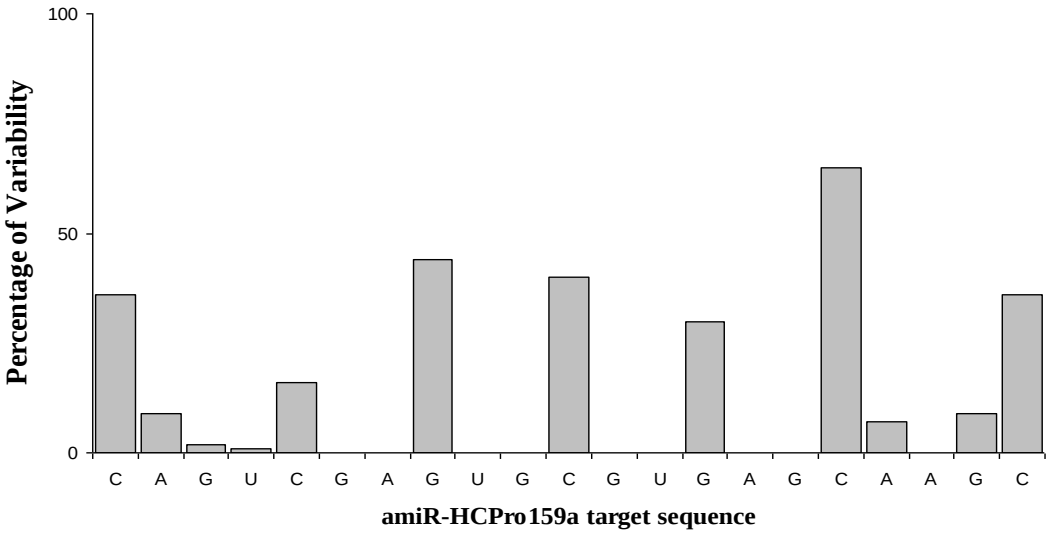


Figure 42. Natural nucleotide variability observed among the 100 natural TuMV isolates along the amiR-HCPro159 target sequence.

2.2. **Objectives**

In the experimental evolution of the TuMV all lineages were able to breakdown the amiR resistance. An analysis of the target revealed always at least one mutation over the 21-nt sequence. Phylogenetic analyses have shown that each gene differ in rate of changes at the nucleotide and amino acid levels. The least variable sequence of the TuMV is the CP, and looking for strictly conserved region within it could enhance the durability of amiR.

A better understanding of factors governing resistance-breaking of amiR such as critical positions highlighted in the two precedent studies has allowed us in developing a second generation of amiRs as overlapping amiRs.

After generation of new transgenic plants, we want to know if they have a better capacity to be resistant when challenged with TuMV variants.

Ultra-deep sequencing of the different lines generated by Lafforgue et al. (2011) shows that viral evolution in fully susceptible plants results from an equilibrium between mutation and genetic drift (Martinez et al., 2012). Diversity in the target sequence was generated and maintained along evolutionary time and the escape variant was present in the evolving population from the very beginning of the evolution experiment (Martinez et al., 2012).

As Martínez et al. (2012) demonstrated since the first passage all diversity encountered along the experiment was generated. So instead to challenge 20 plants at each passage we decided to challenge a large bunch of plants. With virus from only one passage in 12-4 or 10-4 *A. thaliana* plants, we challenge new generation of amiR and test if virus escape mutants are still able to break resistance on new transgenic plants.

2.3. **Design of a new generation of amiRNAs**

In the first experiment we encountered amiR characteristics which could conduct to a loss of the resistance. The first one is the variability of the target region. The phylogenetic and diversity analysis performed with the 100 natural isolates of TuMV gave us precious information on natural variability and conserved regions. One advantage of generating amiRs is that it is only necessary to cover only 21 nucleotides, which, in principle should be easier to find a strictly conserved nucleotide region. Despite this apparent relaxed constraint, only few regions (Table 10) were putative candidates given their strict conservation among the 100 TuMV isolates. To further enhance the protection and durability of the resistance, we inserted two amiRs rather than only one, as done in our first study. Indeed, Niu et al. (2006) demonstrated that a dimeric amiR precursor in *A. thaliana* could be effective against two viruses. To take advantage of the precedent 12-4 fully resistant plant, we inserted a second amiR targeting the CP region on the 12-4 genetic background (hereafter named as TuHC-CP plants). We also observed that some position are more critical for the breakdown of the resistance as others so in addition, we also generated a second double-amiR transgenic line with an overlap double amiR targeting the CP region (TuCPa-b plants). The sequence used to generate the transgenic plants begins at the position 9323 (Fig. 29). By combining multiple amiR into a single plant, the likelihood of resistance breaking should drop exponentially.

Table 10. Exhaustive list of conserved nucleotide sequences of more than 10 nt among the 100 natural TuMV isolates.

Domain	Conserved sequence among natural TuMv isolates	Length (>10 nt)
HC-Pro	1833-UUCAUAUGGG-1843	11
P3	3207-GAUUUAGGCGGCA	13
CP	9204-AACGGAAUGUGGGU-9217	14
	9237-CAGGUGGAAUUCCC-9250	14
	9318-GCUGAAGCGUACAUUGAAAAGCGUAACCA-8346	29
	9366-CGAUAUGGUCUUAACCA-8346	17
	9468-CACAUCCAGAUGAAAGC-9484	17
	9510-AAUUUGUUCGG-9520	11
	9540-ACAACGGUAGAGAACACGGA-955995	20
	9582-AAUCGGAACAUGCA-9595	14
3’-UTR	9665-UAGUAUUCUCGC-9676	12
	9678-UAUGGGAAAUA-9688	11
	9765-UAUUGGUGUUA-9775	11
	9802-UUGCCUUAACAUUU-9815	14

2.4. **Results**

Plants from the genotypes wild-type, fully resistant 12-4 and the two new double-amiR transgenic lines (TuHC-CP and TuCPa-b) were inoculated with saps previously obtained from infected *A. thaliana* wild-type or 10-4 (partially resistant) as previously described. 15 dpi, plants were recorded for their symptoms. Table 11 presents the infectivity data on each genotype and source of inoculum.

Table 11. Number of infected plants of each genotype (wild-type, 12-4, TuHC-CP and TuCPa-b) inoculated with saps obtained from wild-type or 10-4 sap previously infected with TuMV.

Nature of *A.thaliana* plants	Wild-Type sap		10-4 partially resistant sap	
	Infected	Inoculated	Infected	Inoculated
Wild-type	166	301	155	311
12-4	9	522	26	537
TuHC-CP	0	539	0	475
TuCPa-b	0	535	0	533

A total of 2082 double transgenic plants were inoculated for this experiment. As negative control 933 *A. thaliana* Col-0 wild-type plants were inoculated. No transgenic plant showed symptoms, whereas 52.4% of the wild-type plants were infected. The response of transgenic and wild-type is significantly different (Fisher test: $P < 0.0001$). Using the Laplace estimator for the p parameter of the Binomial distribution, we found that the best estimation for the probability of breaking the resistance in the double transgenic plants cannot be larger than 0.098%.

As expected, there are no differences between saps from wild-type and 10-4 plants when inoculated in wild-type plants (Mann-Whitney test, $P = 0.394$) but a significant difference exist when inoculated in fully resistant plants (Mann-Whitney test, $P = 0.003$), with virus isolated from 10-4 plants breaking more frequently the 12-4 resistance, as shown before (Lafforgue et al., 2011). Furthermore, the

probability of infecting 12-4 plants using sap made out from wild-type plants is only about 3%, whereas it rises up to 10% if the sap is produced from 10-4 plants. In this case the difference between simple and double transgenic plants is also significant (Fisher test: $P < 0.0001$), indicating a more durable resistance in the case of double transgenic plants.

With our sample size we failed to detect any significant difference between the two strategies of generating double-amiR transgenic plants, namely having each one targeting different TuMV genomic regions or both overlapping over the same genomic sequence.

TuMV mutation rate *in vivo* of the target sequence of the amiR159-HC-Pro was determined at 6×10^{-5} per replication event (de la Iglesia et al., 2012). An escape mutant able to break the double amiR159-TuHC-CP protection should simultaneous present at least two mutations, one for each amiR target. Using the mutation rate estimated by de la Iglesia et al. (2012), the likelihood of having the at least two mutations necessary for breaking the resistance is $(6\times10^{-5})^2 = 36\times10^{-11}$, assuming that the two regions had the same mutation rate. This very low number implies that to observe at least one event of resistance-breaking for a double-amiR resistant plant we should have to inoculate 2.78×10^{8} plants!

2.5. **Discussion.**

The target sequences chosen in the CP for the new generation of amiR (amiR159-TuHC-CP and amiR159a-CPa-b) correspond to non-variable regions over one hundred of isolates. In our case the CP region chosen have secondary structures which are necessary for replication of the virus (Haldeman-Cahill et al., 1998). It is so more difficult for the virus to mutate as a double function is given by this region. Choosing a conserved RNA motif has allowed us to significantly diminish the spontaneous emergence of escape mutants by drastically increase the number of mutational events needed to

breakdown the resistance. The antiviral strategy increased in efficiency and security with the use of more than one amiR.

Many studies reported resistance breakdown for qualitative gene-for-gene systems (Harrison, 2002; Kang et al., 2005; Lecoq et al., 2004). Whereas evidence to adaptation to quantitative polygenic resistance is more limited but it gives clues on directional selection and local adaptation (Andrivon et al., 2007) or maladaptation (Kaltz et al., 1999; Zhan et al., 2002). Indeed, we can consider two amiRs as a sort of polygenic resistance. A recent study on PVY shows that adaptation to a quantitative resistance is costly for the virus (Montarry et al., 2012). In this case the virus was able to breakdown the polygenic resistance. Unlike for pepper resistance to PVY, the amiR strategy does not allow the virus to multiply in the resistant host. If we consider multiple amiRs as a polygenic resistance this could implies a cost to the virus to breakdown the resistance and so escape variants in susceptible host will be under purifying selection.

In some cases we can couple two types of resistance, indeed some monogenic or recessive resistance can be overcome by known specific mutation (Ayme et al., 2006; Traore et al., 2010). Using amiR designed for these specific mutations could anticipate virus evolution and hijacks the opportunity to breaking down classical resistance.

In this study we brought new insights to better design the next generation of vaccine and protection against plant viruses. Choosing precisely the region in the virus genome to be targeted and knowing its potential to tolerate changes and combining multiple amiRs can lead to a better and more durable protection.

III.2 EXPERIMENTAL EVOLUTION OF A VIRAL GENOME CARRYING TWO RNA SILENCING SUPPRESSORS

Gene duplication is the main source of genetic redundancy which generates multiplicity in gene copy number. A second and less frequent source of functional redundancy are convergent evolutionary processes leading to genes that are close in function but unrelated in sequence (Galperin et al., 1998). Viral genomes are compact and the entire sequence encodes only for functions that are absolutely indispensable for completing viral life cycle. Viral proteins can also have several functions highlighting the need to have the maximum information in the shortest possible genome. The effect of adding a new protein, with a function which is already encoded in the viral genome, could modify the viral proprieties of virulence and pathogenicity. In particular, in the context of the arm races between plants and viruses, which impact may have adding a second VSR to a viral genome on plant defense? Indeed one particular virus, *Citrus tristeza virus* (CTV), presents three different VSR. This is a good example illustrating that it is possible for a viral genome to have functional redundancy.

1. Objectives

Functional redundancy is a term used to describe situations where a given biochemical function is redundantly encoded by two or more genes. In this work, we addressed the question of whether a viral genome carrying functional redundancy encoded by different proteins will evolve to specialize each protein into different function or, alternatively, one of the proteins will be removed. To do so we will explore two genetic contexts by adding a new protein with a redundant function in the viral genome. Firstly, in an otherwise fully functional viral genome we could expect two phenomena upon adding an exogenous gene that provides functional redundancy: either

elimination of the new gene added or, alternatively, subfunctionalization and specialization of both the original gene and the new genes. Secondly, if the original gene shows a subfunctional activity, adding a new function can complement the deficiency and potentially, facilitating its retention and leading to subfunctionalization and specialization. Given its importance in the life cycle of TuMV, and the multifunctional role of HC-Pro, we have chosen to duplicate the VSR function. Two alternative VSRs were added: CMV 2b or TBSV p19. While HC-Pro and p19 act at similar levels in the suppression of RNA silencing pathway, 2b inhibits the amplification and spreading of the signal (Guo and Ding, 2002).

2. Experimental design

TuMV HC-Pro is a multifunctional protein involved in aphid transmission, cell-to-cell movement, polyprotein processing, and in RNA silencing suppression (Brigneti et al., 1998). Mutations in the conserved motif FRNK (e.g., to FKRK) of HC-Pro abolish the function as VSR without affecting the other functions. This mutated TuMV, labeled as TuMV-K presents severely attenuated symptoms and infection dynamic in *Chenopodium quinoa* (Fig. 43) and in *N. tabacum* (Fig. 44). Indeed in *C. quinoa* wild-type TuMV is present in the entire lesion except in the necrotic part, whereas TuMV-K is detected only at the periphery of the lesion and did not provoke HR. In *N. tabacum*, wild-type TuMV expands in the entire leaf. The TuMV-K-GFP has to face to the host defense and have more difficulties to expand itself.

We have inserted in TuMV wild-type or in TuMV-K genomes a second RNA silencing suppressor (either CMV 2b or TBSV p19) or the GFP between the NIb and CP cistrons, a position where potyviruses easily admit insertion of additional cistrons (Lin et al., 2009).

To characterize the patterns of molecular evolution of HC-Pro and of the added genes, the six different genomes were used to initiate

10 independent evolutionary lineages in *N. benthamiana*. Passages were done at two different regimes: every 14 days or every 28 days for 10 or 5 passages respectively (for five months) (Fig. 32).

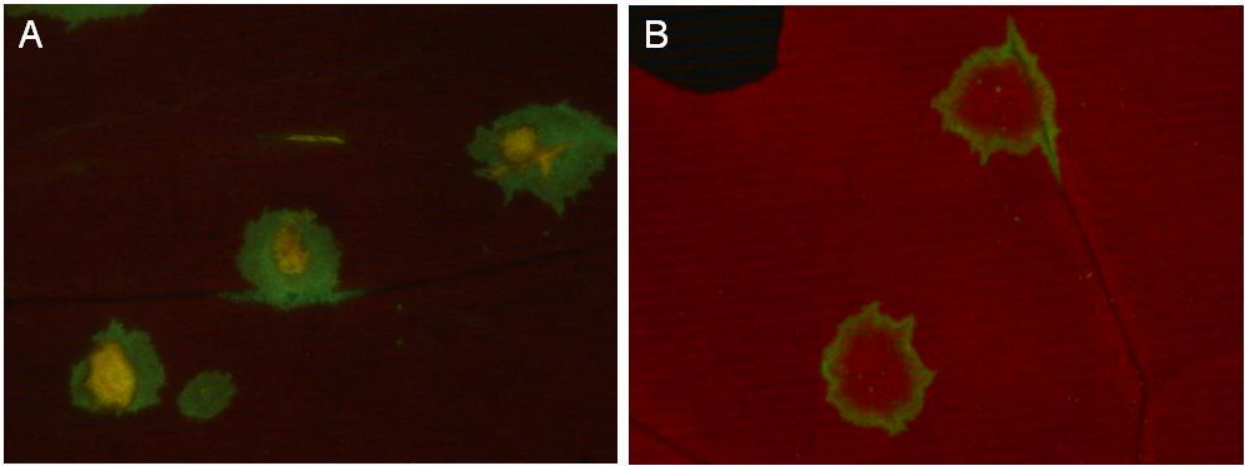


Figure 43. Shape of the local lesions produced in *C. quinoa* by wild-type TuMV-GFP **(A)** and TuMV-K-GFP **(B)**.

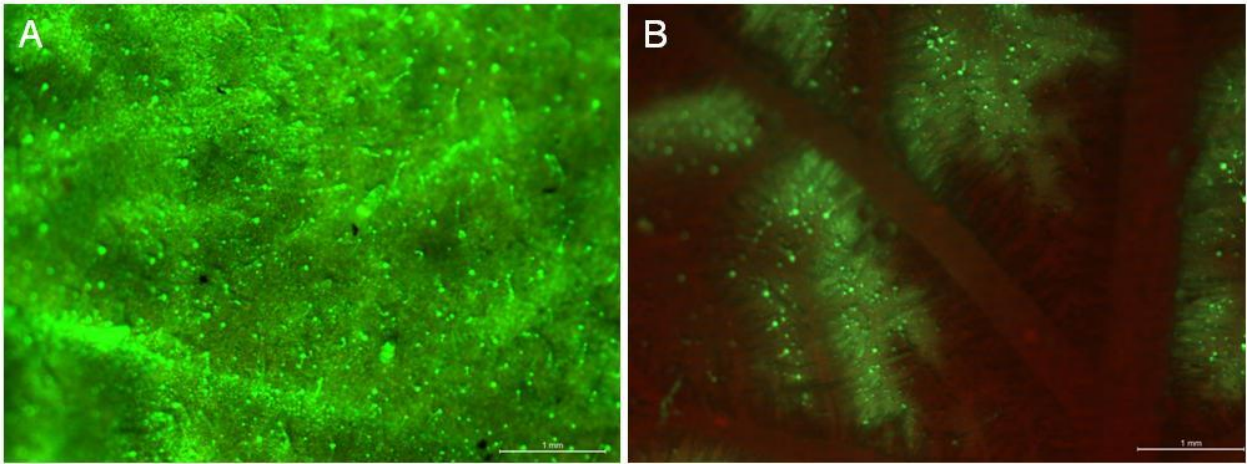


Figure 44. Aspect of *N. benthamiana* leaves infected by wild-type TuMV-GFP **(A)** and TuMV-K-GFP **(B)**.

3. Results

3.1. Characterization of modified genomes

At the phenotypic level at 28 dpi, *N. benthamiana* plants infected with TuMV-K-GFP show a reduction of weight (55%) and size (23%). TuMV-K phenotype is partially restored with the insertion of 2b and closer to the wild-type with p19. However, the constructions carrying 2b or p19 proteins are more virulent (reduction of host size and weight) than the wild-type (Fig 45). Indeed a principal component analysis (PCA) shows that 93% of observed variance depends on the inoculated genome.

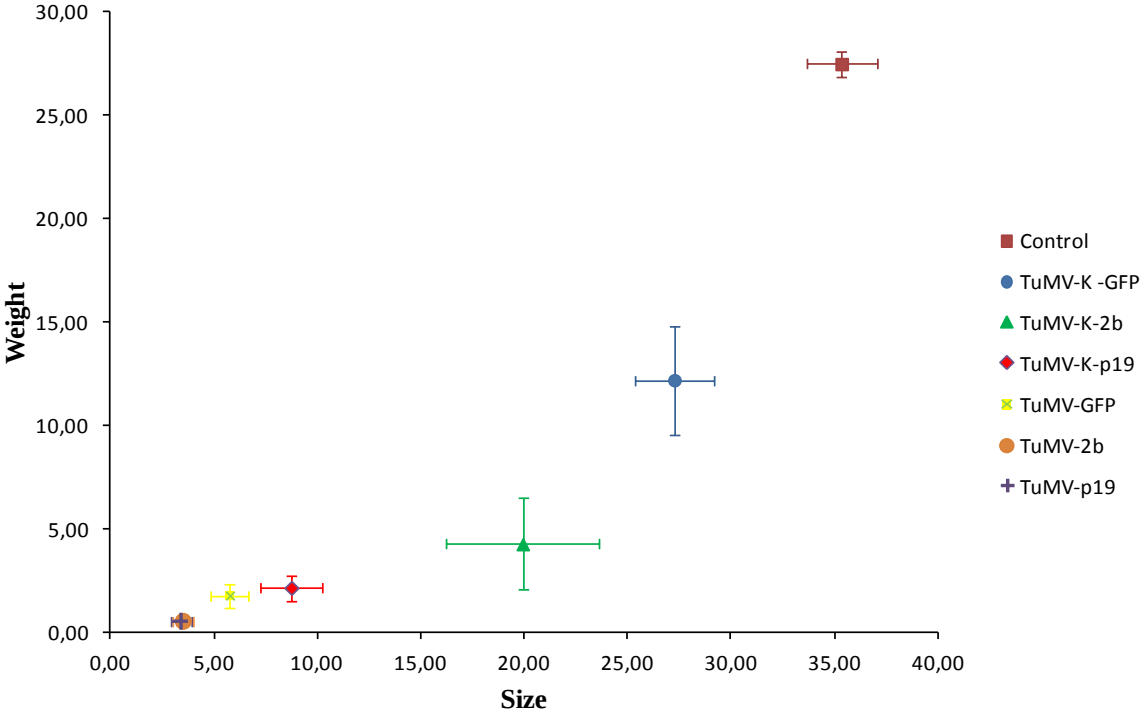


Figure 45. Weight and size at 28 dpi of *N. benthamiana* plants infected with each of the six different TuMV genomes.

HC-Pro suppressor activity is related to symptomatology inhibiting both endogenous miRNA- and siRNA-mediated gene regulations, which resulted in the overexpression of miRNA targeted genes in transgenic plants expressing HC-Pro. Thus, transgenic plants expressing HC-Pro result in a phenotype resembling that characteristic of TuMV-infected plants (Dunoyer et al., 2004; Kasschau et al., 2003). TuMV-K infection results in asymptomatic leaves on *N. benthamiana*. Similarly, though TuMV-K-2b induces symptoms, these are weaker than those produced by TuMV-K-p19, which are closer to those produced by wild-type TuMV. Finally, although the symptoms produced by wild-type TuMV and TuMV-2b are similar; symptoms produced by TuMV-p19 are much stronger, including systemic necrosis on 90% of the plants at 28 dpi (Fig. 46).

3.2. **Infectivity efficiency of modified genomes**

All viral constructions survived when passages were done every 14 dpi. However, some were lost when passages were carried out every 28 dpi; in the case of TuMV-K-p19 20% of the initiated lines were lost, 30% in the case of TuMV-K-2b, and 80% for the TuMV-K lines. This result underlines the deleterious effect of the mutation FRNK to FKNK in the conserved site of HC-Pro suppressor activity, which makes the corresponding TuMV genomes vulnerable to the host defense silencing responses.

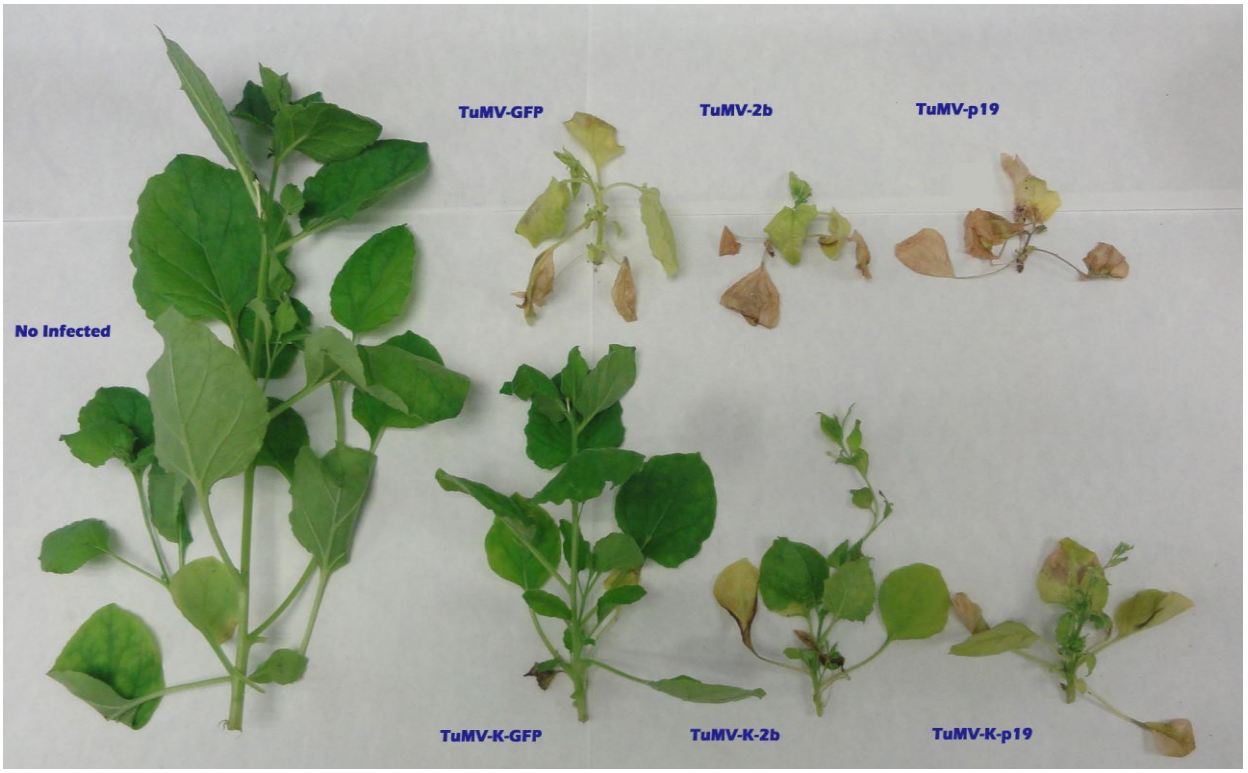


Figure 46. Picture at 28 dpi of healthy and infected *N. benthamiana* plants. On the left a non-infected *N. benthamiana*. The three upper plants were infected with TuMV containing the wild-type HC-Pro with GFP, 2b or p19 proteins whereas the three plants on the bottom contain the mutated HC-Pro with GFP, 2b and p19.

3.3. **Evolution of similar sequences evolution in different genomic contexts**

After experimental evolution in *N. benthamiana*, evolved viruses were sequenced for the HC-Pro and the second suppressor cistrons. All mutations encountered are listed in Table 12. We do not detect any nucleotide changes in the FRNK region. Point mutations appeared along the HC-Pro sequence in 38.5% of lineages evolved at the 14 dpi passage regime and in 24.4% of the lineages passaged at the 28 dpi regime. For the GFP sequence, point mutations were observed in 50% of the lineages passaged every 14 dpi. In addition, one line contained a 501 nt deletion in the GFP sequence. Interestingly, this deletion did not affect the downstream reading frame, thus producing a functional CP. For the 12 lineages evolved by passages every 28 dpi, 41.6% HC-Pro-GFP lineages and the only two HC-Pro-K-GFP lineages that reached the end of the experiment contained point mutations. Regarding the genomes carrying the 2b VSR, 23% of the lineages passaged every 14 dpi had point mutations, while we did not detect any mutations in those lineages passaged every 28 dpi. Finally, for lineages carrying the additional p19 VSR, 23.5% of the lineages passaged every 14 dpi carried point mutations, whereas the frequency of mutated genomes dropped to 11% when lineages were passaged every 28 dpi.

Table 12. Mutations found both in the HC-Pro or in the additional genes in lineages evolved by passages every 14 dpi (**A**) or every 28 dpi (**B**). The number following the name of the genome analyzed is the number of lineages that have survived until the end of the experiment.

A

HC-Pro	Position	Type of mutation	Number mutation
TuMV-GFP (9)	1917	C --> A (s)	1
	2178	C --> T (s)	1
TuMV-K-GFP (10)	1663	G --> A (E --> K)	1
	1845	T --> C (s)	1
	1890	G --> A (s)	1
	2023	C --> T (P--> S)	1
TuMV-2b (10)	1367	G --> A (G --> E)	1
	1392	A --> G (s)	1
	1459	C --> T (R --> C)	1
	1545	A --> T (N --> I)	1
	1560	T --> C (s)	1
	1702	G --> A (A --> T)	1
TuMV-K-2b (3)	1255	T --> A (C-->S)	1
TuMV-p19 (10)	1367	G --> A (G-->E)	1
TuMV-K-p19 (7)	2178	C --> T (s)	1

Foreign gene	Position	Type of mutation	Number mutation
TuMV-GFP (9)	9020	C --> T (S--> F)	2
	8768 - 9269	Deletion	1
TuMV-K-GFP (10)	8595	A --> G (s)	1
	8782	C --> T (s)	3
	9255	C --> T (s)	1
	9343	C --> T (s)	1
	9423	G --> A (s)	1
TuMV-2b (10)	8600	G --> A (S --> N)	1
	8625	T --> C (s)	1
TuMV-K-2b (3)	9226	A --> G (K -->R)	1
TuMV-p19 (10)	9148	C --> T (s)	1
	9364	A --> G (R --> G)	1
	9370	A --> G (K --> E)	1
TuMV-K-p19 (7)	9071	G --> T (G --> V)	1

B

HC-Pro	Position	Type of mutation	Number mutation
TuMV-GFP (10)	1245	C --> T (s)	1
	1353	T --> C (s)	1
TuMV-K-GFP (2)	/	/	/
TuMV-2b (8)	1893	A --> G (I --> M)	1
	2050	G --> A (E-->K)	1
	2178	C --> T (s)	1
TuMV-K-2b (7)	1255	T --> A (C --> S)	1
TuMV-p19 (10)	1980	C --> T (s)	2
	2178	C --> T (s)	1
TuMV-K-p19 (8)	1770	A --> C (s)	1
	2014	A --> G (K--> E)	1

Foreign gene	Position	Type of mutation	Number mutation
TuMV-GFP (10)	8737	G --> A (E-->K)	1
	8922	G --> A (s)	1
	9168	C --> T (s)	1
	9266	G --> A (R--> H)	2
TuMV-K-GFP (2)	/	/	/
TuMV-2b (8)	/	/	/
TuMV-K-2b (7)	/	/	/
TuMV-p19 (10)	8935	T --> C (s)	1
TuMV-K-p19 (8)	8911	G --> A (s)	1

Due to the few mutations encountered at each gene, we detected in most cases only one type (synonymous or non-symonymous) of mutation (Table 13). Among the fifty mutations encountered along the whole experiment, 54% were synonymous. If we compare mutations between HC-Pro and the inserted, this rate is the same.

Table 13. Ratio of synonymous/non-synonymous mutations encountered in lineages evolved by passages every 14 dpi (**A**) or every 28 dpi (**B**).

A	**HC-Pro**	**Foreign Gene**	B	**HC-Pro**	**Foreign Gene**
TuMV-GFP	1	0	TuMV-GFP	1	0,5
TuMV-K-GFP	0,5	1	TuMV-K-GFP	/	/
TuMV-2b	0,33	0,5	TuMV-2b	0,33	/
TuMV-K-2b	0	0	TuMV-K-2b	0	/
TuMV-p19	0	0,33	TuMV-p19	1	1
TuMV-K-p19	1	0	TuMV-K-p19	0,5	1

We have calculated the mutation frequency per site as $\mu = N_\mu/(L \times N_s)$, where N_μ is the number of mutations found, L the length of the fragment sequenced and N_s the number of sequences analyzed (Table 14). All sequences have fairly equal mutation rate around $2.93 \times 10^{-4} \pm 1 \times 10^{-5}$[SEM] per site.

Table 14. Mutation frequencies for lines evolved by means of serial passages each 14 dpi (**A**) or every 28 dpi (**B**).

A	**HC-Pro**	**Foreign Gene**	B	**HC-Pro**	**Foreign Gene**
TuMV-GFP	$2{,}04 \times 10^{-4}$	$4{,}22 \times 10^{-4}$	TuMV-GFP	$1{,}82 \times 10^{-4}$	$6{,}33 \times 10^{-4}$
TuMV-K-GFP	$3{,}64 \times 10^{-4}$	$8{,}86 \times 10^{-4}$	TuMV-K-GFP	0	0
TuMV-2b	$5{,}45 \times 10^{-4}$	5×10^{-4}	TuMV-2b	$3{,}41 \times 10^{-4}$	0
TuMV-K-2b	$3{,}03 \times 10^{-4}$	$6{,}25 \times 10^{-4}$	TuMV-K-2b	$1{,}3 \times 10^{-4}$	0
TuMV-p19	$9{,}09 \times 10^{-5}$	$5{,}12 \times 10^{-4}$	TuMV-p19	$2{,}73 \times 10^{-4}$	$1{,}71 \times 10^{-4}$
TuMV-K-p19	$1{,}3 \times 10^{-4}$	$2{,}84 \times 10^{-4}$	TuMV-K-p19	$2{,}27 \times 10^{-4}$	$2{,}13 \times 10^{-4}$

The effect of gene (main factor) and passage type (covariable), on mutation rate were assessed by an ANOVA. There was no difference between HC-Pro and the foreign gene ($F_{3,1}$ = 1.637, P = 0.214). The only effect detected is between the evolution mode (14 or 28 dpi serial passages) ($F_{1,3}$ = 7.716, P = 0.012). This can be explained by the loss of many sequences during the 28 dpi serial passages (most of the TuM-K-GFP).

4. Discussion

Functional redundancy could be an opportunity to evolve for a given genome. Indeed, as genomic redundancy, the virus can exploit functional redundancy to modify its genome and explore new functions without losing the old ones. Results of our evolution experiments indicate few modifications in the genome. This sharply contrasts with theory predictions of faster evolution for redundant genes. One explanation can be a low mutation or multiplication rate diminishing the possibilities for evolution. It was observed that populations of TuMV from Europe and Australia, probably separated more than 12000 years ago, differed by less than 1% (Blok et al., 1987), indicating strong constraints to modify TuMV genome. Despite the reduced number of observed mutational changes, some have a strong phenotypic effect. During the experiment, we have observed that development of infected plants is strongly affected and finally

stopped; few new leaves are produced and they show a minimal expansion. In this condition the virus quickly lacks new territories to expand, multiply and, consequently, to evolve. Even in an extreme condition, TuMV-p19 provokes necrosis and kills the plants. In the case of the HC-Pro-K, where the TuMV has to face the plant defense system in a situation of disadvantage, the virus replicates at a low level, thus having fewer possibilities to evolve.

Functional redundancy means that two or more genes are performing the same function and that inactivation of one of these genes has little or no effect on the biological phenotype. In the TuMV genomes engineered for this study, we have two unrelated sequences that encode for a similar function: the HC-Pro and a second suppressor 2b or p19. If the new gene was truly redundant, then it would not be protected against the accumulation of deleterious mutations. So if adding the 2b or the p19 protein gives a particular advantage such as increasing the suppressor function or related to any other of the functions performed by these multifunctional proteins (movement, replication, enhancing translation…) (Ding et al., 1995; Qi et al., 2004), we are not any more in the strict functional redundancy. In this case evolution of the viral genome depends on the positive selection and not only strictly on genetic drift.

A functional redundancy situation appears naturally in CTV, that encodes three VRS in its genome, proteins p20, p23 and CP (Lu et al., 2004). It was suggested that suppression of RNA silencing at multiple steps of the silencing pathway may be essential for viruses with large RNA genomes such as CTV to complete their infectious cycle successfully (Lu et al., 2004). As we know the 2b protein acts at a late stage of the RNA silencing machinery (Guo et al., 2005), during amplification and spread of the signal across the host, and that HC-Pro or p19 act at different early steps of the antiviral pathway (Brigneti et al., 1998; Lakatos et al., 2004). Adding a second VSR in the TuMV genome maybe is not a strict functional redundancy for the virus. Indeed similarly to CTV, the additional VSR enhance the function by

acting in a different specific step of the RNA silencing, the modified TuMV is fitter than the wild-type.

Finally, it has been shown that mixed infection of CMV and TuMV induces more severe symptoms on *N. benthamiana* than single infection (Takeshita et al., 2012). This enhancement of virulence is due to the effect of the VSR of the two viruses showing that synergism can act in separate genomes or in a single one as in our modified TuMV.

Taken together, these results suggest that the second suppressor provides some sort of fitness advantage, as an enhancement of the suppressor activity or maybe other functions performed by the exogenous multifunctional proteins added.

IV. CONCLUSIONS

Plants have many different mechanisms to counter viral infections. Natural resistances are mostly specific to a plant/virus pathosystem. Breeders need to get durable solution against the pathogen that has a major threat on their crop. Unfortunately, nature did not provide resistance to all pathogen, and when this is the case, virus can develop strategies to avoid some resistance. The aim of this study was to explore the possibilities of a new biotechnical approach based on the transgenic expression of amiR. Aware of the possible limitations of the technique, as breakdown of resistance, we engaged on an experimental evolution to predict the chances of resistance breaking in field conditions as well as to propose new specific rules to design more durable amiRs.

Our first experiment, revealed important critical points in the breakdown process. First, the choice of the amiR position along the viral sequence and the amount of natural variation existing for that sequence has a major impact on the durability of resistance. TuMV has a natural variability in the region chosen by Qiu et al. (2007) to design the amiR. Although transgenic plants were totally resistant to the TuMV wild-type, only one mutation in the 21 nt target sequence was enough to break resistance, with higher efficiency for particular positions within the target sequence. We also demonstrated that the context of selection was particularly important in the delay of breakdown of resistance. Likewise Lin et al. (2009), in this study we have also shown the emergence of escape alleles containing mutations in the target that are likely to affect the binding of the amiR, thus reducing the slicing efficiency by RISC. In contrast to the research by Lin et al. (2009), ours correspond to a viral coding region and not a neutral heterologous sequence.

To explore deeper the evolution process of TuMV in the *A. thaliana*, the population diversity of each passage of several lineages was determined by next-generation sequencing (NGS) (Martinez et

al., 2011). In particular in a long lineage, it appears that 21 potential escape alleles were present in the evolving populations from the very beginning of the evolution experiment. These escape mutations represented 0.02% of haplotypes across passages. This indicates that breaking the resistance in our case is not correlated with a steady increase of escape mutant frequency or accumulation of different escape alleles until some critical value was reached. It appears that a successful infection in 12-4 plants is stochastic and depends of random picks in the population. Another observation was that the TuMV ancestral sequence was still present in 12-4 plants at a frequency of 0.69%, ranking the 25^{th} in this population. Even if the ancestral TuMV is not the most abundant variant in the 12-4 plants, Martínez et al. (2011) explain its presence by three non-mutually exclusive hypotheses. First, if the efficiency of the RNA silencing machinery in detecting allelic variation in the target depends on a threshold concentration, the TuMV genome carrying the ancestral amiR159-HCPro target can survive only when it represents a minor fraction of the population. Second, the ancestral TuMV can take advantage of expression of the HC-Pro from escape mutant interfering with the RNA silencing machinery and blocks its action. In this case the ancestral TuMV cannot expand itself without the escape mutant and replicate only in already infected cells. Third, new TuMV ancestral genome is reconstitute by back mutation from escapes mutant population. This third possibility is less plausible than the others, indeed using a recent estimate of TuMV mutation rate in *A. thaliana* (de la Iglesia et al., 2012) of ~ 6×10^{-5} per replication event and the fact that TuMV replication mostly proceeds by a stamping machine (Martinez et al., 2011), Martínez et al (2012) expected population frequency of reversion mutants produced by backward mutation during replication of the numerically dominant resistant haplotype should be ~0.006%, a value about 115-fold smaller than the observed frequency. To conclude the observed frequency of ancestral TuMV apparently do not comes from only backward mutations and more due to inefficient RNA silencing machinery and/or of

complementation with escape mutants during cell coinfection. A last remark on the detection of the ancestral TuMV in the 12-4 plants is to remember that in general 99.8% of the population inoculated carries the ancestral amiR159-HCPro target. Fate of the ancestral TuMV in resistant plants can be a complete elimination even with rescue mechanism as complementation or backward mutation. To confirm the selection pressure in 10-4 plants, TuMV populations showed a quantitatively important difference from populations evolved in susceptible plants. Indeed the frequency of escape mutant was more than one order of magnitude higher. These results are compatible with a mutation–selection–drift model of evolution. The rates of synonymous and nonsynonymous substitutions were equivalent and breakdown of the resistance is a consequence in nucleotide changes and not in amino acid composition of the HC-Pro.

In a follow up study, new amiRs were generated, including two amiRs designed to be complementary to a highly conserved region within the CP of ca. 100 natural TuMV isolates. Results indicates that if amiRs are well designed and target different parts of the virus, the probability of breaking the resistance is exponentially reduced and can be close to zero, increasing the durability of resistance. If this method gives an efficient resistance to virus, plants have to produce the amiR at a sufficient level to block directly the virus life cycle. If not, the amount of amiR will act as a selective force as in the 10-4 plants and favors the appearance and rise in frequency of escape mutants even if the amiR is designed for a high conserved region. It is hence, important to have a constant and homogeneous production of amiR along the plant's life and in all organs.

In the last part of this PhD memory, we have analyzed the effect of adding different additional VSRs to TuMV genome and its impact on genomic evolution. We have seen differences between genetic redundancy, which allow modification or deterioration of one gene, and functional redundancy, which can act at different level and have accumulative effects. As it has been well established, many viral proteins have multiple functions, and thus adding a second VSR to

TuMV can also imply increasing the number of viral functions. For example the fact that the 2b protein is also associated to movement may allow a better cell-to-cell and systemic spread of TuMV. In this study we also showed the effect of adding p19 and 2b proteins on TuMV virulence, which may be related to their VSR function. CTV with its three suppressor of RNA silencing shows a possible way for TuMV to acquire other VSR. This incorporation in TuMV sequence could occur in mixed infection with other virus coupled to recombination.

REFERENCES

Abel, P.P., R.S. Nelson, B. De, N. Hoffmann, S.G. Rogers, R.T. Fraley, and R.N. Beachy. 1986. Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science*. 232:738-743.

Adams, M.J., J.F. Antoniw, and C.M. Fauquet. 2005. Molecular criteria for genus and species discrimination within the family Potyviridae. *Arch Virol*. 150:459-479.

Ai, T., L. Zhang, Z. Gao, C.X. Zhu, and X. Guo. 2011. Highly efficient virus resistance mediated by artificial microRNAs that target the suppressor of PVX and PVY in plants. *Plant Biol (Stuttg)*. 13:304-316.

Alamillo, J.M., P. Saenz, and J.A. Garcia. 2006. Salicylic acid-mediated and RNA-silencing defense mechanisms cooperate in the restriction of systemic spread of plum pox virus in tobacco. *Plant J*. 48:217-227.

Ali, A., H. Li, W.L. Schneider, D.J. Sherman, S. Gray, D. Smith, and M.J. Roossinck. 2006. Analysis of genetic bottlenecks during horizontal transmission of Cucumber mosaic virus. *J Virol*. 80:8345-8350.

Andrivon, D., F. Pilet, J. Montarry, M. Hafidi, R. Corbiere, H. Achbani el, R. Pelle, and D. Ellisseche. 2007. Adaptation of Phytophthora infestans to Partial Resistance in Potato: Evidence from French and Moroccan Populations. *Phytopathology*. 97:338-343.

Arabidopsis, G., Initiative. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*. 408:796-815.

Aukerman, M.J., and H. Sakai. 2003. Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell*. 15:2730-2741.

Ayme, V., S. Souche, C. Caranta, M. Jacquemond, J. Chadoeuf, A. Palloix, and B. Moury. 2006. Different mutations in the genome-linked protein VPg of potato virus Y confer virulence on the pvr2(3) resistance in pepper. *Mol Plant Microbe Interact*. 19:557-563.

Barr, J.N., and R. Fearns. 2010. How RNA viruses maintain their genome integrity. *J Gen Virol*. 91:1373-1387.

Belsham, G.J., and P. Normann. 2008. Dynamics of picornavirus RNA replication within infected cells. *J Gen Virol*. 89:485-493.

Bendahmane, A., B.A. Kohn, C. Dedi, and D.C. Baulcombe. 1995. The coat protein of potato virus X is a strain-specific elicitor of Rx1-mediated virus resistance in potato. *Plant J*. 8:933-941.

Bendahmane, A., M. Querci, K. Kanyuka, and D.C. Baulcombe. 2000. Agrobacterium transient expression system as a tool for the isolation of disease resistance genes: application to the Rx2 locus in potato. *Plant J*. 21:73-81.

Bennett, M.D., and I.J. Leitch. 2005. Nuclear DNA amounts in angiosperms: progress, problems and prospects. *Ann Bot*. 95:45-90.

Betancourt, M., A. Fereres, A. Fraile, and F. Garcia-Arenal. 2008. Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus. *J Virol*. 82:12416-12421.

Bishop, K.N., R.K. Holmes, A.M. Sheehy, and M.H. Malim. 2004. APOBEC-mediated editing of viral RNA. *Science*. 305:645.

Blok, J., A. Mackenzie, P. Guy, and A. Gibbs. 1987. Nucleotide sequence comparisons of turnip yellow mosaic virus isolates from Australia and Europe. *Arch Virol*. 97:283-295.

Boden, D., O. Pusch, F. Lee, L. Tucker, and B. Ramratnam. 2003. Human immunodeficiency virus type 1 escape from RNA interference. *J Virol*. 77:11531-11535.

Bortolamiol, D., M. Pazhouhandeh, K. Marrocco, P. Genschik, and V. Ziegler-Graff. 2007. The Polerovirus F box protein P0 targets ARGONAUTE1 to suppress RNA silencing. *Curr Biol*. 17:1615-1621.

Boyes, D.C., A.M. Zayed, R. Ascenzi, A.J. McCaskill, N.E. Hoffman, K.R. Davis, and J. Gorlach. 2001. Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell*. 13:1499-1510.

Bremermann, H.J., and H.R. Thieme. 1989. A competitive exclusion principle for pathogen virulence. *J Math Biol*. 27:179-190.

Brigneti, G., O. Voinnet, W.X. Li, L.H. Ji, S.W. Ding, and D.C. Baulcombe. 1998. Viral pathogenicity determinants are suppressors of transgene silencing in Nicotiana benthamiana. *EMBO J*. 17:6739-6746.

Brodersen, P., L. Sakvarelidze-Achard, M. Bruun-Rasmussen, P. Dunoyer, Y.Y. Yamamoto, L. Sieburth, and O. Voinnet. 2008. Widespread translational inhibition by plant miRNAs and siRNAs. *Science*. 320:1185-1190.

Brommonschenkel, S.H., A. Frary, and S.D. Tanksley. 2000. The broad-spectrum tospovirus resistance gene Sw-5 of tomato is a homolog of the root-knot nematode resistance gene Mi. *Mol Plant Microbe Interact*. 13:1130-1138.

Brosnan, C.A., and O. Voinnet. 2011. Cell-to-cell and long-distance siRNA movement in plants: mechanisms and biological implications. *Curr Opin Plant Biol*. 14:580-587.

Brown, D.W. 1997. Threat to Humans from Virus Infections of Non-human Primates. *Rev Med Virol*. 7:239-246.

Bucher, E., T. Sijen, P. De Haan, R. Goldbach, and M. Prins. 2003. Negative-strand tospoviruses and tenuiviruses carry a gene for a suppressor of gene silencing at analogous genomic positions. *J Virol*. 77:1329-1336.

Bull, J.J., M.R. Badgett, H.A. Wichman, J.P. Huelsenbeck, D.M. Hillis, A. Gulati, C. Ho, and I.J. Molineux. 1997. Exceptional convergent evolution in a virus. *Genetics*. 147:1497-1507.

Burch, C.L., S. Guyader, D. Samarov, and H. Shen. 2007. Experimental estimate of the abundance and effects of nearly neutral mutations in the RNA virus phi 6. *Genetics*. 176:467-476.

Burdon, J.J., and P.H. Thrall. 2008. Pathogen evolution across the agro-ecological interface/ implications for disease managment. *Evolutionary Appilcations*. 1:57-65.

Burgyan, J., and Z. Havelda. 2011. Viral suppressors of RNA silencing. *Trends Plant Sci*. 16:265-272.

Carrasco, P., F. de la Iglesia, and S.F. Elena. 2007. Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco Etch virus. *J Virol*. 81:12979-12984.

Carrington, J.C., P.E. Jensen, and M.C. Schaad. 1998. Genetic evidence for an essential role for potyvirus CI protein in cell-to-cell movement. *Plant J*. 14:393-400.

Carrington, J.C., K.D. Kasschau, and L.K. Johansen. 2001. Activation and suppression of RNA silencing by plant viruses. *Virology*. 281:1-5.

Castresana, J. 2007. Topological variation in single-gene phylogenetic trees. *Genome Biol*. 8:216.

Chao, L., C.U. Rang, and L.E. Wong. 2002. Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage phi6. *J Virol*. 76:3276-3281.

Chare, E.R., and E.C. Holmes. 2006. A phylogenetic survey of recombination frequency in plant RNA viruses. *Arch Virol*. 151:933-946.

Charlesworth, D., B. Charlesworth, and M.T. Morgan. 1995. The pattern of neutral molecular variation under the background selection model. *Genetics*. 141:1619-1632.

Charron, C., M. Nicolai, J.L. Gallois, C. Robaglia, B. Moury, A. Palloix, and C. Caranta. 2008. Natural variation and functional analyses provide

evidence for co-evolution between plant eIF4E and potyviral VPg. *Plant J*. 54:56-68.

Chen, C.C., C.H. Chao, C.C. Chen, S.D. Yeh, H.T. Tsai, and C.A. Chang. 2003. Identification of Turnip mosaic virus isolates causing yellow stripe and spot on calla lily. *Plant Dis.* 87:901-905.

Chen, Y.K., D. Lohuis, R. Goldbach, and M. Prins. 2004. High frequency induction of RNA-mediated resistance against Cucumber mosaic virus using inverted repeat constructs. *Mol Breeding*. 14:215-226.

Chisholm, S.T., S.K. Mahajan, S.A. Whitham, M.L. Yamamoto, and J.C. Carrington. 2000. Cloning of the Arabidopsis RTM1 gene, which controls restriction of long-distance movement of tobacco etch virus. *Proc Natl Acad Sci U S A*. 97:489-494.

Chung, B.Y., W.A. Miller, J.F. Atkins, and A.E. Firth. 2008. An overlapping essential gene in the Potyviridae. *Proc Natl Acad Sci U S A*. 105:5897-5902.

Coleman, J.R., D. Papamichail, S. Skiena, B. Futcher, E. Wimmer, and S. Mueller. 2008. Virus attenuation by genome-scale changes in codon pair bias. *Science*. 320:1784-1787.

Conticello, S.G., C.J. Thomas, S.K. Petersen-Mahrt, and M.S. Neuberger. 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol*. 22:367-377.

Cook, R.J. 2006. From the Academy: Colloquium perspective. Toward cropping systems that enhance productivity and sustainability. *Proc Natl Acad Sci U S A*. 103:18389-18394.

Cooley, M.B., S. Pathirana, H.J. Wu, P. Kachroo, and D.F. Klessig. 2000. Members of the Arabidopsis HRT/RPP8 family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell*. 12:663-676.

Cooper, B., M. Lapidot, J.A. Heick, J.A. Dodds, and R.N. Beachy. 1995. A defective movement protein of TMV in transgenic plants confers resistance to multiple viruses whereas the functional analog increases susceptibility. *Virology*. 206:307-313.

Cooper, I., and R.A. Jones. 2006. Wild plants and viruses: under-investigated ecosystems. *Adv Virus Res*. 67:1-47.

Cosson, P., V. Schurdi-Levraud, Q.H. Le, O. Sicard, M. Caballero, F. Roux, O. Le Gall, T. Candresse, and F. Revers. 2012. The RTM resistance to potyviruses in Arabidopsis thaliana: natural variation of the RTM genes and evidence for the implication of additional genes. *PLoS One*. 7:e39169.

Cosson, P., L. Sofer, Q.H. Le, V. Leger, V. Schurdi-Levraud, S.A. Whitham, M.L. Yamamoto, S. Gopalan, O. Le Gall, T. Candresse, J.C. Carrington, and F. Revers. 2010. RTM3, which controls long-distance movement of potyviruses, is a member of a new plant gene family encoding a meprin and TRAF homology domain-containing protein. *Plant Physiol*. 154:222-232.

Couce, A., and J. Blazquez. 2009. Side effects of antibiotics on genetic variability. *FEMS Microbiol Rev*. 33:531-538.

Csorba, T., A. Bovi, T. Dalmay, and J. Burgyan. 2007. The p122 subunit of Tobacco Mosaic Virus replicase is a potent silencing suppressor and compromises both small interfering RNA- and microRNA-mediated pathways. *J Virol*. 81:11768-11780.

Cuevas, J.M., P. Domingo-Calap, and R. Sanjuan. 2012. The fitness effects of synonymous mutations in DNA and RNA viruses. *Mol Biol Evol*. 29:17-20.

Cuevas, J.M., S.F. Elena, and A. Moya. 2002. Molecular basis of adaptive convergence in experimental populations of RNA viruses. *Genetics*. 162:533-542.

Cullen, B.R. 2006. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J Virol*. 80:1067-1076.

Das, A.T., T.R. Brummelkamp, E.M. Westerhout, M. Vink, M. Madiredjo, R. Bernards, and B. Berkhout. 2004. Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. *J Virol*. 78:2601-2605.

de la Iglesia, F., F. Martinez, J. Hillung, J.M. Cuevas, P.J. Gerrish, J.A. Daros, and S.F. Elena. 2012. Luria-delbruck estimation of turnip mosaic virus mutation rate in vivo. *J Virol*. 86:3386-3388.

de la Pena, M., S.F. Elena, and A. Moya. 2000. Effect of deleterious mutation-accumulation on the fitness of RNA bacteriophage MS2. *Evolution*. 54:686-691.

Desbiez, C., and H. Lecoq. 2004. The nucleotide sequence of Watermelon mosaic virus (WMV, Potyvirus) reveals interspecific recombination between two related potyviruses in the 5' part of the genome. *Arch Virol*. 149:1619-1632.

Di Nicola-Negri, E., A. Brunetti, M. Tavazza, and V. Ilardi. 2005. Hairpin RNA-mediated silencing of Plum pox virus P1 and HC-Pro genes for efficient and predictable resistance to the virus. *Transgenic Res*. 14:989-994.

Diaz-Pendon, J.A., F. Li, W.X. Li, and S.W. Ding. 2007. Suppression of antiviral silencing by cucumber mosaic virus 2b protein in

Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell*. 19:2053-2063.

Diaz, J.A., C. Nieto, E. Moriones, V. Truniger, and M.A. Aranda. 2004. Molecular characterization of a Melon necrotic spot virus strain that overcomes the resistance in melon and nonhost plants. *Mol Plant Microbe Interact*. 17:668-675.

Ding, S.W., W.X. Li, and R.H. Symons. 1995. A novel naturally occurring hybrid gene encoded by a plant RNA virus facilitates long distance virus movement. *EMBO J*. 14:5762-5772.

Ding, S.W., and O. Voinnet. 2007. Antiviral immunity directed by small RNAs. *Cell*. 130:413-426.

Dolja, V.V., R. Haldeman-Cahill, A.E. Montgomery, K.A. Vandenbosch, and J.C. Carrington. 1995. Capsid protein determinants involved in cell-to-cell and long distance movement of tobacco etch potyvirus. *Virology*. 206:1007-1016.

Domingo-Calap, P., J.M. Cuevas, and R. Sanjuan. 2009. The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet*. 5:e1000742.

Dong, X., R. van Wezel, J. Stanley, and Y. Hong. 2003. Functional characterization of the nuclear localization signal for a suppressor of posttranscriptional gene silencing. *J Virol*. 77:7026-7033.

Doron-Faigenboim, A., A. Stern, I. Mayrose, E. Bacharach, and T. Pupko. 2005. Selecton: a server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics*. 21:2101-2103.

Drake, J.W., B. Charlesworth, D. Charlesworth, and J.F. Crow. 1998. Rates of spontaneous mutation. *Genetics*. 148:1667-1686.

Drake, J.W., and J.J. Holland. 1999. Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A*. 96:13910-13913.

Duan, C.G., C.H. Wang, R.X. Fang, and H.S. Guo. 2008. Artificial MicroRNAs highly accessible to targets confer efficient virus resistance in plants. *J Virol*. 82:11084-11095.

Duarte, E., D. Clarke, A. Moya, E. Domingo, and J. Holland. 1992. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc Natl Acad Sci U S A*. 89:6015-6019.

Duffy, S., L.A. Shackelton, and E.C. Holmes. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet*. 9:267-276.

Dunoyer, P., C.H. Lecellier, E.A. Parizotto, C. Himber, and O. Voinnet. 2004. Probing the microRNA and small interfering RNA pathways with

virus-encoded suppressors of RNA silencing. *Plant Cell*. 16:1235-1250.

Edgar, R.C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 5:113.

Elbashir, S.M., W. Lendeckel, and T. Tuschl. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev*. 15:188-200.

Elena, S.F. 1999. Little evidence for synergism among deleterious mutations in a nonsegmented RNA virus. *J Mol Evol*. 49:703-707.

Elena, S.F., S. Bedhomme, P. Carrasco, J.M. Cuevas, F. de la Iglesia, G. Lafforgue, J. Lalic, A. Prosper, N. Tromas, and M.P. Zwart. 2011. The evolutionary genetics of emerging plant RNA viruses. *Mol Plant Microbe Interact*. 24:287-293.

Elena, S.F., and R. Sanjuan. 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. *J Virol*. 79:11555-11558.

Elena, S.F., R.V. Sole, and J. Sardanyes. 2010. Simple genomes, complex interactions: epistasis in RNA virus. *Chaos*. 20:026106.

Escarmis, C., M. Davila, N. Charpentier, A. Bracho, A. Moya, and E. Domingo. 1996. Genetic lesions associated with Muller's ratchet in an RNA virus. *J Mol Biol*. 264:255-267.

Eyre-Walker, A., and P.D. Keightley. 2007. The distribution of fitness effects of new mutations. *Nat Rev Genet*. 8:610-618.

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

Fargette, D., A. Pinel, M. Rakotomalala, E. Sangu, O. Traore, D. Sereme, F. Sorho, S. Issaka, E. Hebrard, Y. Sere, Z. Kanyeka, and G. Konate. 2008. Rice yellow mottle virus, an RNA plant virus, evolves as rapidly as most RNA animal viruses. *J Virol*. 82:3584-3589.

Fisher, R.A. 1930. The Genetical Theory of Natural Selection. *Oxford, Clarendon*

Press.

Force, A., M. Lynch, F.B. Pickett, A. Amores, Y.L. Yan, and J. Postlethwait. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 151:1531-1545.

Fraile, A., J.L. Alonso-Prados, M.A. Aranda, J.J. Bernal, J.M. Malpica, and F. Garcia-Arenal. 1997. Genetic exchange by recombination or reassortment is infrequent in natural populations of a tripartite RNA plant virus. *J Virol*. 71:934-940.

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

Frank, S.A. 1992. A kin selection model for the evolution of virulence. *Proc Biol Sci*. 250:195-197.

Froissart, R., D. Roze, M. Uzest, L. Galibert, S. Blanc, and Y. Michalakis. 2005. Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. *PLoS Biol*. 3:e89.

Fuchs, M., and D. Gonsalves. 2007. Safety of virus-resistant transgenic plants two decades after their introduction: lessons from realistic field risk assessment studies. *Annu Rev Phytopathol*. 45:173-202.

Furio, V., A. Moya, and R. Sanjuan. 2005. The cost of replication fidelity in an RNA virus. *Proc Natl Acad Sci U S A*. 102:10233-10237.

Fusaro, A.F., L. Matthew, N.A. Smith, S.J. Curtin, J. Dedic-Hagan, G.A. Ellacott, J.M. Watson, M.B. Wang, C. Brosnan, B.J. Carroll, and P.M. Waterhouse. 2006. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep*. 7:1168-1175.

Gago, S., S.F. Elena, R. Flores, and R. Sanjuan. 2009. Extremely high mutation rate of a hammerhead viroid. *Science*. 323:1308.

Galperin, M.Y., D.R. Walker, and E.V. Koonin. 1998. Analogous enzymes: independent inventions in enzyme evolution. *Genome Res*. 8:779-790.

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

Garcia-Arenal, F., and B.A. McDonald. 2003. An analysis of the durability of resistance to plant viruses. *Phytopathology*. 93:941-952.

Garcia, S., A. Billecocq, J.M. Crance, M. Prins, D. Garin, and M. Bouloy. 2006. Viral suppressors of RNA interference impair RNA silencing induced by a Semliki Forest virus replicon in tick cells. *J Gen Virol*. 87:1985-1989.

Gavrilets, S. 2010. Rapid transition towards the Division of Labor via evolution of developmental plasticity. *PLoS Comput Biol*. 6:e1000805.

Ge, Q., M.T. McManus, T. Nguyen, C.H. Shen, P.A. Sharp, H.N. Eisen, and J. Chen. 2003. RNA interference of influenza virus production by directly targeting mRNA for degradation and indirectly inhibiting all viral RNA transcription. *Proc Natl Acad Sci U S A*. 100:2718-2723.

Gerlach, W.L., D. Llewellyn, and J. Haseloff. 1987. Construction of a plant disease resistance gene from the satellite RNA of tobacco ringspot virus. *Nature*. 328:802-806.

Gibbs, A., and K. Ohshima. 2010. Potyviruses and the digital revolution. *Annu Rev Phytopathol*. 48:205-223.

Gibbs, A.J., D. Fargette, F. Garcia-Arenal, and M.J. Gibbs. 2010. Time--the emerging dimension of plant virus studies. *J Gen Virol*. 91:13-22.

Gibbs, M.J., and G.F. Weiller. 1999. Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proc Natl Acad Sci U S A*. 96:8022-8027.

Gitlin, L., J.K. Stone, and R. Andino. 2005. Poliovirus escape from RNA interference: short interfering RNA-target recognition and implications for therapeutic approaches. *J Virol*. 79:1027-1035.

Goodrich, D.W., and P.H. Duesberg. 1990. Retroviral recombination during reverse transcription. *Proc Natl Acad Sci U S A*. 87:2052-2056.

Greene, A.E., and R.F. Allison. 1994. Recombination between viral RNA and transgenic plant transcripts. *Science*. 263:1423-1425.

Guo, H.S., and S.W. Ding. 2002. A viral protein inhibits the long range signaling activity of the gene silencing signal. *EMBO J*. 21:398-407.

Guo, H.S., Q. Xie, J.F. Fei, and N.H. Chua. 2005. MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *Plant Cell*. 17:1376-1386.

Haas, G., J. Azevedo, G. Moissiard, A. Geldreich, C. Himber, M. Bureau, T. Fukuhara, M. Keller, and O. Voinnet. 2008. Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. *EMBO J*. 27:2102-2112.

Haasnoot, J., E.M. Westerhout, and B. Berkhout. 2007. RNA interference against viruses: strike and counterstrike. *Nat Biotechnol*. 25:1435-1443.

Haldeman-Cahill, R., J.A. Daros, and J.C. Carrington. 1998. Secondary structures in the capsid protein coding sequence and 3' nontranslated region involved in amplification of the tobacco etch virus genome. *J Virol*. 72:4072-4079.

Hall, J.S., R. French, T.J. Morris, and D.C. Stenger. 2001. Structure and temporal dynamics of populations within wheat streak mosaic virus isolates. *J Virol*. 75:10231-10243.

Hamilton, A.J., and D.C. Baulcombe. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*. 286:950-952.

Harrison, B.D. 1956. The infectivity of extracts made from leaves at intervals after inoculation with viruses. *J Gen Microbiol*. 15:210-220.

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

Haydon, D., N. Knowles, and J. McCauley. 1998. Methods for the detection of non-random base substitution in virus genes: models of synonymous nucleotide substitution in picornavirus genes. *Virus Genes*. 16:253-266.

Hayes, A.J., S.C. Jeong, M.A. Gore, Y.G. Yu, G.R. Buss, S.A. Tolin, and M.A. Maroof. 2004. Recombination within a nucleotide-binding-site/leucine-rich-repeat gene cluster produces new variants conditioning resistance to soybean mosaic virus in soybeans. *Genetics*. 166:493-503.

Heath, M.C. 2000. Nonhost resistance and nonspecific plant defenses. *Curr Opin Plant Biol*. 3:315-319.

Hedrick, P.W. 2004. Genetics of populations. *Jones and Bartlett Publishers, Sudbury, MA.*:376-379.

Hong, Y., K. Saunders, M.R. Hartley, and J. Stanley. 1996. Resistance to geminivirus infection by virus-induced expression of dianthin in transgenic plants. *Virology*. 220:119-127.

Huson, D.H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol*. 23:254-267.

Iino, Y., A. Sugimoto, and M. Yamamoto. 1991. S. pombe pac1+, whose overexpression inhibits sexual development, encodes a ribonuclease III-like RNase. *EMBO J*. 10:221-226.

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

Johnson, R. 1979. Concept of durable resistance. *Phytopathology*. 69:198-199.

Jones, D.A., and D. Takemoto. 2004. Plant innate immunity - direct and indirect recognition of general and specific pathogen-associated molecules. *Curr Opin Immunol*. 16:48-62.

Jones, R.A. 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.

Jridi, C., J.F. Martin, V. Marie-Jeanne, G. Labonne, and S. Blanc. 2006. Distinct viral populations differentiate and evolve independently in a single perennial host plant. *J Virol*. 80:2349-2357.

Kalantidis, K., S. Psaradakis, M. Tabler, and M. Tsagris. 2002. The occurrence of CMV-specific short Rnas in transgenic tobacco expressing virus-

derived double-stranded RNA is indicative of resistance to the virus. *Mol Plant Microbe Interact*. 15:826-833.

Kaltz, O., S. Gandon, Y. Michalakis, and J.A. Shykoff. 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.

Kang, B.C., I. Yeam, and M.M. Jahn. 2005. Genetics of plant virus resistance. *Annu Rev Phytopathol*. 43:581-621.

Kasschau, K.D., Z. Xie, E. Allen, C. Llave, E.J. Chapman, K.A. Krizan, and J.C. Carrington. 2003. P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA unction. *Dev Cell*. 4:205-217.

Kennedy, J.S., M.F. Day, and V.F. Eastop. 1962. A conspectus of aphids as vectors of plant viruses *London : Commonwealth Institute of Entomology*.

Khatchikian, D., M. Orlich, and R. Rott. 1989. Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus. *Nature*. 340:156-157.

Kimura, M. 1991. The neutral theory of molecular evolution: a review of recent evidence. *Jpn J Genet*. 66:367-386.

King, A.M.Q., M.J. Adams, E.B. Carstens, and E.J. Lefkowitz. 2012. Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. *San Diego: Elsevier*:1046-1089.

Klingler, J., R. Creasy, L. Gao, R.M. Nair, A.S. Calix, H.S. Jacob, O.R. Edwards, and K.B. Singh. 2005. Aphid resistance in Medicago truncatula involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs. *Plant Physiol*. 137:1445-1455.

Knapp, S., M. Chase, and J.J. Clarkson. 2004. Nomenclatural changes and a new sectional classification in Nicotiana (Solanaceae). *Taxon*. 53:73-82.

Koonin, E.V., and Y.I. Wolf. 2012. Evolution of microbes and viruses: a paradigm shift in evolutionary biology? *Front Cell Infect Microbiol*. 2:119.

Kouyos, R.D., O.K. Silander, and S. Bonhoeffer. 2007. Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol Evol*. 22:308-315.

Kronke, J., R. Kittler, F. Buchholz, M.P. Windisch, T. Pietschmann, R. Bartenschlager, and M. Frese. 2004. Alternative approaches for

efficient inhibition of hepatitis C virus RNA replication by small interfering RNAs. *J Virol*. 78:3436-3446.

Kubota, K., S. Tsuda, A. Tamai, and T. Meshi. 2003. Tomato mosaic virus replication protein suppresses virus-targeted posttranscriptional gene silencing. *J Virol*. 77:11016-11026.

Lafforgue, G., F. Martinez, J. Sardanyes, F. de la Iglesia, Q.W. Niu, S.S. Lin, R.V. Sole, N.H. Chua, J.A. Daros, and S.F. Elena. 2011. Tempo and mode of plant RNA virus escape from RNA interference-mediated resistance. *J Virol*. 85:9686-9695.

Lakatos, L., T. Csorba, V. Pantaleo, E.J. Chapman, J.C. Carrington, Y.P. Liu, V.V. Dolja, L.F. Calvino, J.J. Lopez-Moya, and J. Burgyan. 2006. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J*. 25:2768-2780.

Lakatos, L., G. Szittya, D. Silhavy, and J. Burgyan. 2004. Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *EMBO J*. 23:876-884.

Lalic, J., J.M. Cuevas, and S.F. Elena. 2011. Effect of host species on the distribution of mutational fitness effects for an RNA virus. *PLoS Genet*. 7:e1002378.

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

Langenberg, W.G., L. Zhang, D.L. Court, L. Giunchedi, and A. Mitra. 1997. Transgenic tobacco plants expressing the bacterial mc gene resist virus infection *Mol Breeding*. 3:391-399.

Lapidot, M., R. Gafny, B. Ding, S. Wolf, W.J. Lucas, and R.N. Beachy. 1993. A dysfunctional movement protein of tobacco mosaic virus that partially modifies the plasmodesmata and limits virus spread in transgenic plants. *Plant J*. 4:959-970.

Larsen, R.C., P.N. Miklas, K.L. Druffel, and S.D. Wyatt. 2005. NL-3 K Strain Is a Stable and Naturally Occurring Interspecific Recombinant Derived from Bean common mosaic necrosis virus and Bean common mosaic virus. *Phytopathology*. 95:1037-1042.

Lecoq, H., B. Moury, C. Desbiez, A. Palloix, and M. Pitrat. 2004. Durable virus resistance in plants through conventional approaches: a challenge. *Virus Res*. 100:31-39.

Legg, J.P., and J.M. Thresh. 2000. Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. *Virus Res*. 71:135-149.

Lenski, R.E. 1991. Quantifying fitness and gene stability in microorganisms. *Biotechnology*. 15:173-192.

Lentini, Z., I. Lozano, E. Tabares, L. Fory, J. Dominguez, M. Cuervo, and L. Calvert. 2003. Expression and inheritance of hypersensitive resistance to rice hoja blanca virus mediated by the viral nucleocapsid protein gene in transgenic rice. *Theor Appl Genet*. 106:1018-1026.

Leonard, S., D. Plante, S. Wittmann, N. Daigneault, M.G. Fortin, and J.F. Laliberte. 2000. Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. *J Virol*. 74:7730-7737.

Leonard, S., C. Viel, C. Beauchemin, N. Daigneault, M.G. Fortin, and J.F. Laliberte. 2004. Interaction of VPg-Pro of turnip mosaic virus with the translation initiation factor 4E and the poly(A)-binding protein in planta. *J Gen Virol*. 85:1055-1063.

Lin, S.S., H.W. Wu, S.F. Elena, K.C. Chen, Q.W. Niu, S.D. Yeh, C.C. Chen, and N.H. Chua. 2009. Molecular evolution of a viral non-coding sequence under the selective pressure of amiRNA-mediated silencing. *PLoS Pathog*. 5:e1000312.

Lindbo, J.A., and W.G. Dougherty. 2005. Plant pathology and RNAi: a brief history. *Annu Rev Phytopathol*. 43:191-204.

Lindbo, J.A., L. Silva-Rosales, W.M. Proebsting, and W.G. Dougherty. 1993. Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *Plant Cell*. 5:1749-1759.

Llave, C., K.D. Kasschau, M.A. Rector, and J.C. Carrington. 2002. Endogenous and silencing-associated small RNAs in plants. *Plant Cell*. 14:1605-1619.

Lodge, J.K., W.K. Kaniewski, and N.E. Tumer. 1993. Broad-spectrum virus resistance in transgenic plants expressing pokeweed antiviral protein. *Proc Natl Acad Sci U S A*. 90:7089-7093.

Loewe, L. 2006. Quantifying the genomic decay paradox due to Muller's ratchet in human mitochondrial DNA. *Genet Res*. 87:133-159.

Lopez-Moya, J.J., and J.A. Garcia. 1999. Potyviruses (Potyviridae). *In Encyclopedia of Virology. Edited by A. Granoff & R. G. Webster. San Diego: Academic Press*. 2nd edn,:1369-1375.

Lovisolo, O., R. Hull, and O. Rosler. 2003. Coevolution of viruses with hosts and vectors and possible paleontology. *Adv Virus Res*. 62:325-379.

Lozsa, R., T. Csorba, L. Lakatos, and J. Burgyan. 2008. Inhibition of 3' modification of small RNAs in virus-infected plants require spatial

and temporal co-expression of small RNAs and viral silencing-suppressor proteins. *Nucleic Acids Res*. 36:4099-4107.

Lu, C., and N. Fedoroff. 2000. A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. *Plant Cell*. 12:2351-2366.

Lu, R., A. Folimonov, M. Shintaku, W.X. Li, B.W. Falk, W.O. Dawson, and S.W. Ding. 2004. Three distinct suppressors of RNA silencing encoded by a 20-kb viral RNA genome. *Proc Natl Acad Sci U S A*. 101:15742-15747.

MacFarlane, S.A. 1997. NATURAL RECOMBINATION AMONG PLANT VIRUS GENOMES : EVIDENCE FROM TOBRAVIRUSES *Seminars in virology*. 8:25-31.

Malim, M.H., and M. Emerman. 2001. HIV-1 sequence variation: drift, shift, and attenuation. *Cell*. 104:469-472.

Mallory, A.C., S. Mlotshwa, L.H. Bowman, and V.B. Vance. 2003. The capacity of transgenic tobacco to send a systemic RNA silencing signal depends on the nature of the inducing transgene locus. *Plant J*. 35:82-92.

Malpica, J.M., A. Fraile, I. Moreno, C.I. Obies, J.W. Drake, and F. Garcia-Arenal. 2002. The rate and character of spontaneous mutation in an RNA virus. *Genetics*. 162:1505-1511.

Malyshenko, S.I., O.A. Kondakova, J.V. Nazarova, I.B. Kaplan, M.E. Taliansky, and G.F.J. Atabekov. 1993. Reduction of tobacco mosaic virus accumulation in transgenic plants producing non-functional viral transport proteins. *J Gen Virol*. 74:1149-1156.

Mansky, L.M. 1998. Retrovirus mutation rates and their role in genetic variation. *J Gen Virol*. 79 (Pt 6):1337-1345.

Marin, J., and R.V. Solé. 1999. Macroevolutionary algorithms: a new optimization method on fitness landscapes. *IEEE Trans. Evol. Comput.* 3:272-286.

Marsh, G.A., R. Rabadan, A.J. Levine, and P. Palese. 2008. Highly conserved regions of influenza a virus polymerase gene segments are critical for efficient viral RNA packaging. *J Virol*. 82:2295-2304.

Martinez-Picado, J., M.P. DePasquale, N. Kartsonis, G.J. Hanna, J. Wong, D. Finzi, E. Rosenberg, H.F. Gunthard, L. Sutton, A. Savara, C.J. Petropoulos, N. Hellmann, B.D. Walker, D.D. Richman, R. Siliciano, and R.T. D'Aquila. 2000. Antiretroviral resistance during successful therapy of HIV type 1 infection. *Proc Natl Acad Sci U S A*. 97:10948-10953.

Martinez, F., G. Lafforgue, M.J. Morelli, F. Gonzalez-Candelas, N.H. Chua, J.A. Daros, and S.F. Elena. 2012. Ultradeep Sequencing Analysis of Population Dynamics of Virus Escape Mutants in RNAi-Mediated Resistant Plants. *Mol Biol Evol*.

Martinez, F., J. Sardanyes, S.F. Elena, and J.A. Daros. 2011. Dynamics of a plant RNA virus intracellular accumulation: stamping machine vs. geometric replication. *Genetics*. 188:637-646.

Martinez, J., A. Patkaniowska, H. Urlaub, R. Luhrmann, and T. Tuschl. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*. 110:563-574.

Mascia, T., F. Cillo, V. Fanelli, M.M. Finetti-Sialer, A. De Stradis, P. Palukaitis, and D. Gallitelli. 2010. Characterization of the interactions between Cucumber mosaic virus and Potato virus Y in mixed infections in tomato. *Mol Plant Microbe Interact*. 23:1514-1524.

Maule, A.J., C. Caranta, and M.I. Boulton. 2007. Sources of natural resistance to plant viruses: status and prospects. *Mol Plant Pathol*. 8:223-231.

Merai, Z., Z. Kerenyi, S. Kertesz, M. Magna, L. Lakatos, and D. Silhavy. 2006. Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. *J Virol*. 80:5747-5756.

Merai, Z., Z. Kerenyi, A. Molnar, E. Barta, A. Valoczi, G. Bisztray, Z. Havelda, J. Burgyan, and D. Silhavy. 2005. Aureusvirus P14 is an efficient RNA silencing suppressor that binds double-stranded RNAs without size specificity. *J Virol*. 79:7217-7226.

Miralles, R., A. Moya, and S.F. Elena. 1999. Effect of population patchiness and migration rates on the adaptation and divergence of vesicular stomatitis virus quasispecies populations. *J Gen Virol*. 80 (Pt 8):2051-2059.

Missiou, A., K. Kalantidis, A. Boutla, S. Tzortzakaki, M. Tabler, and M. Tsagris. 2004. Generation of transgenic potato plants highly resistant to Potato virus Y (PVY) through RNA silencing. *Mol Breeding*. 14:185-197.

Mlotshwa, S., O. Voinnet, M.F. Mette, M. Matzke, H. Vaucheret, S.W. Ding, G. Pruss, and V.B. Vance. 2002. RNA silencing and the mobile silencing signal. *Plant Cell*. 14 Suppl:S289-301.

Monci, F., S. Sanchez-Campos, J. Navas-Castillo, and E. Moriones. 2002. A natural recombinant between the geminiviruses Tomato yellow leaf curl Sardinia virus and Tomato yellow leaf curl virus exhibits a novel pathogenic phenotype and is becoming prevalent in Spanish populations. *Virology*. 303:317-326.

Montarry, J., E. Cartier, M. Jacquemond, A. Palloix, and B. Moury. 2012. Virus adaptation to quantitative plant resistance: erosion or breakdown? *J Evol Biol*. 25:2242-2252.

Mosquera, J., and F.R. Adler. 1998. Evolution of virulence: a unified framework for coinfection and superinfection. *J Theor Biol*. 195:293-313.

Moury, B., F. Fabre, and R. Senoussi. 2007. Estimation of the number of virus particles transmitted by an insect vector. *Proc Natl Acad Sci U S A*. 104:17891-17896.

Moury, B., C. Morel, E. Johansen, L. Guilbaud, S. Souche, V. Ayme, C. Caranta, A. Palloix, and M. Jacquemond. 2004. Mutations in potato virus Y genome-linked protein determine virulence toward recessive resistances in Capsicum annuum and Lycopersicon hirsutum. *Mol Plant Microbe Interact*. 17:322-329.

Mukai, T., S. Chigusa, and I. Yoshikawa. 1964. The Genetic Structure of Natural Populations of Drosophila Melanogaster. Ii. Overdominance of Spontaneous Mutant Polygenes Controlling Viability in Homozygous Genetic Background. *Genetics*. 50:711-715.

Muller, H.J. 1964. The Relation of Recombination to Mutational Advance. *Mutat Res*. 106:2-9.

Mysore, K.S., and C.M. Ryu. 2004. Nonhost resistance: how much do we know? *Nonhost resistance: how much do we know?* 9:97-104.

Napoli, C., C. Lemieux, and R. Jorgensen. 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell*. 2:279-289.

Neher, R.A., and T. Leitner. 2010. Recombination rate and selection strength in HIV intra-patient evolution. *PLoS Comput Biol*. 6:e1000660.

Nicolas, O., S.W. Dunnington, L.F. Gotow, T.P. Pirone, and G.M. Hellmann. 1997. Variations in the VPg protein allow a potyvirus to overcome va gene resistance in tobacco. *Virology*. 237:452-459.

Niu, Q.W., S.S. Lin, J.L. Reyes, K.C. Chen, H.W. Wu, S.D. Yeh, and N.H. Chua. 2006. Expression of artificial microRNAs in transgenic Arabidopsis thaliana confers virus resistance. *Nat Biotechnol*. 24:1420-1428.

Nora, T., C. Charpentier, O. Tenaillon, C. Hoede, F. Clavel, and A.J. Hance. 2007. Contribution of recombination to the evolution of human immunodeficiency viruses expressing resistance to antiretroviral treatment. *J Virol*. 81:7620-7628.

Novella, I.S., J. Quer, E. Domingo, and J.J. Holland. 1999. Exponential fitness gains of RNA virus populations are limited by bottleneck effects. *J Virol*. 73:1668-1671.

Nowak, M.A., and R.M. May. 1994. Superinfection and the evolution of parasite virulence. *Proc Biol Sci*. 255:81-89.

Ohrt, T., J. Muetze, P. Svoboda, and P. Schwille. 2012. Intracellular localization and routing of miRNA and RNAi pathway components. *Curr Top Med Chem*. 12:79-88.

Ohshima, K., Y. Yamaguchi, R. Hirota, T. Hamamoto, K. Tomimura, Z. Tan, T. Sano, F. Azuhata, J.A. Walsh, J. Fletcher, J. Chen, A. Gera, and A. Gibbs. 2002. Molecular evolution of Turnip mosaic virus: evidence of host adaptation, genetic recombination and geographical spread. *J Gen Virol*. 83:1511-1521.

Ohtsuki, A., and A. Sasaki. 2006. Epidemiology and disease-control under gene-for-gene plant-pathogen interaction. *J Theor Biol*. 238:780-794.

Oldroyd, G.E., and B.J. Staskawicz. 1998. Genetically engineered broad-spectrum disease resistance in tomato. *Proc Natl Acad Sci U S A*. 95:10300-10305.

Orr, H.A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution*. 52:935-949.

Orr, H.A. 2000. Adaptation and the cost of complexity. *Evolution*. 54:13-20.

Padidam, M., S. Sawyer, and C.M. Fauquet. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology*. 265:218-225.

Palukaitis, P., and F. Garcia-Arenal. 2003. Cucumoviruses. *Adv Virus Res*. 62:241-323.

Palukaitis, P., and M. Zaitlin. 1997. Replicase-mediated resistance to plant virus disease. *Adv Virus Res*. 48:349-377.

Patwa, Z., and L.M. Wahl. 2008. The fixation probability of beneficial mutations. *J R Soc Interface*. 5:1279-1289.

Peck, J.R., G. Barreau, and S.C. Heath. 1997. Imperfect genes, Fisherian mutation and the evolution of sex. *Genetics*. 145:1171-1199.

Pfeffer, S., P. Dunoyer, F. Heim, K.E. Richards, G. Jonard, and V. Ziegler-Graff. 2002. P0 of beet Western yellows virus is a suppressor of posttranscriptional gene silencing. *J Virol*. 76:6815-6824.

Phillips, P.C. 2008. Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet*. 9:855-867.

Pinel-Galzi, A., M. Rakotomalala, E. Sangu, F. Sorho, Z. Kanyeka, O. Traore, D. Sereme, N. Poulicard, Y. Rabenantoandro, Y. Sere, G. Konate, A. Ghesquiere, E. Hebrard, and D. Fargette. 2007. Theme and variations in the evolutionary pathways to virulence of an RNA plant virus species. *PLoS Pathog*. 3:e180.

Pirone, T.P., and S. Blanc. 1996. Helper-dependent vector transmission of plant viruses. *Annu Rev Phytopathol*. 34:227-247.

Pita, J.S., J.R. de Miranda, W.L. Schneider, and M.J. Roossinck. 2007. Environment determines fidelity for an RNA virus replicase. *J Virol*. 81:9072-9077.

Plisson, C., M. Drucker, S. Blanc, S. German-Retana, O. Le Gall, D. Thomas, and P. Bron. 2003. Structural characterization of HC-Pro, a plant virus multifunctional protein. *J Biol Chem*. 278:23753-23761.

Plotkin, J.B., and G. Kudla. 2011. Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*. 12:32-42.

Poulicard, N., A. Pinel-Galzi, E. Hebrard, and D. Fargette. 2010. Why Rice yellow mottle virus, a rapidly evolving RNA plant virus, is not efficient at breaking rymv1-2 resistance. *Mol Plant Pathol*. 11:145-154.

Price, R.N., A.C. Uhlemann, A. Brockman, R. McGready, E. Ashley, L. Phaipun, R. Patel, K. Laing, S. Looareesuwan, N.J. White, F. Nosten, and S. Krishna. 2004. Mefloquine resistance in Plasmodium falciparum and increased pfmdr1 gene copy number. *Lancet*. 364:438-447.

Pruss, G., X. Ge, X.M. Shi, J.C. Carrington, and V. Bowman Vance. 1997. Plant viral synergism: the potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell*. 9:859-868.

Pruss, G.J., C.B. Lawrence, T. Bass, Q.Q. Li, L.H. Bowman, and V. Vance. 2004. The potyviral suppressor of RNA silencing confers enhanced resistance to multiple pathogens. *Virology*. 320:107-120.

Puigbo, P., S. Garcia-Vallve, and J.O. McInerney. 2007. TOPD/FMTS: a new software to compare phylogenetic trees. *Bioinformatics*. 23:1556-1558.

Qi, Y., X. Zhong, A. Itaya, and B. Ding. 2004. Dissecting RNA silencing in protoplasts uncovers novel effects of viral suppressors on the silencing pathway at the cellular level. *Nucleic Acids Res*. 32:e179.

Qu, J., J. Ye, and R. Fang. 2007. Artificial microRNA-mediated virus resistance in plants. *J Virol*. 81:6690-6699.

Rana, T.M. 2007. Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol*. 8:23-36.

Reed, J.C., K.D. Kasschau, A.I. Prokhnevsky, K. Gopinath, G.P. Pogue, J.C. Carrington, and V.V. Dolja. 2003. Suppressor of RNA silencing encoded by Beet yellows virus. *Virology*. 306:203-209.

Register, J.C., 3rd, and R.N. Beachy. 1988. Resistance to TMV in transgenic plants results from interference with an early event in infection. *Virology*. 166:524-532.

Ren, T., F. Qu, and T.J. Morris. 2000. HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. *Plant Cell*. 12:1917-1926.

Ribeiro, R.M., S. Bonhoeffer, and M.A. Nowak. 1998. The frequency of resistant mutant virus before antiviral therapy. *AIDS*. 12:461-465.

Robaglia, C., and C. Caranta. 2006. Translation initiation factors: a weak link in plant RNA virus infection. *Trends Plant Sci*. 11:40-45.

Roossinck, M.J. 2002. Evolutionary history of Cucumber mosaic virus deduced by phylogenetic analyses. *J Virol*. 76:3382-3387.

Rotondo, G., J.Y. Huang, and D. Frendewey. 1997. Substrate structure requirements of the Pac1 ribonuclease from Schizosaccharmyces pombe. *RNA*. 3:1182-1193.

Ryabov, E.V., G. Fraser, M.A. Mayo, H. Barker, and M. Taliansky. 2001. Umbravirus gene expression helps potato leafroll virus to invade mesophyll tissues and to be transmitted mechanically between plants. *Virology*. 286:363-372.

Sabariegos, R., M. Gimenez-Barcons, N. Tapia, B. Clotet, and M.A. Martinez. 2006. Sequence homology required by human immunodeficiency virus type 1 to escape from short interfering RNAs. *J Virol*. 80:571-577.

Sacristan, S., and F. Garcia-Arenal. 2008. The evolution of virulence and pathogenicity in plant pathogen populations. *Mol Plant Pathol*. 9:369-384.

Sacristan, S., J.M. Malpica, A. Fraile, and F. Garcia-Arenal. 2003. Estimation of population bottlenecks during systemic movement of tobacco mosaic virus in tobacco plants. *J Virol*. 77:9906-9911.

Saenz, P., B. Salvador, C. Simon-Mateo, K.D. Kasschau, J.C. Carrington, and J.A. Garcia. 2002. Host-specific involvement of the HC protein in the long-distance movement of potyviruses. *J Virol*. 76:1922-1931.

Sanjuan, R. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed

mutagenesis studies. *Philos Trans R Soc Lond B Biol Sci*. 365:1975-1982.

Sanjuan, R., P. Agudelo-Romero, and S.F. Elena. 2009. Upper-limit mutation rate estimation for a plant RNA virus. *Biol Lett*. 5:394-396.

Sanjuan, R., A. Moya, and S.F. Elena. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc Natl Acad Sci U S A*. 101:8396-8401.

Sardanyes, J., R.V. Sole, and S.F. Elena. 2009. Replication mode and landscape topology differentially affect RNA virus mutational load and robustness. *J Virol*. 83:12579-12589.

Savolainen-Kopra, C., and S. Blomqvist. 2010. Mechanisms of genetic variation in polioviruses. *Rev Med Virol*. 20:358-371.

Schenk, P.M., K. Kazan, I. Wilson, J.P. Anderson, T. Richmond, S.C. Somerville, and J.M. Manners. 2000. Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. *Proc Natl Acad Sci U S A*. 97:11655-11660.

Schneider, W.L., and M.J. Roossinck. 2001. Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *J Virol*. 75:6566-6571.

Seo, Y.S., M.R. Rojas, J.Y. Lee, S.W. Lee, J.S. Jeon, P. Ronald, W.J. Lucas, and R.L. Gilbertson. 2006. A viral resistance gene from common bean functions across plant families and is up-regulated in a non-virus-specific manner. *Proc Natl Acad Sci U S A*. 103:11856-11861.

Sequeira, L. 1983. Mechanisms of induced resistance in plants. *Annu Rev Microbiol*. 37:51-79.

Shams-Bakhsh, M., T. Canto, and P. Palukaitis. 2007. Enhanced resistance and neutralization of defense responses by suppressors of RNA silencing. *Virus Res*. 130:103-109.

Shriner, D., A.G. Rodrigo, D.C. Nickle, and J.I. Mullins. 2004. Pervasive genomic recombination of HIV-1 in vivo. *Genetics*. 167:1573-1583.

Shukla, D.D., C.W. Ward, and A.A. Brunt. 1994. The Potyviridae. . *Wallingford : CAB International*.

Silberstein, L., I. Kovalski, Y. Brotman, C. Perin, C. Dogimont, M. Pitrat, J. Klingler, G. Thompson, V. Portnoy, N. Katzir, and R. Perl-Treves. 2003. Linkage map of Cucumis melo including phenotypic traits and sequence-characterized genes. *Genome*. 46:761-773.

Silhavy, D., and J. Burgyan. 2004. Effects and side-effects of viral RNA silencing suppressors on short RNAs. *Trends Plant Sci*. 9:76-83.

Silhavy, D., A. Molnar, A. Lucioli, G. Szittya, C. Hornyik, M. Tavazza, and J. Burgyan. 2002. A viral protein suppresses RNA silencing and binds

silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *EMBO J*. 21:3070-3080.

Simon-Loriere, E., and E.C. Holmes. 2011. Why do RNA viruses recombine? *Nat Rev Microbiol*. 9:617-626.

Sire, C., M. Bangratz-Reyser, D. Fargette, and C. Brugidou. 2008. Genetic diversity and silencing suppression effects of Rice yellow mottle virus and the P1 protein. *Virol J*. 5:55.

Sniegowski, P.D., and P.J. Gerrish. 2010. Beneficial mutations and the dynamics of adaptation in asexual populations. *Philos Trans R Soc Lond B Biol Sci*. 365:1255-1263.

Stamatakis, A., T. Ludwig, and H. Meier. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics*. 21:456-463.

Stanley, J., T. Frischmuth, and S. Ellwood. 1990. Defective viral DNA ameliorates symptoms of geminivirus infection in transgenic plants. *Proc Natl Acad Sci U S A*. 87:6291-6295.

Staskawicz, B.J., F.M. Ausubel, B.J. Baker, J.G.J. Ellis, and J.D. ones. 1995. Molecular genetics of plant disease resistance. *Science*. 5:661-667.

Steinhauer, D.A., E. Domingo, and J.J. Holland. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*. 122:281-288.

Stoltzfus, A. 1999. On the possibility of constructive neutral evolution. *J Mol Evol*. 49:169-181.

Takahashi, H., M. Suzuki, K. Natsuaki, T. Shigyo, K. Hino, T. Teraoka, D. Hosokawa, and Y. Ehara. 2001. Mapping the virus and host genes involved in the resistance response in cucumber mosaic virus-Infected Arabidopsis thaliana. *Plant Cell Physiol*. 42:340-347.

Takeda, A., K. Sugiyama, H. Nagano, M. Mori, M. Kaido, K. Mise, S. Tsuda, and T. Okuno. 2002. Identification of a novel RNA silencing suppressor, NSs protein of Tomato spotted wilt virus. *FEBS Lett*. 532:75-79.

Takeshita, M., E. Koizumi, M. Noguchi, K. Sueda, H. Shimura, N. Ishikawa, H. Matsuura, K. Ohshima, T. Natsuaki, S. Kuwata, N. Furuya, K. Tsuchiya, and C. Masuta. 2012. Infection dynamics in viral spread and interference under the synergism between Cucumber mosaic virus and Turnip mosaic virus. *Mol Plant Microbe Interact*. 25:18-27.

Tan, Z., Y. Wada, J. Chen, and K. Ohshima. 2004. Inter- and intralineage recombinants are common in natural populations of Turnip mosaic virus. *J Gen Virol*. 85:2683-2696.

Tang, R.S., D.J. Barton, J.B. Flanegan, and K. Kirkegaard. 1997. Poliovirus RNA recombination in cell-free extracts. *RNA*. 3:624-633.

Taucher, C., A. Berger, and C.W. Mandl. 2010. A trans-complementing recombination trap demonstrates a low propensity of flaviviruses for intermolecular recombination. *J Virol*. 84:599-611.

Tavladoraki, P., E. Benvenuto, S. Trinca, D. De Martinis, A. Cattaneo, and P. Galeffi. 1993. Transgenic plants expressing a functional single-chain Fv antibody are specifically protected from virus attack. *Nature*. 366:469-472.

Thomas, C.L., V. Leh, C. Lederer, and A.J. Maule. 2003. Turnip crinkle virus coat protein mediates suppression of RNA silencing in Nicotiana benthamiana. *Virology*. 306:33-41.

Tomimura, K., A.J. Gibbs, C.E. Jenner, J.A. Walsh, and K. Ohshima. 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*. 12:2099-2111.

Tomimura, K., J. Spak, N. Katis, C.E. Jenner, J.A. Walsh, A.J. Gibbs, and K. Ohshima. 2004. Comparisons of the genetic structure of populations of Turnip mosaic virus in West and East Eurasia. *Virology*. 330:408-423.

Traore, O., A. Pinel-Galzi, S. Issaka, N. Poulicard, J. Aribi, S. Ake, A. Ghesquiere, Y. Sere, G. Konate, E. Hebrard, and D. Fargette. 2010. The adaptation of Rice yellow mottle virus to the eIF(iso)4G-mediated rice resistance. *Virology*. 408:103-108.

Tromas, N., and S.F. Elena. 2010. The rate and spectrum of spontaneous mutations in a plant RNA virus. *Genetics*. 185:983-989.

Tugume, A.K., S.B. Mukasa, N. Kalkkinen, and J.P. Valkonen. 2010. Recombination and selection pressure in the ipomovirus sweet potato mild mottle virus (Potyviridae) in wild species and cultivated sweetpotato in the centre of evolution in East Africa. *J Gen Virol*. 91:1092-1108.

Tumer, N.E., D.J. Hwang, and M. Bonness. 1997. C-terminal deletion mutant of pokeweed antiviral protein inhibits viral infection but does not depurinate host ribosomes. *Proc Natl Acad Sci U S A*. 94:3866-3871.

Turturo, C., A. Friscina, S. Gaubert, M. Jacquemond, J.R. Thompson, and M. Tepfer. 2008. Evaluation of potential risks associated with recombination in transgenic plants expressing viral sequences. *J Gen Virol*. 89:327-335.

Uknes, S., B. Mauch-Mani, M. Moyer, S. Potter, S. Williams, S. Dincher, D. Chandler, A. Slusarenko, E. Ward, and J. Ryals. 1992. Acquired resistance in Arabidopsis. *Plant Cell*. 4:645-656.

Urbanowicz, A., M. Alejska, P. Formanowicz, J. Blazewicz, M. Figlerowicz, and J.J. Bujarski. 2005. Homologous crossovers among molecules of brome mosaic bromovirus RNA1 or RNA2 segments in vivo. *J Virol*. 79:5732-5742.

Urcuqui-Inchima, S., A.L. Haenni, and F. Bernardi. 2001. Potyvirus proteins: a wealth of functions. *Virus Res*. 74:157-175.

Vallad, G.E., and R.M. Goodman. 2004. Systemic acquired resistance and induced systemic resistance in conventional agriculture. *Crop Science*. 44:1920-1934.

Valli, A., J.J. Lopez-Moya, and J.A. Garcia. 2007. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. *J Gen Virol*. 88:1016-1028.

van den Bosch, F., and C.A. Gilligan. 2003. Measures of durability of resistance. *Phytopathology*. 93:616-625.

van der Walt, E., E.P. Rybicki, A. Varsani, J.E. Polston, R. Billharz, L. Donaldson, A.L. Monjane, and D.P. Martin. 2009. Rapid host adaptation by extensive recombination. *J Gen Virol*. 90:734-746.

Vanitharani, R., P. Chellappan, J.S. Pita, and C.M. Fauquet. 2004. Differential roles of AC2 and AC4 of cassava geminiviruses in mediating synergism and suppression of posttranscriptional gene silencing. *J Virol*. 78:9487-9498.

Varallyay, E., A. Valoczi, A. Agyi, J. Burgyan, and Z. Havelda. 2010. Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *EMBO J*. 29:3507-3519.

Varkonyi-Gasic, E., R. Wu, M. Wood, E.F. Walton, and R.P. Hellens. 2007. Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods*. 3:12.

Vaucheret, H., and M. Fagard. 2001. Transcriptional gene silencing in plants: targets, inducers and regulators. *Trends Genet*. 17:29-35.

Vaucheret, H., F. Vazquez, P. Crete, and D.P. Bartel. 2004. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev*. 18:1187-1197.

Vazquez, J., M.F. Correa de Adjounian, C. Sumners, A. Gonzalez, C. Diez-Freire, and M.K. Raizada. 2005. Selective silencing of angiotensin receptor subtype 1a (AT1aR) by RNA interference. *Hypertension*. 45:115-119.

Vera Cruz, C.M., J. Bai, I. Ona, H. Leung, R.J. Nelson, T.W. Mew, and J.E. Leach. 2000. Predicting durability of a disease resistance gene based on an assessment of the fitness loss and epidemiological consequences of avirulence gene mutation. *Proc Natl Acad Sci U S A*. 97:13500-13505.

Vidal, S., H. Cabrera, R.A. Andersson, A. Fredriksson, and J.P. Valkonen. 2002. Potato gene Y-1 is an N gene homolog that confers cell death upon infection with potato virus Y. *Mol Plant Microbe Interact*. 15:717-727.

Vijayapalani, P., M. Maeshima, N. Nagasaki-Takekuchi, and W.A. Miller. 2012. Interaction of the trans-frame potyvirus protein P3N-PIPO with host protein PCaP1 facilitates potyvirus movement. *PLoS Pathog*. 8:e1002639.

Vincenzetti, S., A. Cambi, J. Neuhard, K. Schnorr, M. Grelloni, and A. Vita. 1999. Cloning, expression, and purification of cytidine deaminase from Arabidopsis thaliana. *Protein Expr Purif*. 15:8-15.

Voinnet, O., C. Lederer, and D.C. Baulcombe. 2000. A viral movement protein prevents spread of the gene silencing signal in Nicotiana benthamiana. *Cell*. 103:157-167.

Voinnet, O., Y.M. Pinto, and D.C. Baulcombe. 1999. Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc Natl Acad Sci U S A*. 96:14147-14152.

von Eije, K.J., O. ter Brake, and B. Berkhout. 2008. Human immunodeficiency virus type 1 escape is restricted when conserved genome sequences are targeted by RNA interference. *J Virol*. 82:2895-2903.

Vos, P., G. Simons, T. Jesse, J. Wijbrandi, L. Heinen, R. Hogers, A. Frijters, J. Groenendijk, P. Diergaarde, M. Reijans, J. Fierens-Onstenk, M. de Both, J. Peleman, T. Liharska, J. Hontelez, and M. Zabeau. 1998. The tomato Mi-1 gene confers resistance to both root-knot nematodes and potato aphids. *Nat Biotechnol*. 16:1365-1369.

Vu, T.V., N. Roy Choudhury, and S.K. Mukherjee. 2013. Transgenic tomato plants expressing artificial microRNAs for silencing the pre-coat and coat proteins of a begomovirus, Tomato leaf curl New Delhi virus, show tolerance to virus infection. *Virus Res*. 172:35-45.

Wang, H., K.J. Buckley, X. Yang, R.C. Buchmann, and D.M. Bisaro. 2005. Adenosine kinase inhibition and suppression of RNA silencing by geminivirus AL2 and L2 proteins. *J Virol*. 79:7410-7418.

Wang, H., L. Hao, C.Y. Shung, G. Sunter, and D.M. Bisaro. 2003. Adenosine kinase is inactivated by geminivirus AL2 and L2 proteins. *Plant Cell*. 15:3020-3032.

Wang, M.B., D.C. Abbott, and P.M. Waterhouse. 2000. A single copy of a virus-derived transgene encoding hairpin RNA gives immunity to barley yellow dwarf virus. *Mol Plant Pathol*. 1:347-356.

Warthmann, N., H. Chen, S. Ossowski, D. Weigel, and P. Herve. 2008. Highly specific gene silencing by artificial miRNAs in rice. *PLoS One*. 3:e1829.

Wassenegger, M., and T. Pelissier. 1998. A model for RNA-mediated gene silencing in higher plants. *Plant Mol Biol*. 37:349-362.

Watanabe, Y., T. Ogawa, H. Takahashi, I. Ishida, Y. Takeuchi, M. Yamamoto, and Y. Okada. 1995. Resistance against multiple plant viruses in plants mediated by a double stranded-RNA specific ribonuclease. *FEBS Lett*. 372:165-168.

Weber, H., and A.J. Pfitzner. 1998. Tm-2(2) resistance in tomato requires recognition of the carboxy terminus of the movement protein of tomato mosaic virus. *Mol Plant Microbe Interact*. 11:498-503.

Wege, C., and D. Siegmund. 2007. Synergism of a DNA and an RNA virus: enhanced tissue infiltration of the begomovirus Abutilon mosaic virus (AbMV) mediated by Cucumber mosaic virus (CMV). *Virology*. 357:10-28.

Wernersson, R., and A.G. Pedersen. 2003. RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res*. 31:3537-3539.

Westerhout, E.M., and B. Berkhout. 2007. A systematic analysis of the effect of target RNA structure on RNA interference. *Nucleic Acids Res*. 35:4322-4330.

Westerhout, E.M., M. Ooms, M. Vink, A.T. Das, and B. Berkhout. 2005. HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Res*. 33:796-804.

White, P.S., F. Morales, and M.J. Roossinck. 1995. Interspecific reassortment of genomic segments in the evolution of cucumoviruses. *Virology*. 207:334-337.

Whitelam, G.C., and W. Cockburn. 1996. Antibody expression in transgenic plants. *Trends Plant Sci*. 1:268-272.

Whitham, S., S. McCormick, and B. Baker. 1996. The N gene of tobacco confers resistance to tobacco mosaic virus in transgenic tomato. *Proc Natl Acad Sci U S A*. 93:8776-8781.

Whitham, S.A., R.J. Anderberg, S.T. Chisholm, and J.C. Carrington. 2000. Arabidopsis RTM2 gene is necessary for specific restriction of tobacco etch virus and encodes an unusual small heat shock-like protein. *Plant Cell*. 12:569-582.

Wichman, H.A., M.R. Badgett, L.A. Scott, C.M. Boulianne, and J.J. Bull. 1999. Different trajectories of parallel evolution during viral adaptation. *Science*. 285:422-424.

Wright, S. 1931. Evolution in Mendelian Populations. *Genetics*. 16:97-159.

Wu, Q., X. Wang, and S.W. Ding. 2010. Viral suppressors of RNA-based viral immunity: host targets. *Cell Host Microbe*. 8:12-15.

Yelina, N.E., E.I. Savenkov, A.G. Solovyev, S.Y. Morozov, and J.P. Valkonen. 2002. Long-distance movement, virulence, and RNA silencing suppression controlled by a single protein in hordei- and potyviruses: complementary functions between virus families. *J Virol*. 76:12981-12991.

Zaccomer, B., F. Cellier, J.C. Boyer, A.L. Haenni, and M. Tepfer. 1993. Transgenic plants that express genes including the 3' untranslated region of the turnip yellow mosaic virus (TYMV) genome are partially protected against TYMV infection. *Gene*. 136:87-94.

Zhan, J., C.C. Mundt, M.E. Hoffer, and B.A. McDonald. 2002. Local adaptation and effect of host genotype on the rate of pathogen evolution: an experimental test in a plant pathosystem. *J. Evol. Biol.* 15:634-647.

Zhang, Y. 2003. Evolution by gene duplication: an update. *Trends in Ecology & Evolution* 18:292-298.

Zhuang, J., S. Mukherjee, Y. Ron, and J.P. Dougherty. 2006. High rate of genetic recombination in murine leukemia virus: implications for influencing proviral ploidy. *J Virol*. 80:6706-6711.

Zilberman, D., X. Cao, and S.E. Jacobsen. 2003. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science*. 299:716-719.

Zwart, M.P., J.A. Daros, and S.F. Elena. 2011. One is enough: in vivo effective population size is dose-dependent for a plant RNA virus. *PLoS Pathog*. 7:e1002122.

Plants have many different mechanisms to counter viral infections. Natural resistances are mostly specific to a plant/virus pathosystem. Breeders need to get durable solution against the pathogen that has a major threat on their crop. Unfortunately, nature did not provide resistance to all pathogen, and when this is the case, virus can develop strategies to avoid some resistance. The aim of this study is to explore the possibilities of a new biotechnical approach based on the transgenic expression of amiR. Aware of the possible limitations of the technique, as breakdown of resistance, we engaged on an experimental evolution to predict the chances of resistance breaking in field conditions as well as to propose new specific rules to design more durable amiRs.